Tensor Analysis and Curvilinear Coordinates

Phil Lucht
Rimrock Digital Technology, Salt Lake City, Utah 84103
last update: May 19, 2016
rimrock@xmission.com

Maple code is available upon request. Comments and errata are welcome.

The graphics look ratty in Windows Adobe PDF viewers when not scaled up, but look just fine in this excellent freeware viewer: http://www.tracker-software.com/pdf-xchange-products-comparison-chart .

The table of contents has live links. Most PDF viewers provide these links as bookmarks on the left.

Overview and Summary

In this lengthy monograph, tensor analysis (also known as tensor algebra or tensor calculus) is developed starting from Square Zero which is an arbitrary invertible continuous transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ in N dimensions.

The subject was "exposed" by Gregorio Ricci in the late 1800's under the name "absolute differential calculus". He and his student Tullio Levi-Civita published a masterwork on the subject in 1900 (see References). Christoffel and others had laid the groundwork a few decades earlier.

Three somewhat different applications of tensor analysis are treated concurrently.

Our primary concern is the subject of curvilinear coordinates in N dimensions. All the basic expressions for the standard differential operators in general curvilinear coordinates are derived from scratch (in several ways). These results are often stated but not so often derived.

The second application involves transformations connecting "frames of reference". These transformations could be spatial rotations, Galilean transformations, the Lorentz transformations of special relativity, or the transformations involving the effects of gravity in general relativity. Beyond establishing the tensor analysis formalism, not much is said about this set of applications.

The third application deals with material flows in continuum mechanics.

The first six Chapters develop the theory of tensor analysis in a simple developmental notation where all indices are subscripts, just as in normal college physics. After providing motivation, Chapter 7 translates this developmental notation into the Standard Notation in use today. Chapter 8 treats transformations of length, area and volume and then the curvilinear differential operator expressions are derived, one per Chapter, with a summary in the penultimate Chapter 14. The final Chapter 15 rederives all the same results using the notion of covariance and associated covariant derivatives.

The information is presented informally as if it were a set of lectures. Little attention is paid to strict mathematical rigor. There is no attempt to be concise: examples are given, tangential remarks are inserted, almost all claims are derived in line, and there is a certain amount of repetition. The material is presented in a planned sequence to minimize the need for forward references, but the sequence is not perfect. The interlocking pieces of tensor analysis do seem to exhibit a certain logical circularity.

The reader will find in this document very many worked-out, detailed calculations, derivations and proofs. The intent is to provide the reader with "hands-on experience" in working with all the tensor tools.

The only real prerequisites for the reader are a knowledge of calculus of several variables (such as the chain rule and meaning of ∇^2) and of basic linear algebra (matrices and determinants).

The following Table of Contents (TOC) highlights in bold the main topics which are harder to see in the detailed TOC presented above. In most PDF viewers, the clickable TOC entries appear also as clickable bookmarks to the left of the text. Our equation references are not clickable, but the constant drumbeat of sequential equation numbers on the right should make them easy to locate. (**Split-screen PDF viewing** can be simulated by loading up two copies of the same PDF under different names, a very useful trick.)

Here then is a brief summary of each Chapter and each Appendix.

Chapter 1 introduces the notion of the general invertible transformation $\mathbf{x'} = \mathbf{F}(\mathbf{x})$ as a mapping between x-space and x'-space. The range and domain of this mapping are considered in the familiar examples of polar and spherical coordinates. These same examples are used to illustrate the general ideas of coordinate lines and level surfaces. Certain Pictures are introduced to allow different names for the two inter-mapped spaces, for the function F, and for its associated objects.

Chapter 2 introduces the linear transformations R and $S=R^{-1}$ which approximate the (generally non-linear) $\mathbf{x'} = \mathbf{F}(\mathbf{x})$ in the local neighborhood of a point $\mathbf{x}$. It is shown that two types of vectors naturally arise in the context of this linearization, called contravariant and covariant, and an overbar is used to distinguish a covariant vector. Vector fields are defined and their transformations stated. The idea of scalars and vectors as tensors of rank 0 and rank 1 is presented.

Chapter 3 defines the tangent base vectors $\mathbf{e}_n(\mathbf{x})$ which are tangent to the x'-coordinate lines in x-space. In the example of polar coordinates it is shown that $\mathbf{e}_r = \hat{\mathbf{r}}$ and $\mathbf{e}_\theta = r\,\hat{\boldsymbol{\theta}}$**.** The vectors $\mathbf{e}_n$ exist in x-space and form there a complete basis which in general is non-orthogonal. The tangent base vectors $\mathbf{u}'_n(\mathbf{x}')$ of the inverse transformation $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$ are also defined.

Chapter 4 gives a brief review of the notions of norm, metric and scalar product in Cartesian Space.

Chapter 5 addresses the metric tensor, called $\bar{g}$ in x-space and $\bar{g}'$ in x'-space. The metric tensor is first defined as a matrix object $\bar{g}$, and then $g \equiv \bar{g}^{-1}$. A definition is given for two kinds of (pure) rank-2 tensors (both matrices), and it is then shown that $\bar{g}$ transforms as a covariant rank-2 tensor while g is a contravariant rank-2 tensor. It is demonstrated how $\bar{g}$ applied to a contravariant vector $\mathbf{V}$ produces a vector that is covariant $\bar{\mathbf{V}} = \bar{g}\,\mathbf{V}$, and conversely $g\,\bar{\mathbf{V}} = \mathbf{V}$. In Cartesian space $g = 1$, so the two types of vectors coincide. The role of the metric tensor in the covariant vector dot product is stated, and the metric tensor is related to the tangent base vectors of Chapter 3. The Jacobian J and associated functions are defined, though the significance of J is deferred to Chapter 8. The last Sections briefly discuss the connection between tensor algebra and special relativity (with a mention of spinor algebra), general relativity, and continuum mechanics.

Chapter 6 introduces the reciprocal (dual) base vectors $\mathbf{E}_n$ which are later called $\mathbf{e}^n$ in the Standard Notation. Of special interest are the covariant dot products among the $\mathbf{e}_n$ and $\mathbf{E}_n$. It is shown how an arbitrary vector can be expanded onto different basis sets. It is found that when a contravariant vector in x-space is expanded on the tangent base vectors $\mathbf{e}_n$, the vector components in that expansion are in fact those of the contravariant vector in x'-space, $V'_i = R_{ij}V_j$. This fact proves useful in later Chapters which express differential operators in x-space in terms of curvilinear coordinates and objects of x'-space. The reciprocal base vectors $\mathbf{U}'_n$ of the inverse transformation are also discussed.

Chapter 7 motivates and then makes the transition from the developmental notation to the Standard Notation where contravariant indices are up and covariant ones are down. Although such a transition might seem completely trivial, many confusing issues do arise. Once a matrix can have up and down indices, matrix multiplication and other matrix operations become hazy: a matrix becomes four different matrices. The matrices R and S act like tensors, but are not tensors, and in fact are not even located in a well-defined space. Section 7.15 discusses the significance of tensor analysis with respect to physics in terms of covariant equations, and Section 7.16 broaches the topic of the covariant derivative of a vector field with its associated Christoffel symbols. Section 7.17 describes how to expand tensors of any rank in various bases and notations, and finally Section 7.18 summarizes information about the basis vectors.

The focus then fully shifts to curvilinear coordinates as an application of tensor analysis. The final Chapters are all written in the Standard Notation.

Chapter 8 shows how differential length, area and volume transform under $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. It considers the inverse mapping of a differential orthogonal N-piped (N dimensional parallelepiped) in x'-space to a skewed one in x-space. It is shown how the scale factors $h'_n = \sqrt{g'_{nn}}$ describe that ratio of N-piped edges, while the Jacobian $J = \sqrt{\det(g'_{nn})}$ describes the ratio of N-piped volumes. The relationship between the vector areas of the N-pipeds is more complicated, and it is found that the ratio of vector area magnitudes is $\sqrt{\mathrm{cof}(g'_{nn})}$. Heavy use is made of the results of Appendices A and B, as outlined below.

Chapters 9 through 13 use the information of Chapter 8 and earlier material to derive expressions for all the standard differential operators expressed in general non-orthogonal curvilinear coordinates: divergence, gradient, Laplacian, curl, and vector Laplacian. The last two operators are treated only in N=3 dimensions where the curl has a vector representation, but then the curl is generalized to N dimensions. For each differential operator, simplified results for orthogonal coordinates are also stated.

Chapter 14 gathers all the differential operator expressions into a set of tables, and revisits the polar coordinates example one last time to illustrate a reasonably clean and practical curvilinear notation.

Chapter 15 rederives the general results of Chapters 9 through 13 using the ideas of covariance and covariant differentiation. These derivations are elegantly brief, but lean heavily on the idea of tensor densities (Appendix D) and on the implications of covariance of tensor objects involving covariant derivatives (Appendix F).

About half of our content is contained in a set of Appendices.

Appendix A develops an alternative expression for the reciprocal base vector $\mathbf{E_n}$ as a generalized cross product of the tangent base vectors $\mathbf{e_n}$, applicable when x-space is Cartesian. This alternate $\mathbf{E_n}$ is shown to match the $\mathbf{E_n}$ defined in Chapter 6, and the covariant dot products involving $\mathbf{E_n}$ and $\mathbf{e_n}$ are verified.

Appendix B presents the geometry of a parallepiped in N dimensions (called an N-piped). Using the alternate expression for $\mathbf{E_n}$ developed in Appendix A, it is shown that the vector area of the n^{th} pair of faces on an N-piped spanned by the $\mathbf{e_n}$ is given by $\pm\, \mathbf{A_n}$, where $\mathbf{A_n} = |\det(S)|\, \mathbf{E_n}$, revealing a geometric significance of the reciprocal base vectors. Scaled by differentials so $d\mathbf{A^n} = |\det(S)|\, \mathbf{E_n}(\Pi_{i\neq n}\, dx'^i)$, this equation is then used in Chapter 9 where the divergence of a vector field is defined as the total flux of that field flowing out through all the faces of the skewed differential N-piped in x-space divided by its volume. This same $d\mathbf{A^n}$ appears in Chapter 8 with regard to the transformation of N-piped face vector areas.

Appendix C presents a case study of an N=2 non-orthogonal coordinate system, elliptical polar coordinates. Both the forward and inverse coordinate lines are displayed. The meaning of the curvilinear (x'-space) component V'^n of a contravariant vector is explored in the context of this system, and the difficulties of drawing such components in non-Cartesian (curvilinear) x'-space are pondered. Finally, the Jacobian Integration Rule for changing integration variables is derived.

Appendix D discusses tensor densities and their rules of the road. Special attention is given to the Levi-Civita ε tensor, including a derivation of all the $\varepsilon\varepsilon$ contraction formulas and their covariant statements. It is noted that the curl of a vector is a vector density.

Appendix E describes direct product and polyadic notations (including dyadics) and shows how to expand tensors (and tensor densities) of arbitrary rank on an arbitrary basis.

Appendix F deals with covariant derivatives and the affine connection Γ which tells how the tangent base vectors $\mathbf{e_n}(\mathbf{x}')$ change as $\mathbf{x}'$ changes. Everything is derived from scratch and the results provide the horsepower to make Chapter 15 go.

The next four appendices provide demonstrations of most ideas presented in this paper. In each Appendix, a connection is made to continuum mechanics, and the results are then derived by "brute force", by the covariant technique enabled by Appendix F, or by both methods. These Appendices were motivated by the continuum mechanics text of Lai, Rubin and Krempl (referred to as "Lai", see References). In each Appendix, it is shown how to express the object of interest in arbitrary curvilinear coordinates. Maple code is provided for the general calculations, and that code is then checked to make sure it accurately replicates the results quoted in Lai for spherical and cylindrical coordinates.

Appendix G treats the dyadic object $(\nabla\mathbf{v})$, where $\mathbf{v}$ is a vector.

Appendix H does the same for the vector object **divT** where T is a rank-2 tensor.

Appendix I deals with the vector Laplacian ✡**B** where **B** is a vector.

Appendix J treats the object (∇T), where T is a rank-2 tensor.

Appendix K (following Lai) discusses deformation tensors used in continuum mechanics. Those that are truly tensors may be used to construct covariant constitutive equations describing continuous materials. The covariant time derivatives $\overset{o}{T}_n$, $\overset{\Delta}{T}_n$ and $\overset{\nabla}{T}_n$ are derived (n = 1 are objective Cauchy stress rates). Some examples of covariant constitutive equations involving these tensor objects are listed.

Notations

RHS, LHS refer to the right hand side and left hand side of an equation

QED = which was to be demonstrated ("thus it has been proved")

a ≡ stuff means a is defined by stuff.

// indicates a comment on something shown to the left of //

diag(a,b,c..) means a diagonal matrix with diagonal elements a,b,c..

det(A), A^T = determinant of the matrix A, transpose of a matrix A

Maple = a computer algebra system similar to Mathematica and MATLAB

vectors (like **v**) are indicated in bold; all other tensors (like T) are unbolded.

$\hat{\mathbf{n}} = \mathbf{u}_n$ = unit vector pointing along the n^{th} positive axis of some coordinate system

$\partial_i = \partial/\partial x^i$ and $\partial_t = \partial/\partial t$ are partial derivatives

$d_t \equiv D_t \equiv D/Dt \equiv d/dt$ is a total time derivative, as in $d_t f(\mathbf{x},t) = \partial_t f(\mathbf{x},t) + [\partial f(\mathbf{x},t)/\partial x^i][\partial x^i/\partial t]$

$V_{,a}$ means $\partial_a V$ which means $\partial V/\partial x^a$, and $V_{;a}$ refers to the corresponding covariant derivative

$\mathbf{x} \bullet \mathbf{y}$ is a covariant dot product = $g_{ab}x^a y^b$, except where otherwise indicated

when an equation is repeated after its initial appearance, its equation number is shown in italics

Figures are treated as if they were equations in terms of equation numbering

Components of a vector are written $V_i = (\mathbf{V})_i$ (bold used only in the second form)

Chapter N consists of Sections N.1, N.2 .....

1. The Transformation F: invertibility, coordinate lines, and level surfaces

If $\mathbf{x}$ and $\mathbf{x}'$ are elements of the vector space R^N (N-dimensional reals), one can specify a **mapping**

$$\mathbf{x}' = \mathbf{F}(\mathbf{x}) \qquad\qquad F: R^N \rightarrow R^N \tag{1.1}$$

defined by a set of N continuous (C^2) real functions F_i , each of N real variables,

$$\begin{aligned} x'_1 &= F_1(x_1, x_2, x_3... x_N) \\ x'_2 &= F_2(x_1, x_2, x_3... x_N) \\ &... \\ x'_N &= F_N(x_1, x_2, x_3... x_N) . \end{aligned} \tag{1.2}$$

If all functions F_i are linear in all of their arguments, then the mapping $F: R^N \rightarrow R^N$ is a linear mapping. Otherwise the mapping is non-linear.

A mapping is often referred to as a **transformation**. We shall be interested only in transformations which are 1-to-1 and are therefore invertible. For such transformations,

$$\begin{aligned} &\mathbf{x}' = \mathbf{F}(\mathbf{x}) \quad\; \mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}') \qquad\qquad \text{or} \\ &\mathbf{x}' = \mathbf{x}'(\mathbf{x}) \quad\; \mathbf{x} = \mathbf{x}(\mathbf{x}') . \end{aligned} \tag{1.3}$$

In the transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$, if $\mathbf{x}$ roams over the entire R^N of x-space (the domain is R^N), we may find that $\mathbf{x}'$ roams over only some subset of R^N in x'-space. The 1-to-1 invertible mapping is then between the domain of mapping F which is all of R^N, and the range of mapping F which is this subset.

As just noted, it will be assumed that $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ is essentially invertible so $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$ exists for any $\mathbf{x}'$. By essentially is meant there may be a few problem points in the transformation which can be "fixed up" in some reasonable manner so that $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ is invertible (see examples below).

The functions F_i must be C^1 continuous to support the linearization derivatives appearing in Chapter 2, and they must be C^2 continuous to support some of the differential operators expressed in curvilinear coordinates in Chapters 9-14.

Example 1: Polar coordinates (N=2)

(a) The transformation from Cartesian to polar coordinates is given by,

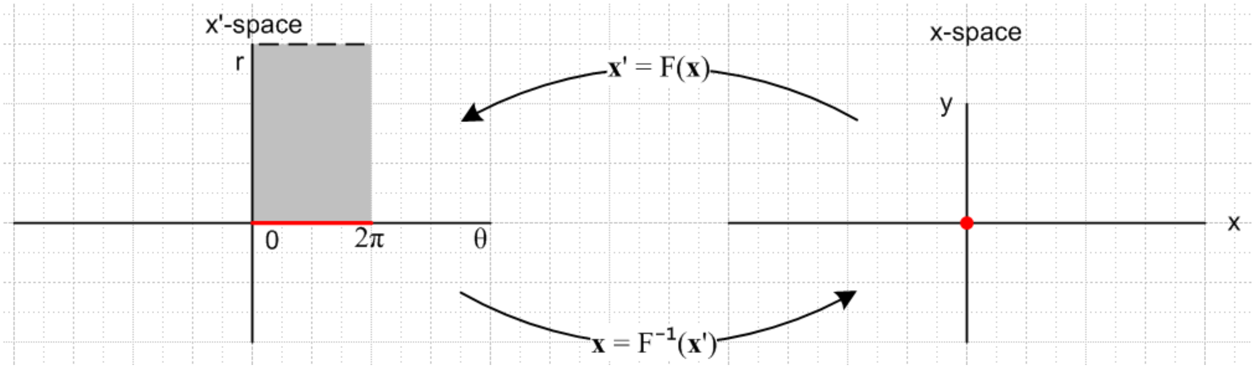
$$\mathbf{x} = (x_1, x_2) = (x,y)$$
$$\mathbf{x}' = (x_1', x_2') = (\theta,r) \qquad // \text{ note that } r = x_2'$$

$$\mathbf{x} = F^{-1}(\mathbf{x}') \quad \leftrightarrow \quad \begin{array}{ll} x = r\cos(\theta) & x_1 = x_2' \cos(x_1') \\ y = r\sin(\theta) & x_2 = x_2' \sin(x_1') \end{array}$$

$$\mathbf{x}' = \mathbf{F}(\mathbf{x}) \quad \leftrightarrow \quad \begin{array}{ll} r = \sqrt{x^2+y^2} & x_2' = \sqrt{x_1{}^2+x_2{}^2} \\ \theta = \tan^{-1}(y/x) & x_1' = \tan^{-1}(x_2/x_1)\ . \end{array} \qquad (1.4)$$

(b) The transformation is non-linear because at least one component function (e.g., $r = \sqrt{x^2+y^2}$) is not of the form r = Ax + By. In this transformation all functions are non-linear.

(c) Here is a drawing showing the nature of this mapping:



(1.5)

The domain of **x**' = **F**(**x**) (in x-space on the right) is all of R^2, but the range in x'-space is the semi-infinite vertical strip shown in gray. Imitating the language of complex variables, we can regard this gray strip as depicting the principle branch of the multi-variable function **x**' = **F(x)**. Other branches are obtained by shifting the gray rectangle left or right by multiples of 2π. Still other branches are obtained by taking the other branch of the real function $r = \sqrt{x^2+y^2}$ which produces down-facing strips. The principle branch plus all the other branches then fill up the E^2 of x'-space, but we care only about the principle branch range shown in gray (this strip could be any strip of width 2π, such as one from -π to π).

(d) This mapping illustrates a "problem point" involving $\theta = \tan^{-1}(y/x)$. This occurs when both x and y are 0, indicated by the red dot on the right. The inverse mapping takes the entire red line segment into this red origin point, so we have a lack of 1-to-1 going on here, meaning that formally the function F is not invertible. This can be fixed up by eliminating the red line segment from the range of F, retaining only the point at its left end. Another problem is that both the left and right vertical edges of the gray strip map into the real axis in x-space, and that is fixed by removing the right edge. Thus, by doing a suitable trimming of the range, F can be made fudlly invertible. No one has ever had major problems using polar coordinates due to these minor issues.

Example 2: Spherical coordinates (N=3)

(a) The transformation from Cartesian to spherical coordinates is given by,

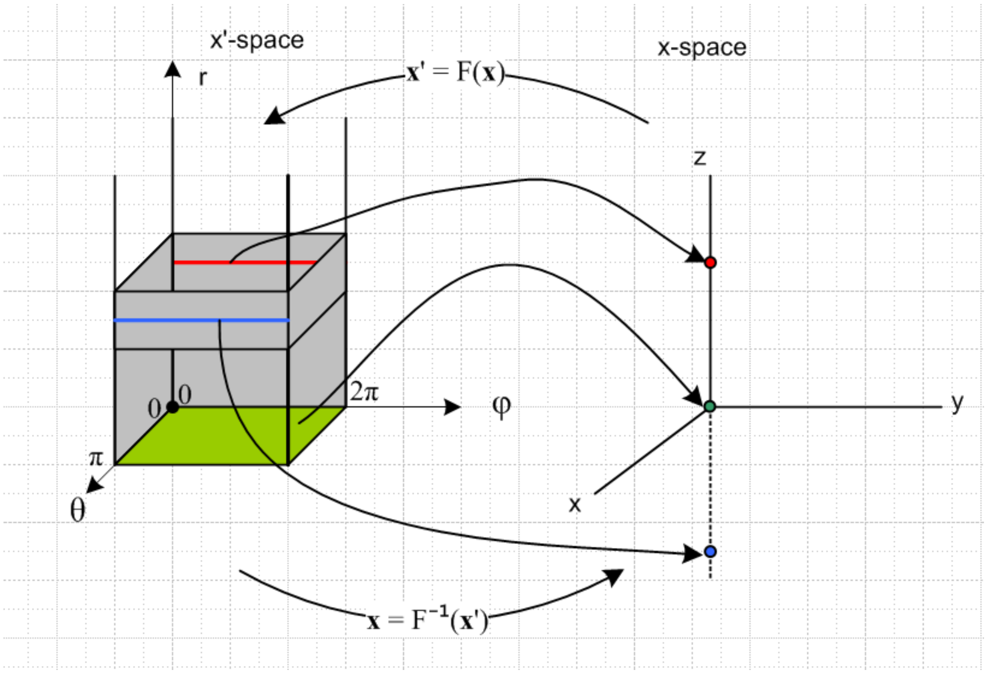
$$\mathbf{x} = (x_1,x_2,x_3\,) = (x,y,z)$$
$$\mathbf{x}' = (x_1',x_2',x_3') = (r,\theta,\varphi)$$

$$\begin{array}{lll} \mathbf{x} = F^{-1}(\mathbf{x}') \quad \leftrightarrow & x = r\sin(\theta)\cos(\varphi) & x_1 = x_1'\sin(x_2')\cos(x_3') \\ & y = r\sin(\theta)\sin(\varphi) & x_2 = x_1'\sin(x_2')\sin(x_3') \\ & z = r\cos(\theta) & x_3 = x_1'\cos(x_2') \end{array}$$

$$\begin{array}{lll} \mathbf{x}' = \mathbf{F}(\mathbf{x}) \quad \leftrightarrow & r = \sqrt{x^2+y^2+z^2} & x_1' = \sqrt{{x_1}^2+{x_2}^2+{x_3}^2} \\ & \theta = \cos^{-1}(z/\sqrt{x^2+y^2+z^2}\,) & x_2' = \cos^{-1}(x_3/\sqrt{{x_1}^2+{x_2}^2+{x_3}^2}\,) \\ & \varphi = \tan^{-1}(y/x) & x_3' = \tan^{-1}(x_2/x_1) \end{array} \qquad (1.6)$$

(b) The transformation is non-linear because at least one component function(e.g., $r = \sqrt{x^2+y^2+z^2}$) is not of the form $r = Ax + By + Cz$. In this transformation, all three functions are non-linear.

(c) Here is a drawing showing the nature of this mapping



(1.7)

The domain of $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ (in x-space on the right) is all of R^3, but the range in x'-space is the interior of an infinitely tall rectangular solid on the left we shall call an "office building". We could regard this office building as depicting the principle branch of the multi-variable function $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. Other branches are obtained by shifting the building left and right by multiples of 2π, or fore and aft by multiples of π, or by flipping it vertically, taking the other branch of $r = \sqrt{x^2+y^2+z^2}$. The principle branch plus all the other branch offices then fill up the R^3 of x-space, but we care only about the principle branch office building whose walls are mostly shown in gray.

(d) This mapping illustrates some "problem points". One is that entire green office building main floor (r=0) maps into the origin in x-space. This problem is fixed by trimming away the main floor keeping only the origin point of the bottom face of the office building. Another problem is that the entire red line segment ($\theta = 0$) maps into the red point shown in x-space. This is fixed by throwing out the back wall of the office building, retaining only a line going up the left edge of the back wall. A similar problem happens on the front wall ($\theta = \pi$, blue) and we fix it the same way: throw out the wall but maintain a thin line which is the left edge of this front wall (this line is missing its bottom point). Thus, by doing a suitable trimming of the range, F is made fully invertible.

Cartesian Space and Quasi-Cartesian Space

(a) Cartesian Space. For the purposes of this document, a Cartesian Space in N dimensions is "the usual" Hilbert Space R^N (or E^N) in which the distance between two vectors is given by the formula

$$d(\mathbf{x},\mathbf{y}) = \sqrt{\Sigma_{i=1}^N (x_i-y_i)^2} \quad \Rightarrow \quad [d(\mathbf{x}+d\mathbf{x},\mathbf{x})]^2 = \Sigma_{i=1}^N (dx_i)^2$$

$$\text{metric tensor} = \text{diag}(1,1,1....1) \qquad (1.8)$$

as discussed in Chapter 4 below.

The θ-r space in the above Example 1 would be a Cartesian space if it were declared that the distance between two points there were $D'^2 = (\theta-\theta')^2 + (r-r')^2$, but that is not the usual intent in using that space. As shown below, the metric tensor used there is $g = \text{diag}(r^2,1)$ and not diag(1,1).

One might argue that our Cartesian Space is in fact a Euclidean space (hence E^N) having Cartesian coordinates. A *non*-Cartesian space is sometimes referred to as a "curved space" (non-Euclidean) and the coordinates in such a space as "curvilinear coordinates". An example is the θ-r space above.

With the Cartesian Space metric tensor as $g^C = 1 = \text{diag}(1,1....1)$, the above equations can be written

$$d^2(\mathbf{x},\mathbf{y}) = g^C_{ij}(x_i-y_i)(x_j-y_j) \quad \text{and} \quad [d(\mathbf{x}+d\mathbf{x},\mathbf{x})]^2 = g^C_{ij}\, dx_i\, dx_j \equiv (ds)^2 \qquad (1.9)$$

$$\text{metric tensor} = g^C = 1 = \text{diag}(1,1,1....1), \quad g^C_{ij} = \delta_{i,j}$$

where repeated indices are implicitly summed, called **the Einstein summation convention**.

(b) Quasi-Cartesian Space. We now define a Quasi-Cartesian Space (not an official term) as one which has a diagonal metric tensor G whose diagonal elements are independently +1 or -1 instead of all +1 as with g^C. In a Quasi-Cartesian Space the two equations above become

$$d^2(\mathbf{x},\mathbf{y}) = G_{ij}(x_i-y_i)(x_j-y_j) \quad \text{and} \quad [d(\mathbf{x}+d\mathbf{x},\mathbf{x})]^2 = G_{ij}\, dx_i\, dx_j \equiv (ds)^2 \qquad (1.10)$$

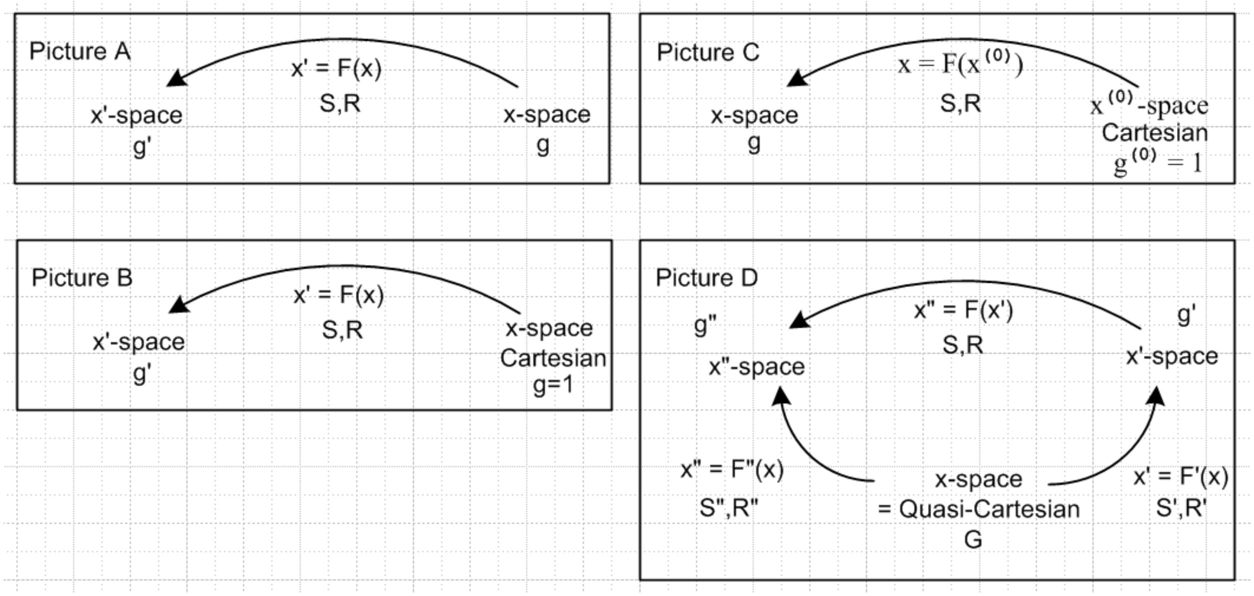
and of course this allows for the possibility of a negative distance-squared (see Section 5.10).

Notice that $G^{-1} = G$ for any distribution of the ±1's in G. As shown later, this means that that covariant and contravariant versions of G are the same.

The motivation for introducing this Quasi-Cartesian Space is to cover the case of special relativity which involves 4 dimensional linear transformations with G = diag(1,-1,-1,-1).

Pictures A,B,C and D

We shall usually work with one of four different "pictures" involving transformations. In each picture the spaces and transformations (and their associated objects) have certain names that prove useful in certain situations :



(1.11)

The matrices R and S are associated with transformation F as described in Chapter 2 below, while G and the g's are metric tensors.

Systems not marked Cartesian *could* of course be Cartesian, but we think of them as general "curved" systems with strange metric tensors. And in general, all the full transformations might be non-linear.

The polar coordinates example above was presented in the context of Picture B. Picture B is the right picture for studying curvilinear coordinates where for example x-space = Cartesian coordinates and x'-space = toroidal coordinates. Picture C is useful for making statements applying to objects in curved x-space where we don't want lots of primes floating around. Pictures A and D are appropriate for consideration of general transformations, as well as linear ones like rotations and Lorentz transformations. In Chapters 9-14 Picture M&S (Moon & Spencer) is introduced for the special purpose of displaying the differential operator expressions. This is Picture B with x'→ u and g'→g on the left side.

The entire rest of this Section uses the Picture B context.

Coordinate Lines

Suppose in x'-space one *varies* a single coordinate, say x'_i, keeping all the other coordinates fixed. In x'-space the locus of points thus created is just a straight line parallel to the x'_i axis, or for a principle branch situation like that of the above examples, a straight line segment. When such a straight line or segment is mapped into x-space, the result is a curve known as a **coordinate line**. A coordinate line is associated with a specific x'-space coordinate x'_i, so one might refer to the " x'_i -coordinate line", x'_i being a label.

In N dimensions, a point **x** in x-space lies on a *unique set* of N coordinate lines with respect to a transformation F. Remember that each such line is associated with one of the x'_i coordinates. In x'-space, a point **x**' lies on a unique intersection of straight lines or segments, and then this all gets mapped into x-space where point $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$ then lies on a unique intersection of coordinate lines.

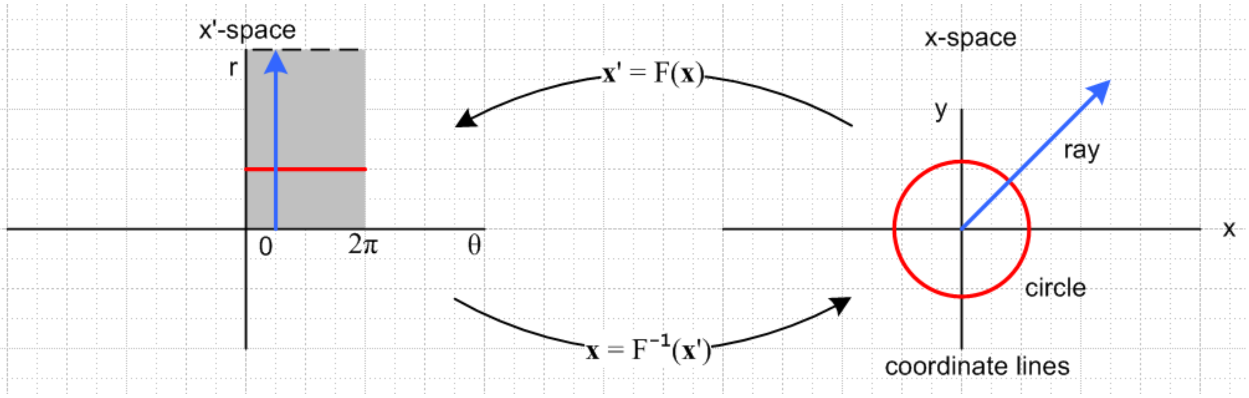
For example, in spherical coordinates we start with some (x,y,z) in x-space and compute the x'_i = (r,θ,φ) in x'-space. Our point **x** in x-space then lies on the r-coordinate line whose label is r, it lies on the θ-coordinate line whose label is θ, and it lies on the φ-coordinate line whose label is φ (see below).

In general a coordinate "line" is some non-planar curve in N-dimensional x-space, meaning that a coordinate line might not lie on an N-1 dimensional plane. In the 2D polar coordinates example below, the red coordinate line does not lie on a 1-dimensional plane (line). In the next example of 3D spherical coordinates, it happens that every coordinate line does lie on a 2-dimensional plane. But in ellipsoidal coordinates, another 3D orthogonal system, every coordinate line does not lie on a 2-dimensional plane.

Some authors refer to coordinate lines as **level curves**, especially in two dimensions mapping the real and imaginary part of analytic functions w = f(z) (Ahlfors p 89).

Example 1: Polar coordinates, coordinate lines

Here are some coordinate lines for our prototype N=2 non-linear transformation, Cartesian to polar coordinates:

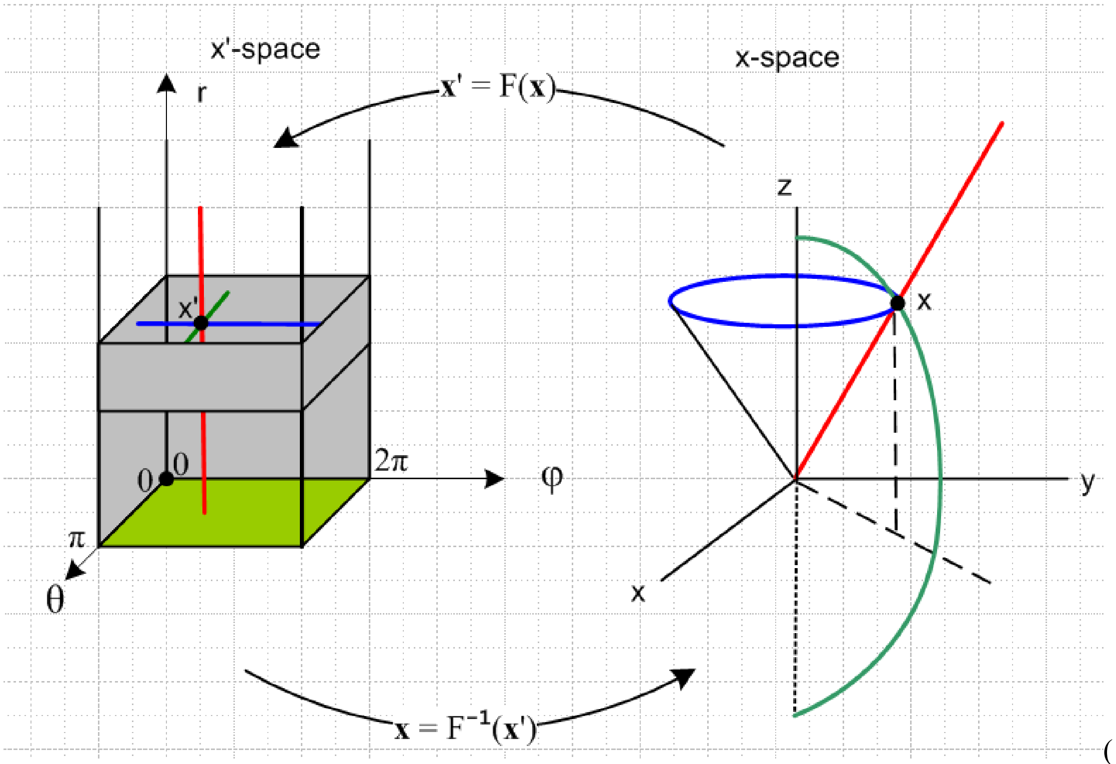


(1.12)

The red circle is a θ-coordinate line, and the blue ray is an r-coordinate line

Example 2: Spherical coordinates, coordinate lines

These coordinate lines are generated exactly as described above. In x'-space one holds two coordinates fixed while allowing one to vary. The locus in x'-space is a line segment or a half line (in the case of varying r). In x-space, the corresponding coordinate lines are as shown.



(1.13)

The green coordinate line is a θ-coordinate line, since only θ is varying.
The red coordinate line is an r-coordinate line, since only r is varying.
The blue coordinate line is a φ-coordinate line, since only φ is varying.

The point **x** indicated by a black dot in x-space lies on the unique set of coordinates lines shown.

Appendix C gives an example of coordinate lines for a non-orthogonal 2D coordinate system.

Level Surfaces

(a) Suppose in x'-space one *fixes* one coordinate, say x'_i, and varies all the *other* coordinates. In x'-space the locus of points thus created is just an (N-1 dimensional) plane perpendicular to the x_i axis, or for a principle branch situation like that above, a rectangle or half strip in the case of r. Mapping this planar surface in x'-space into x-space produces a surface in x-space (of dimension N-1) called a **level surface**. The equations of the N different x_i level surface types are

$$a'^{(n)}_i = F_i(x_1, x_2.....x_N) \qquad i = 1,2...N \qquad (1.14)$$

where $a'_i{}^{(n)}$ is some constant value selected for fixed coordinate x'_i. By taking some set of closely spaced values for this constant, $\{ a'_i{}^{(1)}, a'_i{}^{(2)} .....\}$, one obtains a family of level surfaces all of the same general shape which are closely spaced. For some different value of i, the shapes of such a family of level surfaces will in general be different. In general if $f(x_1, x_2.....x_N) = k$, the set of points **x** which make this equation true for some fixed k is called a **level set**, so a level set is a surface of dimension N-1. Thus, all our level curves are also level sets.

In the polar coordinates example, since there are only 2 coordinates, there is no distinction between a level surface and a coordinate line.

In the spherical coordinates example, there is a distinction.

If one fixes r and varies θ and φ over their horizontal rectangle inside the office building, the level surface in x-space is a **sphere**.

If one fixes θ and varies r and φ over a left-right vertical strip inside the office building, the level surface in x-space is a sphere is a polar **cone**

If one fixes φ and varies r and θ over a fore-aft vertical strip inside the office building, the level surface in x-space is a **half plane** at azimuth φ.

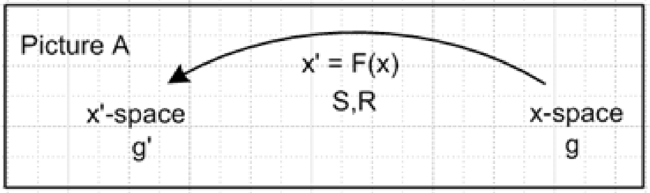
(b) In N dimensions there will be N level surfaces in x-space, each formed by holding some x'_i fixed. The intersection of N-1 level surfaces (omitting say the x_3' level surface) will have all of the x'_i fixed except x'_3. But this describes the x'_3 coordinate line. Thus, each coordinate line can be considered as the intersection of the N-1 level surfaces associated with the other coordinates. One can see this happening on the spherical coordinates example:

The green coordinate line is the intersection of two level surfaces: half-plane and sphere.
The red coordinate line is the intersection of two level surfaces: half-plane and cone.
The blue coordinate line is the intersection of two level surfaces: sphere and cone.

2. Linear Local Transformations associated with F : scalars and two kinds of vectors

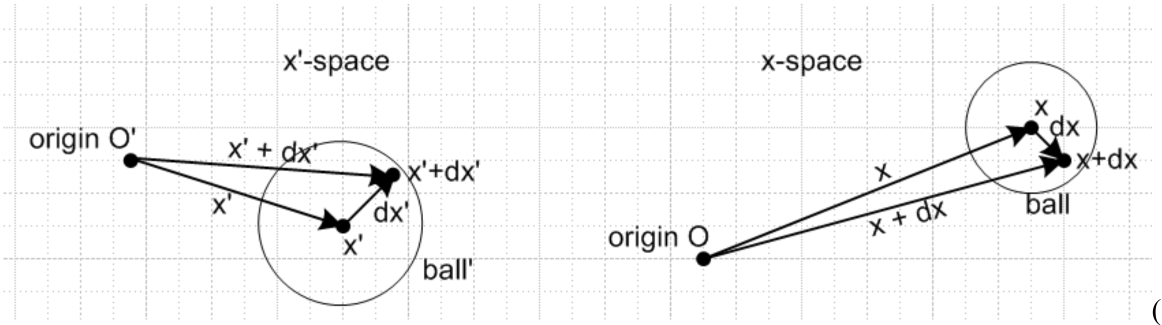
2.1 Linear Local Transformations

We now shift to the **Picture A** context, where x-space is not necessarily Cartesian,



(2.1.1)

Consider again the possibly non-linear transformation **x**' = **F**(**x**) mapping F: $R^N \rightarrow R^N$. Imagine a very small neighborhood around the point **x** in x-space, a "ball" around **x**. Where the mapping is continuous in both directions, one expects a tiny x-space ball around **x** to map into a tiny x'-space ball around **x**' and vice versa. Here is a picture of this situation,



(2.1.2)

where everything in one picture is the mapping of the corresponding thing in the other picture.

In particular, we show a small vector in x-space called d**x** which maps into a small vector in x'-space called d**x**'. Since F was assumed invertible, it must be invertible locally in these two balls. That is, given a d**x** above, one can determine d**x**', and vice versa. (Anticipating a few lines below, this means that the matrices S and R will be invertible so neither can have zero determinant.)

How are these two differential vectors related? For a linear approximation,

$$x'_i + dx'_i = F_i(\mathbf{x} + d\mathbf{x}) \approx F_i(\mathbf{x}) + \Sigma_k(\,\partial F_i(\mathbf{x})/\partial x_k)\, dx_k$$

$$\Rightarrow\ dx'_i = \Sigma_k(\,\partial F_i(\mathbf{x})/\partial x_k)\, dx_k \,. \qquad (2.1.3)$$

The last line shows an equals sign in the limit that dx_k is a vanishing differential. Since $F_i(\mathbf{x}) = x'_i$,

$$dx'_i = \Sigma_k(\partial x'_i/\partial x_k)\, dx_k = \Sigma_k R_{ik}\, dx_k \qquad \text{where} \quad R_{ik} \equiv (\partial x'_i/\partial x_k)\,. \qquad (2.1.4)$$

Doing the same operation in the other direction gives

$$dx_i = \Sigma_k(\,\partial x_i/\partial x'_k)\, dx'_k = \Sigma_k S_{ik}\, dx'_k \qquad \text{where} \quad S_{ik} \equiv (\partial x_i/\partial x'_k)\,. \qquad (2.1.5)$$

One can regard R_{ik} and S_{ik} as elements of NxN matrices R and S. In vector notation then,

$$\begin{aligned} d\mathbf{x}' &= R(\mathbf{x})\, d\mathbf{x} \qquad R_{ik}(\mathbf{x}) \equiv (\partial x'_i/\partial x_k) \qquad R = S^{-1} \qquad // \ dx'_i = R_{ij}\, dx_j \\ d\mathbf{x} &= S(\mathbf{x}')\, d\mathbf{x}' \qquad S_{ik}(\mathbf{x}') \equiv (\partial x_i/\partial x'_k) \qquad S = R^{-1} \qquad // \ dx_i = S_{ij}\, dx'_j\ . \end{aligned} \tag{2.1.6}$$

It is obvious that matrices R and S are inverses of each other, just staring at the above two vector equations. One can verify this fact from the definitions of R and S using the chain rule

$$(RS)_{ij} = \Sigma_k R_{ik}S_{kj} = \Sigma_k (\partial x'_i/\partial x_k)\,(\partial x_k/\partial x'_j) = \Sigma_k \frac{\partial x'_i}{\partial x_k}\frac{\partial x_k}{\partial x'_j} = \frac{\partial x'_i}{\partial x'_j} = \delta_{i,j}\ . \tag{2.1.7}$$

We could get rid of one of these matrices right now, perhaps keeping R and replacing $S = R^{-1}$, but keeping both simplifies expressions encountered later, so for now both are kept.

The letter R does *not* imply that matrix R is a rotation matrix, although it could be. According to the polar decomposition theorem (Lai p 110), any matrix R (detR ≠ 0) can be uniquely written in the form $R = \mathcal{R}U = V\mathcal{R}$ where $\mathcal{R}$ is a rotation matrix (the same one in $\mathcal{R}U$ and $V\mathcal{R}$) and U and V are symmetric positive definite matrices (called right and left stretch tensors) related by $U = \mathcal{R}^T V\mathcal{R}$. Matrix S could of course be written in a similar manner.

Matrices $R(\mathbf{x})$ and $S(\mathbf{x}')$ are in general functions of a point in space $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. As one moves around in space, all the elements of matrices R and S are likely to change. So R and S represent point-dependent linear transformations which are valid for the differentials shown.

One might wonder at this point how the vector $d\mathbf{x}$ is related to its components dx_i and the same question for $d\mathbf{x}'_i$ and dx'_i. As will be shown later in (6.6.9) and (6.6.15),

$$\begin{aligned} d\mathbf{x} &= \Sigma_n dx_n\, \mathbf{u}_n && \text{where the } \mathbf{u}_n \text{ are x-space axis-aligned basis vectors of the form } \mathbf{u}_1 = (1,0,0,..0) \\ d\mathbf{x}' &= \Sigma_n dx'_n\, \mathbf{e}'_n && \text{where the } \mathbf{e}'_n \text{ are x'-space axis-aligned basis vectors of the form } \mathbf{e}'_n = (1,0,0,..0) \end{aligned} \tag{2.1.8}$$

If x-space and x'-space were Cartesian, one could write $\mathbf{u}_n = \hat{\mathbf{n}}$ and $\mathbf{e}'_n = \hat{\mathbf{n}}'$, but in general the $\mathbf{u}_n$ and $\mathbf{e}'_n$ vectors do not have (covariant) unit length, as shown later in (6.5.3) and (6.4.1).

The reader familiar with covariant "up and down" indices will notice that all indices are peacefully sitting "down" in the presentation so far (subscripts, no superscripts). As we carry out our various developmental tasks, that is where all indices shall remain until Chapter 7, whereupon they will start frantically bobbing up and down, seemingly at will. [Since rules are made to be violated, we have violated this one in some examples below where non-standard notation would be hard to swallow.]

Are there any "useful objects" that can be constructed from differentials $d\mathbf{x}$ and which might then transform according by R or S? The answer is yes, but first we discuss scalars.

2.2 Scalars

A quantity is a **scalar** with respect to transformation F if it is the same in both spaces. Thus, any constant like π would be a scalar under any transformation. The mass m of a potato would be a constant under transformations that are rotations or translations. A function of space $\varphi(\mathbf{x})$ is a "field" and it would be a "scalar field" if $\varphi'(\mathbf{x}') = \varphi(\mathbf{x})$. For example, temperature would be a scalar field under rotations. Notice that φ is evaluated at **x**, while φ' is evaluated at **x'** = **F**(**x**). As noted in Section 2.12, one could be more precise by referring to the objects described here as a "tensorial scalar" and a "tensorial scalar field".

2.3 Contravariant vectors

If **x** and **x'** are spatial coordinates (time t is not one of the x_i), then consider the familiar velocity vector,

$$v_i = dx_i/dt \qquad \Rightarrow \qquad \mathbf{v} = d\mathbf{x}/dt \ . \tag{2.3.1}$$

Since dt transforms as a constant (scalar) under our selected transformation type, it seems pretty clear that velocity in x'-space can be related to velocity in x-space using the $d\mathbf{x}' = R(\mathbf{x})\, d\mathbf{x}$ rule above, so

$$\mathbf{v}' = R(\mathbf{x})\, \mathbf{v} \ . \tag{2.3.2}$$

Even though the matrix R(**x**) changes as we move around, this linear transformation R is valid at any point **x** when applied to velocity. Momentum **p** = m**v** would work the same way, since mass m is a scalar (Newtonian mechanics).

In contrast, unless R(**x**) is a constant in space (meaning from (2.8.7) that **F**(**x**) is linear), $\mathbf{x}' \neq R(\mathbf{x})\, \mathbf{x}$, so in general **x** itself is *not* a contravariant vector although d**x** is.

Any vector that transforms according to **V'** = R(**x**)**V** with respect to a transformation F (such as Newtonian velocity and momentum with respect to rotations) is called a **contravariant vector**.

2.4 Covariant vectors

Much of physics is described by differential equations involving the gradient operator (the reason for the overbar is explained in the next Section)

$$\overline{\nabla}_i = \overline{\partial}_i = \partial/\partial x_i \tag{2.4.1}$$

which involves an "upside down" differential. Here is how this *operator* transforms going from x-space to x'-space, again according to the chain rule (implied sum on k, T means transpose)

$$\overline{\nabla}'_i = \overline{\partial}'_i = \frac{\partial}{\partial x'_i} = \frac{\partial x_k}{\partial x'_i} \frac{\partial}{\partial x_k} = S_{ki}\overline{\partial}_k = S^T{}_{ik}\, \overline{\partial}_k = S^T{}_{ik}\, \overline{\nabla}_k = (S^T\, \overline{\nabla})_i$$

$$\Rightarrow \ \overline{\nabla}' = S^T\, \overline{\nabla} \ . \tag{2.4.2}$$

One can think of $\overline{\nabla}$ as acting on a scalar field φ(x) = φ'(x'), and then the above becomes

$$\overline{\nabla}'_i\, \varphi'(x') = \frac{\partial}{\partial x'_i}\varphi'(x') = \frac{\partial x_k}{\partial x'_i}\, \frac{\partial}{\partial x_k}\, \varphi(x) = S^T{}_{ik}\, \overline{\nabla}_k\, \varphi(x)$$

$$\Rightarrow \ \overline{\nabla}'\varphi'(x') = S^T\, \overline{\nabla}\varphi(x) \ . \tag{2.4.3}$$

Since the differential is "upside down", one might expect $\overline{\nabla}$ to transform according to $S = R^{-1}$ instead of R, but it is really S^T that does the job. One could write $\overline{\nabla}' = \overline{\nabla}\, S$ in terms of row vectors.

Vectors that transform according to $\mathbf{V}' = S^T(\mathbf{x})\, \mathbf{V}$ such as the gradient operator $\overline{\nabla}$ are called **covariant vectors** with respect to transformation F.

An example of a covariant vector is the electrostatic electric field obtained from the potential Φ

$$\overline{\mathbf{E}} = -\overline{\nabla}\,\Phi \qquad\qquad \overline{\mathbf{E}}_i = -\overline{\partial}_i\Phi = -\partial\Phi/\partial x_i \qquad\qquad (2.4.4)$$

2.5 Bar notation

In order to distinguish a contravariant from a covariant vector, we shall (for a while) adopt this **bar convention**: contravariant vectors shall be written $\mathbf{V}$ with components V_i and covariant vectors shall be written $\overline{\mathbf{V}}$ with components $\overline{V}_i$. This is why overbars were placed on $\overline{\nabla}$ and $\overline{\partial}_i$ and $\overline{\mathbf{E}}$ in the previous Section. We call this our "developmental notation", as distinct from the Standard Notation introduced in Chapter 7.

The transformation rules for the two vector types can now be written this way:

$$\begin{array}{llllll} \mathbf{V}' = R\,\mathbf{V} & \text{contravariant} & R_{ik}(\mathbf{x}) \equiv (\partial x'_i/\partial x_k) & & & R = S^{-1} \\ \overline{\mathbf{V}}' = S^T\,\overline{\mathbf{V}} & \text{covariant} & S_{ik}(\mathbf{x}') \equiv (\partial x_i/\partial x'_k) & = & S^T{}_{ki}(\mathbf{x}') & \end{array} \qquad (2.5.1)$$

One could imagine replacing S with some Q^T to make the second equation more like the first, but of course then $RQ^T = 1$ instead of $RS = 1$. In the Standard Notation, where there are four versions of the matrix R, we shall see in Section 7.5 that $R \rightarrow R^i{}_j$ and $S \rightarrow S^i{}_j = R_j{}^i$ and S can be removed from the picture altogether.

It is possible to do without *either* R or S and just *write out* all partial derivatives like $\frac{\partial x'_i}{\partial x^k}$ in place of R_{ik}. Many authors do this, including Weinberg. We feel that the forms R_{ik} and later $R^i{}_k$ are much more compact (3 symbols in place of 8) and reveal more clearly the matrix (but not tensor) nature of R and the fact that it is a linear transformation that locally approximates the full $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. Later we shall have long strings of R objects which become extremely cluttered when derivatives are fully written out, see for example Section F.2 and following. Finally, R_{ik} is a easy to type!

2.6 Origin of the names contravariant and covariant

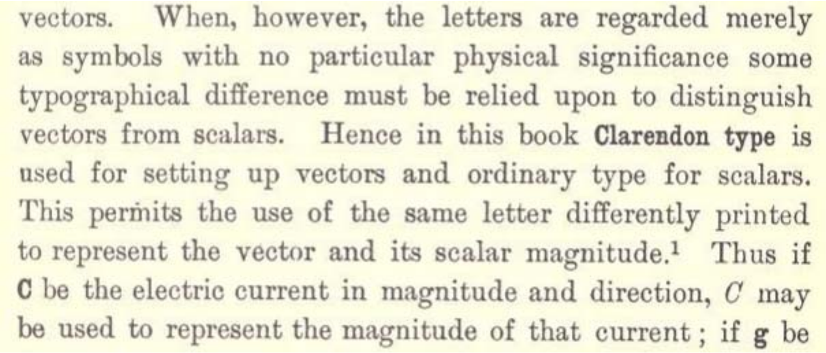
A justification of the terms covariant and contravariant is presented in Section 7.14, since the idea is more easily presented there than here.

It seems that these terms were first used in 1851 (a half century before special relativity) in a paper (see Refs.) by J.J. Sylvester of Sylvester's Law of Inertia fame. Sylvester uses the words covariant and contravariant to describe the relations between a pair of "transformations". In much simpler notation than he uses, if those "transformations" (functions) are F($\mathbf{x}$) and G($\mathbf{x}$) and if A is an 3x3 matrix, then

the pair F(A**x**) and G(A**x**) are said to be covariant (or concurrent)
the pair F(A**x**) and G(A^{-1}**x**) are said to be contravariant (or reciprocal)

The idea is that in comparing the way two things transform, if they both move the same way, then it is covariant, and if they move in opposite directions it is contravariant. In Section 7.14 this idea is applied to the transformation of two objects, where one object is the component of a vector like $V_{\mathbf{n}}$ and the other object is a basis vector onto which a vector is expanded. The connection is a bit distant, but the underlying concept carries through.

Notations like **y** = **F**(A**x**) would have mystified Sylvester in 1851, although in this same paper he introduced two-dimensional arrays of letters and referred to them as "matrices". According to a web piece by John Aldrich of the University of Southampton, J.W. Gibbs in 1881 was the first person to use a single letter to represent a vector (he used Greek letters). It was not until 1901 when his student E.B. Wilson published Gibb's lectures in a Vector Analysis book that the idea was propagated to a wider circle. Wilson converted those Greek letters to bolded ones,



vectors. When, however, the letters are regarded merely as symbols with no particular physical significance some typographical difference must be relied upon to distinguish vectors from scalars. Hence in this book **Clarendon type** is used for setting up vectors and ordinary type for scalars. This permits the use of the same letter differently printed to represent the vector and its scalar magnitude.[1] Thus if **C** be the electric current in magnitude and direction, *C* may be used to represent the magnitude of that current; if **g** be

(2.6.1)

The Wilson/Gibbs book was reprinted seven times, the last being 1943. In 1960 it continued as a Dover book and is now available online as a public domain document (above from page 4).

2.7 Other vector types?

Are there any *other* kinds of vectors with respect to a transformation F? There might be, but only the two types mentioned above are of interest to us in this document. They are both called rank-1 tensors, and there are no other rank-1 tensor types in "tensor analysis" (for rank-n tensors, see Section 7.10). Some authors refer to the **rank** of a tensor as the **order** of a tensor, and we shall sometimes use that term.

In the Standard Notation introduced later, where contravariant vector components are written with indices up and covariant vectors with indices down, and where the notation is so slick and smooth and automatic, one sometimes imagines there are two kinds of vectors *because* there are two places to put indices, up and down. It is of course the other way around: the up/down notation was adopted because there are two rank-1 tensor types.

Two particular (linear) transformation types of interest are rotations and Lorentz transformations, each of which has a certain number of continuous parameters (3 and 6). As the parameters are allowed to vary over their ranges, the set of transformations can be viewed as elements of a continuous group (SO(3) and SO(3,1)). Each of these groups has exactly *one* "vector representation" ("1" and "(1/2)⊕(1/2)"). One should not imagine that somehow the "two-ness" of vector types under general transformations F is connected to there being two vector representations of some particular group. It *happens* that the Lorentz group does have two "spinor representations" (1/2)⊕0 and 0⊕(1/2), but this has nothing at all to do with our general notion of two kinds of vectors. This subject is discussed in more detail in Section 5.14.

2.8 Linear transformations

First, consider a scalar function $f : R^n \rightarrow R$. The function $f(\mathbf{x})$ is *linear* iff

$$f(\mathbf{x}+\mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y}) \qquad \text{for all } \mathbf{x} \text{ and } \mathbf{y} \text{ in } R^n$$
$$f(a\mathbf{x}) = af(\mathbf{x}) \qquad \text{for all } \mathbf{x} \text{ in } R^n \text{ and for all a in R} . \qquad (2.8.1)$$

These conditions imply that the function $f(\mathbf{x})$ must have the form,

$$f(\mathbf{x}) = \Sigma_i \, a_i \, x_i \ . \qquad \text{where } a_i \text{ are constants independent of } \mathbf{x} \ . \qquad (2.8.2)$$

One way to reach this conclusion is to assume a Taylor expansion for $f(\mathbf{x})$ about $\mathbf{x} = 0$,

$$f(\mathbf{x}) = f(\mathbf{0}) + \Sigma_i[\partial_i f]^{\mathbf{x}=0} x_i + \Sigma_{ij}[\partial_{ij} f]^{\mathbf{x}=0} x_i x_j + \text{ higher terms} . \qquad (2.8.3)$$

The (2.8.1) requirement that $f(2\mathbf{x}) = 2f(\mathbf{x})$ eliminates the quadratic and higher terms and for $\mathbf{x} = 0$ eliminates the first term $f(\mathbf{0})$ since $f(\mathbf{0}) = 2f(\mathbf{0})$, resulting in an expression of the form (2.8.2).

When this same discussion is applied to the components $F_i(\mathbf{x})$ of the function $\mathbf{F} : R^n \rightarrow R^n$ one concludes that the form of $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ for linear $\mathbf{F}$ is given by

$$x'_i = F_i(\mathbf{x}) = \Sigma_j F_{ij} x_j \qquad \text{where } F_{ij} \text{ are } \textit{constants} \text{ independent of } \mathbf{x} \ . \qquad (2.8.4)$$

Applied to the differential vector $d\mathbf{x}$ this says

$$dx'_i = \Sigma_j \, F_{ij} \, dx_j \ . \qquad (2.8.5)$$

Comparison with (2.1.4) shows that $F_{ij} = R_{ik}$ so

$$\mathbf{F} \text{ linear} \qquad \Rightarrow \qquad R = F \quad \text{and} \quad S = F^{-1} \ . \qquad (2.8.6)$$

We then arrive at this miniature theorem and its converse:

Theorem:

(a) If $R(\mathbf{x}) = R$ does not vary with $\mathbf{x}$, then $d\mathbf{x}' = R d\mathbf{x} \Rightarrow \mathbf{x}' = R\mathbf{x}$ so $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ is a linear transformation, namely $\mathbf{F}(\mathbf{x}) = R\mathbf{x}$.

(b) If $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ is a linear transformation, then $\mathbf{F}(\mathbf{x}) = R\mathbf{x}$ where R is a matrix of constants which do not vary with $\mathbf{x}$. (2.8.7)

This is the situation for global rotations and Lorentz transformations.

2.9 Vectors that are contravariant by definition

A contravariant vector has been defined above as any N-tuple which transforms the same way that d**x** transforms with respect to F, namely, d**x**' = R(**x**) d**x**. One might state this as

$$\{ d\mathbf{x}', d\mathbf{x} \} \qquad d\mathbf{x}' = R(\mathbf{x})\, d\mathbf{x} \qquad \text{contravariant vector} \ . \tag{2.9.1}$$

Suppose we start with an *arbitrary* N-tuple **V** and simply *define* **V**' ≡ R**V**. One would have to conclude that the pair { **V**', **V** } transforms as a contravariant vector.

$$\{ \mathbf{V}', \mathbf{V} \} \qquad \mathbf{V}' \equiv R(\mathbf{x})\mathbf{V} \qquad \text{contravariant vector} \ . \tag{2.9.2}$$

Conversely, one could start with some given **V**' and define **V** ≡ S(x)**V**' (recall $S = R^{-1}$), and again one would conclude that { **V**', **V** } represents a vector that transforms as a contravariant vector.

We refer to either process as producing a vector which is "contravariant by definition". Creating a contravariant vector in this fashion is a fine thing to do, as long as the defined vector does not conflict with something that already exists.

Example 1: We know that if F is non-linear, the vector **x** does *not* transform as a contravariant vector, because **x**' = R(**x**)**x** is not true, where **x**' = **F**(**x**). If we start with **x** and try to force {**x**', **x**} to be "contravariant by definition" by defining **x**' ≡ R(x) **x**, this **x**' conflicts with the existing **x**' = **F**(**x**), so the method of contravariant by definition is unacceptable. (2.9.3)

Example 2: As another example, consider an N-tuple in x'-space of three masses **V**' = (m_1,m_2,m_3). The transformation is taken in this example to be regular rotations. Since masses are rotational scalars with respect to such rotations, we know that in an x-space rotated frame of reference we would find **V** = (m_1,m_2,m_3). We could attempt to set up { **V**', **V** } as a vector that is "contravariant by definition" by defining **V** ≡ S**V**', but this conflicts with the existing fact that **V** = (m_1,m_2,m_3), so the method of contravariant by definition is again unacceptable. (2.9.4)

Example 3: This time F is a general transformation and we start with **V**' = $\mathbf{e}'_n$ which are a set of axis-aligned basis vectors in x'-space. We define vectors **V** = $\mathbf{e}_n$ according to $\mathbf{e}_n \equiv S\mathbf{e}'_n$. Then { $\mathbf{e}'_n$, $\mathbf{e}_n$ } form a vector which is "contravariant by definition" and $\mathbf{e}'_n = R\, \mathbf{e}_n$ ($R = S^{-1}$). Since the newly defined vector $\mathbf{e}_n$ does not conflict with some already-existing vector in x-space, the method of contravariant by definition in this example is acceptable. This is exactly what is done in the next Section with the tangent base vectors $\mathbf{e}_n$. (2.9.5)

2.10 Vector Fields

We considered above vectors like position **x** (and d**x**) and velocity **v** and the vector operator $\overline{\nabla}$, and we referred to a generic vector as **V**. Many vectors of interest (in fact, most) are functions of **x**, which is to say, they are vector fields. Examples are the electric and magnetic fields **E**(**x**) and **B**(**x**), or the average velocity of a small region of fluid **V**(**x**) or a current density **J**(**x**). Another example is the transformation **F**(**x**).

We already mentioned scalar fields, such as temperature $T(\mathbf{x})$ or electrostatic potential $\Phi(\mathbf{x})$. The way a scalar temperature field transforms going from x-space to x'-space is this

$$T'(\mathbf{x}') = T(\mathbf{x}) \qquad \text{where} \qquad \mathbf{x}' = \mathbf{F}(\mathbf{x}) \ . \tag{2.10.1}$$

If the transformation is a 3D rotation from frame S to frame S', then T' is the temperature measured in frame S' at point **x**' and T is the temperature measured at the corresponding point **x** in frame S and of course there is only one temperature at that point so the numbers are equal. In x'-space one needs the prime on T' because the functional form (how T' depends on the x'_i) is not the same as that of T (how T depends on the x_i). For example, if transformation F is from 2D Cartesian to polar coordinates, then

$$T'(r,\theta) = T(x,y) = T(r\cos\theta, r\sin\theta) \neq T(r,\theta) \ . \tag{2.10.2}$$

Contravariant and covariant vector fields transform as described above, but now one must show the argument for each field in its own space, and again $\mathbf{x}' = \mathbf{F}(\mathbf{x})$:

$$\begin{array}{llll} \mathbf{V}'(\mathbf{x}') = R\,\mathbf{V}(\mathbf{x}) & \text{contravariant} & R_{ik}(\mathbf{x}) \equiv (\partial x'_i/\partial x_k) & R = S^{-1} \\ \overline{\mathbf{V}}'(\mathbf{x}') = S^T\,\overline{\mathbf{V}}(\mathbf{x}) & \text{covariant} & S_{ik}(\mathbf{x}') \equiv (\partial x_i/\partial x'_k) = S^T{}_{ki}(\mathbf{x}') \ . & \end{array} \tag{2.10.3}$$

Similar transformation rules apply to tensors of any rank. For example (as we shall see in Chapter 5) the metric tensor g_{ab} (developmental notation) is a rank-2 contravariant tensor field and the transformation rule is this (implied summation on repeated indices),

$$g'_{ab}(\mathbf{x}') = R_{aa'}R_{bb'}g_{a'b'}(\mathbf{x}) \qquad \text{or} \qquad g'_{ab} = R_{aa'}R_{bb'}g_{a'b'} \tag{2.10.4}$$

Often the coordinate dependence of g is suppressed, just as it is for R and S, as shown on the right above.

Jumping momentarily into Standard Notation, in special relativity one has $x'^{\mu} = \Lambda^{\mu}{}_{\nu}x^{\nu}$ where $F = R = \Lambda$ is a linear transformation, and one would then specify the transformation of a contravariant vector field as

$$V'^{\mu}(x'^{\alpha}) = \Lambda^{\mu}{}_{\nu}\, V^{\nu}(x^{\alpha}) \qquad\qquad x'^{\mu} = \Lambda^{\mu}{}_{\nu}x^{\nu} \tag{2.10.5}$$

Comment. The word field in "vector field **V(x)**" is unrelated to the word field being an algebraic object with • and + properties, such as the field of the real numbers or the finite field {0,1}. However, the components of **V** and **x** in **V(x)** are normally elements of the real number field.

2.11 Names and symbols

The matrix $R_{ik}(\mathbf{x}) = (\partial x'_i/\partial x_k)$ is called the **Jacobian matrix** for the transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$, while the matrix $S_{ik}(\mathbf{x}') = (\partial x_i/\partial x'_k)$ is then the Jacobian matrix of the inverse transformation $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$. The determinant of the Jacobian matrix S will be shown in Section 8.6 to have a certain significance, and that determinant is called "**the Jacobian**" $= \det(S(\mathbf{x}')) \equiv J(\mathbf{x}')$.

Matrix R is sometimes called "the differential dF" of transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ and is. Matrix S is then the differential of the inverse transformation $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$.

The author has anguished over what names to give the matrices R and S = R^{-1}. One option was to use R = L, where L stands for the fact that this matrix is describing a Local coordinate system at point **x**, or a Linearized transformation. But L is commonly used for differential operators and angular momentum, so that got rejected. R is often called Λ in special relativity, but why go Greek so early? Another option is to use R = J for Jacobian matrix, but J looks too much like "an integer" or angular momentum or "the Jacobian". T for Transformation might have been confused with the full transformation F, or Cauchy stress T. Our chosen notation R makes one think perhaps R is a Rotation, but that won't in general be the case. For the moment we will continue to use R and S, where recall RS = 1. We shall refer to R simply as "the R matrix for transformation **F**".

The fact that vectors are processed by NxN matrices R and S puts that part of the subject into the field of linear algebra, and that may be the origin of the name **tensor algebra** as a generalization of this idea (tensors as objects of direct product algebras). Of course the differential calculus aspect of the subject is already highly visible, there are ∂ symbols everywhere (hence the name **tensor calculus**).

2.12 Definition of the words "scalar", "vector" and "tensor"

In Section 2.2 a "scalar" was defined as something that is invariant under some transformation F, and this was identified with a "rank-0 tensor". Similarly, a "vector" is either a contravariant vector or a covariant vector and both of these are "rank-1 tensors". In Section 5.6 certain "rank-2" tensors will appear -- they are matrices that transform in a certain way under a transformation F. In Section 7.10 tensors of rank-n will appear, and these are objects with n indices which transform in a certain manner under F.

To be more precise and to provide protection against the vagaries of "the literature", these objects probably should have been defined with the word "tensorial" in front of them.

"tensorial scalar" ≡ rank-0 tensor with respect to some transformation F
"tensorial vector" ≡ rank-1 tensor with respect to some transformation F
"tensorial tensor" ≡ rank-n tensor with respect to some transformation F (2.12.1)

As has been emphasized several times, a "tensorial tensor" is linked to a particular underlying transformation F, and one should really use the more precise term "tensorial tensor under F".

In this paper, we generally omit the word "tensorial" when discussing the above objects. This brings us into conflict with the following definitions which are often used: (Here, we use the term "expression" to indicate a number, a variable, or some combination of same.)

- A "scalar" is a single expression, a 1-tuple. No invariance under any transformation is implied.

- A "vector" is an N-tuple of expressions. No transformation rule is implied.

- A "second order tensor" is a matrix of expressions. No transformation rule is implied. (2.12.2)

- A "tensor" is an object with n indices, n = 2,3,4... which includes the previous item. A tensor is therefore a collection of expressions which are labeled by n indices each of which goes 1 to N. No transformation rule is implied.

To these definitions we can add another list:

- A "scalar field" is a single function of **x** (the x-space coordinates). No implication of invariance.

- A "vector field" is an N-tuple of functions of **x** -- an N-tuple of scalar fields. No transform implied.

- A "tensor field" of order n is a set of N^n scalar functions, for example, $T_{abc...}(\mathbf{x})$. Same. (2.12.3)

In any discussion which includes relativity (special or general), the words scalar, vector and tensor would always imply the tensorial definitions of these words. Continuum mechanics, however, seems to use the above alternate list of definitions, so that any matrix is called a tensor. Usually such matrices are functions of space and should be called tensor fields, but everybody knows what is meant.

In Section 7.15 we shall discuss the notion of an equation being **covariant**, which means it has the exact same form in different frames of reference which are related by a transformation. For example, one might have **F** = m**a** in frame S, and **F**' = m'**a**' in frame S', where these frames are related by a static rotation. **F** and **a** are tensorial vectors with respect to this rotation, and m and m' are tensorial scalars, and m = m' for that reason. Both sides of **F** = m**a** transform as tensorial vectors. Since rotations are an invariance of Newtonian physics, any valid equation of motion *must* be "covariant", and this applies of course to particle, rigid body and continuum mechanics.

In the latter field, continuum mechanics, one naturally seeks out model equations which are covariant. In order to do this properly, one must know which tensors are tensorial tensors, and which tensors are just tensors with either no transformation rule, or some transformation rule that does not match the tensor. Continuum mechanics has evolved special words to handle this situation. If a tensor is a tensorial tensor, it is said to be **objective**, or **indifferent**. In continuum mechanics an equation which is covariant is said to be **frame-indifferent**.

In this document we shall follow the time-honored tradition of being inconsistent in our use of the words scalar, vector and tensor, but the reader is now at least warned. To further promote that tradition, we shall sometimes refer to tensorial tensors as "true" tensors.

The notion of tensor *densities* described in Appendix D further complicates the nomenclature. One can have scalar densities and vector densities of various weights, for example.

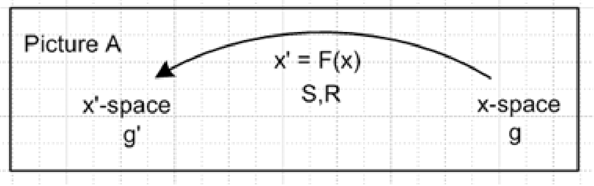
Appendix K explores a few commonly used tensors in continuum mechanics and determines which of these tensors actually transform as tensors (are objective), and which tensors do not transform as tensors (are non-objective).

Comment on "rank". As noted in Section 2.7, some authors refer to a rank-n tensor as an order-n tensor. The word rank has another common use which is unrelated to tensor rank. The rank of a square matrix is the number of linearly independent rows or columns. If an NxN matrix M has rank less than N, then det(M) = 0. Since our R and S matrices do not have zero determinant, they are both of full rank N. The same is true of the metric tensors of Chapter 5 below. Since $g \equiv \det(\overline{g}_{ij}) \neq 0$, g has full rank N.

3. Tangent Base Vectors $\mathbf{e}_n$ and Inverse Tangent Base Vectors $\mathbf{u}'_n$

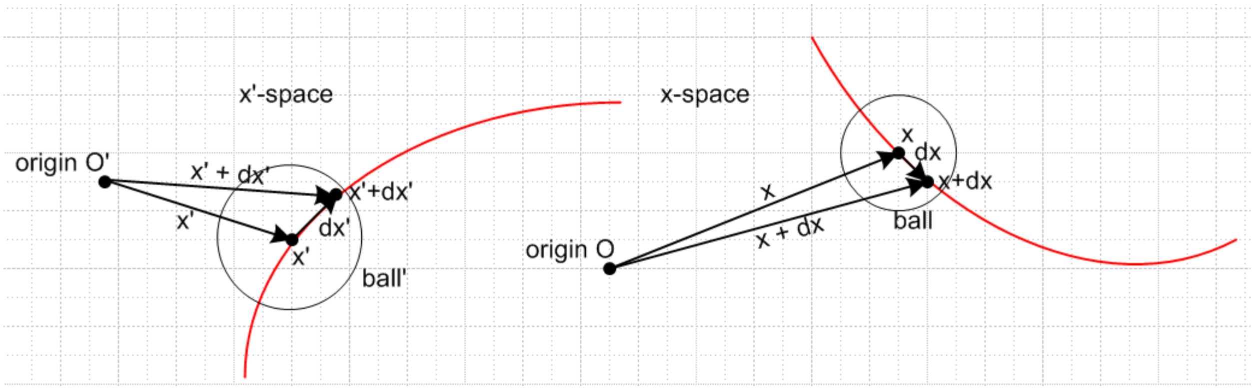
3.1 Differential Displacements

This entire Section is in the context of **Picture A**,



(3.1.1)

In Fig (2.1.2) above showing d**x** and d**x**', one has much freedom to "try out" different differential vectors. For any d**x** one picks at point **x**, one gets some d**x**' according to d**x**' = R(**x**) d**x**. Consider this slightly enhanced version of that figure (red curves added)

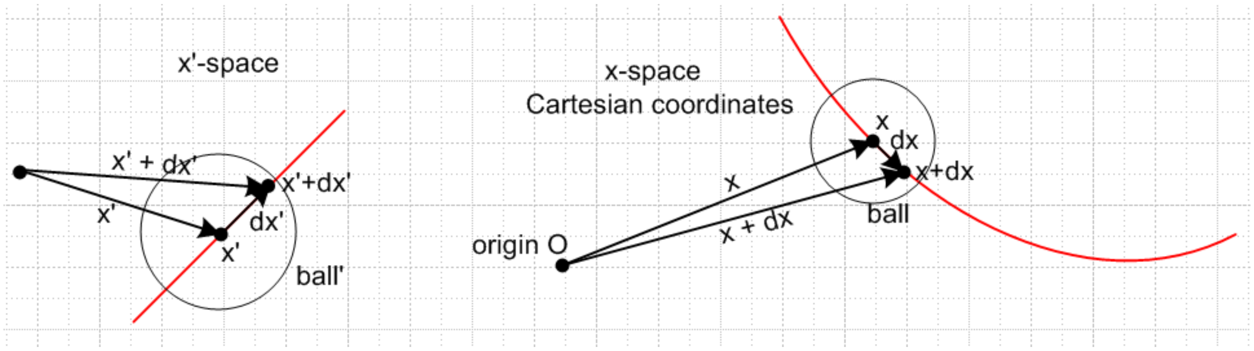


(3.1.2)

The point **x** in x-space (right side) can be regarded as lying on some arbitrary 1-dimensional curve in R^N shown on the right in red. Select d**x** to be the tangent to this curve at point x. That curve will then map into some (probably very different) curve in x'-space which passes through the point **x**'. The tangent to this curve at the point **x**' must be d**x**' = R(**x**) d**x**. A similar statement can be made starting instead with an arbitrary curve in x'-space. The tangent d**x**' there then maps into d**x** = S(**x**') d**x**' in x-space.

The curves are in N-dimensional space and are in general non-planar and the tangents are of course N dimensional tangents, so this 2D picture is mildly misleading.

We now specialize such that the red curve on the left is a straight line parallel to an x'-space axis, which means the curve on the right is a coordinate line,



(3.1.3)

Admittedly the drawing does not strongly suggest that the red line segment on the left is parallel to an axis in x'-space, but since those axes are not drawn, one cannot complain too strenuously.

3.2 Definition of the $\mathbf{e}_n$; the $\mathbf{e}_n$ are the columns of S

First, define a set of N basis vectors in x'-space which point along the positive axes of x'-space,

$$\mathbf{e'}_n \ , \quad n = 1,2...N \qquad (\mathbf{e'}_n)_i = \delta_{n,i} \qquad \mathbf{e'}_1 = (1,0,0...) \text{ etc } . \tag{3.2.1}$$

Assume that the d**x**' arrow above points in this $\mathbf{e'}_n$ direction so that

$$d\mathbf{x}' = \mathbf{e'}_n \, dx'_n \qquad \text{// no implied sum on n} \tag{3.2.2}$$

where dx'_n is a positive differential variation of coordinate x'_n along the $\mathbf{e'}_n$ axis in x'-space. The corresponding d**x** in x-space will be,

$$d\mathbf{x} = \ S \, d\mathbf{x}' = S \, [\mathbf{e'}_n \, dx'_n] = [\ S\mathbf{e'}_n] \, dx'_n \ \equiv \ \mathbf{e}_n \, dx'_n \tag{3.2.3}$$

where this last equality serves as the definition of $\mathbf{e}_n$,

$$\mathbf{e}_n \equiv S\mathbf{e'}_n \ . \tag{3.2.4}$$

Vector $\mathbf{e}_n \ = \mathbf{e}_n(\mathbf{x})$ points along d**x** in x-space and is tangent to the x'_n- coordinate line there at point **x**. This vector $\mathbf{e}_n$ is generally not a unit vector, hence no hat ^ . Writing out the i^{th} component of (3.2.4),

$$(\mathbf{e}_n)_i = \Sigma_j \, S_{ij} \, (\mathbf{e'}_n)_j \ = \ \Sigma_j \, S_{ij} \, \delta_{n,j} \ = \ S_{in} \tag{3.2.5}$$

so that, with (2.1.5),

$$(\mathbf{e}_n)_i = S_{in} = \partial x_i/\partial x'_n \qquad \text{or} \qquad \mathbf{e}_n = \partial\mathbf{x}/\partial x'_n \ = \partial'_n\mathbf{x} \ . \tag{3.2.6}$$

This fact that $(\mathbf{e}_n)_i = S_{in}$ says that the vectors $\mathbf{e}_n$ are the columns of the matrix S:

$$S = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \ .... \ \mathbf{e}_N \ ] \qquad \text{matrix = N columns} \tag{3.2.7}$$

We shall call these $\mathbf{e}_n$ vectors the **tangent base vectors**. The vectors exist in x-space and point along the various coordinate lines that pass through a point $\mathbf{x}$.

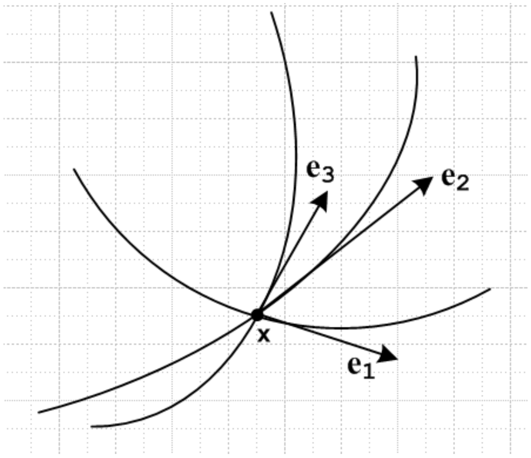
If the points on the x'_n-coordinate line were labeled with the values of x'_n from which they came, one would find that $\mathbf{e}_n$ points in the direction in which those labels increase.

As one moves from $\mathbf{x}$ to some nearby point, the tangent base vectors all change slightly because in general $S = S(\mathbf{x}'(\mathbf{x}))$ and the $\mathbf{e}_n = \mathbf{e}_n(\mathbf{x})$ are the columns of S. Any set of basis vectors which depends on $\mathbf{x}$ in this way is called a local basis. In contrast, the corresponding x'-basis $\mathbf{e}'_n$ shown above with $(\mathbf{e}'_n)_i = \delta_{n,i}$ is a global basis in x'-space since it is the same at any point $\mathbf{x}'$ in x'-space.

Since $\det(S) \neq 0$ due to our assumption that F is invertible, the tangent base vectors are linearly independent and provide a basis for E^N.

One can of course normalize each of the $\mathbf{e}_n$ to be a unit vector $\hat{\mathbf{e}}_n$ according to $\hat{\mathbf{e}}_n = \mathbf{e}_n / |\mathbf{e}_n|$.

Here is a traditional N=3 picture showing the tangent base vectors pointing along three generic coordinate lines in x-space all of which pass through the point $\mathbf{x}$:



(3.2.8)

<u>Comment on notation.</u> Some authors refer to our $\mathbf{e}_n$ as $\mathbf{g}_n$ or $\mathbf{R}_n$ or other. Later it will be shown that $\mathbf{e}_n \bullet \mathbf{e}_m = \overline{g}'_{nm}$ where $\overline{g}'_{nm}$ is the covariant metric tensor for x'-space, so admittedly this provides a reasonable argument for using $\mathbf{g}_n$ so that $\mathbf{g}_n \bullet \mathbf{g}_m = \overline{g}'_{nm}$. But then the primes don't match which is confusing: the $\mathbf{g}_n$ are vectors in x-space, while $\overline{g}'$ is a metric tensor in x'-space. We shall be using yet another g in the form $g = \det(\overline{g}_{nm})$ and a corresponding g'. Due to this proliferation of g objects, we stick with $\mathbf{e}_n$, the notation used by Margenau and Murphy (p 193). A **g**-oriented reader can replace $\mathbf{e} \rightarrow \mathbf{g}$ as needed anywhere in this document. As for unit vector versions of the $\mathbf{e}_n$, we use the notation $\hat{\mathbf{e}}_n \equiv \mathbf{e}_n/|\mathbf{e}_n|$. Morse and Feshbach use $\mathbf{a}_n$ for this purpose (Vol I p 22). A **g**-person might use $\hat{\mathbf{g}}_n$.

A related issue is what symbols to use for the "usual" basis vectors in Cartesian x-space. As noted above in (2.1.8), we are using $\mathbf{u}_n$ with $(\mathbf{u}_n)_i = \delta_{n,i}$ as "axis-aligned basis vectors" in x-space. If $\overline{g} = 1$ for x-space, then these are the usual Cartesian unit vectors (see (6.5.3) below that $\mathbf{u}_n \bullet \mathbf{u}_m = \overline{g}_{nm}$). Many authors use the notation $\mathbf{e}_n$ for *these* vectors which then conflicts with our use of $\mathbf{e}_n$ as the tangent base vectors. Morse and Feshbach use the symbols **i**, **j**, **k** for our Cartesian $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$. Other authors use $\hat{\mathbf{1}}$, $\hat{\mathbf{2}}$, $\hat{\mathbf{3}}$ so then $\mathbf{u}_n = \hat{\mathbf{n}}$.

Often the notation $\mathbf{e}_n$ is used by authors to represent some generic arbitrary set of basis vectors. For this purpose, we shall use the notation $\mathbf{b}_n$.

3.3 $\mathbf{e}_n$ as a contravariant vector

The situation described above is this,

$$d\mathbf{x}' = \mathbf{e}'_n \, dx'_n \qquad \text{x'-space} \qquad \textit{(3.2.2)} \qquad // \text{ no implied sum on n}$$
$$d\mathbf{x} = \mathbf{e}_n \, dx'_n \qquad \text{x-space} \qquad \textit{(3.2.3)} \qquad // \text{ no implied sum on n} \qquad (3.3.1)$$

and the full transformation F maps $d\mathbf{x}$ into $d\mathbf{x}'$. Since $d\mathbf{x}$ is a contravariant vector, the linear transformation R also maps $d\mathbf{x}$ into $d\mathbf{x}'$. Thus

$$d\mathbf{x}' = R(\mathbf{x}) \, d\mathbf{x} \qquad \textit{(2.1.6)}$$

so

$$[\mathbf{e}'_n \, dx'_n] = R(\mathbf{x}) \, [\mathbf{e}_n \, dx'_n] \qquad \textit{(3.3.1)}$$

so

$$\mathbf{e}'_n = R(\mathbf{x}) \, \mathbf{e}_n \ . \qquad (3.3.2)$$

We can regard the last line as a statement that the vector $\mathbf{e}_n$ transforms as a contravariant vector under F. Written out in components one gets

$$(\mathbf{e}'_n)_i = \Sigma_j R_{ij} \, (\mathbf{e}_n)_j \qquad \Rightarrow \qquad \delta_{n,i} = \Sigma_j R_{ij} S_{jn} \qquad (3.3.3)$$

recovering the fact that RS = 1. This is an example of a vector being "contravariant by definition", as discussed in (2.9.5).

The two expansions (3.3.2) and (3.2.4) are easy to verify by showing that the components of both sides are the same:

$$\mathbf{e}'_n \equiv R\mathbf{e}_n = \Sigma_i \, R_{in} \, \mathbf{e}_i \qquad \text{since} \qquad (\mathbf{e}'_n)_j = \Sigma_i \, R_{in} \, (\mathbf{e}_i)_j = \Sigma_i \, R_{in} \, S_{ji} = (SR)_{jn} = \delta_{j,n} = (\mathbf{e}'_n)_j$$

$$\mathbf{e}_n \equiv S\mathbf{e}'_n = \Sigma_i \, S_{in} \, \mathbf{e}'_i \qquad \text{since} \qquad (\mathbf{e}_n)_j = \Sigma_i \, S_{in} \, (\mathbf{e}'_i)_j = \Sigma_i \, S_{in} \, \delta_{i,j} = S_{jn} = (\mathbf{e}_n)_j \qquad (3.3.4)$$

3.4 A semantic question: unit vectors

Above it was noted that $\mathbf{e}'_1 = (1,0,0....)$. Should this be called "a unit vector" ? It will be seen in (6.2.7) that in fact $|\mathbf{e}'_1| = \sqrt{\bar{g}'_{11}} \neq 1$ where $\bar{g}'$ is the covariant metric tensor in x'-space, and $|\mathbf{e}'_1|$ is the covariant length of $\mathbf{e}'_1$. So $\mathbf{e}'_n$ is a unit vector in the sense that it has a single 1 in its column vector definition, but it is not a unit vector in the sense that it does not (in general) have unit magnitude (it would if x'-space were Cartesian with $g'=1$).We take the magnitude = 1 requirement as the proper definition of a unit vector. For this reason, we refer to the $\mathbf{e}'_n$ in x'-space as just "axis-aligned basis vectors" and they have no "hats". One wonders how such a vector should be depicted in a drawing, see Example 1 (b) below and also Section C.5

Example 1: Polar coordinates, tangent base vectors

(a) The first step is to compute the matrix $S_{ik}(\mathbf{x}') \equiv (\partial x_i/\partial x'_k)$ from the inverse equations:

$$\mathbf{x} = (x_1, x_2) = (x,y)$$
$$\mathbf{x}' = (x_1', x_2') = (\theta,r) \qquad \text{// note that r chosen as the second variable } x_2'$$

$$\mathbf{x} = F^{-1}(\mathbf{x}') \quad \leftrightarrow \qquad \begin{array}{l} x = r\cos(\theta) \\ y = r\sin(\theta) \end{array} \qquad \begin{array}{l} x_1 = x_2' \cos(x_1') \\ x_2 = x_2' \sin(x_1') \end{array} \qquad \textit{(1.4)}$$

So

$$S_{11} = (\partial x/\partial\theta) = -r\sin\theta$$
$$S_{12} = (\partial x/\partial r) = \cos\theta \qquad\qquad S_{ik} \equiv (\partial x_i/\partial x'_k)$$
$$S_{21} = (\partial y/\partial\theta) = r\cos\theta$$
$$S_{22} = (\partial y/\partial r) = \sin\theta$$

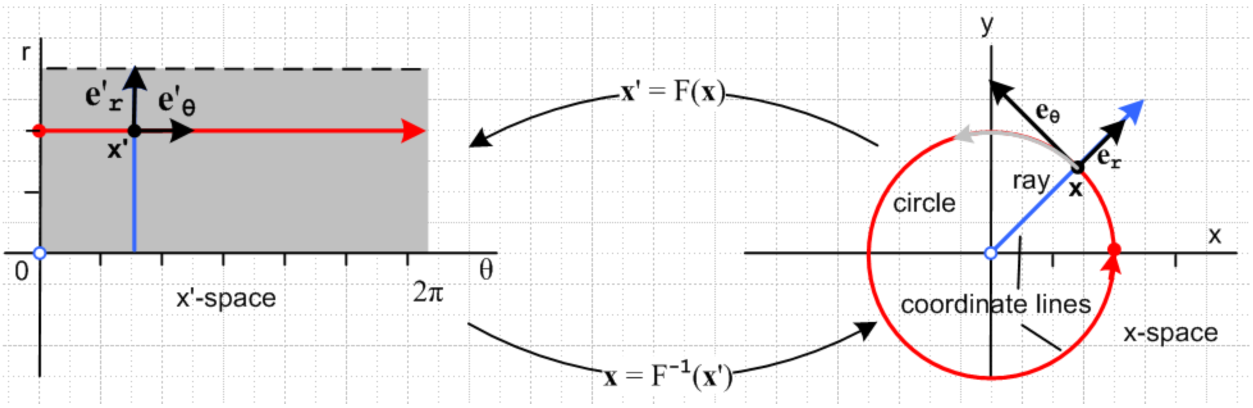
$$S = \begin{pmatrix} -r\sin\theta & \cos\theta \\ r\cos\theta & \sin\theta \end{pmatrix} \qquad \text{//} \quad \det(S) = -r \qquad R = S^{-1} = \begin{pmatrix} -\sin\theta/r & \cos\theta/r \\ \cos(\theta) & \sin\theta \end{pmatrix} . \qquad (3.4.1)$$

[Note: The above S and R are stated for the ordering 1,2 = θ,r . For the more usual ordering 1,2 = r,θ the colums of S should be swapped, and the rows of R should be swapped. In the usual ordering, det(S) = +r.]

The tangent base vectors $\mathbf{e_n}$ can be read off as the columns of S according to (3.2.7),

$$\mathbf{e_1} = r(-\sin\theta,\cos\theta) \quad = \mathbf{e_\theta} \quad = r\,\hat{\mathbf{e}}_\theta \qquad\qquad \text{//} = r\,\hat{\boldsymbol{\theta}}$$
$$\mathbf{e_2} = (\cos\theta,\sin\theta) \quad = \mathbf{e_r} \quad = \hat{\mathbf{e}}_r \; . \qquad\qquad \text{//} = \hat{\mathbf{r}} \qquad (3.4.2)$$

Notice that $\mathbf{e_\theta}$ in this case is not a unit vector. Below is a properly scaled drawing showing the location of the two x'-space basis vectors on the left, and the two tangent base vectors on the right. As just shown, the length of $\mathbf{e_r}$ is 1, while the length of $\mathbf{e_\theta}$ is 2.



(3.4.3)

The tangent base vectors are fairly familiar animals, since $\mathbf{e_r} = \hat{\mathbf{r}}$ and $\mathbf{e_\theta} = r\,\hat{\boldsymbol{\theta}}$ in usual parlance. If one moves radially outward from point **x**, the $\mathbf{e_r}$ base vector stays the same, but $\mathbf{e_\theta}$ grows longer. If one moves azimuthally from **x** to some larger angle θ+Δθ, both vectors stay the same length but they rotate together staying perpendicular.

(b) This is a good place to point out that vectors drawn in a non-Cartesian space can have magnitudes which do not equal the length of the drawn arrows. The "graphical arrow length" of a vector **v** is $(v_x^2 + v_y^2)^{1/2}$, but that is not the right expression for |**v**| in a non-Cartesian space. For example, as will be shown below in (5.10.5), $|\mathbf{e}_\theta'| = |\mathbf{e}_\theta|$, so the magnitude of the vector $\mathbf{e}'_\theta$ shown on the left above is in fact $|\mathbf{e}_\theta'| = r = 2$ and not 1, but the graphical length of the arrow is 1 since $\mathbf{e}'_\theta = (1,0)$. See Section C.5 for further discussion of this topic with a specific 2D non-orthogonal coordinate system.

(c) In this example, two basis vectors $\mathbf{e}'_n$ in x'-space on the left map into the two $\mathbf{e}_n$ vectors on the right according to $\mathbf{e}_n \equiv S\mathbf{e}'_n$. If one were to apply the *full* mapping $\mathbf{x} = F^{-1}(\mathbf{x}')$ to each point along the arrows $\mathbf{e}'_n$, for some general non-linear F one would find that these arrows map into warped arrows on the right whose bases are tangent to those of the $\mathbf{e}_n$. Those warped arrows lie on the coordinate lines. For this particular mapping, $\mathbf{e}'_\theta$ maps under F^{-1} into the warped gray arrow, while $\mathbf{e}'_r$ maps into $\mathbf{e}_r$.

Example 2: Spherical Coordinates, tangent base vectors

$$\mathbf{x} = (x_1, x_2, x_3) = (x,y,z)$$
$$\mathbf{x}' = (x_1', x_2', x_3') = (r,\theta,\varphi)$$

$$\mathbf{x} = F^{-1}(\mathbf{x}') \quad \leftrightarrow \quad \begin{aligned} x &= r\sin\theta\cos\varphi \\ y &= r\sin\theta\sin\varphi \\ z &= r\cos\theta \end{aligned} \qquad \textit{(1.6)}$$

$$S_{ik} \equiv (\partial x_i/\partial x'_k)$$

$$\begin{aligned}
S_{11} &= (\partial x/\partial r) = \sin\theta\cos\varphi \\
S_{12} &= (\partial x/\partial \theta) = r\cos\theta\cos\varphi \\
S_{13} &= (\partial x/\partial \varphi) = -r\sin\theta\sin\varphi \\
S_{21} &= (\partial y/\partial r) = \sin\theta\sin\varphi \\
S_{22} &= (\partial y/\partial \theta) = r\cos\theta\sin\varphi \\
S_{23} &= (\partial y/\partial \varphi) = r\sin\theta\cos\varphi \\
S_{31} &= (\partial z/\partial r) = \cos\theta \\
S_{32} &= (\partial z/\partial \theta) = -r\sin\theta \\
S_{33} &= (\partial z/\partial \varphi) = 0
\end{aligned} \qquad (3.4.4)$$

$$S = \begin{pmatrix} \sin\theta\cos\varphi & r\cos\theta\cos\varphi & -r\sin\theta\sin\varphi \\ \sin\theta\sin\varphi & r\cos\theta\sin\varphi & r\sin\theta\cos\varphi \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix} \qquad R = \begin{pmatrix} \sin\theta\cos\varphi & \sin\theta\sin\varphi & \cos\theta \\ \cos\theta\cos\varphi/r & \cos\theta\sin\varphi/r & -\sin\theta/r \\ -\sin\varphi/(r\sin\theta) & \cos\varphi/(r\sin\theta) & 0 \end{pmatrix}$$

where Maple computes R as S^{-1} and finds as well that: $\det(S) = r^2 \sin\theta$.

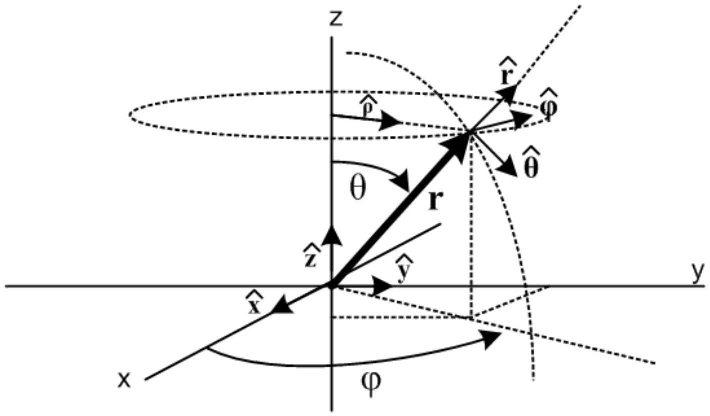
Again, from (3.2.7) the tangent base vectors are the columns of S, so

$\mathbf{e}_r = (\sin\theta\cos\varphi, \sin\theta\sin\varphi,\cos\theta) \qquad |e_r| = 1 \equiv h'_r$
$e_\theta = r(\cos\theta\cos\varphi,\cos\theta\sin\varphi,-\sin\theta) \qquad |e_\theta| = r \equiv h'_\theta$
$e_\varphi = r\sin\theta(-\sin\varphi,\cos\varphi,0) \qquad |e_\varphi| = r\sin\theta \equiv h'_\varphi$ (3.4.5)

and unit vector versions are then

$\hat{\mathbf{e}}_r = (\sin\theta\cos\varphi, \sin\theta\sin\varphi,\cos\theta) = \hat{\mathbf{r}} \qquad \mathbf{e}_r = \hat{\mathbf{r}}$
$\hat{e}_\theta = (\cos\theta\cos\varphi,\cos\theta\sin\varphi,-\sin\theta) = \hat{\boldsymbol{\theta}} \qquad \mathbf{e}_\theta = r\,\hat{\boldsymbol{\theta}}$
$\hat{e}_\varphi = (-\sin\varphi,\cos\varphi,0) = \hat{\boldsymbol{\varphi}} \qquad \mathbf{e}_\varphi = r\sin\theta\,\hat{\boldsymbol{\varphi}}$ (3.4.6)

The unit vectors can be displayed in this standard picture,



(3.4.7)

Notice that $(\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\varphi}}) = (\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3)$ form a right-handed coordinate system at the point $\mathbf{x} = \mathbf{r}$.

3.5 The inverse tangent base vectors $\mathbf{u'_n}$ and inverse coordinate lines

A complete swap $\mathbf{x}' \leftrightarrow \mathbf{x}$ for a mapping $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ of course produces the "inverse mapping". This has the effect of causing R ↔ S in the above discussion. The tangent base vectors for the inverse mapping would then be the columns of matrix R instead of S. We shall denote these inverse tangent base vectors which exist in x'-space by the symbol $\mathbf{u'_n}$. Then:

$(\mathbf{e_n})_i = S_{in} = \partial x_i/\partial x'_n$ // the tangent base vectors as above *(3.2.6)*
$S = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \; .... \; \mathbf{e}_N]$ // are the columns of S *(3.2.7)*

$(\mathbf{u'_n})_i = R_{in} = \partial x'_i/\partial x_n$ // inverse tangent base vectors
$R = [\mathbf{u'}_1, \mathbf{u'}_2, \mathbf{u'}_3 \; .... \; \mathbf{u'}_N]$ // are the columns of R (3.5.1)

By varying only x_n in x-space holding all the other x_i = constant, one generates the x_n-coordinate lines in x'-space, just the reverse of the earlier discussion of this subject. Then inverse tangent base vectors $\mathbf{u'_n}$ will then be tangent to these inverse coordinate lines. An example is given just below and another appears in Appendix C.

The vector $\mathbf{e_n}$ was shown to transform as a contravariant vector into an axis-aligned basis vector $\mathbf{e'_n}$ in x'-space

$$\overset{(3.3.2)}{\mathbf{e'_n} = R\,\mathbf{e_n}} \qquad \overset{(3.3.3)}{(\mathbf{e'_n})_i = \Sigma_j R_{ij}\,(\mathbf{e_n})_j} \qquad \overset{(3.2.6)}{(\mathbf{e_n})_i = S_{in}} \qquad \overset{(3.2.1)}{(\mathbf{e'_n})_i = \delta_{n,i}} \qquad (3.5.2)$$

The same thing happens here, only in reverse :

$$\mathbf{u_n} = S\,\mathbf{u'_n} \qquad (\mathbf{u_n})_i = \Sigma_j S_{ij}\,(\mathbf{u'_n})_j \qquad (\mathbf{u'_n})_i = R_{in} \qquad (\mathbf{u_n})_i = \delta_{n,i} \qquad (3.5.3)$$

where now the $\mathbf{u_n}$ are axis-aligned basis vectors in x-space. A prime on an object indicates which space it inhabits.

The inverse tangent base vectors $\mathbf{u'_n}$ are not the same as the reciprocal base vectors $\mathbf{E_n}$ introduced in Chapter 6 below.

Example 1: Polar coordinates: inverse tangent base vectors and inverse coordinate lines

It was shown earlier for polar coordinates that

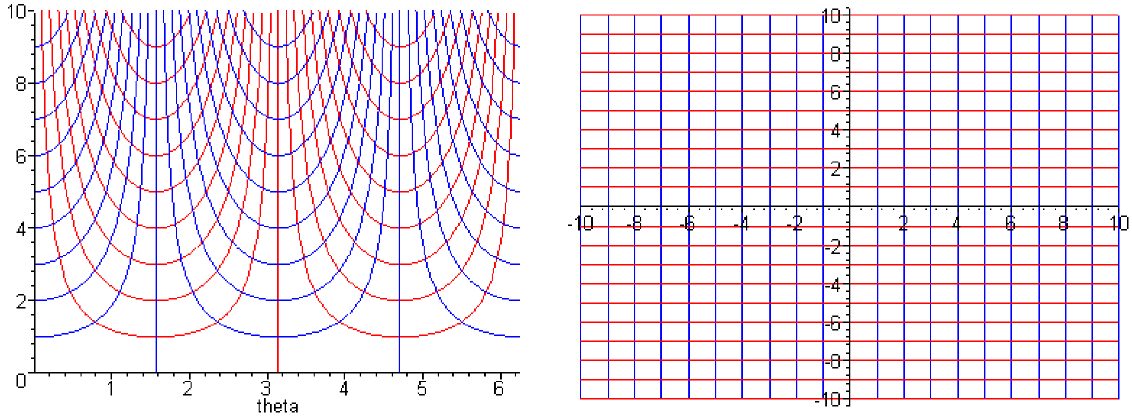
$$R = S^{-1} = \begin{pmatrix} -\sin\theta/r & \cos\theta/r \\ \cos(\theta) & \sin\theta \end{pmatrix} \qquad (3.4.1)$$

so the inverse tangent base vectors (expressed here as row vectors as usual to save space) are given by the columns of R as per (3.5.1),

$$\mathbf{u'_x} = (\,-\sin\theta/r, \cos\theta) \qquad \text{// note near } \theta = 0 \text{ that } \mathbf{u'_x} \text{ indicates a large negative slope}$$
$$\mathbf{u'_y} = (\cos\theta/r, \sin\theta)\;. \qquad \text{// note near } \theta = 0 \text{ that } \mathbf{u'_y} \text{ indicates a small positive slope} \qquad (3.5.4)$$

One expects $\mathbf{u'_x}$ to be tangent to an inverse coordinate line in x'-space which maps to a line in x-space along which only x is varying, which is a horizontal line at fixed y (red below). Looking at the small θ region of the left graph in Fig (3.5.5) below, one sees slopes as just described above.

For the polar coordinates mapping discussed near Fig (3.4.3), horizontal (red) and vertical (blue) lines in x'-space mapped into circles (red) and rays (blue) in x-space, and the tangent base vectors in x-space were tangent to the coordinate lines there. If one instead takes horizontal (red) and vertical (blue) lines in x-space and maps them back into coordinate lines in x'-space, the picture is a bit more complicated. Since $y = r\sin\theta$, the plot of an x-coordinate line (x is varying, y fixed at y_i) in x'-space has the form $r = y_i/\sin\theta$, where y_i denotes some selected y value (a red horizontal line), so plotting $r = y_i/\sin\theta$ in x'-space for various values of y_i displays a set of inverse x-coordinate lines (red). Similarly $r = x_i/\cos\theta$ gives some y-coordinate lines (blue). Here is a Maple plot:



x'-space (θ,r) x-space (x,y)

(3.5.5)

Another example is given in Appendix C.

4. Notions of length, distance and scalar product in Cartesian Space

This Section can be interpreted in either **Picture B** or **Picture D** where the x-space is Cartesian, G=1.

Up to this point, we have dealt only with the *vector space* R^N (a vector space is sometimes called a linear space), and have not "endowed" it with a norm, metric or a scalar product. Quantities like $d\mathbf{x}_i$ above were just little vectors and $\mathbf{x} + d\mathbf{x}$ was vector addition.

Now, for the first time (officially), we discuss length and distance, such as they are in a Cartesian Space, as defined in Chapter 1.

For R^N one first defines a norm which determines the "length" of a vector, the first notion of distance in a limited sense. The "usual" norm is the L^2 norm given by

$$\text{norm of } \mathbf{x} = ||\ \mathbf{x}\ || \equiv (\ x_1^2 + x_2^2 + ... + x_N^2\ )^{1/2} \equiv |\ \mathbf{x}\ | \ . \qquad (4.1)$$

Now we have a normed linear space.

One next defines the notion of the distance *between* two vectors. Although this can be done in many ways, just as there are many possible norms, for R^N the "natural metric" is defined in terms of the above L^2 norm, so that

$$\begin{aligned} \text{distance between } \mathbf{x} \text{ and } \mathbf{y} = \text{metric} = d(\mathbf{x},\mathbf{y}) &\equiv ||\ \mathbf{x} - \mathbf{y}\ || \\ &= (\ [x_1-y_1]^2 + [x_2-y_2]^2 + ... + [x_N-y_N]^2\ )^{1/2} \ . \end{aligned} \qquad (4.2)$$

Now our space is both a normed linear space and a metric space, a combo known as a Banach Space.

One finally adds the notion of a scalar product (inner product) in this way

$$(x,y) \equiv \Sigma_i x_i y_i \equiv \mathbf{x} \bullet \mathbf{y} \qquad // = \Sigma_{ij}\ \delta_{i,j}\ x_i\ y_j \qquad (4.3)$$

which of course implies this special case,

$$(x,x) = \mathbf{x} \bullet \mathbf{x} = \Sigma_i x_i^2 = ||x||^2 = |\ \mathbf{x}\ |^2 \ . \qquad (4.4)$$

Our space has now ascended to the higher level of being a real Hilbert Space of N dimensions. All this structure is implied by the notation R^N, our "Cartesian Space".

The length of the vector $d\mathbf{x}$ in R^N is given by

$$\text{length of } d\mathbf{x} = \text{distance between vectors } \mathbf{x}+d\mathbf{x} \text{ and } \mathbf{x} \equiv ds \equiv ||\ d\mathbf{x}\ || = \sqrt{\Sigma_i (dx_i)^2} \ . \qquad (4.5)$$

To avoid dealing with the square root, one usually writes

$$\begin{aligned} (ds)^2 &\equiv ||\ d\mathbf{x}\ ||^2 = \Sigma_i (dx_i)^2 = (dx_1)^2 + (dx_2)^2 + ... + (dx_N)^2 \\ &= \Sigma_i\ dx_i\ dx_i = \Sigma_{ij}\ \delta_{i,j}\ dx_i\ dx_j \ . \end{aligned} \qquad (4.6)$$

As shown in the next Section, one can interpret $\delta_{i,j}$ as the metric tensor in Cartesian Space.

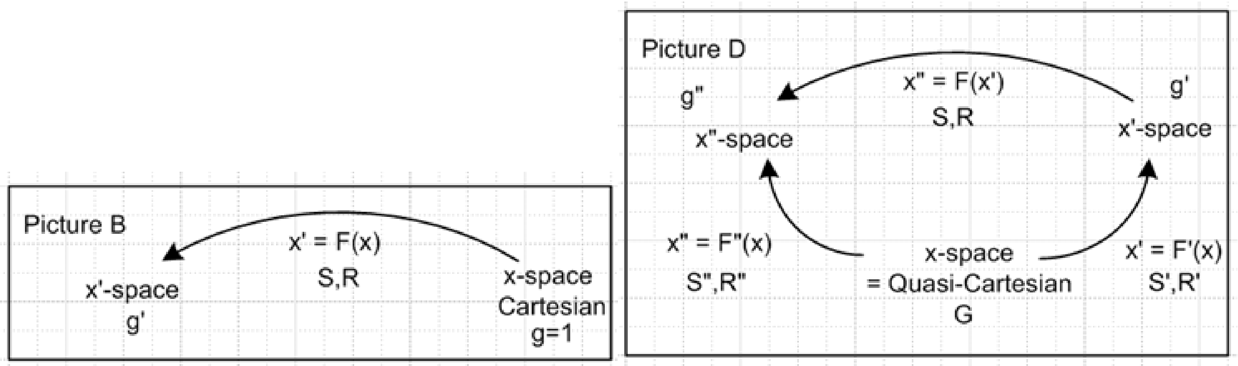
The cursory discussion of this Section is fleshed out in Chapter 2 of Stakgold where the concepts of linear spaces, norms, metrics and inner products are defined with precision. Stakgold compares our N dimensional Cartesian Hilbert Space to the N=∞ dimensional Hilbert Spaces used in functional analysis, where basis vectors might be Legendre polynomials $P_n(z)$ on (-1,1), n = 0,1,2...∞. He has little to say, however, about curvilinear coordinate spaces in this particular book.

5. The Metric Tensor

5.1 The Picture D Context

The metric tensor is the heart of the machine of tensor analysis and we shall have a lot to say about it in this Chapter. Each Section is best presented in the context of one of our Pictures, and there will be some jumping around between pictures. We apologize for this inconvenience and ask forbearance. Hopefully the Sections below will give the reader some experience with typical nitty-gritty manipulations. One advantage of the developmental notation over the standard notation is that matrix methods are easy to use, and they will be used below. Unless otherwise specified, repeated indices are implicitly summed over (**Einstein convention**). But sometimes we do show sum symbols Σ where extra clarity is needed.

We now go to the **Picture D** context. Comparison with Picture B shows that primes must be placed on objects F, R and S related to the transformation from x-space to x'-space:



(5.1.1)

The various partial derivatives are determined from their definitions,

$$R'_{ik} \equiv (\partial x'_i/\partial x_k) \qquad R''_{ik} \equiv (\partial x''_i/\partial x_k) \qquad R_{ik} \equiv (\partial x''_i/\partial x'_k)$$

$$S'_{ik} \equiv (\partial x_i/\partial x'_k) \qquad S''_{ik} \equiv (\partial x_i/\partial x''_k) \qquad S_{ik} \equiv (\partial x'_i/\partial x''_k) \ . \qquad (5.1.2)$$

The unprimed S,R can be expressed in terms of the primed objects this way (chain rule, implied Σ on a),

$$R_{ik} \equiv (\partial x''_i/\partial x'_k) = (\partial x''_i/\partial x_a)(\partial x_a/\partial x'_k) = R''_{ia} S'_{ak} \quad \Rightarrow \quad R = R'' S'$$

$$S_{ik} \equiv (\partial x'_i/\partial x''_k) = (\partial x'_i/\partial x_a)(\partial x_a/\partial x''_k) = R'_{ia} S''_{ak} \quad \Rightarrow \quad S = R' S'' \ . \qquad (5.1.3)$$

5.2 Definition of the metric tensor

The metric or distance between vectors **x** and **x**+d**x** can be specified as done in Chapter 4 in terms of the norm of differential vector d**x**,

$$\text{metric}(\mathbf{x}+d\mathbf{x}, \mathbf{x}) = \text{norm}([\mathbf{x}+d\mathbf{x}] - \mathbf{x}) = \text{norm}(d\mathbf{x}) \equiv |\, d\mathbf{x}\,| \equiv ds \ , \qquad (5.2.1)$$

with the caveat that for non-Cartesian spaces this may not be an official norm, see Section 5.10 below. The squared distance $(ds)^2$ must be a linear combination of products $dx_i dx_j$ just on dimensional grounds. The coefficients in this linear combination form a matrix called the **metric tensor** (later we show this matrix really is a tensor). Here for an N-dimensional space,

$$(ds)^2 = \Sigma_{i=1}{}^N \Sigma_{j=1}{}^N \text{ [metric tensor]}_{ij}\, dx_i\, dx_j . \tag{5.2.2}$$

One "endows" a space with a certain metric tensor, and this in turn determines the distance ds = norm(d**x**) for any differential vector d**x** in that space. A metric tensor is specific to a space; it is a property of the space; it is part of the space's definition.

Recall from the text near (1.10) that our special "quasi-Cartesian" space has a *diagonal* metric tensor G whose diagonal elements are independently either +1 or -1. Then distance ds is determined by

$$(ds)^2 = \Sigma_i\Sigma_j[G_{ij}]dx_idx_j = \Sigma_i\Sigma_j[G_{ii}\delta_{i,j}]dx_idx_j = \Sigma_i G_{ii}\, dx_idx_i = \Sigma_i G_{ii}\, (dx_i)^2. \tag{5.2.3}$$

How might one express this same ds in terms of the the coordinates x'_i of x'-space in Picture D? We found in (2.1.6) that, for Picture A of Fig (2.1.1), d**x** = S(**x**')d**x**' where S is a matrix which represents a general transformation **x**' = **F**(**x**) locally at a point. Translating this to the right side of Picture D, we have d**x** = S'(**x**') d**x**' since S is primed there (as is F and R). Therefore

$$(ds)^2 = \Sigma_iG_{ii}dx_idx_i = \Sigma_i G_{ii}\ (\Sigma_kS'_{ik}\, dx'_k)\ (\Sigma_mS'_{im}\, dx'_m)$$

$$= \Sigma_k\Sigma_m \ \{ \ \Sigma_i G_{ii}\ \ S'_{ik}S'_{im} \ \} \ dx'_k\, dx'_m \quad . \tag{5.2.4}$$

Now let's start all over again and define a metric tensor in x'-space just as we did in x-space,

$$(ds')^2 = \Sigma_k\Sigma_m \text{ [metric tensor]}'_{km}\, dx'_k\, dx'_m .$$

$$= \Sigma_k\Sigma_m \ \overline{g}'_{km}\, dx'_k\, dx'_m . \tag{5.2.5}$$

We call the metric tensor in x'-space g' because x' is primed, and we use an overbar because it will be shown below that $\overline{g}'_{km}$ are the components of a covariant rank-2 tensor $\overline{g}'$ in x'-space.

We now make an important hypothesis: the length of vector d**x** is the same as the length of the vector d**x**', which is to say, we assume that ds = ds'. We are assuming that ds is a scalar under transformation F', so

$$(ds)^2 = (ds')^2 . \qquad \text{// hypothesis, } (ds)^2 \text{ is a scalar} \tag{5.2.6}$$

Since equations (5.2.4) and (5.2.5) must be valid for *any* choice of d**x**', comparison of the two equations yields this relationship between the two metric tensors $\overline{g}'$ and G :

$$\overline{g}'_{km} \equiv \{ \ \Sigma_iG_{ii}S'_{ik}S'_{im} \ \} \ . \tag{5.2.7}$$

Using $G_{ij} = G_{ii}\delta_{i,j}$ we obtain the matrix relationship between metric tensors $\overline{g}'$ and G :

$$\overline{g}'_{km} \equiv \{ \Sigma_i G_{ii} S'_{ik} S'_{im} \} = \Sigma_i \Sigma_j \; S'^{T}_{ki} G_{ij} S'_{jm} \qquad \Rightarrow$$

$$\overline{g}' = S'^{T} G \, S' \, . \qquad (5.2.8)$$

We now repeat all the above steps for the transformation from x-space to x"-space on the left side of Picture D in Fig (5.1.1). We make the hypothesis that $(ds)^2 = (ds")^2$ and in doing so, we obtain a relationship between the metric tensor G for x-space and the metric tensor $\overline{g}"$ for x"-space:

$$\overline{g}" = S"^{T} G \, S" \, . \qquad (5.2.9)$$

To summarize, there are three metric tensors for the three spaces in Picture D :

$$\overline{g} = G \qquad \overline{g}' = S'^{T} \, G \, S' \qquad \overline{g}" = S"^{T} \, G \, S" \, . \qquad (5.2.10)$$

where we have made the hypothesis that ds = ds' = ds".

<u>Concerning the invariance of (ds).</u> In the above discussion, it was assumed in (5.2.6) that distance $(ds)^2$ is the same in x'-space as it is in x-space. As shown below in Section 5.10, this is part of a larger hypothesis that the covariant dot product of any two vectors gives the same number regardless of which space is used to compute the dot product: **A** • **B** = **A'** • **B'**. This in turn implies that |**A**| = |**A'**| and in particular |d**x**| = |d**x'**| or ds = ds'.

In our major application, where x-space is Cartesian and x'-space is that of some curvilinear coordinates, it is a *requirement* that | **A** | = | **A'** |. The length of a vector in Cartesian physical space does not change simply because we choose to express that length in some curvilinear coordinates. Imagine that **A** is a velocity vector **v**. The speed |**v**| of an object is the same number whether one represents **v** in Cartesian or spherical coordinates.

In special relativity it is a very well-verified *hypothesis* that dot products are scalars and that $(ds)^2$ is a scalar under Lorentz transformations. In this context, ds is often written dτ (the so-called proper time).

There are, however, applications of transformations where the scalarity of $(ds)^2$ is not valid and in fact it is crucial that $(ds)^2$ can change under a transformation. For example, in continuum mechanics one can think of the flow of a tiny bit of continuous matter as being modeled by a transformation $\mathbf{x} = \mathcal{F}(\mathbf{x}',t)$. In general $\mathcal{F}$ ($= F^{-1}$) is non-linear. At time t = 0 the geometry of this little blob is described by x'-space coordinates, and after a flow at time t = t it is described by x-space coordinates. If once traces during this flow a little "dumbbell" vector between two very closely spaced particles in the blob, one finds that d**x'** at time t = 0 becomes d**x** = Sd**x'** at time t = t where S is not simply a rotation. The whole point here is that during the flow, the distance vector between two close particles rotates and stretches in some manner, and in general (due to this stretch) |d**x**| ≠ |d**x'**|, so the *geometric* $(ds)^2$ is definitely not invariant under the flow (ie, under the transformation $\mathcal{F}$). On the other hand, as we shall see below, a certain *curvilinear* $(ds)^2$ does remain invariant under the flow. The flow application is considered more in Section 5.16 below, and still further in Appendix K.

5.3 Inverse of the metric tensor

The inverses of the three metric tensors in (5.2.10) shall be indicates *without* an overbar, and we shall eventually show these matrices to be "contravariant" matrices and thus deserve no overbar. We thus now define three new g matrices as these inverses, and compute the inverses:

$$g \equiv \bar{g}^{-1} = G^{-1} = G \qquad \text{// remember G just has +1 and -1 diagonal elements}$$

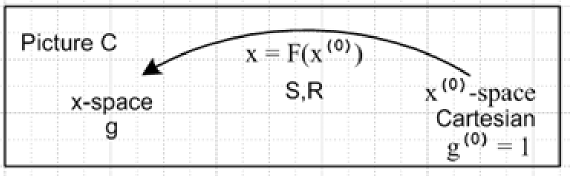
$$g' \equiv \bar{g}'^{-1} = (S'^T G S')^{-1} = S'^{-1} G (S'^T)^{-1} = R' G R'^T$$

$$g'' \equiv \bar{g}''^{-1} = (S''^T G S'')^{-1} = S''^{-1} G (S''^T)^{-1} = R'' G R''^T . \qquad (5.3.1)$$

Here are the collected facts from above:

$$\begin{array}{llll} g = G & g' = R'G R'^T & g'' = R'' G R''^T & S = R' S'' \\ \bar{g} = G & \bar{g}' = S'^T G S' & \bar{g}'' = S''^T G S'' & R = R'' S' \\ \bar{g}g = 1 & \bar{g}'g' = 1 & \bar{g}''g = 1 & \end{array} \qquad (5.3.2)$$

<u>Comment:</u> In the **Picture C** context but with a Quasi-Cartesian $x^{(0)}$-space ($g^{(0)}$=G), one could take the second column above and write it this way,



$$g = RGR^T$$
$$\bar{g} = S^T GS$$
$$\bar{g}g = 1 \qquad (ds)^2 = \bar{g}_{km}\, dx_k\, dx_m \qquad (5.3.3)$$

where now the clutter of primes is gone. If $x^{(0)}$-space is Cartesian so $g^{(0)} = G = 1$, then

$$g = RR^T$$
$$\bar{g} = S^T S. \qquad \text{// if } x^{(0)}\text{-space is Cartesian in Picture C} \qquad (5.3.4)$$

But we continue with **Picture D** shown in (5.1.1).

5.4 A metric tensor is symmetric

Consider the following, where A is any matrix and D is a diagonal matrix :

$$N = ADA^T \quad \Rightarrow \quad N^T = (ADA^T)^T = AD^T A^T = ADA^T = N \quad \Rightarrow \ N \text{ symmetric} \qquad (5.4.1)$$

Replacing A with A^T and N with M gives

$$M = A^T DA \quad \Rightarrow \quad M^T = (A^T DA)^T = A^T D^T (A^T)^T \ = \ A^T DA = M \ \Rightarrow M \text{ symmetric} \qquad (5.4.2)$$

Looking at (5.3.3), since g has the form of N and $\overline{g}$ the form of M with G diagonal, we conclude that both g and $\overline{g}$ are symmetric matrices, and the same is of course true for g' and g". Thus

$$\begin{array}{ll} g_{ab} = g_{ba} & \overline{g}_{ab} = \overline{g}_{ba} \\ g'_{ab} = g'_{ba} & \overline{g}'_{ab} = \overline{g}'_{ba} \end{array} \qquad (5.4.3)$$

5.5 det(g) and g_{nn} of a Cartesian-generated metric tensor are non-negative

If we arrive at x'-space by a transformation F from a Cartesian x-space (as opposed to a Quasi-Cartesian one), we refer to the metric tensor g' in this x'-space as being "Cartesian generated". In this case G = 1 and the metric tensors above are $g = RR^T$ and $\overline{g} = S^T S$ as in (5.3.4). Any matrix of either of these forms has positive diagonal elements and positive determinant:

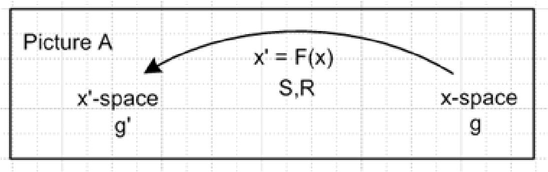
$$(A^T A)_{aa} = \Sigma_b \ (A^T)_{ab} A_{ba} = \Sigma_b \ (A)_{ba} A_{ba} = \Sigma_b \ (A_{ba})^2 \ \geq 0 \qquad \text{// diagonal elements} \geq 0$$

$$\det(A^T A) \ = \det(A^T) \det(A) = \det(A) \det(A) \ = [\ \det(A)\ ]^2 \ \geq 0 \qquad \text{// det} \geq 0 \ . \qquad (5.5.1)$$

To show these results for the AA^T form, just replace $A \rightarrow A^T$ everywhere. Recall that transformation F maps $R^N \rightarrow R^N$ so the coefficients of the linearized matrices R and S are real, and elements of the metric tensor must therefore also be real. For a Quasi-Cartesian-generated metric tensor, these proofs are invalid since then $g = RGR^T$ and $\overline{g} = S^T GS$ and $G \neq 1$.

5.6 Definition of two kinds of rank-2 tensors

We now switch to **Picture A**,



(5.6.1)

Recall the vector transformation rules from (2.5.1),

$$\begin{array}{lllll} \mathbf{V}' = R\,\mathbf{V} & \text{contravariant} & R_{ik}(\mathbf{x}) \equiv (\partial x'_i/\partial x_k) & & R = S^{-1} \\ \overline{\mathbf{V}}' = S^T\,\overline{\mathbf{V}} & \text{covariant} & S_{ik}(\mathbf{x}') \equiv (\partial x_i/\partial x'_k) & = S^T_{ki}(\mathbf{x}') & \end{array} \qquad \mathit{(2.5.1)}$$

which can be written out in components (implied sum on a'),

$$V'_a = R_{aa'} V_{a'} \qquad \text{contravariant} \qquad R_{ik}(\mathbf{x}) \equiv (\partial x'_i/\partial x_k) \qquad R = S^{-1}$$
$$\bar{V}'_a = S^T_{aa'}\bar{V}_{a'} \qquad \text{covariant} \qquad S_{ik}(\mathbf{x}') \equiv (\partial x_i/\partial x'_k) \quad = \quad S^T_{ki}(\mathbf{x}') \,. \qquad (5.6.2)$$

A rank-1 tensor is defined to be a vector which transforms in one of the two ways shown above. Similarly, a (non-mixed) **rank-2 tensor** is defined as a matrix which transforms in one of these two ways:

$$M'_{ab} = R_{aa'} R_{bb'} M_{a'b'} \qquad \text{// contravariant rank-2 tensor}$$
$$\bar{M}'_{ab} = S^T_{aa'} S^T_{bb'} \bar{M}_{a'b'} \qquad \text{// covariant rank-2 tensor} \qquad (5.6.3)$$

and again we put a bar over the covariant objects.

Digression**:** Proof that $(A^{-1})^T = (A^T)^{-1}$ for any invertible matrix A:

(5.6.4)

- $\det(A) = \det(A^T)$
- $\mathrm{cof}(A^T) = [\,\mathrm{cof}(A)]^T$ since $[\mathrm{cof}(A^T)]_{ab} = \mathrm{cof}\,(\,A^T_{ab}) = \mathrm{cof}(A_{ba}) = [\mathrm{cof}(A)]_{ba} = [\mathrm{cof}(A)]^T_{ab}$
- $(A^{-1})^T = \{\, [\mathrm{cof}(A)]^T / \det(A)\,\}^T = [\mathrm{cof}(A^T)]^T / \det(A^T) = (A^T)^{-1}$

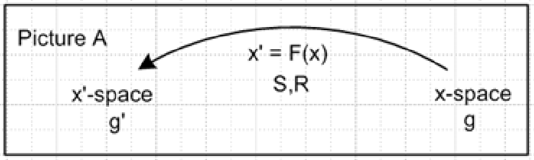
This fact is used many times in the manipulations below.

5.7 Proof that the metric tensor and its inverse are both rank-2 tensors

The above rank-2 tensor transformation rules (5.6.3) can be written in the following matrix form (something not possible with higher-rank tensors),

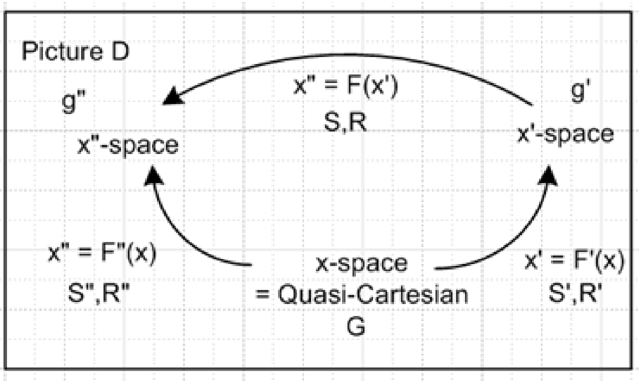
$$M' = R\, M\, R^T \qquad \text{// contravariant rank-2 tensor}$$
$$\bar{M}' = S^T\, \bar{M}\, S \qquad \text{// covariant rank-2 tensor} \qquad (5.7.1)$$

where recall



. (5.7.2)

But we now switch these rules to the **Picture D** context (upper arrow) where F maps x'-space to x"-space,



(5.1.1)

to obtain

$M'' = R\, M'\, R^T$ // contravariant rank-2 tensor (5.7.3)

$\bar{M}'' = S^T\, \bar{M}'\, S$. // covariant rank-2 tensor

Consider then this sequence of steps:

$1 * G * 1 = 1 * G * 1$

$(S''R'')\, G\, (S''R'')^T = (S'R')\, G\, (S'R')^T$ // $S''R'' = 1$ and $S'R' = 1$

$S''\, (R''G\, R''^T)\, S''^T = S'(R'G\, R'^T)\, S'^T$ // regroup

$S''\, g''\, S''^T = S'\, g'\, S'^T$ // since $g'' = R''G\, R''^T$ and $g' = R'G\, R'^T$ by (5.3.2)

$g''\, S''^T = R''\, S'\, g'\, S'^T$ // left multiply by $S''^{-1} = R''$

$g'' = R''\, S'\, g'\, S'^T\, R''^T$ // right multiply by $S''^{T,-1} = R''^T$ (5.7.4)

$g'' = (R''\, S')\, g'\, (S'^T\, R''^T)$ // regroup

$g'' = (R''\, S')\, g'\, (R''\, S')^T$ // $(AB)^T = B^T A^T$

$g'' = R\, g'\, R^T$ // using long forgotten (5.1.3)

With definition (5.7.3) this last result then shows that g' is a *contravariant* rank-2 tensor with respect to the transformation F taking x'-space to x"-space. Continuing on,

$g'' = R\, g'\, R^T$

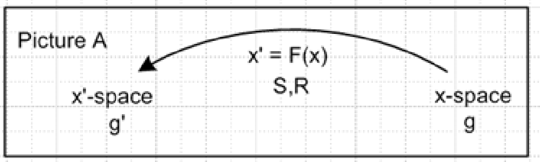
$g''^{-1} = (R\, g'\, R^T)^{-1}$

$g''^{-1} = S^T\, g'^{-1}\, S$ // $R^{T,-1} = S^T$ etc

$\bar{g}'' = S^T\, \bar{g}'\, S$ // $\bar{g}' = g'^{-1}$ (5.7.5)

and this last result shows that $\bar{g}'$ is a *covariant* rank-2 tensor with respect to the transformation F taking x'-space to x"-space, again from (5.7.3). This is why we put a bar over this g from the start.

These two metric tensor transformation statements can be converted to the **Picture A** context,



$$g' = R\, g\, R^T \qquad g'_{ab} = R_{aa'} R_{bb'} g_{a'b'} \qquad \text{// } g \text{ is a contravariant rank-2 tensor}$$
$$\bar{g}' = S^T \bar{g}\, S \qquad \bar{g}'_{ab} = S^T{}_{aa'} S^T{}_{bb'} \bar{g}_{a'b'} \qquad \text{// } \bar{g} \text{ is a covariant rank-2 tensor} \qquad (5.7.6)$$

Since RS = 1, the equations can be inverted to get

$$g = S\, g'\, S^T \qquad g_{ab} = S_{aa'} S_{bb'} g'_{a'b'}$$
$$\bar{g} = R^T\, \bar{g}'\, R \qquad \bar{g}_{ab} = R^T{}_{aa'} R^T{}_{bb'} \bar{g}'_{a'b'} = R_{a'a} R_{b'b} \bar{g}'_{a'b'} \qquad (5.7.7)$$

Further variations of the above are obtained using $RS = 1$, $g\bar{g} = g'\bar{g}' = 1$ and $g = g^T$ (etc.) :

$$Rg = g'\, S^T \qquad \bar{g}'\, R\, g = S^T \qquad g'\, S^T \bar{g} = R$$
$$\bar{g}\, S = R^T \bar{g}' \qquad g\, R^T \bar{g}' = S \qquad \bar{g}\, S\, g' = R^T\ . \qquad (5.7.8)$$

Finally, if x-space is Cartesian so $g = \bar{g} = 1$, one has, analogous to (5.3.4) for Picture C,

$$g' = RR^T \qquad \text{// } g = 1$$
$$\bar{g}' = S^T S \qquad (5.7.9)$$

<u>Comment</u>: If x-space and x'-space are both quasi-Cartesian, so $g = g' = G$, then (5.7.6) says $G = RGR^T$. This does not imply that R must be independent of **x**. One could have $G = R(\mathbf{x})GR(\mathbf{x})^T$. In particular, if both spaces are Cartesian so $g = g' = 1$, then one can have $1 = R(\mathbf{x})R(\mathbf{x})^T$ if $R(\mathbf{x})$ is a real orthogonal matrix for every value of **x**, $R^{-1}(\mathbf{x}) = R^T(\mathbf{x})$. When $R^{-1} = R^T$ we refer to R as a "rotation", where we include possible parity transformations. Then if $g = 1$ and $g' = 1$ one can have $\mathbf{x}' = \mathbf{F}(\mathbf{x}) = R(\mathbf{x})\,\mathbf{x}$ where we refer to $R(\mathbf{x})$ as a **local rotation**. The rotation can be different at every point in space. An example would be

$$R(\mathbf{x}) = R_z(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad \text{where } \theta = \theta(\mathbf{x})\ . \qquad (5.7.10)$$

For N = 4 dimensions in special relativity where $g = g' = G = \mathrm{diag}(-1,1,1,1)$, one can have 4x4 boost and rotation matrices of the form shown below in (5.14.3) and (5.14.4) where the parameters b and r could be $b(\mathbf{x})$ and $r(\mathbf{x})$.

5.8 Metric tensor converts vector types

We continue in **Picture A**. Suppose $\mathbf{V}$ is a contravariant vector so $\mathbf{V}' = \mathrm{R}\mathbf{V}$. Construct a new vector $\mathbf{W}$ with the following properties

$$\begin{aligned} \mathbf{W} &= \bar{g}\,\mathbf{V} && \text{x-space} \\ \mathbf{W}' &= \bar{g}'\,\mathbf{V}' && \text{x'-space} \ . \end{aligned} \qquad (5.8.1)$$

Is vector $\mathbf{W}$ one of our two vector types, or is it neither? One must examine how it transforms under F:

$$\mathbf{W}' = \bar{g}'\,\mathbf{V}' = (S^T\,\bar{g}\,S)\,(R\mathbf{V}) = S^T\,\bar{g}\,(SR)\mathbf{V} = S^T\,\bar{g}\,\mathbf{V} = S^T\,\mathbf{W} \ . \qquad \text{// using (5.7.6)} \qquad (5.8.2)$$

Therefore from (2.5.1) this new vector $\mathbf{W}$ is a *covariant* vector under F, so it should have an overbar,

$$\bar{\mathbf{W}} \equiv \bar{g}\ \mathbf{V} \quad \text{and} \quad \bar{\mathbf{W}}' = S^T\,\bar{\mathbf{W}}\,. \qquad (5.8.3)$$

This covariant vector $\bar{\mathbf{W}}$ can be regarded as the covariant partner of contravariant vector $\mathbf{V}$.

This shows the general idea that applying $\bar{g}$ to any contravariant vector produces a covariant vector! So this is one way to construct covariant vectors if we have a supply of contravariant ones. Conversely, starting with a known covariant vector $\bar{\mathbf{W}}$, one can construct a contravariant vector $\mathbf{V} \equiv g\,\bar{\mathbf{W}}$. Thus, every vector of either type can be thought of as having a partner vector of the other type.

An obvious notation is to write $\bar{\mathbf{W}}$ as $\bar{\mathbf{V}}$ so no extra letter is needed. Then one has

$$\begin{aligned} \bar{\mathbf{V}} &= \bar{g}\,\mathbf{V} && \mathbf{V} = g\,\bar{\mathbf{V}} && \text{// } \bar{g} = g^{-1} \text{ from (5.3.1)} \\ \bar{V}_i &= \bar{g}_{ij}V_j && V_i = g_{ij}\bar{V}_j \ . \end{aligned} \qquad (5.8.4)$$

Then the vectors $\mathbf{V}$ and $\bar{\mathbf{V}}$ are the partner vectors, the first contravariant, the second covariant.

5.9 Vectors in Cartesian space

Theorem: There is no distinction between a contravariant and a covariant vector in Cartesian space. (5.9.1)

Proof: Pick a contravariant vector $\mathbf{V}$. Since $\bar{g} = 1$, $\bar{\mathbf{V}} \equiv \bar{g}\ \mathbf{V} = \mathbf{V}$. But $\bar{\mathbf{V}}$ is a covariant vector. Since $\bar{\mathbf{V}} = \mathbf{V}$, every contravariant vector is identical to its covariant partner *in x-space*. The vector components are numerically equal. The transformation rules (2.5.1) in this case are

$$\begin{aligned} \mathbf{V}' &= R\,\mathbf{V} \\ \bar{\mathbf{V}}' &= S^T\,\bar{\mathbf{V}} = S^T\,\mathbf{V} \ . \end{aligned} \qquad (5.9.2)$$

Since in general one does not have $R = S^T$, one sees that in general $\mathbf{V}' \neq \bar{\mathbf{V}}'$, so the two partner vectors are in general *not* identical in x'-space even though they are identical in Cartesian x-space. In the special case that R is a rotation, $R = R^{-1,T}$ (real orthogonal) and then $R = R^{-1,T} = S^T$ so $\mathbf{V}' = \bar{\mathbf{V}}'$.

5.10 The covariant dot product A • B and norm |A|

For a Cartesian space, Chapter 4 defined the norm as the length of a vector, the metric as the distance between two vectors, and the scalar product (inner product) as the projection of one vector on another. The *official* definitions of norm, metric and scalar product *require* non-negativity: $|\mathbf{x}| \geq 0$, $d(\mathbf{x},\mathbf{y}) \geq 0$, and $\mathbf{x} \bullet \mathbf{x} \geq 0$. For non-Cartesian spaces, the logical extensions of these three concepts can result in all three quantities being negative. Nevertheless, we shall use the term "covariant scalar product" with notation **A • B** as defined below, as well as the notation $|\mathbf{A}|^2 \equiv \mathbf{A} \bullet \mathbf{A}$ where |**A**| will be called the length, magnitude or norm of **A,** even though these objects are not true scalar products or norms. In the curvilinear application of tensor analysis, where x-space is Cartesian, since the norm and scalar product are tensorial scalars, and since they are non-negative in Cartesian x-space, the problem of negative norms does not arise in either space.

How do authors handle this problem? Some authors refer to **A • A** as "the norm" of **A** (e.g., Messiah last line of p 878 discussing special relativity), which is our $|\mathbf{A}|^2$. For a general 4-vector **A** in special or general relativity, most authors just write **A • A** ($A_\mu A^\mu$ in standard notation), they note that the quantity is invariant under transformations, but don't give it a name.

Whereas we use the bold • for this covariant dot product, most special relativity authors prefer to reserve this bold dot for a 3D spatial dot product, and then the 4D dot product is written with some "less bold dot" such as A.B or A•B. Typical usage then in standard notation would be $p\cdot p = p^\mu p_\mu = p_0^2 - \mathbf{p}\bullet\mathbf{p}$ (see for example Bjorken and Drell p 281).

Without further ado, we define the "covariant scalar product" of two contravariant vectors (a new and different use of the word "covariant", but the same as appears in Section 7.15) as

$$\mathbf{A} \bullet \mathbf{B} \equiv A_a \overline{B}_a . \qquad \text{// implied sum on a} \qquad (5.10.1)$$

The important fact about the dot product of two tensorial vectors is that it is a tensorial scalar, as we now show using (5.9.2) for the transformations of **A** and **B** and (2.1.6) that SR = 1 :

$$\mathbf{A'} \bullet \mathbf{B'} \equiv A'_a \overline{B}_a' = (RA)_a(S^T\overline{B})_a = (R_{ai}A_i)(S_{ja}\overline{B}_j) = (S_{ja}R_{ai})A_i\overline{B}_j = (SR)_{ji}A_i\overline{B}_j$$

$$= \delta_{j,i}A_i\overline{B}_j = A_i\overline{B}_i = \mathbf{A} \bullet \mathbf{B} . \qquad (5.10.2)$$

There are various equivalent ways to write the dot product of (5.10.1) using (5.8.4) that $\overline{\mathbf{V}} = \overline{g}\,\mathbf{V}$ and conversely that $\mathbf{V} = g\,\overline{\mathbf{V}}$, and also the fact (5.4.3) that $g_{ab} = g_{ba}$:

$$\mathbf{A} \bullet \mathbf{B} \ \equiv\ A_a\bar{B}_a = A_a(\bar{g}_{ab}B_b) = \bar{g}_{ab}A_aB_b = (\bar{g}_{ba}A_a)B_b = \bar{A}_bB_b = \bar{A}_aB_a$$

$$\mathbf{A} \bullet \mathbf{B} \ \equiv\ A_a\bar{B}_a = (g\bar{A})_a\bar{B}_a = g_{ab}\,\bar{A}_b\,\bar{B}_a = g_{ab}\,\bar{A}_b\,\bar{B}_a \ .$$

To summarize

$$\mathbf{A} \bullet \mathbf{B} = A_a\bar{B}_a = \bar{A}_aB_a = \bar{g}_{ab}A_aB_b = g_{ab}\,\bar{A}_b\,\bar{B}_a \ . \tag{5.10.3}$$

In the special case that **A** = **B**, we use the shorthand norm notation (with caveat as noted above) and (5.10.2) to obtain,

$$|\mathbf{A}|^2 \ \equiv\ \mathbf{A} \bullet \mathbf{A} = \mathbf{A'} \bullet \mathbf{A'} = |\mathbf{A'}|^2 \ . \tag{5.10.4}$$

Going back to Chapter 3 and the vectors $\mathbf{e'}_n$ and $\mathbf{e}_n$, a claim made at the start of Section 3.4 can now be verified:

$$|\mathbf{e'}_n|^2 = \mathbf{e'}_n \bullet \mathbf{e'}_n = \mathbf{e}_n \bullet \mathbf{e}_n = |\mathbf{e}_n|^2 \qquad \Rightarrow \qquad |\mathbf{e'}_n| = |\mathbf{e}_n| \ . \tag{5.10.5}$$

Comment on Notation

Consider again the definition (5.10.1) and the alternate form $\bar{A}_aB_a$ shown in (5.10.3)

$$\mathbf{A} \bullet \mathbf{B} \equiv A_a\bar{B}_a = \bar{A}_aB_a \ . \tag*{(5.10.3)}$$

The dot product involves the contravariant components of one vector and the covariant components of the other vector. In the dot product *notation* **A • B** , we have indicated each vector by its contravariant name just as a convention. We could just as well have indicated one or both vectors by its covariant name, but the dot product indicated by whatever name would be the same: contravariant components of one vector and the covariant components of the other vector. Thus,

$$\mathbf{A} \bullet \mathbf{B} \equiv A_i\bar{B}_i = \bar{A}_iB_i = \bar{\mathbf{A}} \bullet \mathbf{B} = \mathbf{A} \bullet \bar{\mathbf{B}} = \bar{\mathbf{A}} \bullet \bar{\mathbf{B}} \ . \tag{5.10.6}$$

We shall always use the first notation **A • B** since it is the simplest. Similarly, when **B** = **A**,

$$|\mathbf{A}|^2 \equiv \mathbf{A} \bullet \mathbf{A} = |\bar{\mathbf{A}}|^2 \ . \tag{5.10.7}$$

We thus have the interesting fact that, if **A** and **B** are tensorial vectors for *general* **x'** = **F**(**x**), then

$$\mathbf{A} \neq \bar{\mathbf{A}} \qquad \text{but} \qquad \mathbf{A} \bullet \mathbf{B} = \bar{\mathbf{A}} \bullet \mathbf{B} \qquad \text{for any } \mathbf{B} \ . \tag{5.10.8}$$

If **B** = **A**, this says

$$\mathbf{A} \neq \bar{\mathbf{A}} \qquad \text{but} \qquad \mathbf{A} \bullet \mathbf{A} = \bar{\mathbf{A}} \bullet \mathbf{A} = |\mathbf{A}|^2 = A_i\bar{A}_i = g_{ij}A_iA_j \ . \tag{5.10.9}$$

We mention this notational issue to head off the following incorrect notion:

$$\mathbf{A} \bullet \mathbf{B} = \bar{g}_{ab}A_aB_b \quad \Rightarrow \quad \bar{\mathbf{A}} \bullet \bar{\mathbf{B}} = \bar{g}_{ab}\bar{A}_a\bar{B}_b \neq \mathbf{A} \bullet \mathbf{B} \ . \qquad \text{// wrong!!!}$$

<u>A dot product application</u>

In applications in which $(ds)^2$ is regarded as a scalar with respect to transformation F we have

$$(ds')^2 = d\mathbf{x}' \bullet d\mathbf{x}' = (ds)^2 = d\mathbf{x} \bullet d\mathbf{x} \qquad (5.10.10)$$

and ds = ds' is called "the invariant distance". Such applications include curvilinear coordinate transformations and relativity transformations.

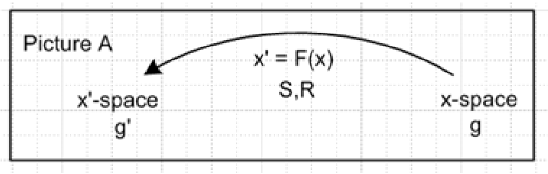
In special relativity, using the Bjorken and Drell notation noted above where $g'_{\mu\nu}$ = diag(1,-1,-,1,-1) and c = 1, one writes (Standard Notation),

$$(d\tau)^2 = g'_{\mu\nu}\, dx'^{\mu}dx'^{\nu} = dx'_{\mu}dx'^{\mu} = dx' \bullet dx' = dx \bullet dx = \text{a Lorentz scalar} = (dt)^2 - d\mathbf{x} \bullet d\mathbf{x}\ , \ \ x^{\mu} = (t,\mathbf{x}) \qquad (5.10.11)$$

and dτ is called "the proper time", a particular case of the invariant distance ds. Notice that $(d\tau)^2 < 0$ for a spacelike 4-vector dx^{μ}, meaning one that lies outside the future and past lightcones ($|d\mathbf{x}| > |dt|$). [We now restore • to our covariant definition after temporarily using it above for a 3-space Cartesian dot product.]

5.11 Metric tensor and tangent base vectors: scale factors and orthogonal coordinates

The context of Picture A continues,



(5.11.1)

Recall this fact from (3.2.7),

$$S = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \ .... \ \mathbf{e}_N] \qquad \textit{(3.2.7)}$$

where the columns of S are the tangent base vectors. It follows from (5.7.6) that

$$\bar{g}' = S^T\ \bar{g}\ S = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \ .... \ \mathbf{e}_N]^T\ \bar{g}\ [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \ .... \ \mathbf{e}_N] \qquad (5.11.2)$$

so

$$\bar{g}' = \begin{array}{lllll} e_1\bullet e_1 & e_1\bullet e_2 & e_1 \bullet e_3 & \ldots\ldots & e_1\bullet e_N \\ e_2\bullet e_1 & e_2\bullet e_2 & e_2 \bullet e_3 & \ldots\ldots & e_2\bullet e_N \\ e_3\bullet e_1 & e_3\bullet e_2 & e_3 \bullet e_3 & \ldots\ldots & e_3\bullet e_N \\ \ldots\ldots\ldots & & & & \\ e_N\bullet e_1 & e_N\bullet e_2 & e_N \bullet e_3 & \ldots\ldots & e_N\bullet e_N \end{array} \qquad (5.11.3)$$

since,

$$\mathbf{e}_n^T\, \bar{g}\; \mathbf{e}_m \;=\; (\mathbf{e}_n)_i\, \bar{g}_{ij}\, (\mathbf{e}_m)_j \;= \bar{g}_{ij}(\mathbf{e}_n)_i(\mathbf{e}_m)_j \;= \mathbf{e}_n \bullet \mathbf{e}_m \qquad (5.11.4)$$

using the covariant dot product shown in (5.10.3). Taking the m,n component of (5.11.3) one gets

$$\bar{g}'_{mn} = \mathbf{e}_m \bullet \mathbf{e}_n \qquad \text{or} \qquad \bar{g}'_{mn} = \partial'_m\mathbf{x} \bullet \partial'_n\mathbf{x} \qquad \text{// using (3.2.6)} \qquad (5.11.5)$$

which makes a direct connection between the covariant metric tensor in x'-space and the tangent base vectors $\mathbf{e}_n$ in x-space. A less graphical derivation of this fact uses (5.7.6) and (3.2.5),

$$\bar{g}'_{nm} = (S^T\bar{g}S)_{nm} = S^T{}_{na}\; \bar{g}_{ab}\; S_{bm} = \bar{g}_{ab}\; S_{an}\, S_{bm} \;=\; \bar{g}_{ab}\, (\mathbf{e}_n)_a(\mathbf{e}_b)_n \;\equiv\; \mathbf{e}_n \bullet \mathbf{e}_m \;\; . \qquad (5.11.6)$$

Scale Factors and Orthgonal Coordinates

We showed just above in (5.11.5) that $\mathbf{e}_m \bullet \mathbf{e}_n = \bar{g}'_{mn}$. Regardless of whether or not the coordinates are orthogonal, one can define the scale factor h'_n and unit vector $\hat{\mathbf{e}}_n$ as follows,

$$h'^2_n \equiv\; |\mathbf{e}_n|^2 = \mathbf{e}_n \bullet \mathbf{e}_n = \bar{g}'_{nn}$$

$$\Rightarrow \qquad h'_n = |\mathbf{e}_n| = \sqrt{\bar{g}'_{nn}}$$

$$\Rightarrow \qquad \hat{\mathbf{e}}_n \equiv \mathbf{e}_n/h'_n = \text{"unit vector"}, \quad |\hat{\mathbf{e}}_n| \;= 1 \;\; . \qquad (5.11.7)$$

If the tangent base vectors $\mathbf{e}_n$ are **orthogonal**, then clearly the metric tensor $\bar{g}'_{mn} = \; \mathbf{e}_m \bullet \mathbf{e}_n$ must be diagonal. In fact, based on (5.11.7), the covariant metric tensor must be

$$\bar{g}'_{mn} = h'^2_m\, \delta_{m,n} \;.$$

We claim that the contravariant metric tensor g'_{nm} is then *also* diagonal with diagonal elements

$$g'_{nn} \;= h'^{-2}_n \;.$$

To verify this claim, we confirm that $\bar{g}'\, g' = 1$:

$$\Sigma_m\; \bar{g}'_{nm}\, g'_{mk} \;= \Sigma_m\; [\delta_{n,m}\, h'^2_m][\delta_{m,k} h'^{-2}_m] \;= \Sigma_m\, \delta_{n,m}\delta_{m,k} = \delta_{n,k} \;= (1)_{nk}$$

Since the inverse matrix is unique as long as det(g')≠0, $g'_{mk} = \delta_{m,k}h'^{-2}_m$ must be it!

We have then established this simple fact:

$$\text{orthogonal coordinates} \quad \Leftrightarrow \quad \begin{array}{ll} \bar{g}'_{mn} = h'^{\,2}_{m}\ \delta_{m,n} & \text{covariant} \\ g'_{mn} = h'^{\,-2}_{m}\delta_{m,n} & \text{contravariant .} \end{array} \qquad (5.11.8)$$

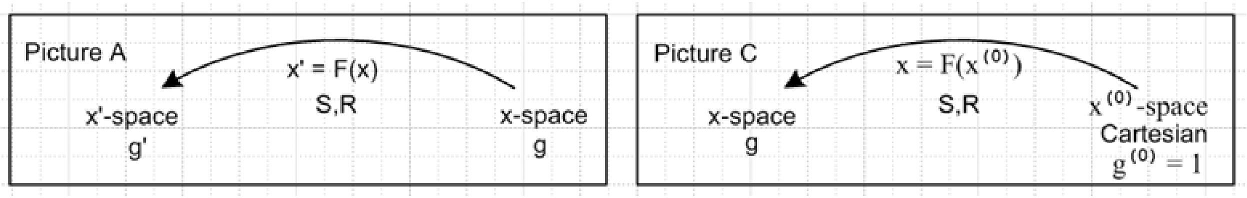
For an orthogonal coordinate system, at any point **x** in x-space, the tangents $\mathbf{e_n}$ to the N coordinate lines passing through that point are orthogonal. Most examples below will involve such systems, with Appendix C providing a non-orthogonal example.

In the Standard Notation of Chapter 7, the above association becomes

$$\text{orthogonal coordinates} \quad \Leftrightarrow \quad \begin{array}{ll} g'_{mn} = h'^{\,2}_{m}\ \delta_{m,n} & \text{covariant} \\ g'^{mn} = h'^{\,-2}_{m}\delta_{m,n} & \text{contravariant .} \end{array} \qquad (5.11.9)$$

The reason we put a prime on h'_n is because it is associated with x'-space and its metric tensor g' in Picture A or B of Fig (1.11). The curvilinear coordinates are **x**'.

When working with Picture C, however, we would call this scale factor h_n because it is then associated with the metric tensor g and x-space. In Picture C, the curvilinear coordinates are **x**.



In other Pictures, the curvilinear "space on the left" might be called ξ-space, and since this does not carry a prime, again one would write h_n.

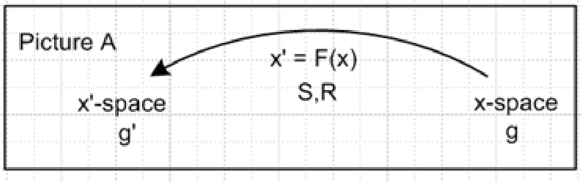
Section 5.5 above showed that $\bar{g}'_{nn} \geq 0$ when x-space is Cartesian. This is the usual case for the curvilinear coordinates application, and so in this case the scale factors h'_n are always real and positive.

Note 1. Different authors use different symbols for scale factors. For example, Margenau and Murphy refer to them as the Q_n. Morse and Feshbach and most modern works use h_n. They never use a prime, because they always use Picture C or equivalent where **x** is the curvilinear coordinate.

Note 2: Some authors refer to the scale factors h'_n as the **Lamé coefficients**, while other authors refer to R_{ij} as the Lamé coefficients which they call h_{ji}. (Lame, for PDF search with no accent mark)

5.12 The Jacobian J

The context of Picture A continues,



(5.12.1)

First of all, note that since RS = 1,

$$\det(S) = 1/\det(R)\ . \tag{5.12.2}$$

Recall from (2.1.6) our definitions of R_{ik} and S_{ik} :

$R_{ik}(\mathbf{x}) = (\partial x'_i/\partial x_k)$ is called the Jacobian matrix for the transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$
$S_{ik}(\mathbf{x}') = (\partial x_i/\partial x'_k)$ is then the Jacobian matrix of the inverse transformation $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$. (5.12.3)

There are two Jacobian matrices here, which are inverses of each other since RS = 1, and for each of these matrices we could define "a Jacobian" as the determinant of that matrix:

$$J(\mathbf{F}) = \det(R) \qquad 1/ J(\mathbf{F}) = 1/\det(R) = \det(S) = J(\mathbf{F}^{-1})$$
$$J(\mathbf{F}^{-1}) = \det(S)\ . \tag{5.12.4}$$

Because our major interest is in curvilinear coordinates $\mathbf{x}'$, and because curvilinear coordinates are almost always defined by equations of the inverse form $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$ such as the following for polar coordinates,

$$x = r\cos\theta \qquad \mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}') \tag{5.12.5}$$
$$y = r\sin\theta,$$

we shall officially define "*the* Jacobian" to be $J = J(F^{-1})$. So,

$$\text{the Jacobian} \equiv J(\mathbf{x}') \equiv \det(S(\mathbf{x}')) = \det(\partial x_i/\partial x'_k) = 1/\det(R(\mathbf{x}(\mathbf{x}')) = 1/ \det(\partial x'_i/\partial x_k)\ . \tag{5.12.6}$$

Note 1: Objects which relate to the transformation between x-space and x'-space cannot themselves be tensors because tensor objects must be associated with a specific space, the way $\mathbf{V}(\mathbf{x})$ is a vector in x-space and $\mathbf{V}'(\mathbf{x}')$ is a vector in x'-space. Thus $S_{ij}(\mathbf{x}') = \partial x_i/\partial x'_k$, although a matrix, is not a rank-2 tensor. Similarly, $J(\mathbf{x}')$, while a "scalar" function, is not a rank-0 tensorial scalar. One does not ask how S and J themselves "transform" in going from x-space to x'-space.

Note 2: An alternative notation used by some authors is this

$$\mathcal{J}(\mathbf{x},\mathbf{x}') \equiv \det(S(\mathbf{x},\mathbf{x}')) = \det(\partial x_i/\partial x'_k) \tag{5.12.7}$$

as if **x** and **x**' were independent variables. In our presentation, **x** is not an independent variable but is determined by $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$. Just as one might write $f'(x') = \partial f/\partial x'$, we write $J(\mathbf{x}') = \det(\partial x_i/\partial x'_k)$. The connection would be $J(\mathbf{x}') = \mathcal{J}(\mathbf{x}{=}F^{-1}(\mathbf{x}'),\mathbf{x}')) = \mathcal{J}(\mathbf{x}(\mathbf{x}'),\mathbf{x}')$.

<u>Note 3:</u> Other sources often use the notation | M | to indicate the determinant of a matrix. We shall use the notation det(M), and reserve | | to indicate the magnitude of some quantity, such as |J| below.

The determinant of any NxN matrix S may be written ($\varepsilon_{abc..}$ is the permutation tensor, see Section 7.7),

$$\det(S) = \varepsilon_{abc...x}\, S_{a1}\, S_{b2}\, ...\, S_{xN}\ . \qquad // = \varepsilon_{abc...x}\, S_{1a}\, S_{2b}\, ...\, S_{Nx} \qquad (5.12.8)$$

For our particular S with $S_{in} = (\mathbf{e}_n)_i$ from (3.2.5) this becomes

$$\det(S) = \varepsilon_{abc...x}\ (\mathbf{e}_1)_a(\mathbf{e}_2)_b ....\ (\mathbf{e}_N)_x \qquad (5.12.9)$$

so J is related to the tangent base vectors by

$$J = \ \varepsilon_{abc...x}\ (\mathbf{e}_1)_a(\mathbf{e}_2)_b ....\ (\mathbf{e}_N)_x\ . \qquad (5.12.10)$$

It was shown in (5.7.6) that $\bar{g}' = S^T\ \bar{g}\ S$ and $g' = R\ g\ R^T$, these being the transformation rules for covariant and contravariant rank-2-tensors. Therefore

$$\det(\bar{g}') = \det(S^T\bar{g}S) = \det(S^T)\det(\bar{g})\det(S) = \det(S)\det(S)\det(\bar{g}) = J^2 \det(\bar{g})$$
$$\det(g') = \det(RgR^T) = \det(R)\det(g)\det(R^T) = \det(R)\det(R)\det(g) = J^{-2} \det(g)$$

or

$$\det(\bar{g}') = J^2 \det(\bar{g}) \qquad \Rightarrow \qquad J^2 = \det(\bar{g}') / \det(\bar{g}) = [\det(S)]^2$$
$$\det(g') = J^{-2}\det(g)\ . \qquad (5.12.11)$$

It is a tradition to define certain scalar (but not tensorial scalar) objects with *the same name* g and g',

$$g(\mathbf{x}) \ \equiv\ \det(\bar{g}(\mathbf{x})) \ = 1/\det(g(\mathbf{x})) \qquad // \text{ in x-space}$$
$$g'(\mathbf{x}') \equiv\ \det(\bar{g}'(\mathbf{x}')) \ = 1/\det(g'(\mathbf{x}')) \qquad // \text{ in x'-space} \qquad (5.12.12)$$

so that

$$J^2(\mathbf{x}') = \det(\bar{g}'(\mathbf{x}')) / \det(\bar{g}(\mathbf{x})) = g'(\mathbf{x}') / g(\mathbf{x})\ . \qquad (5.12.13)$$

Normally the argument dependence is suppressed and one then writes

$$g \equiv \det(\bar{g}) = 1/\det(g)$$
$$g' \equiv \det(\bar{g}') = 1/\det(g')$$
$$J^2 = \det(\bar{g}')/ \det(\bar{g}) = g'/g \quad \Rightarrow \qquad g' = J^2\ g\ . \qquad (5.12.14)$$

As explained in Section D.1, the equation $g' = J^2\ g$ says that g, instead of being a tensorial scalar, is a scalar *density* of weight -2. A tensorial scalar s has weight 0: $s' = J^0 s = s$.

Warning: One must be careful to distinguish the scalars g and g' from the the matrices g and g'. It is usually clear from the context which meaning is implied.

It is convenient to make the following definition, called the **signature** of the metric tensor,

$$s = \text{sign}[\det(\bar{g})] = \text{sign}(g) \qquad \text{// that is, } s = \text{either } +1 \text{ or } -1 \; . \tag{5.12.15}$$

Since $g\,\bar{g} = 1$ by (5.3.2), one has $\det(g)\det(\bar{g}) = 1$ or so that $\text{sign}[\det(\bar{g})] = \text{sign}[\det(g)]$.
Since $\det(\bar{g}') / \det(\bar{g}) = [\det(S)]^2$ by (5.12.11), one has $\text{sign}[\det(\bar{g})] = \text{sign}[\det(\bar{g}')]$.
Since $g'\,\bar{g}' = 1$ by (5.3.2), one has $\det(g')\det(\bar{g}') = 1$ or so that $\text{sign}[\det(\bar{g}')] = \text{sign}[\det(g')]$.

Therefore:

$$s = \text{sign}[\det(\bar{g})] = \text{sign}[\det(g)] = \text{sign}[\det(\bar{g}')] = \text{sign}[\det(g')] = \text{sign}(g) = \text{sign}(g') \; . \tag{5.12.16}$$

Since transformation F is assumed invertible in its domain and range, one cannot have $\det(S)=0$ anywhere except perhaps on a boundary. Since $\det(\bar{g}') = [\det(S)]^2\det(\bar{g})$ by (5.12.11), if we assume $\det(\bar{g})$ is non-vanishing in the x-space domain of F, then $\det(\bar{g}') \neq 0$ everywhere in the range of F. The conclusion with this assumption is that the signature s is always well-defined.

Obviously, the quantities sg and sg' are both positive, and since $J^2 = g'/g$ one can write

$$|J| = \sqrt{sg'} / \sqrt{sg} = |\det(S)| = \sqrt{g'/g} \; . \tag{5.12.17}$$

For the curvilinear coordinates application, x-space is Cartesian, $\det(\bar{g}) = g = 1$, and thus $s = 1$ and then

$$|J| = \sqrt{g'} = |\det(S)| \; . \qquad \text{// curvilinear} \tag{5.12.18}$$

For the relativity application, x-space is Minkowski space with $\det(\bar{g}) = g = -1$ so $s = -1$ and

$$|J| = \sqrt{-g'} = |\det(S)| \; . \qquad \text{// relativity} \tag{5.12.19}$$

Here then is a summary of the results of this Section:

$$J(\mathbf{x}') \equiv \det(S(\mathbf{x}')) \;=\; \det(\partial x_i/\partial x'_k) \;=\; 1/\det(R(\mathbf{x}(\mathbf{x}'))) = 1/\det(\partial x'_i/\partial x_k)$$

$$g \equiv \det(\overline{g}) \qquad\qquad g' \equiv \det(\overline{g}')$$

$$g' = J^2 g \qquad\qquad \Rightarrow \qquad g \text{ is a scalar density of weight -2 (Section D.1)}$$

$$s \equiv \text{sign}[\det(\overline{g})] = \text{sign}[\det(g)] \;=\; \text{sign}[\det(\overline{g}')] = \text{sign}[\det(g')] \;=\; \text{sign}(g) = \text{sign}(g')$$

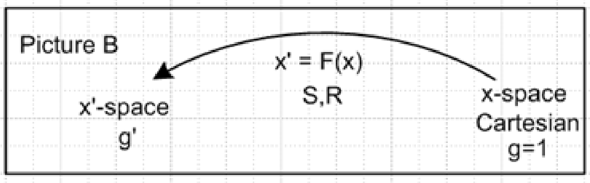
$$|J| \;= \sqrt{sg'} \,/ \sqrt{sg} \;=\; |\det(S)| = \sqrt{g'/g} \tag{5.12.20}$$

<u>Note</u>: Weinberg p 98 (4.4.1) defines $g = -\det(g_{ij})$. This is the only one of Weinberg's conventions that we have not adopted, so in this paper it is always true that $g \equiv + \det(g_{ij})$ even though this is -1 in the application to special relativity.

<u>Carl Gustav Jacob Jacobi (1804 –1851)</u>. German, Berlin PhD 1825 then went to Konigsberg, did much in a short life. Elucidated the whole world of elliptic integrals and functions, such as F(x,k) and sn(x;k), which occur even in simple problems like the 2D pendulum. Wiki claims he promoted Legendre's ∂ symbol for partial derivatives (used throughout this document) and made it a standard. Among many other contributions, he saw the significance of the object J which now bears his name: "the Jacobian". The Jacobi Identity is another familiar item, a rule for non-commuting operators $[x,[y,z]] + [z,[x,y]] + [y,[z,x]] = 0$ which finds use with quantum mechanical operators and matrices, and more generally with Lie group generators -- examples of which appear in Section 5.14 below.

5.13 Some relations between g, R and S in Pictures B and C (Cartesian x-space).

In Picture B, which is Picture A with g = 1,



(5.13.1)

the statement of the rank-2 tensor transformation of g' and $\overline{g}'$ becomes, as shown in (5.7.9),

$$\begin{aligned} g' &= RR^T \\ \overline{g}' &= S^T S \end{aligned} \tag{5.13.2}$$

which can be written in a variety of ways,

$$\begin{aligned} R^T = (SR)R^T \;&= S(RR^T) = S\,g' \quad \Rightarrow \quad R = g'\,S^T \quad \Rightarrow \quad 1 = S\,g'\,S^T \\ S^T = S^T(R^T S^T) \;&= (S^T S)R \;= \overline{g}'\,R \quad \Rightarrow \quad S = R^T\,\overline{g}' \quad \Rightarrow \quad 1 = R^T\,\overline{g}'\,R\,. \end{aligned} \tag{5.13.3}$$

In summary:

$$\begin{array}{llll} g' = RR^T & R^T = S\, g' & R = g'\ S^T & 1 = S\, g'\, S^T \\ \bar{g}' = S^TS & S^T = \bar{g}'\, R & S = \ R^T\ \bar{g}' & 1 = R^T\, \bar{g}'\, R\ . \end{array} \qquad (5.13.4)$$

The diagonal elements of $\bar{g}'$ and g' are given by

$$\begin{aligned} \bar{g}'_{nn} &= \Sigma_n\ S^T{}_{ni}S_{in} = \Sigma_n\,(S_{in}{}^2) = \ \Sigma_i\,(\partial x_i/\partial x'_n)^2 \\ g'_{nn} &= \Sigma_n\ R_{ni}R^T{}_{in} = \Sigma_n\,(R_{ni}{}^2)\ = \Sigma_i\,(\partial x'_n/\partial x_i)^2\ . \end{aligned} \qquad (5.13.5)$$

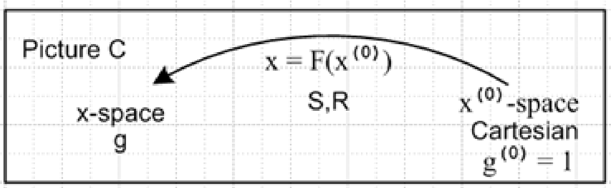
If the x'_i are *orthogonal* coordinates, then $\bar{g}'_{nm} = h'_n{}^2\delta_{n,m}$ and $g'_{nm} = h'_n{}^{-2}\delta_{n,m}$ as shown in (5.11.8) where the h'_n are the scale factors defined in (5.11.7). These scale factors may then be expressed as,

$$h'_n{}^2 = \bar{g}'_{nn} = \ \Sigma_i\,(\partial x_i/\partial x'_n)^2 \qquad\qquad h'_n{}^{-2} = g'_{nn} = \ \Sigma_i\,(\partial x'_n/\partial x_i)^2\ . \qquad (5.13.6)$$

With notational changes $x'_n \to \xi_n$ (curvilinear coordinates for M&F) and $h'_n \to h_n$, these last two equations appear (for three dimensions) in Morse & Feshbach Vol I p 24 equation (1.3.4):

$$h_n^2 = \left(\frac{\partial x}{\partial \xi_n}\right)^2 + \left(\frac{\partial y}{\partial \xi_n}\right)^2 + \left(\frac{\partial z}{\partial \xi_n}\right)^2 = \left[\left(\frac{\partial \xi_n}{\partial x}\right)^2 + \left(\frac{\partial \xi_n}{\partial y}\right)^2 + \left(\frac{\partial \xi_n}{\partial z}\right)^2\right]^{-1} \quad (1.3.4) \qquad (5.13.7)$$

In Picture A (5.13.1) the curvilinear coordinates are called x'_n and the Cartesian coordinates are x_n. In Picture C, the curvilinear coordinates are called x_n and the Cartesian coordinates are $x_n{}^{(0)}$. To simplify notation, we use ξ_n in place of $x_n{}^{(0)}$ for the Cartesian coordinates (these are *totally different* ξ_n from the ξ_n shown above in the M&F quote).



$$\xi_n \equiv x_n{}^{(0)} \qquad (5.13.8)$$

In this Picture C, equations (5.13.2,4,5,6) appear as follows (no primes on g's) :

$$\begin{aligned} g &= RR^T \\ \bar{g} &= S^TS \end{aligned}$$

$$\begin{array}{llll} g = RR^T & R^T = S\, g & R = g\, S^T & 1 = S\, gS^T \\ \bar{g} = S^TS & S^T = \bar{g}\, R & S = \ R^T\ \bar{g} & 1 = R^T\, \bar{g}\, R \end{array}$$

$$\begin{aligned} \bar{g}_{nn} &= \Sigma_n\ S^T{}_{ni}S_{in} = \Sigma_n\,(S_{in}{}^2) = \ \Sigma_i\,(\partial \xi_i/\partial x_n)^2 \\ g_{nn} &= \Sigma_n\ R_{ni}R^T{}_{in} = \Sigma_n\,(R_{ni}{}^2)\ = \Sigma_i\,(\partial x_n/\partial \xi_i)^2 \end{aligned}$$

$$h_n{}^2 = \bar{g}_{nn} = \ \Sigma_i\,(\partial \xi_i/\partial x_n)^2 \qquad\qquad h_n{}^{-2} = g_{nn} = \ \Sigma_i\,(\partial x_n/\partial \xi_i)^2 \qquad (5.13.9)$$

In Picture B of (5.13.1), we have primes on things like g' and h_n' *because* these primes match the names of the curvilinear coordinates which are x'_n and which live in x'-space. In Picture C, since the curvilinear coordinates are now x_n, objects like g and h_n have no primes to match the fact that x_n have no primes and live in x-space. And in Picture C, one has $S_{in} = \partial\xi_i/\partial x_n$, where x_n is the curvilinear coordinate and ξ_i the Cartesian coordinate. It is admittedly a bit confusing, but it is just notation.

Example 1: Polar coordinates: metric tensor and Jacobian

Picture C (5.13.8) continues (so now $\theta = x_1$ and $r = x_2$) and the metric tensor for polar coordinates will be computed in two ways. It was shown in (3.4.1) and (3.4.2) that

$$S = \begin{pmatrix} -r\sin\theta & \cos\theta \\ r\cos\theta & \sin\theta \end{pmatrix} = [\,\mathbf{e}_1, \mathbf{e}_2\,] \qquad \mathbf{e}_1 = r(-\sin\theta, \cos\theta) \qquad \mathbf{e}_2 = (\cos\theta, \sin\theta)\ . \qquad (5.13.10)$$

One way to compute $\bar{g}$ is this, using (5.13.9): (1=θ, 2=r)

$$\bar{g} = S^T S = \begin{pmatrix} -r\sin\theta & r\cos\theta \\ \cos\theta & \sin\theta \end{pmatrix}\begin{pmatrix} -r\sin\theta & \cos\theta \\ r\cos\theta & \sin\theta \end{pmatrix} = \begin{pmatrix} r^2 & 0 \\ 0 & 1 \end{pmatrix} \qquad \Rightarrow \qquad \bar{g}_{\theta\theta} = r^2 \qquad \bar{g}_{rr} = 1$$

$$h_\theta = r \qquad h_r = 1$$

Another method uses (5.11.3) (but in Picture C, so no prime on $\bar{g}$),

$$\bar{g} = \begin{pmatrix} \mathbf{e}_1\bullet\mathbf{e}_1 & \mathbf{e}_1\bullet\mathbf{e}_2 \\ \mathbf{e}_2\bullet\mathbf{e}_1 & \mathbf{e}_2\bullet\mathbf{e}_2 \end{pmatrix} = \begin{pmatrix} r^2 & 0 \\ 0 & 1 \end{pmatrix} \qquad // \det(\bar{g}) = r^2\ . \qquad (5.13.11)$$

Notice that this metric tensor is in fact symmetric, and that one of its elements is a function of the coordinates. The length^2 of a small vector d**x** can be written using (5.2.5) (but in Picture C),

$$(ds)^2 = \bar{g}_{km}\, dx_k\, dx_m = \bar{g}_{\theta\theta}\, d\theta\, d\theta + \bar{g}_{rr}\, dr\, dr = r^2\, (d\theta)^2 + (dr)^2\ . \qquad (5.13.12)$$

The Jacobian is given by (5.12.6) (but in Picture C so J(**x**') → J(**x**)),

$$J(r,\theta) = \det(S(r,\theta)) = \det\begin{pmatrix} -r\sin\theta & r\cos\theta \\ \cos\theta & \sin\theta \end{pmatrix} = -r \ \text{ so } \ |J| = r \quad \text{and} \quad g = J^2 = r^2,\ \sqrt{g} = r\ . \qquad (5.13.13)$$

Note that $J^2 = g'/g \rightarrow g/1 = g$ converting (5.12.14) to Picture C.

Example 2: Spherical coordinates: metric tensor and Jacobian

As with Example 1, Picture C is used, this time with $(x_1, x_2, x_3) = (r,\theta,\varphi)$.

In (3.4.4) it was found that

$$S = \begin{pmatrix} \sin\theta\cos\varphi & r\cos\theta\cos\varphi & -r\sin\theta\sin\varphi \\ \sin\theta\sin\varphi & r\cos\theta\sin\varphi & r\sin\theta\cos\varphi \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix} . \tag{3.4.4}$$

Comment: The matrix S looks the same in all Pictures. What differs in different Pictures are the *names* of the coordinates that go with the rows and columns of the matrix. Looking at Picture A of Fig (2.1.1) and equations (2.1.6), one sees that the matrix elements of S in Picture A are $S_{ik} = (\partial x_i/\partial x'_k)$ where x_i are the Cartesian coordinates x,y,z and x'_i are the curvilinear coordinates r,θ,φ. But in Picture C of Fig (5.13.8), one has instead $S_{ik} = \partial\xi_i/\partial x_k$ where ξ_i are the Cartesian coordinates x,y,z and x_i are the curvilinear coordinates r,θ,φ.

The metric tensor from (5.13.9) is then given by a Maple matrix calculation as

$$\bar{g} = S^T S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2\theta \end{pmatrix} \qquad \det(\bar{g}) = r^4\sin^2\theta \tag{5.13.14}$$

so that

$$\begin{aligned} \bar{g}_{11} &= \bar{g}_{rr} = 1 & h_1 &= h_r = \sqrt{\bar{g}_{rr}} = 1 \\ \bar{g}_{22} &= \bar{g}_{\theta\theta} = r^2 & h_2 &= h_\theta = \sqrt{\bar{g}_{\theta\theta}} = r \\ \bar{g}_{33} &= \bar{g}_{\varphi\varphi} = r^2\sin^2\theta & h_3 &= h_\varphi = \sqrt{\bar{g}_{\varphi\varphi}} = r\sin\theta \ . \end{aligned} \tag{5.13.15}$$

The Jacobian using (5.12.6) is found by Maple to be,

$$J(r,\theta,\varphi) = \det(S) = r^2\sin\theta \ . \tag{5.13.16}$$

An alternate calculation uses (5.12.14) converted to Picture C where $J^2 = g'/g \rightarrow g/1 = g$. Then,

$$J = \sqrt{g} = \sqrt{\det(\bar{g})} = \sqrt{r^4\sin^2\theta} = r^2\sin\theta. \tag{5.13.17}$$

Differential distance using (5.2.5) (but in Picture C) and (5.13.15) is:

$$(ds)^2 = \bar{g}_{km}\, dx_k\, dx_m = (dr)^2 + r^2(d\theta)^2 + r^2\sin^2(d\varphi)^2 \tag{5.13.18}$$

and if $d\varphi = 0$, this agrees with the polar coordinates result (5.13.12).

5.14 Special Relativity and its Metric Tensor: vectors and spinors

In this Section the Standard Notation introduced below in Chapter 7 is used. In that notation R_{ij} is written $R^i{}_j$, contravariant vectors V_i are written V^i, and covariant vectors $\overline{V}_j$ are written V_j. It is a tradition in special and general relativity to use Greek letters for 4-vector indices and Latin letters for spatial 3-vector indices.

The (Quasi-Cartesian) metric tensor of special relativity is frequently taken as G = diag(1,-1,-1,-1) and the ordering of 4-vectors as x^μ = (t,x,y,z) where c=1 (speed of light) and μ= 0,1,2,3 (Bjorken and Drell p 281). General relativity people often use G = diag(-1,1,1,1) ≡ η instead (Weinberg p 26). Still other authors use G = 1 and x^μ = (it,x,y,z) where i is the imaginary i, but this approach does not easily fit into our tensor framework which is based on real numbers.

A Lorentz transformation is a *linear* transformation $F^\mu{}_\nu$ (equation below has an implied sum on ν)

$$x'^\mu = F^\mu{}_\nu\, x^\nu = R^\mu{}_\nu\, x^\nu \qquad \Rightarrow \qquad x^\nu \text{ is a contravariant vector} \qquad // \ \mathbf{x}' = \mathbf{F}(\mathbf{x}) \qquad (5.14.1)$$

and a theory requirement is that invariant length be preserved, x'**.**x' = x**.**x = scalar. Special relativity also requires that the metric tensor G be the same in all frames, since no frame is special, so G' = G. But this says, in our old notation, that R G R^T = G (which is (5.7.6) with g' = g = G). This condition restricts the (proper) Lorentz transformations to be rotations, boosts (velocity transformations), or any combination of the two. In particular,

$$\begin{aligned} & R\,G\,R^T = G \ \Rightarrow\ \det(R\,G\,R^T) = \det(G) \\ & \Rightarrow\ \det(R)\det(G)\det(R^T) = \det(G)\ \Rightarrow \qquad [\det(R)]^2\,(-1) = (-1) \\ & \Rightarrow\ \det(R) = \pm 1\ . \end{aligned} \qquad (5.14.2)$$

"Proper" Lorentz transformations have det(R) = det(F) = +1, and here are two examples. First, a boost transformation in the x direction,

$$F^\mu{}_\nu = \begin{pmatrix} \cosh(b) & \sinh(b) & 0 & 0 \\ \sinh(b) & \cosh(b) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \exp(-ibK_1) \qquad \text{where } (K_1)^\mu{}_\nu = \begin{pmatrix} 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad (5.14.3)$$

and second, a rotation transformation about the x axis,

$$F^\mu{}_\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos(r) & -\sin(r) \\ 0 & 0 & \sin(r) & \cos(r) \end{pmatrix} = \exp(-irJ_1) \qquad \text{where } (J_1)^\mu{}_\nu = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix} . \qquad (5.14.4)$$

The matrices K_1 and J_1 are called generators and are a part of a set of six 4x4 matrices J_i and K_i for i = 1,2,3. These 6 generator matrices satisfy a set of commutation relations known as a Lie Algebra,

$$\begin{aligned} &[\,J_i, J_j] = +i\,\varepsilon_{ijk}J_k && //\ [\,A,B\,] \equiv AB - BA \\ &[\,J_i, K_j] = +i\,\varepsilon_{ijk}\,K_k && \\ &[\,K_i, K_j] = -i\,\varepsilon_{ijk}J_k\ . && (5.14.5) \end{aligned}$$

In these commutators, the generators J_i and K_i can be regarded as abstract non-commuting operators, while the specific 4x4 matrices shown above for J_1 and K_1 are just a "representation" of these abstract operators as 4x4 matrices. The six 4x4 generator matrices J_i and K_i are ($g = G = \mathrm{diag}(1,-1,-1,-1)$)

$$\begin{aligned} (J^{\mu\nu})^{\alpha}{}_{\beta} = i\,(\,g^{\mu\alpha}\delta^{\nu}{}_{\beta} - g^{\nu\alpha}\delta^{\mu}{}_{\beta}) \quad (J_1)^{\alpha}{}_{\beta} &\equiv (J^{23})^{\alpha}{}_{\beta} = i\,(\,g^{2\alpha}\delta^{3}{}_{\beta} - g^{3\alpha}\delta^{2}{}_{\beta}) \ \text{ and cyclic 123} \\ (K_1)^{\alpha}{}_{\beta} &\equiv (J^{01})^{\alpha}{}_{\beta} = i\,(\,g^{0\alpha}\delta^{1}{}_{\beta} - g^{1\alpha}\delta^{0}{}_{\beta}) \ \text{ and cyclic 123} \qquad (5.14.6) \end{aligned}$$

where $(-i)(J^{\mu\nu})^{\alpha\beta} = (\,g^{\mu\alpha}g^{\nu\beta} - g^{\nu\alpha}g^{\mu\beta})$ is a rank-4 tensor, antisymmetric under $\mu\leftrightarrow\nu$ and $\alpha \leftrightarrow \beta$.

An arbitrary Lorentz transformation can be represented as $F^{\mu}{}_{\nu}(\mathbf{r},\mathbf{b}) = [\,\exp\{-\,i\,(\,\mathbf{r}\bullet\mathbf{J} + \mathbf{b}\bullet\mathbf{K})\}\,]^{\mu}{}_{\nu}$ where the 6 numbers **r** and **b** are called parameters (rotation and boost) and this F is a combined boost/rotation transformation (note that $e^{A+B} \neq e^{A}e^{B}$ for non-commuting matrices A,B). The product of two such Lorentz transformations is also a Lorentz transformation, and in fact the transformations form a continuous group known as the Lorentz Group, which then has 6 parameters.

The first two commutators shown above (all J and all K) are each associated with a 3 parameter continuous group called the rotation group. The abstract generators of this group can be "represented" as matrices of any dimension, and are labeled by a number j such that 2j+1 is the matrix dimension. For example, the 2x2 matrix representation of the rotation group is labeled by j = 1/2, and is called the spinor representation and is associated in physics with the "intrinsic spin" of particles of spin 1/2 such as electrons. The vectors (spinors) in this case have two elements, and (1,0) and (0,1) are "up" and "down".

Representations of the *Lorentz* group have labels $\{j_1, j_2\}$, where j_1 is for the J-generated rotation subgroup, and j_2 for the K-generated rotation subgroup, and are usually denoted $j_1\otimes j_2$. Such a representation then has vectors containing $(2j_1+1)(2j_2+1)$ elements. In the case $1/2\otimes1/2$ there are 2*2=4 elements in a vector, and when these elements are linearly combined in a certain manner, they form the 4-vector object which one writes as A^{μ} such as x^{μ}. This is the "vector representation" of the Lorentz group upon which is built the entire edifice of special relativity tensor algebra.

Two other basic representations of the Lorentz group are these: $1/2 \otimes 0$ and $1/2 \otimes 0$. These are 2x2 matrix representations and they are *different* 2x2 representations. For each representation one can construct a whole tensor analysis world based on 2-vectors. Just as with the 4-vectors, one has contravariant and covariant 2-vectors. The two representations $1/2 \otimes 0$ and $1/2 \otimes 0$ are called spinor representations since they are each 2-dimensional. Since there are two distinct spinor representations, one needs some way of distinguishing them from each other. One representation might be called "undotted" and the other "dotted" and then there are *four* 2-vector types to worry about, which transform this way,

$$\begin{aligned} &V'^{a} = R^{a}{}_{b}V^{b} && V'^{\dot{a}} = R^{\dot{a}}{}_{\dot{b}}V^{\dot{b}} && \text{contravariant 2-vectors} \\ &V'_{a} = R_{a}{}^{b}V_{b} && V'_{\dot{a}} = R_{\dot{a}}{}^{\dot{b}}V_{\dot{b}} && \text{covariant 2-vectors} \qquad (5.14.7) \end{aligned}$$

where now dots on the indices indicate which Lorentz group representation that index belongs to. The 2x2 matrices $R^{a}{}_{b}$ and $R^{\dot{a}}{}_{\dot{b}}$ are not the same. A typical rank-2 tensor would transform this way,

$$X'^{a\dot{b}} = R^{a}{}_{a'}R^{\dot{b}}{}_{\dot{b}'}\, X^{a'\dot{b}'} \quad . \qquad (5.14.8)$$

This then is the subject of what is sometimes called Spinor Algebra as opposed to Tensor Algebra, but it is really just regular tensor algebra with respect to the two spinor representations of the Lorentz group.

We have inserted this blatant digression just to show that the general subject of tensor analysis includes all this spinor stuff under its general umbrella.

In closing, Maple shows that the metric tensor G is indeed preserved under boosts and rotations. In Maple,

evalm(Bx &* G &* transpose(Bx)) means $B_x\, G\, B_x^T$

and Maple is just verifying that $B_x\, G\, B_x^T = G$ and similarly $R_x\, G\, R_x^T = G$:

```
G := matrix(4,4,[1,0,0,0,0,-1,0,0,0,0,-1,0,0,0,0,-1]);
```

$$G := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

```
Bx := matrix(4,4,[cosh(b),sinh(b),0,0, sinh(b),cosh(b),0,0, 0,0,1,0, 0,0,0,1]);
```

$$Bx := \begin{bmatrix} \cosh(b) & \sinh(b) & 0 & 0 \\ \sinh(b) & \cosh(b) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

```
evalm(Bx &* G &* transpose(Bx));
```

$$\begin{bmatrix} \cosh(b)^2 - \sinh(b)^2 & 0 & 0 & 0 \\ 0 & \sinh(b)^2 - \cosh(b)^2 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

```
simplify(%);
```

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

```
Rx := matrix(4,4,[1,0,0,0, 0,1,0,0, 0,0,cos(r),-sin(r), 0,0,sin(r),cos(r)]);
```

$$Rx := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos(r) & -\sin(r) \\ 0 & 0 & \sin(r) & \cos(r) \end{bmatrix}$$

```
evalm(Rx &* G &* transpose(Rx));
```

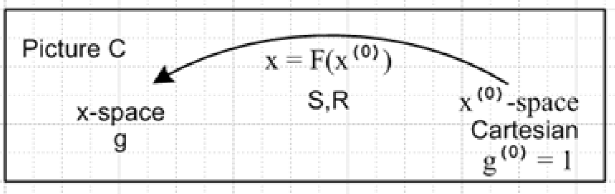
$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -\cos(r)^2 - \sin(r)^2 & 0 \\ 0 & 0 & 0 & -\cos(r)^2 - \sin(r)^2 \end{bmatrix}$$

```
simplify(%);
```

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \qquad (5.14.9)$$

5.15 General Relativity and its Metric Tensor

In general relativity a Picture of interest is Picture C



$$x^{(0)} \to \xi, \quad g^{(0)} \to \eta \qquad (5.15.1)$$

but the $x^{(0)}$-space is replaced by a Quasi-Cartesian [see (1.10)] space with coordinates ξ^i with the metric tensor of special relativity which is now called η.

This ξ-space represents a "freely-falling" coordinate system in which the laws of *special* relativity apply and the metric tensor is taken to be $G = \overline{G} = \mathrm{diag}(-1,1,1,1) \equiv \eta$.

The x^i are the coordinates of *some other* coordinate system. There is some transformation $\mathbf{x} = \mathbf{F}(\xi)$ which defines the relationship between these two systems. The covariant metric tensor in x-space is written $\overline{g} = S^TGS = S^T\eta S$. This is (5.7.6) for Picture C with $g^{(0)} = G = \eta$, "quasi-Cartesian". Using the Standard Notation introduced in Chapter 7 below, this is usually written as

$\overline{g} = S^T\eta S$ // translation of (5.7.6) to our quasi-Cartesian Picture C

$g_{dn} = S^T \eta_{dn} S$ // developmental notation, g_{dn} has "down" indices

$g_{\mu\nu} = (S^T)^{\mu}{}_{\alpha}\, \eta_{\alpha\beta}\, S^{\beta}{}_{\nu} = S^{\alpha}{}_{\mu}\, \eta_{\alpha\beta}\, S^{\beta}{}_{\nu}$ // Standard Notation as in Chapter 7

$$g_{\mu\nu} = (\partial\xi^\alpha/\partial x^\mu)\ \eta_{\alpha\beta}\ (\partial\xi^\beta/\partial x^\nu) \qquad // \text{ e.g., } S^\alpha{}_\mu = (\partial\xi^\alpha/\partial x^\mu)$$

$$g_{\mu\nu}(x) = (\partial\xi^\alpha/\partial x^\mu)\ (\partial\xi^\beta/\partial x^\nu)\ \eta_{\alpha\beta} \qquad // \text{ Weinberg p 71 (3.2.7) .} \qquad (5.15.2)$$

The last line then *defines* the gravitational metric tensor in x-space based on the transformation $\xi = F^{-1}(\mathbf{x}) = \xi(\mathbf{x})$. (This and the following references are from the book of Weinberg, see References.)

Newton's Second Law m**a** = **f** appears this way in general relativity,

$$m\ (\partial^2 x^\mu/\partial\tau^2) = f^\mu - m\ \Gamma^\mu{}_{\nu\lambda}\ (\partial x^\nu/\partial\tau)\ (\partial x^\lambda/\partial\tau) \qquad // \text{ p 123 (5.1.11 following)} \qquad (5.15.3)$$

where f^μ is an externally applied force, but there is then an extra bilinear velocity-dependent term which represents an effective gravitational force (it acts on mass m) arising from spacetime itself. The object $\Gamma^\mu{}_{\nu\lambda}$ is called the affine connection and is related to the metric tensor in this way.

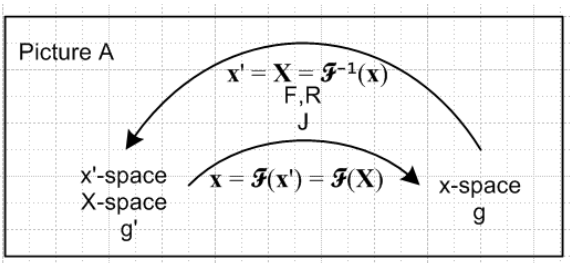
$$\Gamma^\mu{}_{\nu\lambda} = (1/2)\ g^{\mu\sigma}(\ \partial_\nu g_{\lambda\sigma} + \partial_\lambda g_{\nu\sigma} - \partial_\sigma g_{\nu\lambda}\ )\ . \qquad // \text{ p 75 (3.3.7)} \qquad (5.15.4)$$

These very brief comments are only meant to convince the reader that the equations of general relativity also have their place under the general umbrella of tensor analysis as discussed in this document. The fact that $\Gamma^\mu{}_{\nu\lambda}$ is not a mixed rank-3 tensor is demonstrated for example in (F.6.3).

5.16 Continuum Mechanics and its Metric Tensors

One can describe (Lai) the forward "flow" of a continuous blob of matter by **x** = **x**(**X**,t) where **X** = **x**(**X**,t_0). A "particle" of matter (imagine a tiny cube) that starts at location **X** at time t_0 ends up at **x** at time t. Two points in the flow separated by d**X** at t_0 end up separated by some d**x** at t. The relation between these separation vectors is given by d**x** = F d**X** where F is called the **deformation gradient**. F describes how a particle starting say with a cubic shape at t_0 gets deformed into some parallelepiped (3-piped) shape at t as shown in Fig (5.16.7) below. If we examine d**x** = F d**X** we find that | d**x** | ≠ | d**X** | since the vector d**X** typically gets rotated and stretched as d**X** → d**x** during the flow. [This flow is further described in Appendix K.]

The finite-duration flow **x** = **x**(**X**,t) from time t_0 to time t can be thought of as a (generally non-linear) transformation of the form **x** = $\mathcal{F}$(**X**,t), where we can regard t as an added parameter. Recall from Chapter 2 that a general transformation was annotated as **x**' = **F**(**x**) or **x** = $\mathbf{F}^{-1}$(**x**') and the linearized transformation was d**x**' = R d**x** or d**x** = S d**x**', as in (2.1.6). To be compatible with Lai notation which uses symbol F for the deformation gradient, and to make things clearer, we shall replace our Chapter 2 transformation name **F** by the name $\mathcal{F}^{-1}$, and we shall identify x'-space with X-space. We then have this modified version of Picture A,

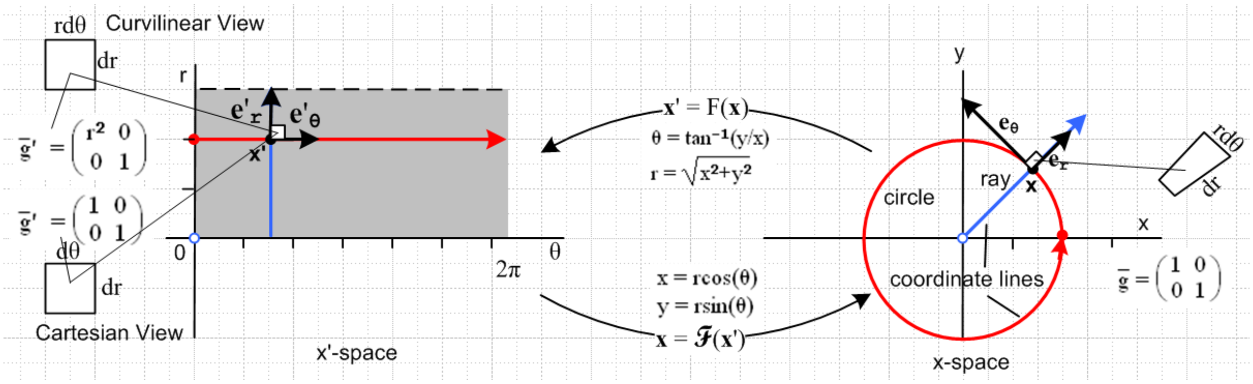


(5.16.1)

In this flow we assume Cartesian coordinates in both x-space and X-space, so the metric tensors which determine physical distance in these spaces are both 1. Before continuing, it is helpful to have an example of a fluid flow situation.

Example: A 2D Fluid Flow based on the Polar Coordinates Transformation

We first quote Fig (3.4.3) with a few enhancements ($\mathcal{F} \equiv \mathbf{F}^{-1}$) :



(3.4.3) (5.16.2)

The tiny white square on the left maps into a tiny white 2-piped shown on the right under this transformation. The edges of this 2-piped in x-space (on the right) are dr and rdθ, there is no confusion about that, and the area is dA = rdrdθ. The question at hand is how one should interpret the edges of the tiny white square on the *left*, and what is the x'-space metric tensor $\overline{g}'$ on the left? It is shown in Section C.5 that there are two distinct interpretations and each has its own use. We call the two interpretations the Cartesian View and the Curvilinear View.

The view we are familiar with is the **Curvilinear View** where $\overline{g}' = \begin{pmatrix} r^2 & 0 \\ 0 & 1 \end{pmatrix}$. In this view, the top edge of the white square on the left has length ds' = rdθ and this matches the corresponding edge of the 2-piped on the right, which is ds = rdθ. This is the view in which hypothesis (5.2.6) is valid (that ds' = ds), so all our tensor machinery is applicable, such as the fact (5.7.9) that $\overline{g}' = S^TS$. One might call this the covariant view, since distance ds is a scalar under the transformation. Here is a bit more detail.

For the upper edge of the white square on the left, we can set $d\mathbf{x}' = d\theta\, \mathbf{e}'_\theta = (d\theta\, r)\, \hat{\mathbf{e}}'_\theta$. This maps into the northeast edge of the 2-piped on the right which is then $d\mathbf{x} = (rd\theta)\, \hat{\mathbf{e}}_\theta$. One then verifies that

$$(ds')^2 = d\mathbf{x}' \bullet d\mathbf{x}' = g'_{\theta\theta}\, d\theta\, d\theta = r^2\, d\theta\, d\theta = (rd\theta)^2 \qquad \text{// Curvilinear View}$$

$$(ds)^2 = d\mathbf{x} \bullet d\mathbf{x} = (rd\theta)\, \hat{\mathbf{e}}_\theta \bullet (rd\theta)\, \hat{\mathbf{e}}_\theta = (rd\theta)^2 \qquad (5.16.3)$$

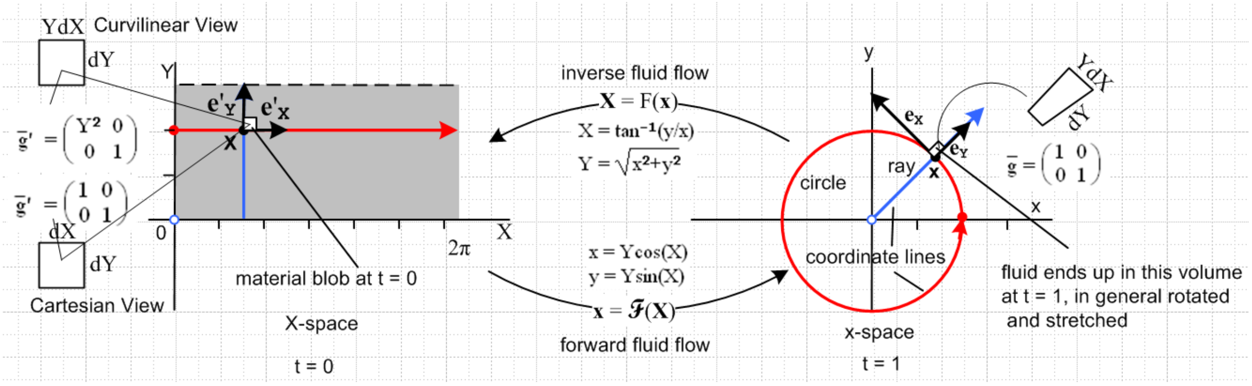
In the **Cartesian View**, we instead use $\bar{g}' = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ on the left. In this case, the edges of the white square are dr and dθ, just as if these were regular Cartesian coordinates. As explained in Appendix C, one advantage of the Cartesian View is that, for a general coordinate system, it is the only view that can be drawn on a piece of paper (2D) allowing the reader to have some comprehension of what is going on (and this is even more true in 3D). A disadvantage of the Cartesian View is that ds ≠ ds' so it is "non-covariant". In the Cartesian View we find, for our same top edge vector,

$$(ds')^2 = d\mathbf{x}' \bullet d\mathbf{x}' = 1\, d\theta\, d\theta = (d\theta)^2 \qquad \text{// Cartesian View}$$

$$(ds)^2 = d\mathbf{x} \bullet d\mathbf{x} = (rd\theta)\, \hat{\mathbf{e}}_\theta \bullet (rd\theta)\, \hat{\mathbf{e}}_\theta = (rd\theta)^2 \qquad (5.16.4)$$

and so ds ≠ ds' which says $|\, d\mathbf{x}\, | \neq |d\mathbf{x}'|$.

To arrive at our Example model for a fluid flow, we redraw Fig (5.16.2) renaming **x**' = (θ,r) to be **X** = (X,Y), and make corresponding changes elsewhere in the figure:



(5.16.5)

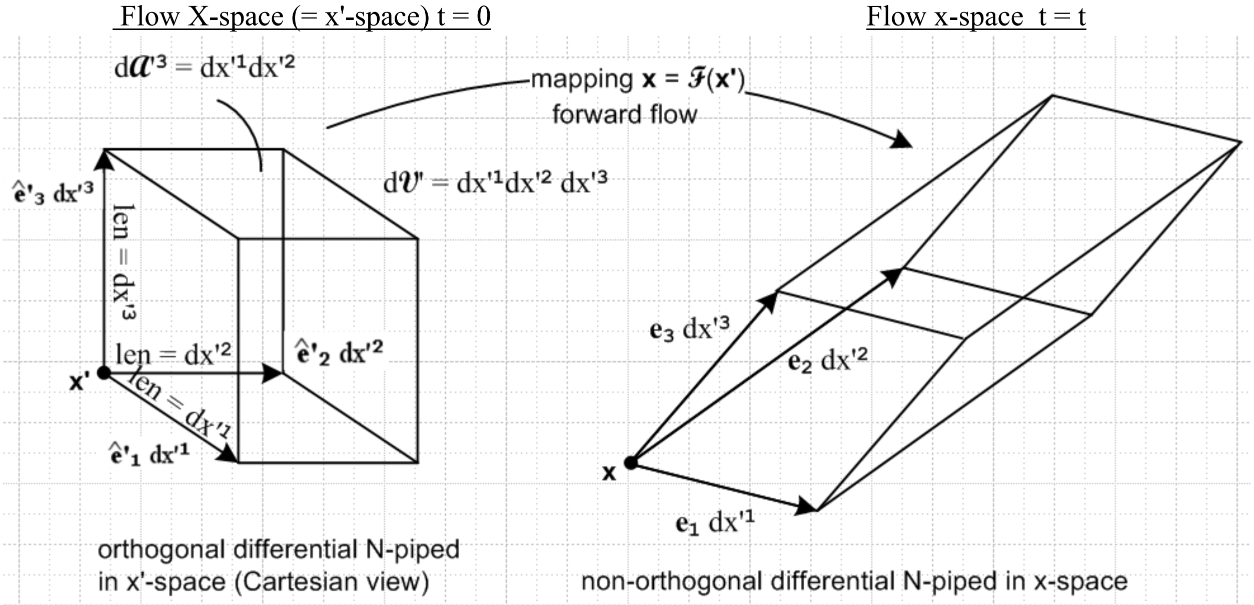
The tiny white square on the left is a blob of (2D) fluid at time t = 0 and after flowing a while it ends up in the location of the white 2-piped on the right at time t = 1. This is just a simple 2D flow example making use of a transformation we are already familiar with (see Comment below). For a general 3D forward flow transformation $\mathbf{x} = \mathcal{F}(\mathbf{X},t)$ one finds that the differential white cubic blob at t = 0 ends up in a 3-piped at t = t which is both rotated and stretched relative to the starting cube. Due to this stretch, one finds in general that $|d\mathbf{x}| \neq |d\mathbf{X}|$ where these are vector lengths relative to the Cartesian View metric tensor, which is in fact the metric tensor that determines physical distance in both X-space and x-space. Thus, the $\bar{g}' = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ metric tensor for X space has this well-defined meaning -- it relates to actual physical distance in X-space. However, it is the curvilinear view metric tensor $\bar{g}' = \begin{pmatrix} r^2 & 0 \\ 0 & 1 \end{pmatrix} = S^TS$ whose matrix elements determine such things as how much edges stretch in the flow and how much areas and volumes change during the flow, which we shall examine below.

<u>Comment:</u> The Example presented above really makes no flow sense since there is no time parameter in the flow transformation! The purpose of the Example is merely to illustrate the idea of two metric tensors in X-space, and to do so with a transformation that is familiar from our "curvilinear coordinates" study, and to tie in the notions of the two Views from Appendix C. The Example could be made more reasonable using a contrived transformation like this,

$$\begin{array}{llll} & \underline{\text{general } t} & \underline{t=0} & \underline{t=1} \\ \mathbf{x} = \mathcal{F}(\mathbf{X},t): & x = (1-t)X + t\,Y\cos(X) & x = X & x = Y\cos(X) \\ & y = (1-t)Y + t\,Y\sin(X) & y = Y & y = Y\sin(X)\ . \end{array} \qquad (5.16.6)$$

In this case, at $t = 0$ we have the correct $\mathbf{x} = \mathcal{F}(\mathbf{X},t=0) = \mathbf{X}$ and at the end of the flow ($t = 1$) we have the transformation equations shown in Fig (5.16.5).

Having worked through this contrived example, we can now draw a figure showing the general mapping of a differential blob during a flow from time $t = 0$ to time $t = t$. This figure is taken from Chapter 8 where the X-space left side is called x'-space as in (15.6.1) and the initial "blob" is shown drawn in the Cartesian View: (this picture happens to use the Chapter 7 Standard Notation for coordinates, contravariant = upper index)



(8.2.2) (5.16.7)

We might be interested in knowing how much the dx'_3 edge is stretched going from time t=0 to time t=t, and by how much the volume changes, and so on. All these geometric questions are posed and answered in Chapter 8, and we shall use some of those results below.

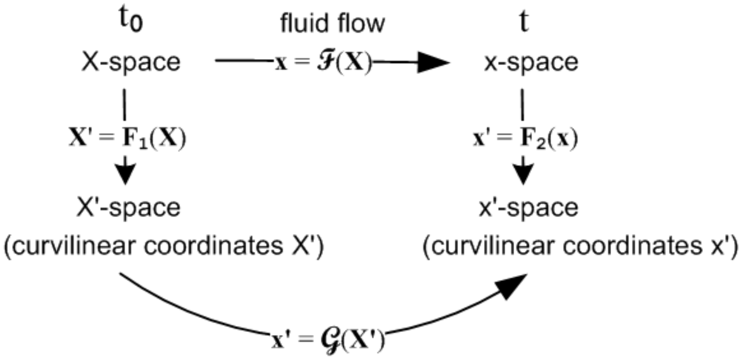
Here then is a translation table comparing the continuum mechanics notation of Lai to the tensor notation used earlier in our document (and in the above Example).

continuum mechanics		our document	(Forward Flow $\mathbf{X} \rightarrow \mathbf{x}$)
$\mathbf{X}, \mathbf{x}$	↔	$\mathbf{x}', \mathbf{x}$	
$\mathbf{x} = \mathbf{x}(\mathbf{X},t) = \mathcal{F}(\mathbf{X},t)$	↔	$\mathbf{x} = \mathcal{F}(\mathbf{x}',t) = \mathbf{F}^{-1}(\mathbf{x}',t)$	// Lai p 70 (3.1.4)
$d\mathbf{x} = F\, d\mathbf{X}$	↔	$d\mathbf{x} = S\, d\mathbf{x}'$	// as in (2.1.6) // Lai p 86 (3.7.6) or p 105 (3.18.3)
F	↔	S	
F^{-1}	↔	R	// $R = S^{-1}$
$\mathbf{x}$ = Cartesian	↔	$\bar{g} = 1$	
$\mathbf{X}$ = Cartesian	↔	$\bar{g}' = 1$	// Cartesian View
$C = F^T F$	↔	$\bar{g}' = S^T S$	// Curvilinear View, as in (5.7.9) // Lai p 114 (3.23.2)
$C^{-1} = F^{-1}(F^{-1})^T$	↔	$g' = RR^T$	// since $g' = \bar{g}'^{-1}$ and $S^{-1} = R$, as in (5.7.9)

(5.16.8)

Thus, the deformation gradient F is just the S matrix associated with transformation $\mathbf{F} = \mathcal{F}^{-1}$. The Curvilinear View metric tensor $\bar{g}' = S^T S$ appears in Lai as $C = F^T F$ which is known as the right Cauchy-Green deformation tensor (manifestly symmetric, so a viable metric tensor). [The *left* Cauchy-Green deformation tensor is $B = FF^T$].

Given the above flow situation, it is then possible to add two more transformations $\mathbf{F}_1$ and $\mathbf{F}_2$ which take X-space and x-space to independent sets of curvilinear coordinates $\mathbf{X}'$ and $\mathbf{x}'$: (a new $\mathbf{x}'$ here)



(5.16.9)

and we then have an interesting triple application of the notions of Chapter 2 to a real-world situation. This drawing is the implicit subject of Section 3.29 (p131-138) of Lai. The flow transformation of interest here is

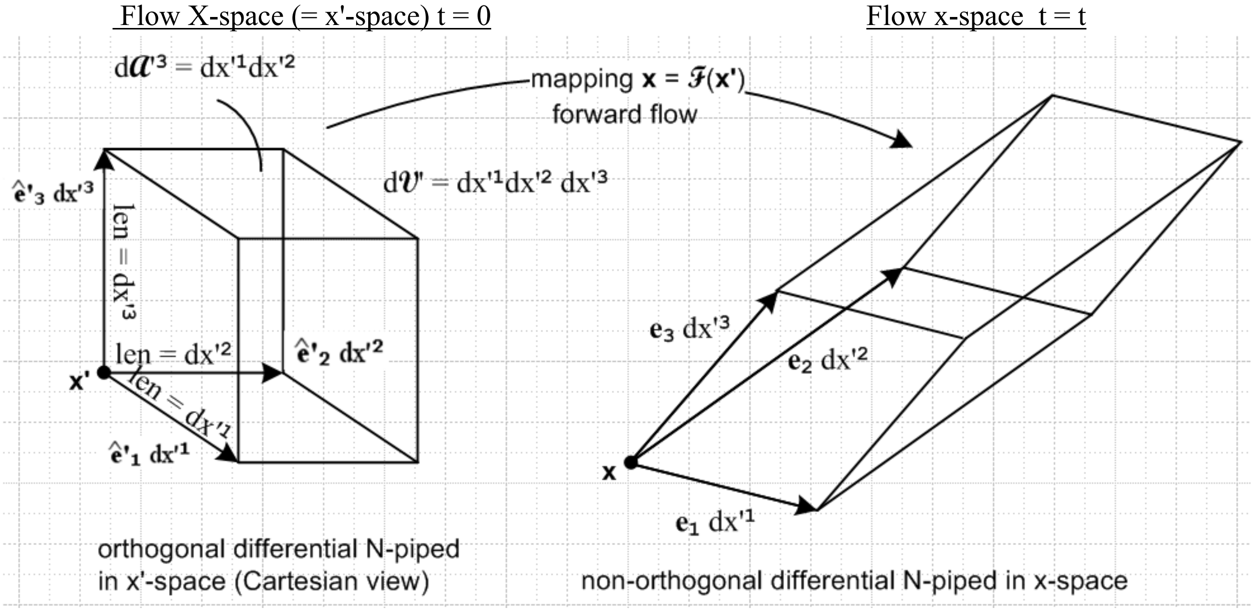
$$\mathbf{x}' = \mathbf{F}_2(\mathbf{x}) = \mathbf{F}_2(\mathcal{F}(\mathbf{X})) = \mathbf{F}_2(\mathcal{F}(\mathbf{F}_1^{-1}(\mathbf{X}'))) \equiv \mathcal{G}(\mathbf{X}') . \tag{5.16.10}$$

In (reverse) dyadic notation the deformation gradient is written $F = (\nabla\mathbf{x})$ where ∇ means $\nabla^{(\mathbf{X})}$ so that

$$d\mathbf{x} = F\, d\mathbf{X} = (\nabla\mathbf{x})\, d\mathbf{X} \qquad F_{ij} = (\nabla\mathbf{x})_{ij} = \partial_j^{(\mathbf{X})} x_i = \partial x_i / \partial X_j . \tag{5.16.11}$$

The (∇**x**) notation is explained in (E.4.4), and in Appendix G the object (∇**v**) for an arbitrary vector field **v**(**x**) is expressed in general curvilinear coordinates.

Consider again Fig (5.16.7) displayed above:



(8.2.2) *(5.16.7)*

We can identify the mapping shown in this picture with our Flow situation of table (5.16.8). Chapter 8 discusses in much detail how length, volume and area transform under a general transformation. The length, area and volume magnitudes on the left are called $d\mathcal{L}'^{n} = dx'^{(n)}$, $d\mathcal{A}'^{n}$ and $d\mathcal{V}'$, while the corresponding quantities on the right are called $d\mathbf{x}^{(n)}$, $\overline{d\mathbf{A}}^{(n)}$ and dV, where the first two items are vectors. Cribbing the results of (8.4.g.2) and converting them from standard notation to developmental notation, we have

$$|\, d\mathbf{x}^{(n)}| / d\mathcal{L}'^{n} = h'_n \;=\; [\bar{g}'_{nn}]^{1/2} \;=\; \text{the scale factor for edge } d\mathbf{x}^{(n)}$$

$$|\, \overline{d\mathbf{A}}^{(n)}| / d\mathcal{A}'^{n} \;=\; (1/h'_n)\,|J| \;=\; (1/h'_n)\, g'^{1/2} = [g'_{nn}\, g']^{1/2} \;=\; [\mathrm{cof}(\bar{g}'_{nn})]^{1/2}$$

$$|dV| / d\mathcal{V}' = |J| = g'^{1/2} \qquad // \text{ where } g' \equiv \det(\bar{g}'_{ij}) \;=\; J^2\,, \qquad \bar{g}' = S^T S \qquad \textit{(8.4.g.2)} \qquad (5.16.12)$$

We can then translate these three lines into our Flow context:

$$|\, d\mathbf{x}^{(n)} | \,/\, |\, d\mathbf{X}^{(n)} | \;=\; h'_n = [\bar{g}'_{nn}]^{1/2} = \;[(F^TF)_{nn}]^{1/2} = \;[C_{nn}]^{1/2} \qquad // \text{ Lai p 114 (3.23.6-8)}$$

$$d\mathbf{A}^{n}| \,/\, |d\mathcal{A}_0{}^{n}| = [g'_{nn} g']^{1/2} \;=\; [\mathrm{cof}(\bar{g}'_{nn})]^{1/2} = [\mathrm{cof}(F^TF)_{nn})]^{1/2} \;=\; [\mathrm{cof}\, C_{nn})]^{1/2} \qquad // \text{ Lai p 129 (3.27.11)*}$$

$$|dV| \,/\, |dV_0| = |J| = g'^{1/2} = [\det(\bar{g}'_{ij})]^{1/2} = [\det(F^TF)]^{1/2} \;=\; |\det(F)| \qquad // \text{ Lai p 130 (3.28.3)} \qquad (5.16.13)$$

where

	edge	area	volume		
X-space :	$d\mathbf{X}^{(n)}$	$d\boldsymbol{a}_0{}^{n}$	dV_0	time t_0	
x-space :	$d\mathbf{x}^{(n)}$	$d\mathbf{A}^{n}$	dV	time t	(5.16.14)

Thus, for example, the volume change of a "flowing" particle of continuous matter is given by the Jacobian |J| = |detF| associated with the deformation gradient tensor F. We put quotes on "flowing" only because this might be a particle of solid steel that is momentarily moving and deforming a very small amount during an oscillation or in response to an applied stress.

In the middle line of (5.16.13) we state that $|\ d\mathbf{A}^{n}| \ / \ |\ d\boldsymbol{a}_0{}^{n}| = \ [\mathrm{cof}(F^TF)_{nn})]^{1/2}$ and quote Lai p 129 (3.27.11) for verification. However, what Lai (3.27.11) actually says (slightly translated to our notation) is this:

$$dA^{(n)}/d\boldsymbol{a}^{(n)}{}_0 \ = \ \det(F)\ |\ (F^{-1})^T\ \mathbf{u}_n\ | \qquad\qquad \mathbf{u}_n = \text{unit base vector},\ \ (\mathbf{u}_n)_i = \delta_{n,i} \qquad (5.16.15)$$

which seems a far cry from our result $[\mathrm{cof}(F^TF)_{nn})]^{1/2}$. But consider, using $F^{-1} = R$ from table (5.16.8),

$$|\ (F^{-1})^T\ \mathbf{u}_n\ |^2 = |\ R^T\ \mathbf{u}_n\ |^2 \ = [R^T\mathbf{u}_n]_i[R^T\mathbf{u}_n]_i = R_{ni}\ R_{ni} = (RR^T)_{nn} \ = g'_{nn} \qquad (5.16.16)$$

so

$$\det(F)\ |\ (F^{-1})^T\ \mathbf{u}_n\ | \ = g'^{1/2}\ g'_{nn}{}^{1/2} \ = [\mathrm{cof}(\bar{g}'_{nn})]^{1/2} \ = \ [\mathrm{cof}(F^TF)_{nn})]^{1/2}\ . \qquad (5.16.17)$$

Here det(F) = det(S) = $g'^{1/2}$ from (5.12.19) and we have used Theorem 1 (8.4.f.1) converted to developmental notation,

$$g'\ g'_{nn} \ = [\mathrm{cof}(\bar{g}'_{nn})]\ . \qquad // \text{ more generally, } \ (\det A)\ A^{-1} = \mathrm{cof}(A) \ \text{ if } A = A^T \qquad (5.16.18)$$

It might be noted that the Lai book does in fact use our "developmental notation" in that all indices are written "down" (when indices are shown), but no overbars mark covariant objects. Here are a few examples:

Lai notation		Developmental notation	Standard Notation
$d\mathbf{A}_0 = d\mathbf{X}^{(1)} \times d\mathbf{X}^{(2)}$	(3.27.1)	$\overline{d\mathbf{A}}_0 = d\mathbf{X}^{(1)} \times d\mathbf{X}^{(2)}$	$(dA_0)_i = \varepsilon_{ijk}\ [d\mathbf{X}^{(1)}]^j\ [d\mathbf{X}^{(1)}]^k$
$[\mathrm{div}T]_i = \partial_j T_{ij}$	(4.7.3)	$[\mathrm{div}T]_i = \bar{\partial}_j T_{ij}$	$[\mathrm{div}T]^i = \partial_j T^{ij}$

(5.16.19)

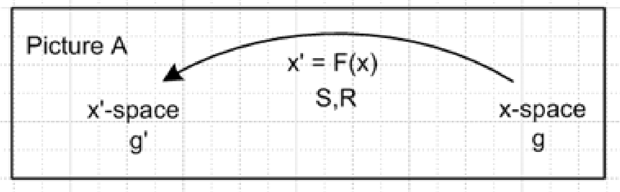
Of course when Cartesian coordinates are assumed, the up and down position makes no difference.

Lai writes tensors in bold face such as **F** for the deformation gradient noted above, or **T** for the stress tensor. Perhaps this is done to emphasize the notion of a tensor as an operator, as in our Section E.7. Lai writes a specific matrix as [**T**], but a matrix element is T_{ij}. Notation is an ongoing burden and each area of physics seems to have its own accepted conventions.

6. Reciprocal Base Vectors $\mathbf{E}_n$ and Inverse Reciprocal Base Vectors $\mathbf{U}'_n$

6.1 Definition of the $\mathbf{E}_n$

This entire Chapter uses the Picture A context,



(6.1.1)

Although various definitions are possible, we shall define the reciprocal tangent vectors $\mathbf{E}_n$ in the following manner [implied sum on i, and $\mathbf{e}_n = \partial'_n\mathbf{x}$ from (3.2.6)]

$$\mathbf{E}_n \equiv g'_{ni}\,\mathbf{e}_i = g'_{ni}\,\partial'_i\mathbf{x} \qquad \Rightarrow \qquad \mathbf{e}_n = \bar{g}'_{ni}\mathbf{E}_i\ . \qquad // \text{ since } \bar{g}' = g'^{-1} \qquad (6.1.2)$$

<u>Comment</u>: Notice how this differs in structure from the rule for forming a covariant vector from a contravariant one,

$$\bar{V}_n = \bar{g}'_{ni}V_i\ . \qquad (6.1.3)$$

In the last equation, the right side is a linear combination of vector *components* V_i, while in the previous equation (6.1.2) the right side is a linear combination of *vectors* $\mathbf{e}_i$. In this case, i is a label on $\mathbf{e}_i$, whereas in the other case i is an index on V_i. Labels and indices are different.

Since the tangent base vectors $\mathbf{e}_i$ are contravariant vectors in x-space by (3.3.2), and since $\mathbf{E}_n$ is a linear combination of the $\mathbf{e}_i$, the $\mathbf{E}_n$ are also contravariant vectors in x-space. Notice that in the definition $\mathbf{E}_n \equiv g'_{ni}\,\mathbf{e}_i$, these two x-space vectors are related by the metric tensor of the *other* space, x'-space.

One can express the components of $\mathbf{E}_n$ in two ways,

$$\begin{aligned}(\mathbf{E}_n)_k &\equiv g'_{ni}\,(\mathbf{e}_i)_k = S_{ki}g'_{ni} && // \ (\mathbf{e}_i)_k \equiv S_{ki} \text{ by (3.2.5)}\\ &= g'_{ni}S_{ki} = R_{ni}g_{ik} && // \ R_{ni}g_{ik} = g'_{ni}(S^T)_{ik} = g'_{ni}S_{ki} \text{ by (5.7.8) } Rg = g'\,S^T \qquad (6.1.4)\end{aligned}$$

Taking i→c and then k→i (and g is symmetric),

$$(\mathbf{E}_n)_i = g'_{nc}S_{ic} = g_{ic}R_{nc}\ . \qquad // \text{ sum on second indices} \qquad (6.1.5)$$

Applying R to both sides of $\mathbf{E}_n \equiv g'_{ni}\,\mathbf{e}_i$ in (6.1.2) gives $\mathbf{E}_n$ transformed into x'-space,

$$\mathbf{E}'_n = g'_{nk}\,\mathbf{e}'_k$$

so

$$(\mathbf{E}'_n)_i = g'_{nk}\,(\mathbf{e}'_k)_i = g'_{nk}\delta_{k,i} = g'_{ni}\ . \qquad (6.1.6)$$

Using the axis-aligned basis vectors $\mathbf{u}_n$ in x-space, $(u_n)_m = \delta_{n,m}$, we also have this dot product,

$$\mathbf{E}_n \bullet \mathbf{u}_m = R_{nm} \ . \tag{6.1.7}$$

<u>Proof</u>: $\mathbf{E}_n \bullet \mathbf{u}_m = \overline{g}_{ab}(\mathbf{E}_n)_a(u_m)_b = \overline{g}_{ab}\,(g_{ac}R_{nc})(\delta_{m,b}) = (\overline{g}_{ma}g_{ac})\,R_{nc} = \delta_{m,c}\,R_{nc} = R_{nm}$.

6.2 The $\mathbf{e}_n$ and $\mathbf{E}_n$ Dot Products and Reciprocity (Duality)

Three covariant dot products are of great interest. The first is this (see (5.7.6) for last step)

$$\mathbf{e}_n \bullet \mathbf{e}_m = \overline{g}_{ij}\,(\mathbf{e}_n)_i\,(\mathbf{e}_m)_j = \overline{g}_{ij}\,S_{in}\,S_{jn} = S^T{}_{ni}\,\overline{g}_{ij}\,S_{jm} = (\,S^T\,\overline{g}\,S)_{nm} = \overline{g}'_{nm} \ . \tag{6.2.1}$$

The second is

$$\mathbf{E}_n \bullet \mathbf{e}_m = \overline{g}_{ij}\,(\mathbf{E}_n)_i\,(\mathbf{e}_m)_j = \overline{g}_{ij}\,g_{ic}\,R_{nc}\,S_{jm} = \delta_{j,c}\,R_{nc}\,S_{jm} = R_{nj}\,S_{jm} = (RS)_{nm} = \delta_{n,m} \tag{6.2.2}$$

and the third is

$$\mathbf{E}_n \bullet \mathbf{E}_m = \overline{g}_{ij}\,(\mathbf{E}_n)_i\,(\mathbf{E}_m)_j = \overline{g}_{ij}\,g_{ic}\,R_{nc}\,g_{jb}\,R_{mb} = \delta_{j,c}\,R_{nc}\,g_{jb}\,R_{mb} = R_{nj}\,g_{jb}\,R_{mb}$$

$$= R_{nj}\,g_{jb}\,R^T{}_{bm} = (R\,g\,R^T)_{nm} = g'_{nm} \ . \quad \text{// last step is (5.7.6)} \tag{6.2.3}$$

To summarize the last three results and (6.1.2),

$$\begin{array}{llll}
\mathbf{e}_n \bullet \mathbf{e}_m = \overline{g}'_{nm} & \Rightarrow \quad |\mathbf{e}_n| = \sqrt{\overline{g}'_{nn}} = h'_n \quad \text{(scale factor)} & & \mathbf{E}_n = g'_{ni}\,\mathbf{e}_i \\
\mathbf{E}_n \bullet \mathbf{e}_m = \delta_{n,m} & & & \mathbf{e}_n = \overline{g}'_{ni}\,\mathbf{E}_i \\
\mathbf{E}_n \bullet \mathbf{E}_m = g'_{nm} & \Rightarrow \quad |\mathbf{E}_n| = \sqrt{g'_{nn}} \quad . & &
\end{array} \tag{6.2.4}$$

When the relation between two sets of vectors $\mathbf{e}_n$ and $\mathbf{E}_m$ is as shown in (6.2.4), the vectors are said to be **reciprocal** or **dual** to each other. Since the $\mathbf{e}_n$ already have the name "tangent base vectors", we refer to the $\mathbf{E}_n$ as the "reciprocal base vectors". The notion of reciprocal or dual vector sets is studied in more detail in the notes just below.

Notice that the two relations on the right in (6.2.4) are implied by the three dot product relations on the left. For example, since the $\mathbf{e}_n$ are a complete set, we can expand $\mathbf{E}_n$ with temporarily unknown coefficients α_{nk} to get $\mathbf{E}_n = \alpha_{nk}\mathbf{e}_k$. Dotting both sides with $\mathbf{E}_m$ then reveals the coefficients:

$$g'_{nm} = \mathbf{E}_n \bullet \mathbf{E}_m = (\alpha_{nk}\mathbf{e}_k) \bullet \mathbf{E}_m = \alpha_{nk}\,(\mathbf{e}_k \bullet \mathbf{E}_m) = \alpha_{nk}\,\delta_{k,m} = \alpha_{nm}$$

and therefore we find $\mathbf{E}_n = \alpha_{nk}\mathbf{e}_k = g'_{nk}\mathbf{e}_k = g'_{ni}\mathbf{e}_i$ as shown top right in (6.2.4).

Whereas $\mathbf{e}_n$ and $\mathbf{E}_m$ exist in x-space, $\mathbf{e}'_n = R\mathbf{e}_n$ and $\mathbf{E}'_m = R\mathbf{E}_m$ exist in x'-space -- these are just the transformed contravariant vectors using the rule (2.3.2), $\mathbf{v}' = R(\mathbf{x})\mathbf{v}$. Recall from (3.2.1) that the $\mathbf{e}'_n$ are axis-aligned unit vectors in x'-space, $(\mathbf{e}'_n)_i = \delta_{n,i}$.

Using the fact (6.1.6) that $(\mathbf{E}'_n)_i = g'_{nk}\,(\mathbf{e}'_k)_i$ we find that,

$$\mathbf{E}'_n \bullet \mathbf{e}'_m = \bar{g}'_{ij}\,(\mathbf{E}'_n)_i(\mathbf{e}'_m)_j = \bar{g}'_{ij}\ g'_{ni}\,\delta_{m,j} = \bar{g}'_{im}\ g'_{ni} = \delta_{n,m}\ . \tag{6.2.5}$$

This is consistent with (6.2.4) and (5.10.2) that $\mathbf{A}' \bullet \mathbf{B}' = \mathbf{A} \bullet \mathbf{B}$,

$$\mathbf{E}'_n \bullet \mathbf{e}'_m = \mathbf{E}_n \bullet \mathbf{e}_m = \delta_{n,m}\ . \tag{6.2.6}$$

The other two dot products above work this same way. We can then rewrite (6.2.4) for the primed vectors,

$$\begin{array}{llll}
\mathbf{e}'_n \bullet \mathbf{e}'_m = \bar{g}'_{nm} & \Rightarrow & |\mathbf{e}'_n| = \sqrt{\bar{g}'_{nn}} = h'_n \quad \text{(scale factor)} & \mathbf{E}'_n = g'_{ni}\,\mathbf{e}'_i \\
\mathbf{E}'_n \bullet \mathbf{e}'_m = \delta_{n,m} & & & \mathbf{e}'_n = \bar{g}'_{ni}\ \mathbf{E}'_i \\
\mathbf{E}'_n \bullet \mathbf{E}'_m = g'_{nm} & \Rightarrow & |\mathbf{E}'_n| = \sqrt{g'_{nn}}\ . &
\end{array} \tag{6.2.7}$$

As an exercise, we can verify the top left equation in this manner, knowing $(\mathbf{e}'_n)_i = \delta_{n,i}$,

$$\mathbf{e}'_n \bullet \mathbf{e}'_m = \bar{g}'_{ij}\,(\mathbf{e}'_n)_i\,(\mathbf{e}'_m)_j = \bar{g}'_{ij}\,\delta_{n,i}\,\delta_{m,j} = \bar{g}'_{nm}$$

The vectors $\mathbf{e}_n$ and $\mathbf{E}_n$ are examples of the general notion of dual vector sets which we now describe.

<u>Notes on Reciprocity (Duality)</u>

1. Suppose some set of vectors $\mathbf{b}_n$ forms a complete basis for x-space. Can one find a set of vectors $\mathbf{B}_n$ that have the property

$$\mathbf{B}_m \bullet \mathbf{b}_n = \delta_{m,n}\ ? \qquad\qquad \text{// duality relation; } \mathbf{B}_a \text{ and } \mathbf{b}_a \text{ are reciprocal} \tag{6.2.8}$$

As shown below, the answer is normally "yes", and the vectors $\mathbf{B}_n$ are uniquely determined by the $\mathbf{b}_n$. One says that the set $\{\mathbf{B}_n\}$ is "dual to" the set $\{\mathbf{b}_n\}$ and vice versa. If we regard $\mathbf{b}_n$ as a basis, then $\mathbf{B}_n$ is the "dual basis", and $\mathbf{B}_m \bullet \mathbf{b}_n = \delta_{m,n}$ is the "duality relation". Another terminology is that the vectors $\mathbf{B}_n$ are "reciprocal to" the vectors $\mathbf{b}_n$ and vice versa.

2. It was shown in (6.2.2) that $\mathbf{E}_n \bullet \mathbf{e}_m = \delta_{n,m}$ so the $\mathbf{E}_n$ and the $\mathbf{e}_n$ vectors are dual to each other. The $\mathbf{e}_n$ are the tangent base vectors, and the $\mathbf{E}_n$ are "reciprocal to" the $\mathbf{e}_n$ which is why we call them the reciprocal base vectors.

3. One can solve for the $\mathbf{B}_n$ in terms of the $\mathbf{b}_n$. Each $\mathbf{B}_n$ has N components, so there are N^2 unknowns. The duality relation $\mathbf{B}_m \bullet \mathbf{b}_n = \delta_{m,n}$ is a set of N^2 equations. This is basically a Cramer's Rule problem in

N^2 variables. Since the $\mathbf{b_n}$ form a complete basis, one can expand $\mathbf{B_m}$ on the $\mathbf{b_n}$ with some coefficients we will call w'_{mn} (at this point w'_{mn} is unknown),

$$\mathbf{B_m} = w'_{mn}\mathbf{b_n} \,. \tag{6.2.9}$$

Then from (6.2.8),

$$\delta_{m,k} = \mathbf{B_m} \bullet \mathbf{b_k} = w'_{mn}\ \mathbf{b_n} \bullet \mathbf{b_k} \quad . \qquad // \ \mathbf{b_n} \bullet \mathbf{b_k} = \bar{g}_{ij}(b_n)_i(b_k)_j \ . \tag{6.2.10}$$

Define matrix W' by,

$$W'_{nk} \equiv \mathbf{b_n} \bullet \mathbf{b_k} \tag{6.2.11}$$

and note that W'_{nk} is symmetric. Then (6.2.10) says

$$\delta_{m,k} = w'_{mn}W'_{nk} \quad \text{or} \quad w'W' = 1 \quad \text{or} \quad w' = W'^{-1} \,. \tag{6.2.12}$$

Assuming for the moment that $\det W' \neq 0$, the solution is given by $w' = W'^{-1}$. The (5.6.4) "Digression" showed that $(A^{-1})^T = (A^T)^{-1}$ for invertible A, so $w'^T = (W'^{-1})^T = (W'^T)^{-1} = W'^{-1} = w'$ and therefore w' is symmetric as well. Since W' is known from (6.2.11), $w' = W'^{-1}$ and the $\mathbf{B_m} = w'_{mn}\mathbf{b_n}$ of (6.2.9) have been found. Finally,

$$\mathbf{B_m} \bullet \mathbf{B_n} = w'_{mi}\mathbf{b_i} \bullet w'_{nj}\mathbf{b_j} = w'_{mi}w'_{nj}\ \mathbf{b_i} \bullet \mathbf{b_j} = w'_{mi}w'_{nj}W'_{ji} = w'_{mi}\delta_{ni} = w'_{nm} \,. \tag{6.2.13}$$

4. In the case that $\mathbf{b_n} = \mathbf{e_n}$ and $\mathbf{B_n} = \mathbf{E_n}$, one finds that $W'_{nk} = \mathbf{e_n} \bullet \mathbf{e_k} = \bar{g}'_{nk}$, the covariant metric tensor. Then $w' = W'^{-1}$ must be the contravariant metric tensor $w'_{mn} = g'_{mn} = \mathbf{E_m} \bullet \mathbf{E_n}$. Then $\mathbf{B_m} = w'_{mn}\mathbf{b_n}$ says $\mathbf{E_n} = g'_{mn}\ \mathbf{e_n}$ which agrees with our original definition (6.1.2) of the $\mathbf{E_n}$. Moreover, $\det(W') = \det(\bar{g}'_{ab}) = g'$ of (5.12.14), so as long as $g' \neq 0$, one has $\det W' \neq 0$.

5. For the general $\mathbf{b_n}$ case, one can imagine that the $\mathbf{b_n}$ are the tangent base vectors for *some* transformation $\mathbf{x}'' = \mathbf{F_b}(\mathbf{x})$ (with some linearized $R_b \neq R$ where R goes with $\mathbf{F}$), then the issue of $\det W' \neq 0$ boils down to $g' \neq 0$ for that transformation. In the case that x-space is Cartesian, we know that $J^2 = \det(\bar{g}')/\det(\bar{g}) = g'/g = g'$. If this transformation $\mathbf{F_b}$ is invertible, then $J \neq 0$ everywhere, so $g' \neq 0$ and then $\det W \neq 0$.

6. If the $\mathbf{b_m}$ are true tensorial vectors, then the $\mathbf{B_m} = w'_{mn}\mathbf{b_n}$ will be as well (linear combination) and then $\mathbf{B_n} \bullet \mathbf{b_m}$ is a tensorial scalar. Therefore if $\mathbf{B_n} \bullet \mathbf{b_m} = \delta_{n,m}$ in x-space, then so also $\mathbf{B'_n} \bullet \mathbf{b'_m} = \delta_{n,m}$ in x'-space, where $\mathbf{b_m}' = R\mathbf{b_m}$ and $\mathbf{B'_n} = R\mathbf{B_n}$. For example, $\mathbf{E'_n} \bullet \mathbf{e'_m} = \delta_{n,m}$ in x'-space where $\mathbf{e_m}' = R\mathbf{e_m}$ and $\mathbf{E'_n} = R\mathbf{E_n}$.

7. In Section 6.5 we shall encounter another dual pair $\mathbf{U_n} \bullet \mathbf{u_m} = \mathbf{U'_n} \bullet \mathbf{u'_m} = \delta_{n,m}$ which is associated with the inverse transformation $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$.

8. One major significance of the equation $\mathbf{B}_n \bullet \mathbf{b}_m = \delta_{n,m}$ is that it allows the following expansions:

$$\mathbf{V} = \Sigma_n \; k_n \, \mathbf{B}_n \qquad \text{where } k_m = \mathbf{V} \bullet \mathbf{b}_m$$
$$\mathbf{V} = \Sigma_n \; c_n \, \mathbf{b}_n \qquad \text{where } c_m = \mathbf{V} \bullet \mathbf{B}_m \qquad (6.2.14)$$

so that for example $\mathbf{b}_m \bullet \mathbf{V} = \mathbf{b}_m \bullet [\Sigma_n \; k_n \, \mathbf{B}_n] = \Sigma_n \; k_n \, \mathbf{b}_m \bullet \mathbf{B}_n = \Sigma_n \; k_n \, \delta_{m,n} = k_m$. These expansions are explored in Section 6.6 below for the two dual sets $\mathbf{E}_n$, $\mathbf{e}_n$ and $\mathbf{U}_n$, $\mathbf{u}_n$.

9. Saying that the basis $\mathbf{b}_n$ is *complete* implies that any vector $\mathbf{V}$ can be expanded in this basis. Consider the second expansion shown in (6.2.14). Installing $c_n = \mathbf{V} \bullet \mathbf{B}_n$ into the expansion one finds,

$$\mathbf{V} = \Sigma_n \; (\mathbf{V} \bullet \mathbf{B}_n) \, \mathbf{b}_n$$

so

$$V_b = \Sigma_n \; [\Sigma_a V_a (\overline{\mathbf{B}}_n)_a] \, (\mathbf{b}_n)_b \qquad \text{// from (5.10.3)}$$

$$= \Sigma_a V_a \; \{ \, \Sigma_n (\overline{\mathbf{B}}_n)_a \, (\mathbf{b}_n)_b \, \} \; .$$

The last equation can only be valid if

$$\Sigma_n (\overline{\mathbf{B}}_n)_a \, (\mathbf{b}_n)_b = \delta_{a,b} \qquad \text{// completeness relation for the dual set } \{\mathbf{b}_n, \mathbf{B}_n\} \qquad (6.2.15)$$

and this is the formal statement that the basis $\mathbf{b}_n$ is complete.

We can verify this completeness claim for $\{\mathbf{e}_n, \mathbf{E}_n\}$ as follows

$$\Sigma_n \, (\overline{\mathbf{E}}_n)_a (\mathbf{e}_n)_b = \Sigma_n \, (\bar{g} \, \mathbf{E}_n)_a (\mathbf{e}_n)_b = \Sigma_{nc} \; \bar{g}_{ac} (\mathbf{E}_n)_c (\mathbf{e}_n)_b \qquad \text{// from (5.8.1)}$$

$$= \Sigma_{nc} \; \bar{g}_{ac} \; [\, \Sigma_k \, g_{ck} R_{nk}] \, [S_{bn}] \qquad \text{// from (6.1.5) and (3.2.6)}$$

$$= \Sigma_{nck} \; S_{bn} R_{nk} \, g_{kc} \; \bar{g}_{ca} = [SRg\bar{g}]_{ba} = [\; 1 \; 1 \;]_{ba} = 1_{ba} = \delta_{b,a} \; . \qquad (6.2.16)$$

10. As a final step, we want to write the completeness relation (6.2.15) in matrix form. To do this, we first take a short digression.

If $\mathbf{b}$ is a column vector, and if $\mathbf{b}^{\mathbf{T}}$ is the corresponding row vector version of $\mathbf{b}$, then $(\mathbf{b}^{\mathbf{T}})_i = b_i$.

If $\mathbf{a}$ and $\mathbf{b}$ are vectors of dimension N, and if it happens that

$$(\mathbf{a})_i (\mathbf{b})_j = c_{ij} \quad , \qquad (6.2.17)$$

then one can write

$$(\mathbf{a})_i (\mathbf{b}^{\mathbf{T}})_j = c_{ij} \; . \qquad (6.2.18)$$

This equation can be written in matrix form as,

$$\mathbf{ab}^{\mathrm{T}} = c \ . \tag{6.2.19}$$

Graphically, here is an example for dimension N=2 which shows how $\mathbf{ab}^{\mathrm{T}}= c$ agrees with $(\mathbf{a})_i(\mathbf{b})_j = c_{ij}$:

$$\mathbf{ab}^{\mathrm{T}} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} (b_1 \ b_2) = \begin{pmatrix} a_1b_1 & a_1b_2 \\ a_2b_1 & a_2b_2 \end{pmatrix} = \text{a matrix} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} . \tag{6.2.20}$$

This notion can be trivially generalized to apply to a sum of vectors. Let $\mathbf{a}^{[n\}}$ and $\mathbf{b}^{[n]}$ each be a set of N vectors labeled by n. Then the following component equation

$$\Sigma_n \ (\mathbf{a}^{[n]})_i(\mathbf{b}^{[n]})_j = c_{ij} \tag{6.2.21}$$

can be written in matrix form as

$$\Sigma_n \, \mathbf{a}^{[n]} \ (\mathbf{b}^{[n]})^{\mathrm{T}} = c \ . \tag{6.2.22}$$

Now, since the completeness relation (6.2.15) has the form of (6.2.21)

$$\Sigma_n(\overline{\mathbf{B}}_n)_a \ (\mathbf{b}_n)_b = \delta_{a,b} \qquad \text{// completeness relation for the dual set } \{\mathbf{b}_n, \mathbf{B}_n \} \qquad \textit{(6.2.15)}$$

it can be expressed in the following compact matrix form from (6.2.22), where c = 1 is the identity matrix,

$$\Sigma_n \, \overline{\mathbf{B}}_n \, \mathbf{b}_n^{\mathrm{T}} = 1 \qquad \text{// completeness (matrix form) .} \tag{6.2.23}$$

In the special case that $\mathbf{b}_n = \mathbf{e}_n$ this reads

$$\Sigma_n \, \overline{\mathbf{E}}_n \, \mathbf{e}_n^{\mathrm{T}} = 1 \qquad \text{// completeness (matrix form) .} \tag{6.2.24}$$

6.3 Covariant partners for $\mathbf{e}_n$ and $\mathbf{E}_n$

A covariant partner vector is formed according to (5.8.4) which says $\overline{\mathbf{V}} = \overline{g} \, \mathbf{V}$ or $\overline{V}_i = \overline{g}_{ij}V_j$, so

$$\begin{aligned} \overline{\mathbf{e}}_n = \overline{g} \, \mathbf{e}_n \quad &\Rightarrow \quad (\overline{\mathbf{e}}_n)_i = \overline{g}_{ij}(\mathbf{e}_n)_j = \overline{g}_{ij}S_{jn} && \text{// (3.2.6)} \\ &\qquad = [\, \overline{g}S \,]_{in} = [R^{\mathrm{T}} \, \overline{g}']_{in} = R_{ji}\overline{g}'_{jn} && \text{// (5.2.8)} \end{aligned} \tag{6.3.1}$$

$$\overline{\mathbf{E}}_n = \overline{g} \, \mathbf{E}_n \quad \Rightarrow \quad (\overline{\mathbf{E}}_n)_i = \overline{g}_{ij}(\mathbf{E}_n)_j = \overline{g}_{ij}g_{jc}R_{nc} = \delta_{i,c}R_{nc} = R_{ni} \qquad \text{// (6.1.5)} \tag{6.3.2}$$

The components of the four vectors are then

$$\begin{array}{llll} (\mathbf{e}_n)_i = S_{in} & \textit{(3.2.6)} & (\mathbf{E}_n)_i = g_{ic}R_{nc} = g'_{nc}S_{ic} & \textit{(6.1.5)} \\ (\overline{\mathbf{e}}_n)_i = \overline{g}_{ij}S_{jn} = R_{ji}\overline{g}'_{jn} & \textit{(6.3.1)} & (\overline{\mathbf{E}}_n)_i = R_{ni} & \textit{(6.3.2)} \end{array} \tag{6.3.3}$$

In (3.2.7) we wrote $S = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \,....\, \mathbf{e}_N\,]$ as a representation of the $(\mathbf{e}_n)_j = S_{jn}$, showing that the tangent base vectors are the columns of matrix S. In the equation $(\overline{\mathbf{E}}_n)_i = R_{ni}$ the index order is not reversed as it is in $(\mathbf{e}_n)_j = S_{jn}$, so one concludes that the $\overline{\mathbf{E}}_n$ as are the *rows* of matrix R

$$R = \begin{bmatrix} \overline{\mathbf{E}}_1 \\ \overline{\mathbf{E}}_2 \\ \overline{\mathbf{E}}_3 \\ \ldots \\ \overline{\mathbf{E}}_N \end{bmatrix} = [\overline{\mathbf{E}}_1, \overline{\mathbf{E}}_2, \overline{\mathbf{E}}_3 \,....\, \overline{\mathbf{E}}_N\,]^T \qquad (6.3.4)$$

Now recall how vectors transform,

$$\begin{array}{llllll} \mathbf{V}' = R\,\mathbf{V} & \text{contravariant} & R_{ik}(\mathbf{x}) \equiv (\partial x'_i/\partial x_k) & & R = S^{-1} & \\ \overline{\mathbf{V}}' = S^T\,\overline{\mathbf{V}} & \text{covariant} & S_{ik}(\mathbf{x}') \equiv (\partial x_i/\partial x'_k) & = & S^T{}_{ki}(\mathbf{x}') & \quad \mathit{(2.5.1)} \end{array}$$

Here we regard both $\mathbf{V}$ and $\overline{\mathbf{V}}$ as existing in x-space, while $\mathbf{V}'$ and $\overline{\mathbf{V}}'$ are the corresponding vectors in x'-space.

As shown in the example (3.4.3), tangent base vectors $\mathbf{e}_n$ are vectors that one draws in x-space. They exist in x-space. The relation $\mathbf{E}_n \equiv g'_{ni}\,\mathbf{e}_i$. of (6.1.2) says the $\mathbf{E}_n$ vectors are just linear combinations of the $\mathbf{e}_n$ vectors, so the $\mathbf{E}_n$ vectors also exist in x-space and that is where one would draw them.

Based on this fact and the previous paragraph, we conclude the perhaps obvious fact that the vectors $\mathbf{e}_n$, $\overline{\mathbf{e}}_n$, $\mathbf{E}_n$,$\overline{\mathbf{E}}_n$ all exist in x-space.

What are the corresponding vectors in x'-space that one gets from transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$? We just apply the transformation rules quoted above to find out:

$$\mathbf{e}_n' = R\mathbf{e}_n \quad\Rightarrow\quad (\mathbf{e}_n')_i = R_{ij}\,(\mathbf{e}_n)_j = R_{ij}S_{jn} = [RS]_{in} = [1]_{in} = \delta_{i,n} \qquad (6.3.5)$$

in agreement with (3.2.1). Next,

$$\mathbf{E}_n' = R\mathbf{E}_n \quad\Rightarrow\quad (\mathbf{E}_n')_i = R_{ij}\,(\mathbf{E}_n)_j = R_{ij}g_{jc}R_{nc} = [RgR^T]_{in} = g'_{in} \qquad // \text{ (5.7.6)} \qquad (6.3.6)$$

For the covariant vectors we find

$$\overline{\mathbf{e}}_n' = S^T\overline{\mathbf{e}}_n \Rightarrow \quad (\overline{\mathbf{e}}_n')_i = S_{ji}\,(\overline{\mathbf{e}}_n)_j = S_{ji}\,\overline{g}_{jk}S_{kn} = [S^T\overline{g}S]_{in} = \overline{g}'_{in} \qquad // \text{ (5.7.6)} \qquad (6.3.7)$$

$$\overline{\mathbf{E}}_n' = S^T\overline{\mathbf{E}}_n \quad\Rightarrow\quad (\overline{\mathbf{E}}_n')_i = S_{ji}\,(\overline{\mathbf{E}}_n)_j = S_{ji}\,R_{nj} = [RS]_{ni} = \delta_{n,i} \qquad (6.3.8)$$

The conclusions are

$$\begin{array}{ll} (\mathbf{e}_n')_i = \delta_{i,n} & \qquad (\mathbf{E}_n')_i = g'_{in} \\ (\overline{\mathbf{e}}_n')_i = \overline{g}'_{in} & \qquad (\overline{\mathbf{E}}_n')_i = \delta_{n,i} \end{array} \qquad (6.3.9)$$

6.4 Summary of the basic facts about $\mathbf{e}_n$ and $\mathbf{E}_n$

For use below, we package various results above into a single block. Here we use i,j for indices, and n.m for basis vector labels,

$$(\bar{\mathbf{e}}_n)_i = \bar{g}_{ij}\,(\mathbf{e}_n)_j \qquad (\bar{\mathbf{E}}_n)_i = \bar{g}_{ij}\,(\mathbf{E}_n)_j \qquad [\ \bar{\mathbf{e}}_n = \bar{g}\,\mathbf{e}_n \qquad \bar{\mathbf{E}}_n = \bar{g}\,\mathbf{E}_n] \qquad \textit{(5.8.4)}$$

$$(\mathbf{e}'_n)_i = R_{ij}(\mathbf{e}_n)_j \qquad (\mathbf{E}'_n)_i = R_{ij}(\mathbf{E}_n)_j \qquad [\ \mathbf{e}'_n = R\,\mathbf{e}_n \qquad \mathbf{E}'_n = R\,\mathbf{E}_n\ ]$$

$$(\bar{\mathbf{e}}'_n)_i = S_{ji}(\bar{\mathbf{e}}_n)_j \qquad (\bar{\mathbf{E}}'_n)_i = S_{ji}(\bar{\mathbf{E}}_n)_j \qquad [\ \bar{\mathbf{e}}'_n = S^T\,\bar{\mathbf{e}}_n \qquad \bar{\mathbf{E}}'_n = S^T\,\bar{\mathbf{E}}_n\ ] \qquad \textit{(2.5.1)}$$

$$(\mathbf{e}_n)_i = S_{in} \qquad (\mathbf{E}_n)_i = g_{ij}R_{nj} = g'_{nj}S_{ij} \qquad (\mathbf{e}_n')_i = \delta_{i,n} \qquad (\mathbf{E}_n')_i = g'_{ni}$$

$$(\bar{\mathbf{e}}_n)_i = \bar{g}_{ij}S_{jn} = R_{ji}\bar{g}'_{jn} \qquad (\bar{\mathbf{E}}_n)_i = R_{ni} \qquad (\bar{\mathbf{e}}_n')_i = \bar{g}'_{ni} \qquad (\bar{\mathbf{E}}_n')_i = \delta_{n,i}$$

(6.3.3) *(6.3.9)*

$$\mathbf{e}_n \bullet \mathbf{e}_m = \bar{g}'_{nm} \quad \Rightarrow \quad |\mathbf{e}_n| = \sqrt{\bar{g}'_{nn}} = h'_n \quad \text{(scale factor)} \qquad \mathbf{E}_n = g'_{ni}\,\mathbf{e}_i$$

$$\mathbf{E}_n \bullet \mathbf{e}_m = \delta_{n,m} \qquad\qquad \mathbf{e}_n = \bar{g}'_{ni}\ \mathbf{E}_i$$

$$\mathbf{E}_n \bullet \mathbf{E}_m = g'_{nm} \quad \Rightarrow \quad |\mathbf{E}_n| = \sqrt{g'_{nn}} \quad . \qquad \textit{(6.2.4)}$$

$$\mathbf{e'}_n \bullet \mathbf{e'}_m = \bar{g}'_{nm} \quad \Rightarrow \quad |\mathbf{e'}_n| = \sqrt{\bar{g}'_{nn}} = h'_n \quad \text{(scale factor)} \qquad \mathbf{E'}_n = g'_{ni}\,\mathbf{e'}_i$$

$$\mathbf{E'}_n \bullet \mathbf{e'}_m = \delta_{n,m} \qquad\qquad \mathbf{e'}_n = \bar{g}'_{ni}\ \mathbf{E'}_i$$

$$\mathbf{E'}_n \bullet \mathbf{E'}_m = g'_{nm} \quad \Rightarrow \quad |\mathbf{E'}_n| = \sqrt{g'_{nn}} \quad . \qquad \textit{(6.2.7)}$$

$$(\bar{\mathbf{E}}_n)_i(\mathbf{e}_n)_j = \delta_{i,j} \qquad [\ \Sigma_n\,\bar{\mathbf{E}}_n\,\mathbf{e}_n^T = 1\ ] \qquad \textit{(6.2.16) and (6.2.24)} \qquad (6.4.1)$$

6.5 Repeat the above for the inverse transformation: definition of the U'_n

Section 3.5 introduced the *inverse* tangent base vectors called $\mathbf{u}'_n$. That is to say, the $\mathbf{u}'_n$ are the tangent base vectors of the inverse transformation $\mathbf{x} = F^{-1}(\mathbf{x}')$. Whereas the forward tangent base vectors $\mathbf{e}_n$ exist in x-space, the inverse tangent base vectors $\mathbf{u}'_n$ exist in x'-space, so that is why we have put a prime on the $\mathbf{u}'_n$.

Recall from (6.1.2) that $\mathbf{E}_n \equiv g'_{ni}\,\mathbf{e}_i$ where $\mathbf{E}_n$ were the reciprocals of $\mathbf{e}_n$. For the inverse transformation we define in exact analogy (g ↔ g') some vectors $\mathbf{U}'_n$ reciprocal to the $\mathbf{u}'_n$,

$$\mathbf{U}'_n \equiv g_{ni}\,\mathbf{u}'_i \qquad \Rightarrow \qquad \mathbf{u}'_n = \bar{g}_{ni}\,\mathbf{U}'_i \ . \qquad // \text{ since } \bar{g} = g^{-1} \qquad (6.5.1)$$

This is just another application of the generic duality discussion of Section 6.2 with generic basis vectors called $\mathbf{b}_n$ and $\mathbf{B}_n$. Everything goes along as in the previous Sections, but with the following changes to account for the fact that we are now doing the *inverse* transformation $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$:

$$g' \leftrightarrow g \qquad R \leftrightarrow S \qquad \mathbf{e}_n \rightarrow \mathbf{u}'_n \qquad \mathbf{e}'_n \rightarrow \mathbf{u}_n \qquad \mathbf{E}_n \rightarrow \mathbf{U}'_n \qquad \mathbf{E}'_n \rightarrow \mathbf{U}_n \qquad (6.5.2)$$

and similarly for barred vectors. We can then manually translate the block of data presented in (6.4.1) with the above rules to get:

$$(\bar{\mathbf{u}}'_n)_i = \bar{g}'_{ij}\,(\mathbf{u}'_n)_j \quad (\bar{\mathbf{U}}'_n)_i = \bar{g}'_{ij}\,(\mathbf{U}'_n)_j \qquad [\ \bar{\mathbf{u}}'_n = \bar{g}'\,\mathbf{u}'_n \qquad \bar{\mathbf{U}}'_n = \bar{g}'\,\mathbf{U}'_n]$$

$$(\mathbf{u}_n)_i = S_{ij}(\mathbf{u}'_n)_j \qquad (\mathbf{U}_n)_i = S_{ij}(\mathbf{U}_n')_j \qquad [\ \mathbf{u}_n = S\,\mathbf{u}'_n \qquad \mathbf{U}_n = S\,\mathbf{U}_n'\ ]$$
$$(\bar{\mathbf{u}}_n)_i = R_{ji}(\bar{\mathbf{u}}'_n)_j \qquad (\bar{\mathbf{U}}_n)_i = R_{ji}(\bar{\mathbf{U}}'_n)_j \qquad [\ \bar{\mathbf{u}}_n = R^T\,\bar{\mathbf{u}}'_n \qquad \bar{\mathbf{U}}_n = R^T\,\bar{\mathbf{U}}'_n\ ]$$

$$(\mathbf{u}'_n)_j = R_{jn} \qquad (\mathbf{U}'_n)_i = g'_{ij}S_{nj} = g_{nj}R_{ij} \qquad (\mathbf{u}_n)_i = \delta_{i,n} \qquad (\mathbf{U}_n)_i = g_{ni}$$
$$(\bar{\mathbf{u}}'_n)_i = \bar{g}'_{ij}R_{jn} = S_{ji}\bar{g}_{jn} \qquad (\bar{\mathbf{U}}'_n)_i = S_{ni} \qquad (\bar{\mathbf{u}}_n)_i = g_{ni} \qquad (\bar{\mathbf{U}}_n)_i = \delta_{n,i}$$

$$\mathbf{u}'_n \bullet \mathbf{u}'_m = \bar{g}_{nm} \qquad \Rightarrow \qquad |\mathbf{u}'_n| = \sqrt{\bar{g}_{nn}} = h_n \quad \text{(scale factor)} \qquad \mathbf{U}'_n = g_{ni}\,\mathbf{u}'_i$$
$$\mathbf{U}'_n \bullet \mathbf{u}'_m = \delta_{n,m} \qquad \qquad \mathbf{u}'_n = \bar{g}_{ni}\ \mathbf{U}'_i$$
$$\mathbf{U}'_n \bullet \mathbf{U}'_m = g_{nm} \qquad \Rightarrow \qquad |\mathbf{U}'_n| = \sqrt{g_{nn}}$$

$$\mathbf{u}_n \bullet \mathbf{u}_m = \bar{g}_{nm} \qquad \Rightarrow \qquad |\mathbf{u}_n| = \sqrt{\bar{g}_{nn}} = h_n \quad \text{(scale factor)} \qquad \mathbf{U}_n = g_{ni}\ \mathbf{u}_i$$
$$\mathbf{U}_n \bullet \mathbf{u}_n = \delta_{n,m} \qquad \qquad \mathbf{u}_n = \bar{g}_{ni}\ \mathbf{U}_i$$
$$\mathbf{U}_n \bullet \mathbf{U}_m = g_{nm} \qquad \Rightarrow \qquad |\mathbf{U}_n| = \sqrt{g_{nn}}$$

$$(\bar{\mathbf{U}}'_n)_a(\mathbf{u}'_n)_b = \delta_{a,b} \quad \text{or} \qquad \Sigma_n\,\bar{\mathbf{U}}'_n\,\mathbf{u}'_n{}^T = 1 \tag{6.5.3}$$

It is helpful to keep all these eight vector symbol names in mind (and each has a covariant partner)

(6.5.4)

	x'-space	x-space		
axis-aligned basis vectors	$\mathbf{e}'_n$	$\mathbf{u}_n$	$(\mathbf{e}'_n)_i = \delta_{n,i}$	$(\mathbf{u}_n)_i = \delta_{n,i}$
dual partners to the above	$\mathbf{E}'_n$	$\mathbf{U}_n$	$(\mathbf{E}'_n)_i = g'_{ni}$	$(\mathbf{U}_n)_i = g_{ni}$
tangent base vectors	$\mathbf{u}'_n$	$\mathbf{e}_n$	$(\mathbf{u}'_n)_i = R_{in}$	$(\mathbf{e}_n)_i = S_{in}$
reciprocal base vectors	$\mathbf{U}'_n$	$\mathbf{E}_n$	$(\mathbf{U}'_n)_i = g'_{ia}S_{na}$	$(\mathbf{E}_n)_i = g_{ia}R_{na}$
			$= g_{na}R_{ia}$	$= g'_{na}S_{ia}$

and recall that $\mathbf{X}_n \bullet \mathbf{x}_m = \delta_{n,m}$ for each of the four dual pairs (two primed, two unprimed).

6.6 Expanding vectors on different sets of basis vectors

x-space expansions on $\mathbf{u}_n$ and $\mathbf{U}_n$

Assume that $\mathbf{V}$ is some generic N-tuple $\mathbf{V} = (V_1,V_2....V_N)$. There are various ways to expand $\mathbf{V}$ onto basis vectors. One way is to expand on the axis-aligned basis vectors $\mathbf{u}_n$, which recall live in x-space,

$$\mathbf{V} = V_1\,\mathbf{u}_1 + V_2\,\mathbf{u}_2 + ... = \Sigma_n V_n \mathbf{u}_n \qquad \text{where} \qquad \mathbf{U}_n \bullet \mathbf{V} = V_n\ . \tag{6.6.1}$$

The components V_n are $\mathbf{U}_n \bullet \mathbf{V}$ because $\mathbf{U}_n \bullet \mathbf{u}_m = \delta_{n,m}$. From (5.8.4) one can write $V_n = g_{nm}\bar{V}_m$

(regarded here as a definition of the $\bar{V}_m$) so one finds that

$$\mathbf{V} = \Sigma_n V_n \mathbf{u}_n = \Sigma_n g_{nm}\bar{V}_m \, \mathbf{u}_n = \Sigma_n \bar{V}_m \, g_{mn} \, \mathbf{u}_n = \Sigma_n \bar{V}_m \, \mathbf{U}_m \qquad (6.6.2)$$

and thus another expansion for **V** is this

$$\mathbf{V} = \bar{V}_1 \, \mathbf{U}_1 + \bar{V}_2 \, \mathbf{U}_2 + .. = \Sigma_n \bar{V}_n \mathbf{U}_n \quad \text{where} \quad \mathbf{u}_n \bullet \mathbf{V} = \bar{V}_n \ . \qquad (6.6.3)$$

Comments:

1. If **V** *is not* a contravariant vector, one can still define $\bar{V}_n = \bar{g}_{nm}V_m$, but $\bar{V}_n$ won't be a covariant vector. A familiar example is that x_n is never a contravariant vector if F is non-linear, but we can still talk about the components $\bar{x}_n \equiv \bar{g}_{nm}x_m$. In Standard Notation, $x_n \rightarrow x^n$ and $\bar{x}_n \rightarrow x_n$ and we do not hesitate to use these two objects even though they are not tensorial vectors.

2. If **V** *is* a contravariant vector, the expansion above $\mathbf{V} = \Sigma_n V_n \mathbf{u}_n$ displays the contravariant components of **V**. The second expansion $\mathbf{V} = \Sigma_n \bar{V}_m \, \mathbf{U}_m$ is still an expansion for contravariant vector **V**, but it displays the components of the covariant vector $\bar{\mathbf{V}}$ which is the "partner" to **V** by $\bar{V}_n = \bar{g}_{nm}V_m$. It would be incorrect to write this second expansion as $\bar{\mathbf{V}} = \Sigma_n \bar{V}_m \, \mathbf{U}_m$ since that would say $\bar{g}\mathbf{V} = \Sigma_n \bar{V}_m \, \mathbf{U}_m$ which is just not true. We comment later on how this situation changes a bit in the Standard Notation.

x-space expansions on $\mathbf{e}_n$ and $\mathbf{E}_n$

Another possibility is to expand **V** on the tangent basis vectors $\mathbf{e}_n$, and we denote the components just momentarily as α_n,

$$\mathbf{V} = \alpha_1 \, \mathbf{e}_1 + \alpha_2 \, \mathbf{e}_2 + ... = \Sigma_n \, \alpha_n \, \mathbf{e}_n \quad . \qquad (6.6.4)$$

Using $\mathbf{e}_n \bullet \mathbf{E}_m = \delta_{n,m}$ one finds that

$$\alpha_n = \mathbf{E}_n \bullet \mathbf{V} = (\mathbf{E}_n)_k \, V_k = R_{nk} \, V_k = V'_n \qquad // \ \bar{g} = 1 \ \text{ so } \mathbf{A} \bullet \mathbf{B} = \bar{g}_{ab} A_a B_b = A_k B_k \qquad (6.6.5)$$

Therefore, the expansion is

$$\mathbf{V} = V'_1 \mathbf{e}_1 + V'_2 \mathbf{e}_2 + ... = \Sigma_n \, V'_n \, \mathbf{e}_n \qquad \text{where} \quad \mathbf{E}_n \bullet \mathbf{V} = V'_n \ . \qquad (6.6.6)$$

If it happens that the N-tuple $\mathbf{V} = (V_1, V_2 .... V_N)$ transforms as a contravariant vector, then V_n are the contravariant components of that vector, and V'_n are the contravariant components of **V**' in x'-space. On the other hand, if **V** is not a tensorial vector, so V_n are not components of a contravariant vector, we can still define $V'_n \equiv R_{nk} \, V_k$, but then the V'_n are not the contravariant components of **V**'.

Writing $V'_n = g'_{nm}\bar{V}'_m$ the above expansion can be expressed as

$$\mathbf{V} = \Sigma_n V'_n \mathbf{e}_n = \Sigma_{n,m}\ g'_{nm}\overline{V}'_m \mathbf{e}_n\ = \Sigma_m \overline{V}'_m \Sigma_n\ g'_{mn} \mathbf{e}_n\ = \Sigma_m \overline{V}'_m\ \mathbf{E}_m \tag{6.6.7}$$

so

$$\mathbf{V} = \ \overline{V}'_1\mathbf{E}_1 + \overline{V}'_2\mathbf{E}_2 + ... \quad = \ \Sigma_n \overline{V}'_n \mathbf{E}_n \qquad \text{where} \quad \mathbf{e}_n \bullet \mathbf{V}\ =\ \overline{V}'_n \tag{6.6.8}$$

<u>Summary of x-space expansions:</u>

$$\begin{array}{llll}
\mathbf{V} = V_1\, \mathbf{u}_1 + V_2\, \mathbf{u}_2 + ... & = \ \Sigma_n V_n\ \mathbf{u}_n & \text{where} \ \ \mathbf{U}_n \bullet \mathbf{V}\ =\ V_n & \mathbf{U}_n = g_{ni}\, \mathbf{u}_i \\
\mathbf{V} = \overline{V}_1\, \mathbf{U}_1 + \overline{V}_2\, \mathbf{U}_2 + ... & = \ \Sigma_n \overline{V}_n\ \mathbf{U}_n & \text{where} \ \ \mathbf{u}_n \bullet \mathbf{V}\ =\ \overline{V}_n & \\
\mathbf{V} = V'_1\ \mathbf{e}_1 + V'_2\ \mathbf{e}_2 + ... & = \ \Sigma_n\, V'_n\, \mathbf{e}_n & \text{where} \ \ \mathbf{E}_n \bullet \mathbf{V}\ =\ V'_n & \mathbf{E}_n = g'_{ni}\, \mathbf{e}_i \\
\mathbf{V} = \ \overline{V}'_1\ \mathbf{E}_1 + \overline{V}'_2\ \mathbf{E}_2 + ... & = \ \Sigma_n\, \overline{V}'_n\, \mathbf{E}_n & \text{where} \ \ \mathbf{e}_n \bullet \mathbf{V}\ =\ \overline{V}'_n &
\end{array} \tag{6.6.9}$$

<u>Expanding on unit vectors.</u> The covariant lengths of the different basis vectors are given by

$$\begin{array}{ll}
|\mathbf{e}_n|\ = |\mathbf{e}'_n|\ = \sqrt{g'_{nn}} \qquad & |\mathbf{u}_n|\ = |\mathbf{u}'_n| = \sqrt{g_{nn}} \\
|\mathbf{E}_n| = |\mathbf{E}'_n| = \sqrt{g'_{nn}} \qquad & |\mathbf{U}_n| = |\mathbf{U}'_n| = \sqrt{g_{nn}}\ .
\end{array} \tag{6.6.10}$$

Using these lengths, one can define unit vector versions of all the basis vectors and then rewrite the above expansions as expansions on the unit vectors with lower case coefficients. For example, using

$$\hat{\mathbf{e}}_n \equiv \mathbf{e}_n/\sqrt{g'_{nn}} \tag{6.6.11}$$

the third expansion above becomes (script font for unit-vector components)

$$\mathbf{V} = \mathcal{V}'_1\hat{\mathbf{e}}_1 + \mathcal{V}'_2\hat{\mathbf{e}}_2 + ... \ = \ \Sigma_n\, \mathcal{V}'_n\, \hat{\mathbf{e}}_n \quad \text{where} \qquad \sqrt{g'_{nn}}\ \mathbf{E}_n \bullet \mathbf{V}\ =\ \mathcal{V}'_n = \sqrt{g'_{nn}}\, V'_n\ . \tag{6.6.12}$$

An example of a case where this last expansion would be useful is in the use of spherical curvilinear coordinates, where for example $\hat{\mathbf{e}}_1 = \hat{\mathbf{r}}$.

The N-tuple ($\mathcal{V}'_1$, $\mathcal{V}'_2$ ... $\mathcal{V}'_N$), although related to contravariant vector **V** ($V'_n = R_{nk}\ V_k$), is not itself a contravariant vector since it does not obey the rule $\mathcal{V}'_n = R_{nk}\ \mathcal{V}_k$. In fact

$$V'_n = R_{nk}\, V_k \ \Rightarrow\ (1/\sqrt{g'_{nn}}\,)\ \mathcal{V}'_n\ = R_{nk}\ (1/\sqrt{g_{kk}}\,)\ \mathcal{V}_k\ \Rightarrow\ \mathcal{V}'_n\ = R_{nk}\ (\sqrt{g'_{nn}}/\sqrt{g_{kk}}\,)\mathcal{V}_k \tag{6.6.13}$$

<u>x'-space expansions</u>

Having done x-space expansions, we turn now to x'-space expansions. These can be obtained from the x-space expansions by this set of rules,

$$g' \leftrightarrow g \quad R \leftrightarrow S \quad \mathbf{u}'_n \rightarrow \mathbf{e}_n \quad \mathbf{u}_n \rightarrow \mathbf{e}'_n \quad \mathbf{U}'_n \rightarrow \mathbf{E}_n \quad \mathbf{U}_n \rightarrow \mathbf{E}'_n \quad V'_n \leftrightarrow V_n \quad \overline{V}'_n \leftrightarrow \overline{V}_n \tag{6.6.14}$$

and here then are the x'-space expansions:

$$\begin{array}{llll}
\mathbf{V'} = V'_1\, \mathbf{e'}_1 + V'_2\, \mathbf{e'}_2 + ... & = \Sigma_n V'_n\, \mathbf{e'}_n & \text{where } \mathbf{E'}_n \bullet \mathbf{V'} = V'_n & \mathbf{E'}_n = g'_{ni}\, \mathbf{e'}_i \\
\mathbf{V'} = \overline{V}'_1\, \mathbf{E'}_1 + \overline{V}'_2\, \mathbf{E'}_2 + ... & = \Sigma_n \overline{V}'_n\, \mathbf{E'}_n & \text{where } \mathbf{e'}_n \bullet \mathbf{V'} = \overline{V}'_n & \\
\mathbf{V'} = V_1\, \mathbf{u'}_1 + V_2\, \mathbf{u'}_2 + ... & = \Sigma_n V_n\, \mathbf{u'}_n & \text{where } \mathbf{U'}_n \bullet \mathbf{V'} = V_n & \mathbf{U'}_n = g_{ni}\, \mathbf{u'}_i \\
\mathbf{V'} = \overline{V}_1\, \mathbf{U'}_1 + \overline{V}_2\, \mathbf{U'}_2 + ... & = \Sigma_n \overline{V}_n\, \mathbf{U'}_n & \text{where } \mathbf{u'}_n \bullet \mathbf{V'} = \overline{V}_n & (6.6.15)
\end{array}$$

6.7 Another way to write the $\mathbf{E}_n$

The reciprocal base vectors were defined above as linear combinations of the tangent base vectors, all in the general Picture A context,

$$\mathbf{E}_k \equiv g'_{ki}\, \mathbf{e}_i \ . \qquad (6.7.1)$$

It is rather remarkable that there is another way to write $\mathbf{E}_k$ in terms of the $\mathbf{e}_i$ that looks completely different. In this other way, it turns out that $\mathbf{E}_k$ is expressed in terms of all the $\mathbf{e}_i$ *except* $\mathbf{e}_k$ and is given by (only valid in Picture B where g=1)

$$\mathbf{E}_k = \det(R)\, (-1)^{k-1}\, \mathbf{e}_1 \text{ x } \mathbf{e}_2 \text{ x } ......\text{x } \mathbf{e}_N \qquad // \ \mathbf{e}_k \text{ missing} \qquad k = 1,2,3...N \qquad (6.7.2)$$

This is a generalized cross product (see Section A.4) of N-1 vectors, since $\mathbf{e}_k$ is missing, so there are N-2 "crosses". The above multi-cross-product equation is a shorthand for

$$(\mathbf{E}_k)_\alpha \equiv \det(R)\, (-1)^{k-1} \varepsilon_{\alpha abc...x}\, (\mathbf{e}_1)_a (\mathbf{e}_2)_b \ ...... \ (\mathbf{e}_N)_x \qquad // \ \kappa \text{ and } (\mathbf{e}_\kappa)_K \text{ are missing} \qquad (6.7.3)$$

Here ε is the totally antisymmetric tensor with N indices. If κ is the k^{th} letter of the alphabet ($k = 2 \Rightarrow \kappa = b$), then κ is missing from the list of summation indices of ε, and the factor $(\mathbf{e}_k)_\kappa$ is missing from the product of factors, so there are then N-1 factors.

This cross product expression for $\mathbf{E}_n$ is derived in **Appendix A**.

This is all fairly obscure sounding, but can be brought down to earth by writing things out for N = 3, where the formula reduces to this cyclic set of equations,

$$\begin{aligned}
\mathbf{E}_1 &= \det(R)\, \mathbf{e}_2 \text{ x } \mathbf{e}_3 \\
\mathbf{E}_2 &= \det(R)\, \mathbf{e}_3 \text{ x } \mathbf{e}_1 \\
\mathbf{E}_3 &= \det(R)\, \mathbf{e}_1 \text{ x } \mathbf{e}_2 \ . \qquad (6.7.4)
\end{aligned}$$

These equations can be verified in a simple manner. To show an equation is true, if suffices to show that the projections of both sides on the three $\mathbf{e}_n$ are the same, since the $\mathbf{e}_n$ form a complete basis as noted earlier. For the first equation one needs then to show that

$$\mathbf{E}_1 \bullet \mathbf{e}_n = \det(R)\, \mathbf{e}_2 \text{ x } \mathbf{e}_3 \bullet \mathbf{e}_n \qquad \text{for } n = 1,2,3 \ . \qquad (6.7.5)$$

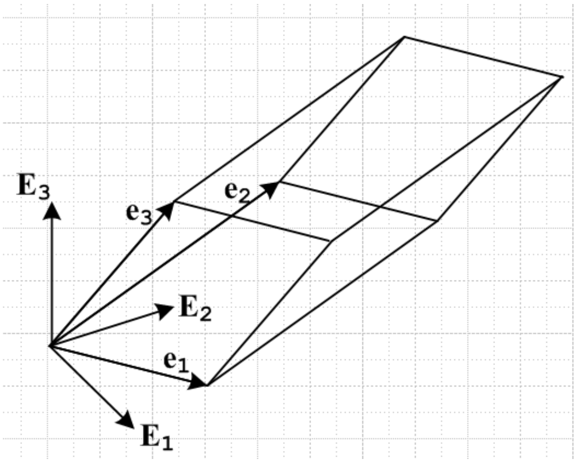
If n=2 or n=3, both sides vanish, according to $\mathbf{e}_n \bullet \mathbf{E}_m = \delta_{n,m}$ on the left, and according to geometry on the right, which leaves just the case n = 1. In this case the LHS = 1, so one has to show that $\mathbf{e}_2 \text{ x } \mathbf{e}_3 \bullet \mathbf{e}_1 = 1/\det(R) = \det(S)$. But

$\mathbf{e}_1 \bullet \mathbf{e}_2 \times \mathbf{e}_3 = (\mathbf{e}_1)_i (\mathbf{e}_2 \times \mathbf{e}_3)_i = (\mathbf{e}_1)_i \varepsilon_{ijk} (\mathbf{e}_2)_j(\mathbf{e}_3)_k = \varepsilon_{ijk} (\mathbf{e}_1)_i(\mathbf{e}_2)_j(\mathbf{e}_3)_k$

$= \varepsilon_{ijk} S_{i1}S_{j2}S_{k3} = \det(S)$ QED // (3.2.5) says $(\mathbf{e}_n)_i = S_{in}$ (6.7.6)

The other two equations of the set can be verified in the same manner.

Here is a picture (N=3) drawn in x-space for a non-orthogonal coordinate system. The vectors shown here form a distorted right handed coordinate system which has det(R) > 0.



(6.7.7)

The reader is invited to exercise his or her right hand to confirm the directions of the arrows,

$\mathbf{E}_1 = \det(R)\, \mathbf{e}_2 \times \mathbf{e}_3 \qquad \mathbf{E}_2 = \det(R)\, \mathbf{e}_3 \times \mathbf{e}_1 \qquad \mathbf{E}_3 = \det(R)\, \mathbf{e}_1 \times \mathbf{e}_2$. (6.7.8)

6.8 Comparison of $\bar{\mathbf{e}}_n$ and $\mathbf{E}_n$

Recall from (6.3.1) that $\bar{\mathbf{e}}_n$ is the covariant partner of $\mathbf{e}_n$ where $\bar{\mathbf{e}}_n = \bar{g}\, \mathbf{e}_n$.
Recall from (6.1.2) that $\mathbf{E}_n$ is the reciprocal of $\mathbf{e}_n$ where $\mathbf{E}_n \equiv g'_{ni}\, \mathbf{e}_i$.

The comparison of these two defining equations is interesting,

$(\bar{\mathbf{e}}_n)_i \equiv \bar{g}_{ik}(\mathbf{e}_n)_k$ // matrix acts on vector index (6.8.1)

$(\mathbf{E}_n)_i \equiv g'_{nk}\, (\mathbf{e}_k)_i$. // matrix acts on $\mathbf{e}_k$ label (6.8.2)

If x-space is Cartesian, then $\bar{\mathbf{e}}_n = \mathbf{e}_n$ as usual, but of course $\mathbf{E}_n \neq \mathbf{e}_n$ in this case since normally $g' \neq 1$. We mention this to head off a possible confusion when the Standard Notation is introduced in the next Chapter and the above two equations become (see (7.4.1) and (7.13.1)),

$(\mathbf{e}_n)_i \equiv g_{ik}(\mathbf{e}_n)^k$ // Standard Notation, g lowers an index (6.8.3)

$(\mathbf{e}^n)^i \equiv g'^{nk}\, (\mathbf{e}_k)^i$. // Standard Notation, k is a label on $\mathbf{e}_k$, not an index (6.8.4)

The mapping to standard notation does *not* include $\overline{\mathbf{e}}_{\mathbf{n}} \to \mathbf{e}^{\mathbf{n}}$, for example. One fact about the standard notation is that, unlike the developmental notation, one cannot look at a vector in bold like $\mathbf{e}_{\mathbf{n}}$ and determine whether it is contravariant or covariant. Only when the index is displayed can one tell. One can think of $\mathbf{e}_{\mathbf{n}}$ as representing both its contravariant self and its covariant partner ($\mathbf{e}_{\mathbf{n}}$ is a tensorial vector).

6.9 Handedness of coordinate systems: the $\mathbf{e}_{\mathbf{n}}$, the sign of det(S), and Parity

1. Handedness of a Coordinate System. Let $\mathbf{b}_{\mathbf{n}}$ be a complete set of basis vectors in an N dimensional vector space, where the $\mathbf{b}_{\mathbf{n}}$ are not necessarily of unit length, and are not necessarily orthogonal. Consider this quantity

$$B \equiv \det(\mathbf{b}_1, \mathbf{b}_2.....\mathbf{b}_n) \;=\; \varepsilon_{abc..x}\,(\mathbf{b}_1)_a\,(\mathbf{b}_2)_b....\,(\mathbf{b}_N)_x \;=\; \mathbf{b}_1\bullet[\mathbf{b}_2 \text{ x } \mathbf{b}_3......\text{x } \mathbf{b}_N] \qquad (6.9.1)$$

where the generalized cross product is discussed in Section A.4. This basis $\mathbf{b}_{\mathbf{n}}$ defines a "coordinate system" in that we can expand a position vector as follows

$$\mathbf{x} = \Sigma_n\, x^{(b)}{}_n\, \mathbf{b}_n \qquad (6.9.2)$$

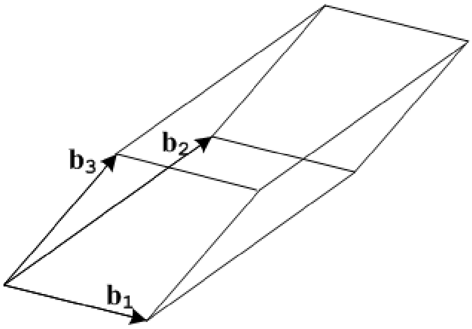
where the $x^{(b)}{}_n$ are the "coordinates" of point $\mathbf{x}$ in this coordinate system. We make the following definition:

system $\mathbf{b}_{\mathbf{n}}$ is a "right handed coordinate system" iff $B > 0$
system $\mathbf{b}_{\mathbf{n}}$ is a "left handed coordinate system" iff $B < 0$. (6.9.3)

One of course wants to show that for N = 3 this definition corresponds to one's intuition about right and left handed systems. For N= 3 ,

$$B \equiv \det(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3) \;=\; \varepsilon_{abc}\,(\mathbf{b}_1)_a\,(\mathbf{b}_2)_b(\mathbf{b}_3)_c \;=\; \mathbf{b}_1\bullet[\mathbf{b}_2 \text{ x } \mathbf{b}_3] \qquad (6.9.4)$$

Suppose the $\mathbf{b}_{\mathbf{n}}$ are arranged as shown in this picture, where the visual intention is that the corner nearest the label $\mathbf{b}_1$ is closest to the viewer.



(6.9.5)

With one's high-school-trained right hand, one can see that $\mathbf{b}_2$ x $\mathbf{b}_3$ points in the general direction of $\mathbf{b}_1$ (certainly $\mathbf{b}_2$ x $\mathbf{b}_3$ lies somewhere in the half space of the $\mathbf{b}_2$, $\mathbf{b}_3$ face plane which contains $\mathbf{b}_1$), and so the quantity B = $\mathbf{b_1} \bullet [\mathbf{b_2} \times \mathbf{b_3}] > 0$. So this is an example of a right-handed coordinate system. The figure shown is a skewed 3-piped which can be regarded as a distortion of an orthogonal 3-piped for which the $\mathbf{b_n}$ would span an orthogonal coordinate system in which one would have $\hat{\mathbf{b}}_1 = \hat{\mathbf{b}}_2 \times \hat{\mathbf{b}}_3$.

2. The x'-space $\mathbf{e'_n}$ coordinate system is always right handed. In our standard picture of x-space and x'-space, the coordinate system in x'-space is spanned by a set of basis vectors $\mathbf{e'_n}$ where $(\mathbf{e'_n})_i = \delta_{n,i}$, as discussed in Chapter 3 (a). This system is "right handed" because

$$B = \varepsilon_{abc..x} (\mathbf{e'_1})_a (\mathbf{e'_2})_b .... (\mathbf{e'_N}) = \varepsilon_{abc..x} \delta_{1a}\delta_{2b}....\delta_{xN} = \varepsilon_{123...N} = +1 > 0 \qquad (6.9.6)$$

where we use the standard normalization of the ε tensor as shown. Notice that this conclusion is independent of the metric tensor g' in x-space.

3. The x-space $\mathbf{u_n}$ coordinate system is always right handed. The basis $\mathbf{u_n}$ where $(\mathbf{u_n})_i = \delta_{n,i}$ in x-space is right handed for the same reason as shown in the above paragraph, and for any g. When g=1 in x-space, the $\mathbf{u_n}$ form the usual Cartesian right-handed orthonormal basis in x-space. See Section 3.4 concerning the meaning of "unit vector".

4. The x-space $\mathbf{e_n}$ coordinate system handedness is determined by the sign of det(S). Our x-space coordinate system of great interest is that spanned by the $\mathbf{e_n}$ basis vectors, where $(\mathbf{e_n})_i = S_{in}$ as in (3.2.5). One has

$$B = \varepsilon_{abc..x} (\mathbf{e_1})_a (\mathbf{e_2})_b .... (\mathbf{e_N}) = \varepsilon_{abc..x} S_{a1} S_{b2} .... S_{xN} = \det(S) \ . \qquad (6.9.7)$$

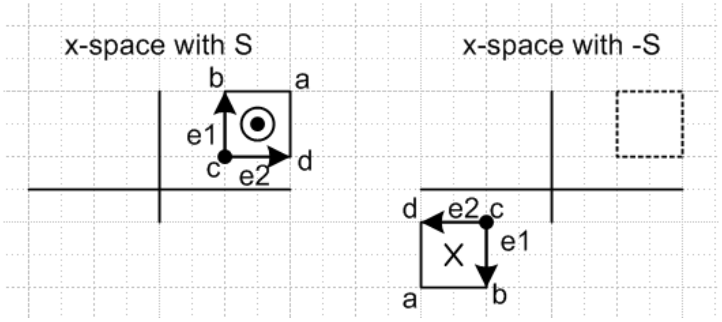
Therefore, using σ ≡ sign(detS) and the Jacobian J = det(S) of (5.12.6), one has

$$\begin{array}{llll} \text{system } \mathbf{e_n} \text{ is a "right handed coordinate system"} & \text{iff} & \det(S) = J > 0 & \text{or } \ \sigma = +1 \\ \text{system } \mathbf{e_n} \text{ is a "left handed coordinate system"} & \text{iff} & \det(S) = J < 0 & \text{or } \ \sigma = -1 \ . \end{array} \qquad (6.9.8)$$

5. The Parity Transformation. The identity transformation F = 1 results in matrix $S_I = I$ with $\det S_I = +1$. In this case the $\mathbf{e_n}$ form a right-handed coordinate system, and in fact $\mathbf{e_n} = \mathbf{u_n}$. The parity transformation F = -1, on the other hand, results in $S_P = -I$ with $\det(S_P) = (-1)^N$ and $\mathbf{e_n} = -\mathbf{u_n}$. When N is odd, the parity transformation converts the right-handed $\mathbf{u_n}$ system to a left-handed $\mathbf{e_n}$ system. If S is some matrix which does not change handedness, meaning detS > 0, then $S' = SS_P$ *does* change handedness for odd N, since in this case $\det S' = \det S \det S_P = (-1)^N \det S$. So given some S' that changes handedness for N=odd, one can regard it as "containing the parity transformation" which, if removed, would result in no handedness change. When N is even, handedness stays the same under F = -1.

6. N=3 Parity Inversion Example. Since $\mathbf{x'} = -\mathbf{x}$ under the parity transform F = -1, parity is a reflection of all position vectors through the origin. Objects sitting in x-space, such as N-pipeds, whether or not the origin lies inside the object, are "turned inside out" by the parity transformation, but the inside of the object still maps to the inside of the parity-transformed object under this transformation.

Consider this crude picture which shows on the left a right-handed 3-piped in x-space where $\mathbf{e}_1$ and $\mathbf{e}_2$ happen to be perpendicular, and the back part of the 3-piped is not drawn. This 3-piped is associated with some transformation S [$(\mathbf{e}_n)_i = S_{in}$] with detS > 0.



(6.9.9)

Now consider S' = $SS_P = S_PS$ = -S. For this S', the 3-piped appears as shown on the right. The two pipeds here are related by a parity transformation, all points inverting through the origin. On the right, the volume of the 3-piped lies *toward* the viewer from the plane shown. The circled dot on the left represents the out-facing normal vector of the 3-piped face area which is facing the viewer, and this normal is in the direction – $\mathbf{e}_1 \mathrm{x} \mathbf{e}_2$. This is called a "near face" in Appendix B since it touches the tails of the $\mathbf{e}_n$. After the parity transformation, this same face has become the *back* face on the inverted 3-piped shown on the right, with out-facing normal indicated by the X. The direction of this normal is + $\mathbf{e}_1 \mathrm{x} \mathbf{e}_2$. In general, an out-facing "near face" area points in the $-\mathbf{E}_n$ direction, and Appendix A shows that $\mathbf{E}_3 = \mathbf{e}_1 \mathrm{~x~} \mathbf{e}_2 / \det(S)$. On the left we have $\mathbf{E}_3 = \mathbf{e}_1 \mathrm{~x~} \mathbf{e}_2 / |\det(S)|$ so the face there just mentioned points in the $-\mathbf{E}_3 = -\ \mathbf{e}_1 \mathrm{x} \mathbf{e}_2$ direction. On the right we have $\mathbf{E}_3 = \mathbf{e}_1 \mathrm{~x~} \mathbf{e}_2 / \det(S') = -\ \mathbf{e}_1 \mathrm{~x~} \mathbf{e}_2 /|\det(S)|$, so the face there points in the $-\mathbf{E}_3 = +\mathbf{e}_1 \mathrm{x} \mathbf{e}_2$ direction.

7. The sign of det(S) in the curvilinear coordinates application. For a given ordering of the **x'** coordinates, det(S) will have a certain sign. By changing the x'_i ordering to any odd permutation of the original ordering (for example, swap two coordinates), det(S) will negate because two columns of $S_{ik}(\mathbf{x}') \equiv (\partial x_i/\partial x'_k)$ will be swapped. In the curvilinear coordinates application it is therefore always possible to select the ordering of the x'_i coordinates to cause det(S) to be positive. One always starts with a right-handed Cartesian system $\hat{\mathbf{n}}$ for x-space, and det(S)>0 then guarantees that the $\mathbf{e}_n$ will form a right-handed system there as well. Since the underlying transformation F is assumed invertible, one cannot have det(S)=0 anywhere in the domain **x** (or range **x'**) of **x'** = **F**(**x**), and therefore det(S) cannot change sign anywhere in the space of interest.

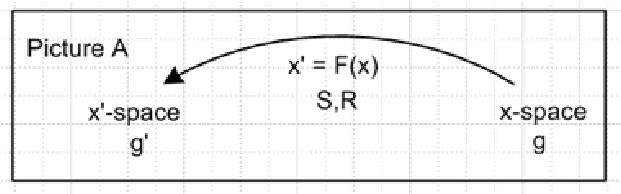
For graphical reasons, we have selected coordinates in the "wrong order" in both the polar coordinates examples (called Example 1) and in the elliptic polar system studied in Appendix C, which is why detS < 0 for both these systems.

7. Translation to the Standard Notation

In this Chapter we discuss the "translation" from our developmental notation (all lower indices; overbars for covariant objects) to the Standard Notation used in tensor analysis.

The developmental notation has served well in the discussion of scalars and vectors, tensors of rank-0 and rank-1. For pure (unmixed) tensors of rank-2 it does especially well, allowing the use of matrix algebra to leverage the use of familiar matrix theorems such as det(ABC) = det(A)det(B)det(C) and $A^{-1} = \text{cof}(A^T)/\det(A)$. The transformation of the contravariant metric tensor is cleanly expressed as $g' = R\, g\, R^T$, and so on. The notation in fact works fine for unmixed tensors of any rank, but runs into big trouble with "mixed" tensors as shown in the next Sections.

We continue to use Picture A:



7.1 Outer Products

It is possible to form larger tensors from smaller ones using the "outer product" method. For example, consider,

$$T_{ab} \equiv U_a V_b \tag{7.1.1}$$

where U and V are assumed to be contravariant vectors. One then has

$$T'_{ab} = U'_a V'_b = (R_{aa'}U_{a'})\,(R_{bb'}V_{b'}) = R_{aa'}\, R_{bb'}\, U_{a'}V_{b'} = R_{aa'}\, R_{bb'}\, T_{a'b'} \tag{7.1.2}$$

so in this way a contravariant rank-2 tensor like M in (5.6.3) has been successfully constructed from two contravariant vectors. Similarly,

$$\bar{T}_{ab} \equiv \bar{U}_a \bar{V}_b \qquad \Rightarrow \qquad \bar{T}'_{ab} = S^T_{aa'}\, S^T_{bb'}\, \bar{T}_{a'b'} \tag{7.1.3}$$

so the outer product of two covariant vectors transforms as a covariant rank-2 tensor, again as in (5.6.3).

More generally, one can combine any number of contravariant tensors in an outer product fashion to obtain a new tensor. Here are a few examples:

$$T_{abc} = U_a A_{ab}$$

$$T_{abcd} \equiv A_{ab} B_{cd}$$

$$T_{abcde} \equiv A_{ab} B_{cd} U_e \tag{7.1.4}$$

7.2 Mixed Tensors and Notation Issues

Suppose we take the "outer product" of a contravariant vector with a covariant vector,

$$[\ ...\ ]_{ab} \ \equiv\ U_a\bar{V}_b \tag{7.2.1}$$

where we are not sure what to call this thing, so we just call it [...]. Here is how this new object transforms (always: with respect to the underlying transformation **x**' = **F**(**x**))

$$[\ ...\ ]'_{ab} \ =\ U'_a\bar{V}'_b \ = (R_{aa'}U_{a'})\,(S^T{}_{bb'}\bar{V}_{b'}) \ = R_{aa'}\ S^T{}_{bb'}\ U_{a'}\bar{V}_{b'} \ = R_{aa'}\ S^T{}_{bb'}\ [...]_{ab}\ . \tag{7.2.2}$$

This object transforms as a contravariant vector on the first index (ignoring the second), and as a covariant vector on the second index (ignoring the first). This is an example of a "mixed" rank-2 tensor. Extending this outer product idea, one can make elaborate tensor objects with an arbitrary mixture of "contravariant indices" and "covariant indices". For example

$$[.....]_{abcd} \ = U_a\bar{V}_b\,X_c\bar{Y}_d \ \ . \tag{7.2.3}$$

To write down the transformation rule for such an object, one must know which indices are contravariant and which are covariant. It is totally clear how the object transforms, looking at the right side of (7.2.3), but somehow this information has to be embedded in the notation $[.....]_{abcd}$ because once this object is defined, the right hand side might not be immediately available for inspection. Worse, there may *be* no right hand side for a mixed tensor, because not all mixed tensors are outer products of vectors (they just transform as if they were).

Just as we can use the idea $\bar{\mathbf{V}} \equiv \bar{g}\ \mathbf{V}$ to convert a contravariant vector to its covariant partner, we can similarly use $\bar{g}$ to convert the 1st or 3rd index on $[.....]_{abcd}$ from contravariant to covariant. We could apply *two* $\bar{g}$'s with the proper linkage of indices to convert them both at once.

So given the ability of $\bar{g}$ to change any index one way, and g to change it the other way, one can think of the 4-index object $[.....]_{abcd}$ as a family of 16 different 4-index objects, each corresponding to a certain choice for the indices being one type or the other. We know how to interconvert between these 16 objects just applying g or $\bar{g}$ factors.

So how does one annotate *which* of the 16 objects [.....] one is staring at for some choice of index types? Here is a somewhat facetious possibility, the Morse Code method

$$\overset{\;\;-}{W}_{ab} \ \equiv\ U_a\bar{V}_b$$

$$\overset{-\;-}{W}_{abcd} \ = U_a\bar{V}_b\,X_c\bar{Y}_d \ \ . \tag{7.2.4}$$

Instead of having a bar over the entire object, in the first case the bar it is placed just over the right side of the W to indicate that b is a covariant index, while no bar means the first index is contravariant. The second example shows how horrible such a notation would be.

We really want to put some kind of notation *on the individual indices*, not on the object! Here is a notation that is slightly better than the Morse code option, though similar to it,

$$W_{a\bar{b}c\bar{d}} \ = U_aV_{\bar{b}}\,X_cY_{\bar{d}} \tag{7.2.5}$$

Here overbars on covariant indices distinguish them. Now one can dispense with the overbars on covariant vectors as well, putting the overbar on the index, for example $\overline{V}_{a} = \overline{g}_{ab}V_{b} \rightarrow V_{\bar{a}} = g_{\bar{a}\bar{b}}V_{b}$.

There are several problems with this scheme. One is that in the spinor application of tensor analysis used in special relativity, dots are placed on certain indices, such as in (5.14.8), and these would conflict with the proposed overbars. A more substantial reason is that this last notation is hard to type (or typeset, as one used to say), it looks cluttered with all the overbars, and the subscripts are already hard to read without extra decorations since they are in a smaller font than the main text.

7.3 The up/down bell goes off

This is where a bell went off somewhere, perhaps in the mind of Gregorio Ricci in the 1880-1900 time frame (1900 snippet quoted in (7.10.8) below). Someone *might* have said: suppose, instead of using overbars on indices or some other decoration, we distinguish covariant indices by making them be *superscripts* instead of subscripts. Superscripts are as easy to type as subscripts, and the result is fairly easy to read and totally unambiguous. We would then have for our ongoing example of (7.2.4 and 5),

$$W_{a}{}^{b}{}_{c}{}^{d} = U_{a}V^{b}X_{c}Y^{d} \qquad \text{// a path not taken}$$

This is *almost* what happened, but the up/down decision went the other way and we now have:

superscripts = contravariant = up
subscript = covariant = down (7.3.1)

and then we get this translation for (7.2.4),

$$W_{a\bar{b}c\bar{d}} = U_{a}V_{\bar{b}}\,X_{c}Y_{\bar{d}} \qquad \rightarrow \qquad W^{a}{}_{b}{}^{c}{}_{d} = U^{a}V_{b}X^{c}Y_{d} \qquad \text{// the path taken} \qquad (7.3.2)$$

and this has become **The Standard Notation**. Perhaps the reason for this choice was that the covariant gradient ∂_{n} object appeared more commonly in equations than idealized objects such as d**x**, and ∂_{n} already used a lower index.

A downside of this particular up/down decision is that every student has be be confused by the fact that his or her familiar position, velocity and momentum vectors that always had subscripts suddenly have superscripts in the Standard Notation. The silver lining is that this shocking change alerts the student to the fact that whatever subject is being studied is going to have two kinds of vectors.

Despite appearances, it is not completely obvious how one should translate the whole world as presented in the previous six Chapters into this new notation. There are quite a few subtle details that will be discussed in the following Sections.

7.4 Some Preliminary Translations; raising and lowering tensor indices with g

In the rest of this entire Chapter, anything to the left of a → arrow is in "developmental notation", while anything to the right of → is in "Standard Notation".

1. Basic translations.

So we start translating some of the results above:

$s \rightarrow s$ // a scalar

$V_a \rightarrow V^a$ // a contravariant rank-1 tensor (vector)
$\bar{V}_a \rightarrow V_a$ // a covariant rank-1 tensor (vector)

$M_{ab} \rightarrow M^{ab}$ // a contravariant rank-2 tensor
$\bar{M}_{ab} \rightarrow M_{ab}$ // a covariant rank-2 tensor

$g_{ab} \rightarrow g^{ab}$ // the contravariant rank-2 metric tensor g (and similarly for g')
$\bar{g}_{ab} \rightarrow g_{ab}$ // the covariant rank-2 metric tensor $\bar{g}$
g is inverse of $\bar{g}$ → g^{ab} is inverse of g_{ab} . (7.4.1)

As noted earlier, one "feature" of the Standard Notation is that it is no longer sufficient to specify an object by a single letter. One has to somehow indicate the index nature by showing index positions. Thus, "g" stands for all four metric tensors g_{ab} , g^{ab}, $g^a{}_b$ and $g_a{}^b$. The pure covariant metric tensor is g_{ab} or perhaps g_{**} . At first this seems a disadvantage of the notation, but one then realizes that the true object really is "g", and it has four different "representations" and the notation makes this very clear.

In this same spirit, we have the following translation for a bolded vector:

$\mathbf{V} \rightarrow \mathbf{V}$
$\bar{\mathbf{V}} \rightarrow \mathbf{V}$. (7.4.2)

The reason is that the overbar is no longer used to denote covariancy. The above lines show a subtle change in the interpretation of the bolded symbol **V** in the standard notation: the single symbol **V** stands for both the developmental vector **V** *and* for its developmental covariant partner vector $\bar{\mathbf{V}}$. The new symbol **V** is both contravariant with components V^n and it is covariant with components V_n.

(7.4.3)

2. Raising and Lowering Tensor Indices with g

As for converting a vector from one type to the other, the relations (5.8.4) become,

$\bar{V}_a = \bar{g}_{ab} V_b \rightarrow V_a = g_{ab} V^b$ // g_{ab} "lowers" a contravariant index
$V_a = g_{ab} \bar{V}_b \rightarrow V^a = g^{ab} V_b$ // g^{ab} "raises" a covariant index (7.4.4)

In the Standard Notation, the nature of a vector (contravariant or covariant) is determined by the up or down position of the index. Looking at $g_{ab}V^b = V_a$, we see that " g_{ab} lowers the index on V^b to give V_a "

which of course just means that $\overline{g}_{ab}V_b = \overline{V}_a$ in the developmental notation. It means nothing more, and nothing less. Similarly, $g^{ab} V_b = V^a$ says that " g^{ab} raises the index on V_b to give V^a " and this is just the Standard Notation way of saying $g_{ab} \overline{V}_b = V_a$.

The key idea here is that g_{ab} lowers a tensor index of a tensor, and g^{ab} raises a tensor index (both with an implied sum as above). We know from the outer product idea shown in (7.1.1) that we can construct a covariant rank-2 tensor from two covariant rank-1 tensors in this way.

$$T_{ab} \equiv U_aV_b \qquad \textit{(7.1.1)}$$

This is in Standard Notation now, so unlike (7.1.1), all three objects are covariant tensors. We can then use g^{ca} to raise the a index on both sides of this equation,

$$g^{ca}T_{ab} = T^c{}_b \qquad \text{// implied sum on a -- always implied sums on repeated indices!}$$

$$g^{ca}(U_aV_b) = (g^{ca}U_a)V_b = U^cV_b \quad \Rightarrow$$

$$T^c{}_b = U^cV_b \ . \qquad (7.4.5)$$

Here, $T^c{}_a$ is a "mixed" rank-2 tensor, the first index is contravariant, the second index is covariant. This fact matches the nature of the right side of the equation U^cV_b. We could have but did not deal with mixed tensors in the developmental notation because the notation could not really handle it well, as shown in (7.2.4). So in developmental notation, we dealt only with "pure" rank-2 tensors like M_{ab} and $\overline{M}_{ab}$.

As a next step, we can apply g^{db} to both sides of (7.4.5),

$$g^{db}\, T^c{}_b = T^{cd}$$

$$g^{db}\, (U^cV_b) = U^c(g^{db}V_b) = U^cV^d \quad \Rightarrow$$

$$T^{cd} = U^cV^d \qquad (7.4.6)$$

We end up with a contravariant rank-2 tensor that is the outer product of two contravariant rank-1 tensors. This is in fact the equation (7.1.1) expressed in Standard Notation.

Now let's do both index raising operations at the same time:

$$T_{ab} \equiv U_aV_b$$

$$g^{ca}\, g^{db}\, T_{ab} = T^{cd}$$

$$g^{ca}\, g^{db}\, U_aV_b = (g^{ca}U_a)(g^{db}V_b) = U^cV^d \quad \Rightarrow$$

$$T^{cd} = U^cV^d \qquad (7.4.7)$$

This shows that we are free to raise or lower corresponding indices on both sides of a tensor equation at will, and we don't have to write out the details. Thus if the first of these is true, then the other three forms are also valid:

$$T_{ab} \equiv U_aV_b \qquad \Rightarrow \qquad T^a{}_b = U^aV_b \quad T_a{}^b = U_aV^b \quad T^{ab} = U^aV^b \tag{7.4.8}$$

This idea works with *any* rank-2 tensor *because* any rank-2 tensor transforms under **x**' = **F**(**x**) in the same way that the outer product of two vectors transforms. If M is some arbitrary contravariant rank-2 tensor, then $g^{ca}M_{ab} = M^c{}_b$ and so on, repeating all of the above.

The idea applies as well to rank-3 tensors and rank-n tensors as we shall describe below in Section 7.10. For example, if $T_{abc} = U_aV_bW_c$, or if M is some arbitrary rank-3 tensor M_{abc}, then

$$T_{abc} \equiv U_aV_bW_c \qquad\qquad M_{abc}$$

$$g^{da}T_{abc} = T^d{}_{bc} \qquad\qquad g^{da}M_{abc} = M^d{}_{bc}$$

$$(g^{da}U_a)V_bW_c = U^dV_bW_c \Rightarrow$$

$$T^d{}_{bc} = U^dV_bW_c . \tag{7.4.9}$$

In this case, the validity of the first equation implies the validity of the 7 others:

$$T_{abc} \equiv U_aV_bW_c \quad \Rightarrow \quad T^a{}_{bc} \equiv U^aV_bW_c \qquad T_a{}^b{}_c \equiv U_aV^bW_c \qquad T_{ab}{}^c \equiv U_aV_bW^c$$

$$T^{ab}{}_c \equiv U^aV^bW_c \qquad T_a{}^{bc} \equiv U_aV^bW^c \qquad T^a{}_b{}^c \equiv U^aV_bW^c$$

$$\text{and finally } T^{abc} \equiv U^aV^bW^c \tag{7.4.10}$$

To summarize, in this new Standard Notation, the covariant metric tensor g_{ab} becomes an "index lowering operator" and the contravariant metric tensor g^{ab} becomes an "index raising operator". This is a huge advantage of the Standard Notation. It pretty much eliminates the need to think, something universally appreciated. In a certain obscure sense, it is like double entry accounting (credits and debits), where the notation itself serves as a check on the accuracy of bookkeeping entries.

The g raising and lower rules may be represented generically in this manner,

$$g_{aa'} [\text{----}^{a'}\text{---}] = [\text{----}_a\text{---}]$$
$$g^{aa'} [\text{----}_{a'}\text{---}] = [\text{----}^a\text{---}] \tag{7.4.11}$$

where $[\text{----}^{a'}\text{---}]$ represents an arbitrary tensor expression (perhaps just one tensor) with lots of tensor indices indicated by dashes. Each of these dash indices could be up or down, it does not matter. An example of the second line is $g^{aa'}M_{a'bc} = M^a{}_{bc}$.

3. The invariant distance and covariant dot product:

Based on the discussion so far, we can write the following translations, starting from (5.2.5) and (5.10.3),

$$dx_i \rightarrow dx^i$$

$$(ds)^2 = \bar{g}_{ab}\, dx_a\, dx_b \rightarrow g_{ab} dx^a dx^b$$

$$\mathbf{A \bullet B} = \bar{g}_{ab}\, A_a\, B_b \rightarrow \mathbf{A \bullet B} = g_{ab}\, A^a\, B^b\ . \qquad (7.4.12)$$

Using the raising and lowering idea above, we can write the dot product in other ways:

$$\mathbf{A \bullet B} = g_{ab}\, A^a\, B^b = A^a(\, g_{ab}B^b) = A^aB_a$$

$$\mathbf{A \bullet B} = g_{ab}\, A^a\, B^b = g_{ba}\, A^a\, B^b = (g_{ba}\, A^a)B^b = A_bB^b \qquad \text{// g is symmetric}$$

$$\mathbf{A \bullet B} = A_bB^b = A_b\,(g^{bc}B_c) = g^{bc}A_bB_c\ . \qquad (7.4.13)$$

To summarize :

$$\mathbf{A \bullet B} = g_{ab}\, A^a\, B^b = A^aB_a = A_aB^a = g^{ab}A_aB_b = \mathbf{B \bullet A}\ . \qquad (7.4.14)$$

The general idea is this: *any* tensor index on *any* tensor object can be raised by g^{ab} and can be lowered by g_{ab}. Remember that a tensor object lives in some space like x-space, so we shall have to ponder what to do for our matrices S_{ab} and R_{ab} which live half in x-space and half in x'-space, a subject we defer to Section 7.5.

4. Contraction Rules

When an index is summed with one index down and the other up, one says that the two indices are **contracted**. For example, in the expressions A^aB_a or A_aB^a, index a is contracted. In $g_{ab}A^aB^b$ both indices a and b are contracted.

Tilt Reversal Theorem. The tilt of any pair of contracted indices can be reversed without affecting anything. (7.4.15)

A proof is given below in Section 7.11. An example is $A^aB_a = A_aB^a$ as shown above in (7.4.14).

Contraction Theorem. In analyzing the transformation nature of a tensor expression, one can simply ignore all contracted indices and look only at remaining indices to determine the tensor nature of the object. (7.4.16)

A proof is given below in Section 7.12. As examples, in the cases A^aB_a or $g_{ab}A^aB^b$, if we ignore contracted indices, there are no indices left, so these objects must transform as a tensor with no indices, which is a tensorial scalar. Thus **A•B** above is a scalar, something we already knew from (5.10.2). The object A^aM_{ab} transforms as a covariant vector, and the object $M^a{}_a$ transforms as a scalar.

5. What about having g raise and lower indices *on itself*?

After all, g is a valid tensor. We find

$$g_{ab}g^{bc} = g_a{}^c \qquad \text{// lower the first index of rank-2 tensor } g^{bc}$$

$$g_{ac}g^{bc} = g^b{}_a \qquad \text{// lower the second index}$$

$$g_{cb}g^b{}_a = g_{ca} \ . \qquad \text{// lower both indices (lower second, then lower first)} \qquad (7.4.17)$$

We know from (7.4.1) above that g_{ab} and g^{ab} are inverses. Thus, it must be true that

$$g_{ab}g^{bc} = (1)_a{}^c = \delta_a{}^c = \delta_{a,c}$$

$$g^{bc}g_{ac} = (1)^b{}_a = \delta^b{}_a = \delta_{b,a} \ . \qquad (7.4.18)$$

Here $\delta_a{}^c$ and $\delta^b{}_a$ are just cosmetically nice ways to write the Kronecker deltas shown. Comparing (7.4.17) and (7.4.18) we learn that

$$g_a{}^c(\mathbf{x}) = \delta_a{}^c = \delta_{a,c}$$

$$g^b{}_a(\mathbf{x}) = \delta^b{}_a = \delta_{b,a} \ . \qquad (7.4.19)$$

Thus, the mixed forms of the metric tensor are trivial and are not functions of the nature of x-space.

7.5 Dealing with the matrices R and S ; various Rules and Theorems

1. Translations of R and S

Consider the translation of this partial derivative into the new up/down notation. Since the differential d**x** element is contravariant and is now written dx^i :

$$(\partial x'_i/\partial x_k) \qquad \rightarrow \qquad (\partial x'^i/\partial x^k) \ . \qquad (7.5.1)$$

In terms of "existence", this object has one leg in each space of Picture A of (1.11). The gradient operator $\partial/\partial x^k$ is an x-space object, while x'^i is an x'-space object. Since $(\partial x'^i/\partial x^k)$ does not live in x-space or in x'-space exclusively, but straddles the two spaces, it cannot possibly be a tensor of any kind. Recall that a tensor object must be entirely within a space, it cannot have body parts hanging out into other spaces. *Nevertheless*, it seems clear that each of the two indices has a well-defined nature. We showed in (2.4.2) that the gradient transforms as a covariant vector, so we regard k as a covariant index and write $\partial/\partial x^k$ =

∂_k. And of course dx'^i is a contravariant vector, so i is a contravariant index. Here then is the proper translation starting with (2.5.1),

$$R_{ik} \equiv (\partial x'_i/\partial x_k) \qquad \rightarrow \qquad R^i{}_k \equiv (\partial x'^i/\partial x^k) = \partial_k x'^i \ . \tag{7.5.2}$$

To summarize, $R^i{}_k$ is *not* a mixed rank-2 tensor, though it looks just like one. Therefore, $R^i{}_k$ can never appear in a tensor equation -- it just appears in the equations that show how tensors transform. However, each of the two indices of R has a well-defined transformational nature, and we place them up and down in the proper manner.

It is not atypical for an object to have up and down indices but the object is not a tensor. As was noted below (2.3.2), the canonical example is that for a non-linear transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$, x^i has a contravariant index but is not (does not transform under **F** as) a contravariant vector.

Consider now the translation of the transformation rule for a contravariant vector from (2.5.1),

$$V'_a = R_{ab}V_b \qquad \rightarrow \qquad V'^a = R^a{}_b V^b \ . \qquad \text{// contravariant} \tag{7.5.3}$$

Even though R is not a tensor, we see that index b is contracted and is thus neutralized from the evaluation of the tensor nature of the RHS. This leaves upper index a as the only free index, indicating that the RHS is a contravariant vector, and this of course then matches the LHS. However, it is a contravariant vector in x'-space, which will be further discussed in (7.5.9) below. Main point: we can deal with the indices on R just as we deal with indices on true tensors.

Notice that, even though both sides of $V'^a = R^a{}_b V^b$ have the same "tensor nature" (both sides are a contravariant vector component in x'-space) one cannot ask how the equation $V'^a = R^a{}_b V^b$ "transforms" under a transformation. That question can only be asked about equations constructed of objects all of which are tensors in the same space. Here V and half of R are in x-space, and V' and the other half of R are in a x'-space. There is no object called R', as if R were in x-space and R' were in x'-space.

We can repeat the above discussion for S instead of R. We omit the words and just show the translations, quoting expressions from (2.1.6) and (2.5.1) :

$$(\partial x_i/\partial x'_k) \qquad \rightarrow \qquad (\partial x^i/\partial x'^k)$$

$$S_{ik} \equiv (\partial x_i/\partial x'_k) \qquad \rightarrow \qquad S^i{}_k \equiv (\partial x^i/\partial x'^k) = \partial'_k x^i$$

$$\bar{V}'_a = S^T{}_{ab}\bar{V}_b = S_{ba}\,\bar{V}_b \qquad \rightarrow \qquad V'_a = S^b{}_a V_b \qquad \text{// covariant} \tag{7.5.4}$$

<u>2, Transformation rule for rank-2 tensors</u>

For a contravariant rank-2 tensor, the translation of the transformation rules (5.6.3) are,

$$\begin{aligned} M'_{ab} &= R_{aa'}R_{bb'}M_{a'b'} && \rightarrow && M'^{ab} = R^a{}_{a'}R^b{}_{b'}M^{a'b'} && \text{contravariant rank-2 tensor} \\ \bar{M}'_{ab} &= S_{a'a}S_{b'b}\bar{M}_{a'b'} && \rightarrow && M'_{ab} = S^{a'}{}_a S^{b'}{}_b M_{a'b'} \ . && \text{covariant rank-2 tensor} \end{aligned} \tag{7.5.5}$$

Applying this to the metric tensor M = g we find these translations of (5.7.6),

$$\begin{array}{llll} g' = R\, g\, R^T & \Rightarrow g'_{ab} = R_{aa'}R_{bb'}g_{a'b'} & \rightarrow & g'^{ab} = R^a{}_{a'}R^b{}_{b'}g^{a'b'} \\ \overline{g}' = S^T\, \overline{g}\, S & \Rightarrow \overline{g}'_{ab} = S_{a'a}S_{b'b}\,\overline{g}_{a'b'} & \rightarrow & g'_{ab} = S^{a'}{}_a S^{b'}{}_b\, g_{a'b'}\ . \end{array} \tag{7.5.6}$$

The inverses of the above equations appear in (5.7.7) and translate this way ,

$$\begin{array}{llll} g = S\, g'\, S^T & \Rightarrow g_{ab} = S_{aa'}S_{bb'}g'_{a'b} & \rightarrow & g^{ab} = S^a{}_{a'}S^b{}_{b'}g'^{a'b} \\ \overline{g} = R^T\, \overline{g}'\, R & \Rightarrow \overline{g}_{ab} = R_{a'a}R_{b'b}\overline{g}'_{a'b'} & \rightarrow & g_{ab} = R^{a'}{}_a R^{b'}{}_b\, g'_{a'b'}\ . \end{array} \tag{7.5.7}$$

By raising and lowering indices on the first result in (7.5.5), and using the tilt reversal rule (7.4.15), we obtain these four transformation rules :

$$\begin{aligned} M'^{ab} &= R^a{}_{a'}\ R^b{}_{b'}\ M^{a'b'} \\ M'^a{}_b &= R^a{}_{a'}\ R_b{}^{b'}\ M^{a'}{}_{b'} \\ M'_a{}^b &= R_a{}^{a'}\ R^b{}_{b'}\ M_{a'}{}^{b'} \\ M'_{ab} &= R_a{}^{a'}\ R_b{}^{b'}\ M_{a'b'}\ . \end{aligned} \tag{7.5.8}$$

One sees then a family of four tensors associated with M. The first is contravariant, the last is covariant, and the other two are mixed. Comparison of the last line of (7.5.8) to the last line of (7.5.5) suggests that $R_a{}^{a'} = S^{a'}{}_a$, and we will formally prove this fact in (7.5.13) below.

3. Raising and lowering indices on R and S

Although R and S are not tensors, one can still raise and lower their two indices using metric tensors, but things are a little different from the tensor situation, the reason being that R and S each have one foot in x-space and the other foot in x'-space.

Object $R^a{}_b = (\partial x'^a/\partial x^b)$ was considered above. One could lower the a index using g'_{**} since x'^a is in x'-space and is an up index. The index in $\partial/\partial x^b = \partial_b$ is really a lower index (gradient), so one could in effect raise it using g^{**} (no prime) because $\partial/\partial x^b$ is in x-space. So when raising and lowering indices on $R^a{}_b$ one has the *unusual situation* that one must use g' when acting on the first index, and g when acting on the second. With this in mind, we can now write three other index configurations of $R^a{}_b$:

$$\begin{array}{lll} R^a{}_b & = (\partial x'^a/\partial x^b) & \text{// original object (formerly } R_{ab}) \\ R^{ab} = R^a{}_{b'}\ g^{b'b} & = (\partial x'^a/\partial x_b) & \text{// g pulls up the second index of } R^a{}_b \\ R_{ab} = g'_{aa'}R^{a'}{}_b & = (\partial x'_a/\partial x^b) & \text{// g' pulls down the first index of } R^a{}_b \\ R_a{}^b = g'_{aa'}R^{a'}{}_{b'}\ g^{b'b} & = (\partial x'_a/\partial x_b)\ . & \text{// both actions at once} \end{array} \tag{7.5.9}$$

Although the g and g' factors can be placed anywhere, we have put g' factors on the left of R, and g factors on the right, each next to its appropriate leg of R.

In each case, examination of the corresponding partial derivative shows that that the index sense matches on both sides. For example, in $R^{ab} = (\partial x'^a/\partial x_b) = \partial^b x'^a$, both indices are contravariant on both sides. Remember that R^{ab} is not a contravariant rank-2 tensor due to its dual-space nature.

In the same manner, we arrive at these index configurations for $S^a{}_b$:

$$\begin{aligned}
&S^a{}_b(\mathbf{x}) &&= (\partial x^a/\partial x'^b) &&\text{// original object (formerly } S_{ab}\text{)}\\
&S^{ab} \equiv S^a{}_{b'}\, g'^{b'b} &&= (\partial x^a/\partial x'_b) &&\text{// g' pulls the second index up}\\
&S_{ab} \equiv g_{aa'} S^{a'}{}_b &&= (\partial x_a/\partial x'^b) &&\text{// g pulls the first index down}\\
&S_a{}^b \equiv g_{aa'} S^{a'}{}_{b'}\, g'^{b'b} &&= (\partial x_a/\partial x'_b) &&\text{// both actions at once} \qquad (7.5.10)
\end{aligned}$$

For object S, the metric tensor g raises or lowers the first index of S, while g' raises or lowers the second index of S. This is just the reverse of what happens for object R as reported above. This is not surprising since one gets S↔R when one swaps x-space ↔ x'-space.

4. Inverse of R and S (and Comment on Transpose)

We have showed the translation $R_{ab} \rightarrow R^a{}_b$. In the standard notation, imagine that there is some inverse R^{-1} *defined* by $(R^{-1})^c{}_a R^a{}_b = \delta^c{}_b$. The chain rule says that

$$(\partial x^c/\partial x'^a)\,(\partial x'^a/\partial x^b) = \delta^c{}_b \qquad \text{or} \qquad S^c{}_a\, R^a{}_b = \delta^c{}_b$$

$$(\partial x'^c/\partial x^a)\,(\partial x^a/\partial x'^b) = \delta^c{}_b \qquad \text{or} \qquad R^c{}_a\, S^a{}_b = \delta^c{}_b\,. \qquad (7.5.11)$$

Therefore from the first line it must be that $(R^{-1})^c{}_a = S^c{}_a$. The second line shows that $(S^{-1})^c{}_a = R^c{}_a$. Using the above rules for raising and lowering indices on both sides of an equation, we obtain these four versions of the developmental notation fact that $R^{-1} = S$ and $S^{-1} = R$:

$$\begin{aligned}
&(R^{-1})^{ik} = S^{ik} &&(S^{-1})^{ik} = R^{ik}\\
&(R^{-1})^i{}_k = S^i{}_k &&(S^{-1})^i{}_k = R^i{}_k\\
&(R^{-1})_i{}^k = S_i{}^k &&(S^{-1})_i{}^k = R_i{}^k\\
&(R^{-1})_{ik} = S_{ik} &&(S^{-1})_{ik} = R_{ik}\ . \qquad (7.5.12)
\end{aligned}$$

In (7.5.11) we see a certain "up tilt" matrix multiplication which will be explained below in Section 7.8.

5. The R-S Tilt Theorem: $S^a{}_b = R_b{}^a$ which is the same as: $(\partial x^a/\partial x'^b) = (\partial x'_b/\partial x_a)$ (7.5.13)

Notice in this rule that the indices are reflected in a vertical line running between the indices: $S^a{}_{|b} = R_{b|}{}^a$.

Proof: This proof is a bit long-winded, but brings in many earlier results and is a good exercise:

$\delta_b{}^{a"} S^a{}_{a"} = S^a{}_b$ // introduce a δ . Remember all g's are symmetric.

$(g'_{bb'}\, g'^{a"b'})\; S^a{}_{a"} = S^a{}_b$ // since g'_{ab} and g'^{ab} are inverses, see (7.4.18)

$g'_{bb'}\, \delta^{b'}{}_{b"}\, S^a{}_{a"}\, g'^{a"b"} = S^a{}_b$ // reorder and use $g'^{a"b'} = \delta^{b'}{}_{b"} g'^{a"b"}$

$g'_{bb'}\, (R^{b'}{}_{a'}\, S^{a'}{}_{b"})\, S^a{}_{a"}\, g'^{a"b"} = S^a{}_b$ // $\delta^{b'}{}_{b"} = (R^{b'}{}_{a'}\, S^{a'}{}_{b"})$ from (7.5.11)

$g'_{bb'}\, R^{b'}{}_{a'}\, (S^a{}_{a"}\, S^{a'}{}_{b"}\, g'^{a"b"}) = S^a{}_b$ // regroup (7.5.14)

$g'_{bb'}\, R^{b'}{}_{a'}\, (g^{aa'}) = S^a{}_b$ // $g^{aa'} = S^a{}_{a"}\, S^{a'}{}_{b"}\, g'^{a"b"}$ from (7.5.7)

$(g'_{bb'}\, R^{b'}{}_{a'}\, g^{aa'}) = S^a{}_b$ // regroup

$R_b{}^a = S^a{}_b$. // use last line of (7.5.9)

Notice that the above theorem says

$$S^a{}_b = (\partial x^a/\partial x'^b) = (\partial x'_b/\partial x_a) = R_b{}^a. \tag{7.5.15}$$

Similar results can be derived for other index positions (or we can just raise and lower indices!) to get

$$\begin{array}{ll} S^a{}_b = R_b{}^a & = (\partial x^a/\partial x'^b) = (\partial x'_b/\partial x_a) \\ S_{ab} = R_{ba} & = (\partial x_a/\partial x'^b) = (\partial x'_b/\partial x^a) \\ S^{ab} = R^{ba} & = (\partial x^a/\partial x'_b) = (\partial x'^b/\partial x_a) \\ S_a{}^b = R^b{}_a & = (\partial x_a/\partial x'_b) = (\partial x'^b/\partial x^a)\ . \end{array} \tag{7.5.16}$$

Here index a is always in x-space, while index b is in x'-space.

6. Determinants of R and S.

We showed earlier that $R_{ij} \rightarrow R^i{}_j$ and $S_{ij} \rightarrow S^i{}_j$. The determinants det(R) and det(S) translate as follows:

$$\det(R) = \varepsilon_{abc...} R_{1a} R_{2b} .... R_{Nx} \quad \rightarrow \quad \det(R^*{}_*) = \varepsilon_{abc...} R^1{}_a R^2{}_b .... R^N{}_x$$
$$\det(S) = \varepsilon_{abc...} S_{1a} S_{2b} ..... S_{Nx} \quad \rightarrow \quad \det(S^*{}_*) = \varepsilon_{abc...} S^1{}_a S^2{}_b .... S^N{}_x$$

or

$$\det(R) = \varepsilon_{abc...}R_{a1}R_{b2}....R_{xN} \quad \rightarrow \quad \det(R^*{}_*) = \varepsilon_{abc...}R^a{}_1R^b{}_2....R^x{}_N$$
$$\det(S) = \varepsilon_{abc...}S_{a1}S_{b2}.....S_{xN} \quad \rightarrow \quad \det(S^*{}_*) = \varepsilon_{abc...}S^a{}_1S^b{}_2....S^x{}_N \qquad (7.5.17)$$

where ε is the bookkeeping permutation tensor discussed in Section 7.7 below.

Reader Exercise: Write $\det(R^*{}_*) = (\varepsilon_{abc}R^1{}_aR^2{}_bR^3{}_c)$ and $\det(S^*{}_*) = (\varepsilon_{a'b'c'}S^{a'}{}_1S^{b'}{}_2S^{b'}{}_3)$. We know that in either developmental or standard notation det(R)det(S) = det(RS) = det(1) = 1. The exercise is to verify this directly using the N=3 first expression in (D.10.37) for $\varepsilon_{abc}\varepsilon_{a'b'c'}$ along with (7.5.11).

7. Jacobian and related

Here are translations of a few equations from Section 5.12:

(5.12.2) $$\det(S) = 1/\det(R) \quad \rightarrow \quad \det(S^i{}_j) = 1/\det(R^i{}_j) \qquad (7.5.18)$$

(5.12.6) $$J = \det(S) \quad \rightarrow \quad J = \det(S^i{}_j) \qquad (7.5.19)$$

(5.12.12) $$g = \det(\overline{g}) \quad \rightarrow \quad g = \det(g_{ij})$$
$$g' = \det(\overline{g}') \quad \rightarrow \quad g' = \det(g'_{ij}) \qquad (7.5.20)$$

(5.12.14) $$1/g = \det(g) \quad \rightarrow \quad 1/g = \det(g^{ij})$$
$$1/g' = \det(g') \quad \rightarrow \quad 1/g' = \det(g'^{ij}) \qquad (7.5.21)$$

(5.12.14) $$g' = J^2\, g \quad \rightarrow \quad g' = J^2\, g$$
$$|J| = \sqrt{g'/g} \quad \rightarrow \quad |J| = \sqrt{g'/g} \qquad (7.5.22)$$

7.6 Orthogonality Rules, Inversion Rules, Cancellation Rules

1. Orthogonality Rules:

Theorem (7.5.13) that $S^a{}_b = R_b{}^a$ can be used to eliminate S in various forms of SR = 1:

$$SR = 1 \;\rightarrow\; S^a{}_b\, R^b{}_c = \delta^a{}_c \;\;\Rightarrow\;\; R_b{}^a\, R^b{}_c = \delta^a{}_c \;\;\Rightarrow\;\; R_b{}^a\, R^b{}_c = g^a{}_c \text{ from (7.4.19)} \quad . \qquad (7.6.1)$$

Although R is not a tensor, the object $R_b{}^a\, R^b{}_c$ is a tensor of the mixed type $M^a{}_c$. Thus, both sides of this last equation transform as this type of mixed rank-2 tensor.

In this last equation, we can lower index a on both sides, raise index c on both sides, and reverse the tilt of the b contraction to get this result, again using (7.4.19),

$$R^b{}_a\, R_b{}^c = g_a{}^c = \delta_a{}^c \, . \qquad (7.6.2)$$

We now repeat this process starting instead with RS = 1:

RS = 1 → $R^a{}_b S^b{}_c = \delta^a{}_c$ ⇒ $R^a{}_b R_c{}^b = \delta^a{}_c$ ⇒ $R^a{}_b R_c{}^b = g^a{}_c$

$R_a{}^b R^c{}_b = g_a{}^c = \delta_a{}^c$. (7.6.3)

Collecting the four forms written with δ we obtain the four **orthogonality rules** for R:

1: $R_b{}^a R^b{}_c = \delta^a{}_c$ 2: $R^b{}_a R_b{}^c = \delta_a{}^c$ // sum is on first index

3: $R^a{}_b R_c{}^b = \delta^a{}_c$ 4: $R_a{}^b R^c{}_b = \delta_a{}^c$ // sum is on second index (7.6.4)

For any general transformation **x**' = **F**(**x**), these rules are valid. Going the reverse direction, if we replace all $R_x{}^y$ with $S^y{}_x$, the four orthogonality rules of (7.6.4) can be rederived in this manner,

1: $R_b{}^a R^b{}_c = R_b{}^a S_c{}^b = S_c{}^b R_b{}^a = (SR)_c{}^a = (1)_c{}^a = \delta_c{}^a = \delta^a{}_c$ // down-tilt

2: $R^b{}_a R_b{}^c = R^b{}_a S^c{}_b = S^c{}_b R^b{}_a = (SR)^c{}_a = (1)^c{}_a = \delta^c{}_a = \delta_a{}^c$ // up-tilt

3: $R^a{}_b R_c{}^b = R^a{}_b S^b{}_c = (RS)^a{}_c = (1)^a{}_c = \delta^a{}_c$ // up-tilt

4: $R_a{}^b R^c{}_b = R_a{}^b S_b{}^c = (RS)_a{}^c = (a)_a{}^c = \delta_a{}^c$ // down-tilt

All these rules just say RS = SR = 1. Notice in the last four lines the notion of up-tilt and down-tilt matrix multiplication. This subject is discussed in Section 7.8 below.

Comment: The orthogonality rules are easy to memorize using these few facts:

- The sum is either on both first indices, or on both second indices
- the tilts of the two R's are opposite each other.

2. Inversion Rules.

The first rule makes this claim:

$[----^a-----] = R^a{}_b [--------^b------]$ ⇔ $R_a{}^b [----^a-----] = [--------^b------]$ (7.6.5)

sum on 2nd index b sum on 1st index a

so if one wants to invert the left equation for $[--------^b------]$, the result is as shown on the right. Notice that the R indices are reflected through a horizontal line segment between the indices. This is different from the fact (7.5.13) that $S^a{}_b = R_b{}^a$ where indices are reflected through a vertical line segment between the indices.

Proof: Start with $[----^a-----] = R^a{}_{b'} [--------^{b'}------]$. Apply $R_a{}^b$ to both sides, then use (7.6.4) :

$R_a{}^b[----^a-----] = R_a{}^b R^a{}_{b'} [--------^{b'}------] = \delta^b{}_{b'} [--------^{b'}------] = [--------^b------]$ QED

The rule is also valid if all tilts are reversed. Thus,

$$[\text{----}_a\text{-----}] = R_a{}^b\,[\text{--------}_b\text{------}] \quad \Leftrightarrow \quad R^a{}_b\,[\text{----}_a\text{-----}] = [\text{--------}_b\text{------}] \qquad (7.6.6)$$

sum on 2nd index b　　sum on 1st index a

Proof: Start with $[\text{----}_a\text{-----}] = R_a{}^{b'}\,[\text{--------}_{b'}\text{------}]$. Apply $R^a{}_b$ to both sides, then use (7.6.4) :

$$R^a{}_b[\text{----}_a\text{-----}] = R^a{}_b\,R_a{}^{b'}\,[\text{--------}_{b'}\text{------}] = \delta_b{}^{b]}\,[\text{--------}_{b'}\text{------}] = [\text{--------}_b\text{------}] \quad \text{QED}$$

Example of (7.6.5): $V'^a = R^a{}_b V^b \quad \Leftrightarrow \quad R_a{}^b V'^a = V^b \quad \text{or} \quad V^b = R_a{}^b V'^a$ (7.6.7)

Example of (7.6.6): $V'_a = R_a{}^b V_b \quad \Leftrightarrow \quad R^a{}_b V'_a = V_b \quad \text{or} \quad V_b = R^a{}_b V'_a$ (7.6.8)

Example: Rank-3 tensor transformation inversion :

$$M'^{abc} = R^a{}_i R^b{}_j R^c{}_k M^{ijk} \qquad \text{// rank-3 transformation rule (contravariant)}$$
$$M^{abc} = R_i{}^a R_j{}^b R_k{}^c M'^{ijk} \qquad \text{// its inverse} \qquad (7.6.9)$$

Notice that, in the inverse transformation, the R indices are up-tilted and the summation indices are strangely on the left side. Using the raising and lowering rules of (7.5.9), one sees that the corresponding covariant forms are obtained just by moving indices vertically on both sides. Thus.

$$M'_{abc} = R_a{}^i R_b{}^j R_c{}^k M_{ijk} \qquad \text{// rank-3 transformation rule (covariant)}$$
$$M_{abc} = R^i{}_a R^j{}_b R^k{}_c M'_{ijk} \qquad \text{// its inverse} \qquad (7.6.10)$$

3. Cancellation Rules

The first rule makes this claim:

$$R^a{}_b\,[\text{----}^b\text{-----}] = R^a{}_b\,[\text{--------}^b\text{------}] \quad \Leftrightarrow \quad [\text{----}^b\text{-----}] = [\text{--------}^b\text{------}] \qquad (7.6.11)$$

where the two R factors can be cancelled out.

Proof: Start with $R^a{}_{b'}\,[\text{----}^{b'}\text{-----}] = R^a{}_{b'}\,[\text{--------}^{b'}\text{------}]$. Apply $R_a{}^b$ to both sides, then use (7.6.4) :

$$R_a{}^b R^a{}_{b'}\,[\text{----}^{b'}\text{-----}] = R_a{}^b R^a{}_{b'}\,[\text{--------}^{b'}\text{------}] \quad \Rightarrow$$

$$\delta^b{}_{b'}\,[\text{----}^{b'}\text{-----}] = \delta^b{}_{b'}\,[\text{--------}^{b'}\text{------}] \quad \Rightarrow \quad [\text{----}^b\text{-----}] = [\text{--------}^b\text{------}] \quad \text{QED}$$

The rule is also valid if all tilts are reversed. Thus,

$$R_a{}^b\,[\text{----}_b\text{-----}] = R_a{}^b\,[\text{--------}_b\text{------}] \quad \Leftrightarrow \quad [\text{----}_b\text{-----}] = [\text{--------}_b\text{------}] \qquad (7.6.12)$$

The proof is similar to that shown above.

7.7 About δ and ε

The Kronecker δ is sometimes written in different ways to make things "look nice",

$$\delta_a{}^b = \delta^a{}_b = \delta_b{}^a = \delta^b{}_a = \delta_{a,b} \ . \qquad (7.7.1)$$

Eq (7.4.19) showed that that one can regard the above sequence of equalities as saying

$$g_a{}^b = g^a{}_b = g_b{}^a = g^b{}_a = \delta_{a,b} \qquad (7.7.2)$$

where these g objects are mixed versions of the symmetric rank-2 metric tensor g_{ab}. There is no "δ tensor", it is the g metric tensor, but tradition is to write the diagonal objects using the δ symbol.

The object $\varepsilon_{abc...}$ is a bit more complicated. It can at first be regarded as a mere bookkeeping device, in which context it is usually called "the permutation tensor". It appears for example in the expansion of a determinant in N dimensions,

$$\det(M) = \varepsilon_{abc...x}M_{1a}M_{2b}.....M_{Nx} = \varepsilon_{abc...x}M_{a1}M_{b2}.....M_{xN} \qquad (7.7.3)$$

or in an ordinary cross product in Cartesian space,

$$A_a = \varepsilon_{abc}B_bC_c \ . \qquad (7.7.4)$$

This permutation tensor has the usual properties that $\varepsilon_{123...N} = +1$, that ε changes sign when any two indices are swapped, and that ε vanishes if two or more indices are the same. This permutation "tensor" is not really a tensor since one would regard it as being the same in x-space or x'-space. Whether indices are written up or down on this ε is immaterial.

At another level, however, $\varepsilon_{abc...x}$ with N indices (the same ε symbol is used) is a covariant rank-N tensor density of weight -1 known as the Levi-Civita tensor. This subject is addressed in Appendix D in much detail. In what we call the Weinberg convention, individual indices of ε can be raised and lowered by g as discussed in Section 7.4 subsection 2, just as with any tensor. Therefore, in Cartesian space with $g = 1$, indices on ε are raised and lowered with no consequence [see (5.9.1) or (7.4.4)], and then one can identify any form of ε as being the permutation tensor. For example, $\varepsilon^{abc} = \varepsilon_{abc} = \varepsilon_a{}^b{}_c$ and so on. In a non-Cartesian x-space, however, one would say that $\varepsilon_a{}^b{}_c = g^{bb'}\varepsilon_{ab'c} \neq \varepsilon_{abc}$. In the Weinberg convention, one sets $\varepsilon^{123..N} = \varepsilon'^{123..N} = 1$ and $\varepsilon^{abc..x} = \varepsilon'^{abc..x}$ has the properties of the permutation tensor described above and these properties are the same in x-space as in x'-space. Then for general $g\neq 1$, $\varepsilon_{abc..x}$ (lower indices) is NOT the permutation tensor. The bottom line is that one must be aware of the space in which one is working (the Picture). The ε appearing above in the determinant expansion (7.7.3) is always just the permutation tensor, but in the cross product that is not the case, and one would properly write

$$A_a = \varepsilon_{abc}B^bC^c \qquad (7.7.5)$$

and conclude that the cross product of two ordinary contravariant vectors is a covariant vector density (Section D.8). Again, in Cartesian space where one often works, this would be the same as $A_a = \varepsilon^{abc}B^bC^c = \varepsilon_a{}^b{}_cB^bC^c$, but the "properly tilted form" $A_a = \varepsilon_{abc}B^bC^c$ reveals the tensor nature of the object A_a. As mentioned below in Section 7.15, this "covariant" equation would appear as $A'_a = \varepsilon'_{abc}B'^bC'^c$ in x'-space, but since A'_a is a covariant vector density, $A'_a \neq R_a{}^bA_b$, and in fact $A'_a = J\ R_a{}^bA_b$.

The permutation tensor $\varepsilon_{abc...x}$ and the contravariant Levi-Civita tensor $\varepsilon^{abc...x}$ are both "totally antisymmetric" which just means ε changes sign if any pair of indices is swapped. In fact, as discussed in Section D.3, there IS only one antisymmetric tensor of rank N apart from a multiplicative scalar factor, and $\varepsilon^{abc...x}$ is it. This fact simplifies various calculations. Technically, $\varepsilon^{abc...x}$ is a totally antisymmetric tensor *density*, but normally it is just called "the totally antisymmetric tensor". As implied by (D.5.1), the covariant Levi-Civita tensor $\varepsilon_{abc...x}$ is also totally antisymmetric and is therefore a multiple of $\varepsilon^{abc...x}$.

The reader is invited to peruse Appendix D at some appropriate time for more about tensor densities and the ε tensor.

7.8 Covariance and Matrix Multiplication

Before continuing the process of translation from developmental to standard notation, we digress momentarily to consider the notion of covariance in developmental notation.

As we shall discuss in more detail below in Section 7.15, an equation is said to be covariant under the transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ if it has "the same form" in both x-space and x'-space. The "same form" means that the equation looks the same but everything is primed in x'-space.

Example 1: Newton's Law $\mathbf{F} = m\mathbf{a}$ is covariant under rotations ($\mathbf{x}' = \mathbf{F}(\mathbf{x}) = R\mathbf{x}$), and in x'-space this law takes the form $\mathbf{F}' = m'\mathbf{a}'$ which has the same form as the equation in x-space $\mathbf{F} = m\mathbf{a}$. Once we know that **F** and **a** are contravariant vectors and m is a scalar, this conclusion is automatic from (2.3.2),

$$\mathbf{F} = m\mathbf{a} \Rightarrow \mathbf{F}' = m'\mathbf{a}' \text{ proof:} \qquad \mathbf{F}' = R\mathbf{F} = R(m\mathbf{a}) = m\,R\mathbf{a} = m\,\mathbf{a}' = m'\,\mathbf{a}' \qquad (7.8.1)$$

where m = m' follows since mass is a scalar under rotation. Thus, Newton's Law has the same form when it is examined in two frames of reference related by a rotation. It is covariant under rotations.

Example 2: Consider the equation $\mathbf{A}\bullet\mathbf{B} = \pi$ where **A** and **B** are contravariant vectors and • is the covariant dot product shown in (5.10.3), $\mathbf{A}\bullet\mathbf{B} \equiv \bar{g}_{ab}A_aB_b$. It was shown in (5.10.2) that the quantity **A•B** transforms as a scalar under general transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ so that $\mathbf{A}'\bullet\mathbf{B}' = \mathbf{A}\bullet\mathbf{B}$. Since the number π is also a scalar under any transformation (it is a constant), one could say that π' = π (it is the same number 3.14159 in x'-space and x-space), so

$$\mathbf{A}\bullet\mathbf{B} = \pi \Rightarrow \mathbf{A}'\bullet\mathbf{B}' = \pi' \qquad \text{, equation is covariant.} \qquad (7.8.2)$$

What we see here is that an equation is covariant IFF both sides of the equation transform as the same tensorial tensor type under the transformation of interest. In Example 1, both sides of $\mathbf{F} = m\mathbf{a}$ transform as contravariant vectors under rotations, and in Example 2 both sides of **A•B** = π transform as scalars under a *general* transformation.

Example 3: Consider the outer product equation (7.1.1) $T_{ab} = U_a V_b$ where U and V are contravariant vectors. We show in (7.1.2) that T_{ab} transforms as a contravariant rank-2 tensor. Both sides of this equation transform in this way, so in x'-space the equation becomes $T'_{ab} = U'_a V'_b$. The equation is therefore covariant under the transformation **x**' = **F**(**x**) with d**x**' = Rd**x**.

Approaching this example in a slightly different manner, suppose we *define* $T_{ab} \equiv U_a V_b$ where U and V are contravariant vectors. We then ask: Is T_{ab} a contravariant rank-2 tensor? Line (7.1.2) shows that the answer is yes,

$$T'_{ab} = U'_a V'_b = (R_{aa'} U_{a'})\,(R_{bb'} V_{b'}) = R_{aa'}\, R_{bb'}\, U_{a'} V_{b'} = R_{aa'}\, R_{bb'}\, T_{a'b'} \qquad (7.8.3)$$

which matches the transformation rule as stated in (5.6.3).

Example 4: Suppose A and B are tensorial contravariant rank-2 tensors. Is the equation AB = C covariant? If it were, we would have to show that in x'-space we have A'B' = C' where C is a contravariant rank-2 tensor. To investigate, we use the rule (5.7.1) which states how a contravariant rank-2 tensor transforms in terms of Picture A shown in (5.7.2) :

$$A'B' = (RAR^T)(RBR^T) = RA(R^TR)BR^T \qquad \text{// since A and B are contra rank-2 tensors} \qquad (7.8.4)$$

$$C' = RCR^T = RABR^T \qquad \text{// assuming C is also a contra rank-2 tensor and AB = C}$$

If it were true that $R^TR = 1$, one would find from the first line above that $A'B' = RABR^T = RCR^T = C'$ and the answer would be yes, the equation AB = C is covariant. However, for a general Picture A transformation with metric tensor g in x-space and g' in x'-space, what we know about R comes from (5.7.6) : $g' = R\, g\, R^T$. Even if g = 1 so x-space is Cartesian, this says $g' = RR^T$, but this tells us nothing about R^TR. So for a general transformation, we have $R^TR \neq 1$ and so the equation AB = C is NOT covariant. [In the *special case* that R is a rotation, so $R^T = R^{-1}$ (real orthogonal), then $R^TR= R^{-1}R = 1$.]

As with Example 3, we can reformulate the current example in a different manner. Suppose we define C ≡ AB and specify that both A and B are contravariant rank-2 tensors. In this case, is C a contravariant rank-2 tensor? If it were, we would have to have $(A'B') = R(AB)R^T$ from (5.7.1). But we showed above that, since $R^TR \neq 1$, we end up with $(A'B') \neq R(AB)R^T$. Therefore, C ≡ AB is not a contravariant rank-2 tensor.

Could C be a *covariant* rank-2 tensor? If it were, we would need to have $(A'B') = S^T(AB)S$ from (5.7.1). But above we show that $A'B' = RA(R^TR)BR^T$ and this is completely different from $(A'B') = S^T(AB)S$. Thus, C is not a covariant rank-2 tensor.

Since C has two indices, the only way it could be a tensorial tensor is if it is either a contravariant or a covariant rank-2 tensor, but we have just ruled out both these possibilities.

Therefore C ≡ AB is not a tensorial tensor of any kind whatsoever, even though A and B are tensorial tensors.

Matrix Rule #1. In developmental notation, if A and B are contravariant rank-2 tensors, the matrix product AB is (in general) not a rank-2 tensor and is in fact not any kind of tensor. The equation C = AB is not covariant. Mimicking the above discussion, the reader can show that the same conclusion applies to $C = \overline{A}B$, $C = A\overline{B}$ and $C = \overline{A}\,\overline{B}$: in none of these cases is C a tensor of any kind, and all these equations are non-covariant. Similarly, the Rule applies to X = ABC or X = ABCD and so on. (7.8.5)

For this reason, we shall never ask how to transform an equation like X = ABC... from developmental to standard notation. Equations which are non-covariant are simply of no interest, and can never describe a physical relationship, as explained below in Section 7.15.

The attentive reader might ask: What about the equation $g' = R\, g\, R^T$ which has the form X = ABC. And if g = 1, what about $g' = RR^T$ whose form is X = AB? In both these cases, the left hand side *is* a contravariant rank-2 tensor. These equations do not violate the Matrix Rule #1 above because the matrices R and R^T are not tensors of any kind, as noted below (7.5.2). Furthermore, one does not ask whether $g' = R\, g\, R^T$ is covariant or not because it is an equation relating objects in different spaces and not all objects in the equation are tensors.

We now consider the notion of matrix multiplication using mixed rank-2 tensors. Since we never introduced such mixed tensors in our developmental notation, we have this discussion entirely in the Standard Notation. Consider

$$C^i{}_j = A^i{}_k B^k{}_j \,. \qquad \text{// implied sum on k} \tag{7.8.6}$$

The indices k have the right adjacency so one could think of this as being a matrix equation C = AB where all three objects are "down-tilt rank-2 mixed tensors". Down-tilt just means the two indices are tilting down like ${}^i{}_j$. We can ask again our questions of Example 4. If A and B are rank-2 tensors, is C = AB covariant? And if we define C ≡ AB, is C a rank-2 tensor?

The answer to both questions is yes.

To show that AB = C is covariant, we start with the second line of (7.5.8) with $R_b{}^{b'} = S^{b'}{}_b$ from (7.5.13) :

$$M'^a{}_b = R^a{}_{a'}\ S^{b'}{}_b\, M^{a'}{}_{b'} \qquad\qquad \textit{(7.5.8)}$$

and we apply this transformation rule to both A and B to get

$$\begin{aligned} A'^i{}_k B'^k{}_j &= (R^i{}_{a'}\ S^{b'}{}_k\, A^{a'}{}_{b'})\,(R^k{}_{a''}\ S^{b''}{}_j\, B^{a''}{}_{b''}) \\ &= R^i{}_{a'}\ S^{b'}{}_k\ R^k{}_{a''}\ S^{b''}{}_j\ A^{a'}{}_{b'}B^{a''}{}_{b''} \ = R^i{}_{a'}\ (S^{b'}{}_k\ R^k{}_{a''})\ S^{b''}{}_j\ A^{a'}{}_{b'}B^{a''}{}_{b''} \\ &= R^i{}_{a'}\ (SR)^{b'}{}_{a''}\ S^{b''}{}_j\ A^{a'}{}_{b'}B^{a''}{}_{b''} = R^i{}_{a'}\ \delta^{b'}{}_{a''}\ S^{b''}{}_j\ A^{a'}{}_{b'}B^{a''}{}_{b''} \\ &= R^i{}_{a'}\ S^{b''}{}_j\ A^{a'}{}_{b'}B^{b'}{}_{b''} = R^i{}_{a'}\ S^{b''}{}_j\ (AB)^{a'}{}_{b''} = R^i{}_{a'}\ S^{b''}{}_j\ C^{a'}{}_{b''} \\ &= C'^i{}_j \,. \qquad \text{// using (7.5.8) a third time with M = C} \end{aligned} \tag{7.8.7}$$

Thus we have shown that AB = C $\Rightarrow$ A'B' = C' so our down-tilt matrix equation is covariant.

If we define C ≡ AB where A and B are down-tilt mixed rank-2 tensors, then C will be a rank-2 down-tilt tensor providing we can show that $(A'B')'^{a}{}_{b} = R^{a}{}_{a'}\ S^{b'}{}_{b}\,(AB)^{a'}{}_{b} = R^{a}{}_{a'}\ S^{b'}{}_{b}\,C^{a'}{}_{b}$. But this is just what was shown above (albeit with different indices), so yes, C is also a down-tilt mixed rank-2 tensor.

One way to clarify the intention of C = AB is to write the matrix equation as $C_{dt} = A_{dt}B_{dt}$ where the notation A_{dt} means the down-tilt mixed rank-2 tensor having components $A^{i}{}_{j}$. In other words,

$$C^{i}{}_{j} = A^{i}{}_{k}B^{k}{}_{j} \qquad \Leftrightarrow \qquad (C_{dt})_{ij} = (A_{dt})_{ik}(B_{dt})_{kj} \qquad (7.8.8)$$

where the matrices $C_{dt},\ A_{dt}$ and B_{dt} can be regarded as ordinary linear algebra matrices. For example, the identity det(XY) = det(X)det(Y) then says

$$\det(C_{dt}) = \det(A_{dt})\det(B_{dt}) \qquad \text{or} \qquad \det(C^{*}{}_{*}) = \det(A^{*}{}_{*})\det(B^{*}{}_{*}) \qquad (7.8.9)$$

where we use wildcards to indicate the position of the indices of the elements which appear in the matrix whose determinant we are computing.

It is easy to show that the conclusions reached above apply similarly to an all *up-tilt* matrix equation

$$C_{i}{}^{j} = A_{i}{}^{k}\ B_{k}{}^{j}\ . \qquad // \text{ implied sum on k} \qquad (7.8.10)$$

We thus arrive at:

Matrix Rule #2. In Standard Notation, it is reasonable to use matrix notation in the following two situations involving mixed rank-2 tensors:

$$C^{i}{}_{j} = A^{i}{}_{k}B^{k}{}_{j} \qquad C_{dt} = A_{dt}B_{dt} \qquad dt = \text{down-tilt}$$
$$C_{i}{}^{j} = A_{i}{}^{k}B_{k}{}^{j} \qquad C_{ut} = A_{ut}B_{ut} \qquad ut = \text{up-tilt} \qquad (7.8.11)$$

In special relativity the down-tilt matrix form is most often used, and one just writes C = AB without bothering with the dt clarifying subscripts. This is consistent with the usual statement $x'^{\mu} = \Lambda^{\mu}{}_{\nu}x^{\nu}$ to describe a Lorentz transformation acting on the contravariant vector x^{ν} (where Λ = our R).

Note: In the matrix equation $C^{i}{}_{j} = A^{i}{}_{k}B^{k}{}_{j}$ containing "down-tilt" mixed rank-2 tensors, the k index sum "tilts up".

7.9 Matrix Inverse, Transpose and Determinant

Matrix Inverses

Consider the standard notation matrix equation AB=1 where (assuming det(A) ≠ 1) we can write $B = A^{-1}$. As demonstrated above, AB = 1 can only be a covariant equation if A and B are both down-tilt or both up-tilt mixed rank-2 tensors. Then we are talking about either $A_{dt}B_{dt} = 1_{dt}$ or $A_{ut}B_{ut} = 1_{ut}$, and the corresponding $B_{dt} = (A^{-1})_{dt}$ and $B_{ut} = (A^{-1})_{ut}$, all these being matrix equations . Thus

$$A_{dt}B_{dt} = 1_{dt} \qquad B_{dt} = (A^{-1})_{dt} \qquad A^i{}_k B^k{}_j = \delta^i{}_j \qquad B^i{}_j = (A^{-1})^i{}_j$$

$$A_{ut}B_{ut} = 1_{ut} \qquad B_{ut} = (A^{-1})_{ut} \qquad A_i{}^k B_k{}^j = \delta_i{}^j \qquad B_i{}^j = (A^{-1})_i{}^j \ . \qquad (7.9.1)$$

In this context, we have shown that if A is a mixed rank-2 tensor, then (A^{-1}) is a mixed rank-2 tensor as well, assuming it exists.

In (7.5.11) we considered $S = R^{-1}$ and $R = S^{-1}$ and reached the conclusions shown in (7.9.1) for the cases A = S and R = B. It happens that in this special case, S and R are not tensors, but the results are still valid.

It was shown in (7.4.19) that 1_{dt} is really g_{dt}, the down-tilt form of the metric tensor g, and similarly for 1_{ut} :

$$1_{dt} = g_{dt} \qquad (1)^i{}_j = g^i{}_j = \delta^i{}_j = \delta_{i,j}$$

$$1_{ut} = g_{ut} \qquad (1)_i{}^j = g_i{}^j = \delta_i{}^j = \delta_{i,j} \ . \qquad (7.9.2)$$

Thus first equation in (7.9.1) can be written in covariant form $A_{dt}B_{dt} = g_{dt}$ where all three matrices are down-tilt mixed rank-2 tensors. And this is also true for $A_{ut}B_{ut} = g_{ut}$.

Transpose Matrices

It is possible to define the "covariant transpose" M^{T} (italic *T*) of a matrix M in Standard Notation, and things work out as follows, where M is any rank-2 tensor,

$$\begin{array}{lll} (M^{T})^{ab} = M^{ba} & (R^{T})^{ab} = R^{ba} & (S^{T})^{ab} = S^{ba} \\ (M^{T})^a{}_b = M_b{}^a & (R^{T})^a{}_b = R_b{}^a & (S^{T})^a{}_b = S_b{}^a \\ (M^{T})_a{}^b = M^b{}_a & (R^{T})_a{}^b = R^b{}_a & (S^{T})_a{}^b = S^b{}_a \\ (M^{T})_{ab} = M_{ba} & (R^{T})_{ab} = R_{ba} & (S^{T})_{ab} = S_{ba} \ . \end{array} \qquad \text{// covariant transpose} \qquad (7.9.3)$$

The general rule is that, in each equation, the indices are reflected in a vertical line separating them.

The notation is "covariant" in that one can raise and lower indices at will in each equation and thereby create a new valid equation. The drawback of this notation is that, when the indices are tilted, M^{T} in general differs from what we might call the "matrix transpose" M^{T} of a matrix M. When one transposes a matrix, one normally means to swap the rows and columns, and that means to swap the indices. One would then write

$$\begin{array}{lll} (M^{\mathrm{T}})^{ab} = M^{ba} & (R^{\mathrm{T}})^{ab} = R^{ba} & (S^{\mathrm{T}})^{ab} = S^{ba} \\ (M^{\mathrm{T}})^a{}_b = M^b{}_a & (R^{\mathrm{T}})^a{}_b = R^b{}_a & (S^{\mathrm{T}})^a{}_b = S^b{}_a \\ (M^{\mathrm{T}})_a{}^b = M_b{}^a & (R^{\mathrm{T}})_a{}^b = R_b{}^a & (S^{\mathrm{T}})_a{}^b = S_b{}^a \\ (M^{\mathrm{T}})_{ab} = M_{ba} & (R^{\mathrm{T}})_{ab} = R_{ba} & (S^{\mathrm{T}})_{ab} = S_{ba} \ . \end{array} \qquad \text{// matrix transpose} \qquad (7.9.4)$$

The middle two lines (tilted indices) are very non-covariant in form since for example $(M^{\mathbf{T}})^{a}{}_{b} = M^{b}{}_{a}$ has the up and down sense of index a and b not even matching. The *use* of the $M^{\mathbf{T}}$ matrix transpose is in theorems like $\det(M) = \det(M^{\mathbf{T}})$ where rows and columns are swapped without changing the determinant.

Comments:

1. The issue here is that one might have

$$(M^{\boldsymbol{T}})^{a}{}_{b} = M_{b}{}^{a} \neq (M^{\mathbf{T}})^{a}{}_{b} = M^{b}{}_{a} \quad .$$

If M is an x-space tensor and if the x-space metric tensor is g = 1, then up and down index positions don't matter and our problem goes away,

$$(M^{\boldsymbol{T}})_{ab} = M_{ba} = (M^{\mathbf{T}})_{ab} = M_{ba}$$

so then $M^{\boldsymbol{T}} = M^{\mathbf{T}}$. (7.9.5)

2. For R and S: As shown in (7.5.9), g' acts on the first index of R while g acts on the second index. Conversely as shown in (7.5.10), g acts on the first index of S while g' acts on the second index. So up and down indices are the same for R and S only if both g = 1 and g' = 1. Only in this special case, which implies a (local) rotation for $\mathbf{x}' = \mathbf{F}(\mathbf{x})$, does one have $R^{\boldsymbol{T}} = R^{\mathbf{T}}$ and $S^{\boldsymbol{T}} = S^{\mathbf{T}}$. This situation arises in Appendix E where we have $\mathbf{x}'' = \mathbf{F}_{\mathbf{M}}(\mathbf{x})$ and R and S (there called M and N) are rotations. (7.9.6)

3. As it turns out, regardless of g and g', one has

$$\det(M) = \det(M^{\mathbf{T}}) = \det(M^{\boldsymbol{T}}) \tag{7.9.7}$$

for any index positions on the matrices. The traditional fact that $\det(M) = \det(M^{\mathbf{T}})$ is true for any kind of square matrix, while the fact that $\det(M) = \det(M^{\boldsymbol{T}})$ is proven in (D.12.20). Since *T* and T are the same when indices are both up or both down, the interesting case is when indices are tilted. Here is an outline,

$$\det(M^{\boldsymbol{T}}) = \det([M^{\boldsymbol{T}}]^{i}{}_{j}) = \det(M_{j}{}^{i}) = \det(g_{jj'}M^{j'}{}_{i'}g^{i'i}) = \det(g_{dn}M_{dt}g_{up})$$

$$= \det(g_{dn})\det(M_{dt})\det(g_{up}) = g \det(M_{dt})\,(1/g) = \det(M_{dt}) = \det(M^{i}{}_{j}) = \det(M) \quad .$$

4. As shown in the next few paragraphs, in either up-tilt or down-tilt notations one can write,

$$RR^{\boldsymbol{T}} = R^{\boldsymbol{T}}R = 1 \qquad SS^{\boldsymbol{T}} = S^{\boldsymbol{T}}S = 1 \qquad RS = SR = 1$$

$$R^{\boldsymbol{T}} = R^{-1} = S \qquad\qquad S^{\boldsymbol{T}} = S^{-1} = R \quad . \tag{7.9.8}$$

Thus, both the R matrix and S matrix are "covariant real-orthogonal" for any underlying transformations $\mathbf{x}' = \mathbf{F}(\mathbf{x})$, even when R and S are not rotation matrices. Since the matrix elements are real, R and S are also "covariant unitary".

Restatement of Orthgonality Conditions

We shall generally avoid using the covariant transpose notation since it leads to results which can be confusing although correct, but there will be times when we use it. To restate the potential confusion, in developmental notation the statement $RR^{\mathrm{T}} = 1$ (or $R^{\mathrm{T}} = R^{-1}$, meaning "real orthogonal") implies that R is a "rotation", where we include in this term the possibility of axis reflections. But in Standard Notation, $RR^{\mathit{T}} = 1$ is valid for *any* matrix R associated with a general transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$, rotations and non-rotations alike. To see why this is so, just write out orthogonality rules from (7.6.4) :

$$\#4 \quad R_a{}^b R^c{}_b = \delta_a{}^c \quad \Rightarrow \quad R_a{}^b (R^{\mathit{T}})_b{}^c = \delta_a{}^c \quad \Rightarrow \quad RR^{\mathit{T}} = 1 \ \text{(up-tilt)} .$$

$$\#3 \quad R^a{}_b R_c{}^b = \delta^a{}_c \quad \Rightarrow \quad R^a{}_b (R^{\mathit{T}})^b{}_c = \delta^a{}_c \quad \Rightarrow \quad RR^{\mathit{T}} = 1 \ \text{(down-tilt)} .$$

$$\#2 \quad R^b{}_a R_b{}^c = \delta_a{}^c \quad \Rightarrow \quad (R^{\mathit{T}})_a{}^b R_b{}^c = \delta_a{}^c \quad \Rightarrow \quad R^{\mathit{T}}R = 1 \ \text{(up-tilt)} .$$

$$\#1 \quad R_b{}^a R^b{}_c = \delta^a{}_c \quad \Rightarrow \quad (R^{\mathit{T}})^a{}_b R^b{}_c = \delta^a{}_c \quad \Rightarrow \quad R^{\mathit{T}}R = 1 \ \text{(down-tilt)} . \qquad (7.9.9)$$

The equations on the right are just restatements of the orthogonality rules. The same equations are valid with R→ S, for example

$$RR^{\mathit{T}} = 1 \Rightarrow (RR^{\mathit{T}})^{-1} = 1 \Rightarrow (R^{\mathit{T}})^{-1}R^{-1} = 1 \Rightarrow (R^{-1})^{\mathit{T}}R^{-1} = 1 \Rightarrow S^{\mathit{T}}S = 1 .$$

To totally distinguish the above situations, we could use this rather unpleasant notation,

$[RR^{\mathrm{T}} = 1]_{\mathrm{DN}} \Leftrightarrow$ R is a "rotation" DN = Developmental Notation

$[RR^{\mathit{T}} = 1]_{\mathrm{SN,ut}}$ for *any* R associated with F SN,ut = Standard Notation, up-tilt

$[RR^{\mathit{T}} = 1]_{\mathrm{SN,dt}}$ for *any* R associated with F SN,dt = Standard Notation, down-tilt (7.9.10)

But such notations are not necessary if one understands that an equation like $RR^{\mathit{T}} = 1$ implies all up-tilt or all down-tilt matrices.

How to recognize a rotation

How then does one recognize a rotation matrix directly in Standard Notation? We translate using the rule $R_{ik} \rightarrow R^i{}_k$ shown in (7.5.2),

$$[RR^{\mathrm{T}} = 1]_{\mathrm{DN}} \Rightarrow R_{ik}R_{jk} = \delta_{i,j} \quad \rightarrow \quad R^i{}_k R^j{}_k = \delta_{i,j} .$$

A simpler alternate form for recognizing a rotation can be obtained as follows,

$$R^i{}_k R^j{}_k = \delta_{i,j} \qquad \text{// now apply } S^a{}_i \text{ to both sides}$$

$$S^a{}_i R^i{}_k R^j{}_k = S^a{}_i \delta_{i,j}$$

$$(SR)^a{}_k R^j{}_k = S^a{}_j \qquad \text{// but SR = 1}$$

$$\delta^a{}_k R^j{}_k = S^a{}_j = R_j{}^a \qquad \text{// right side from (7.5.13)}$$

$$R^j{}_a = R_j{}^a \qquad \text{// if R is a rotation}$$

Therefore, in Standard Notation a "rotation" can be identified in either of these two ways,

$$R^i{}_k R^j{}_k = \delta_{i,j} \qquad \text{or} \qquad R^j{}_a = R_j{}^a \qquad \text{// "rotation" .} \tag{7.9.11}$$

Matrix form of the transformation rule for a rank-2 tensor

Instead of trying to use this covariant transpose *T* superscript in the standard notation, we can get rid of T in the developmental notation and *then* do our translation to standard notation. For example we start in developmental notation,

$$M' = R\,M\,R^T \qquad \Rightarrow \qquad M'_{ad} = R_{ab}M_{bc}(R^T)_{cd} = R_{ab}M_{bc}R_{dc} = R_{ab}R_{dc}M_{bc} \tag{7.9.12}$$

Then we make the conversion using (7.5.2) $R_{ik} \rightarrow R^i{}_k$ and (7.4.1) $M_{ab} \rightarrow M^{ab}$:

$$M'_{ad} = R_{ab}R_{dc}M_{bc} \qquad \rightarrow \qquad (M')^{ad} = R^a{}_b R^d{}_c M^{bc} \tag{7.9.13}$$

and this then is the Standard Notation rule for the way a contravariant rank-2 tensor transforms, as was shown in the first line of (7.5.8).

However, notice that, where we make use of the "tilt reversal theorem" (7.4.15),

$$(M')^{ad} = R^a{}_b R^d{}_c M^{bc} \qquad \text{// covariant, so can lower index d on both sides}$$

so

$$(M')^a{}_d = R^a{}_b R_{dc} M^{bc} = R^a{}_b R_d{}^c M^b{}_c = R^a{}_b M^b{}_c R_d{}^c = R^a{}_b M^b{}_c (R^{\boldsymbol{T}})^c{}_d = (RMR^{\boldsymbol{T}})^a{}_d\,. \tag{7.9.14}$$

Thus we can write $M' = RMR^{\boldsymbol{T}}$ in standard notation provided we assume that we are using all down-tilt matrices. The same equation results if all matrices are up-tilt:

$$(M')^{ad} = R^a{}_b R^d{}_c M^{bc}$$

so

$$(M')_a{}^d = R_{ab} R^d{}_c M^{bc} = R_{ab} M^{bc} R^d{}_c = R_a{}^b M_b{}^c (R^{\boldsymbol{T}})_c{}^d = (RMR^{\boldsymbol{T}})_a{}^b \tag{7.9.15}$$

We then arrive at this conclusion:

Fact: The transformation rule for a rank-2 tensor M can be written in matrix notation in the Standard Notation as

$$M' = RMR^{\boldsymbol{T}} \qquad \text{all up-tilt or all down-tilt} \qquad (7.9.16)$$

provided it is understood that all matrices are up-tilt or all are down-tilt. The result is very similar to what it was in Developmental Notation: $M' = RMR^{\mathbf{T}}$. The use of the covariant transpose thus causes this equation to maintain its form when we go to Standard Notation.

Transpose in Dirac Notation

In the Dirac notation dot products are represented as $\mathbf{a} \bullet \mathbf{b} = \langle\mathbf{a} \mid \mathbf{b}\rangle = \mathbf{b} \bullet \mathbf{a} = \langle\mathbf{b} \mid \mathbf{a}\rangle$. In general Dirac theory (as used in quantum mechanics) one has $\langle\mathbf{a}|\mathbf{b}\rangle = \langle\mathbf{b}|\mathbf{a}\rangle^*$, but all our scalar products are real so we can ignore complex conjugation. Consider then,

$$\langle\mathbf{a} \mid M \mid \mathbf{b}\rangle \equiv \langle\mathbf{a} \mid M\,\mathbf{b}\rangle = \mathbf{a} \bullet (M\mathbf{b}) = a^i(M\mathbf{b})_i = a^i[\,M_i{}^j b_j] = a^i\,M_i{}^j\,b_j$$

$$= b_j\,M_i{}^j\,a^i = b_j\,(M^{\boldsymbol{T}})^j{}_i\,a^i = b_j\,[M^{\boldsymbol{T}}\mathbf{a}]^j = \mathbf{b} \bullet (M^{\boldsymbol{T}}\mathbf{a}) = \langle\mathbf{b}|\,M^{\boldsymbol{T}}\mathbf{a}\rangle = \langle\mathbf{b}|\,M^{\boldsymbol{T}}\,|\mathbf{a}\rangle\,. \qquad (7.9.17)$$

This proves:

Fact: In Dirac notation $\langle\mathbf{a} \mid M \mid \mathbf{b}\rangle = \langle\mathbf{b} \mid M^{\boldsymbol{T}} \mid \mathbf{a}\rangle$ for any rank-2 tensor M, where *T* is the covariant transpose. (7.9.18)

If **a** and **b** are vectors in Cartesian space with g = 1, then $\langle\mathbf{a} \mid M \mid \mathbf{b}\rangle = \langle\mathbf{b}|\,M^{\mathbf{T}}\,|\mathbf{a}\rangle$ since then $M^{\boldsymbol{T}} = M^{\mathbf{T}}$.

A fuller description of Dirac notation is presented in (E.7.4).

Summary of Covariant Transpose Equations

We summarize the above facts which relate to the covariant transpose in Standard Notation,

$\det(M) = \det(M^{\boldsymbol{T}})$	determinant matrix transpose theorem *(7.9.7)*
$RS = SR = 1 \quad \text{so } R^{-1} = S \text{ and } S^{-1} = R$	relation between S and R from their definition *(7.5.11)*
$RR^{\boldsymbol{T}} = R^{\boldsymbol{T}}R = 1 \Rightarrow R^{\boldsymbol{T}} = R^{-1} = S$	orthogonality rules (R is covariant real orthog.) *(7.9.8)*
$SS^{\boldsymbol{T}} = S^{\boldsymbol{T}}S = 11 \Rightarrow S^{\boldsymbol{T}} = S^{-1} = R$	orthogonality rules (S is covariant real orthog.) *(7.9.8)*
$M' = RMR^{\boldsymbol{T}}$	how a (tilted) rank-2 tensor transforms *(7.9.16)*
$\langle\mathbf{a} \mid M \mid \mathbf{b}\rangle = \langle\mathbf{b} \mid M^{\boldsymbol{T}} \mid \mathbf{a}\rangle$	Dirac matrix elements of operator M *(7.9.17)* (7.9.19)

In these equations $M^{\boldsymbol{T}}$ is the covariant transpose shown in (7.9.3). It is assumed that all matrix products involve matrices which are either all down-tilt or all up-tilt. The various orthogonality rules like $RR^{\boldsymbol{T}} = 1$ are valid for any R matrix, not just for a rotation.

Determinant of a Matrix

In developmental notation one writes

$$\det(A) = \varepsilon_{abc...}\ A_{1a}A_{2b}A_{3c}..... \quad = \varepsilon_{abc...}\ A_{a1}\ A_{b2}A_{c3}..... \tag{7.9.20}$$

where the A_{ij} are components of the contravariant rank-2 tensor A and where ε is the permutation tensor discussed above in Section 7.7.

We have argued above that the notion of a rank-2 tensor being a matrix in Standard Notation is only viable for mixed rank-2 tensor of either the down-tilt or up-tilt variety. Thus, the matrix determinants of interest in Standard Notation would be these:

$$\det(A_{dt}) = \det(A^*{}_*) = \varepsilon_{abc...}\ A^1{}_a\ A^2{}_bA^3{}_c..... \quad = \varepsilon_{abc...}\ A^a{}_1\ A^b{}_2A^c{}_3..... \tag{7.9.21}$$

$$\det(A_{ut}) = \det(A_*{}^*) = \varepsilon_{abc...}\ A_1{}^a\ A_2{}^bA_3{}^c..... \quad = \varepsilon_{abc...}\ A_a{}^1\ A_b{}^2A_c{}^3..... \tag{7.9.22}$$

Determinants of a rank-2 tensor A rarely appear in this document, but they do appear for the non-tensor objects R and S as shown in (7.5.17), as needed for the Jacobian in (5.12.6).

For more information on determinants of rank-2 tensors see Section D.12.

7.10 Tensors of Rank n, direct products, Lie groups, symmetry and Ricci-Levi-Civita

The most general **tensor of rank n** (aka **order n**) will have some number s of contravariant indices and then some number n-s of covariant indices. If s = n, the tensor is pure contravariant, and if s = 0, it is pure covariant, otherwise it is "mixed" (as opposed to "pure"). The transformation of the tensor under F will show a factor $R^a{}_{a'}$ for each contravariant index, and a factor $S^{a'}{}_a$ ($= R_a{}^{a'}$ by (7.5.13)) for each covariant index, as illustrated by this example :

$$\begin{aligned} T'^{abc}{}_{de} &= R^a{}_{a'}\ R^b{}_{b'}\ R^c{}_{c'}\ S^{d'}{}_d\, S^{e'}{}_e\, T^{a'b'c'}{}_{d'e'} \\ T'^{abc}{}_{de} &= R^a{}_{a'}\ R^b{}_{b'}\ R^c{}_{c'}\ R_d{}^{d'}\ R_e{}^{e'}\ T^{a'b'c'}{}_{d'e'} \quad . \end{aligned} \tag{7.10.1}$$

Note that for $R^a{}_{a'}$ and $R_a{}^{a'}$ the second index is the summation index, but for $S^{a'}{}_a$ it is the first index.

A rank-n tensor always transforms the way an outer product of n vectors transforms if those vectors have indices which type-match those of the tensor. In the above case, an object that would transform the same as $T^{abc}{}_{de}$ would be

$$A^aB^bC^cD_dE_e\ . \tag{7.10.2}$$

A tensor of rank-n has 2^n tensor objects in its family since each index can be up or down. For example, the tensor T above is one of $2^5 = 32$ tensors one can form. Of these, one is pure covariant and one is pure contravariant and 30 are mixed.

If any of these tensors is a tensor field, such as $T^{abc}{}_{de}(\mathbf{x})$, then of course all family members are tensor fields.

Direct Products. Consider again the outer product of vectors $A^aB^bC^cD_dE_e$. The transformation $A'^a = R^a{}_{a'}A^{a'}$ occurs in an N-dimensional contravariant vector space we shall call $\mathcal{R}$. In this space one could establish a set of basis vectors, and of course there are rules for adding vectors and so on. Transformation $B'^a = R^a{}_cB^c$ occurs in an identical copy of the space $\mathcal{R}$, but transformation $D'_d = S^{d'}{}_dD_{d'} = R_d{}^{d'}D_{d'}$ occurs in a covariant version of $\mathcal{R}$ we call $\overline{\mathcal{R}}$. Since dot products (inner products) have been established for vectors in these spaces, they can be regarded as full blown Hilbert Spaces with the caveats of Section 5.10.

The transformation of the outer product object, as already noted, is given by

$$A'^aB'^bC'^cD'_dE'_e = R^a{}_{a'}\, R^b{}_{b'}\, R^c{}_{c'}\, R_d{}^{d'}\, R_e{}^{e'}\, A^{a'}B^{b'}C^{c'}D_{d'}E_{e'} \tag{7.10.3}$$

and one can consider the operator $R^a{}_{a'}\, R^b{}_{b'}\, R^c{}_{c'}\, R_d{}^{d'}\, R_e{}^{e'}$ as a transformation element in a so-called direct product space which in this case would be written

$$\mathcal{R}_{dp} = \mathcal{R}\otimes\mathcal{R}\otimes\mathcal{R}\otimes\overline{\mathcal{R}}\otimes\overline{\mathcal{R}} \tag{7.10.4}$$

One could then define

$$(\mathcal{R}_{dp})^{abc}{}_{de}\,;\,{}_{a'b'c'}{}^{d'e'} \equiv R^a{}_{a'}\, R^b{}_{b'}\, R^c{}_{c'}\, R_d{}^{d'}\, R_e{}^{e'} \tag{7.10.5}$$

so that

$$A'^aB'^bC'^cD'_dE'_e = (\mathcal{R}_{dp})^{abc}{}_{de}\,;\,{}_{a'b'c'}{}^{d'e'}\, A^{a'}B^{b'}C^{c'}D_{d'}E_{e'} \tag{7.10.6}$$

and of course this would apply to any tensor of the same index configuration, such as

$$T'^{abc}{}_{de} = \mathcal{R}_{dp}{}^{abc}{}_{de}\,;\,{}_{a'b'c'}{}^{d'e'}\, T^{a'b'c'}{}_{d'e'} \quad . \tag{7.10.7}$$

This suggests a definition of "tensor" as follows" : tensors are those objects that are transformed by all possible direct product representations formable from the two fundamental vector representations $\mathcal{R}$ and $\overline{\mathcal{R}}$. To this set of spaces one would add the identity space ***1*** to handle tensorial scalars.

Appendix E continues this direct product discussion in terms of the basis vectors that form a complete set for a direct product space such as $\mathcal{R}_{dp}$ and shows how to expand tensors on such bases.

Lie Groups. The direct product notion is just a formalism, but the formalism has some implications when the space $\mathcal{R}$ is associated with a "representation" of a Lie group. In this case, a direct product $\mathcal{R}_{dp} = \mathcal{R} \otimes \mathcal{R}$ can be written as a sum of "irreducible" representations of that group. What this means is that the transformation elements of $\mathcal{R}_{dp}$ and the objects T^{ab} can be shuffled around with linear combinations so that $(\mathcal{R}_{dp})^{ab}{}_{a'b'}$, when thought of as a matrix with columns labeled by N^2 ab possibilities and rows labeled by the N^2 a'b' possibilities, appears in "block diagonal form" with all zeros outside the blocks. In this case, the shuffled components of tensor T^{ab} can be regarded as a non-interacting assembly of pieces each of which transforms according to one of those blocks of the shuffled $(\mathcal{R}_{dp})^{ab}{}_{a'b'}$.

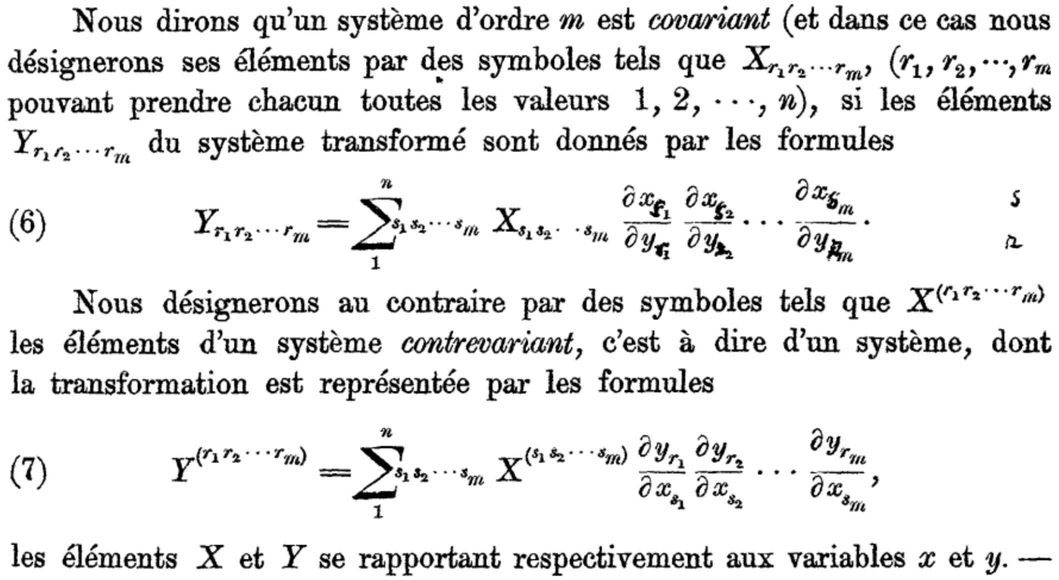
The most famous example occurs with N=3 and the rotation group SU(2) in which case $\mathcal{R}^{(1)} \equiv \mathcal{R}$ can be decomposed according to $\mathcal{R}^{(1)} \otimes \mathcal{R}^{(1)} = \mathcal{R}^{(2)} \oplus \mathcal{R}^{(1)} \oplus \mathcal{R}^{(0)}$ where the ⊕ symbols indicate this block diagonal form. In this case the blocks are 5x5, 3x3 and 1x1, fitting onto the diagonal of the 9x9 matrix area. The numbers L = 0,1,2 here label the rotation group representations and that label is associated with angular momentum. The elements of the 5x5 block are called $D^{(2)}{}_{M,M'}(\varphi,\theta,\psi)$ where M,M' = 2,1,0,-1.-2, and where φ,θ,ψ are the "Euler angles" which serve to label a particular rotation. This $D^{(2)}$ object *is* the L=2 matrix representation of the rotation group. Taking two vectors **A** and **B**, one can identify **A•B** as the combination transforming according to $\mathcal{R}^{(0)}$ ("scalar") and **A**x**B** (linearly combined) as that transforming as $\mathcal{R}^{(1)}$ ("vector"). The traceless symmetric matrix $A_iB_j - \delta_{i,j}\mathbf{A}\bullet\mathbf{B}$ has 5 independent elements associated with $\mathcal{R}^{(2)}$ ("quadrupole").

This whole reduction idea can be applied to larger direct products such as $\mathcal{R}^{(1)} \otimes \mathcal{R}^{(1)} \otimes \mathcal{R}^{(1)}$ and tensor components T^{abc}.

The Standard Model of elementary particle physics is chock full of direct products of this nature, where the idea of rotational symmetry is extended to other kinds of "internal" symmetry, spin and isospin being two examples. Representations of the Lie symmetry group SU(3) are associated with quarks which are among of the fundamental building blocks of the Standard Model.

The group discussion above can be applied generally to quantum physics. The basic idea is that if "the physics" (the Hamiltonian or Lagrangian) describing some quantum object is invariant under a certain symmetry group (such as rotational symmetry or perhaps some discrete crystal symmetry), then the quantum states of that object can be classified according to the representations of that group. The Bohr hydrogen atom "physics" $H \sim \nabla^2 - 1/|\mathbf{r}|$ has perfect rotation group symmetry and is also symmetric about the axis (angle ψ) from center to electron (no spin). The representation functions then must have M' = 0, and then $D^{(L)}{}_{M,0}(\varphi,\theta,\psi) \sim Y_{LM}(\theta,\varphi)$, the famous spherical harmonics that describe the "orbitals" which have mystified first-year chemistry students for the last 100 years.

<u>Historical Note:</u> Ricci and Levi-Civita (see Refs) referred to rank-n tensors as "systems of order n" and did not include mixed tensors in their 1900 paper. Nor did they use the Einstein summation convention, since Einstein thought of that later on. They *did* use the up and down index notation pretty much as it is used today, though the up indices are enclosed in parenthesis. Here is a direct quote from the paper where the nature of the contravariant and covariant tensors is described (with crude translation below for non-French readers). Equation (6) had typos which some thoughtful reader corrected: the y subscripts should be r's and the x subscripts should be s's. In our notation $\partial x_s/\partial y_r \rightarrow \partial x^s/\partial x'^r = S^s{}_r = R_r{}^s$.



Nous dirons qu'un système d'ordre *m* est *covariant* (et dans ce cas nous désignerons ses éléments par des symboles tels que $X_{r_1 r_2 \cdots r_m}$, ($r_1, r_2, \cdots, r_m$ pouvant prendre chacun toutes les valeurs $1, 2, \cdots, n$), si les éléments $Y_{r_1 r_2 \cdots r_m}$ du système transformé sont donnés par les formules

$$(6) \qquad Y_{r_1 r_2 \cdots r_m} = \sum_{1}^{n}{}_{s_1 s_2 \cdots s_m} X_{s_1 s_2 \cdots s_m} \frac{\partial x_{s_1}}{\partial y_{r_1}} \frac{\partial x_{s_2}}{\partial y_{r_2}} \cdots \frac{\partial x_{s_m}}{\partial y_{r_m}}.$$

s

n

Nous désignerons au contraire par des symboles tels que $X^{(r_1 r_2 \cdots r_m)}$ les éléments d'un système *contrevariant*, c'est à dire d'un système, dont la transformation est représentée par les formules

$$(7) \qquad Y^{(r_1 r_2 \cdots r_m)} = \sum_{1}^{n}{}_{s_1 s_2 \cdots s_m} X^{(s_1 s_2 \cdots s_m)} \frac{\partial y_{r_1}}{\partial x_{s_1}} \frac{\partial y_{r_2}}{\partial x_{s_2}} \cdots \frac{\partial y_{r_m}}{\partial x_{s_m}},$$

les éléments X et Y se rapportant respectivement aux variables x et y. —

(7.10.8)

We will say that a system of order m is covariant (and in this case we will designate its elements by the symbol $X_{r1,r2....}$) (r_1,r_2.... can each take all the values 1...n), if the elements $Y_{r1,r2....}$ of the transformed system are given by the formulas (6) .

We will designate on the contrary by the symbols $X^{(r1,r2....)}$ the elements of a contravariant system, which is to say of a system where the transformation is represented by the formulas (7) ,

the elements X and Y being related respectively to (presumably "are functions of") the variables x and y.

Their y is our **x**', and their n is our N. Notice that the indices on the coordinates themselves are taken down, contrary to current usage. They do not explain why the words contravariant and covariant are used.

In our notation, the two equations above would be written

$$M'_{abc...} = R_a{}^{a'} R_b{}^{b'} .... M_{a'b'c'...} = \Sigma_{primed} M_{a'b'c'...} R_a{}^{a'} R_b{}^{b'} .... \qquad (6)$$

$$M'^{abc...} = R^a{}_{a'} R^b{}_{b'} .... M^{a'b'c'...} = \Sigma_{primed} M^{a'b'c'...} R^a{}_{a'} R^b{}_{b'}{}' .... \qquad (7) \qquad (7.10.9)$$

7.11 The Contraction Tilt-Reversal Rule

In some complicated combination of multiple tensors, imagine there is somewhere a pair of summed indices where one is up and the other is down. As noted above (7.4.15), such a sum is called a **contraction**. The contracted indices could be on the same object or they could be on different objects. We depict this situation with the following symbolic notation,

$$[\text{-----}^{a}\text{---------}_{a}\text{----}] \qquad (7.11.1)$$

where the dashes indicate indices that we don't care about and which won't change -- each one could be up or down. We know we can reverse the tilt this way, as in (7.4.11),

$$[\text{-----}^{a}\text{---------}_{a}\text{----}] = g^{ab}\, g_{ac}\, [\text{-----}_{b}\text{---------}^{c}\text{----}] \qquad (7.11.2)$$

where the first g raises the index b to a, and the second g lowers the index c to a. But the two g's are inverses as in (7.4.1), $g^{ab}\, g_{ac} = g^{ba}\, g_{ac} = \delta_{b,c}$, which at once gives the desired result (index b→a)

$$[\text{-----}^{a}\text{---------}_{a}\text{----}] = [\text{-----}_{a}\text{---------}^{a}\text{----}] \ . \qquad \text{// the Contraction Tilt-Reversal Rule} \qquad (7.11.3)$$

A notable example of course is this:

$$A^{a}Y_{a} = A_{a}Y^{a} \quad = \quad \text{" } \mathbf{A} \bullet \mathbf{Y} \text{ "} \qquad \text{// or perhaps " A.Y " as noted in Section 5.10} \qquad (7.11.4)$$

<u>When is index tilt-reversal allowed and when is it not allowed?</u>

It is always allowed when both indices of the tilted contraction are valid tensor indices. Consider these four examples to be discussed below:

$$R_{a}{}^{b}A_{b} = R_{ab}A^{b} \qquad \text{Proof: } R_{a}{}^{b}A_{b} = R_{ac}g^{cb}\, g_{bd}A^{d} = g^{cb}g_{bd}R_{ac}A^{d} = \delta^{c}{}_{d}R_{ac}A^{d} = R_{ac}A^{c}$$

$$A^{a}R_{a}{}^{b} \neq A_{a}R^{ab} \qquad \text{Proof: } A^{a}R_{a}{}^{b} = g^{ac}A_{c}\, R^{db}\, g'_{da} = g^{ac}\, g'_{ad}\, A_{c}\, R^{db} \neq A_{a}R^{ab}$$

$$A^{a}\partial_{a} = A_{a}\partial^{a} \qquad \text{Proof: } A^{a}\partial_{a}f = g^{ac}A_{c}\, g_{ad}\partial^{d}f = g^{ac}g_{ad}A_{c}\partial^{d}f = \delta^{c}{}_{d}A_{c}\partial^{d}f = A_{c}\partial^{c}f \qquad (7.11.5)$$

$$\partial_{a}A^{a} \neq \partial^{a}A_{a} \qquad \text{Proof: } \partial_{a}A^{a} = (g_{ac}\partial^{c})(g^{ad}A_{d}) = g_{ac}g^{ad}(\partial^{c}A_{d}) + g_{ac}(\partial^{c}g^{ad})A_{d}$$

$$= \delta_{c}{}^{d}(\partial^{c}A_{d}) + g_{ac}(\partial^{c}g^{ad})A_{d} = \partial^{c}A_{c} + (\partial_{a}g^{ad})A_{d} \neq \partial^{c}A_{c}$$

In the first example, since $R_{a}{}^{b}$ is not a tensor, one is on dangerous ground doing the tilt reversal, but it happens to work because the second index is associated with metric tensor g_{ij} which is the same metric tensor that raises and lowers indices of A_{b}. In the second example, the tilt-reversal fails because the first index of $R_{a}{}^{b}$ is associated with the x'-space metric tensor g'_{ij}. In the third example, both indices are valid tensor indices (with the same metric tensor).

The fourth example shows a failure of the tilt-reversal rule and this example is very important. The inequality becomes an equality only if the underlying transformation F is linear so that R and S and g are then constants independent of position. For general F, such as the F involved in curvilinear coordinate transformations, the object $\partial_{a}A^{b}$ is not a rank-2 tensor and so the object $\partial_{a}A^{a}$ does not represent contraction of two true tensor indices and therefore the "contraction tilt-reversal rule" does not apply. The neutralization rule of Section 7.12 below also does not apply for this same reason, so $\partial_{a}A^{a}$ does not transform as a scalar under general F.

Here is one more example along the lines of the fourth example above that shows the potential danger of reversing a tilt when it is not justified. In this example, we use notation to be introduced in Section 7.16 below. Consider the equation,

$$V^{a} = \varepsilon^{abc}B_{b;c} \qquad (1) \qquad \text{// valid}$$

where

B_b = a tensorial vector
$B_{b;c} = \partial_c B_b - \Gamma^n{}_{bc}B_n$ = the covariant derivative of vector B_b = a rank-2 tensor
ε^{abc} = a rank-3 tensor density (weight -1) (the Levi-Civita tensor)
V^a = a vector density (weight -1) (7.11.6)

With regard to $B_{b;c}$: (a) in comma notation one writes $\partial_c B_b = B_{b,c}$; (b) $\Gamma^n{}_{bc} = \Gamma^n{}_{cb}$.

Since all indices on equation (1) are tensor indices, one can lower index a and reverse the b and c tilts to get

$$V_a = \varepsilon_{abc}B^{b;c} \quad . \qquad (2) \qquad // \text{ valid} \qquad (7.11.7)$$

Now go back to equation (1). Because ε^{abc} is antisymmetric on b and c, whereas $\Gamma^n{}_{bc}$ is symmetric on b and c, one can write $\varepsilon^{abc}B_{b;c} = \varepsilon^{abc}B_{b,c}$ since the Γ term vanishes by symmetry. Thus one gets

$$V^a = \varepsilon^{abc}B_{b,c} \quad . \qquad (3) \qquad // \text{ valid} \qquad (7.11.8)$$

Were one to blindly lower a and reverse the b and c tilts, one would get

$$V_a = \varepsilon_{abc}B^{b,c} \quad . \qquad (4) \qquad // \text{ NOT valid} \qquad (7.11.9)$$

The reason for "not valid" is that the reversal of the b tilt is not justified. That is,

$$V^a = \varepsilon^{abc}B_{b,c} = \varepsilon^{abc}\partial_c B_b = g^{bb'}\,\varepsilon^a{}_{b'}{}^c\,\partial_c(g_{bb''}B^{b''})$$

$$\Rightarrow V_a = g^{bb'}\,\varepsilon_{ab'}{}^c\,\partial_c(g_{bb''}B^{b''}) = g^{bb'}\,\varepsilon_{ab'c}\,\partial^c(g_{bb''}B^{b''}) \qquad // \text{ c tilt reversal is OK}$$

$$= g^{bb'}g_{bb''}\,\varepsilon_{ab'c}(\partial^c B^{b''}) + g^{bb'}\,\varepsilon_{ab'c}\,(\partial^c g_{bb''})\,B^{b''}$$

$$= \delta^{b'}{}_{b''}\,\varepsilon_{ab'c}(\partial^c B^{b''}) + g^{bb'}\,\varepsilon_{ab'c}\,(\partial^c g_{bb''})\,B^{b''}$$

$$= \varepsilon_{abc}(\partial^c B^b) + g^{bb'}\,\varepsilon_{ab'c}\,(\partial^c g_{bb''})\,B^{b''}$$

$$= \varepsilon_{abc}B^{b,c} + g^{bb'}\,\varepsilon_{ab'c}\,(\partial^c g_{bb''})\,B^{b''} = \varepsilon_{abc}B^{b,c} + \text{ extra term!} \qquad (7.11.10)$$

It is due to this extra term that (4) is not valid. Basically this is the same as the fourth example above, but the situation is embedded in a more complicated environment (extra tensors, tensor densities, comma notation, covariant derivatives, other tilted indices, etc). One way to summarize the example is this:

$$(V^a = \varepsilon^{abc}B_{b,c}) \Leftrightarrow (V^a = \varepsilon^{abc}B_{b;c}) \Leftrightarrow (V_a = \varepsilon_{abc}B^{b;c}) \quad \not\Leftrightarrow \quad (V_a = \varepsilon_{abc}B^{b,c}) \qquad (7.11.11)$$

7.12 The Contraction Neutralization Rule

A contracted index pair plays no role in how a tensor object transforms; the two contracted indices neutralize each other, as we now show. (7.12.1)

First, recall that the indices on a general rank-n tensor (perhaps formed from several tensors) transform the same way that an outer product of n vectors transforms, where the vector index types match those of the tensor. The vectors transform this way,

$V'^{a} = R^{a}{}_{b}V^{b} \qquad\qquad V'_{a} = S^{b}{}_{a}V_{b}$. *(7.5.3) and (7.5.4)*

We take the same "big object" $[\text{-----}^{a}\text{---------}_{a}\text{----}]$ appearing in (7.11.1) and now ask how it transforms.

Below the X's represent either R or S factors for the dash indices (each of which might be up or down):

$$\begin{aligned}
[\text{-----}^{a}\text{---------}_{a}\text{----}]' &= \text{XXXXX}\ R^{a}{}_{b}\ \text{XXXXXXXXX}\ S^{c}{}_{a}\ \text{XXXX}\ [\text{-----}^{b}\text{---------}_{c}\text{----}] \\
&= S^{c}{}_{a}\ R^{a}{}_{b}\ \text{XXXXX}\ \ \text{XXXXXXXXX}\ \ \text{XXXX}\ [\text{-----}^{b}\text{---------}_{c}\text{----}] \\
&= \delta^{c}{}_{b}\ \text{XXXXX}\ \ \text{XXXXXXXXX}\ \ \text{XXXX}\ [\text{-----}^{b}\text{---------}_{c}\text{----}] \qquad // \text{RS} = 1 \\
&= \text{XXXXX}\ \ \text{XXXXXXXXX}\ \ \text{XXXX}\ [\text{-----}^{a}\text{---------}_{a}\text{----}] \qquad (7.12.2)
\end{aligned}$$

where now the only X's left are for the *other* indices. Again we look at our canonical example,

$A^{a}Y_{a} = A_{a}Y^{a} = \mathbf{A}\bullet\mathbf{Y}$. *(7.4.14)*

The contracted vector indices neutralize each other and the resulting object transforms as a scalar.

Here are some examples of tensor transformations with 0,1 and 2 index pairs contracted:

$T'^{abc}{}_{de} = R^{a}{}_{a'}\ R^{b}{}_{b'}\ R^{c}{}_{c'}\ S^{d'}{}_{d}\ S^{e'}{}_{e}\ T^{a'b'c'}{}_{d'e'}$ // no pairs contracted
$T'^{abc}{}_{ae} = R^{b}{}_{b'}\ R^{c}{}_{c'}\ S^{e'}{}_{e}\ T^{a'b'c'}{}_{a'e'}$ // index a contracted
$T'^{abc}{}_{ab} = R^{c}{}_{c'}\ T^{a'b'c'}{}_{a'e'}$ // index a and index b contracted
$Q' = Q$ where $Q = A^{a}Y_{a}$ and $Q' = A'^{a}Y'_{a}$ // index a contracted (7.12.3)

This shows the idea that one can take a larger tensor like $T^{abc}{}_{de}$ and form from it smaller (lower rank) tensors by contracting tilted pairs of indices. In the above example list we really have

$D^{bc}{}_{e} \equiv T^{abc}{}_{ae}$ = a mixed rank-3 tensor

$E^{c} = T^{abc}{}_{ab}$ = a contravariant vector (rank-1 tensor)

$Q = A^{a}Y_{a}$ = a scalar (rank-0 tensor). (7.12.4)

It is similarly possible to build larger tensors from smaller ones, for example

$$Z^{abc}{}_{de} = V^{a}\ W_{e}\ g^{ab}\ L^{c} \qquad (7.12.5)$$

which goes under the same rubric "outer product" mentioned earlier.

7.13 The tangent and reciprocal base vectors and expansions on same

Tangent and reciprocal base vectors

Here are some basic translations. In the first four lines, n is a label, i is a tensor index:

$$(\mathbf{e}_n)_i \quad\rightarrow\quad (\mathbf{e}_n)^i \qquad \text{// contravariant index i}$$

$$(\bar{\mathbf{e}}_n)_i \quad\rightarrow\quad (\mathbf{e}_n)_i \qquad \text{// covariant index i}$$

$$(\mathbf{E}_n)_i \quad\rightarrow\quad (\mathbf{e}^n)^i \qquad \text{// contravariant index i}$$

$$(\bar{\mathbf{E}}_n)_i \quad\rightarrow\quad (\mathbf{e}^n)_i \qquad \text{// covariant index i}$$

$$(\mathbf{e}_n)_i = S_{in} \ \ \mathit{(3.2.5)} \quad\rightarrow\quad (\mathbf{e}_n)^i = S^i{}_n = R_n{}^i \qquad \text{// contravariant index i}$$

$$(\bar{\mathbf{E}}_n)_i = R_{ni} \ \ \mathit{(6.3.2)} \quad\rightarrow\quad (\mathbf{e}^n)_i = S_i{}^n = R^n{}_i \qquad \text{// covariant index i}$$

$$(\mathbf{E}_n)_i = R_{nk}g_{ki} \ \ \mathit{(6.1.4)} \quad\rightarrow\quad (\mathbf{e}^n)^i = R^n{}_k g^{ki} = R^{ni} \qquad \text{// contravariant index i}$$

$$\Sigma_n(\bar{\mathbf{E}}_n)_a(\mathbf{e}_n)_b = \delta_{b,a} \ \mathit{(6.2.16)} \quad\rightarrow\quad \Sigma_n(\mathbf{e}^n)_a(\mathbf{e}_n)_b = \delta_{b,a} \qquad \text{// completeness}$$

$$\Sigma_n \, \bar{\mathbf{E}}_n \, \mathbf{e}_n{}^T = 1 \ \mathit{(6.2.24)} \quad\rightarrow\quad \Sigma_n \, \mathbf{e}^n \, \mathbf{e}_n{}^T = 1 \qquad \text{// completeness (matrix form)} \qquad (7.13.1)$$

As noted earlier, writing a vector in bold such as $\mathbf{e}_n$ is not enough to say whether the vector is contravariant or covariant. Here we use an index position marker * on the right. For example, $(\mathbf{e}_1)^*$ is a column vector of contravariant components, matching $S^i{}_n = (\mathbf{e}_n)^i$ from (7.13.1).

$$S = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \ .... \ \mathbf{e}_N] \ \mathit{(3.2.7)} \quad\rightarrow\quad S^*{}_* = [(\mathbf{e}_1)^*, (\mathbf{e}_2)^*, (\mathbf{e}_3)^* \ .... \ (\mathbf{e}_N)^*] \qquad (7.13.2)$$

$$R = [\bar{\mathbf{E}}_1, \bar{\mathbf{E}}_2, \bar{\mathbf{E}}_3 \ .... \ \bar{\mathbf{E}}_N]^T \ \mathit{(6.3.3)} \quad\rightarrow\quad R^*{}_* = [(\mathbf{e}^1)_*, (\mathbf{e}^2)_*, (\mathbf{e}^3)_* \ .... \ (\mathbf{e}^N)_*]^T \qquad (7.13.3)$$

The relationship between $\mathbf{e}^n$ and $\mathbf{e}_n$ is very simple. Start with (6.1.2) and use (7.4.1) $g'_{ab} \rightarrow g'^{ab}$ to get the first result below. Then apply g'_{mn} to the first result so $g'_{mn}\mathbf{e}^n = g'_{mn} \, g'^{ni} \, \mathbf{e}_i = \delta_m{}^i \, \mathbf{e}_i = \mathbf{e}_n$ which is the second result:

$$\mathbf{E}_n \equiv g'_{ni} \, \mathbf{e}_i \quad\rightarrow\quad \mathbf{e}^n = g'^{ni} \, \mathbf{e}_i \quad \text{and} \quad \mathbf{e}_n = g'_{ni} \, \mathbf{e}^i \ . \qquad (7.13.4)$$

For either contravariant or covariant indices (indices are not shown!), g'^{ni} raises the *label* on $\mathbf{e}_i$, and inverting one finds that g'_{ni} lowers the *label* on $\mathbf{e}^i$. This fact makes things easy to remember. Using the fact (3.2.6) that $\mathbf{e}_i = \partial'_i\mathbf{x}$, one has $g'^{ni} \, \mathbf{e}_i = g'^{ni} \, \partial'_i\mathbf{x} = \partial'^n\mathbf{x}$ so the above line can be expressed as

$$\mathbf{E}_n \equiv g'_{ni} \, \mathbf{e}_i \quad\rightarrow\quad \mathbf{e}^n = \partial'^n\mathbf{x} \quad \text{and} \quad \mathbf{e}_n = \partial'_n\mathbf{x} \ \ . \qquad (7.13.5)$$

The dot products (6.2.4) translate this way:

$$\mathbf{e}_n \bullet \mathbf{e}_m = \bar{g}'_{nm} \qquad \rightarrow \qquad \mathbf{e}_n \bullet \mathbf{e}_m = g'_{nm} \qquad = \partial'_n\mathbf{x} \bullet \partial'_m\mathbf{x} \qquad |\mathbf{e}_n| = \sqrt{g'_{nn}} = h'_n$$

$$\mathbf{E}_n \bullet \mathbf{e}_m = \delta_{n,m} \qquad \rightarrow \qquad \mathbf{e}^n \bullet \mathbf{e}_m = \delta^n{}_m \qquad = \partial'^n\mathbf{x} \bullet \partial'_m\mathbf{x}$$

$$\mathbf{E}_n \bullet \mathbf{E}_m = g'_{nm} \qquad \rightarrow \qquad \mathbf{e}^n \bullet \mathbf{e}^m = g'^{nm} \qquad = \partial'^n\mathbf{x} \bullet \partial'^m\mathbf{x} \qquad |\mathbf{e}^n| = \sqrt{g'^{nn}} \qquad (7.13.6)$$

The "labels" on the base vectors behave in this dot product structure the same way that up and down "indices" behave. This is the motivation for $\mathbf{E}_n \rightarrow \mathbf{e}^n$. Thus, the three final equations can be regarded as the same equation $\mathbf{e}_n \bullet \mathbf{e}_m = g'_{nm}$ where we can raise either or both indices/labels to get the other equations. For example, $\mathbf{e}^n \bullet \mathbf{e}_m = g'^n{}_m = \delta^n{}_m$.

The dot product (6.1.7) becomes

$$\mathbf{E}_n \bullet \mathbf{u}_m = R_{nm} \quad \rightarrow \quad \mathbf{e}^n \bullet \mathbf{u}_m = R^n{}_m \qquad // = \langle \mathbf{e}^n \,|\, \mathbf{u}_m \rangle \text{ in bra-ket notation (App E (g))} \qquad (7.13.7)$$

The matrix $R^n{}_m$ is a "basis change matrix" between basis $\mathbf{u}_k$ and basis $\mathbf{e}^k$.

<u>Inverse tangent and reciprocal base vectors</u>

Using the rules given above in (6.5.2),

$$g' \leftrightarrow g \qquad R \leftrightarrow S \qquad \mathbf{e}_n \rightarrow \mathbf{u}'_n \qquad \mathbf{e}'_n \rightarrow \mathbf{u}_n \qquad \mathbf{E}_n \rightarrow \mathbf{U}'_n \qquad \mathbf{E}'_n \rightarrow \mathbf{U}_n \qquad \textit{(6.5.2)}$$

we can obtain from (7.13.1) and (7.13.6) the corresponding results for the inverse tangent and reciprocal base vectors:

$(\mathbf{u}'_n)_i \quad \rightarrow \quad (\mathbf{u}'_n)^i$ // contravariant index i
$(\bar{\mathbf{u}}'_n)_i \quad \rightarrow \quad (\mathbf{u}'_n)_i$ // covariant index i

$(\mathbf{U}'_n)_i \quad \rightarrow \quad (\mathbf{u}'^n)^i$ // contravariant index i
$(\bar{\mathbf{U}}'_n)_i \quad \rightarrow \quad (\mathbf{u}'^n)_i$ // covariant index i

$(\mathbf{u}'_n)_i = R_{in} \quad \rightarrow \quad (\mathbf{u}'_n)^i = R^i{}_n = S_n{}^i$ // contravariant index i

$(\bar{\mathbf{U}}'_n)_i = S_{ni} \quad \rightarrow \quad (\mathbf{u}'^n)_i = R_i{}^n = S^n{}_i$ // covariant index i

$(\mathbf{U}'_n)_i = S_{nk}g'_{ki} \quad \rightarrow \quad (\mathbf{u}'^n)^i = S^n{}_k g'^{ki} = S^{ni}$ // contravariant index i

$\Sigma_n(\bar{\mathbf{U}}'_n)_a(\mathbf{u}'_n)_b = \delta_{b,a}$ *(6.2.16)* $\rightarrow \quad \Sigma_n(\mathbf{u}'^n)_a(\mathbf{u}'_n)_b = \delta_{b,a}$ // completeness

$R = [\mathbf{u}'_1, \mathbf{u}'_2, \mathbf{u}'_3 \;....\; \mathbf{u}'_N] \quad \rightarrow \quad R^*{}_* = [(\mathbf{u}'_1)^*, (\mathbf{u}'_2)^*, (\mathbf{u}'_3)^* \;....\; (\mathbf{u}'_N)^*]$
$S = [\bar{\mathbf{U}}'_1, \bar{\mathbf{U}}'_2, \bar{\mathbf{U}}'_3 \;....\; \bar{\mathbf{U}}'_N]^T \quad \rightarrow \quad S^*{}_* = [(\bar{\mathbf{U}}'^1)_*, (\bar{\mathbf{U}}'^2)_*, (\bar{\mathbf{U}}'^3)_* \;....\; (\bar{\mathbf{U}}'^N)_*]^T$

$\mathbf{U}'_n \equiv g_{ni}\,\mathbf{u}'_n \quad \rightarrow \quad \mathbf{u}'^n = g^{ni}\,\mathbf{u}'_i$ and $\mathbf{u}'_n = g_{ni}\,\mathbf{u}'^i$

$$\begin{aligned} \mathbf{u}'_n \bullet \mathbf{u}'_m &= \bar{g}_{nm} & &\rightarrow & \mathbf{u}'_n \bullet \mathbf{u}'_m &= g_{nm} \\ \mathbf{U}'_n \bullet \mathbf{u}'_m &= \delta_{n,m} & &\rightarrow & \mathbf{u}'^n \bullet \mathbf{u}'_m &= \delta^n{}_m \\ \mathbf{U}'_n \bullet \mathbf{U}'_m &= g_{nm} & &\rightarrow & \mathbf{u}'^n \bullet \mathbf{u}'^m &= g^{nm} \end{aligned} \qquad (7.13.8)$$

<u>Summary table.</u>

The summary table given in (6.5.9) was this

	x'-space	x-space		*(6.5.9)*
axis-aligned basis vectors	$\mathbf{e}'_n$	$\mathbf{u}_n$	$(\mathbf{e}'_n)_i = \delta_{n,i}$	$(\mathbf{u}_n)_i = \delta_{n,i}$
dual partners to the above	$\mathbf{E}'_n$	$\mathbf{U}_n$	$(\mathbf{E}'_n)_i = g'_{ni}$	$(\mathbf{U}_n)_i = g_{ni}$
tangent base vectors	$\mathbf{u}'_n$	$\mathbf{e}_n$	$(\mathbf{u}'_n)_i = R_{in}$	$(\mathbf{e}_n)_i = S_{in}$
reciprocal base vectors	$\mathbf{U}'_n$	$\mathbf{E}_n$	$(\mathbf{U}'_n)_i = g'_{ia}\,S_{na}$	$(\mathbf{E}_n)_i = g_{ia}\,R_{na}$
			$= g_{na}R_{ia}$	$= g'_{na}S_{ia}$

We translate this entire table into Standard Notation:

	x'-space	x-space		(7.13.9)
axis-aligned basis vectors	$\mathbf{e}'_n$	$\mathbf{u}_n$	$(\mathbf{e}'_n)^i = \delta_n{}^i$	$(\mathbf{u}_n)^i = \delta_n{}^i$
dual partners to the above	$\mathbf{e}'^n$	$\mathbf{u}^n$	$(\mathbf{e}'^n)^i = g'^{ni}$	$(\mathbf{u}^n)^i = g^{ni}$
tangent base vectors	$\mathbf{u}'_n$	$\mathbf{e}_n$	$(\mathbf{u}'_n)^i = R^i{}_n$	$(\mathbf{e}_n)^i = S^i{}_n = R_n{}^i$
reciprocal base vectors	$\mathbf{u}'^n$	$\mathbf{e}^n$	$(\mathbf{u}'^n)^i = g'^{ia}\,S^n{}_a$	$(\mathbf{e}^n)^i = g^{ia}\,R^n{}_a$
			$(\mathbf{u}'^n)_i = S^n{}_i$	$(\mathbf{e}^n)_i = R^n{}_i$

x-space expansions

The x-space expansions of (6.6.9)

$$
\begin{array}{llll}
\mathbf{V} = V_1\,\mathbf{u}_1 + V_2\,\mathbf{u}_2 + \ldots & = \Sigma_n V_n\,\mathbf{u}_n & \text{where}\quad \mathbf{U}_n \bullet \mathbf{V} = V_n & \mathbf{U}_n = g_{ni}\,\mathbf{u}_i \\
\mathbf{V} = \overline{V}_1\,\mathbf{U}_1 + \overline{V}_2\,\mathbf{U}_2 + \ldots & = \Sigma_n \overline{V}_n\,\mathbf{U}_n & \text{where}\quad \mathbf{u}_n \bullet \mathbf{V} = \overline{V}_n & \\
\mathbf{V} = V'_1\,\mathbf{e}_1 + V'_2\,\mathbf{e}_2 + \ldots & = \Sigma_n V'_n\,\mathbf{e}_n & \text{where}\quad \mathbf{E}_n \bullet \mathbf{V} = V'_n & \mathbf{E}_n = g'_{ni}\,\mathbf{e}_i \\
\mathbf{V} = \overline{V}'_1\,\mathbf{E}_1 + \overline{V}'_2\,\mathbf{E}_2 + \ldots & = \Sigma_n \overline{V}'_n\,\mathbf{E}_n & \text{where}\quad \mathbf{e}_n \bullet \mathbf{V} = \overline{V}'_n & \qquad\qquad \mathit{(6.6.9)}
\end{array}
$$

translate into the following :

$$
\begin{array}{llll}
\mathbf{V} = V^1\,\mathbf{u}_1 + V^2\,\mathbf{u}_2 + \ldots & = \Sigma_n V^n\,\mathbf{u}_n & \text{where}\quad \mathbf{u}^n \bullet \mathbf{V} = V^n & \mathbf{u}^n = g^{ni}\,\mathbf{u}_i \\
\mathbf{V} = V_1\,\mathbf{u}^1 + V_2\,\mathbf{u}^2 + \ldots & = \Sigma_n V_n\,\mathbf{u}^n & \text{where}\quad \mathbf{u}_n \bullet \mathbf{V} = V_n & \\
\mathbf{V} = V'^1\mathbf{e}_1 + V'^2\mathbf{e}_2 + \ldots & = \Sigma_n V'^n\,\mathbf{e}_n & \text{where}\quad \mathbf{e}^n \bullet \mathbf{V} = V'^n & \mathbf{e}^n = g'^{ni}\,\mathbf{e}_i \\
\mathbf{V} = V'_1\mathbf{e}^1 + V'_2\,\mathbf{e}^2 + \ldots & = \Sigma_n V'_n\,\mathbf{e}^n & \text{where}\quad \mathbf{e}_n \bullet \mathbf{V} = V'_n & \qquad\qquad (7.13.10)
\end{array}
$$

x'-space expansions

Similarly, the x'-space expansions of (6.6.15)

$$
\begin{array}{llll}
\mathbf{V}' = V'_1\,\mathbf{e}'_1 + V'_2\,\mathbf{e}'_2 + \ldots & = \Sigma_n V'_n\,\mathbf{e}'_n & \text{where}\quad \mathbf{E}'_n \bullet \mathbf{V}' = V'_n & \mathbf{E}'_n = g'_{ni}\,\mathbf{e}'_i \\
\mathbf{V}' = \overline{V}'_1\,\mathbf{E}'_1 + \overline{V}'_2\,\mathbf{E}'_2 + \ldots & = \Sigma_n \overline{V}'_n\,\mathbf{E}'_n & \text{where}\quad \mathbf{e}'_n \bullet \mathbf{V}' = \overline{V}'_n & \\
\mathbf{V}' = V_1\mathbf{u}'_1 + V_2\mathbf{u}'_2 + \ldots & = \Sigma_n V_n\,\mathbf{u}'_n & \text{where}\quad \mathbf{U}'_n \bullet \mathbf{V}' = V_n & \mathbf{U}'_n = g_{ni}\,\mathbf{u}'_i \\
\mathbf{V}' = \overline{V}_1\mathbf{U}'_1 + \overline{V}_2\mathbf{U}'_2 + \ldots & = \Sigma_n \overline{V}_n\,\mathbf{U}'_n & \text{where}\quad \mathbf{u}'_n \bullet \mathbf{V}' = \overline{V}_n & \qquad\qquad \mathit{(6.6.15)}
\end{array}
$$

translate into the following :

$$
\begin{array}{llll}
\mathbf{V}' = V'^1\,\mathbf{e}'_1 + V'^2\,\mathbf{e}'_2 + \ldots & = \Sigma_n V'^n\,\mathbf{e}'_n & \text{where}\quad \mathbf{e}'^n \bullet \mathbf{V}' = V'^n & \mathbf{e}'^n = g'^{ni}\,\mathbf{e}'_i \\
\mathbf{V}' = V'_1\,\mathbf{e}'^1 + V'_2\,\mathbf{e}'^2 + \ldots & = \Sigma_n V'_n\,\mathbf{e}'^n & \text{where}\quad \mathbf{e}'_n \bullet \mathbf{V}' = V'_n & \\
\mathbf{V}' = V^1\mathbf{u}'_1 + V^2\mathbf{u}'_2 + \ldots & = \Sigma_n V^n\,\mathbf{u}'_n & \text{where}\quad \mathbf{u}'^n \bullet \mathbf{V}' = V^n & \mathbf{u}'^n = g^{ni}\,\mathbf{u}'_i \\
\mathbf{V}' = V_1\mathbf{u}'^1 + V_2\mathbf{u}'^2 + \ldots & = \Sigma_n V_n\,\mathbf{u}'^n & \text{where}\quad \mathbf{u}'_n \bullet \mathbf{V}' = V_n & \qquad\qquad (7.13.11)
\end{array}
$$

Summary of all expansions:

Using implied sum notation, we can now summarize the eight expansions above, plus the unit vector expansion onto $\hat{\mathbf{e}}_n$, on just two lines :

$$
\begin{array}{lll}
\mathbf{V} = V^n\,\mathbf{u}_n = V_n\,\mathbf{u}^n = V'^n\,\mathbf{e}_n = V'_n\,\mathbf{e}^n = \mathcal{V}'^n\,\hat{\mathbf{e}}_n & \text{// x-space expansions,} \Rightarrow\ \mathcal{V}'^n = h'_n V'^n & \\
\mathbf{V}' = V'^n\,\mathbf{e}'_n = V'_n\,\mathbf{e}'^n = V^n\,\mathbf{u}'_n = V_n\,\mathbf{u}'^n\,. & \text{// x'-space expansions} & (7.13.12)
\end{array}
$$

In all cases one sees a tilted index summation where one index is a vector index and the other is a basis vector label. Half the forms shown above can be obtained from the others by just "reversing the tilt". The power of the Standard Notation makes itself felt in relations like these.

Due to this tilt situation, sometimes a basis like $\{\mathbf{e}_n\}$ appearing in $\mathbf{V} = V'^n \mathbf{e}_n$ is called a "covariant basis" while the basis $\{\mathbf{e}^n\}$ appearing in $\mathbf{V} = V'_n \mathbf{e}^n$ is called a "contravariant basis".

Corresponding expansions of higher rank tensors are presented in Section 7.17 below.

If **V** is a tensor density of weight W (see Appendix D and E) the rule for adjusting the above expansions is to make the replacement $V'^n \rightarrow J^W V'^n$ and $V'_n \rightarrow J^W V'_n$ where J is the Jacobian of (5.12.6).

7.14 Comment on Covariant versus Contravariant

Consider this expansion for a vector **V** in x-space,

$$\mathbf{V} = V^n \mathbf{b}_n \qquad\qquad V^n = \mathbf{V} \bullet \mathbf{b}^n \qquad (7.14.1)$$

where $\mathbf{b}_n$ is some basis having dual basis $\mathbf{b}^n$ where as usual $\mathbf{b}_n \bullet \mathbf{b}^m = \delta_n{}^m$, see Section 6.2. Imagine taking $V^n \rightarrow V'^n = R^n{}_m V^m$ and $\mathbf{b}_i \rightarrow \mathbf{b}_i' = Q_i{}^j \mathbf{b}_j$. What Q would cause the following to be true?

$$\mathbf{V} = V^n \mathbf{b}_n = V'^n \mathbf{b}'_n \ . \qquad (7.14.2)$$

In other words, how does one transform that basis $\mathbf{b}_n$ such that the vector **V** remains unchanged if V^n is transformed contravariantly? The answer to this question is that $Q_i{}^j = R_i{}^j$ since then (using R orthogonality rule #2 of (7.6.4))

$$V'^n \mathbf{b}'_n = [R^n{}_m V^m][R_n{}^j \mathbf{b}_j] = (R^n{}_m R_n{}^j) V^m \mathbf{b}_j = \delta_m{}^j V^m \mathbf{b}_j = V^j \mathbf{b}_j = V^n \mathbf{b}_n \ . \qquad (7.14.3)$$

Compare then the transformation of V^n with that of the basis $\mathbf{b}_n$:

$$V'^n = R^n{}_m V^m$$
$$\mathbf{b}_n' = R_n{}^m \mathbf{b}_m \qquad (7.14.4)$$

The V^m vector *components* transform with $R^n{}_m$ but the basis *vectors* have to transform with $R_n{}^m$ to maintain the invariance of the vector **V**. One varies with the down-tilt R, while the other varies with the up-tilt R, so the two objects are varying against each other in this tilt sense. They are "contra-varying", so one refers to the components V^m as **contravariant** components with respect to the basis $\mathbf{b}_m$.

If one starts over with V_n components and the $\mathbf{b}^n$ "dual" (reciprocal) expansion vectors and asks for a solution to this corresponding problem,

$$\mathbf{V} = V_n \mathbf{b}^n = V'_n \mathbf{b}'^n \qquad (7.14.5)$$

one finds not surprisingly that the dual basis must vary as $\mathbf{b}^{n'} = R^n{}_m \mathbf{b}^m$ and then one has

$$V'_n = R_n{}^m V_m$$
$$b^{n'} = R^n{}_m b^m \qquad (7.14.6)$$

which is the previous result with all indices up↔down. Comparing the tilts, one would say that the V_m again "contra vary" with the way the $\mathbf{b}^m$ vary to maintain invariance of **V**. But one does not care about the dual basis, one cares about the *basis*, so relative to the basis $\mathbf{b}_n$ one has

$$V'_n = R_n{}^m V_m$$
$$\mathbf{b}_n' = R_n{}^m \mathbf{b}_m \ . \qquad (7.14.7)$$

If the basis $\mathbf{b}_m$ is varied as shown in (7.14.7), then the dual basis $\mathbf{b}^m$ varies as in (7.14.6) and **V** of (7.14.5) remains invariant. Comparing now the way the V_m transform with the way the basis vectors $\mathbf{b}_m$ transform in (7.14.7), one sees that both equations have the same tilted $R_n{}^m$. They are "co-varying", so one refers to the components V_m as **covariant** components with respect to the basis $\mathbf{b}_m$.

7.15 The Significance of Tensor Analysis

"Why is tensor analysis important?", the reader might ask in the midst of this storm of index shuffling. Now is a good time to answer the question. Consider the following sample equation in x-space, where the fields Q, H, T and B may or may not be tensor fields:

$$Q_a{}^d{}_c(\mathbf{x}) = H_{ab}(\mathbf{x}) T^b{}_c(\mathbf{x}) \, B^d(\mathbf{x}) \ . \qquad (7.15.1)$$

Notice that when contracted indices are ignored, the remaining indices have the same type on both sides. If the various objects really were tensors, one would say this was a "valid tensor equation" based on the index structure just described.

One says that an equation is "covariant with respect to transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$" if the equation has exactly the *same form* in x'-space that it has in x-space , which for our example would be

$$Q'_a{}^d{}_c(\mathbf{x}') = H'_{ab}(\mathbf{x}') T'^b{}_c(\mathbf{x}') \, B'^d(\mathbf{x}') \quad . \qquad (7.15.2)$$

Here the word "covariant" has a new meaning, different from its being a type of vector or index. The meaning is related in the sense that, comparing the above two equations, everything has "moved" in the same manner ("co-varied") under the transformation. (Some authors think the word "invariant" is more appropriate; Ricci and Levi-Civita used the term "absolute"; continuum mechanics uses the term "frame-indifferent".)

If the objects Q, H, T and B are *tensors* under F, then covariance of any valid tensor equation like the one shown above *is guaranteed*!!

The reason is that, once the contracted indices on the two sides are ignored according to the "contraction neutralization rule" (7.12.1), the objects on the two sides of the equation have the same indices which are of the same type, so both sides are tensors of the same type, and therefore both sides transform from x-space to x'-space in the same way. If one starts, for example, with the primed equation and installs the known transformations for all the pieces, one ends up with the unprimed equation.

If this explanation is not convincing, a brute force demonstration can perhaps help out. The following is also a good exercise is using the two tilt forms of the R matrix. Recall from (7.5.13) that $S^b{}_a = R_a{}^b$ and that SR = 1 is replaced by the various orthogonality rules (7.6.4).

We shall process the primed equation into the unprimed one :

$$Q'_a{}^d{}_c(\mathbf{x}') = H'_{ab}(\mathbf{x}')T'^b{}_c(\mathbf{x}')\, B'^d(\mathbf{x}') \qquad \text{// x'-space equation} \tag{7.15.3}$$

$$[R_a{}^{a'}R^d{}_{d'}R_c{}^{c'}Q_{a'}{}^{d'}{}_{c'}(\mathbf{x})] \;=\; [R_a{}^{a'}R_b{}^{b'}\, H_{a'b'}(\mathbf{x})]\, [R^b{}_{b''}R_c{}^{c'}\; T^{b''}{}_{c'}(\mathbf{x})\,]\, [R^d{}_{d'}B^{d'}(\mathbf{x})]$$

$$=\; R_a{}^{a'}\, R^d{}_{d'}\;\, R_c{}^{c'}(R_b{}^{b'}\; R^b{}_{b''})\;\, H_{a'b'}(\mathbf{x})\, T^{b''}{}_{c'}(\mathbf{x})B^{d'}(\mathbf{x})\;\; . \tag{7.15.4}$$

Using orthogonality rule #1 of (7.6.4) we continue

$$=\; R_a{}^{a'}\, R^d{}_{d'}\;\, R_c{}^{c'}(\delta^{b'}{}_{b''})\;\, H_{a'b'}(\mathbf{x})\, T^{b''}{}_{c'}(\mathbf{x})\;\, B^{d'}(\mathbf{x})$$
$$=\; R_a{}^{a'}\, R^d{}_{d'}\;\, R_c{}^{c'} \qquad\quad H_{a'b'}(\mathbf{x})\, T^{b'}{}_{c'}(\mathbf{x})\;\, B^{d'}(\mathbf{x}) \tag{7.15.5}$$

so that, using the fact that Q is a tensor to replace Q' on the left side of (7.15.3),

$$(R_a{}^{a'}R^d{}_{d'}R_c{}^{c'})\, Q_{a'}{}^{d'}{}_{c'}(\mathbf{x}) \;=(R_a{}^{a'}\; R^d{}_{d'}\;\; R_c{}^{c'})\;\, H_{a'b'}(\mathbf{x})\, T^{b'}{}_{c'}(\mathbf{x})\;\, B^{d'}(\mathbf{x})\, . \tag{7.15.6}$$

Now apply the Cancellation Rule (7.6.12) three times to conclude that

$$Q_{a'}{}^{d'}{}_{c'}(\mathbf{x}) \;= H_{a'b'}(\mathbf{x})\, T^{b'}{}_{c'}(\mathbf{x})\;\, B^{d'}(\mathbf{x}) \tag{7.15.7}$$

and then remove all primes on indices to get

$$Q_a{}^d{}_c(\mathbf{x}) = H_{ab}(\mathbf{x})T^b{}_c(\mathbf{x})\, B^d(\mathbf{x})\;. \qquad \text{// x-space equation} \tag{7.15.8}$$

Thus it has been shown that, if all the objects transform as tensors, the equation is covariant.

Tensor density equations are also covariant. As discussed in Appendix D, a tensor density of weight W is a generalization of a tensor which has the same transformation rule as a regular tensor, but there is an extra factor of J^{-W} on the right hand side of the rule, where J is the Jacobian J = detS. For example,

$$Q'_a{}^d{}_c(\mathbf{x}') = J^{-W_Q}\; R_a{}^{a'}R^d{}_{d'}R_c{}^{c'}Q_{a'}{}^{d'}{}_{c'}(\mathbf{x}) \tag{7.15.9}$$

would indicate that Q was a tensor density of weight W_Q. If $W_Q = 0$, then Q is a regular tensor. With this definition in mind, it is easy to generalize the notion of a "covariant equation" to include tensor densities. Consider some arbitrary tensor equation which we represent by our example above,

$$Q_a{}^d{}_c(\mathbf{x}) = H_{ab}(\mathbf{x})T^b{}_c(\mathbf{x})\, B^d(\mathbf{x})\;. \tag{7.15.8}$$

Suppose all four objects Q, H, T, B are tensor densities with weights W_Q, W_H, W_T, W_B. *If* the four objects Q, H, T, B are tensor densities, and *if* the up/down free indices match on both sides (the non-contracted indices), and *if* $W_Q = W_H + W_T + W_B$, then this is a "valid tensor density equation" and covariance is guaranteed, so it follows that

$$Q'_a{}^d{}_c(\mathbf{x}') = H'_{ab}(\mathbf{x}')T'^b{}_c(\mathbf{x}')\, B'^d(\mathbf{x}') \ . \qquad\qquad (7.15.3)$$

It is trivial to edit the above proof by just adding weight factors in the right places and then of course they cancel out on the two sides.

<u>Examples of covariant tensor equations:</u> In special relativity, which happens to involve *linear* Lorentz transformations, a fundamental principle is that any "equation of motion" describing anything at all (particles, EM fields, etc) must be covariant with respect to Lorentz transformations, or it cannot be a valid equation of motion (ignoring general relativity). An equation of motion must look the same in a reference frame which is rotated, boosted, or related by any combination of boosts and rotations to some original frame of reference (see Section 5.14)).

As was noted earlier, the tradition is to write 4-vector indices as Greek letters and 3-vector spatial indices as Latin letters. For example, we can define the "electromagnetic field-strength tensor" (rank-2) this way in terms of the 4-vector "vector potential" A^μ:

$$F^{\mu\nu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu \qquad\qquad \text{with } A^\mu = (\frac{1}{c}\varphi, \mathbf{A}) \qquad\qquad (7.15.10)$$

where ∂^μ means $g^{\mu\alpha}\partial_\alpha$, the contravariant form of the gradient operator. The components are then

$$F^{\mu\nu} = \begin{bmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{bmatrix} \qquad\qquad (7.15.11)$$

where c is the speed of light and of course **E** and **B** are the electric and magnetic fields. Maxwell's two *in*homogeneous equations (that is, the two with sources) are, in SI units where $\varepsilon_0\mu_0 = 1/c^2$,

$$\partial_\nu F^{\mu\nu} = \mu_0 J^\mu \qquad\qquad \text{with } J^\mu = (c\rho, \mathbf{J}) \qquad\qquad (7.15.12)$$

while the two homogeneous equations become

$$\partial_\alpha F_{\mu\nu} + \partial_\mu F_{\nu\alpha} + \partial_\nu F_{\alpha\mu} = 0 \qquad\qquad \text{or} \qquad\qquad \partial_\alpha F_{\mu\nu} + \text{cyclic} = 0 \ . \qquad\qquad (7.15.13)$$

<u>Comment</u>: With some effort, one can show that (7.15.12) says curl $\mathbf{B} - \mu_0\varepsilon_0\partial_t\mathbf{E} = \mu_0\mathbf{J}$ and div $\mathbf{E} = \rho/\varepsilon_0$. Similarly, (7.15.13) says that curl $\mathbf{E} + \partial_t\mathbf{B} = 0$ and div $\mathbf{B} = 0$, where $\mathbf{B}$ = curl $\mathbf{A}$. **A** is the vector potential, φ is the scalar potential, **J** is current density, ρ is charge density, all in SI units.

One can see that each of the above equations involves only tensors and we expect that in x'-space these equations will take the form

$$\partial'_\nu F'^{\mu\nu} = \mu_0 J'^\mu \qquad \text{with } J'^\mu = (c\rho',\mathbf{J}') \text{ and } \quad A'^\mu = (\frac{1}{c}\varphi', \mathbf{A}')$$

$$\partial'_\alpha F'_{\mu\nu} + \partial'_\mu F'_{\nu\alpha} + \partial'_\nu F'_{\alpha\mu} = 0 \qquad \text{or} \qquad \partial'_\alpha F'_{\mu\nu} + \text{cyclic} = 0 \,. \qquad (7.15.14)$$

Objects like $\partial_\nu F^{\mu\nu}$ and $\partial_\alpha F_{\mu\nu}$ are true rank-3 tensors because the transformation F is linear. Notice that both sides of (7.15.12) transform as a contravariant vector, while both sides of (7.15.13) transform as a covariant rank-3 tensor (the right side is the all-zero rank-3 tensor).

Covariance of tensor equations involving derivatives with *non-linear* F. A tensor equation which involves derivatives of tensors is non-covariant under transformations F which are non-linear. The reason is that the derivative of a tensor is, in that case, not a tensor, as shown in the next Section. Such tensor equations *can be made* covariant by replacement of all derivatives by covariant derivatives (which are indicated by a semicolon). In general relativity, this is known as the Principle of General Covariance (Weinberg p 106). A simple example is the tensor equation $g_{ab;c} = 0$ shown in (F.9.13). Examples relating to the transformation from Cartesian to curvilinear coordinates appear in Chapter 15.

7.16 The Christoffel Business: covariant derivatives

This subject is treated in full detail in Appendix F, but here we provide some motivation. It should be noted that a normal derivative is sometimes written $\partial_a V_b = V_{b,a}$ with a comma, whereas the covariant derivative discussed below is written $V_{b;a}$ with a semicolon.

When a transformation F is non-linear, the matrix $R^a{}_b$ is a function of **x**. Thus one gets the following transformation for a lower index derivative of a covariant vector field component $\partial_a V_b(\mathbf{x})$, where a "second term" quite logically appears,

$$(\partial'_a V'_b) = (R_a{}^d \partial_d)\,(R_b{}^c V_c) = R_a{}^d R_b{}^c\,(\partial_d V_c) + R_a{}^d(\partial_d R_b{}^c)V_c \,. \qquad (7.16.1)$$

This second term did not arise in (2.4.3) where we looked at ∂_a on a *scalar* field $\varphi'(\mathbf{x}') = \varphi(\mathbf{x})$,

$$(\partial'_a \varphi') = (R_a{}^d \partial_d)\,\varphi = R_a{}^d\,(\partial_d\,\varphi). \qquad (7.16.2)$$

In special relativity, for example, where transformations are linear, $\partial_d R_b{}^c = 0$, there is no second term, and the object $\partial_a V_b$ transforms as a covariant rank-2 tensor,

$$(\partial'_a V'_b) = R_a{}^d R_b{}^c\,(\partial_d V_c) \quad . \qquad // \text{ F is a linear transformation} \qquad (7.16.3)$$

But in the general case the second term is present, so $\partial_d V_c$ fails to transform as a rank-2 covariant tensor. In this case, one defines a certain "covariant derivative" which itself has an extra piece

$$V_{b;a} \equiv \partial_a V_b - \Gamma^k{}_{ab}\, V_k \qquad \Rightarrow \qquad V_{d;c} \equiv \partial_c V_d - \Gamma^k{}_{cd}\, V_k \qquad (7.16.4)$$

where $\Gamma^c{}_{ab}$ is a certain function of the metric tensor g. One then finds that

$$V'_{b;a} = R_a{}^d R_b{}^c V_{d;c} \qquad \text{// the notation } \nabla_c V_d \equiv V_{d;c} \text{ is also commonly used}$$

or

$$[\partial'_a V'_b - \Gamma'^k{}_{ab} V'_k] = R_a{}^d R_b{}^c [\partial_d V_c - \Gamma^k{}_{cd} V_k] \qquad (7.16.5)$$

so that this *covariant* derivative of a covariant vector field V_c transforms as a covariant rank-2 tensor even with non-linear transformation F (see Christoffel Ref., 1869). This issue arises in general relativity and elsewhere. The object $\Gamma^c{}_{ab}$ (sometimes called the "Christoffel connection") is given by

$$\Gamma^c{}_{ab} \equiv \{ab,c\} \equiv \left\{ \begin{matrix} c \\ ab \end{matrix} \right\} \equiv g^{cd} [ab,d] = ½ g^{cd}(\partial_a g_{bd} + \partial_b g_{ad} - \partial_d g_{ab}) \qquad \text{// Christoffel 2nd kind}$$

$$(7.16.6)$$

$$\Gamma_{dab} \equiv [ab,d] \equiv ½ (\partial_a g_{bd} + \partial_b g_{ad} - \partial_d g_{ab}) \qquad \text{// Christoffel 1st kind}$$

and this is where the various "Christoffel symbols" come into play. In general relativity and elsewhere, $\Gamma^c{}_{ab}$ is known as the "affine connection" which represents the effect of "curved space" appearing as a force which acts on a mass (that is to say, a gravitational force), see Section 5.15.

Warning: There is a differently defined version of $\Gamma^c{}_{ab}$ floating around in the literature. The version used above and everywhere in this document is that of Weinberg and is the most common form.

The derivative of *any* tensor field other than a scalar field shows this same complication when the underlying transformation F is non-linear. For example, $\partial_a g_{bd}(\mathbf{x})$ does not transform as a rank-3 tensor,

$$\partial'_a g'_{bd}(\mathbf{x}') = (R_a{}^d \partial_d)(R_b{}^{b'} R_d{}^{d'} g_{b'd'}) = R_a{}^d R_b{}^{b'} R_d{}^{d'} (\partial_d g_{b'd'}) + \text{other terms} \qquad (7.16.7)$$

and therefore neither of the Christoffel symbols Γ_{dab} or $\Gamma^c{}_{ab}$ transforms as a tensor in this case.

See Appendix F for more detail.

7.17 Expansions of higher order tensors

Appendix E clarifies the use of direct product and polyadic notations for describing the basis vector combinations onto which higher order tensors can be expanded in a simple generalization of the vector expansions presented in Section 7.13 above, summarized in (7.13.12). There it was shown that a vector **A** can be expanded in two interesting ways :

$$\mathbf{A} = \Sigma_n A^n \mathbf{u}_n \qquad A^n \text{ are the contravariant components of A in x-space}$$

$$\mathbf{A} = \Sigma_n A'^n \mathbf{e}_n . \qquad A'^n \text{ are the contravariant components of A in x'-space} \qquad (7.17.1)$$

In the first, $\mathbf{u}_n$ are axis-aligned basis vectors, and in the second $\mathbf{e}_n$ are the tangent base vectors. If A is instead a tensor of rank n, these expansions are replaced by

$$A = \Sigma_{ijk...} A^{ijk...} (\mathbf{u}_i \otimes \mathbf{u}_j \otimes \mathbf{u}_k ...) \qquad A^{ijk...} \text{ are the contravariant components of A in x-space}$$

$$A = \Sigma_{ijk...} A'^{ijk...} (\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k ...) \qquad A'^{ijk...} \text{ are the contravariant components of A in x'-space}$$

$$(7.17.2)$$

where there are n indices in each sum, n factors in the direct products, and n contravariant indices on the components of tensors A in x-space and in x'-space. In the polyadic notation the direct-product crosses are eliminated giving

$$A = \Sigma_{ijk...} A^{ijk...} \mathbf{u}_i\mathbf{u}_j\mathbf{u}_k... \qquad A^{ijk...} \text{ are the contravariant components of A in x-space}$$
$$A = \Sigma_{ijk...} A'^{ijk...} \mathbf{e}_i\mathbf{e}_j\mathbf{e}_k... \qquad A'^{ijk...} \text{ are the contravariant components of A in x'-space} \ . \qquad (7.17.3)$$

In the case of rank-2 tensors, a product like $\mathbf{u}_i \otimes \mathbf{u}_j = \mathbf{u}_i\mathbf{u}_j$ is called a dyadic (see Section E.4). In this case (only) the product can be visualized as $\mathbf{u}_i\mathbf{u}^T{}_j$ which is a matrix constructed from a column vector to the left of a row vector. Thus one can write

$$A = \Sigma_{ij} A^{ij} \mathbf{u}_i\mathbf{u}^T{}_j \qquad A^{ij} \text{ are the contravariant components of A in x-space}$$
$$A = \Sigma_{ij} A'^{ij} \mathbf{e}_i\mathbf{e}^T{}_j \ . \qquad A'^{ij} \text{ are the contravariant components of A in x'-space} \ . \qquad (7.17.4)$$

Section E.7 promotes the interpretation of a rank-2 tensor A as an operator in a Hilbert space, where the matrices A^{ij} and A'^{ij} are matrices *associated with* the operator A in different bases,

$$A^{nm} = \langle u^n \mid A \mid u^m \rangle \ = \text{the x-space components of tensor A}$$
$$A'^{nm} = \langle e^n \mid A \mid e^m \rangle \ = \text{the x'-space components of tensor A} \ . \qquad (7.17.5)$$

These expansion methods are used in Appendices G and H to derive curvilinear expressions for two objects that play a role in continuum mechanics, $(\nabla\mathbf{v})$ and div(T) (where T is a tensor).

7.18 Collection of Facts about basis vectors $\mathbf{e}_n$, $\mathbf{u'}_n$ and $\mathbf{b}_n$.

Recall this Developmental Notation block of information regarding the $\mathbf{e}_n$ family of basis vectors, (6.4.1),

$$(\overline{\mathbf{e}}_n)_i = \overline{g}_{ij}\,(\mathbf{e}_n)_j \qquad (\overline{\mathbf{E}}_n)_i = \overline{g}_{ij}\,(\mathbf{E}_n)_j \qquad [\ \overline{\mathbf{e}}_n = \overline{g}\ \mathbf{e}_n \qquad \overline{\mathbf{E}}_n = \overline{g}\ \mathbf{E}_n] \qquad \textit{(5.8.4)}$$

$$(\mathbf{e'}_n)_i = R_{ij}(\mathbf{e}_n)_j \qquad (\mathbf{E'}_n)_i = R_{ij}(\mathbf{E}_n)_j \qquad [\ \mathbf{e'}_n = R\ \mathbf{e}_n \qquad \mathbf{E'}_n = R\ \mathbf{E}_n\ ]$$
$$(\overline{\mathbf{e}}'_n)_i = S_{ji}(\overline{\mathbf{e}}_n)_j \qquad (\overline{\mathbf{E}}'_n)_i = S_{ji}(\overline{\mathbf{E}}_n)_j \qquad [\ \overline{\mathbf{e}}'_n = S^T\,\overline{\mathbf{e}}_n \qquad \overline{\mathbf{E}}'_n = \ S^T\,\overline{\mathbf{E}}_n\ ] \qquad \textit{(2.5.1)}$$

$$(\mathbf{e}_n)_i = S_{in} \qquad (\mathbf{E}_n)_i = g_{ij}R_{nj} \ = g'_{nj}S_{ij} \qquad (\mathbf{e}_n')_i = \delta_{i,n} \qquad (\mathbf{E}_n')_i = g'_{ni}$$
$$(\overline{\mathbf{e}}_n)_i = \overline{g}_{ij}S_{jn} = R_{ji}\overline{g}'_{jn} \qquad (\overline{\mathbf{E}}_n)_i \ = \ R_{ni} \qquad (\overline{\mathbf{e}}_n')_i = \overline{g}'_{ni} \qquad (\overline{\mathbf{E}}_n')_i = \delta_{n,i}$$
$$\textit{(6.3.3)} \qquad\qquad \textit{(6.3.9)}$$
$$\mathbf{e}_n \bullet \mathbf{e}_m = \overline{g}'_{nm} \qquad \Rightarrow \qquad |\mathbf{e}_n| \ = \sqrt{\overline{g}'_{nn}} \ = \ h'_n \quad \text{(scale factor)} \qquad \mathbf{E}_n = g'_{ni}\,\mathbf{e}_i$$
$$\mathbf{E}_n \bullet \mathbf{e}_m = \delta_{n,m} \qquad\qquad \mathbf{e}_n = \ \overline{g}'_{ni}\ \mathbf{E}_i$$
$$\mathbf{E}_n \bullet \mathbf{E}_m = g'_{nm} \qquad \Rightarrow \qquad |\mathbf{E}_n| \ = \sqrt{g'_{nn}} \quad . \qquad \textit{(6.2.4)}$$

$$\mathbf{e'_n} \bullet \mathbf{e'_m} = \bar{g}'_{nm} \quad \Rightarrow \quad |\mathbf{e'_n}| = \sqrt{\bar{g}'_{nn}} = h'_n \quad \text{(scale factor)} \qquad \mathbf{E'_n} = g'_{ni}\, \mathbf{e'_i}$$
$$\mathbf{E'_n} \bullet \mathbf{e'_m} = \delta_{n,m} \qquad\qquad \mathbf{e'_n} = \bar{g}'_{ni}\, \mathbf{E'_i}$$
$$\mathbf{E'_n} \bullet \mathbf{E'_m} = g'_{nm} \quad \Rightarrow \quad |\mathbf{E'_n}| = \sqrt{g'_{nn}} \quad . \qquad \textit{(6.2.7)}$$

$$(\bar{\mathbf{E}}_\mathbf{n})_i(\mathbf{e_n})_j = \delta_{i,j} \qquad [\ \Sigma_n \bar{\mathbf{E}}_\mathbf{n}\, \mathbf{e_n}^T = 1\ ] \qquad \textit{(6.2.16) and (6.2.24)} \qquad \text{DN} \qquad \textit{(6.4.1)}$$

We shall now manually translate this entire block into Standard Notation. In doing so, we use

$$g_{ab} \to g^{ab} \qquad \text{// the contravariant rank-2 metric tensor g (and similarly for g')}$$
$$\bar{g}_{ab} \to g_{ab} \qquad \text{// the covariant rank-2 metric tensor } \bar{g} \qquad \textit{(7.4.1)}$$

$$R_{ij} \to R^i{}_j \qquad \textit{(7.5.2)}$$
$$S_{ik} \to S^i{}_k \qquad \textit{(7.5.4)}$$

$$S^i{}_j = R_j{}^i \qquad \textit{(7.5.13)}$$

and (7.13.1) for everything else. Matrix equations are ignored and only their component forms are translated. Here is the result,

$$(\mathbf{e_n})_i = g_{ij}\,(\mathbf{e_n})^j \qquad (\mathbf{e^n})_i = g_{ij}\,(\mathbf{e^n})^j \qquad \textit{(5.8.4)}$$

$$(\mathbf{e'_n})^i = R^i{}_j(\mathbf{e_n})^j \qquad (\mathbf{e'^n})^i = R^i{}_j(\mathbf{e^n})^j$$
$$(\mathbf{e'_n})_i = S^j{}_i(\mathbf{e_n})_j = R_i{}^j(\mathbf{e_n})_j \qquad (\mathbf{e'^n})_i = S^j{}_i(\mathbf{e^n})_j = R_i{}^j(\mathbf{e^n})_j \qquad \textit{(2.5.1)}$$

$$(\mathbf{e_n})^i = S^i{}_n = R_n{}^i \qquad (\mathbf{e^n})^i = g^{ij}R^n{}_j = g'^{nj}S^i{}_j \qquad (\mathbf{e'_n})^i = \delta_n{}^i \qquad (\mathbf{e'^n})^i = g'^{ni}$$
$$(\mathbf{e_n})_i = g_{ij}S^j{}_n = R^j{}_i g'_{jn} \qquad (\mathbf{e^n})_i = R^n{}_i \qquad (\mathbf{e'_n})_i = g'_{ni} \qquad (\mathbf{e'^n})_i = \delta^n{}_i$$
$$\textit{(6.3.3)} \qquad\qquad \textit{(6.3.9)}$$

$$\mathbf{e_n} \bullet \mathbf{e_m} = g'_{nm} \quad \Rightarrow \quad |\mathbf{e_n}| = \sqrt{g'_{nn}} = h'_n \quad \text{(scale factor)} \qquad \mathbf{e^n} = g'^{ni}\, \mathbf{e_i}$$
$$\mathbf{e^n} \bullet \mathbf{e_m} = \delta^n{}_m \qquad\qquad \mathbf{e_n} = g'_{ni}\, \mathbf{e^i}$$
$$\mathbf{e^n} \bullet \mathbf{e^m} = g'^{nm} \quad \Rightarrow \quad |\mathbf{e^n}| = \sqrt{g'^{nn}} \quad . \qquad \textit{(6.2.4)}$$

$$\mathbf{e'_n} \bullet \mathbf{e'_m} = g'_{nm} \quad \Rightarrow \quad |\mathbf{e'_n}| = \sqrt{g'_{nn}} = h'_n \quad \text{(scale factor)} \qquad \mathbf{e'^n} = g'^{ni}\, \mathbf{e'_i}$$
$$\mathbf{e'^n} \bullet \mathbf{e'_m} = \delta^n{}_m \qquad\qquad \mathbf{e'_n} = g'_{ni}\, \mathbf{e'^i}$$
$$\mathbf{e'^n} \bullet \mathbf{e'^m} = g'^{nm} \quad \Rightarrow \quad |\mathbf{e'^n}| = \sqrt{g'^{nn}} \quad . \qquad \textit{(6.2.7)}$$

$$(\mathbf{e^n})_i(\mathbf{e_n})^j = \delta_i{}^j \qquad \textit{(6.2.16)} \qquad \text{// completeness} \qquad \text{SN} \qquad (7.18.1)$$

Recall that $\{\mathbf{E_n}, \mathbf{e_n}\}$ are a dual pair in the sense of Section 6.2, For the inverse transformation, we also had $\{\mathbf{U_{n'}}, \mathbf{u'_n}\}$ as a dual pair in Section 6.5, and the block of data about the $\mathbf{u'_n}$ family of vectors was stated in (6.5.3) in developmental notation. We could translate that into standard notation as we did the above, but it is easier to translate (7.18.1) using the rules (6.5.2)

$$g' \leftrightarrow g \qquad R \leftrightarrow S \qquad \mathbf{e_n} \to \mathbf{u'_n} \qquad \mathbf{e'_n} \to \mathbf{u_n} \qquad \mathbf{E_n} \to \mathbf{U'_n} \qquad \mathbf{E'_n} \to \mathbf{U_n} \qquad \textit{(6.5.2)}$$

translated to Standard Notation

$$g' \leftrightarrow g \qquad R \leftrightarrow S \qquad \mathbf{e}_n \rightarrow \mathbf{u}'_n \qquad \mathbf{e}'_n \rightarrow \mathbf{u}_n \qquad \mathbf{e}^n \rightarrow \mathbf{u}'^n \qquad \mathbf{e}'^n \rightarrow \mathbf{u}^n \qquad (7.18.2)$$

Applying these rules one at a time to (7.18.1) we obtain this data block for the $\mathbf{u}'_n$ family,

$$(\mathbf{u}'_n)_i = g'_{ij}\,(\mathbf{u}'_n)^j \qquad (\mathbf{u}'^n)_i = g'_{ij}\,(\mathbf{u}'^n)^j \qquad \textit{(5.8.4)}$$

$$(\mathbf{u}_n)^i = S^i{}_j(\mathbf{u}'_n)^j \qquad (\mathbf{u}^n)^i = S^i{}_j(\mathbf{u}'^n)^j$$
$$(\mathbf{u}_n)_i = R^j{}_i(\mathbf{u}'_n)_j = S_i{}^j(\mathbf{u}'_n)_j \qquad (\mathbf{u}^n)_i = R^j{}_i(\mathbf{u}'^n)_j = S_i{}^j(\mathbf{u}'^n)_j \qquad \textit{(2.5.1)}$$

$$(\mathbf{u}'_n)^i = R^i{}_n = S_n{}^i \qquad (\mathbf{u}'^n)^i = g'^{ij}S^n{}_j = g^{nj}R^i{}_j \qquad (\mathbf{u}_n)^i = \delta_n{}^i \qquad (\mathbf{u}^n)^i = g^{ni}$$
$$(\mathbf{u}'_n)_i = g'_{ij}R^j{}_n = S^j{}_i g_{jn} \qquad (\mathbf{u}'^n)_i = S^n{}_i = R_i{}^n \qquad (\mathbf{u}_n)_i = g_{ni} \qquad (\mathbf{u}^n)_i = \delta^n{}_i$$
$$\textit{(6.3.3)} \qquad\qquad \textit{(6.3.9)}$$

$$\mathbf{u}'_n \bullet \mathbf{u}'_m = g_{nm} \quad \Rightarrow \quad |\mathbf{u}'_n| = \sqrt{g_{nn}} = h_n \quad \text{(scale factor)} \qquad \mathbf{u}'^n = g^{ni}\,\mathbf{u}'_i$$
$$\mathbf{u}'^n \bullet \mathbf{u}'_m = \delta^n{}_m \qquad\qquad \mathbf{u}'_n = g_{ni}\,\mathbf{u}'^i$$
$$\mathbf{u}'^n \bullet \mathbf{u}'^m = g^{nm} \quad \Rightarrow \quad |\mathbf{u}'^n| = \sqrt{g^{nn}} \quad . \qquad \textit{(6.2.4)}$$

$$\mathbf{u}_n \bullet \mathbf{u}_m = g_{nm} \quad \Rightarrow \quad |\mathbf{u}_n| = \sqrt{g_{nn}} = h_n \quad \text{(scale factor)} \qquad \mathbf{u}^n = g^{ni}\,\mathbf{u}_i$$
$$\mathbf{u}^n \bullet \mathbf{u}_m = \delta^n{}_m \qquad\qquad \mathbf{u}_n = g_{ni}\,\mathbf{u}^n$$
$$\mathbf{u}^n \bullet \mathbf{u}^m = g^{nm} \quad \Rightarrow \quad |\mathbf{u}^n| = \sqrt{g^{nn}} \quad . \qquad \textit{(6.2.7)}$$

$$(\mathbf{u}'^n)_i(\mathbf{u}'_n)^j = \delta_i{}^j \quad \textit{(6.2.16)} \qquad \text{// completeness} \qquad \text{SN} \qquad (7.18.3)$$

Finally, we shall restate (7.18.1) for the generic dual pair $\{\mathbf{B}_n, \mathbf{b}_n\}$ of Section 6.2. Without detailed proof, we claim that one can in general construct a transformation $\mathbf{x}' = \mathbf{F_b}(\mathbf{x})$ for which some generic complete set of vectors $\mathbf{b}_n(\mathbf{x})$ *are* the tangent base vectors like $\mathbf{e}_n(\mathbf{x})$ are for $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. For this transformation $\mathbf{F_b}$, one would have $(\mathbf{b}_n)_i = (S_b(\mathbf{x}))_{in}$ from (3.2.6), so matrix $S_b(\mathbf{x})$ would be known. Then $R_b = (S_b)^{-1}$ is also known, and one would then integrate (2.1.6) $d\mathbf{x}' = R_b(\mathbf{x})\, d\mathbf{x}$ to obtain a viable $\mathbf{F_b}(\mathbf{x})$.

As discussed in Section 6.2, for this given set of $\{\mathbf{B}_n, \mathbf{b}_n\}$, we defined $W'_{nk} \equiv \mathbf{b}_n \bullet \mathbf{b}_k$ in developmental notation which is analogous to $\overline{g}'_{nk} = \mathbf{e}_n \bullet \mathbf{e}_k$ for $\mathbf{F}(\mathbf{x})$. Thus suggests we rename $W' \rightarrow \overline{w}'$ and then we have $\overline{g}'_{nk} = \mathbf{e}_n \bullet \mathbf{e}_k$ for $\mathbf{F}(\mathbf{x})$ and $\overline{w}'_{nk} = \mathbf{b}_n \bullet \mathbf{b}_k$ for $\mathbf{F_b}(\mathbf{x})$, where $\overline{w}' = W$. In Standard Notation these equations become $g'_{nk} = \mathbf{e}_n \bullet \mathbf{e}_k$ and $w'_{nk} = \mathbf{b}_n \bullet \mathbf{b}_k$ since

$$W'_{nk} = \overline{w}'_{nk} \quad \rightarrow \quad w'_{nk} \qquad \text{like} \qquad \overline{g}'_{nk} \quad \rightarrow \quad g'_{nk}$$
$$w'_{nk} \quad \rightarrow \quad w'^{nk} \qquad\qquad g'_{nk} \quad \rightarrow \quad g'^{nk} \; . \qquad (7.18.4)$$

So the rules for translating (7.18.1) from $\mathbf{e}_n$ to generic $\mathbf{b}_n$ are are very simple :

$$\mathbf{e} \rightarrow \mathbf{b} \qquad // \text{ that is, } \mathbf{e}_n \rightarrow \mathbf{b}_n,\ \mathbf{e}^n \rightarrow \mathbf{b}^n \text{ and the same for primed basis vectors}$$
$$g' \rightarrow w' \qquad // \text{ that is, } g'_{ij} \rightarrow w'_{ij} \text{ and } g'^{ij} \rightarrow w'^{ij} \qquad (7.18.5)$$

Note that the arrows → in (7.18.4) are for DN→SN, whereas those in (7.18.5) are SN **e** → SN **b**.

The result is

$$(\mathbf{b}_n)_i = g_{ij}\,(\mathbf{b}_n)^j \qquad (\mathbf{b}^n)_i = g_{ij}\,(\mathbf{b}^n)^j \qquad \textit{(5.8.4)}$$

$$(\mathbf{b}'_n)^i = R^i{}_j(\mathbf{b}_n)^j \qquad (\mathbf{b}'^n)^i = R^i{}_j(\mathbf{b}^n)^j$$
$$(\mathbf{b}'_n)_i = S^j{}_i(\mathbf{b}_n)_j = R_i{}^j(\mathbf{b}_n)_j \qquad (\mathbf{b}'^n)_i = S^j{}_i(\mathbf{b}^n)_j = R_i{}^j(\mathbf{b}^n)_j \qquad \textit{(2.5.1)}$$

$$(\mathbf{b}_n)^i = S^i{}_n = R_n{}^i \qquad (\mathbf{b}^n)^i = g^{ij}R^n{}_j = w'^{nj}S^i{}_j \qquad (\mathbf{b}_n')^i = \delta_n{}^i \qquad (\mathbf{b}'^n)^i = w'^{ni}$$
$$(\mathbf{b}_n)_i = g_{ij}S^j{}_n = R^j{}_i w'_{jn} \qquad (\mathbf{b}^n)_i = R^n{}_i \qquad (\mathbf{b}_n')_i = w'_{ni} \qquad (\mathbf{b}'^n)_i = \delta^n{}_i$$
$$\textit{(6.3.3)} \qquad \textit{(6.3.9)}$$

$$\mathbf{b}_n \bullet \mathbf{b}_m = w'_{nm} \quad \Rightarrow \quad |\mathbf{b}_n| = \sqrt{w'_{nn}} = h'_n \quad \text{(scale factor)} \qquad \mathbf{b}^n = w'^{ni}\,\mathbf{b}_i$$
$$\mathbf{b}^n \bullet \mathbf{b}_m = \delta^n{}_m \qquad \mathbf{b}_n = w'_{ni}\,\mathbf{b}^i$$
$$\mathbf{b}^n \bullet \mathbf{b}^m = w'^{nm} \quad \Rightarrow \quad |\mathbf{b}^n| = \sqrt{w'^{nn}} \quad . \qquad \textit{(6.2.4)}$$

$$\mathbf{b}'_n \bullet \mathbf{b}'_m = w'_{nm} \quad \Rightarrow \quad |\mathbf{b}'_n| = \sqrt{w'_{nn}} = h'_n \quad \text{(scale factor)} \qquad \mathbf{b}'^n = w'^{ni}\,\mathbf{b}'_i$$
$$\mathbf{b}'^n \bullet \mathbf{b}'_m = \delta^n{}_m \qquad \mathbf{b}'_n = w'_{ni}\,\mathbf{b}'^i$$
$$\mathbf{b}'^n \bullet \mathbf{b}'^m = w'^{nm} \quad \Rightarrow \quad |\mathbf{b}'^n| = \sqrt{w'^{nn}} \quad . \qquad \textit{(6.2.7)}$$

$$(\mathbf{b}^n)_i(\mathbf{b}_n)^j = \delta_i{}^j \qquad \textit{(6.2.16)} \qquad // \text{ completeness} \qquad \text{SN} \qquad (7.18.6)$$

7.19 More on basis vectors and matrix elements of R and S

Ambiguity

In the summary tables of Section 7.18 we see various equations like these,

$$(\mathbf{u}_n)^i = \delta_n{}^i \qquad (\mathbf{e}_n)^i = R_n{}^i \qquad (\mathbf{e}'_n)^i = \delta_n{}^i \qquad (\mathbf{u}'_n)^i = R^i{}_n\ . \qquad (7.19.1)$$

There is a certain ambiguity in writing, for example, $(\mathbf{e}_n)^i$: it is not clearly stated in what basis component i is evaluated! The basis depends on the expansion in which a component appears. For example, from (7.13.10) we write for a generic vector **V**,

$$\mathbf{V} = \Sigma_i V^i\ \mathbf{u}_i \qquad \mathbf{u}^i \bullet \mathbf{V} = V^i \qquad \textit{(7.13.10)}$$

and we then know that V^i is a component in the u-basis. Applying this to $\mathbf{V} = \mathbf{e}_n$ (an x-space vector) we conclude that

$$\mathbf{e_n} = \Sigma_i\ (\mathbf{e_n})^i\ \mathbf{u_i} \qquad\qquad \mathbf{u^i} \bullet \mathbf{e_n} = (\mathbf{e_n})^i\ . \tag{7.19.2}$$

Thus, $(\mathbf{e_n})^i$ is in fact a component in the u-basis, not for example the e-basis. We would then write

$$(\mathbf{e_n})^i = (\mathbf{e_n})^{(u)i} = R_n{}^i\ . \tag{7.19.3}$$

This is the "natural basis" to use for the components of $\mathbf{e_n}$ and for any x-space vector **V** where we use the natural expansion $\mathbf{V} = \Sigma_i V^i\ \mathbf{u_i}$ shown above.

One could however expand $\mathbf{e_n}$ on the $\mathbf{e_n}$ basis to get

$$\mathbf{e_n} = \Sigma_i\ (\mathbf{e_n})^{(e)i}\ \mathbf{e_i} \qquad\qquad \mathbf{e^i} \bullet \mathbf{e_n} = (\mathbf{e_n})^{(e)i} = \delta^i{}_n\ . \tag{7.19.4}$$

Then with no ambiguity one could write

$$\begin{aligned} (\mathbf{e_n})^{(u)i} &= R_n{}^i && \text{// natural} \\ (\mathbf{e_n})^{(e)i} &= \delta_n{}^i \end{aligned} \tag{7.19.5}$$

Thus, when we write $(\mathbf{e_n})^i$, we imply that we are using the u-basis for the components. Since this applies to any x-space vector, it is also true for $(\mathbf{u_n})^i$, so $(\mathbf{u_n})^i = (\mathbf{u_n})^{(u)i} = \mathbf{u^i} \bullet \mathbf{u_n}$.

What about x'-space vectors? From (7.13.11) we write

$$\mathbf{V'} = \Sigma_i(V')^i\ \mathbf{e'_i} \qquad\qquad \mathbf{e'^i} \bullet \mathbf{V'} = (V')^i\ . \tag*{\textit{(7.13.11)}}$$

Applying this to $\mathbf{V'} = \mathbf{e'_n}$ (an x'-space vector) we conclude that

$$\mathbf{e'_n} = \Sigma_i(\mathbf{e'_n})^i\ \mathbf{e'_i} \qquad\qquad \mathbf{e'^i} \bullet \mathbf{e'_n} = (\mathbf{e'_n})^i = \delta^i{}_n\ . \tag{7.19.6}$$

Thus, $(\mathbf{e'_n})^i$ is in fact a component in the e'-basis, not for example the u'-basis. We would then write

$$(\mathbf{e'_n})^i = (\mathbf{e'_n})^{(e')i} = \delta_n{}^i\ . \tag{7.19.7}$$

This is the "natural basis" to use for the components of $\mathbf{e'_n}$ and for any x'-space vector **V'** where we use the natural expansion $\mathbf{V'} = \Sigma_i(V')^i\ \mathbf{e'_i}$ shown above.

One could however expand $\mathbf{e'_n}$ on the $\mathbf{u'_n}$ basis to get

$$\mathbf{e'_n} = \Sigma_i\ (\mathbf{e'_n})^{(u')i}\ \mathbf{u'_i} \qquad\qquad \mathbf{u'^i} \bullet \mathbf{e'_n} = (\mathbf{e'_n})^{(u')i} = \mathbf{u^i} \bullet \mathbf{e_n} = (\mathbf{e_n})^i = R_n{}^i\ . \tag{7.19.8}$$

Then with no ambiguity one could write

$$\begin{aligned} (\mathbf{e'_n})^{(e')i} &= \delta_n{}^i && \text{// natural} \\ (\mathbf{e'_n})^{(u')i} &= R_n{}^i \end{aligned} \tag{7.19.9}$$

Thus, when we write $(\mathbf{e'}_n)^i$, we imply that we are using the e'-basis for the components. Since this applies to any x'-space vector, it is also true for $(\mathbf{u'}_n)^i$, so $(\mathbf{u'}_n)^i = (\mathbf{u'}_n)^{(e')i} = \mathbf{e'}^i \bullet \mathbf{u'}_n$.

We summarize the above discussion as follows:

Fact: $\mathbf{u}_n$ and $\mathbf{e}_n$ components are by default presented in the u-basis (7.19.10)
$\mathbf{e'}_n$ and $\mathbf{u'}_n$ components are by default presented in the e'-basis

Table of basis vector dot products

Since there are four kinds of basis vectors of interest, $\mathbf{u}_n$, $\mathbf{e}_n$, $\mathbf{u'}_n$ and $\mathbf{e'}_n$ with lower labels, one might imagine there are 16 scalar products of interest,

	u	e	u'	e'
u	u•u	u•e	u•u'	u•e'
e	e•u	e•e	e•u'	e•e'
u'	u'•u	u'•e	u'•u'	u'•e'
e'	e'•u	e'•e	e'•u'	e'•e'

However, since a scalar product only exists within a Hilbert space (x-space or x'-space), the cross space entries in this table make no sense, so we eliminate them to get

	u	e	u'	e'
u	u•u	u•e	-	-
e	e•u	e•e	-	-
u'	-	-	u'•u'	u'•e'
e'	-	-	e'•u'	e'•e'

(7.19.11)

Since u•e = e•u and u'•e' = e'•u', we see that there are only 6 distinct scalar products of interest. Collecting data from the tables (7.13.1) and (7.13.2) we summarize these 6 cases as follows:

Table of basis vector dot products:

$$(\mathbf{u}_n)^i = \mathbf{u}^i \bullet \mathbf{u}_n = <\mathbf{u}^i \mid \mathbf{u}_n> = g^i{}_n = \mathbf{u'}^i \bullet \mathbf{u'}_n = <\mathbf{u'}^i \mid \mathbf{u'}_n>$$

$$(\mathbf{e}_n)^i = \mathbf{u}^i \bullet \mathbf{e}_n = <\mathbf{u}^i \mid \mathbf{e}_n> = S^i{}_n = R_n{}^i$$

$$(\mathbf{e}_n')^i = \mathbf{e'}^i \bullet \mathbf{e'}_n = <\mathbf{e'}^i \mid \mathbf{e'}_n> = g'^i{}_n = \mathbf{e}^i \bullet \mathbf{e}_n = <\mathbf{e}^i \mid \mathbf{e}_n>$$

$$(\mathbf{u}_n')^i = \mathbf{e'}^i \bullet \mathbf{u'}_n = <\mathbf{e'}^i \mid \mathbf{u'}_n> = R^i{}_n = S_n{}^i \qquad (7.19.12)$$

Since $\mathbf{a} \bullet \mathbf{b} = \mathbf{a}' \bullet \mathbf{b}'$, the first and third lines each contain 2 of the 6 cases. We are careful to use the correct tensor notation $g_n{}^i$ and $g'_n{}^i$ for $\delta_n{}^i$ as discussed near equation (7.4.17). Doing this, we may state the following :

Fact: In any of the above equation lines, one may raise the label n and/or lower the label/index i and end up with another valid equation line. (7.19.13)

In this manner, from the 4 lines above one may generate 12 more lines of equations.

Note that we have slipped in the Dirac notation < | > described in (E.7.4) as an alternate way to write the dot products.

As an example of the above Fact, if we raise the label n in the third line we get these valid equations,

$$(\mathbf{e}^{n'})^i = \mathbf{e}'^i \bullet \mathbf{e}^{n'} = <\mathbf{e}'^i \,|\, \mathbf{e}^{n'}> = g'^{in} = \mathbf{e}^i \bullet \mathbf{e}^n = <\mathbf{e}^i \,|\, \mathbf{e}^n>$$

and if we then lower index i, we get another set of valid equations,

$$(\mathbf{e}^{n'})_i = \mathbf{e}'_i \bullet \mathbf{e}^{n'} = <\mathbf{e}'_i \,|\, \mathbf{e}^{n'}> = g'_i{}^n = \mathbf{e}_i \bullet \mathbf{e}^n = <\mathbf{e}_i \,|\, \mathbf{e}^n> .$$

To prove the above Fact, one first observes in which space a vector lies, so one knows which metric tensor raises and lowers the index. Then one uses these facts from (7.18.1) and (7.18.3),

$$\begin{array}{llll} \mathbf{e}^n = g'^{ni}\,\mathbf{e}_i & \mathbf{e}'^n = g'^{ni}\,\mathbf{e}'_i & \mathbf{u}^n = g^{ni}\,\mathbf{u}_i & \mathbf{u}'^n = g^{ni}\,\mathbf{u}'_i \\ \mathbf{e}_n = g'_{ni}\,\mathbf{e}^i & \mathbf{e}'_n = g'_{ni}\,\mathbf{e}'^i & \mathbf{u}_n = g_{ni}\,\mathbf{u}^n & \mathbf{u}'_n = g_{ni}\,\mathbf{u}'^i \quad . \end{array} \tag{7.19.14}$$

For example, consider the third line in (7.19.12),

$$(\mathbf{e}_n)^i = \mathbf{u}^i \bullet \mathbf{e}_n = <\mathbf{u}^i \,|\, \mathbf{e}_n> = R_n{}^i .$$

Apply g'^{mn} to get

$$g'^{mn}(\mathbf{e}_n)^i = g'^{mn}\,\mathbf{u}^i \bullet \mathbf{e}_n = g'^{mn}<\mathbf{u}^i \,|\, \mathbf{e}_n> = g'^{mn}\,R_n{}^i$$

or

$$[g'^{mn}\,\mathbf{e}_n]^i = \mathbf{u}^i \bullet [g'^{mn}\,\mathbf{e}_n] = <\mathbf{u}^i \,|\, [g'^{mn}]\,\mathbf{e}_n> = g'^{mn}\,R_n{}^i$$

or

$$(\mathbf{e}^m)^i = \mathbf{u}^i \bullet [\mathbf{e}^m] = <\mathbf{u}^i \,|\, \mathbf{e}^m> = R^{mi} \qquad \text{// see (7.5.9) about } R^{mi}$$

which shows that the label n can be raised to get another valid set of equations. Now start with the same third line in (7.19.12) and instead apply g_{ji} to get

$$g_{ji}(\mathbf{e}_n)^i = g_{ji}\mathbf{u}^i \bullet \mathbf{e}_n = g_{ji}<\mathbf{u}^i \,|\, \mathbf{e}_n> = g_{ji}R_n{}^i$$

or

$$g_{ji}(\mathbf{e}_n)^i = [g_{ji}\mathbf{u}^i] \bullet \mathbf{e}_n = <g_{ji}\mathbf{u}^i \,|\, \mathbf{e}_n> = g_{ji}R_n{}^i$$

or

$$(\mathbf{e}_n)_j = \mathbf{u}_j \bullet \mathbf{e}_n = <\mathbf{u}_j \,|\, \mathbf{e}_n> = R_{nj} \qquad \text{// see (7.5.9) about } R_{nj}$$

which shows that the index i can be lowered to get another valid set of equations.

Matrix elements of R and S

One can regard the elements $R^i{}_j$ and $S^i{}_j$ as being certain matrix elements of operators R and S, and this is most easily handled in the Dirac bra-ket notation of (E.7.4). Since $\mathbf{V}' = R\mathbf{V}$ (vector transformation rule) one sees that a matrix element <a' | R | b> = <a' | R b> must have | b> being an x-space vector, and <a' | being a transposed x'-space vector.

We shall now compute matrix elements of R in several basis combinations. As examples of $\mathbf{V}' = R\mathbf{V}$ we know that,

$$\mathbf{e}'_n = R\,\mathbf{e}_n$$
$$\mathbf{u}'_n = R\,\mathbf{u}_n$$

or

$$|\mathbf{e}'_n\rangle = R\,|\mathbf{e}_n\rangle$$
$$|\mathbf{u}'_n\rangle = R\,|\mathbf{u}_n\rangle \ . \qquad (7.19.15)$$

We can "close on the left" in various ways. For example,

$$\langle \mathbf{e}'^i\,|\,\mathbf{e}'_n\rangle = \langle \mathbf{e}'^i\,|\,R\,|\,\mathbf{e}_n\rangle = g'^i{}_n$$

$$\langle \mathbf{u}'^i\,|\,\mathbf{e}'_n\rangle = \langle \mathbf{u}'^i\,|\,R\,|\,\mathbf{e}_n\rangle = S^i{}_n = R_n{}^i \qquad (7.19.16)$$

Here we look up the scalar products on the left of each line using table (7.19.12), making use of Fact (7.19.13), and adjusting the various index names and the up and down index sense.

Now start instead with the second equation of (7.19.15) and close two ways, again looking up the basis vector scalar products in (7.19.12),

$$\langle \mathbf{e}'^i\,|\mathbf{u}'_n\rangle = \langle \mathbf{e}'^i\,|\,R\,|\,\mathbf{u}_n\rangle = R^i{}_n = S_n{}^i$$

$$\langle \mathbf{u}'^i\,|\mathbf{u}'_n\rangle = \langle \mathbf{u}'^i\,|\,R\,|\,\mathbf{u}_n\rangle = g^i{}_n \ . \qquad (7.19.17)$$

Here then are four R matrix elements of interest,

$$\langle \mathbf{e}'^i\,|\,R\,|\,\mathbf{e}_n\rangle = g'^i{}_n$$

$$\langle \mathbf{u}'^i\,|\,R\,|\,\mathbf{e}_n\rangle = S^i{}_n = R_n{}^i$$

$$\langle \mathbf{e}'^i\,|\,R\,|\,\mathbf{u}_n\rangle = R^i{}_n = S_n{}^i$$

$$\langle \mathbf{u}'^i\,|\,R\,|\,\mathbf{u}_n\rangle = g^i{}_n \ . \qquad (7.19.18)$$

To get matrix elements of S, we use the fact (7.9.8) that $R = S^T$ and the fact (7.9.17) that $<\mathbf{a}|S|\mathbf{b}> = <\mathbf{b}|S^T|\mathbf{a}> = <\mathbf{b}|R|\mathbf{a}>$, so we translate the above four lines:

$$<\mathbf{e}_n \mid S \mid \mathbf{e}'^i> = <\mathbf{e}'^i \mid R \mid \mathbf{e}_n> = g'^i{}_n$$

$$<\mathbf{e}_n \mid S \mid \mathbf{u}'^i> = <\mathbf{u}'^i \mid R \mid \mathbf{e}_n> = S^i{}_n = R_n{}^i$$

$$<\mathbf{u}_n \mid S \mid \mathbf{e}'^i> = <\mathbf{e}'^i \mid R \mid \mathbf{u}_n> = R^i{}_n = S_n{}^i$$

$$<\mathbf{u}_n \mid S \mid \mathbf{u}'^i> = <\mathbf{u}'^i \mid R \mid \mathbf{u}_n> = g^i{}_n \,. \tag{7.19.19}$$

and the above then is a full statement of all matrix elements of S and R, for a particular up down sense of the basis vector labels. As usual, $g'^i{}_n = g^i{}_n = \delta^i{}_n$ but we maintain the true tensor form to allow for up down modifications as per the following claim:

Fact: In the above equations, one can raise/lower labels on either or both sides to get new valid equations. (7.19.20)

The proof of this claim is the same as the proof of Fact (7.19.13) above using (7.19.14). For example,

$$<\mathbf{u}'^i \mid R \mid \mathbf{u}_n> = g^i{}_n$$

so

$$g_{ji}<\mathbf{u}'^i \mid R \mid \mathbf{u}_n> = g_{ji}g^i{}_n$$

or

$$<\mathbf{u}'_j \mid R \mid \mathbf{u}_n> = g_{jn} \,.$$

Similarly $<\mathbf{e}'^i \mid R \mid \mathbf{u}_n> = R^i{}_n = S_n{}^i \Rightarrow <\mathbf{e}'_i \mid R \mid \mathbf{u}_n> = R_{in} = S_{ni}$.

<u>Interpretation of $R^i{}_n$ and $S^i{}_n$</u>

Since $\mathbf{V}' = R\,\mathbf{V}$, one can regard the operator R as a mapping R : X → X' (x-space → x'-space). R may be regarded as a "cross tensor" having one foot in each space. As shown in (7.19.19),

$$R^i{}_j = <\mathbf{e}'^i \mid R \mid \mathbf{u}_j> = [\,R^{(e',u)}]^i{}_j \tag{7.19.21}$$

so one can regard R as an abstract cross-tensor being expanded as follows in the mixed-basis sense of (E.10.4),

$$R = \Sigma_{ij}\,[\,R^{(e',u)}]^i{}_j\,(\mathbf{e}'_i \otimes \mathbf{u}^j) \,. \tag{7.19.22}$$

As usual, the coefficients (the cross tensor components) can be projected out using

$$(\mathbf{e'^a} \otimes \mathbf{u_b}) \bullet R = \Sigma_{ij} [R^{(e',u)}]^i{}_j (\mathbf{e'^a} \otimes \mathbf{u_b}) \bullet (\mathbf{e'_i} \otimes \mathbf{u^j})$$

$$= \Sigma_{ij} [R^{(e',u)}]^i{}_j (\mathbf{e'^a} \bullet \mathbf{e'_i}) \otimes (\mathbf{u_b} \bullet \mathbf{u^j})$$

$$= \Sigma_{ij} [R^{(e',u)}]^i{}_j \delta^a{}_i \delta_b{}^j$$

$$= [R^{(e',u)}]^a{}_b . \qquad (7.19.23)$$

Similarly, since $\mathbf{V} = S\,\mathbf{V'}$ one can regard the operator S as a mapping S : X' → X where, from (7.19.19),

$$S^i{}_j = <\mathbf{u^i} | S | \mathbf{e'_j}> = [S^{(u,e')}]^i{}_j \qquad (7.19.24)$$

and then the cross-tensor expansion and projection is given by,

$$S = \Sigma_{ij} [S^{(u,e')}]^i{}_j (\mathbf{u_i} \otimes \mathbf{e'^j})$$

$$(\mathbf{u^a} \otimes \mathbf{e'_b}) \bullet S = [S^{(u,e')}]^a{}_b . \qquad (7.19.25)$$

A Pitfall Example

One knows from (7.19.19) that

$$< \mathbf{u'_i} | R\, \mathbf{u^a}> = < \mathbf{u'_i} | R | \mathbf{u^a}> = g_i{}^a = \delta_i{}^a .$$

This is also known from,

$$< \mathbf{u'_i} | R\, \mathbf{u^a}> = < \mathbf{u'_i} | \mathbf{u'^a}> = \delta_i{}^a .$$

Consider now the following slight of hand wherein we get a different result,

$$< \mathbf{u'_i} | R\, \mathbf{u^a}> = [R\, \mathbf{u^a}]_i = R_i{}^j(\mathbf{u^a})_j = R_i{}^j\delta^a{}_j = R_i{}^a \qquad \text{wrong!}$$

Where is the error being made here? Well, the vector $\mathbf{u'^a} = R\mathbf{u^a}$ being in x'-space has a natural e'-type basis as discussed below (7.19.7), but here the component $[R\mathbf{u^a}]_i$ is in the u'-basis. When doing something "unnaturally" one must pay more attention to labels. The correct version of the above is

$$< \mathbf{u'_i} | R\, \mathbf{u^a}> = [R\, \mathbf{u^a}]^{(u')}{}_i = [R^{(u',u)}]^i{}_j (\mathbf{u^a})^{(u)}{}_j = [R^{(u',u)}]^i{}_j \delta^a{}_j$$

$$= [R^{(u',u)}]^i{}_a = <\mathbf{u'^i} | R | \mathbf{u_a}> = g^i{}_a = \delta^i{}_a . \qquad (7.19.26)$$

Now consider instead a different example. From (7.19.19) one knows that,

$$< \mathbf{e'_i} | R\, \mathbf{u^a}> = < \mathbf{e'_i} | R | \mathbf{u^a}> = R_i{}^a .$$

In this case write

$$< \mathbf{e'}_i \,|\, R\, \mathbf{u}^a> \;=\; [R\, \mathbf{u}^a]_i \;=\; R_i{}^j \delta^a{}_j = R_i{}^a$$

and there is no problem because e' is the natural basis for x'-space. In more detail

$$< \mathbf{e'}_i \,|\, R\, \mathbf{u}^a> \;= [R\, \mathbf{u}^a]^{(e')}{}_i \;=\; [R^{(e',u)}]_i{}^j\, (\mathbf{u}^a)^{(u)}{}_j \;=\; [R^{(e',u)}]_i{}^j\, \delta^a{}_j$$

$$= \; [R^{(e',u)}]_i{}^a \; = <\mathbf{e'}_i \,|\, R \,|\, \mathbf{u}^a> \;=\; R_i{}^a\,. \qquad (7.19.27)$$

We find that the Dirac notation is useful because it provides a bulletproof formalism for avoiding ambiguities such as that of the previous example. The above equation sequence can be written (implied sum on j as usual)

$$< \mathbf{e'}_i \,|\, R\, \mathbf{u}^a> \;=\; < \mathbf{e'}_i \,|\, R\, [1]\, \mathbf{u}^a> \;=\; < \mathbf{e'}_i \,|\, R\, [\; |\, \mathbf{u}^j><\mathbf{u}_j\,|\,]\, \mathbf{u}^a>$$

$$= \; < \mathbf{e'}_i \,|\, R \,|\, \mathbf{u}^j><\mathbf{u}_j \,|\, \mathbf{u}^a> = [R^{(e',u)}]_i{}^j\; (\mathbf{u}^a)^{(u)}{}_j \qquad (7.19.28)$$

where we use the (E.7.4) completeness relation $1 = \;|\, \mathbf{u}^j><\mathbf{u}_j\,|$. In fact

$$\text{x-space completeness:} \quad 1 = \;|\, \mathbf{u}^j><\mathbf{u}_j| = \;|\, \mathbf{u}_j><\mathbf{u}^j| \;=\; |\, \mathbf{e}^j><\mathbf{e}_j| = \;|\, \mathbf{e}_j><\mathbf{e}^j|$$

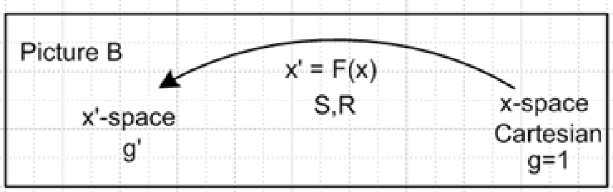
$$\text{x'-space completeness:} \quad 1 = \;|\, \mathbf{u'}^j><\mathbf{u'}_j| = \;|\, \mathbf{u'}_j><\mathbf{u'}^j| \;=\; |\, \mathbf{e'}^j><\mathbf{e'}_j| = \;|\, \mathbf{e'}_j><\mathbf{e'}^j| \qquad (7.19.29)$$

There are just restatements of the (6.2.8) duality idea that $\mathbf{b}^i \bullet \mathbf{b}_j \;= <\mathbf{b}^i \,|\, \mathbf{b}_j > \;=\; \delta^i{}_j$ for any basis.

8. Transformation of Differential Length, Area and Volume

8.1 Overview of Chapter 8

This Chapter and all remaining Chapters use the Standard Notation introduced in Chapter 7.
The term N-piped is short for N dimensional parallelepiped.
The context is Picture B:



(8.1.1)

Since this Chapter is quite lengthy, a brief overview is in order:

The transformation of differential length, area, and volume is first framed in terms of the mapping of an orthogonal differential N-piped in x'-space to a skewed differential N-piped in x-space. The N-piped in x'-space has axis-aligned edges of length dx'^n, while the N-piped in x-space has edges $\mathbf{e_n} dx'^n$ where $\mathbf{e_n}$ are the tangent base vectors introduced in Chapter 3. We want to learn what happens to the edges, face areas and volume as one differential N-piped is mapped into the other by the curvilinear coordinate transformation $\mathbf{x'} = \mathbf{F}(\mathbf{x})$. After solving this problem, we go on to consider the transformations of arbitrary differential vectors, areas and volume.

8.2: The differential N-piped mapping
The differential N-piped mapping is described and various symbols are defined.

8.3: Properties of the finite N-piped spanned by the $\mathbf{e_n}$ in x-space
Results from Appendix B concerning *finite* N-piped geometric properties are quoted. Certain definition changes are made to make the formulas suitable for tensor analysis. The purpose of the lengthy Appendix B is to lend credence to the general formulas for elements of area and volume in N dimensions.

8.4: Back to the differential N-piped mapping: how edges, areas and volume transform
(a) Setup. The finite N-piped edges $\mathbf{e_n}$ are scaled by curvilinear coordinate variations dx'^n to create a differential N-piped in x-space having edges $(\mathbf{e_n} dx'^n)$.
(b) Edge Transformation. The edges $\mathbf{e_n}$ of the x-space N-piped map to axis-aligned edges $\mathbf{e'_n}$ in x'-space.
(c) Area Transformation. Tensor density methods as presented in Appendix D are used here.
(d) Volume Transformation. The volume transformation is computed several different ways.
(e) Covariant Magnitudes. These are $| d\mathbf{x'}^{(n)} |$, $| d\mathbf{A'}^n |$ and $| dV' |$ in the Curvilinear View of x'-space.
(f) Two Theorems. (1) $g'^{nn} g' = \text{cof}(g'_{nn})$ and (2) $|(\Pi^{\mathbf{x}}{}_{i\neq n}\mathbf{e_i})| = \sqrt{\text{cof}(g'_{nn})}$.
(g) Cartesian-View Magnitude Ratios. Appropriate for the continuum mechanics application.
(h) Nested Cofactor Formulas and S^TS notation: Description of all "leaf" areas of an N-piped in both descending and ascending orders.

(i) Transformation of arbitrary differential vectors, areas and volume. Having built confidence with the general form of vector area and volume expressions in the N-piped case, the N-piped is jettisoned and formulas for the transformation of arbitrary vectors, areas and volume are derived.
(j) Concatenation of Transformations. What happens to the transformation of vectors, areas and volumes when two transformations are concatenated? One result is that $J = J_1 J_2$.
(k) Examples of area magnitude transformation for N = 2,3,4
Example 2: Spherical Coordinates: area patches

8.5: Transformation of Differential Volume applied to Integration
The volume transformation obtained in Section 8.4 is related to the traditional notion of the Jacobian changing the "measure" in an integration. The "Jacobian Integration Rule" can then be expressed as a distributional equation.

8.6: Interpretations of the Jacobian

8.7: Volume integration of a tensor field under linear transformations
Under suitable conditions, the volume integration of a tensor field integrand yields a tensor of the same type.

8.2 The differential N-piped mapping

Consider these relations involving a differential distance in the n-direction ($dx'^n > 0$):

$d\mathbf{x}'^{(n)} = \mathbf{e}'_n \, dx'^n$ x'-space axis-aligned differential vector, and $(\mathbf{e}'_n)^i = \delta_n{}^i$ *(3.2.2)*

$d\mathbf{x}^{(n)} = \mathbf{e}_n \, dx'^n$ x-space mapping of the above vector under F^{-1} or R^{-1}

$d\mathbf{x}'^{(n)} = R(\mathbf{x}) \, d\mathbf{x}^{(n)}$ relation of the two differential vectors (contravariant rule) *(2.1.6)*

$\mathbf{e}'_n = R(\mathbf{x}) \, \mathbf{e}_n$ *(3.3.2)* (8.2.1)

A superscript (n) on the differentials makes clear there is no implied sum on n. Note that application of R(x) to the second equation yields the first equation, making use of the last two equations, so the second equation really does contain dx'^n. In more generality, recall $\mathbf{V} = \Sigma_n V'^n \mathbf{e}_n$ from (7.13.10).

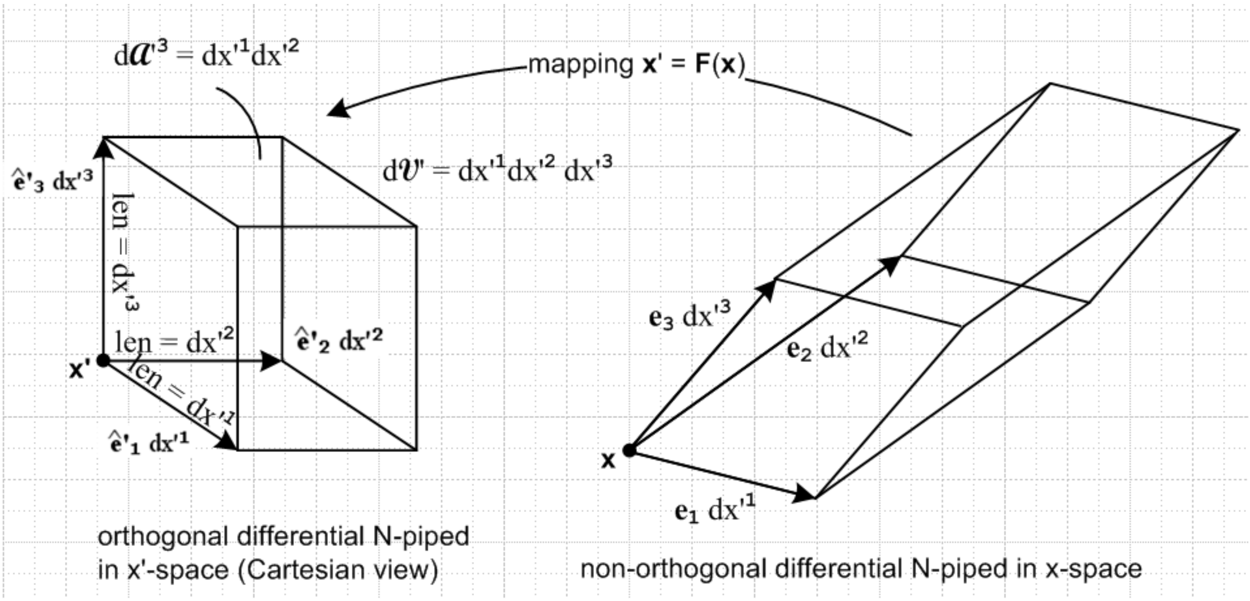
The vectors $d\mathbf{x}^{(n)}$ span a differential N-piped in x-space, while the $d\mathbf{x}'^{(n)}$ span a corresponding differential N-piped in x'-space. The two N-pipeds are related by the mapping $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. Since the regions are differentially small, this mapping is the same as the linearized mapping $d\mathbf{x}' = R \, d\mathbf{x}$.

The metric tensor in x-space will be taken to be g = 1, so it is a Cartesian space.

As discussed at the end of Section 5.16, the x'-space N-piped can be viewed in (at least) two ways depending on how the metric tensor g' is set. For a continuum mechanics flow application, one sets g' = 1 and this gives the Cartesian View of the x'-space N-piped. For such flows dot products and magnitudes of vectors like $d\mathbf{x}^{(n)}$ are not invariant under the transformation. For our curvilinear coordinates application, however, we set $g' = RgR^T = RR^T$ and this causes vector dot products and magnitudes to be invariant and we can talk about such objects as being tensorial scalars. This is the Curvilinear View of x'-space.

When g' ≠ 1, it is impossible to accurately represent the Curvilinear-View picture of the x'-space N-piped as a drawing in physical space (for N=3). This subject is discussed for a sample 2D system in Section C.5. Although the basis vectors $(\mathbf{e}'_n)^i = \delta_n{}^i$ in x'-space are *always* axis-aligned, they are only orthogonal for an orthogonal coordinate system, since $\mathbf{e}'_n \bullet \mathbf{e}'_m = \mathbf{e}_n \bullet \mathbf{e}_m = g'_{nm}$. Nevertheless, even for a non-orthogonal system we draw the axes as if they were orthogonal, which at least provides a representation of the notion of "axis aligned" basis vectors. For N > 3 one at least imagine this kind of drawing.

Due to these graphical difficulties, in the drawing below the Cartesian View of x'-space is shown. Since g' = 1 for this situation, the axis-aligned basis vectors $\mathbf{e}'_n$ are in fact unit vectors $\hat{\mathbf{e}}'_n$ and are orthogonal, so the picture becomes at least comprehensible:



orthogonal differential N-piped
in x'-space (Cartesian view)

non-orthogonal differential N-piped in x-space

(8.2.2)

The orthogonal Cartesian-View N-piped allows *visualization* of these curvilinear coordinate variations, all $dx'^k > 0$ (objects below are primed because they exist in x'-space),

$$d\mathcal{L}'^n \equiv dx'^n \qquad \text{length} \qquad n = 1,2....N$$
$$d\mathcal{A}'^n \equiv \Pi_{i\neq n} dx'^i \qquad \text{area}$$
$$d\mathcal{V}' \equiv \Pi_i dx'^i = d\mathcal{A}'^n\, d\mathcal{L}'^n \ . \qquad \text{volume (no implied sum on n)} \qquad (8.2.3)$$

For example, for N=3 one would have

$$d\mathcal{L}'^1 \equiv dx'^1 \qquad d\mathcal{L}'^2 \equiv dx'^2 \qquad d\mathcal{L}'^3 \equiv dx'^3$$
$$d\mathcal{A}'^1 = dx'^2dx'^3 \qquad d\mathcal{A}'^2 = dx'^3dx'^1 \qquad d\mathcal{A}'^3 = dx'^1dx'^2$$
$$d\mathcal{V}' = dx'^1dx'^2dx'^3 = d\mathcal{A}'^1d\mathcal{L}'^1 = d\mathcal{A}'^2d\mathcal{L}'^2 = d\mathcal{A}'^3d\mathcal{L}'^3 \ . \qquad (8.2.4)$$

The Cartesian-View x'-space N-piped is always orthogonal because the $(\mathbf{e}'_n)$ are orthonormal axis-aligned unit vectors (since g'=1). In contrast, the x-space N-piped is typically rotated and possibly skewed as well (if the coordinates x'^i describe a non-orthogonal coordinate system). The transformation F and its linearized version R map the skewed x-space N-piped into the orthogonal x'-space N-piped. As one moves around in x-space so that point **x** changes, the picture on the left above keeps its shape, just translating itself to the new point **x**', but the picture on the right changes shape and volume because the vectors $\mathbf{e}_n(\mathbf{x})$ are functions of **x**.

It is our goal to write expressions for edges, areas and volumes in these two spaces and to then show how these objects transform between the two spaces. To this end, we shall rely on work done in Appendix B which is summarized in the next Section. Following that, we shall add to each edge a differential distance associated with that edge (such as $\mathbf{e}_r \rightarrow \mathbf{e}_r dr$ in spherical coordinates), and that will bring us back to the differential N-piped picture above.

8.3 Properties of the finite N-piped spanned by the $\mathbf{e}_n$ in x-space

The finite N-piped spanned by the tangent base vectors $\mathbf{e}_n$ in x-space has the following properties (as shown in Appendix B) :

- The N spanning edges are the vectors $\mathbf{e}_n$ which have lengths $|\mathbf{e}_n| = h'_n$ (scale factors, (5.11.7)). (8.3.1)

- There are 2^N vertices. (8.3.2)

- There are N pairs of faces. The two faces of each pair are parallel in N dimensions. One face of each pair touches the point where the tails of all the $\mathbf{e}_n$ vectors meet (the near face) while the other does not touch this meeting point (the far face). (8.3.3)

- Each face of an N-piped is an (N-1)-piped having 2^{N-1} vertices. The faces are planar surfaces of dimension N-1 embedded in an N dimensional space. (8.3.4)

- A face's vector area $\mathbf{A}^n$ is spanned by all the $\mathbf{e}_i$ *except* $\mathbf{e}_n$ and is labeled by this missing $\mathbf{e}_n$ vector. (8.3.5)

- The far face has out-facing vector area $\mathbf{A}^n$, and the near face has out-facing area vector $-\mathbf{A}^n$. These vector areas are normal to the faces. (8.3.6)

- The vector area $\mathbf{A}^n$ is given by several equivalent expressions:

$$\mathbf{A}^n = |\det(S^a{}_b)|\ \mathbf{e}^n$$

$$\mathbf{A}^n = \sigma\,(-1)^{n-1}\,\Pi^x{}_{i\neq n}\ \mathbf{e}_i$$

$$\mathbf{A}^n = \sigma\,(-1)^{n-1}\,\mathbf{e}_1 \times \mathbf{e}_2 \ldots \times \mathbf{e}_N \qquad // \ \mathbf{e}_n \text{ missing} \qquad \sigma \equiv \text{sign}[\det(S^a{}_b)] = \text{sign}[\det(R^a{}_b)]$$

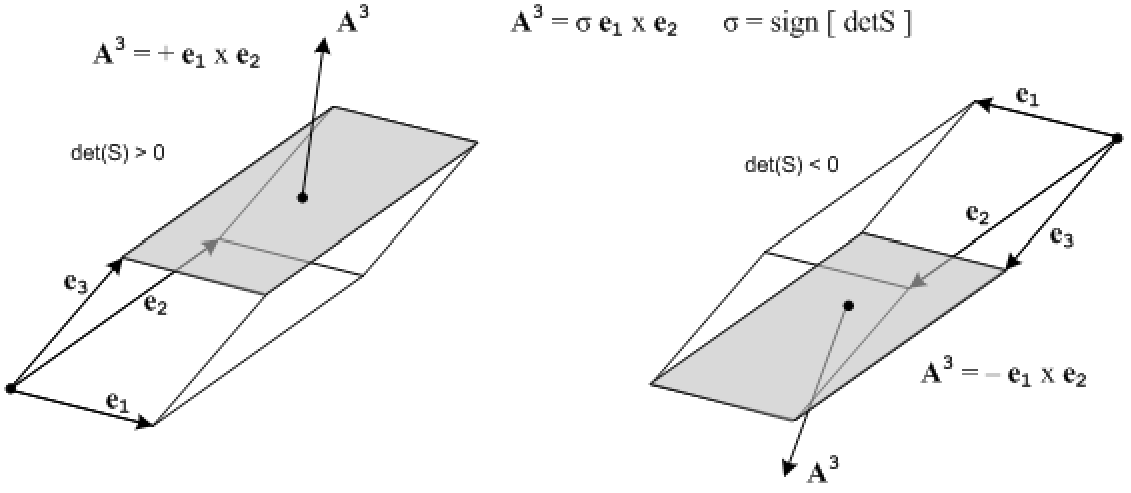
$$(\mathbf{A}^n)_i = \sigma\,(-1)^{n-1}\,\varepsilon_{iabc..x}\,(\mathbf{e}_1)^a(\mathbf{e}_2)^b \ldots .\,(\mathbf{e}_N)^x \qquad // \ e_n \text{ missing} \qquad (8.3.7)$$

- The volume of the N-piped is given by (see Section 5.12 concerning J)

$$V = |\det[\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \ldots \mathbf{e}_N]| = |\det(S^a{}_b)| = g'^{1/2} = |J| \ . \qquad (8.3.8)$$

The mapping picture above does not apply to a finite N-piped. The *finite* N-piped just discussed exists in x-space. One might ponder into what shape it maps in x'-space under the transformation F. In the case of spherical coordinates, Fig (1.13), *all* of x-space maps into a certain orthogonal "office building" in x'-space. A finite N-piped within x-space maps into some very complicated 3D shape within the office building which is bounded by curves which in general are not even coplanar. The point is that a finite N-piped in x-space does *not* map back into some nice orthogonal N-piped in x'-space and the picture drawn in the previous Section does not apply. However, when differentials are added in the next Section, *then,* since the mapped regions are very small, the mapping of the x-space differential N-piped is in fact an "orthogonal" N-piped in x'-space. This is because for a tiny region near some point $\mathbf{x}$, the mapping between $d\mathbf{x}$ and $d\mathbf{x}'$ is described by the linear relation $d\mathbf{x}' = R(\mathbf{x})\, d\mathbf{x}$ of (2.1.6).

Conventions for defining the area vector and volume. In Appendix B (as reported in (8.3.6)) the area vector $\mathbf{A}^\mathbf{n}$ is defined so that the out-facing normal of the x-space N-piped's "far face n" is $\mathbf{A}^\mathbf{n}$, regardless of the sign of det(S). This was done to simplify the computation of the flux of a vector field emerging from the N-piped in the geometric divergence calculation in Chapter 9. That calculation is then valid for either sign of det(S), a sign that we call σ. The following picture illustrates on the left an x-space 3-piped which is obtained by reverse-mapping the x'-space orthogonal 3-piped using an S which has $\det(S) > 0$. On the right is the x-space 3-piped that results for $S \rightarrow -S$. These two 3-pipeds are related by a parity inversion of all points through the origin. If the origin lies far away, these two N-pipeds lie far away from each other, a fact not illustrated in the picture:



(8.3.9)

Notice that $\mathbf{A}^3$ for the "far face 3" is outfacing in both cases.

This definition of the vector area is not suitable for the vector analysis we are about to undertake. Instead of the above situation, we will redefine $\mathbf{A}^3 = \mathbf{e}_1 \times \mathbf{e}_2$ for both pictures, and this will cause the $\mathbf{A}^3$ vector in the right picture to point up into the interior of the N-piped. This new definition allows us to interpret $\mathbf{A}^3 = \mathbf{e}_1 \times \mathbf{e}_2$ as a "valid vector equation" to which we may apply the ideas of covariance and tensor densities.

Notice that under a parity transformation, this newly defined $\mathbf{A}^3$ is a "pseudovector" which is one which does not reverse direction under a parity transformation, since $\mathbf{A}^3 = (-\mathbf{e}_1) \times (-\mathbf{e}_2)$. The subject of parity and handedness and the sign of det(S) is discussed more in Section 6.9.

Here then are the expressions for $\mathbf{A}^n$ with this new definition, where the new forms are obtained from (8.3.7) by multiplying by σ = sign (det(S)) :

$$\mathbf{A}^n = \det(S^a{}_b)\, \mathbf{e}^n = J\, \mathbf{e}^n \tag{8.3.10}$$

$$\mathbf{A}^n = (-1)^{n-1}\, \mathbf{e}_1 \text{ x } \mathbf{e}_2 \ldots \text{ x } \mathbf{e}_N \qquad // \ \mathbf{e}_n \text{ missing} \tag{8.3.11}$$

$$(\mathbf{A}^n)_i = (-1)^{n-1}\, \varepsilon_{iabc..x}\, (\mathbf{e}_1)^a(\mathbf{e}_2)^b \ldots. (\mathbf{e}_N)^x \qquad // \ e_n \text{ missing} \tag{8.3.12}$$

The second line shows that pseudovector areas can only exist for an odd number of dimensions (such as N=3).

A similar redefinition of the volume will now be done. In Appendix B the volume is defined so as to be a positive number regardless of σ with the result V = |det(S)|. We now redefine the volume by multiplication by σ, so that now V = det(S) which is of course a negative number when det(S) < 0, which in turn means the $\mathbf{e}_n$ are forming a left-handed coordinate system as per Section 6.9. So:

$$V = \det[\, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \ldots \mathbf{e}_N] = \det(S^a{}_b) = J\ . \tag{8.3.13}$$

8.4 Back to the differential N-piped mapping: how edges, areas and volume transform

(a) The Setup

If the edges of the finite N-piped described above are *scaled* by positive differentials $dx'^n > 0$, the result is a differential N-piped in x-space which has these properties:

$$d\mathbf{x}^{(n)} = \mathbf{e}_n\, dx'_n \ \ \mathit{(8.2.1)} \qquad // \text{ edges}$$

$$d\mathbf{A}^n = J\, \mathbf{e}^n\ (\Pi_{i\neq n} dx'^i)\ \ \mathit{(8.3.10)} \qquad // \text{ areas}$$

$$d\mathbf{A}^n = (-1)^{n-1}\, (dx'_1\mathbf{e}_1) \text{ x } (dx'_2\mathbf{e}_2) \ldots \text{ x } (dx'_N\mathbf{e}_N)\ \mathit{(8.3.11)} \qquad // \ \mathbf{e}_n \text{ missing from cross product}$$

$$= (-1)^{n-1}\, \mathbf{e}_1 \text{ x } \mathbf{e}_2 \ldots \text{ x } \mathbf{e}_N\, (\Pi_{i\neq n} dx'^i) \qquad // \ \mathbf{e}_n \text{ missing from cross product}$$

$$= (-1)^{n-1}\, (\Pi^x{}_{i\neq n}\mathbf{e}_i)\, (\Pi_{i\neq n} dx'^i) \qquad // \text{ shorthand notation of (A.10.5)}$$

$$(d\mathbf{A}^n)_i = (-1)^{n-1}\, \varepsilon_{iabc..x}\, (\mathbf{e}_1)^a(\mathbf{e}_2)^b \ldots. (\mathbf{e}_N)^x\ \ (\Pi_{i\neq n} dx'^i) \qquad // \ \mathbf{e}_n \text{ factor and index missing, } \mathit{(8.3.12)}$$

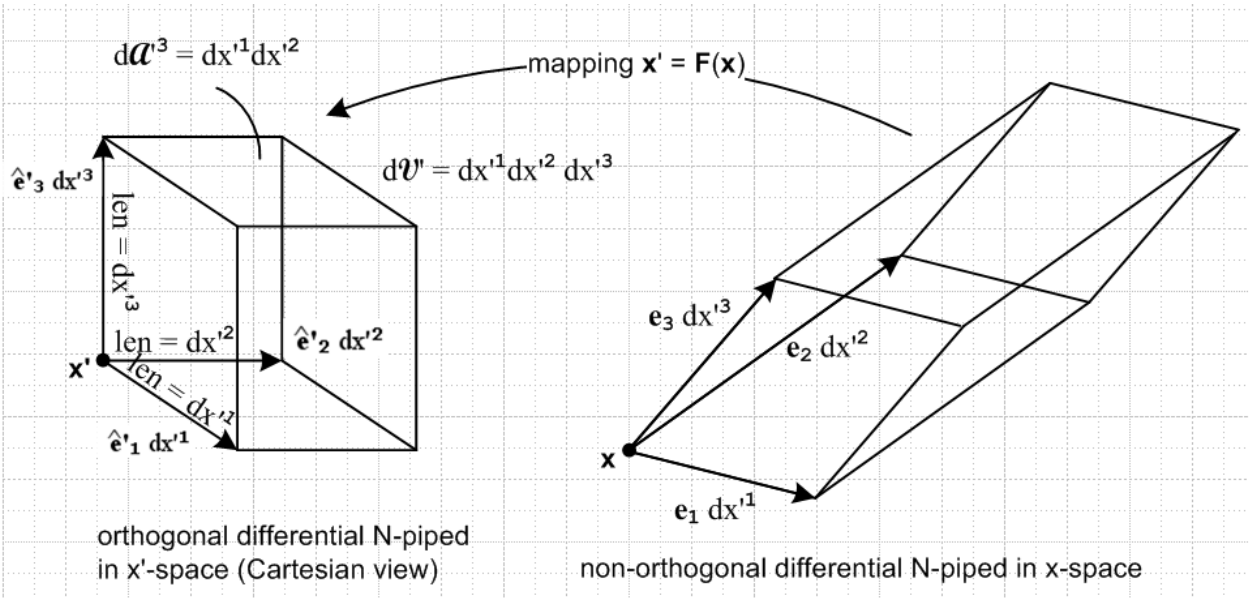
$$dV = \det[dx'\mathbf{e}_1, dx'\mathbf{e}_2, dx'\mathbf{e}_3 \ldots dx'\mathbf{e}_N] = \det[\, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \ldots \mathbf{e}_N]\, (\Pi_i dx'^i) \qquad \mathit{(8.3.13)}$$

$$= \det(S^a{}_b)(\Pi_i dx'^i) = J\, (\Pi_i dx'^i) \qquad // \ (7.13.2) \text{ and } (5.12.6)$$

$$= \varepsilon_{abc..x}\, (dx'\mathbf{e}_1)^a(dx'\mathbf{e}_2)^b \ldots.. (dx'\mathbf{e}_N)^x \qquad // \ (5.12.8)$$

(8.4.a.1)

where d$\mathbf{A}^n$ and dV are obtained from $\mathbf{A}^n$ and V according to the new definitions described above. These equations apply to the right side of the N-piped mapping picture Fig (8.2.2) which is replicated here :



(8.2.2)

For spherical coordinates, the N-piped on the right would be spanned by these vectors, see (3.4.6),

$$\mathbf{e}_r dr = \hat{\mathbf{r}} dr, \qquad \mathbf{e}_\theta d\theta = r\hat{\boldsymbol{\theta}} d\theta \qquad \mathbf{e}_\varphi d\varphi = r\sin\theta\, \hat{\boldsymbol{\varphi}} d\varphi \ . \tag{8.4.a.2}$$

The reader is reminded that the vectors shown on the right of (8.2.2) exist in x-space, and that this x-space is assumed to be Cartesian for this entire Chapter, meaning g = 1 and $V_k = V^k$ for vectors.

(b) Edge Transformation

The edge $d\mathbf{x}^{(n)}$ we know from (2.1.6) transforms as a tensorial vector under transformation F, so

$$d\mathbf{x}'^{(n)} = R\, d\mathbf{x}^{(n)} \qquad \text{where} \qquad d\mathbf{x}^{(n)} = \mathbf{e}_n\, dx'^n \quad \mathit{(8.2.1)}\ . \tag{8.4.b.1}$$

Making use of (7.18.1) and RS = 1, evaluation gives

$$[d\mathbf{x}'^{(n)}]^i = R^i{}_j\, (\mathbf{e}_n)^j\, dx'^n = R^i{}_j\, S^j{}_n\, dx'^n = (RS)^i{}_n\, dx'^n = \delta^i{}_n\, dx'^n$$

$$\Rightarrow\ d\mathbf{x}'^{(n)} = \mathbf{e}'_n\, dx'^n \qquad \text{since } (\mathbf{e}'_n)^i = \delta_n{}^i \tag{8.4.b.2}$$

so this contravariant edge points along the the n axis direction in x'-space of Fig (8.2.2). This fact was stated above in (8.2.1), we are just exercising our notation.

(c) Area Transformation

How does $d\mathbf{A}^n$ transform under F? Looking at the component form stated above in (8.4.a.1),

$$(d\mathbf{A}^n)_i = (-1)^{n-1}\, \varepsilon_{iabc..x}\, (\mathbf{e}_1)^a(\mathbf{e}_2)^b....\, (\mathbf{e}_N)^x \ \ (\Pi_{i\neq n}dx'^i) \qquad // \ \mathbf{e}_n \text{ factor and index missing} \qquad (8.4.c.1)$$

it is seen that dA^n is a combination of tensor objects like $\varepsilon_{iabc..x}$ and $(e_2)^b$. As discussed in Appendix D, a vector density of weight 0 is an ordinary vector such as $\mathbf{e}_2$ or $d\mathbf{x}$. The ε tensor (rank-N) is a tensor density of weight -1. In forming more complicated tensor objects, the rule (D.2.3) is that one adds the weights of the objects being combined. Therefore, one may conclude that the object $d\mathbf{A}^n$ is a vector density of weight -1. This has the *immediate implication* that dA^n transforms to x'-space under F according to the rule (D.1.4),

$$d\mathbf{A}'^n = J\, R\, d\mathbf{A}^n \qquad \text{or} \qquad (d\mathbf{A}'^n)^i = J\, R^i{}_j\, (d\mathbf{A}^n)^j \qquad // \ J^{-W} = J^{-(-1)} = J \qquad (8.4.c.2)$$

where J = det(S) is the Jacobian of Section 5.12. Moreover, the above equation for $(d\mathbf{A}^n)_i$ is a "valid tensor density equation" as per (7.15.9) and is therefore covariant. This means that in x'-space the equation has the exact same form, but tensor objects are primed (the dx'^i are constants),

$$(d\mathbf{A}'^n)_i = (-1)^{n-1}\, \varepsilon'_{iabc..x}\, (e'_1)^a(e'_2)^b....\, (e'_N)^x \ \ (\Pi_{i\neq n}dx'^i) \qquad // \ \mathbf{e}'_n \text{ factor and index missing}$$

(8.4.c.3)

Insertion of (D.5.13) that $\varepsilon'_{iabc..x} = J^2\, \varepsilon_{iabc..x}$ and $(e'_n)^i = \delta_n{}^i$ then gives

$$\begin{aligned} (d\mathbf{A}'^n)_i &= (-1)^{n-1}\, J^2\, \varepsilon_{iabc..x}\, \delta_1{}^a\delta_2{}^b...\delta_N{}^x \ \ (\Pi_{i\neq n}dx'^i) \\ &= (-1)^{n-1}\, J^2\, \varepsilon_{i123..N}\, (\Pi_{i\neq n}dx'^i) \qquad // \text{ index n missing on } \varepsilon \\ &= \delta^n{}_i\, J^2 \ (\Pi_{i\neq n}dx'^i) \qquad (8.4.c.4) \end{aligned}$$

The last step follows from the fact that $\varepsilon_{i123..N}$ with n missing must vanish if $i \neq n$, and if $i=n$ then $\varepsilon_{n123..N} = (-1)^n\ \varepsilon_{123..n..N} = (-1)^n$. The conclusion then is that

$$d\mathbf{A}'^n = J^2\ (\Pi_{i\neq n}dx'^i)\, \mathbf{e}'^n \qquad \text{since } (\mathbf{e}'^n)_i = \delta^n{}_i \ \ \mathit{(7.18.1)} \qquad (8.4.c.5)$$

and the covariant vector area $d\mathbf{A}'^n$ points in the n-axis direction in x'-space.

For the covariance-dubious reader, here is an alternate derivation of (8.4.c.4) from (8.4.c.2) :

$$\begin{aligned} (d\mathbf{A}'^n)_i &= J\, R_i{}^j\, (d\mathbf{A}^n)_j = J\, R_i{}^j\, \{[\ (-1)^{n-1}\, \varepsilon_{iabc..x}\, (\mathbf{e}_1)^a(\mathbf{e}_2)^b....\, (\mathbf{e}_N)^x \ \ (\Pi_{i\neq n}dx'^i)\ \} \qquad // \ \mathbf{e}_n \text{ missing} \\ &= J\, R_i{}^j\, \{[\ (-1)^{n-1}\, \varepsilon_{iabc..x}\, R_1{}^a\, R_2{}^b....\, R_N{}^x \ \ (\Pi_{i\neq n}dx'^i)\ \} \qquad // \ R_n{}^\kappa \text{ missing} \\ &= J\, \{[\ (-1)^{n-1}\, \varepsilon_{iabc..x}\, R_i{}^j\, R_1{}^a\, R_2{}^b....\, R_N{}^x \ \ (\Pi_{i\neq n}dx'^i)\ \} \qquad // \ R_n{}^\kappa \text{ missing} \\ &= J\, \{[\ (-1)^{n-1}\, \varepsilon_{iabc..x}\, S^j{}_i\, S^a{}_1\, S^b{}_2....\, S^x{}_N \ \ (\Pi_{i\neq n}dx'^i)\ \}\ . \qquad // \ S^\kappa{}_n \text{ missing} \qquad (8.4.c.6) \end{aligned}$$

where $R_i{}^j = S^j{}_i$ by (7.5.13). The εSSS..S object is the determinant of an N-1 dimensional matrix. If $i \neq n$, then i must be some index like 1 or 2. But then the determinant has two identical columns, so vanishes. Thus, the result is proportional to $\delta^n{}_i$. Continuing,

$$= \delta^n{}_i\, J \,\{\, (-1)^{n-1}\, \varepsilon_{nabc..x}\, S^j{}_n\, S^a{}_1\, S^b{}_2 .... S^x{}_N\, (\Pi_{i\neq n} dx'^i) \,\} \qquad // \ S^x{}_n \text{ missing in group}$$

$$= \delta^n{}_i\, J \,\{\, (-1)^{n-1}\, \varepsilon_{nabc..x}\, S^a{}_1\, S^b{}_2 .... S^j{}_n .... S^x{}_N\, (\Pi_{i\neq n} dx'^i) \,\}$$

$$= \delta^n{}_i\, J \,\{\, \varepsilon_{abc..n...x}\, S^a{}_1\, S^b{}_2 .... S^j{}_n .... S^x{}_N\, (\Pi_{i\neq n} dx'^i) \,\}$$

$$= \delta^n{}_i\, J \,\{\, \det(S)\, (\Pi_{i\neq n} dx'^i) \,\}$$

$$= \delta^n{}_i\, J^2\, (\Pi_{i\neq n} dx'^i) \qquad (8.4.c.7)$$

which agrees with the result (8.4.c.4) from the covariant x'-space equation.

(d) Volume Transformation

What about the volume dV? Assume for the moment that dV is correctly represented this way, where any n will do:

$$dV = d\mathbf{A}^n \bullet d\mathbf{x}^{(n)} \ = \ (d\mathbf{A}^n)_i\, [d\mathbf{x}^{(n)}]^i \,. \qquad // \text{ no implied sum on n} \qquad (8.4.d.1)$$

Installing our (8.4.a.1) expressions for $d\mathbf{A}^n$ and $d\mathbf{x}^{(n)}$ gives, using (7.18.1) that $\mathbf{e}^n \bullet \mathbf{e}_n = 1$,

$$dV = J\, \mathbf{e}^n\, (\Pi_{i\neq n} dx'^i) \bullet (\mathbf{e}_n\, dx'_n) \ = J\, (\Pi_i dx'^i)\, \mathbf{e}^n \bullet \mathbf{e}_n = \ J\, (\Pi_i dx'^i) \qquad (8.4.d.2)$$

which agrees with (8.4.a.1), verifying our assumption (8.4.d.1). Looking at $dV = (d\mathbf{A}^n)_i [d\mathbf{x}^{(n)}]^i$, dV is seen to be a tensor combination of a vector density of weight -1 with a vector density of weight 0 (an ordinary vector $d\mathbf{x}^{(n)}$), so according to (D.2.3) dV must be scalar density of weight -1. Rule (D.1.4) then says,

$$dV' = J\, dV \ . \qquad (8.4.d.3)$$

Since $dV = J\, (\Pi_i dx'^i)$ from (8.4.a.1). one gets

$$dV' = J^2\, (\Pi_i dx'^i) \,. \qquad (8.4.d.4)$$

Again, one can verify this last result from the x'-space covariant form of the equation:

$$dV' = d\mathbf{A}'^n \bullet d\mathbf{x}'^{(n)} \ = \ \underset{(8.4.c.5)}{\{\, J^2 (\Pi_{i\neq n} dx'^i)\, \mathbf{e}'^n \,\}} \bullet \underset{(8.2.1)}{\{\mathbf{e}'_n\, dx'^n)} \ = J^2 (\Pi_i dx'^i)\, \mathbf{e}'^n \bullet \mathbf{e}'_n \ = J^2 (\Pi_i dx'^i) \qquad (8.4.d.5)$$

since $\mathbf{e}'^n \bullet \mathbf{e}'_n = \mathbf{e}^n \bullet \mathbf{e}_n = 1$.

An alternate derivation of the dV transform rule dV' = J dV comes from just staring at (8.4.a.1),

$$dV = \varepsilon_{abc..x}\, (dx'\mathbf{e}_1)^a (dx'\mathbf{e}_2)^b ..... (dx'\mathbf{e}_N)^x \qquad (8.4.d.6)$$

which by the argument above is seen directly to transform as a tensor density of weight -1.

To complete the circle, we can verify for a second time the claim (8.4.d.1) that $dV = d\mathbf{A}^n \bullet d\mathbf{x}^{(n)}$:

$$d\mathbf{A}^n \bullet d\mathbf{x}^{(n)} = (d\mathbf{A}^n)_i\, [d\mathbf{x}^{(n)}]^i$$

$$= \{(-1)^{n-1}\, \varepsilon_{iabc..x}\, (\mathbf{e}_1)^a (\mathbf{e}_2)^b .... (\mathbf{e}_N)^x\ (\Pi_{k\neq n} dx'^k)\ \}\ (dx'^n \mathbf{e}_n)^i \qquad // \ \mathbf{e}_n \text{ missing .....}$$

$$= \ \varepsilon_{abc..i..x}\, (\mathbf{e}_1)^a (\mathbf{e}_2)^b .. (\mathbf{e}_n)^\kappa .. (\mathbf{e}_N)^x\ (\Pi_k dx'^i) \qquad // \text{ note indices on } \varepsilon$$

$$= \det(S)\ (\Pi_k dx'^i) \qquad // \ (5.12.8) \text{ where } (\mathbf{e}_b)^a = S^a{}_b \text{ from } (7.18.1)$$

$$= dV\ . \qquad // \text{ from } (8.4.a.1) \qquad (8.4.d.7)$$

To summarize the above, by examining the vector density nature of our various objects, we have been able to determine exactly how edges, areas and the volume transform under transformation F:

edges	$d\mathbf{x}'^{(n)} = R\, d\mathbf{x}^{(n)}$	or	$[d\mathbf{x}'^{(n)}]^i = R^i{}_j\, [d\mathbf{x}^{(n)}]^j$	// ordinary vector
areas	$d\mathbf{A}'^n = J\, R\, d\mathbf{A}^n$	or	$(d\mathbf{A}'^n)^i = J\, R^i{}_j\, (d\mathbf{A}^n)^j$	// vector density W = -1
volume	$dV' = J\, dV$			// scalar density W = -1

(8.4.d.8)

(e) Covariant Magnitudes

The x'-space magnitudes here are the "covariant" ones which are associated with the Curvilinear View of x'-space, as discussed above. Since $d\mathbf{x}^{(n)}$ is a vector, it follows that

$$|\, d\mathbf{x}'^{(n)} |^2 = d\mathbf{x}'^{(n)} \bullet d\mathbf{x}'^{(n)} = d\mathbf{x}^{(n)} \bullet d\mathbf{x}^{(n)} = |\, d\mathbf{x}^{(n)} |^2 \qquad \Rightarrow \qquad |\, d\mathbf{x}'^{(n)} | = |\, d\mathbf{x}^{(n)} |\ . \qquad (8.4.e.1)$$

Since $d\mathbf{A}^n$ is a vector density of weight -1, if follows that,

$$|\, d\mathbf{A}'^n |^2 = d\mathbf{A}'^n \bullet d\mathbf{A}'^n = J^2\, d\mathbf{A}^n \bullet d\mathbf{A}^n = J^2 |\, d\mathbf{A}^n |^2 \qquad \Rightarrow \qquad |\, d\mathbf{A}'^n\, | = |J|\ |\, d\mathbf{A}^n\, | \qquad (8.4.e.2)$$

where we note that $d\mathbf{A}^n \bullet d\mathbf{A}^n$, being a combination of two weight -1 vector densities, is a scalar density of weight -2 and thus transforms as shown above. For completeness, we can add from the above table,

$$V' = J\, dV \qquad \Rightarrow \qquad |\, dV'\, | = |J|\ |\, dV\, |\ . \qquad (8.4.e.3)$$

We now gather equations from above and take their absolute values on the right. Note from (7.18.1) that $|\mathbf{e}_n| = h'_n$ and $|\mathbf{e}^n| = \sqrt{g'^{nn}}$.

$$\begin{array}{llll}
d\mathbf{x}^{(n)} = \mathbf{e}_n\, dx'_n & \mathit{(8.2.1)} & \Rightarrow |d\mathbf{x}^{(n)}| = |\mathbf{e}_n|\, dx'_n = h'_n\, dx'_n & \\
d\mathbf{x}'^{(n)} = \mathbf{e}'_n\, dx'_n & \mathit{(8.2.1)} & \Rightarrow |d\mathbf{x}'^{(n)}| = |\mathbf{e}'_n|\, dx'_n = h'_n\, dx'_n & \\
 & & & \\
d\mathbf{A}^n = J\, (\Pi_{i\neq n} dx'^i)\, \mathbf{e}^n & \mathit{(8.4.a.1)} & \Rightarrow |d\mathbf{A}^n| = |J|\, |\mathbf{e}^n|\, (\Pi_{i\neq n} dx'^i) = |J|\sqrt{g'^{nn}}\, (\Pi_{i\neq n} dx'^i) & \\
d\mathbf{A}'^n = J^2 (\Pi_{i\neq n} dx'^i)\, \mathbf{e}'^n & \mathit{(8.4.c.5)} & \Rightarrow |d\mathbf{A}'^n| = J^2\, |\mathbf{e}'^n|\, (\Pi_{i\neq n} dx'^i) = J^2\sqrt{g'^{nn}}\, (\Pi_{i\neq n} dx'^i) & \\
 & & & \\
dV = J\, (\Pi_i dx'^i) & \mathit{(8.4.a.1)} & \Rightarrow |dV| = |J|\, (\Pi_i dx'^i) & \\
dV' = J^2\, (\Pi_i dx'^i) & \mathit{(8.4.d.4)} & \Rightarrow |dV'| = J^2\, (\Pi_i dx'^i)\ . & (8.4.e.4)
\end{array}$$

The above can be written more compactly using the Cartesian-View coordinate variation groupings shown in (8.2.3),

$$\begin{array}{llll}
d\mathbf{x}^{(n)} = \mathbf{e}_n\, d\mathcal{L}'^n & \Rightarrow |d\mathbf{x}^{(n)}| = h'_n\, d\mathcal{L}'^n & d\mathcal{L}'^n \equiv dx'^n & \\
d\mathbf{x}'^{(n)} = \mathbf{e}'_n\, d\mathcal{L}'^n & \Rightarrow |d\mathbf{x}'^{(n)}| = h'_n\, d\mathcal{L}'^n & & \\
 & & & \\
d\mathbf{A}^n = J\, d\mathcal{A}'^n\, \mathbf{e}^n & \Rightarrow |d\mathbf{A}^n| = (|J|\sqrt{g'^{nn}}\,)\, d\mathcal{A}'^n & d\mathcal{A}'^n \equiv \Pi_{i\neq n} dx'^i & \\
d\mathbf{A}'^n = J^2\, d\mathcal{A}'^n\, \mathbf{e}'^n & \Rightarrow |d\mathbf{A}'^n| = (J^2\sqrt{g'^{nn}}\,)\, d\mathcal{A}'^n & & \\
 & & & \\
dV = J\, d\mathcal{V}' & \Rightarrow |dV| = |J|\, d\mathcal{V}' & d\mathcal{V}' \equiv \Pi_i dx'^i & \\
dV' = J^2\, d\mathcal{V}' & \Rightarrow |dV'| = J^2\, d\mathcal{V}'\ . & & (8.4.e.5)
\end{array}$$

The covariant edge magnitude is of course unchanged by the transformation since it is a scalar, while the area and volume magnitudes are magnified by |J| in going from x-space to x'-space. Since $dV = d\mathbf{A}^n \bullet d\mathbf{x}^{(n)}$, it is clear that the area's transformation factor of |J| is passed onto the volume. Notice that dV' is always positive regardless of the sign of J, and this is because x'-space is always a right-handed coordinate system, as in Section 6.9. In contrast, dV can have either sign depending on the sign of J = det(S), and so dV < 0 when the $\mathbf{e}_n$ form a left-handed coordinate system in x-space.

(f) Two Theorems : $g'^{nn} g' = \mathrm{cof}(g'_{nn})$ and $|(\Pi^x{}_{i\neq n} \mathbf{e}_i)| = \sqrt{\mathrm{cof}(g'_{nn})}$

We now pause to prove two small theorems which will be used below.

Theorem 1:
- $g'^{nn} g' = \mathrm{cof}(g'_{nn})$ where $g' = \det(g'_{ij})$ (8.4.f.1)
- $g'^{nn} g' = g' / h'^2_n$ (orthog only)

For orthogonal coordinates $g'^{nn} = 1/h'^2_n$ as shown in (5.11.9), proving the second item The first equality can be shown as follows. First, define these two regular matrices,

$$(g'_{up})_{ab} \equiv g'^{ab} \qquad (g'_{dn})_{ab} \equiv g'_{ab}\ .$$

Then since these metric tensors are inverses, use fact that $A^{-1} = \mathrm{cof}(A^T)/\det A$ to get

$g'_{\mathtt{up}} = (g'_{\mathtt{dn}})^{-1} = \mathrm{cof}(g'_{\mathtt{dn}}{}^{\mathtt{T}})/\det(g'_{\mathtt{dn}}) = \mathrm{cof}(g'_{\mathtt{dn}})/\det(g'_{\mathtt{dn}})$ // $g'_{\mathtt{dn}}$ symmetric

$\Rightarrow\ (g'_{\mathtt{up}})_{\mathtt{nn}} = \mathrm{cof}[(g'_{\mathtt{dn}})_{\mathtt{nn}}]/\det(g'_{\mathtt{dn}})$

or

$g'^{\mathtt{nn}} = \mathrm{cof}[g'_{\mathtt{nn}}] / g'$ QED

Theorem 2: $|(\Pi^{\mathtt{x}}{}_{\mathtt{i}\neq\mathtt{n}}\mathbf{e}_{\mathtt{i}})| = \sqrt{\mathrm{cof}(g'_{\mathtt{nn}})}$ (8.4.f.2)

The quantity on the left is this

$|(\Pi^{\mathtt{x}}{}_{\mathtt{i}\neq\mathtt{n}}\mathbf{e}_{\mathtt{i}})| \equiv |\ \mathbf{e}_1 \text{ x } \mathbf{e}_2 \ ... \text{ x } \mathbf{e}_N\ |$ where $\mathbf{e}_{\mathtt{n}}$ is missing from cross product. (8.4.f.3)

We shall give three quick proofs of Theorem 2, the last being valid only for N=3.

- First, one form for $d\mathbf{A}^{\mathtt{n}}$ from (8.4.a.1) is this,

$d\mathbf{A}^{\mathtt{n}} = (-1)^{\mathtt{n}-1}\ (\Pi^{\mathtt{x}}{}_{\mathtt{i}\neq\mathtt{n}}\mathbf{e}_{\mathtt{i}})\ (\Pi_{\mathtt{i}\neq\mathtt{n}}dx'^{\mathtt{i}}) = (-1)^{\mathtt{n}-1}\ (\Pi^{\mathtt{x}}{}_{\mathtt{i}\neq\mathtt{n}}\mathbf{e}_{\mathtt{i}})\ d\mathcal{A}'^{\mathtt{n}}$

$\Rightarrow |d\mathbf{A}^{\mathtt{n}}| = |(\Pi^{\mathtt{x}}{}_{\mathtt{i}\neq\mathtt{n}}\mathbf{e}_{\mathtt{i}})|\ d\mathcal{A}'^{\mathtt{n}}$. (8.4.f.4)

Comparison of this last result with the third line of (8.4.e.5) shows that the following must be true.

$|(\Pi^{\mathtt{x}}{}_{\mathtt{i}\neq\mathtt{n}}\mathbf{e}_{\mathtt{i}})| = |J|\sqrt{g'^{\mathtt{nn}}} = g'^{1/2}\sqrt{g'^{\mathtt{nn}}} = \sqrt{\mathrm{cof}(g'_{\mathtt{nn}})}$. QED (8.4.f.5)

Note that since g=1 (Cartesian x-space), $J^2 = g'$ from (5.12.14), and the last step follows from Theorem 1.

- Here is a more direct proof: [$\Pi^{\mathtt{x}}{}_{\mathtt{i}\neq\mathtt{n}}\ (\mathbf{e}_{\mathtt{i}})$ is a vector in Cartesian x-space]

$|\ \Pi^{\mathtt{x}}{}_{\mathtt{i}\neq\mathtt{n}}\ (\mathbf{e}_{\mathtt{i}})|^2 = \Pi^{\mathtt{x}}{}_{\mathtt{i}\neq\mathtt{n}}\ (\mathbf{e}_{\mathtt{i}})\ \bullet\ \Pi^{\mathtt{x}}{}_{\mathtt{j}\neq\mathtt{n}}\ (\mathbf{e}_{\mathtt{j}}) = [\ \Pi^{\mathtt{x}}{}_{\mathtt{i}\neq\mathtt{n}}\ (\mathbf{e}_{\mathtt{i}})]_{\mathtt{k}}\ [\ \Pi^{\mathtt{x}}{}_{\mathtt{j}\neq\mathtt{n}}\ (\mathbf{e}_{\mathtt{j}})]_{\mathtt{k}}$ (8.4.f.6)

$= [\ \varepsilon_{\mathtt{kabc...x}}\ (\mathbf{e}_1)^{\mathtt{a}}(\mathbf{e}_2)^{\mathtt{b}}\ ......\ (\mathbf{e}_N)_{\mathtt{x}}]\ [\ \varepsilon_{\mathtt{ka'b'c'...x}}\ (\mathbf{e}_1)^{\mathtt{a'}}(\mathbf{e}_2)^{\mathtt{b'}}\ ......\ (\mathbf{e}_N)^{\mathtt{x'}}]$ // $\mathbf{e}_{\mathtt{n}}$ missing in both

$= \varepsilon_{\mathtt{kabc...x}}\ \varepsilon_{\mathtt{ka'b'c'...x}}\ \{(e_1)^{\mathtt{a}}(e_2)^{\mathtt{b}}\ ......\ (e_N)_{\mathtt{x}}\ \}\ (e_1)^{\mathtt{a'}}(e_2)^{\mathtt{b'}}\ ......\ (e_N)^{\mathtt{x'}}$ // $e_{\mathtt{n}}$ missing

$= \mathbf{e}_1\bullet\mathbf{e}_1\ \mathbf{e}_2\bullet\mathbf{e}_2\ ....\ \mathbf{e}_N\bullet\mathbf{e}_N$ + all signed permutations of the 2nd labels // $\mathbf{e}_{\mathtt{n}}$ missing

$= g'_{11}g'_{22}.....\ g'_{NN}$ + all signed permutations of the 2nd indices, *(7.18.1)* // $\mathbf{e}_{\mathtt{n}}$ missing

But this is last object is the determinant of the $g'_{\mathtt{ij}}$ matrix with $g'_{\mathtt{nn}}$ crossed out, which is to say, it is the minor of $g'_{\mathtt{nn}}$. Since $g'_{\mathtt{nn}}$ is a diagonal element, the minor and cofactor are the same. Thus, this last object is in fact just $\mathrm{cof}(g'_{\mathtt{nn}})$. QED.

- A proof for N=3 uses normal vector algebra. Setting n = 1, for example, one needs to show that

$$| \ \Pi^{x}{}_{i\neq 1}\ \mathbf{e}_i \ |^2 = | \ \mathbf{e}_2 \text{ x } \mathbf{e}_3 \ |^2 \ = \ \mathrm{cof}(g'_{11}) \ \ .$$

To this end, use the vector identity

$$(\mathbf{A} \text{ x } \mathbf{B}) \bullet (\mathbf{A} \text{ x } \mathbf{B}) \ = A^2B^2 - (\mathbf{A}\bullet\mathbf{B})^2$$

and fact (7.18.1) that $\mathbf{e}_n \bullet \mathbf{e}_m = g'_{nm}$ to show that,

$$| \ \mathbf{e}_2 \text{ x } \mathbf{e}_3 \ |^2 \ = (\mathbf{e}_2 \text{ x } \mathbf{e}_3) \bullet (\mathbf{e}_2 \text{ x } \mathbf{e}_3 \ ) \ = \ |\mathbf{e}_2|^2 \ |\mathbf{e}_3|^2 \ - (\mathbf{e}_2\bullet\mathbf{e}_3)^2 \ \ = g'_{22} \ g'_{33} - (g'_{23})^2 \ = \mathrm{cof}(g'_{11}). \qquad (8.4.f.7)$$

and the cases n = 2 and 3 are similar.

(g) Cartesian-View Magnitude Ratios

In the Cartesian View of x'-space one can write the Cartesian x'-space magnitudes as

$$| \ d\mathbf{x}'^{(n)} \ |_c \ = d\mathcal{L}'^n \qquad | \ d\mathbf{A}'^{(n)} \ |_c \ = d\mathcal{A}'^n \qquad | \ dV' \ |_c \ = d\mathcal{V}' \ . \qquad (8.4.g.1)$$

Then from the three x-space magnitude equations in (8.4.e.5) one obtains the following three ratios of x-space objects divided by their corresponding Cartesian-View x'-space objects:

$$| \ d\mathbf{x}^{(n)} \ |/ \ d\mathcal{L}'^n = h'_n \ = \ [g'_{nn}]^{1/2} \ \ = \text{ the scale factor for edge } d\mathbf{x}^{(n)}$$

$$| \ d\mathbf{A}^{(n)} \ |/ \ d\mathcal{A}'^n \ = \sqrt{g'^{nn}} \ \ |J| \ = \sqrt{g'^{nn}} \ \ g'^{1/2} = [\ g'^{nn} \ g']^{1/2} \ = [\mathrm{cof}(g'_{nn})]^{1/2} \qquad // \text{ Theorem 1 above}$$

$$|dV| \ / \ d\mathcal{V}' = |J| = g'^{1/2} \ \ . \qquad // \ g' \equiv \det(g'_{ij}) \ = J^2 \qquad (8.4.g.2)$$

It is convenient to make the definition

$$dA^n \ \equiv \ | \ d\mathbf{A}^{(n)} \ | \qquad (8.4.g.3)$$

and then the above area magnitude ratio relation may be written

$$dA^n = \sqrt{\mathrm{cof}(g'_{nn})} \ \ d\mathcal{A}'^n \qquad (8.4.g.4)$$

and $d\mathcal{A}'^n = \Pi_{i\neq n}dx'^i$ is just a product of the appropriate curvilinear coordinate variations. Since g'_{nn} is a diagonal element of the g' matrix, one can reexpress the cofactor in the above equation in this manner:

$$\mathrm{cof}(g'_{nn}) = \mathrm{minor}(g'_{nn}) \ = \det \ (g' \text{ with row n and column n crossed out}) \qquad (8.4.g.5)$$

On the other hand, one can regard the number $\mathrm{cof}(g'_{nn})$ as one element of the cofactor matrix cof(g') whose elements are given by $[\mathrm{cof}(g')]_{ij} \equiv \ \mathrm{cof}(g'_{ij})$. Then,

$$\mathrm{cof}(g'_{nn}) \ = \ [\mathrm{cof}(g')]_{nn} \ \ . \qquad (8.4.g.6)$$

(h) Nested Cofactor Formulas and S^TS notation

Descending Leaf Hierarchy

The object dA^n is the "area" of a face on a differential N-piped spanned by the full set of vectors $\mathbf{e}_i$ but not including $\mathbf{e}_n$. This face, which is itself an (N-1)-piped, in turn has its own "areas" which are (N-2)-pipeds, and so on, so there is a hierarchy of "areas" of dimensions N-1 all the way down. The area ratios of corresponding areas under transformation F are determined by equations similar to (8.4.g.4). For this purpose we define Cof to be a matrix and cof to be a number,

$$\mathrm{Cof}(M_{ij}) \equiv \text{the submatrix of M obtained by crossing out row i and column j of M}$$

$$\mathrm{cof}(M_{ij}) = (-1)^{i-j}\,\mathrm{minor}(M_{ij}) = (-1)^{i-j}\,\det[\mathrm{Cof}(M_{ij})] = \text{the usual "cofactor"} . \qquad (8.4.h.1)$$

Then

$$\mathrm{Cof}(g'_{nn}) = \text{the submatrix of g' obtained by crossing out row n and column n of g'}$$

$$\mathrm{cof}(g'_{nn}) = \mathrm{minor}(g'_{nn}) = \det[\mathrm{Cof}(g'_{nn})] . \qquad (8.4.h.2)$$

Then for example the m^{th} face of face n of an N-piped has area ratio $\sqrt{\mathrm{cof}[\mathrm{Cof}(g'_{nn})]_{mm}}$. The object inside the radical is the cofactor of the m,m element of the N-1 x N-1 matrix $\mathrm{Cof}(g'_{nn})$. If we refer to this area as $dA^{n,m}$ we can go down the hierarchy in this manner

$$dA^n = \sqrt{\mathrm{cof}(g'_{nn})}\ \Pi_{i\neq n}dx'^i \qquad \text{face n}$$

$$dA^{n,m} = \sqrt{\mathrm{cof}[\mathrm{Cof}(g'_{nn})]_{mm}}\ \Pi_{i\neq n,m}dx'^i \qquad \text{face m of face n}$$

$$dA^{n,m,k} = \sqrt{\mathrm{cof}(\mathrm{Cof}[\mathrm{Cof}(g'_{nn})]_{mm})_{kk}}\ \Pi_{i\neq n,m,k}dx'^i \qquad \text{face k of face m of face n} \qquad (8.4.h.3)$$

and so on. As a very simple example, for N = 3 the faces $dA^{n,m}$ are line segments dx'^i and one has

$$\mathrm{Cof}(g'_{33}) = \begin{pmatrix} g'_{11} & g'_{12} \\ g'_{21} & g'_{22} \end{pmatrix} \Rightarrow \mathrm{cof}[\mathrm{Cof}(g'_{33})]_{22} = g'_{11} = h'^2_1 \Rightarrow \sqrt{\mathrm{cof}[\mathrm{Cof}(g'_{33})]_{22}} = h'_1 \qquad (8.4.h.4)$$

and $dA^{3,2} = h'_1$ is in fact the edge length ratio given above in (8.4.g.2).

There is another way to write these area magnitudes if we assume that x-space is Cartesian, but it is difficult to express in standard notation, so we momentarily revert to our developmental notation. In that notation, if x-space is Cartesian we can write from (7.5.6) and (3.2.7) that $\overline{g}' = S^TS$,

$$\overline{g}' = S^T \overline{g} S = S^T S = \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \dots \\ \mathbf{e}_N \end{pmatrix} (\mathbf{e}_1, \mathbf{e}_2, \dots \mathbf{e}_N) = \begin{bmatrix} \mathbf{e}_1 \bullet \mathbf{e}_1 & \mathbf{e}_1 \bullet \mathbf{e}_2 & \mathbf{e}_1 \bullet \mathbf{e}_3 & \dots\dots & \mathbf{e}_1 \bullet \mathbf{e}_N \\ \mathbf{e}_2 \bullet \mathbf{e}_1 & \mathbf{e}_2 \bullet \mathbf{e}_2 & \mathbf{e}_2 \bullet \mathbf{e}_3 & \dots\dots & \mathbf{e}_2 \bullet \mathbf{e}_N \\ \mathbf{e}_3 \bullet \mathbf{e}_1 & \mathbf{e}_3 \bullet \mathbf{e}_2 & \mathbf{e}_3 \bullet \mathbf{e}_3 & \dots\dots & \mathbf{e}_3 \bullet \mathbf{e}_N \\ \dots\dots\dots & & & & \\ \mathbf{e}_N \bullet \mathbf{e}_1 & \mathbf{e}_N \bullet \mathbf{e}_2 & \mathbf{e}_N \bullet \mathbf{e}_3 & \dots\dots & \mathbf{e}_N \bullet \mathbf{e}_N \end{bmatrix} \qquad (8.4.h.5)$$

as was written earlier in (5.11.3). Since the covariant metric tensor $\overline{g}'_{ij}$ in developmental notation is equal to g'_{ij} in standard notation, we shall use the standard notation below in expressing these tensor elements.

The area (magnitude) dA^n of face n of a differential N-piped is, from (8.4.g.4) and (8.4.e.5),

$$dA^n = \sqrt{\text{cof}(g'_{nn})}\ \Pi_{i\neq n} dx'^i \ . \qquad (8.4.h.6)$$

The object cof (g'_{nn}) is the minor of the above bracketed matrix with row n and column n crossed out. That reduced matrix can in fact be written

$$\text{Cof}(g'_{nn}) = S^{(n)T} S^{(n)}) = \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ .. \\ \mathbf{e}_N \end{pmatrix} (\mathbf{e}_1, \mathbf{e}_2, \dots \mathbf{e}_N) \quad \text{where } \mathbf{e}_n \text{ is missing from both vectors .} \qquad (8.4.h.7)$$

so

$$\text{cof}(g'_{nn}) = \det[S^{(n)T} S^{(n)}] = \det \left[\begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ .. \\ \mathbf{e}_N \end{pmatrix} (\mathbf{e}_1, \mathbf{e}_2, \dots \mathbf{e}_N) \right] \ . \qquad (8.4.h.8)$$

Therefore,

$$dA^n = \sqrt{\text{cof}(g'_{nn})}\ \Pi_{i\neq n} dx'^i$$

$$= \sqrt{\det[S^{(n)T} S^{(n)}]}\ \Pi_{i\neq n} dx'^i$$

$$\text{where } S^{(n)} = (\mathbf{e}_1, \mathbf{e}_2 \dots \mathbf{e}_N) \text{ with } \mathbf{e}_n \text{ missing .} \qquad (8.4.h.9)$$

The same argument results in

$$dA^{n,m} = \sqrt{\text{cof}[\text{Cof}(g'_{nn})]_{mm}}\ \Pi_{i\neq n,m} dx'^i$$

$$= \sqrt{\det[S^{(n,m)T} S^{(n,m)}]}\ \Pi_{i\neq n,m} dx'^i$$

$$\text{where } S^{(n,m)} = (\mathbf{e}_1, \mathbf{e}_2 \dots \mathbf{e}_N) \text{ with } \mathbf{e}_n \text{ and } \mathbf{e}_m \text{ missing .} \qquad (8.4.h.10)$$

Continuing down one more level,

$$dA^{n,m,k} = \sqrt{\mathrm{cof}(\mathrm{Cof}[\mathrm{Cof}(g'_{nn})]_{mm})_{kk}}\ \Pi_{i\neq n,m,k}dx'^i$$

$$= \sqrt{\det[S^{(n,m,k)T}S^{(n,m,k)}]}\ \Pi_{i\neq n,m,k}dx'^i$$

$$\text{where } S^{(n,m,k)} = \ (\mathbf{e}_1,\mathbf{e}_2....\mathbf{e}_N) \text{ with } \mathbf{e}_n, \mathbf{e}_m, \mathbf{e}_k \text{ all missing .} \qquad (8.4.h.11)$$

In this descending hierarchy of n-piped "leaf" areas, we eliminate one tangent base vector $\mathbf{e}_i$ at a time, and the eliminated vectors' labels are used to label the "leaf" whose area is indicated, as in $S^{(n,m)}$.

Ascending Leaf Hierarchy

One can alternatively build up the hierarchy of areas from the bottom, for example using $S^{[n.m]}$ to indicate the area of the 2-piped spanned by $\mathbf{e}_n$ and $\mathbf{e}_m$ in R^N where the top level N-piped lives. One then has

$$dA^{[n]} = \sqrt{\det(S^{[n]T}S^{[n]})}\ dx'^n \text{ where } S^{[n]} = \mathbf{e}_n . \qquad (8.4.h.12)$$

In this case $S^{[n]T}\, S^{[n]} = \mathbf{e}_n \bullet \mathbf{e}_n = h'_n{}^2$ so $dA^{[n]} = h'_n dx'^n$ as expected. Next, for 2-pipeds,

$$dA^{[n.m]} = \sqrt{\det(S^{[n.m]T}S^{[n.m]})}\ dx'^n dx'^m \text{ where } S^{[n.m]} = (\mathbf{e}_n, \mathbf{e}_m) . \qquad (8.4.h.13)$$

In this case

$$\det[S^{[n.m]T}S^{[n.m]}] = \det\ [\begin{pmatrix} \mathbf{e}_n \\ \mathbf{e}_m \end{pmatrix}(\mathbf{e}_n, \mathbf{e}_m)] = \det\begin{pmatrix} \mathbf{e}_n\bullet\mathbf{e}_n & \mathbf{e}_n\bullet\mathbf{e}_m \\ \mathbf{e}_m\bullet\mathbf{e}_n & \mathbf{e}_m\bullet\mathbf{e}_m \end{pmatrix} = \det\begin{pmatrix} g'_{nn} & g'_{nm} \\ g'_{mn} & g'_{mm} \end{pmatrix}$$

$$= g'_{nn}g'_{mm} - (g'_{nm})^2 = h'_n{}^2\, h'_m{}^2 - (\mathbf{e}_n\bullet\mathbf{e}_m)^2 \qquad (8.4.h.14)$$

so

$$dA^{[n.m]} = \sqrt{h'_n{}^2\, h'_m{}^2 - (\mathbf{e}_n\bullet\mathbf{e}_m)^2}\ dx'^n dx'^m . \qquad (8.4.h.15)$$

Doing one more step, at the level of 3-pipeds in the hierarchy buildup one would then have

$$dA^{[n.m.k]} = \sqrt{\det[S^{[n.m.k]T}S^{[n.m.k]}]}\ dx'^n dx'^m dx'^k \text{ where } S^{[n.m.k]} = (\mathbf{e}_n, \mathbf{e}_m, \mathbf{e}_k) \qquad (8.4.h.16)$$

and now

$$\det[S^{[n.m.k]T}S^{[n.m.k]}] = \det\ [\begin{pmatrix} \mathbf{e}_n \\ \mathbf{e}_m \\ \mathbf{e}_k \end{pmatrix}(\mathbf{e}_n, \mathbf{e}_m, \mathbf{e}_k)] = \det\begin{pmatrix} g'_{nn} & g'_{nm} & g'_{nk} \\ g'_{mn} & g'_{mm} & g'_{mk} \\ g'_{kn} & g'_{km} & g'_{kk} \end{pmatrix} . \qquad (8.4.h.17)$$

Comment 1: Notice that the matrices S appearing above in the above S^TS structures are in general not square matrices because the number of components N of the $\mathbf{e}_i$ vectors in R^N generally does not match the number of $\mathbf{e}_i$ vectors in the list which defines S. On the other hand, S^TS is always a square matrix and therefore always has a determinant.

Comment 2: We have used the term "leaf" to refer to any n-piped appearing in the top level N-piped, with n then ranging from N down to 1 (which leaf is just point). In a more precise discussion of this subject, these leaves are called "boundaries" and in addition to having "area", these boundaries have "orientation". We have only discussed the orientation of the top level leaves by using the vector area $d\mathbf{A}^n$. Usually the precise discussion is couched in the language of "differential forms", and by careful consideration of boundaries and their orientations one is able to derive the generalized Stokes's theorem which reads

$$\int_{\mathbf{M}} d\alpha = \int_{\partial\mathbf{M}} \alpha \qquad (8.4.h.18)$$

where α is a differential n-form, dα is an (n+1)-form which is the "exterior derivative" of α, M is a smooth type of surface called a manifold, and ∂M is the boundary of that manifold. This abstract theorem then encompasses all the well-known integral theorems of analysis in any number of dimensions n. See Sjamaar [2015] Sections 5.2 - 5.4 , Section 8.2, and Theorem 9.9 on page 117. The fact that $\sqrt{\det(S^TS)}$ is the area scaling factor appears as Theorem 8.4 on Sjamaar page 101 where it is called $\sqrt{\det(A^TA)}$. See also Appendix F (The Volume of an n-piped embedded in R^m) of Lucht *Tensor Products*.

(i) Transformation of arbitrary differential vectors, areas and volume

The above discussion is geared to the N-piped transform picture and reveals how $d\mathbf{x}^{(n)}$, $d\mathbf{A}^{(n)}$ and dV transform where all these quantities are directly associated with the particular differential N-piped in x-space spanned by the $(\mathbf{e}_n dx'^n)$ vectors.

But suppose d**x** is an *arbitrary* differential vector in x-space. Certainly $d\mathbf{x}' = R\, d\mathbf{x}$, so we know how this d**x** transforms under F. But what about area and volume?

Based on the work in Appendix B as carried through into the N-piped transform discussion above, it seems clear that an *arbitrary* differential area in x-space d**A** can be *represented* as follows : (g=1)

$$d\mathbf{A} = (d\mathbf{x}^{[1]}) \text{ x } (d\mathbf{x}^{[2]}) \ldots \text{ x } (d\mathbf{x}^{[N-1]}) \qquad (8.4.i.1)$$

$$(d\mathbf{A})_i = \varepsilon_{iabc..x} (d\mathbf{x}^{[1]})^a (d\mathbf{x}^{[2]})^b \ldots . (d\mathbf{x}^{[N-1]})^x \quad , \qquad (8.4.i.2)$$

where the $d\mathbf{x}^{[i]}$ are an arbitrary set of N-1 linearly independent differential vectors in x-space. For N=3 one would write $d\mathbf{A} = d\mathbf{x}^{[1]} \text{ x } d\mathbf{x}^{[2]}$. Since ε has weight -1 and all other objects are true vectors (weight 0), one again concludes that d**A** transforms as a tensor density of weight -1, so

$$d\mathbf{A}' = J\, R\, d\mathbf{A} \qquad \text{or} \qquad dA'_i = J\, R_i{}^j\, dA_j \ . \qquad (8.4.i.3)$$

If any of these $d\mathbf{x}^{[k]}$ were a linear combination of the N-2 others, one could say $d\mathbf{x}^{[k]} = \Sigma_j \alpha^k{}_j\, d\mathbf{x}^{[j]}$ and then the above $(d\mathbf{A})_i$ expression would give a sum of terms each of which vanishes by symmetry, resulting in $(d\mathbf{A})_i = 0$.

Finally, given the set of $d\mathbf{x}^{[i]}$ linearly independent vectors shown above for i = 1,2..N-1, we can certainly find one more such that all N are then linearly independent, so then we have a set of N arbitrary differential vectors $d\mathbf{x}^{[i]}$(arbitrary as long as they are linearly independent), and these will form a volume in x-space,

$$dV = \det(d\mathbf{x}^{[1]}, d\mathbf{x}^{[2]}, ... d\mathbf{x}^{[N]}) = \varepsilon_{abc..y}(d\mathbf{x}^{[1]})^a(d\mathbf{x}^{[2]})^b....(d\mathbf{x}^{[N]})^y . \quad (8.4.i.4)$$

By the argument just given, inspection shows that this dV transforms as a tensor density of weight -1 so dV' = JdV.

These then are the transformation rules for arbitrary differential vectors, areas and volumes transforming under F :

$$\begin{array}{lll} d\mathbf{x}' = R\, d\mathbf{x} & |d\mathbf{x}'| = |d\mathbf{x}| & \\ d\mathbf{A}' = J\, R\, d\mathbf{A} & |dA'| = |J|\,|dA| & \\ dV' = JdV & |dV'| = |J|\,|dV| \ . & J = \det(S) \qquad J^2 = g' \end{array} \quad (8.4.i.5)$$

Review and covariant form of the d**A** and dV equations

To review, in Cartesian x-space we have these expressions for area and volume

$$\begin{array}{l} d\mathbf{A} = (d\mathbf{x}^{[1]}) \times (d\mathbf{x}^{[2]}) ... \times (d\mathbf{x}^{[N-1]}) \\ dV = \det[d\mathbf{x}^{[1]}, d\mathbf{x}^{[2]}, ... d\mathbf{x}^{[N]}] \quad . \end{array} \quad (8.4.i.6)$$

Written out in terms of contravariant vector components $(d\mathbf{x}^{[i]})^j$ these expressions appear as,

$$\begin{array}{l} (d\mathbf{A})_i = \varepsilon_{iabc..x}\,(d\mathbf{x}^{[1]})^a(d\mathbf{x}^{[2]})^b....(d\mathbf{x}^{[N-1]})^x \\ dV = \varepsilon_{abc..y}\,(d\mathbf{x}^{[1]})^a(d\mathbf{x}^{[2]})^b....(d\mathbf{x}^{[N]})^y \ , \end{array} \quad (8.4.i.7)$$

where ε is the usual permutation tensor involved in cross products and determinants. As noted earlier, these are both "valid tensor density equations" as in (7.15.9), so they can be written in x'-space as

$$\begin{array}{l} (d\mathbf{A}')_i = \varepsilon'_{iabc..x}\,(d\mathbf{x}'^{[1]})^a(d\mathbf{x}'^{[2]})^b....(d\mathbf{x}'^{[N-1]})^x \\ dV' = \varepsilon'_{abc..y}\,(d\mathbf{x}'^{[1]})^a(d\mathbf{x}'^{[2]})^b....(d\mathbf{x}'^{[N]})^y \ . \end{array} \quad (8.4.i.8)$$

where $d\mathbf{x}'^{[i]} = Rd\mathbf{x}^{[i]}$ and where ε' is the x'-space Levi-Civita tensor. From (D.5.13), each ε' can be written as $\varepsilon' = J^2\varepsilon = (g'/g)\varepsilon = g'\varepsilon$, so

$$\begin{array}{l} (d\mathbf{A}')_i = g'\,\varepsilon_{iabc..x}\,(d\mathbf{x}'^{[1]})^a(d\mathbf{x}'^{[2]})^b....(d\mathbf{x}'^{[N-1]})^x \\ dV' = g'\,\varepsilon_{abc..y}\,(d\mathbf{x}'^{[1]})^a(d\mathbf{x}'^{[2]})^b....(d\mathbf{x}'^{[N]})^y \ . \end{array} \quad (8.4.i.9)$$

Since the permutation tensor now appears, these can be written in terms of cross products,

$$\begin{array}{l} d\mathbf{A}' = g'\,(d\mathbf{x}'^{[1]}) \times (d\mathbf{x}'^{[2]}) ... \times (d\mathbf{x}'^{[N-1]}) \\ dV' = g' \det[d\mathbf{x}'^{[1]}, d\mathbf{x}'^{[2]}, ... d\mathbf{x}'^{[N]}] \quad . \end{array} \quad (8.4.i.10)$$

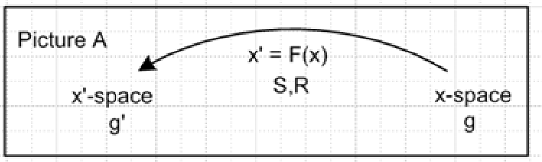
Using contravariant components $[d\mathbf{x}'^{[1]}]^i$, this cross product and determinant are formed just as they are in x-space. These two equations then give at least some feel for "the meaning of d**A**' and dV' in x'-space" .

Going back to x-space, we could have written the equations there using $g = \det(g'_{ij}) = \det(\delta_{ij}) = 1$:

$$d\mathbf{A} = g\,(d\mathbf{x}^{[1]}) \times (d\mathbf{x}^{[2]}) \ldots \times (d\mathbf{x}^{[N-1]})$$
$$dV = g \det (d\mathbf{x}^{[1]}, d\mathbf{x}^{[2]}, \ldots d\mathbf{x}^{[N]}) \quad . \qquad (8.4.i.11)$$

This then is a useful interpretation of the covariant form of these equations. Adding primes to everything in the above two equations yields the previous two equations and only the permutation tensor ε is involved in both sets of equations.

With this understanding, the above transformation rules for differential vectors, areas and volumes can be extended from Picture B to the more general Picture A, where g is an arbitrary metric tensor,



How things look in developmental notation.

Recall that covariant tensor objects get overbars and all indices are down in the developmental notation used in Chapters 1-6 of this document. Here then are some of the above equations expressed in this notation:

$$\overline{d\mathbf{A}} = (d\mathbf{x}^{[1]}) \times (d\mathbf{x}^{[2]}) \ldots \times (d\mathbf{x}^{[N-1]})$$
$$dV = \det [d\mathbf{x}^{[1]}, d\mathbf{x}^{[2]}, \ldots d\mathbf{x}^{[N]}] \qquad (8.4.i.12)$$

$$(\overline{d\mathbf{A}})_i = \overline{\varepsilon}_{iabc..x} \; (d\mathbf{x}^{[1]})_a (d\mathbf{x}^{[2]})_b \ldots. \; (d\mathbf{x}^{[N-1]})_x$$
$$dV = \overline{\varepsilon}_{abc..y} \; (d\mathbf{x}^{[1]})_a (d\mathbf{x}^{[2]})_b \ldots. \; (d\mathbf{x}^{[N]})_y \qquad (8.4.i.13)$$

$$(\overline{d\mathbf{A}'})_i = \overline{\varepsilon}'_{iabc..x} \; (d\mathbf{x}'^{[1]})_a (d\mathbf{x}'^{[2]})^b \ldots. \; (d\mathbf{x}'^{[N-1]})^x$$
$$dV' = \overline{\varepsilon}'_{abc..y} \; (d\mathbf{x}'^{[1]})^a (d\mathbf{x}'^{[2]})^b \ldots. \; (d\mathbf{x}'^{[N]})^y \qquad (8.4.i.14)$$

$$d\mathbf{x}' = R\, d\mathbf{x} \qquad |d\mathbf{x}'| = |d\mathbf{x}|$$
$$\overline{d\mathbf{A}'} = J\, S^T\, \overline{d\mathbf{A}} \qquad |\overline{d\mathbf{A}'}| = |J|\, |\overline{d\mathbf{A}}|$$
$$dV' = JdV \qquad |dV'| = |J|\, |dV| \qquad J = \det(S) \qquad J^2 = g' \qquad (8.4.i.15)$$

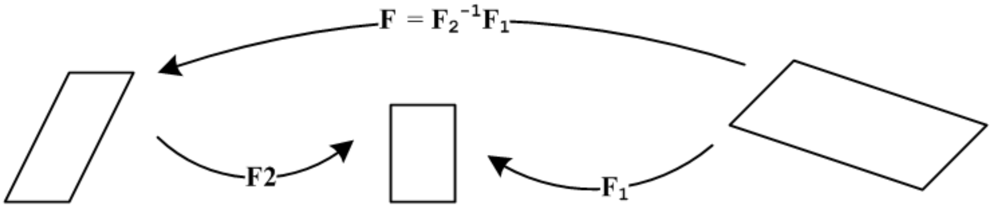
Recall from (2.5.1) that a covariant vector transforms as $\overline{\mathbf{V}}' = S^T\, \overline{\mathbf{V}}$, and a covariant vector density of weight W will then transform as $\overline{\mathbf{V}}' = J^{-W}\, S^T\, \overline{\mathbf{V}}$ and this explains the middle line of the above three. The *contravariant* differential area would transform as $d\mathbf{A}' = J\, R\, d\mathbf{A}$.

(j) Concatenation (Composition) of Transformations

Consider $\mathbf{x}'' = \mathbf{F}_2(\mathbf{x}')$ and $\mathbf{x}' = \mathbf{F}_1(\mathbf{x})$ so that $\mathbf{x}'' = \mathbf{F}_2[\mathbf{F}_1(\mathbf{x})] \equiv \mathbf{F}(\mathbf{x})$. At some point $\mathbf{x}$, the linearization will yield $d\mathbf{x}'' = R_2 R_1 d\mathbf{x}$ as the rule for vector transformation, so the R matrix associated with transformation

F is $R = R_2R_1$, and then $S = R^{-1} = R_1^{-1}R_2^{-1} = S_1S_2$. Each transformation will have an associated Jacobian: $J_1 = \det(S_1)$ and $J_2 = \det(S_2)$. The concatenated transformation then has $J = \det(S) = \det(S_1S_2) = \det(S_1)\det(S_2) = J_1J_2$. The implication is that when one concatenates (composes) two transformations in this manner, the area and volume transformations shown above still apply, where J is taken to be the product of the two underlying Jacobians.

For example, one could consider the mapping between two different skewed N-pipeds, each representing a different curvilinear coordinate system, with our orthogonal N-piped as an intermediary object,



(8.4.j.1)

In this case one has $\mathbf{F} = \mathbf{F}_2^{-1}\mathbf{F}_1$, so $R = R_2^{-1}R_1 = S_2R_1$ and then $S = S_1R_2$ so $J = J_1/J_2$. This J then would be used in the above area and volume transformation rules, for example, $d\mathbf{A}_{\mathbf{left}} = J\,R\,d\mathbf{A}_{\mathbf{right}}$.

In the continuum mechanics flow application, g = 1 on both left and right as well as center, time t_0 is on the right, time t on the left, and the volume transformation is given by $dV_{\mathbf{left}} = J\,dV_{\mathbf{right}}$ where J is associated with the combined F. This J then characterizes the volume change between an initial and final flow particle where each is skewed in some arbitrary manner.

(k) Examples of area magnitude transformation for N = 2,3,4

In (8.4.g.4) it was shown that $dA^n = \sqrt{\mathrm{cof}(g'_{nn})}\ d\boldsymbol{a}'^n$. Since this is a somewhat strange result, some examples are in order. Recall that the dA^n are the areas of the faces of the differential N-piped in x-space, while the $d\boldsymbol{a}'^n$ are the curvilinear coordinate variations one can visualize in the Cartesian-View picture (8.2.2) shown above.

For N=2 the area magnitude transformation results are (for a general non-orthogonal x'-space system)

$$\begin{aligned} dA^1 &= \sqrt{g'_{22}}\ d\boldsymbol{a}'^1 = h'_2\, d\boldsymbol{a}'^1 & \qquad d\boldsymbol{a}'^1 &= dx'^1 = d\mathcal{L}'^1 \\ dA^2 &= \sqrt{g'_{11}}\ d\boldsymbol{a}'^2 = h'_1\, d\boldsymbol{a}'^2 & \qquad d\boldsymbol{a}'^2 &= dx'^2 = d\mathcal{L}'^2\ . \end{aligned} \qquad (8.4.k.1)$$

These equations are simple because the area of a parallelogram "face" is the length of an edge and so these equations just coincide with the length transformation results stated above . Remember that a face is labeled by the index of the vector which does *not* span the face, so h_2' appears in the face 1 equation.

For N=3 the area magnitude transformation results are

$$\begin{aligned} dA^1 &= \sqrt{g'_{22}\ g'_{33} - (g'_{23})^2}\ \ d\boldsymbol{a}'^1 & \qquad d\boldsymbol{a}'^1 &= dx'^2dx'^3 \\ dA^2 &= \sqrt{g'_{33}\ g'_{11} - (g'_{31})^2}\ \ d\boldsymbol{a}'^2 & \qquad d\boldsymbol{a}'^2 &= dx'^3dx'^1 \\ dA^3 &= \sqrt{g'_{11}\ g'_{22} - (g'_{12})^2}\ \ d\boldsymbol{a}'^3 & \qquad d\boldsymbol{a}'^3 &= dx'^1dx'^2\ . \end{aligned} \qquad (8.4.k.2)$$

For an *orthogonal* N=3 system the metric tensor g'_{ab} is diagonal, and then the above simplifies to

$$\begin{array}{llll} dA^1 = \sqrt{g'_{22}\ g'_{33}}\ d\boldsymbol{a}'^1 & = h'_2\ h'_3\, d\boldsymbol{a}'^1 & \qquad d\boldsymbol{a}'^1 = dx'^2 dx'^3 & \\ dA^2 = \sqrt{g'_{33}\ g'_{11}}\ d\boldsymbol{a}'^2 & = h'_3\ h'_1\, d\boldsymbol{a}'^2 & \qquad d\boldsymbol{a}'^2 = dx'^3 dx'^1 & \\ dA^3 = \sqrt{g'_{11}\ g'_{22}}\ d\boldsymbol{a}'^3 & = h'_1\ h'_2\, d\boldsymbol{a}'^3 & \qquad d\boldsymbol{a}'^3 = dx'^1 dx'^2\ . & \qquad (8.4.k.3) \end{array}$$

For an N=4 orthogonal system (g'_{ij} is diagonal with $g'_{ii} = h'_i{}^2$),

$$\begin{array}{llll} dA^1 = \sqrt{\text{cof}(g'_{11})}\ d\boldsymbol{a}'^1 & = h'_2\ h'_3\, h'_4\ d\boldsymbol{a}'^1 & \qquad d\boldsymbol{a}'^1 = dx'^2 dx'^3 dx'^4 & \\ dA^2 = \sqrt{\text{cof}(g'_{22})}\ d\boldsymbol{a}'^2 & = h'_1\ h'_3\, h'_4\ d\boldsymbol{a}'^2 & \qquad d\boldsymbol{a}'^2 = dx'^3 dx'^4 dx'^1 & \\ dA^3 = \sqrt{\text{cof}(g'_{33})}\ d\boldsymbol{a}'^3 & = h'_1\ h'_2\, h'_4\ d\boldsymbol{a}'^3 & \qquad d\boldsymbol{a}'^3 = dx'^4 dx'^1 dx'^2 & \\ dA^4 = \sqrt{\text{cof}(g'_{44})}\ d\boldsymbol{a}'^4 & = h'_1\ h'_2\, h'_3\ d\boldsymbol{a}'^4 & \qquad d\boldsymbol{a}'^4 = dx'^1 dx'^2 dx'^3\ . & \qquad (8.4.k.4) \end{array}$$

Example 2: Spherical Coordinates: area patches

Consider again $dA^n = \sqrt{\text{cof}(g'_{nn})}\ d\boldsymbol{a}'^n$. Since spherical coordinates are orthogonal, the orthogonal N=3 example above may be used. Example 2 of Chapter 5 showed in (5.13.15) that [1,2,3 = r,θ,φ] ,

$$\begin{array}{lll} h'_1 = h'_r\ = 1 & \qquad d\boldsymbol{a}'^1 = dx'^2 dx'^3 = d\theta d\varphi & \\ h'_2 = h'_\theta\ = r & \qquad d\boldsymbol{a}'^2 = dx'^3 dx'^1 = dr d\varphi & \\ h'_3 = h'_\varphi\ = r\sin\theta & \qquad d\boldsymbol{a}'^3 = dx'^1 dx'^2 = dr d\theta & \qquad (8.4.k.5) \end{array}$$

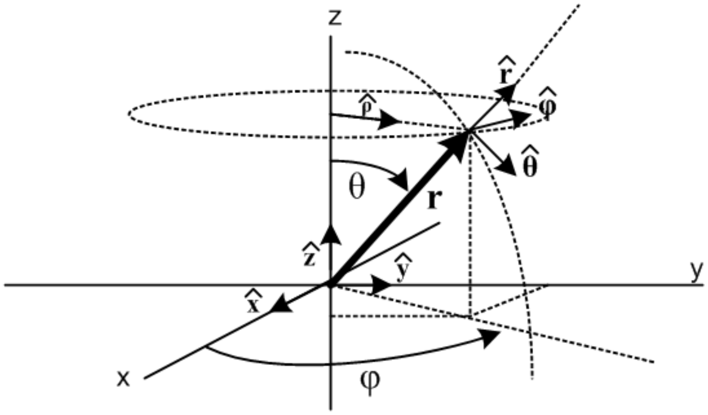
Therefore

$$\begin{array}{llll} d\mathbf{A}^1 = dA^1 \hat{\mathbf{e}}_1 & \Rightarrow\quad d\mathbf{A}^r = dA^r\ \hat{\mathbf{e}}_r\ =\ dA^r\ \hat{\mathbf{r}}\ \text{ with } & dA^r = h'_2\ h'_3\ d\boldsymbol{a}'^1 = r^2 \sin\theta\ d\theta d\varphi & \\ d\mathbf{A}^2 = dA^2 \hat{\mathbf{e}}_2 & \Rightarrow\quad d\mathbf{A}^\theta = dA^\theta\ \hat{\mathbf{e}}_\theta\ =\ dA^\theta\ \hat{\boldsymbol{\theta}}\ \text{ with } & dA^\theta = h'_3\ h'_1\ d\boldsymbol{a}'^2 = r\sin\theta\ dr d\varphi & \\ d\mathbf{A}^3 = dA^3 \hat{\mathbf{e}}_3 & \Rightarrow\quad d\mathbf{A}^\varphi = dA^\varphi\ \hat{\mathbf{e}}_\varphi\ =\ dA^\varphi\ \hat{\boldsymbol{\varphi}}\ \text{ with } & dA^\varphi = h'_1\ h'_2\ d\boldsymbol{a}'^3 = r dr d\theta & \qquad (8.4.k.6) \end{array}$$

so that

$$\begin{array}{llll} d\mathbf{A}^r = r^2 \sin\theta\ d\theta d\varphi\ \hat{\mathbf{r}} & \qquad \rho d\varphi\ \ r d\theta & \qquad \rho\ = r\sin\theta & \\ d\mathbf{A}^\theta = r\sin\theta\ dr d\varphi\ \hat{\boldsymbol{\theta}} & \qquad \rho d\varphi\ \ dr & & \\ d\mathbf{A}^\varphi = r dr d\theta\ \hat{\boldsymbol{\varphi}} & \qquad r d\theta\ \ dr & & \qquad (8.4.k.7) \end{array}$$

where all three vectors are seen to have the correct dimensions length^2. As an exercise in staring, the reader is invited to verify these results from the picture below using the hints shown above on the right,



(8.4.k.8)

8.5 Transformation of Differential Volume applied to Integration

As discussed in Section C.8, the integral $\int_D dV\ h(\mathbf{x})$ is the same regardless of the way the dV elements are chosen, as long as those elements exactly fill the integration region D.

In the discussion above, |dV| (call it dV_a) refers to a positive differential volume element in x-space which is typically not aligned with the axes and for a general transformation F is not in general orthogonal, as shown on the right side of Fig (8.2.2). Moreover, the shape of the differential volume N-piped varies over the region of integration. Nevertheless, this "rag-tag band" of differential volumes, as mentioned below (C.8.3) for the 2D case, fills the integration region perfectly.

Alternatively one could consider |dV| (call it dV_b) to be the usual $dx_1 dx_2 ..... dx_N$ differential volume elements in x-space, and of course this set of differential volume elements also fills the integration space perfectly.

Thinking of these two different differential volumes as dV_a and dV_b, one can see from the definition of the Riemann integral as the limit of a sum,

$$\lim \Sigma_i\, dV_a(\mathbf{x}_i)\, f(\mathbf{x}_i) \;=\lim \Sigma_i\, dV_b(\mathbf{x}_i)\, f(\mathbf{x}_i) \qquad (8.5.1)$$

that

$$\int_D dV_a\ h(\mathbf{x}) \;=\int_D dV_b\ h(\mathbf{x}) \quad . \qquad (8.5.2)$$

There would be little meaning to the statement $dV_a = dV_b$, since no one is claiming there is some particular skewed N-piped of volume dV_a which matches some axis-aligned N-piped of volume dV_b . Nevertheless, one could write $dV_a = dV_b$ as a distributional symbolic equality where the meaning of that symbolic equality is precisely the equivalence of the two integrals above for any domain D and for any reasonable function h(**x**). [Formally h(**x**) might have to be a "test function" φ(**x**). Certainly one would require that both integrals converge.]

It was shown in the second last line of (8.4.e.4) that the volume of the skewed N-piped on the right side of Fig (8.2.2) was given by

$$dV_a \;= |J(\mathbf{x}')|\, d\mathcal{V}' \;= |J(\mathbf{x}')|\, (\, \Pi_{i=1}{}^{N}\, dx'^{i}) \,. \qquad |J(\mathbf{x}')| = \sqrt{g'(\mathbf{x})} \qquad // \ g = +1 \qquad (8.5.3)$$

Combining this with the distributional symbolic equation $dV_{\mathbf{a}} = dV_{\mathbf{b}}$ gives

$$dV_{\mathbf{a}} = dV_{\mathbf{b}}$$

or

$$|J(\mathbf{x}')|\ (\ \Pi_{i=1}{}^{N}\ dx'^i) = \ (\ \Pi_{i=1}{}^{N}\ dx^i)$$

or

$$|J(\mathbf{x}')|\ d\mathcal{V}' = \ dV_{\mathbf{b}}\ . \tag{8.5.4}$$

Now overriding our previous notation, we can make these new commonly used definitions

$$\begin{aligned} dV\ &\equiv (\ \Pi_{i=1}{}^{N}\ dx^i) \\ dV' &\equiv (\ \Pi_{i=1}{}^{N}\ dx'^i) \end{aligned} \tag{8.5.5}$$

and express the distributional result (8.5.4) as

$$|J(\mathbf{x}')|\ dV' = \ dV\ . \qquad\qquad // \ \ |J(\theta,\rho)|\ d\rho d\theta \ = \ dxdy \text{ in (C.8.6)} \tag{8.5.6}$$

We refer to this distributional equality in the example of (C.8.6) as the "Jacobian Integration Rule". The symbolic equation is a shorthand for this equation

$$\int_D dV\ \ h(\mathbf{x})\ =\ \int_{D'} dV'\ |J(\mathbf{x}')|\ \ h(\mathbf{x}) \tag{8.5.7}$$

where on the right $h(\mathbf{x}) = h(\mathbf{x}(\mathbf{x}'))$ and region D' is the same region as D expressed in terms of the x' coordinates. Writing out the volume elements this says

$$\int_D (\ \Pi_{i=1}{}^{N}\ dx^i)\ \ h(\mathbf{x})\ =\ \int_{D'} (\ \Pi_{i=1}{}^{N}\ dx'^i)\ \ |J(\mathbf{x}')|\ \ h(\mathbf{x}(\mathbf{x}')) \tag{8.5.8}$$

and finally, using $J^2 = g'/g = g' = \det(\overline{g}') \rightarrow \det(g'_{ab})$ from (5.12.14),

$$\int_D (\ \Pi_{i=1}{}^{N}\ dx^i)\ \ h(\mathbf{x})\ =\ \int_{D'} (\ \Pi_{i=1}{}^{N}\ dx'^i)\ [\ \sqrt{\det(g'_{ab})}\ ]\ \ h(\mathbf{x}(\mathbf{x}'))\ \ . \tag{8.5.9}$$

For example, when applied to polar and spherical coordinates, one gets

$$\int_D dxdy\ \ h(\mathbf{x})\ =\ \int_{D'} drd\theta\ \ [r]\ h(\mathbf{x}(r,\theta)) \qquad\qquad \sqrt{\det(g'_{ab})}\ = \Pi_i h'_i = r \tag{8.5.10}$$

$$\int_D dxdydz\ \ h(\mathbf{x})\ =\ \int_{D'} drd\theta d\varphi\ \ [\ r^2\sin\theta\ ]\ h(\mathbf{x}(r,\theta,\varphi)) \qquad \sqrt{\det(g'_{ab})}\ = \Pi_i h'_i = \ r^2\ \sin\theta \tag{8.5.11}$$

In the first case $h(\mathbf{x}) = h(x,y)$ and $h(\mathbf{x}(r,\theta)) = h(r\cos\theta,r\sin\theta)$.
In the second case $h(\mathbf{x}) = h(x,y,z)$ and $h(\mathbf{x}(r,\theta,\varphi)) = h(r\sin\theta\cos\varphi,r\sin\theta\sin\varphi,r\cos\theta)$.

8.6 Interpretations of the Jacobian

Using the facts of Section 5.12 (in Standard Notation) and of the above Sections, one can produce various expressions and interpretations for the Jacobian J and its absolute value |J| :

$$J(\mathbf{x}') \equiv \det(S^i{}_j(\mathbf{x}')) = \det(\partial x^i/\partial x'^k) = 1/\det(R^i{}_j(\mathbf{x}(\mathbf{x}')) = 1/\det(\partial x'^i/\partial x^k) \qquad \textit{(5.12.6)}$$

$$|J(\mathbf{x}')| = \sqrt{\det(g'_{ab}(\mathbf{x}'))} = \sqrt{g'(\mathbf{x}')} \qquad // \ g=1 \qquad \textit{(5.12.14)}$$

$$|J(\mathbf{x}')| = \text{the volume of the N-piped in x-space spanned by the } \mathbf{e_n}(\mathbf{x}), \text{ where } \mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}') \qquad \textit{(8.3.8)}$$

$$|J(\mathbf{x}')| = dV_{N\text{-piped}}/d\mathcal{V}' = \text{ratio of differential x-space N-piped volume } / \ (\Pi_{i=1}{}^N dx'^i) \qquad \textit{(8.4.e.5)}$$

$$|J(\mathbf{x}')| = dV/dV' = (\Pi_{i=1}{}^N dx^i)/(\Pi_{i=1}{}^N dx'^i) \quad // \text{ distributional Jacobian Integration Rule} \qquad \textit{(8.5.6)}$$

(8.6.1)

As discussed in Section 6.9, if the curvilinear coordinates are ordered so that the $\mathbf{e_n}$ form a right-handed coordinate system, then det(S)>0, $\sigma = \text{sign}(\det(S)) = +1$, and |J| = J. Some authors define their Jacobians as the inverse of ours, see (5.12.3).

8.7 Volume integration of a tensor field under linear transformations

Recall the distributional statement of (8.5.6),

$$dV' = J^{-1} dV \quad . \qquad (8.7.1)$$

In the language of Appendix D, this says that the volume element transforms from x-space to x'-space as a scalar density of weight +1. Since by (D.1.6) $g^{1/2}$ transforms as a scalar density of weight -1, according to (D.2.3) the object $g^{1/2}dV$ then transforms as an ordinary scalar (weight 0) (see for example Weinberg p 99 (4.4.6)). Then if one were to define

$$T^{ijk...} \equiv \int_D g^{1/2} dV \, A^{ijk..}(\mathbf{x}) \qquad (8.7.2)$$

one might expect that, if $A^{ijk..}(\mathbf{x})$ transforms as a tensor field, then $T^{ijk...}$ might transform as tensor. To investigate this conjecture, consider the above integral in x'-space,

$$T'^{ijk...} \equiv \int_{D'} g'^{1/2} dV' \, A'^{ijk..}(\mathbf{x}') = \int_D g^{1/2} dV \, A'^{ijk..}(\mathbf{x'(x)}) \qquad // \ g^{1/2}dV = \text{scalar}$$

$$= \int_D g^{1/2} dV \, R^i{}_{i'} R^j{}_{j'} ..... A^{i'j'k'...}(\mathbf{x}) \, .$$

If the underlying transformation **x**' = **F**(**x**) is linear (see Section 2.8) then the $R^a{}_b$ are independent of position and we continue the above,

$$= R^i{}_{i'} R^j{}_{j'} ..... \int_D g^{1/2} dV\, A^{i'j'k'\cdots}(\mathbf{x})$$

$$= R^i{}_{i'} R^j{}_{j'} ..... T^{i'j'k'\cdots} \qquad (8.7.3)$$

Therefore, for *linear* transformations **x**' = **F**(**x**), an integral of the form (8.7.2) of a tensor field is itself a tensor of the same type.

Examples: Suppose g = 1 and x' = Rx where R is a global rotation, so we have a linear transformation. For a rotation in developmental notation one has $R^{-1} = R^T$. But (5.7.6) says $g' = R\, g\, R^T = R\, 1\, R^{-1} = 1$, so g' = 1 as well. We may conclude that the volume integral of a tensor field of any type is a tensor of the same type. Here are two simple examples:

$$J^i = \int_D dV\, x^i$$

$$J^{ij} = \int_D dV\, [\, r^2\delta^{ij} - x^i x^j] \ . \qquad (8.7.4)$$

Under rotations, x^i is a true vector, and $r^2\delta^{ij} - x^i x^j$ is a true rank-2 tensor (traceless). It follows that J^i and J^{ij} are also tensors. Under rotations, mass density ρ transforms as a scalar, so the following objects are tensors as well,

$$I^i = \int_D dV\, \rho(\mathbf{x})\, x^i \qquad // \text{ vector}$$

$$I^{ij} = \int_D dV\, \rho(\mathbf{x})\, [\, r^2\delta^{ij} - x^i x^j] \ . \qquad // \text{ rank-2 tensor, } r^2 = \Sigma_i (x^i)^2 \qquad (8.7.5)$$

Vector components I^i are the first moments of a mass distribution, while I^{ij} is the usual inertia tensor.

Since I^{ij} is real and symmetric, it can be diagonalized by a certain rotation R. We can think of this as a transformation from x-space to x'-space where the resulting tensor I'^{ij} is diagonal. In x'-space, the diagonal elements of the tensor I'^{ij} are then its eigenvalues ($\lambda_i = I'^{ii}$), while the eigenvectors are the axis-aligned $\mathbf{e}'_i$ of Chapter 3. Back in x-space, the eigenvectors of the non-diagonal I^{ij} are then $\mathbf{e}_n \equiv S\mathbf{e}'_n$ where $S = R^{-1}$. This can be verified as follows (developmental notation)

$$I' = R\, I\, R^{-1} \quad \Rightarrow \quad I'_{ij} = R_{ii'} I_{i'j'} (R^{-1})_{j'j} = R_{ii'} R_{jj'} I_{i'j} \qquad // \text{ contravariant tensor}$$

$$I'\, \mathbf{e}'_n = \lambda_n\, \mathbf{e}'_n \qquad // \text{ I' is diagonal}$$

so

$$I\, \mathbf{e}_n = [\, R^{-1} I'\, R\,]\, [S\mathbf{e}'_n] = R^{-1}\, I'\, (RS)\, \mathbf{e}'_n = S\, I'\, \mathbf{e}'_n = S\, \lambda_n \mathbf{e}'_n = \lambda_n\, S\mathbf{e}'_n = \lambda_n \mathbf{e}_n \ . \qquad (8.7.6)$$

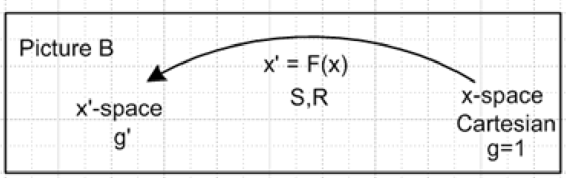
9. The Divergence in curvilinear coordinates

Note: Covariant derivations of all the curvilinear differential operator expressions appear in Chapter 15. In Chapters 9 through 13, we provide more "physical" or "brute force" derivations which are, of necessity, much less compact. That compactness is a testament to the power of the covariant derivative formalism, which might be called "semicolon technology". The formalism is not used in Chapters 9 through 13.

9.1 Geometric Derivation of the Curvilinear Divergence Formula

In Cartesian coordinates div $\mathbf{B} = \nabla \bullet \mathbf{B} = \partial_n B^n$, but expressed in curvilinear coordinates the right side has a more complicated form.

We provide here a geometric derivation (in N dimensions) of the formula for the divergence of a contravariant vector field expressed in curvilinear coordinates, which means x'-space coordinates with Picture B.



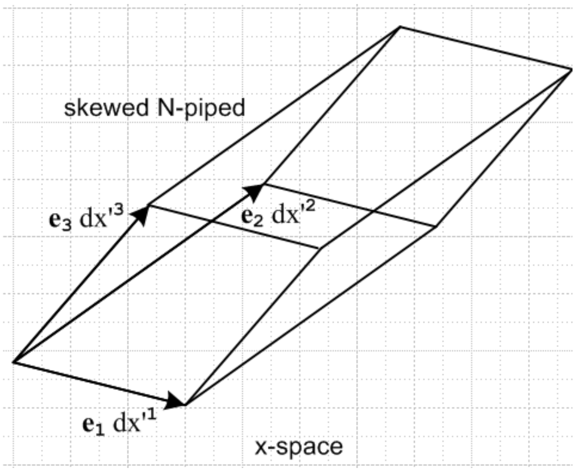
(9.1.1)

This derivation is an exercise in using the transformation results obtained in Chapter 8 above, and in understanding the meaning of the components of a vector, as discussed in Section C.5.

The divergence of a vector field can be computed in a Cartesian x-space by taking the limit of the flux emerging from a closed volume divided by the size of the volume, in the limit that the volume shrinks down around some point **x**. Being a scalar field, the divergence is a property of the vector field at some point **x** and therefore cannot depend on the shape of the closed volume used for the calculation†. If the shape of the volume is taken to be a standard-issue axis-aligned N-piped, the divergence obtained will be expressed in terms of the Cartesian coordinates and in terms of the Cartesian components of the vector field: $[\text{div } \mathbf{B}](\mathbf{x}) = \partial_n B^n(\mathbf{x})$ where $\mathbf{B} = B^n \hat{\mathbf{n}}$. However, if the N-piped shape is the one below, evaluation of this same $[\text{div } \mathbf{B}](\mathbf{x})$ produces an expression which involves only the curvilinear coordinates and the curvilinear components of the vector field, as will now be demonstrated.

> † For example, the total amount of water per second flowing through a mathematical closed boundary surrounding a point-source sprinkler head will not depend on the shape of that boundary. This fact is part of the Divergence Theorem, see e.g. wiki.

We start by considering again our differential non-orthogonal N-piped sitting in x-space, which has edges $\mathbf{e}_n dx'^n$, faces $d\mathbf{A}^n$ and volume dV, as discussed in Section 8.4 above:



(9.1.2)

In order to avoid confusion with volume V or area **A**, we name the vector field **B**. As just noted, the divergence of a vector field **B** is the total flux flowing out through the faces of the N-piped divided by the volume of the N-piped, in the limit that all differentials go to 0. Thus one writes symbolically,

$$[\text{div } \mathbf{B}](\mathbf{x}) \;=\; (1/dV) \int d\mathbf{A}\bullet\mathbf{B} \;=\; (1/dV) \int d\mathbf{A}(\mathbf{x})\bullet\mathbf{B}(\mathbf{x}) \qquad (9.1.3)$$

where the surface integral is over all the faces of the above x-space differential N-piped. Recall that as we move around in space, the shape (and size) of the above N-piped changes, so the d**A** of a face changes, hence d**A**(**x**).

<u>Comment on the div **B** as a scalar.</u> If **B** is a tensorial vector field, then we claim that div **B** is a tensorial scalar field, and one can write

$$[\text{div } \mathbf{B}]'(\mathbf{x}') = [\text{div } \mathbf{B}](\mathbf{x}) \;. \qquad (9.1.4)$$

The operator object $\{(1/dV)\int d\mathbf{A}(\mathbf{x})\bullet\}$ acts as a tensorial vector operator, so that the result of its action on **B** in (9.1.3) is a tensorial scalar. In (8.4.c.2) and (8.4.d.3) it was shown that d**A** and dV are vector and scalar densities of weight -1. This means that dV' = J dV and d**A**' = J Rd**A** so the ratio d**A**/dV in our operator is a tensorial vector.

The fact that div **B** is a tensorial scalar is much more obvious from the alternative divergence derivation given in Section 15.3. There we find that div **B** = $B^i{}_{;i}$ which evaluated in Cartesian space becomes div **B** = $B^i{}_{,i} = \partial_i B^i$. But in this Chapter we are avoiding Appendix F covariant derivatives!

The task is now to compute the integral $\int d\mathbf{A}(\mathbf{x})\bullet\mathbf{B}(\mathbf{x})$ over the N-piped in (9.1.2).

Appendix B shows that the N-piped faces come in parallel pairs, so we start by considering pair n. We now restore the original Appendix B convention that $\mathbf{A}_n$ points outward from the N-piped far face regardless of the sign of det(S). Then adding the differential distances ($\Pi_{i\neq n}$ dx'^{i}) to the first line of (8.3.7) one gets,

$$d\mathbf{A}^n(\mathbf{x}) = |\det(S^i{}_j(\mathbf{x}'))|\ \mathbf{e}^n(\mathbf{x})\ (\Pi_{i\neq n}\, dx'^i) = \sqrt{g'(\mathbf{x}')}\ \mathbf{e}^n(\mathbf{x})\ (\Pi_{i\neq n}\, dx'^i)\ . \tag{9.1.5}$$

From (5.12.6) and (5.12.14) with g = 1 one has $|\det(S)| = |J| = g'^{1/2}$, while $\mathbf{e}^n$ are the reciprocal base vectors called $\mathbf{E}_n$ Chapter 6. Quantities $d\mathbf{A}^n$, S, $\mathbf{e}^n$ and g' are explicitly shown as functions of space, and as usual $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. Again, the out-facing vector area for the far face of pair n is $d\mathbf{A}^n$, while the out-facing vector area for the near face is $- d\mathbf{A}^n$.

From (9.1.5) the contribution to the above divergence integral from "far face n" is, approximately,

$$\begin{aligned}[d\mathbf{A}^n]\bullet\mathbf{B}(\mathbf{x}) &\approx [\sqrt{g'(\mathbf{x}'_{far})}\ \mathbf{e}^n(\mathbf{x}_{far})\ (\Pi_{i\neq n}\, dx'^i)\ ]\bullet\mathbf{B}(\mathbf{x}_{far}) \\ &= \sqrt{g'(\mathbf{x}'_{far})}\ (\Pi_{i\neq n}\, dx'^i)\ \mathbf{e}^n(\mathbf{x}_{far}) \bullet \mathbf{B}(\mathbf{x}_{far})\ ,\end{aligned} \tag{9.1.6}$$

where $\mathbf{x}_{far}$ is taken to be a point at the center of far face n. Recall from the 2nd last line of (7.13.10) that

$$\mathbf{B}(\mathbf{x}) = B'^n(\mathbf{x}')\mathbf{e}_n \qquad \text{where} \qquad B'^n(\mathbf{x}') = \mathbf{e}^n(\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) \tag{9.1.7}$$

which says that, when $\mathbf{B}$ is expanded on the $\mathbf{e}_n$, the coefficients B'^n of the expansion are the contravariant components of vector $\mathbf{B}$ transformed into $\mathbf{B}'$ in x'-space (the curvilinear coordinate space) by $\mathbf{B}' = R\mathbf{B}$. Thus we write,

$$B'^n(\mathbf{x}'_{far}) = \mathbf{e}^n(\mathbf{x}_{far}) \bullet \mathbf{B}(\mathbf{x}_{far})\ . \tag{9.1.8}$$

Inserting this into (9.1.6) gives

$$d\mathbf{A}^n\bullet\mathbf{B}(\mathbf{x}_{far}) \approx \sqrt{g'(\mathbf{x}'_{far})}\ (\Pi_{i\neq n}\, dx'^i)\ B'^n(\mathbf{x}'_{far})\ . \tag{9.1.9}$$

This far-face n contribution to the flux integral is now expressed entirely in terms of x'-space objects and coordinates. A similar expression obtains for the near-face n, but the sign of $d\mathbf{A}^n$ is reversed. Adding the contributions of these two faces of pair n gives

$$\int_{\text{two faces n}} d\mathbf{A}\bullet\mathbf{B}(\mathbf{x}) = \{\ \sqrt{g'(\mathbf{x}'_{far})}\ B'^n(\mathbf{x}'_{far}) - \sqrt{g'(\mathbf{x}'_{near})}\ B'^n(\mathbf{x}'_{near})\ \}\ (\Pi_{i\neq n}\, dx'^i)\ . \tag{9.1.10}$$

In x'-space, if $\mathbf{x}'_{near}$ is a point at the center of the near face of face pair n, then

$$\mathbf{x}'_{far} = \mathbf{x}'_{near} + \mathbf{e}'_n\, dx'^n \qquad \text{where } \mathbf{e}'_n = \text{axis-aligned basis vector in x'-space} \tag{9.1.11}$$

since these two points map into the near and far face-n centers in x-space. For any function f,

$$f(\mathbf{x}'_{far}) - f(\mathbf{x}'_{near}) \approx (\partial'_n\, f(\mathbf{x}'_{near}))\, dx'^n \qquad // \text{ no implied sum on n} \tag{9.1.12}$$

where a change is made only in coordinate x'^n by amount dx'^n. Applying to $f = J\, B'^n$ yields

$$\{\ \sqrt{g'(\mathbf{x}'_{far})}\ B'^n(\mathbf{x}'_{far}) - \sqrt{g'(\mathbf{x}'_{near})}\ B'^n(\mathbf{x}'_{near})\ \} \approx \partial'_n\, [\sqrt{g'(\mathbf{x}'_{near})}\ B'^n(\mathbf{x}'_{near})]\, dx'^n\ . \tag{9.1.13}$$

In the limit $dx'^n \to 0$ one has $\mathbf{x_{near}} \to \mathbf{x}$. Then (9.1.10) with (9.1.13) gives

$$\int_{\text{two faces n}} d\mathbf{A}\bullet\mathbf{B}(\mathbf{x}) \;=\; \partial'_n [\sqrt{g'(\mathbf{x}')}\;\; B'^n(\mathbf{x}')]\, dx'^n \;\; (\Pi_{i\neq n}\, dx'^i)$$

$$= \partial'_n [\sqrt{g'(\mathbf{x}')}\, B'^n(\mathbf{x}')]\, (\Pi_i\, dx'^i) \;\; . \qquad (9.1.14)$$

Now *all* N differentials are present in $(\Pi_i dx'^i)$. The total flux flowing out through all N pairs of faces of the N-piped in x-space is this same result with an implicit sum on n, so

$$\text{total flux} \;=\; \int d\mathbf{A}\bullet\mathbf{B}(\mathbf{x}) \;=\; \partial'_n [\sqrt{g'(\mathbf{x}')}\;\; B'^n(\mathbf{x}')]\, (\Pi_i\, dx'^i) \;=\; \partial'_n [\sqrt{g'(\mathbf{x}')}\, B'^n(\mathbf{x}')]\, d\mathcal{V}\,' \qquad (9.1.15)$$

where $d\mathcal{V}' = \Pi_i\, dx'^i$ is the volume of the differential N-piped in Cartesian-view x'-space as in (8.2.3). The divergence of **B** from the defining symbolic expression is then

$$[\text{div } \mathbf{B}](\mathbf{x}) = \int d\mathbf{A}\bullet\mathbf{B}(\mathbf{x}) \,/\, dV \;=\; \partial'_n [\sqrt{g'(\mathbf{x}')}\;\; B'^n(\mathbf{x}')]\, (d\mathcal{V}'/dV) \qquad (9.1.16)$$

where dV is the volume of the x-space N-piped shown in Fig (8.2.2). In (8.4.e.5) it is shown that

$$dV = \; (g')^{1/2}\, d\mathcal{V}' \qquad \Rightarrow \qquad (d\mathcal{V}'/dV) \;=\; 1/\sqrt{g'(\mathbf{x}')} \qquad // \;(g')^{1/2} \;= J \qquad (9.1.17)$$

so that (9.1.16) becomes

$$[\text{div } \mathbf{B}](\mathbf{x}) = [1/\sqrt{g'(\mathbf{x}')}\,]\, \partial'_n [\sqrt{g'(\mathbf{x}')}\;\; B'^n(\mathbf{x}')] \qquad \text{// all x'-space coordinates and objects} \qquad (9.1.18)$$

$$[\text{div } \mathbf{B}](\mathbf{x}) = \partial_n B^n(\mathbf{x})\, . \qquad \text{// all x-space coordinates and objects} \qquad (9.1.19)$$

The added second line just shows [div **B](x)** expressed in terms of the Cartesian x-space coordinates and objects, while the first line resulting from our derivation shows *the same* [div **B](x)** expressed in terms of only x'-space objects and coordinates. The goal advertised above has been fulfilled.

If **B** is a tensorial vector, then as noted above div **B** is a tensorial scalar,

$$[\text{div } \mathbf{B}](\mathbf{x}) = [\text{div } \mathbf{B}]'(\mathbf{x}') \qquad (9.1.20)$$

and thus the left sides of both equations above could be replaced by [div **B]'(x')**.

9.2 Various expressions for div B

It is shown above that

$$[\text{div } \mathbf{B}](\mathbf{x}) \;= [1/\sqrt{g'(\mathbf{x}')}\,]\, \partial'_n [\sqrt{g'(\mathbf{x}')}\;\; B'^n(\mathbf{x}')] \;\; . \qquad (9.2.1)$$

To obtain div **B** written in terms of *covariant* components B'_n, one sets $B'^n = g'^{nm}\, B'_m$ to get

$$[\mathrm{div}\ \mathbf{B}](\mathbf{x}) = [1/\sqrt{g'(\mathbf{x}')}\ ]\ \partial'_n\, [\sqrt{g'(\mathbf{x}')}\ g'^{nm}(\mathbf{x}')\ B'_m(\mathbf{x}')]\ . \qquad (9.2.2)$$

Recall that the B'_m are the coefficients of **B** when expanded on the $\mathbf{e}^n$, $\mathbf{B} = B'_n\mathbf{e}^n$.

In practical work **B** is expanded on the *unit* vectors $\hat{\mathbf{e}}_n \equiv \mathbf{e}_n/\,|\mathbf{e}_n| = \mathbf{e}_n/h'_n$ so that

$$\mathbf{B} = B'^n\mathbf{e}_n = B'^n\, h'_n\hat{\mathbf{e}}_n = \mathcal{B}'^n\hat{\mathbf{e}}_n \qquad \text{where} \quad \mathcal{B}'^n \equiv B'^n\, h'_n \qquad (9.2.3)$$

and then

$$[\mathrm{div}\ \mathbf{B}](\mathbf{x}) = [1/\sqrt{g'(\mathbf{x}')}\ ]\ \partial'_n\, [\sqrt{g'(\mathbf{x}')}\ \mathcal{B}'^n(\mathbf{x}')/\ h'_n(\mathbf{x}')\ ]\ . \qquad (9.2.4)$$

For example, spherical coordinate work might use $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3 = \hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\varphi}}$. As noted earlier, the components $\mathcal{B}^n(\mathbf{x})$ are not contravariant vector components since they don't quite transform properly:

$$B'^n = R^n{}_m B^m \qquad (\mathcal{B}'^n\,/\,h'_n) = R^n{}_m\ (\mathcal{B}^n\,/\,1) \qquad \mathcal{B}'^n = h'_n\,(R^n{}_m\ \mathcal{B}^n)\ . \qquad (9.2.5)$$

Our Picture B results so far are these, assuming **B** is a tensorial vector,

<u>General:</u> $\quad B'^n(\mathbf{x}') = R^n{}_m B^m(\mathbf{x}) \qquad \mathbf{x} = F^{-1}(\mathbf{x}') \equiv \mathbf{x}(\mathbf{x}') \qquad \mathbf{x}' = \mathbf{F}(\mathbf{x}) \equiv \mathbf{x}'(\mathbf{x})$

$$\begin{aligned}
&[\mathrm{div}\ \mathbf{B}](\mathbf{x}) = [1/\sqrt{g'(\mathbf{x}')}\ ]\ \partial'_n\, [\sqrt{g'(\mathbf{x}')}\ \ B'^n(\mathbf{x}')] && \mathbf{B} = B'^n\mathbf{e}_n \\
&[\mathrm{div}\ \mathbf{B}](\mathbf{x}) = [1/\sqrt{g'(\mathbf{x}')}\ ]\ \partial'_n\, [\sqrt{g'(\mathbf{x}')}\ \mathcal{B}'^n(\mathbf{x}')/\ h'_n(\mathbf{x}')\ ] && \mathbf{B} = \mathcal{B}'^n\ \hat{\mathbf{e}}_n \\
&[\mathrm{div}\ \mathbf{B}](\mathbf{x}) = [1/\sqrt{g'(\mathbf{x}')}\ ]\ \partial'_n\, [\sqrt{g'(\mathbf{x}')}\ g'^{nm}(\mathbf{x}')\ B'_m(\mathbf{x}')] && \mathbf{B} = B'_n\mathbf{e}^n \\
&[\mathrm{div}\ \mathbf{B}](\mathbf{x}) = \partial_n\, B^n(\mathbf{x}) \qquad //\ B^n = \text{Cartesian components of B} && \mathbf{B} = B^n\ \hat{\mathbf{n}} \qquad (9.2.6) \\
&[\mathrm{div}\ \mathbf{B}](\mathbf{x}) = [\mathrm{div}\ \mathbf{B}]'(\mathbf{x}')
\end{aligned}$$

For orthogonal curvilinear coordinates, one has

$$g'_{ij} = h'_i{}^2\,\delta_{i,j} \qquad g'^{ij} = h'_i{}^{-2}\,\delta_{i,j} \quad \det(g'_{ij}) = \Pi_i h'_i{}^2 \quad \sqrt{g'} = (\Pi_i h'_i) = h'_1 h'_2....h'_N \qquad (9.2.7)$$

so the above expressions can be written (the arguments **x'** of the h'_n are now suppressed)

<u>Orthogonal:</u>

$$\begin{aligned}
&[\mathrm{div}\ \mathbf{B}](\mathbf{x}) = [1/(\Pi_i h'_i)]\ \partial'_n\, [(\Pi_i h'_i)\ B'^n(\mathbf{x}')] && \mathbf{B} = B'^n\mathbf{e}_n \\
&[\mathrm{div}\ \mathbf{B}](\mathbf{x}) = [1/(\Pi_i h'_i)]\ \partial'_n\, [(\Pi_i h'_i)\ \mathcal{B}'^n(\mathbf{x}')/\ h'_n\ ] && \mathbf{B} = \mathcal{B}'^n\ \hat{\mathbf{e}}_n \\
&[\mathrm{div}\ \mathbf{B}](\mathbf{x}) = [1/(\Pi_i h'_i)]\ \partial'_n\, [(\Pi_i h'_i)\ B'_n(\mathbf{x}')/h'_n{}^2] && \mathbf{B} = B'_n\mathbf{e}^n \\
&[\mathrm{div}\ \mathbf{B}](\mathbf{x}) = \partial_n\, B^n(\mathbf{x}) \qquad //\ B^n = \text{Cartesian components of B} && \mathbf{B} = B^n\ \hat{\mathbf{n}} \\
&[\mathrm{div}\ \mathbf{B}](\mathbf{x}) = [\mathrm{div}\ \mathbf{B}]'(\mathbf{x}') && (9.2.8)
\end{aligned}$$

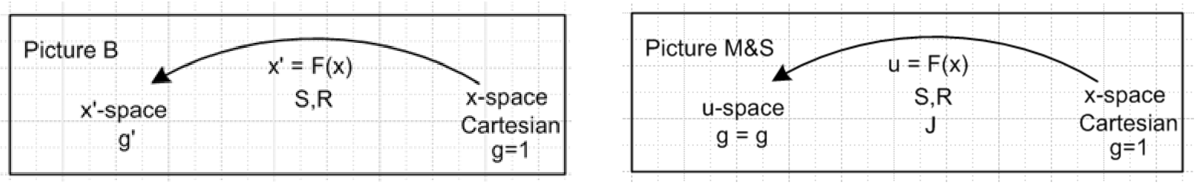
Comment: Are the equations (9.2.6) valid if **B** is *not* a tensorial vector? For such a **B** one might try to make it be tensorial "by definition" as discussed in Section 2.9. One would then go ahead and define $B'^n(\mathbf{x}') \equiv R^n{}_m B^m(\mathbf{x})$ and claim success. If such a definition does not result in an inconsistency, then such a **B** has been moved into the class of tensorial vectors. Example 1 in (2.9.3) shows have such an inconsistency might arise, and it is interesting to see how that plays out here. Suppose F is non-linear so that R and $g' = RR^T$ are functions of **x**' and are not constants. Take $\mathbf{B}(\mathbf{x}) = \mathbf{x}$ (the identity field) and try to make it be contravariant by definition, $\mathbf{x}'^n \equiv R^n{}_m \mathbf{x}^m$. The Cartesian divergence is then div $\mathbf{B} = \partial_n B^n(\mathbf{x}) = \partial_n x^n = 3$. But the first equation of (9.2.6) says

$$\text{div}\,\mathbf{B} = [1/\sqrt{g'(\mathbf{x}')}\,]\,\partial'_n\,[\sqrt{g'(\mathbf{x}')}\;\,x'^n] = \partial'_n\, x'^n + [1/\sqrt{g'(\mathbf{x}')}\,]\,x'^n\partial'_n\sqrt{g'(\mathbf{x}')} = 3 + \text{other stuff} \qquad (9.2.9)$$

and thus the two calculations for div **B** disagree. As noted earlier, $\mathbf{x}' \equiv R\mathbf{x}$ conflicts with $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ in the case of non-linear F.

This comment can be applied to the tensorial character of the differential operators treated in later Chapters.

9.3 Translation from Picture B to Picture M&S



(9.3.1)

Picture M&S reflects the notation used by Moon & Spencer. In order to avoid a symbol conflict with the Cartesian tensor components, the Curvilinear (now u-space) components are displayed in **italics**.

The rules for translation are

- replace **x**' by **u** everywhere (9.3.2)
- replace ∂'_n by ∂_n meaning $\partial/\partial u^n$ (exception: on a "Cartesian" line ∂_n means $\partial/\partial x^n$)
- replace g' by g (both the scalar and the tensor) and h_n' by h_n
- put all primed tensor components (scalar, vector, etc) into unprimed italics (eg, $B'^n \rightarrow B^n$, $f' \rightarrow f$)

After this translation, all unprimed tensor components are functions of **x**, while all italicized tensor components are functions of **u**.

Here then are the translations of the two blocks above: (implied summation everywhere)

General: now $B^n(\mathbf{u}) = \mathrm{R}^n{}_m\mathrm{B}^m(\mathbf{x}) \qquad \mathbf{x} = \mathbf{F}^{-1}(\mathbf{u}) \equiv \mathbf{x}(\mathbf{u}) \qquad \mathbf{u} = \mathbf{F}(\mathbf{x}) \equiv \mathbf{u}(\mathbf{x})$

$$[\mathrm{div}\,\mathbf{B}](\mathbf{x}) = [1/\sqrt{g}\,]\,\partial_n\,[\sqrt{g}\;B^n] \qquad\qquad \mathbf{B} = B^n\mathbf{e}_n$$

$$[\mathrm{div}\,\mathbf{B}](\mathbf{x}) = [1/\sqrt{g}\,]\,\partial_n\,[\sqrt{g}\;\mathscr{B}^n/\,h_n\,] \qquad\qquad \mathbf{B} = \mathscr{B}^n\hat{\mathbf{e}}_n \qquad // \text{ M\&S 1.06}$$

$$[\mathrm{div}\,\mathbf{B}](\mathbf{x}) = [1/\sqrt{g}\,]\,\partial_n\,[\sqrt{g}\;g^{nm}\,B_m] \qquad\qquad \mathbf{B} = B_n\mathbf{e}^n$$

$$[\mathrm{div}\,\mathbf{B}](\mathbf{x}) = \partial_n\mathrm{B}^n \qquad // \ \mathrm{B}^n = \text{Cartesian components of } \mathbf{B} \qquad\qquad \mathbf{B} = \mathrm{B}^n\,\hat{\mathbf{n}} = \mathrm{B}_n\,\hat{\mathbf{n}}$$

$$[\mathrm{div}\,\mathbf{B}](\mathbf{x}) = [\mathit{div}\,\boldsymbol{B}](\mathbf{u}) \qquad // \text{ transformation (scalar)} \qquad\qquad (9.3.3)$$

Orthogonal:

$$[\mathrm{div}\,\mathbf{B}](\mathbf{x}) = [1/(\Pi_i h_i)]\,\partial_n\,[(\Pi_i h_i)\,B^n] \qquad\qquad \mathbf{B} = B^n\mathbf{e}_n$$

$$[\mathrm{div}\,\mathbf{B}](\mathbf{x}) = [1/(\Pi_i h_i)]\,\partial_n\,[(\Pi_i h_i)\,\mathscr{B}^n\,/\,h_n] \qquad\qquad \mathbf{B} = \mathscr{B}^n\,\hat{\mathbf{e}}_n$$

$$[\mathrm{div}\,\mathbf{B}](\mathbf{x}) = [1/(\Pi_i h_i)]\,\partial_n\,[(\Pi_i h_i)\,B_n\,/\,h_n^{\,2}] \qquad\qquad \mathbf{B} = B_n\mathbf{e}^n \qquad\qquad (9.3.4)$$

Notice that the scalar function $[\mathrm{div}\,\mathbf{B}]'(\mathbf{x}') \rightarrow [\mathit{div}\,\boldsymbol{B}](\mathbf{u})$ according to the fourth rule of (9.3.2) above, and that the arguments of all $h_k(\mathbf{u})$ are suppressed.

As an example, for N=3 the second line of (9.3.4) becomes

$$[\mathrm{div}\,\mathbf{B}](\mathbf{x}) = [1/(h_1h_2h_3)]\ \{\ \partial_1[h_2h_3\,\mathscr{B}^1(\mathbf{u})\ ]\ +\text{cyclic}\ \} \qquad\qquad \mathbf{B} = \mathscr{B}^n\,\hat{\mathbf{e}}_n \qquad\qquad (9.3.5)$$

where + cyclic means two other terms with 1,2,3 cyclically permuted.

With the replacements

$$\mathbf{B} \rightarrow \mathbf{E}\,, \qquad \mathscr{B}^n \rightarrow \mathrm{E}_n \qquad h_n \rightarrow \sqrt{g_{nn}}$$

the 2nd equation of (9.3.3) agrees with Moon & Spencer p 2 (1.06).

Comment: We use the "script MT bold" font $\mathscr{B}^n$ for components of vectors expanded onto $\hat{\mathbf{e}}_n$. It had to be something in upper case distinct from B^n and B^n. In practice one can replace $\mathscr{B}^n$ with a different symbol and then $\mathscr{B}^n$ is just a formal notation appearing in formulas. For example, in spherical coordinates (1,2,3) = (r,θ,φ) one can make the replacements $\mathscr{B}^1, \mathscr{B}^2, \mathscr{B}^3 \rightarrow \mathrm{B}_r, \mathrm{B}_\theta, \mathrm{B}_\varphi$ and these then do not conflict with $\mathrm{B}_x, \mathrm{B}_y, \mathrm{B}_z$ or B_r, B_θ, B_φ . See Section 14.7 for another example. (9.3.6)

9.4 Comparison of various authors' notations

Different authors use different symbols for curvilinear coordinates. They usually use x-space as the Cartesian space, and then something like u-space or ξ-space as the curvilinear space:

		Curvilinear coords	Cartesian coords	Curvilinear space	$\sqrt{g_{nn}}$
Picture C		x^n	$x^{(0)n}$	x-space	h_n
Picture B		x'^n	x^n	x'-space	h'_n
Moon & Spencer (M&S)	p 2	u^n	x^n	u-space	$\sqrt{g_{nn}}$
Morse & Feshbach (M&F)	p 115	ξ_n	x_n	ξ-space	h_n
Margenau & Murphy	p 192	q^n	x^n	q-space	Q_n

(9.4.1)

These authors don't use any special notation to distinguish Cartesian from curvilinear components, nor is it always clear whether a component is a coefficient of a unit vector or not, so one must be careful. For example, on page 115 Morse & Feshbach simply say

$$\text{div}\ \mathbf{A} = \nabla \cdot \mathbf{A} = \frac{1}{h_1 h_2 h_3} \sum_n \frac{\partial}{\partial \xi_n}\left(h_1 h_2 h_3 \frac{A_n}{h_n} \right) \tag{9.4.2}$$

which compare to the 2nd line of (9.3.4) above (sum on n now displayed),

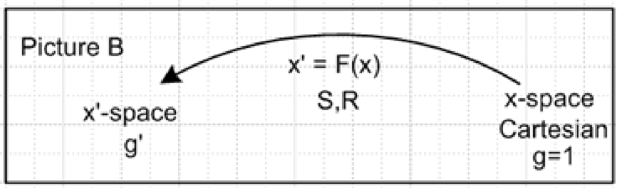
$$[\text{div}\ \mathbf{A}](\mathbf{x}) = [\mathit{div}\ \boldsymbol{A}](\mathbf{u}) = [1/(h_1h_2h_3)]\ \Sigma_n\ \partial_n\ [h_1h_2h_3\ \mathcal{A}^n(\mathbf{u}) / h_n] \qquad \mathbf{A} = \mathcal{A}^n\ \hat{\mathbf{e}}_n \tag{9.4.3}$$

so one should identify the M&F A_n with $\mathcal{A}^n$, the coefficient of $\hat{\mathbf{e}}_n$.

10. The Gradient in curvilinear coordinates

10.1 Expressions for grad f

This Chapter is considered in the Picture B context,



(10.1.1)

The gradient of f is defined in Cartesian space by

$$[\text{grad f}]_n \equiv G_n \equiv \partial_n f(\mathbf{x}) \qquad \text{so} \qquad \mathbf{G} = \text{grad f} = \partial_n f(\mathbf{x})\, \hat{\mathbf{n}} \quad . \tag{10.1.2}$$

Comment: To follow our conventions, we should perhaps write [grad f] as [**grad** f] as in [**∇**f] with a bolded del operator. We decided against this so that grad, div and curl are all written non-bolded. This seems to be a common usage of other authors.

Assuming f is a tensorial scalar field under F, then the $G_n = \partial_n f(\mathbf{x})$ are covariant vector field components under F. This is because,

$$[\partial'_n f'(\mathbf{x'})] = (R_n{}^m \partial_m) f'(\mathbf{x'}) = (R_n{}^m \partial_m) f(\mathbf{x}) = R_n{}^m [\partial_m f(\mathbf{x})] \,, \tag{10.1.3}$$

where recall from (2.10.1) that $f'(\mathbf{x'}) = f(\mathbf{x})$ for a scalar field. [In the technical language of Appendix F, one can write from (F.9.1) that tensorial vector $f_{;n} = f_{,n} = \partial_n f$.]

Since $\partial_n f$ is a tensorial vector, the covariance of Section 7.15 gives us (10.1.2) expressed in x'-space,

$$[\text{grad f}]'_n = G'_n \equiv \partial'_n f'(\mathbf{x'}) \,. \tag{10.1.4}$$

According to the last line of (7.13.10) vector **G** can be expanded as

$$\mathbf{G} = G'_1 \mathbf{e}^1 + G'_2\, \mathbf{e}^2 + \ldots = \Sigma_n G'_n\, \mathbf{e}^n \qquad \text{where} \qquad \mathbf{e}_n \bullet \mathbf{G} = G'_n \tag{10.1.5}$$

where the coefficients G'_n are the covariant components of **G**' in x'-space, which is the transformed **G**. One can thus write

$$\begin{aligned} &\mathbf{G}(\mathbf{x}) = [\text{grad f}](\mathbf{x}) = G'_n\, \mathbf{e}^n = \partial'_n f'(\mathbf{x'})\, \mathbf{e}^n = \mathbf{e}^n\, \partial'_n f'(\mathbf{x'}) \equiv \nabla'_{CL} f'(\mathbf{x'}) \qquad && \nabla'_{CL} \equiv \mathbf{e}^n \partial'_n \\ &\mathbf{G}(\mathbf{x}) = [\text{grad f}](\mathbf{x}) = \partial_n f(\mathbf{x})\, \hat{\mathbf{n}} = \hat{\mathbf{n}}\, \partial_n f(\mathbf{x}) = \nabla f(\mathbf{x}) && \nabla \equiv \hat{\mathbf{n}}\, \partial_n \end{aligned} \tag{10.1.6}$$

where the second line shows the usual Cartesian form of the gradient. The first line shows how one *could* define a "curvilinear gradient" operator $\nabla'_{CL} \equiv \mathbf{e}^n \partial'_n$, but this does not seem particularly useful. To restate,

$$[\text{grad f}](\mathbf{x}) \ = \ \partial'_n f\,'(\mathbf{x}')\, \mathbf{e}^n \qquad\qquad \text{// Curvilinear}$$
$$[\text{grad f}](\mathbf{x}) \ = \ \partial_n f(\mathbf{x})\, \hat{\mathbf{n}} \ . \qquad\qquad \text{// Cartesian} \qquad\qquad (10.1.7)$$

In the first line, the reciprocal vectors $\mathbf{e}^n$ exist in x-space, but the coefficients are expressed entirely in terms of x'-space coordinates and objects. Since $f(\mathbf{x})$ is a scalar field, $f(\mathbf{x}) = f\,'(\mathbf{x}')$, one could regard the derivative appearing in the first line as

$$\partial'_n f\,'(\mathbf{x}') \ = \partial'_n f(\mathbf{x}(\mathbf{x}')) \qquad\qquad \text{where } \ \mathbf{x}(\mathbf{x}') \ = F^{-1}(\mathbf{x}') \ . \qquad\qquad (10.1.8)$$

The *contravariant* components of grad f are then easily obtained as

$$G'^i(\mathbf{x}') = [\text{grad f}]'^i(\mathbf{x}') \ = \partial'^i f\,'(\mathbf{x}') \ = g'^{ij}(\mathbf{x}')\, \partial'_j f'(\mathbf{x}') \qquad\qquad \mathbf{G} = \text{grad f} \ = G'^i(\mathbf{x}')\, \mathbf{e}_i \qquad\qquad (10.1.9)$$

where now the expansion $G'^i(\mathbf{x}')\, \mathbf{e}_i$ is that of the 3rd line of (7.13.10).

If an expansion on unit vectors is desired, the right equation of (10.1.9) can be written, since $\mathbf{e}_i = h'_i \hat{\mathbf{e}}_i$,

$$\mathbf{G} = [\text{grad f}](\mathbf{x}) = G'^i \mathbf{e}_i = (G'^i\, h'_i)\, \hat{\mathbf{e}}_i \quad = \mathcal{G}'^i\, \hat{\mathbf{e}}_i \qquad\qquad \text{where} \quad \mathcal{G}'^i \equiv h'_i\, G'^i \qquad\qquad (10.1.10)$$

so then (10.1.9) reads,

$$\mathcal{G}'^i(\mathbf{x}')/h'_i = [\text{grad f}]'^i(\mathbf{x}') \ = \partial'^i f\,'(\mathbf{x}') \ = g'^{ij}(\mathbf{x}')\, \partial'_j f'(\mathbf{x}') \qquad\qquad \mathbf{G} = \text{grad f} \ = \mathcal{G}'^i\, \hat{\mathbf{e}}_i \ . \qquad\qquad (10.1.11)$$

Gathering up these results one gets

$$G'_i(\mathbf{x}') = [\text{grad f}]'_i(\mathbf{x}') = \ \partial'_i f\,'(\mathbf{x}') \qquad\qquad \mathbf{G} = \text{grad f} \ = \ G'_i(\mathbf{x}')\, \mathbf{e}^i$$
$$G'^i(\mathbf{x}') = [\text{grad f}]'^i(\mathbf{x}') \ = \partial'^i f\,'(\mathbf{x}') \ = g'^{ij}(\mathbf{x}')\, \partial'_j f'(\mathbf{x}') \qquad\qquad \mathbf{G} = \text{grad f} \ = G'^i(\mathbf{x}')\, \mathbf{e}_i$$
$$\mathcal{G}^{\,i}(\mathbf{x}') = h'_i[\text{grad f}]'^i(\mathbf{x}') \ = h'_i\, \partial'^i f\,'(\mathbf{x}') \ = h'_i\, g'^{ij}(\mathbf{x}')\, \partial'_j f'(\mathbf{x}') \qquad\qquad \mathbf{G} = \text{grad f} \ = \mathcal{G}'^i(\mathbf{x}')\, \hat{\mathbf{e}}_i$$
$$(10.1.12)$$

which can be rewritten

$$[\text{grad f}](\mathbf{x}) \ = (\partial'_i f\,'(\mathbf{x}'))\, \mathbf{e}^i$$
$$[\text{grad f}](\mathbf{x}) \ = (\partial'^i f\,'(\mathbf{x}'))\, \mathbf{e}_i \quad\ = g'^{ij}(\mathbf{x}')\, (\partial'_j f'(\mathbf{x}'))\, \mathbf{e}_i$$
$$[\text{grad f}](\mathbf{x}) \ = h'_i\, (\partial'^i f\,'(\mathbf{x}'))\, \hat{\mathbf{e}}_i \ \ = h'_i\, g'^{ij}(\mathbf{x}')\, (\partial'_j f'(\mathbf{x}'))\, \hat{\mathbf{e}}_i$$
$$[\text{grad f}](\mathbf{x}) = (\partial_i f(\mathbf{x}))\, \hat{\mathbf{n}} = \nabla f(\mathbf{x}) \qquad \text{// Cartesian}$$
$$[\text{grad f}]'_i(\mathbf{x}') \ = R_i{}^j[\text{grad f}]_j(\mathbf{x}) \qquad \text{// transformation} \qquad\qquad f\,'(\mathbf{x}') = f(\mathbf{x}) \qquad\qquad (10.1.13)$$

Again, in each of the first three forms above, $\mathbf{G} = [\text{grad f}](\mathbf{x})$ is being expressed as a linear combination of **e** vectors which are in x-space, but the coefficients are given entirely in terms of x'-space coordinates and objects. One can compare the last line of (10.1.13) to the last line of (9.28) that $[\text{div } \mathbf{B}](\mathbf{x}) \ = [\text{div } \mathbf{B}]'(\mathbf{x}')$, Since div **B** is a scalar, there is no mixture of components as in $[\text{grad f}]'_i(\mathbf{x}') \ = R_i{}^j[\text{grad f}]_j(\mathbf{x})$.

For the orthogonal case $g'^{ij} = (1/h'^2_i)\, \delta_{i,j}$ from (5.11.9) so the block above becomes

$[\text{grad f}](\mathbf{x}) = (\partial'_i f'(\mathbf{x}'))\, \mathbf{e}^i$

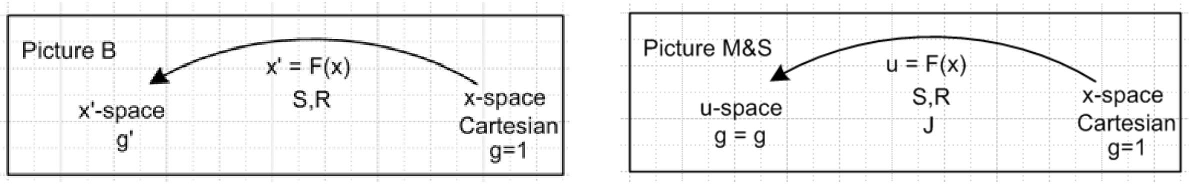
$[\text{grad f}](\mathbf{x}) = (\partial'^i f'(\mathbf{x}'))\, \mathbf{e}_i \qquad = (1/h'^2_i)\, (\partial'_i f'(\mathbf{x}'))\, \mathbf{e}_i$

$[\text{grad f}](\mathbf{x}) = h'_i\, (\partial'^i f'(\mathbf{x}'))\, \hat{\mathbf{e}}_i \ = \ (1/h'_i)\ (\partial'_i f'(\mathbf{x}'))\, \hat{\mathbf{e}}_i$

$[\text{grad f}](\mathbf{x}) = (\partial_i f(\mathbf{x}))\, \hat{\mathbf{n}}$ // Cartesian

$[\text{grad f}]'_i(\mathbf{x}') = R_i{}^j [\text{grad f}]_j(\mathbf{x})$ // transformation $f'(\mathbf{x}') = f(\mathbf{x})$ (10.1.14)

This above equations can be converted from Picture B to Picture M&S using the same rules given in (9.3.2), which we repeat below



(10.1.15)

- replace $\mathbf{x}'$ by $\mathbf{u}$ everywhere
- replace ∂'_n by ∂_n meaning $\partial/\partial u^n$ (exception: on a "Cartesian" line ∂_n means $\partial/\partial x^n$) *(9.3.2)*
- replace g' by g (both the scalar and the tensor) and h_n' by h_n
- put all primed tensor components (scalar, vector, etc) into unprimed italics (eg, $B'^n \rightarrow B^n$, $f' \rightarrow f$)

After this translation, all unprimed tensor components are functions of **x**, while all italicized tensor components are functions of **u**.

The translated results are then (all implied sums)

$[\text{grad f}](\mathbf{x}) = (\partial_i f)\, \mathbf{e}^i$

$[\text{grad f}](\mathbf{x}) = (\partial^i f)\, \mathbf{e}_i \qquad = g^{ij}\, (\partial_j f)\, \mathbf{e}_i$

$[\text{grad f}](\mathbf{x}) = h_i (\partial^i f)\, \hat{\mathbf{e}}_i \ = h_i\, g^{ij}\, (\partial_j f)\, \hat{\mathbf{e}}_i$

$[\text{grad f}](\mathbf{x}) = (\partial_i \text{f})\, \hat{\mathbf{n}}$ // Cartesian

$[\mathit{grad} f]_i(\mathbf{u}) = R_i{}^j [\text{grad f}]_j(\mathbf{x})$ // transformation (vector) $f(\mathbf{u}) = \text{f}(\mathbf{x}) = \text{f}(\mathbf{x}(\mathbf{u}))$ (10.1.16)

Notice that $[\text{grad f}]'_i(\mathbf{x}') \rightarrow [\mathit{grad} f]_i(\mathbf{u})$ according to the fourth rule, meaning $G'_i(\mathbf{x}') \rightarrow G_i(\mathbf{u})$.

For *orthogonal* curvilinear coordinates,

$[\text{grad f}](\mathbf{x}) = (\partial_i f)\, \mathbf{e}^i$

$[\text{grad f}](\mathbf{x}) = (\partial^i f)\, \mathbf{e}_i \qquad = (1/h_i^2)\, (\partial_i f)\, \mathbf{e}_i$

$[\text{grad f}](\mathbf{x}) = h_i (\partial^i f)\, \hat{\mathbf{e}}_i \ = (1/h_i)\ (\partial_i f)\, \hat{\mathbf{e}}_i$ // M&S 1.05 (10.1.17)

One can always make the replacement $f(\mathbf{u}) = \text{f}(\mathbf{x}(\mathbf{u}))$ in any of the above equations (f scalar). And one more time: the various **e** vectors are in x-space, but all the coefficients are expressed in curvilinear u-space coordinates and components. With the replacements

$$f \to \varphi \qquad \hat{\mathbf{e}}_i \to \mathbf{a}_i \qquad h_i \to \sqrt{g_{ii}}$$

the last equation of (10.1.17) agrees with Moon & Spencer p 2 (1.05).

10.2 Expressions for grad f • B

Sometimes one is interested in the following quantity (back to Picture B)

$$\text{grad f} \bullet \mathbf{B} \qquad (10.2.1)$$

where f is a tensorial scalar field and **B** is a tensorial vector field. In this case, since grad f is a tensorial vector field, the quantity grad f • **B** is a tensorial scalar, and so (grad f)' • **B**' = (grad f) • **B** .

This quantity grad f • **B** can be written several ways depending on how **B** is expanded:

$$\begin{aligned} &\mathbf{B} = \Sigma_i B'_i \mathbf{e}^i && \Rightarrow && \text{grad f} \bullet \mathbf{B} = (\partial'^n f')\, \mathbf{e}_n \bullet \Sigma_i B'_i \mathbf{e}^i = (\partial'^n f')\, B'_n \\ &\mathbf{B} = \Sigma_i B'^i \mathbf{e}_i && \Rightarrow && \text{grad f} \bullet \mathbf{B} = (\partial'_n f')\, \mathbf{e}^n \bullet \Sigma_i B'^i \mathbf{e}_i = (\partial'_n f')\, B'^n . \qquad (10.2.2) \end{aligned}$$

Here grad f comes from lines 2 and 1 of (10.1.14) and the **B** expansions from (7.13.10), and there is an implied sum on n. The last line can be written, using $\mathcal{B}'^i \equiv B'^i\ h'_i$,

$$\mathbf{B} = \Sigma_i [B'^i\ h'_i]\, \hat{\mathbf{e}}_i \equiv \Sigma_i \mathcal{B}'^i\, \hat{\mathbf{e}}_i \quad \Rightarrow \quad \text{grad f} \bullet \mathbf{B} = (\partial'_n f')\, B'^n = (\partial'_n f')\, (\mathcal{B}'^n/h'_n) \qquad (10.2.3)$$

To summarize:

$$\begin{aligned} &[\text{grad f}](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) = (\partial'^n f'(\mathbf{x}'))\, B'_n(\mathbf{x}') && \text{for } \mathbf{B} = \Sigma_i B'_i \mathbf{e}^i \\ &[\text{grad f}](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) = (\partial'_n f'(\mathbf{x}'))\, B'^n(\mathbf{x}') && \text{for } \mathbf{B} = \Sigma_i B'^i \mathbf{e}_i \\ &[\text{grad f}](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) = (\partial'_n f'(\mathbf{x}'))\, \mathcal{B}'^n(\mathbf{x}')/h'_n(\mathbf{x}') && \text{for } \mathbf{B} = \Sigma_i \mathcal{B}'^i \hat{\mathbf{e}}_i \\ &[\text{grad f}](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) = (\partial_n f(\mathbf{x}))\, B^n(\mathbf{x}) && \text{for } \mathbf{B} = B^n\, \hat{\mathbf{n}} \qquad \text{// Cartesian} \\ &[\text{grad f}](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) = [\text{grad f}]'(\mathbf{x}') \bullet \mathbf{B}'(\mathbf{x}') && \text{// scalar} \qquad (10.2.4) \end{aligned}$$

Since grad f • **B** is a scalar, these results resemble the divergence results more than the gradient ones. Everything on the right side of the first three equations involves only x'-space coordinates and components.

The conversion from Picture B to Picture M&F is straightforward using rules (9.3.2),

$$\begin{aligned} &[\text{grad f}](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) = (\partial^n f)\, \mathit{B}_n && && \mathbf{B} = \Sigma_i \mathit{B}_i \mathbf{e}^i \\ &[\text{grad f}](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) = (\partial_n f)\, \mathit{B}^n && && \mathbf{B} = \Sigma_i \mathit{B}^i \mathbf{e}_i \\ &[\text{grad f}](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) = (\partial_n f)\, \mathcal{B}^n/h_n && && \mathbf{B} = \Sigma_i \mathcal{B}^i \hat{\mathbf{e}}_i \\ &[\text{grad f}](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) = (\partial_n \text{f})\, B^n && \text{// Cartesian} && \mathbf{B} = B^n\, \hat{\mathbf{n}} \\ &[\text{grad f}](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) = [\boldsymbol{grad}\, f](\mathbf{u}) \bullet \boldsymbol{B}(\mathbf{u}) && \text{// transformation (scalar)} && \qquad (10.2.5) \end{aligned}$$

where once again $\mathit{f}(\mathbf{u}) = \mathrm{f}(\mathbf{x}(\mathbf{u}))$.

<u>Comment:</u> According to the Cartesian line above, one can write

$$\text{grad f} \bullet d\mathbf{x} = \partial_n f(\mathbf{x})\, dx^n = df = f(\mathbf{x}+d\mathbf{x}) - f(\mathbf{x}) \quad . \tag{10.2.6}$$

This equation df = [grad f]($\mathbf{x}$) • d$\mathbf{x}$ is sometimes used as an alternate definition of the gradient. If d$\mathbf{x}$ is selected to be in the direction of grad f, the dot product has its maximum value, and therefore the gradient points in the direction of the maximum change of a scalar function f($\mathbf{x}$). For N=2, in the usual 3D plot of real z = f(x,y) the gradient then points "uphill", and the negative of the gradient then points "downhill".

11. The Laplacian in curvilinear coordinates

We continue to use Picture B, see (11.9) below.

The Laplacian (also known as the Laplace-Beltrami operator) is defined by

$$\text{lap } f = \text{div (grad f)} = \text{div } \mathbf{G} \qquad \text{where} \qquad \mathbf{G} \equiv \text{grad f} \ . \tag{11.1}$$

Since f is (by assumption) a tensorial scalar field, **G** = grad f is a tensorial vector, as shown in (10.1.3). Then, as claimed in (9.1.4), div(grad f) = div **G** is a tensorial scalar, meaning [lap f]($\mathbf{x}$) = [lap f]'($\mathbf{x'}$). [In Chapter 15 we shall find that lap f = $f'^{i}{}_{;i}$ which in Cartesian space is $f'^{i}{}_{,i} = \partial_i\partial^i f = \partial_i{}^2 f$.]

In Cartesian coordinates one writes

$$\text{lap } f = \nabla^2 f = \nabla\bullet\nabla f = \Sigma_n \partial_n{}^2 f \tag{11.2}$$

but this form gets modified when lap f is expressed in curvilinear coordinates. The first line of (9.2.6) showed that,

$$\text{div } \mathbf{G} = [1/\sqrt{g'}\ ]\ \partial'_m [\sqrt{g'}\ G'^{m}] \qquad \text{where} \qquad \mathbf{G} = G'^{n}\mathbf{e}_n \qquad g' = \det(g'_{ab}) \ . \tag{11.3}$$

Meanwhile, the second line of (10.1.13) showed that,

$$\text{grad f} = \mathbf{G} = [g'^{nm} (\partial'_n f ')\ ]\ \mathbf{e}_m = G'^{m}\mathbf{e}_m \ , \quad G'^{m} = g'^{nm} (\partial'_n f ') \ , \quad \partial'_n f' = \partial'_n f(\mathbf{x}(\mathbf{x}')) = \partial'_n f '(\mathbf{x}') \ . \tag{11.4}$$

Therefore,

$$\text{lap f} = \text{div (grad f)} = \text{div } \mathbf{G} = [1/\sqrt{g'}\ ]\ \partial'_m [\sqrt{g'}\ G'^{m}] = [1/\sqrt{g'}\ ]\ \partial'_m [\sqrt{g'}\ g'^{nm} (\partial'_n f ')] \tag{11.5}$$

so the general results can be concisely stated:

$$\begin{aligned}
&[\text{lap f}](\mathbf{x}) = [1/\sqrt{g'(\mathbf{x}')}\ ]\ \partial'_m [\ \sqrt{g'(\mathbf{x}')}\ g'^{nm}(\mathbf{x}')\ (\partial'_n f '(\mathbf{x}'))\ ] && \text{// implied sum on n and m} \\
&[\text{lap f}](\mathbf{x}) = \Sigma_n \partial_n{}^2 f(\mathbf{x}) \qquad \text{// Cartesian} && f '(\mathbf{x}') = f(\mathbf{x}) = f(\mathbf{x}(\mathbf{x}')) \\
&[\text{lap f}](\mathbf{x}) = [\text{lap f}]'(\mathbf{x}') \qquad \text{// scalar} \ .
\end{aligned} \tag{11.6}$$

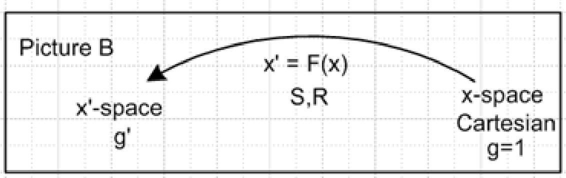
For an *orthogonal* coordinate system we know from (5.11.9) that

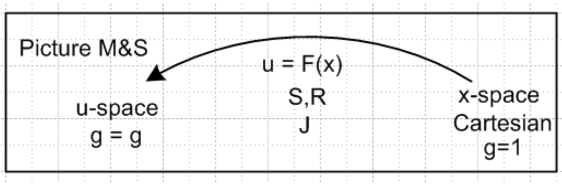
$$g'_{nm} = h'_n{}^2\ \delta_{n,m} \qquad g'^{nm} = (1/h'_n{}^2)\ \delta_{n,m} \qquad \sqrt{g'} = \Pi_i\ h'_i \tag{11.7}$$

so the first line of (11.6) simplifies to

$$[\text{lap f}](\mathbf{x}) = [1/(\Pi_i h'_i)]\ \ \partial'_m [(\Pi_i h'_i)\ (1/h'_m{}^2)\ (\partial'_m f ')\ ] \ . \tag{11.8}$$

Converting from Picture B to Picture MS with rules (9.3.2) gives,





(11.9)

$$[\text{lap f}](\mathbf{x}) = [1/\sqrt{g}\,]\ \partial_m\ [\sqrt{g}\, g^{nm}\, (\partial_n f)\,]$$

$$[\text{lap f}](\mathbf{x}) = \Sigma_n\, \partial_n^{\,2} f(\mathbf{x}) \qquad \text{// Cartesian} \qquad f(\mathbf{u}) = f(\mathbf{x}) = f(\mathbf{x}(\mathbf{u}))$$

$$[\text{lap f}](\mathbf{x}) = [lap\, f\,](\mathbf{u}) \qquad \text{// transformation (scalar)} \qquad (11.10)$$

The first equation simplifies in the orthogonal case to

$$[\text{lap f}](\mathbf{x}) = [1/\sqrt{g}\,]\ \partial_m\ [\sqrt{g}\ \ (1/h_m^{\,2})\, (\partial_m f)\,] \qquad \text{// orthogonal} \qquad \text{// M\&S 1.09} \qquad (11.11)$$

or

$$[\text{lap f}](\mathbf{x}) = [1/(\Pi_i h_i)]\ \partial_m\ [\ (\Pi_i h_i)\ (1/h_m^{\,2})\ (\partial_m f)\ ] \quad . \qquad (11.12)$$

For N=3 this says,

$$[\text{lap f}](\mathbf{x}) = 1/(h_1 h_2 h_3)\ \{\ \partial_1\ [\ (h_2 h_3/h_1)\ \partial_1 f\ \ ]\ \ + \text{cyclic}\ \ \}\ \ . \qquad (11.13)$$

With the replacements

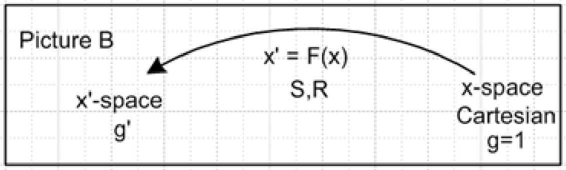
$$f \rightarrow \varphi \qquad\qquad h_m^{\,2} \rightarrow g_{mm}$$

equation (11.11) agrees with Moon & Spencer p 3 (1.09).

12. The Curl in curvilinear coordinates

12.1 Definition of curl B

The vector curl is defined only in N=3 dimensions (but see Section 12.6 below). Picture B is used.



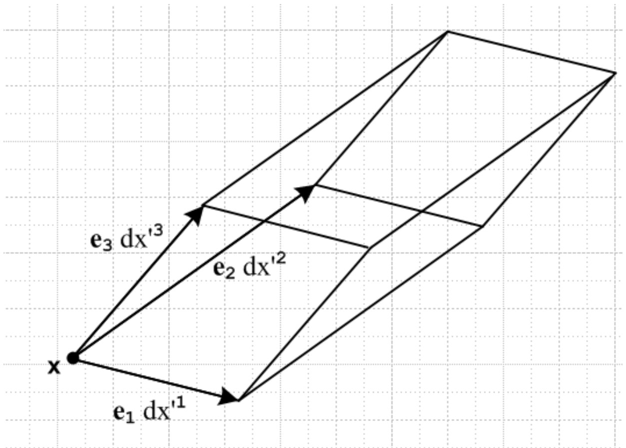
(12.1.1)

In Cartesian coordinates one writes

$$[\text{curl } \mathbf{B}]_i(\mathbf{x}) = [\nabla \text{ x } \mathbf{B}(\mathbf{x})]_i = \varepsilon_{ijk}\partial_j B_k(x) \; , \qquad (12.1.2)$$

but when expressed in terms of curvilinear coordinates and components, the form is different.

Consider the x-space differential 3-piped shown on the right side of Fig (8.2.2),



non-orthogonal differential N-piped in x-space

(12.1.3)

This 3-piped has three pairs of parallel faces. Within each pair, the "near" face touches the point **x** at which the tails of the spanning vectors meet, while the "far" face does not.

According to (8.4.a.1) we can write (all $dx'^i > 0$)

$$d\mathbf{A}^n = J\, \mathbf{e}^n \;\; (\Pi_{i\neq n} dx'^i) \; . \qquad (12.1.4)$$

According to (8.3.6) the vector - $d\mathbf{A}^n$ is the out-facing area vector for a near face. Therefore the in-facing area vector for a near face is $d\mathbf{A}^n$. For example, the area vector $d\mathbf{A}^3$ of the bottom face in (12.1.3) points up, because the bottom face is a near face and in-facing for it means up.

Consider now the line integral of a vector field **B**(x) around the boundary of *near* face n, where the circulation sense of the integral is determined from the right-hand-rule by the direction of $d\mathbf{A_n}$. Denote this line integral by

$$(\oint \mathbf{B}\bullet d\mathbf{x})^n \quad . \tag{12.1.5}$$

Sometimes this line integral is referred to as "the circulation" or "the rotation" of **B** around near face n (and rot **B** is another notation used for curl **B**). This terminology is suggestive of a fluid vortex where the fluid velocity **v** has a large curl component perpendicular to the vortex.

In x-space the quantity $\mathbf{C}(\mathbf{x}) \equiv \text{curl } \mathbf{B(x)}$ is a vector field defined in the following manner in the limit that all the differentials $dx'^n \rightarrow 0$:

$$\mathbf{C} \bullet d\mathbf{A}^n = (\oint \mathbf{B}\bullet d\mathbf{x})^n \qquad\qquad \mathbf{C} \equiv \text{curl } \mathbf{B} \qquad , \tag{12.1.6}$$

Since $d\mathbf{A}^n$ is given in (12.1.4) in terms of $\mathbf{e}^n$, **C** should be expanded on the $\mathbf{e_k}$. Since **C** is a vector density of weight -1, (D.2.9) shows that the proper expansion of **C** onto the e_k is given by

$$\mathbf{C} = J^{-1} \Sigma_{k=1}{}^3 C'^k \mathbf{e_k} \tag{12.1.7}$$

so that, using (12.1.4),

$$\mathbf{C} \bullet d\mathbf{A}^n = J^{-1} (\Sigma_k C'^k \mathbf{e_k}) \bullet (J \mathbf{e}^n (\Pi_{i\neq n} dx'^i)) = C'^n (\Pi_{i\neq n} dx'^i) \ . \tag{12.1.8}$$

Combining this with (12.1.6) gives,

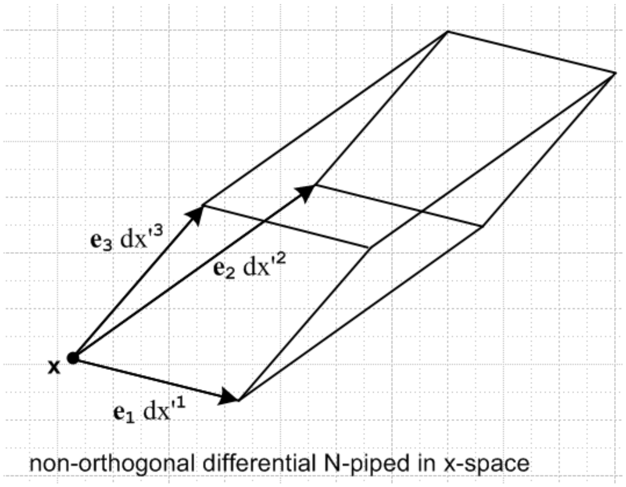
$$C'^n(\mathbf{x}') (\Pi_{i\neq n} dx'^i) = (\oint \mathbf{B}\bullet d\mathbf{x})^n \quad . \tag{12.1.9}$$

Since $(\Pi_{i\neq n}dx'^i) = dA_n$ is the area of face n of our differential N-piped, (12.1.9) says that the n^{th} component of the curl in x'space C'^n equals the line integral of **B** around the boundary of face n, divided by the area of face n, in the limit that all differentials go to zero.

Our task is now to compute this line integral and thereby come up with an expression for $C'^n(\mathbf{x}')$, the components of **C** = curl **B** when **B** is expanded onto the $\mathbf{e_k}$ in x-space.

12.2 Computation of the line integral

This shall be done for the bottom face (n=3) of the x-space differential 3-piped. As noted below (12.1.4), the area vector $d\mathbf{A}^3$ points up in Fig (12.1.3). Our circulation integral by the right hand rule will be counterclockwise around this bottom face boundary as seen in the figure, replicated here,



(12.2.1)

In this picture, it is intended that the upper right face is farthest from the viewer, so it is the "back" face of the 3-piped. It's parallel partner face touching the point **x** is the "front" face. We could call these faces near and far, but front and back seem better here. Then the bottom face about which we are integrating has edges front, right, back, and left.

So, grouping the contributions in pairs, one sees that.

$$(\oint \mathbf{B}\bullet d\mathbf{s})^3 \approx [\mathbf{B}(\mathbf{x}_{\mathtt{front}}) - \mathbf{B}(\mathbf{x}_{\mathtt{back}})] \bullet (\mathbf{e}_1\ dx'^1) + [\mathbf{B}(\mathbf{x}_{\mathtt{right}}) - \mathbf{B}(\mathbf{x}_{\mathtt{left}})] \bullet (\mathbf{e}_2\ dx'^2) \qquad (12.2.2)$$

where $\mathbf{B}(\mathbf{x}_{\mathtt{front}})$ refers to the value of **B** at the center of the front edge of the parallelogram which is the bottom face, and similarly for the other three edges. In the limit that the $dx'^n \rightarrow 0$, this simple approximation of the line integral is "good enough" to produce the desired results.

Motivated by (7.18.1) that $\mathbf{e}_i \bullet \mathbf{e}^j = \delta_i{}^j$, expand vector **B** as in the last line of (7.13.10),

$$\mathbf{B} = B'_j \mathbf{e}^j \qquad \text{where} \qquad B'_j(\mathbf{x}') = \mathbf{B}(\mathbf{x}) \bullet \mathbf{e}_j \qquad (12.2.3)$$

where the B'_j are the covariant components of **B** in x'-space. Inserting this expansion four times into (12.2.2) gives,

$$(\oint \mathbf{B}\bullet d\mathbf{x})^3 = [B'_1(\mathbf{x}'_{\mathtt{front}}) - B'_1(\mathbf{x}'_{\mathtt{back}})]\, dx'^1 + [B'_2(\mathbf{x}'_{\mathtt{right}}) - B'_2(\mathbf{x}'_{\mathtt{left}})]\, dx'^2 \qquad (12.2.4)$$

where $\mathbf{x}'_{\mathtt{front}} = \mathbf{F}(\mathbf{x}_{\mathtt{front}})$ and similarly for the other three points. In x-space the figure shows,

$$d\mathbf{x}_{\mathtt{BF}} \equiv \mathbf{x}_{\mathtt{back}} - \mathbf{x}_{\mathtt{front}} = \mathbf{e}_2 dx'^2$$
$$d\mathbf{x}_{\mathtt{RL}} \equiv \mathbf{x}_{\mathtt{right}} - \mathbf{x}_{\mathtt{left}} = \mathbf{e}_1 dx'^1 . \qquad (12.2.5)$$

Applying matrix R gives the corresponding x'-space equations (recall $d\mathbf{x}' = R d\mathbf{x}$ and $\mathbf{e}'_n = R\mathbf{e}_n$),

$$d\mathbf{x}'_{\mathtt{BF}} \equiv \mathbf{x}'_{\mathtt{back}} - \mathbf{x}'_{\mathtt{front}} = \mathbf{e}'_2 dx'^2 \qquad \mathbf{e}'_2 = (0,1,0...)$$
$$d\mathbf{x}'_{\mathtt{RL}} \equiv \mathbf{x}'_{\mathtt{right}} - \mathbf{x}'_{\mathtt{left}} = \mathbf{e}'_1 dx'^1 \qquad \mathbf{e}'_1 = (1,0,0...) . \qquad (12.2.6)$$

Using the fact that

$$f(\mathbf{x}'+d\mathbf{x}') \approx f(\mathbf{x}') + (\partial f/\partial x'^n)\, d\mathbf{x}'^n$$

for a variation $d\mathbf{x}' = \mathbf{e}'_n dx'^n$ (with no sum on n), one writes,

$$\begin{array}{llll} B'_1(\mathbf{x}'_{back}) \approx B'_1(\mathbf{x}'_{front}) + (\partial B'_1/\partial x'^2)\, dx'^2 & \text{where} \quad d\mathbf{x}' = \mathbf{e}'_2 dx'^2 & n=2 & \\ B'_2(\mathbf{x}'_{right}) \approx B'_2\, (\mathbf{x}'_{left}) + (\partial B'_2/\partial x'^1)\, dx'^1 & \text{where} \quad d\mathbf{x}' = \mathbf{e}'_1 dx'^1 & n=1\ . & \qquad (12.2.7) \end{array}$$

The circulation integral (12.2.4) is then

$$\begin{aligned} (\oint \mathbf{B}\bullet d\mathbf{x})^3 &\approx [-(\partial B'_1/\partial x'^2)\, dx'^2]\, dx'^1 + [(\partial B'_2/\partial x'^1)\, dx'^1]\, dx'^2 \\ &= [-(\partial B'_1/\partial x'^2) + (\partial B'_2/\partial x'^1)]\, dx'^1\, dx'^2 \quad = [-\partial'_2 B'_1 + \partial'_1 B'_2]\, dx'^1\, dx'^2 \\ &= [\partial'_1 B'_2 - \partial'_2 B'_1]\, dx'^1\, dx'^2 \\ &= \varepsilon^{3ab}\, \partial'_a B'_b\, (\Pi_{i\neq 3}\, dx'^i)\ . \end{aligned} \qquad (12.2.8)$$

Repeating this calculation for faces 1 and 2 produces cyclic results, and all three face line integrals can be summarized as (where equality holds in the limit $dx'^i \rightarrow 0$)

$$(\oint \mathbf{B}\bullet d\mathbf{x})^n = \varepsilon^{nab}\, \partial'_a B'_b\, (\Pi_{i\neq n}\, dx'^i)\ . \qquad (12.2.9)$$

Appendix D discusses the tensor ε known as the Levi-Civita ε tensor. In Cartesian space, the up and down position of the indices does not matter, as for any tensor. In non-Cartesian space up and down *does* matter, as with any tensor. The only fact needed here is (D.4.8) that $\varepsilon'^{abc..} = \varepsilon^{abc..}$ where ε' is the tensor in x'-space. In Cartesian space one can regard $\varepsilon^{abc..} = \varepsilon_{abc..}$ as a bookkeeping permutation tensor with the properties stated below (7.7.4) Installing then $\varepsilon^{nab} = \varepsilon'^{nab}$ into (12.2.9) yields,

$$(\oint \mathbf{B}\bullet d\mathbf{x})^n = \varepsilon'^{nab}\, \partial'_a B'_b\, (\Pi_{i\neq n}\, dx'^i) \qquad (12.2.10)$$

and this integral is then given entirely in terms of x'-space coordinates and objects.

12.3 Solving for the curl

Recall the expression (12.1.9) involving the curl component C'^n in x'-space,

$$C'^n(\mathbf{x}')\, (\Pi_{i\neq n}\, dx'^i) = (\oint \mathbf{B}\bullet d\mathbf{x})^n\ . \qquad \mathit{(12.1.9)}$$

Insert (12.2.10) to get

$$C'^{n}(\mathbf{x}')\ (\ \Pi_{i\neq n}\, dx'^{i}) = \ \varepsilon'^{nab}\, \partial'_{a}B'_{b}\ (\ \Pi_{i\neq n}\, dx'^{i})\ \ . \qquad (12.3.1)$$

The differentials cancel out, so *then* take $dx'^{i} \rightarrow 0$ and thus shrink the 3-piped down to the point of interest $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$ so that

$$C'^{n}\ = \varepsilon'^{nab}\, \partial'_{a}B'_{b}\ . \qquad (12.3.2)$$

Then from (12.1.7) we find, since $J = \sqrt{g'}$,

$$\text{curl}\ \mathbf{B}\ =\ \mathbf{C} =\ J^{-1}\ C'^{n}\ \mathbf{e}_{n} \qquad \text{// implied sum on n}$$

$$=\ [(1/\sqrt{g'}\ )\ \varepsilon'^{nab}\, \partial'_{a}B'_{b}\ ]\ \mathbf{e}_{n}\ \ . \qquad (12.3.3)$$

The comparison between the curvilinear- and Cartesian-expressed curls is this:

$$[\text{curl}\ \mathbf{B}](\mathbf{x})\ \ =\ [(1/\sqrt{g'}\ )\ \varepsilon'^{nab}\, \partial'_{a}B'_{b}(\mathbf{x}')\ ]\ \mathbf{e}_{n}\ =\ (1/\sqrt{g'}\ )\ \{\ \ [\partial'_{1}B'_{2}\ -\partial'_{2}B'_{1}]\ \mathbf{e}_{3}\ \ + \text{cyclic}\ \ \}$$

$$[\text{curl}\ \mathbf{B}](\mathbf{x})\ \ =\ \varepsilon^{nab}\, \partial_{a}B_{b}(\mathbf{x})\ \ \hat{\mathbf{n}} \qquad =\ \{\ [\ \partial_{1}B_{2}\ -\ \partial_{2}B_{1}\ ]\ \hat{\mathbf{3}}\ \ \ + \text{cyclic}\ \}\ . \quad (12.3.4)$$

Comment 1. The equation curl $\mathbf{B}\ =\ \mathbf{C} =\ [(1/\sqrt{g'}\ )\ \varepsilon'^{nab}\, \partial'_{a}B'_{b}\ ]\ \mathbf{e}_{n}$ obtained above assumed that **B** was a vector and the expansion $\mathbf{B} =\ B'_{n}\mathbf{e}^{n}$ was used. If **B** were a vector density of weight -1, the expansion would be $\mathbf{B} =\ J^{-1}B'_{n}\mathbf{e}^{n}$ and the result would be curl $\mathbf{B}\ =\ \mathbf{C} =\ [(1/\sqrt{g'}\ )\ \varepsilon'^{nab}\, \partial'_{a}(J^{-1}B'_{b})\ ]$. This situation will arise in consideration of the vector Laplacian in Chapter 13. Once again, $J = \sqrt{g'}$. (12.3.5)

Comment 2. Our "big result" of this Section is really (12.3.2) that $C'^{n}\ =\ \varepsilon'^{nab}\ \partial'_{a}B'_{b}$. We could have obtained this same result starting way back with (12.1.2) that $C^{n} = \varepsilon^{abc}\partial_{a}B_{b}$. The incantation is that this latter equation, being a tensor density equation, is covariant in the sense of (7.15.9) and therefore in x'-space it has the same form with everything primed. But doing it the hard way provides a good physical intuition for the curl of a vector field. (12.3.6)

12.4 Various forms of the curl

In (12.3.2) and (12.3.3) we found that,

$$C'^{n}\ \ = \varepsilon'^{nab}\, \partial'_{a}B'_{b} \qquad \mathbf{C} =\ J^{-1}C'^{n}\ \mathbf{e}_{n} \qquad \mathbf{B} =\ B'_{n}\ \mathbf{e}^{n}$$

$$\text{curl}\ \mathbf{B}\ \ =\ [(1/\sqrt{g'}\ )\, \varepsilon'^{nab}\, \partial'_{a}B'_{b}\ ]\ \mathbf{e}_{n}\ \ . \qquad (12.4.1)$$

If it is desired to have contravariant components of **B**, one gets

$$C'^{n}\ \ = \varepsilon'^{nab}\, \partial'_{a}(g'_{bc}B'^{c}\ ) \qquad \mathbf{C} =\ J^{-1}C'^{n}\ \mathbf{e}_{n} \qquad \mathbf{B} =\ B'^{n}\ \mathbf{e}_{n}$$

$$\text{curl}\ \mathbf{B}\ =\ [(1/\sqrt{g'}\ )\ \ \varepsilon'^{nab}\, \partial'_{a}(g'_{bc}B'^{c}\ )]\ \mathbf{e}_{n}\ . \qquad (12.4.2)$$

For practical applications, one usually wants both vectors expanded on the $\hat{\mathbf{e}}_{n}$ unit vectors in this way

$$\mathbf{C} = J^{-1}C'^{n}\ \mathbf{e_n} = (J^{-1}C'^{n}\, h'_{n})\ \hat{\mathbf{e}}_{n} \equiv \mathcal{C}'^{n}\ \hat{\mathbf{e}}_{n} \qquad \mathcal{C}'^{n} = h'_{n}\, J^{-1}C'^{n}$$
$$\mathbf{B} = B'^{n}\ \mathbf{e_n} = (B'^{n}\ h'_{n})\ \hat{\mathbf{e}}_{n} \equiv \mathcal{B}'^{n}\ \hat{\mathbf{e}}_{n} \qquad \mathcal{B}'^{n} = h'_{n}B'^{n}\ . \qquad (12.4.3)$$

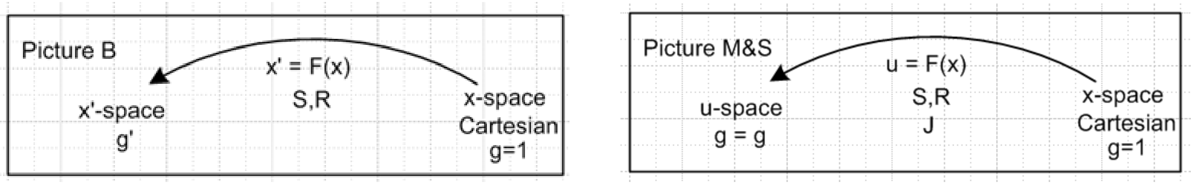
Using then $\mathcal{C}'^{n} = (J^{-1}h'_{n})C'^{n}$ and $\mathbf{C} = (J^{-1}h'_{n})C'^{n}\ \hat{\mathbf{e}}_{n}$ and $J = \sqrt{g'}$ one finds,

$$\mathcal{C}'^{n} = [(1/\sqrt{g'}\,)\, h'_{n}\, \varepsilon'^{nab}\, \partial'_{a}(g'_{bc}\, \mathcal{B}'^{c}/h'_{c}\,)] \qquad \mathbf{C} = \mathcal{C}'^{n}\ \hat{\mathbf{e}}_{n} \qquad \mathbf{B} = \mathcal{B}'^{n}\ \hat{\mathbf{e}}_{n}$$
$$\text{curl}\ \mathbf{B} = [(1/\sqrt{g'}\,)\, h'_{n}\, \varepsilon'^{nab}\, \partial'_{a}(g'_{bc}\, \mathcal{B}'^{c}/h'_{c}\,)]\ \hat{\mathbf{e}}_{n} \qquad \text{curl}\ \mathbf{B} = \mathbf{C}\ \ . \qquad (12.4.4)$$

To summarize: $B'^{c} = R^{c}{}_{d}B^{d}$ $\qquad g' = g'(\mathbf{x}')$ etc.

$$[\text{curl}\ \mathbf{B}](\mathbf{x}) = \varepsilon'^{nab}\, [(1/\sqrt{g'}\,)\, \partial'_{a}B'_{b}\,]\ \mathbf{e_n} \qquad \mathbf{B} = B'_{n}\mathbf{e}^{n} \qquad \textit{(12.4.1)}$$
$$[\text{curl}\ \mathbf{B}](\mathbf{x}) = \varepsilon'^{nab}\, [(1/\sqrt{g'}\,)\ \partial'_{a}(g'_{bc}B'^{c}\,)]\ \mathbf{e_n} \qquad \mathbf{B} = B'^{n}\mathbf{e_n} \qquad \textit{(12.4.2)}$$
$$[\text{curl}\ \mathbf{B}](\mathbf{x}) = \varepsilon'^{nab}\, [(1/\sqrt{g'}\,)\, h'_{n}\, \partial'_{a}(g'_{bc}\, \mathcal{B}'^{c}/h'_{c}\,)]\ \hat{\mathbf{e}}_{n} \qquad \mathbf{B} = \mathcal{B}'^{n}\ \hat{\mathbf{e}}_{n} \qquad \textit{(12.4.4)}$$
$$[\text{curl}\ \mathbf{B}](\mathbf{x}) = \varepsilon^{nab}\, \partial_{a}B_{b}(\mathbf{x})\ \ \hat{\mathbf{n}} \qquad \text{// Cartesian} \qquad \mathbf{B} = B^{n}\ \hat{\mathbf{n}} \qquad (12.4.5)$$

Converting from Picture B to Picture MS with rules (9.3.2),



(12.4.6)

one gets :

$$[\text{curl}\ \mathbf{B}](\mathbf{x}) = \varepsilon^{nab}\, [(1/\sqrt{g}\,)\, \partial_{a}\mathit{B}_{b}\,]\ \mathbf{e_n} \qquad \mathbf{B} = \mathit{B}_{n}\mathbf{e}^{n}$$
$$[\text{curl}\ \mathbf{B}](\mathbf{x}) = \varepsilon^{nab}\, [(1/\sqrt{g}\,)\, \partial_{a}(g_{bc}\mathit{B}^{c}\,)]\ \mathbf{e_n} \qquad \mathbf{B} = \mathit{B}^{n}\mathbf{e_n}$$
$$[\text{curl}\ \mathbf{B}](\mathbf{x}) = \varepsilon^{nab}[(1/\sqrt{g}\,)\, h_{n}\ \ \partial_{a}(g_{bc}\, \mathcal{B}^{c}/h_{c}\,)]\ \hat{\mathbf{e}}_{n} \qquad \mathbf{B} = \mathcal{B}^{n}\hat{\mathbf{e}}_{n}$$
$$[\text{curl}\ \mathbf{B}](\mathbf{x}) = \varepsilon^{nab}\, \partial_{a}B_{b}(\mathbf{x})\ \ \hat{\mathbf{n}} \qquad \text{// Cartesian} \qquad \mathbf{B} = B^{n}\ \hat{\mathbf{n}} \qquad (12.4.7)$$

Warning: The object ε^{nab} in the first three equations is now in u-space which is non-Cartesian, so up and down index positions do matter, but when indices are all up, it continues to be the normal permutation tensor.

Determinant Notation

Consider the following determinant,

$$\det(M) = \varepsilon^{abc}X_{a}Y_{b}Z_{c} = \begin{vmatrix} X_1 & X_2 & X_3 \\ Y_1 & Y_2 & Y_3 \\ Z_1 & Z_2 & Z_3 \end{vmatrix}\ .$$

If one regards X_n as a vector $\mathbf{X}_n$ with components $(\mathbf{X}_n)_i$ one could write the above line as three determinant equations,

$$Q_i = \det(M_i) = \varepsilon^{abc}(X_a)_i Y_b Z_c = \begin{vmatrix} (X_1)_i & (X_2)_i & (X_3)_i \\ Y_1 & Y_2 & Y_3 \\ Z_1 & Z_2 & Z_3 \end{vmatrix} \ , \quad i = 1,2,3 \ .$$

One could then combine these into a single vector equation,

$$\mathbf{Q} = \begin{vmatrix} \mathbf{X}_1 & \mathbf{X}_2 & \mathbf{X}_3 \\ Y_1 & Y_2 & Y_3 \\ Z_1 & Z_2 & Z_3 \end{vmatrix} = \varepsilon^{abc}\, \mathbf{X}_a Y_b Z_c \ .$$

As an example, suppose $\mathbf{X}_n = \mathbf{e}_n$, $Y_n = \partial_n$ (an operator), and $Z_n = B_c$. Then one has

$$\mathbf{Q} = \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ B_1 & B_2 & B_3 \end{vmatrix} = \varepsilon^{abc}\, \mathbf{X}_a Y_b Z_c = \varepsilon^{nab}\, \mathbf{X}_n Y_a Z_b = \varepsilon^{nab}\, \mathbf{e}_n(\partial_a B_b) \ .$$

But the expression on the right is recognized as $\sqrt{g}$ [curl $\mathbf{B}](\mathbf{x})$ from the first line of (12.4.7). It is in this notation that we now rewrite (12.4.7) in the traditional determinant notation:

$$[\text{curl } \mathbf{B}](\mathbf{x}) = (1/\sqrt{g}\,) \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ B_1 & B_2 & B_3 \end{vmatrix} \qquad \mathbf{B} = B_n\, \mathbf{e}^n \tag{12.4.8}$$

$$[\text{curl } \mathbf{B}](\mathbf{x}) = (1/\sqrt{g}\,) \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ g_{1c}B^c & g_{2c}B^c & g_{3c}B^c \end{vmatrix} \qquad \mathbf{B} = B^n\, \mathbf{e}_n \tag{12.4.9}$$

$$[\text{curl } \mathbf{B}](\mathbf{x}) = (1/\sqrt{g}\,) \begin{vmatrix} h_1\, \hat{\mathbf{e}}_1 & h_2\, \hat{\mathbf{e}}_2 & h_3\, \hat{\mathbf{e}}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ (g_{1c}/h_c)\, \mathcal{B}^c & (g_{2c}/h_c)\, \mathcal{B}^c & (g_{3c}/h_c)\, \mathcal{B}^c \end{vmatrix} \qquad \mathbf{B} = \mathcal{B}^n\, \hat{\mathbf{e}}_n \tag{12.4.10}$$

$$[\text{curl } \mathbf{B}](\mathbf{x}) = \begin{vmatrix} \hat{\mathbf{1}} & \hat{\mathbf{2}} & \hat{\mathbf{3}} \\ \partial_1 & \partial_2 & \partial_3 \\ B_1 & B_2 & B_3 \end{vmatrix} \qquad \text{// here } \partial_n = \partial/\partial x^n \text{ and } B_i = B_i(\mathbf{x}) \qquad \mathbf{B} = B^n\, \hat{\mathbf{n}} \tag{12.4.11}$$

12.5 The curl in orthogonal coordinate systems

For such systems $g_{ij} = \delta_{i,j} h_i^2$ and $\det(g_{ab}) = h_1^2 h_2^2 h_3^2$ so $\sqrt{g} = h_1 h_2 h_3$. It is then a simple matter to convert all the above forms and the results are:

Picture B: $\mathrm{B}'^{c}(\mathbf{x}') = \mathrm{R}^{c}{}_{d}\mathrm{B}^{d}(\mathbf{x})$ $\qquad h_i' = h_i'(\mathbf{x}')$ etc.

$$\begin{aligned}
&[\mathrm{curl}\ \mathbf{B}](\mathbf{x}) = \varepsilon'^{nab}\,[(1/(h_1'h_2'h_3')\,\partial'_a \mathrm{B}'_b\,]\,\mathbf{e}_n && \mathbf{B} = \mathrm{B}'_n\mathbf{e}^n \\
&[\mathrm{curl}\ \mathbf{B}](\mathbf{x}) = \varepsilon'^{nab}\,[(1/(h_1'h_2'h_3'))\ \partial'_a(h'_b{}^2\mathrm{B}'^b\,)]\,\mathbf{e}_n && \mathbf{B} = \mathrm{B}'^n\mathbf{e}_n \\
&[\mathrm{curl}\ \mathbf{B}](\mathbf{x}) = \varepsilon'^{nab}\,[(1/(h_1'h_2'h_3'))\,h'_n\,\partial'_a(h'_b\,\mathcal{B}'^b)\,)]\,\hat{\mathbf{e}}_n && \mathbf{B} = \mathcal{B}'^n\,\hat{\mathbf{e}}_n \\
&[\mathrm{curl}\ \mathbf{B}](\mathbf{x}) = \varepsilon^{nab}\,\partial_a \mathrm{B}_b(\mathbf{x})\ \hat{\mathbf{n}} \qquad // \text{ Cartesian} && \mathbf{B} = \mathrm{B}^n\,\hat{\mathbf{n}} \qquad (12.5.1)
\end{aligned}$$

Picture M&S (12.4.6) : $\mathcal{B}^{c}(\mathbf{u}) = \mathrm{R}^{c}{}_{d}\mathrm{B}^{d}(\mathbf{x})$ $\qquad h_i = h_i(\mathbf{u})$ etc.

$$\begin{aligned}
&[\mathrm{curl}\ \mathbf{B}](\mathbf{x}) = \varepsilon^{nab}\,[(h_1h_2h_3)^{-1}\,\partial_a B_b\,]\,\mathbf{e}_n && \mathbf{B} = B_n\mathbf{e}^n \\
&[\mathrm{curl}\ \mathbf{B}](\mathbf{x}) = \varepsilon^{nab}\,[(h_1h_2h_3)^{-1}\,\partial_a(h_b{}^2 B^b\,)]\,\mathbf{e}_n && \mathbf{B} = B^n\mathbf{e}_n \\
&[\mathrm{curl}\ \mathbf{B}](\mathbf{x}) = \varepsilon^{nab}[(h_1h_2h_3)^{-1}\,h_n\ \partial_a(h_b\,\mathcal{B}^b)]\,\hat{\mathbf{e}}_n && \mathbf{B} = \mathcal{B}^n\hat{\mathbf{e}}_n \qquad (12.5.2)
\end{aligned}$$

$$[\mathrm{curl}\ \mathbf{B}](\mathbf{x}) = (h_1h_2h_3)^{-1}\begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ B_1 & B_2 & B_3 \end{vmatrix} \qquad \mathbf{B} = B_n\ \mathbf{e}^n \qquad (12.5.3)$$

$$[\mathrm{curl}\ \mathbf{B}](\mathbf{x}) = (h_1h_2h_3)^{-1}\begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ h_1{}^2B^1 & h_2{}^2B^2 & h_3{}^2B^3 \end{vmatrix} \qquad \mathbf{B} = B^n\ \mathbf{e}_n \qquad (12.5.4)$$

$$[\mathrm{curl}\ \mathbf{B}](\mathbf{x}) = (h_1h_2h_3)^{-1}\begin{vmatrix} h_1\ \hat{\mathbf{e}}_1 & h_2\ \hat{\mathbf{e}}_2 & h_3\ \hat{\mathbf{e}}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ h_1\,\mathcal{B}^1 & h_2\,\mathcal{B}^2 & h_3\,\mathcal{B}^3 \end{vmatrix} \qquad \mathbf{B} = \mathcal{B}^n\ \hat{\mathbf{e}}_n \qquad // \text{ M\&S 1.07a}$$

$$(12.5.5)$$

With the replacements

$$\mathbf{B} \rightarrow \mathbf{E} \qquad \mathcal{B}^n \rightarrow \mathrm{E}_n \qquad \hat{\mathbf{e}}_n \rightarrow \mathbf{a}_n \qquad h_i \rightarrow \sqrt{g_{ii}} \qquad (h_1h_2h_3)^{-1} \rightarrow (1/\sqrt{g}\,)$$

equation (12.5.5) agrees with Moon & Spencer p 3 (1.07a).

12.6 The curl in N > 3 dimensions

Looking at the basic form (12.4.7) of the curl,

$$[\mathrm{curl}\ \mathbf{B}]^n(\mathbf{x}) = \varepsilon^{nab}\,\partial_a \mathrm{B}_b(\mathbf{x}) \qquad // \text{ Cartesian} \qquad (12.6.1)$$

it is hard to imagine a generalization to N>3 dimensions where the curl is still a vector. The only vectors available for construction purposes are ∂_n and B_n. For N=4 one might try out various generalizing forms

$$[\mathrm{curl}\ \mathbf{B}]^n(\mathbf{x}) = (1/\sqrt{g}\,)\ \varepsilon^{nabc}\,\partial_a(\partial_b\,\mathrm{B}_c) = (1/\sqrt{g}\,)\ \varepsilon^{nabc}\,\partial_a\partial_b\,\mathrm{B}_c \qquad ?$$

$$[\mathrm{curl}\ \mathbf{B}]^n(\mathbf{x}) = (1/\sqrt{g}\,)\ \varepsilon^{nabc}\,\partial_a\,(\mathrm{B}_b\mathrm{B}_c)) \qquad ?$$

but these two forms vanish because antisymmetric ε is contracted against something symmetric. Thus the idea of using multiple cross products as used in Appendix A does not prove helpful.

At this point we call upon the Appendix F notion of covariant derivatives and related semicolons.

The rank-2 tensor $B_{b;a} - B_{a;b} = \partial_a B_b - \partial_b B_a$ discussed in (D.8.5) provides the logical extension of the curl to N > 3 dimensions. For N=3 it *happens* that the object can be *associated with* a vector curl,

$$[\mathrm{curl}\ \mathbf{B}]^n = \varepsilon^{nab}\ [B_{b;a} - B_{a;b}]/2 = \varepsilon^{nab} B_{b;a} = \varepsilon^{nab}\ [\partial_a B_b - \partial_b B_a]/2 = \varepsilon^{nab}\partial_a B_b\ . \quad (12.6.2)$$

In relativity where N=4, there *is* no vector curl, and one sees $B_{b;a} - B_{a;b}$ referred to as the covariant curl, and $\partial_a B_b - \partial_b B_a$ as the ordinary curl (Weinberg p 106 (4.7.2)).

Writing the N-dimensional contravariant curl components in this manner in Cartesian x-space,

$$[\mathrm{curl}\ B]^{ij} = (B^{j;i} - B^{i;j}) \quad (12.6.3)$$

one can then ask how this generalized curl would be expressed in terms of x'-space coordinates and objects. (This curl is a regular rank-2 tensor with weight 0, the vector curl has weight -1).

The general issue of expanding tensors is addressed in Appendix E where this general result is obtained in (E.2.11),

$$A = \Sigma_{ijk...} A'^{ijk...} (\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k ...)\ . \qquad A'^{ijk...} = \text{contravariant components of A in x'-space} \quad (12.6.4)$$

Applying this to A = [curl B] one gets

$$[\mathrm{curl}\ B] = \Sigma_{ij}[\mathrm{curl}\ B]'^{ij}\mathbf{e}_i\otimes\mathbf{e}_j = \Sigma_{ij}(B'^{j;i} - B'^{i;j})\mathbf{e}_i\otimes\mathbf{e}_j\ . \quad (12.6.5)$$

The Cartesian components of curl B in x-space can then be expressed in terms of x'-space components and coordinates,

$$\begin{aligned}[\mathrm{curl}\ B]^{ab}(\mathbf{x}) &= \Sigma_{ij}\ [\mathrm{curl}\ B]'^{ij}\ (\mathbf{e}_i)^a(\mathbf{e}_j)^b \\ &= \Sigma_{ij}(B'^{j;i} - B'^{i;j})(\mathbf{e}_i)^a(\mathbf{e}_j)^b \quad (12.6.6)\end{aligned}$$

where the tangent base vectors $\mathbf{e}_n$ exist as usual in x-space, and $B'^i = B'^i(\mathbf{x}')$ where $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. Using (F.9.5) one finds that

$$(B'^{a;b} - B'^{b;a}) = (\ \partial'^a B'^b - \partial'^b B'^a\ ) + (\ g'^{ac}\Gamma'^b{}_{cn} - g'^{bc}\Gamma'^a{}_{cn})B'^n \quad (12.6.7)$$

where Γ is the affine connection. As in previous Sections we could write this in various ways involving components B'_n, B'^n, and $\mathcal{B}'^n$. Since (12.6.7) is covariant, in x-space it takes the same form with no primes

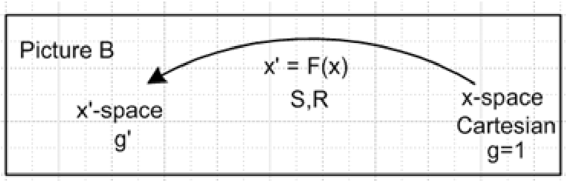
$$\begin{aligned}(B^{a;b} - B^{b;a}) &= (\ \partial^a B^b - \partial^b B^a\ ) + (\ g^{ac}\Gamma^b{}_{cn} - g^{bc}\Gamma^a{}_{cn})B^n \quad (12.6.8) \\ &= (\ \partial^a B^b - \partial^b B^a\ )\ .\end{aligned}$$

The second line is due to having $\Gamma = 0$ by (F.4.16) for x-space being Cartesian with $g = 1$. This is also true for a quasi-Cartesian x-space having $g = G$, as discussed in (1.10), because the x-space basis vectors are then constants and Γ involves their spatial derivatives.

13. The Vector Laplacian in curvilinear coordinates

13.1 Derivation of the Vector Laplacian in general curvilinear coordinates

The vector Laplacian is defined in terms of the vector curl which is only defined for N=3 dimensions. The context is Picture B,



(13.1.1)

The definition of the vector Laplacian of a vector field **B**(**x**) is

$$\nabla^2\mathbf{B} \equiv \text{grad}(\text{div}\,\mathbf{B}) - \text{curl}\,(\text{curl}\,\mathbf{B})\ . \qquad (13.1.2)$$

If **B** is a vector, div **B** is a scalar, grad(div **B**) is a vector, so we expect $\nabla^2\mathbf{B}$ to be a vector. An annoying observation is that, according to Section D.8, curl **B** is a vector density of weight -1 and therefore curl(curl**B**) a vector density of weight -2, and so apples are being subtracted from oranges to make $\nabla^2\mathbf{B}$**.** This tensor conundrum is resolved in Section 15.7 so we ignore it for now and proceed undaunted.

In Cartesian coordinates, one finds that, as proven below (15.7.3),

$$[\nabla^2\mathbf{B}]^i = \nabla^2(B^i) \equiv \Sigma_n\, \partial_n{}^2 B^i\ , \qquad (13.1.3)$$

but expressed in general curvilinear coordinates the form gets modified.

To avoid confusion, some authors use different symbols for the vector Laplacian operator. For example, Moon and Spencer use ✡ in place of ∇^2 and we will honor these authors by using that symbol here, so

$$\text{✡}\mathbf{B} \equiv \text{grad}(\text{div}\,\mathbf{B}) - \text{curl}\,(\text{curl}\,\mathbf{B})\ . \qquad (13.1.4)$$

In order to make use of the results of earlier Sections, define

$$\mathbf{G} \equiv \text{grad}(f) \quad \text{where} \quad f = \text{div}\,\mathbf{B} \qquad (13.1.5)$$

$$\mathbf{V} \equiv \text{curl}\,\mathbf{C} \quad \text{where} \quad \mathbf{C} \equiv \text{curl}\,\mathbf{B} \qquad (13.1.6)$$

so that

$$\text{✡}\mathbf{B} = \mathbf{G} - \mathbf{V}\ . \qquad (13.1.7)$$

The first line of (10.1.13) gives this expression for **G**,

$$\mathbf{G} = \text{grad(f)} = (\partial'_{\mathbf{k}} f\,')\ \mathbf{e}^{\mathbf{k}} \ . \tag{13.1.8}$$

Now replace f 'using the first line of (9.2.6)

$$f\,' = f\,'(\mathbf{x}') = f(\mathbf{x}) = \text{div}\ \mathbf{B} = [1/\sqrt{g'}\ ]\ \partial'_{\mathbf{i}}\ [\sqrt{g'}\ \ B'^{\mathbf{i}}] \tag{13.1.9}$$

so that

$$\begin{aligned} \mathbf{G} &= \partial'_{\mathbf{k}}\{\ (1/\sqrt{g'}\ )\ \partial'_{\mathbf{i}}\ (\sqrt{g'}\ \ B'^{\mathbf{i}})\}\ \ \mathbf{e}^{\mathbf{k}} \\ &= \partial'^{\mathbf{n}}\{\ (1/\sqrt{g'}\ )\ \partial'_{\mathbf{i}}\ (\sqrt{g'}\ \ B'^{\mathbf{i}})\}\ \ \mathbf{e}_{\mathbf{n}}\ . \end{aligned} \tag{13.1.10}$$

The second term **V** is little more complicated. First, from the first line of (12.4.5) write,

$$\begin{aligned} \mathbf{C} &= \text{curl}\ \mathbf{B} = \varepsilon'^{\mathbf{nab}}\ [(1/\sqrt{g'}\ )\ \ \partial'_{\mathbf{a}}\{\ \ \ B'_{\mathbf{b}}\}\ ]\ \mathbf{e}_{\mathbf{n}}\ \ = J^{-1}C'^{\mathbf{n}}\ \mathbf{e}_{\mathbf{n}}\ \ \ = (1/\sqrt{g'}\ )C'^{\mathbf{n}}\ \mathbf{e}_{\mathbf{n}} \qquad //\ J = \sqrt{g'} \\ \mathbf{V} &= \text{curl}\ \mathbf{C} = \varepsilon'^{\mathbf{ncd}}\ [(1/\sqrt{g'}\ )\ \ \partial'_{\mathbf{c}}\{\ J^{-1}C'_{\mathbf{d}}\}\ ]\ \mathbf{e}_{\mathbf{n}}\ \ = J^{-2}V'^{\mathbf{n}}\ \mathbf{e}_{\mathbf{n}}\ \ =\ \ (1/g')\ \ V'^{\mathbf{n}}\ \mathbf{e}_{\mathbf{n}} \end{aligned} \tag{13.1.11}$$

where recall from (D.4.8) that $\varepsilon'^{\mathbf{abc}\cdots} = \varepsilon^{\mathbf{abc}\cdot\cdot}$ = the usual permutation tensor, but written up and primed so as to be in covariant form. The factors of J appearing in the expansions on the far right are due to the fact that **C** is a vector density of weight - 1 and **V** a vector density of weight -2, see (D.2.9).

From (13.1.11) we extract these facts,

$$C'^{\mathbf{n}} = \varepsilon'^{\mathbf{nab}}\ \partial'_{\mathbf{a}}\{\ B'_{\mathbf{b}}\} \tag{13.1.12}$$
$$V'^{\mathbf{n}} = \sqrt{g'}\ \varepsilon'^{\mathbf{ncd}}\ \partial'_{\mathbf{c}}\{\ (1/\sqrt{g'}\ )C'_{\mathbf{d}}\}\ \ . \tag{13.1.13}$$

Now lower the index on $C'^{\mathbf{n}}$ in (13.1.12) to get

$$C'_{\mathbf{d}} = g'_{\mathbf{de}}C'^{\mathbf{e}}\ \ = g'_{\mathbf{de}}\ \varepsilon'^{\mathbf{eab}}\ (\partial'_{\mathbf{a}}B'_{\mathbf{b}})\ , \tag{13.1.14}$$

and insert this into (13.1.13) to get

$$\begin{aligned} V'^{\mathbf{n}} &= \sqrt{g'}\ \varepsilon'^{\mathbf{ncd}}\ \partial'_{\mathbf{c}}\{\ (1/\sqrt{g'}\ )\ g'_{\mathbf{de}}\ \varepsilon'^{\mathbf{eab}}\ (\partial'_{\mathbf{a}}B'_{\mathbf{b}})\ \} \\ &= \sqrt{g'}\ \ \varepsilon'^{\mathbf{ncd}}\ \varepsilon'^{\mathbf{eab}}\ \partial'_{\mathbf{c}}\ \{\ (1/\sqrt{g'}\ )\ g'_{\mathbf{de}}(\partial'_{\mathbf{a}}B'_{\mathbf{b}})\ \}\ . \end{aligned} \tag{13.1.15}$$

Then from (13.1.11) we have this expression for **V**,

$$\mathbf{V} = \text{curl}\ \mathbf{C} =\ (1/g')\ \ V'^{\mathbf{n}}\ \mathbf{e}_{\mathbf{n}}\ \ = [\ (1/\sqrt{g'}\ )\varepsilon'^{\mathbf{ncd}}\ \varepsilon'^{\mathbf{eab}}\ \partial'_{\mathbf{c}}\ \{\ (1/\sqrt{g'}\ )\ g'_{\mathbf{de}}(\partial'_{\mathbf{a}}B'_{\mathbf{b}})\ \}]\ \mathbf{e}_{\mathbf{n}}\ \ . \tag{13.1.16}$$

We can then summarize our results:

$$\text{✡}\mathbf{B} = \mathbf{G} - \mathbf{V} \qquad \mathit{(13.1.7)}$$
$$\mathbf{G} = \partial'^{n}\{ (1/\sqrt{g'})\, \partial'_{i}\, (\sqrt{g'}\ B'^{i})\}\ \mathbf{e_n} \qquad \mathit{(13.1.10)} \qquad \mathbf{B} = B'^{n}\mathbf{e_n}$$
$$\mathbf{V} = [\, (1/\sqrt{g'})\, \varepsilon'^{ncd}\, \varepsilon'^{eab}\, \partial'_{c}\, \{ (1/\sqrt{g'})\, g'_{de}(\partial'_{a}B'_{b}) \}]\, \mathbf{e_n} \qquad \mathit{(13.1.16)} \qquad \mathbf{B} = B'_{n}\mathbf{e^n}$$
$$= [\, (1/\sqrt{g'})\, \varepsilon'^{ncd}\, \varepsilon'^{eab}\, \partial'_{c}\, \{ (1/\sqrt{g'})\, g'_{de}(\partial'_{a}[g'_{bf}B'^{f}]) \}]\, \mathbf{e_n}\ . \qquad \mathbf{B} = B'^{n}\mathbf{e_n} \qquad (13.1.17)$$

Often one wants vectors expanded onto the unit vectors $\hat{\mathbf{e}}_{n}$

$$\mathbf{B} = B'^{n}\ \mathbf{e_n} = B'^{n}\ h'_{n}\, \hat{\mathbf{e}}_{n} = \mathcal{B}'^{n}\hat{\mathbf{e}}_{n} \qquad \Rightarrow \qquad B'^{n} = \mathcal{B}'^{n}/h'_{n} \qquad (13.1.18)$$

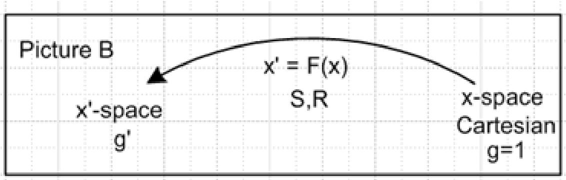
and the above set of equations becomes

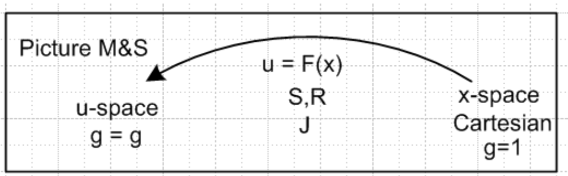
$$\text{✡}\mathbf{B} = \mathbf{G} - \mathbf{V}$$
$$\mathbf{G} = h'_{n}\ \partial'^{n}\{ (1/\sqrt{g'})\, \partial'_{i}\, (\sqrt{g'}\, \mathcal{B}'^{i}/h'_{i})\}\ \hat{\mathbf{e}}_{n} \qquad \mathbf{B} = \mathcal{B}'^{n}\ \hat{\mathbf{e}}_{n}$$
$$\mathbf{V} = h'_{n}\, [\, (1/\sqrt{g'})\, \varepsilon'^{ncd}\, \varepsilon'^{eab}\, \partial'_{c}\, \{(1/\sqrt{g'})\, g'_{de}(\partial'_{a}[g'_{bf}\, \mathcal{B}'^{f}/h'_{f}]) \}]\, \hat{\mathbf{e}}_{n}\ . \qquad (13.1.19)$$

To this list we can add the Cartesian form

$$\text{✡}\mathbf{B} = [\nabla^2 B^{n}\,]\, \hat{\mathbf{n}} \quad . \qquad // \text{ Cartesian, } \hat{\mathbf{n}} = \mathbf{u_n} \qquad \mathbf{B} = B^{n}\, \hat{\mathbf{n}} \qquad (13.1.20)$$

Converting from Picture B to Picture MS using rules (9.3.2) gives,





(13.1.21)

$$\text{✡}\mathbf{B} = \mathbf{G} - \mathbf{V}$$

$$\mathbf{G} = \partial^{n}\{ (1/\sqrt{g})\, \partial_{i}\, (\sqrt{g}\ B^{i})\}\ \mathbf{e_n} \qquad \mathbf{B} = B^{n}\mathbf{e_n}$$
$$\mathbf{V} = [\, (1/\sqrt{g})\, \varepsilon^{ncd}\, \varepsilon^{eab}\, \partial_{c}\, \{\ (1/\sqrt{g})\, g_{de}(\partial_{a}[g_{bf}B^{f}]) \}]\, \mathbf{e_n}$$

$$\mathbf{G} = h_{n}\ \partial^{n}\{ (1/\sqrt{g})\, \partial_{i}\, (\sqrt{g}\, \mathcal{B}^{i}/h_{i})\}\ \hat{\mathbf{e}}_{n} \qquad \mathbf{B} = \mathcal{B}^{n}\ \hat{\mathbf{e}}_{n}$$
$$\mathbf{V} = h_{n}\, [\, (1/\sqrt{g})\, \varepsilon^{ncd}\, \varepsilon^{eab}\, \partial_{c}\, \{\ (1/\sqrt{g})\, g_{de}(\partial_{a}[g_{bf}\mathcal{B}^{f}/h_{f}]) \}]\, \hat{\mathbf{e}}_{n}$$

$$\text{✡}\mathbf{B} = [\nabla^2 B^{n}\,]\, \hat{\mathbf{n}} \qquad // \text{ Cartesian, } \hat{\mathbf{n}} = \mathbf{u_n} \qquad \mathbf{B} = B^{n}\, \hat{\mathbf{n}} \qquad (13.1.22)$$

In the equations above, all the ∂_{i} mean $\partial/\partial u^{i}$ and the argument **u** of all functions is suppressed. The Cartesian form will be verified below.

<u>Comment</u>: The above general statement of $\text{✡}\mathbf{B}$ in curvilinear coordinates is amazingly complicated. Using (D.10.37-39), one could replace $\varepsilon^{ncd}\,\varepsilon^{eab} = \delta^{(ncd;eab)} = \delta^{ncd}_{eab} = \delta_{n,e}\delta_{c,a}\delta_{d,b}$ + 5 similar terms,

but this just expands the **V** term into six terms in place of one. The alternate path of starting with $[\text{✡}\mathbf{B}]^n$ = $B^{'n;j}{}_{;j}$ as shown in Section 15.8 leads to a result involving the affine connection Γ, as well as $\partial_i\Gamma$ and ΓΓ terms and is also quite unpleasant. We must accept the fact that ✡**B** is a complicated object which complicates the study of differential equations (such as Navier-Stokes) in general coordinates.

13.2 The Vector Laplacian in orthogonal curvilinear coordinates

We continue in Picture M&S and shall use only the $\hat{\mathbf{e}}_n$ expansion. Setting $g_{ij} = h_i^2\delta_{i,j}$ and $g^{ij} = (1/h_i)^2\delta_{i,j}$ from (5.11.8), our equation (13.1.22) simplifies somewhat,

$$\begin{aligned}\text{✡}\mathbf{B} = \mathbf{G} - \mathbf{V} = h_n\ &[g^{nk}\,\partial_k\{(1/\sqrt{g})\,\partial_i\,(\sqrt{g}\,\mathcal{B}^i(\mathbf{u})/h_i)\} \\ &- (1/\sqrt{g})\ \varepsilon^{ncd}\,\varepsilon^{eab}\ \partial_c\{(1/\sqrt{g})\,g_{de}\,(\partial_a\,[g_{bf}\,\mathcal{B}^f(\mathbf{u})/h_f])\}\,]\hat{\mathbf{e}}_n\end{aligned}$$

$$\begin{aligned}= h_n\ &[\delta_{k,n}\,\partial_k\{(1/\sqrt{g})\,\partial_i\,(\sqrt{g}\,\mathcal{B}^i(\mathbf{u})/h_i)\}\ (1/h_n)^2 \\ &- (1/\sqrt{g})\ \varepsilon^{ncd}\,\varepsilon^{eab}\ \partial_c\{(1/\sqrt{g})\,h_d{}^2\delta_{d,e}\,(\partial_a\,[h_b{}^2\delta_{b,f}\,\mathcal{B}^f(\mathbf{u})/h_f])\}\,]\ \hat{\mathbf{e}}_n\end{aligned}$$

$$\begin{aligned}= h_n\ &[\,\partial_n\{(1/\sqrt{g})\,\partial_i\,(\sqrt{g}\,\mathcal{B}^i(\mathbf{u})/h_i)\}\ \ (1/h_n)^2 \\ &- (1/\sqrt{g})\ \varepsilon^{ncd}\,\varepsilon^{dab}\ \partial_c\{(1/\sqrt{g})\,h_d{}^2\,(\partial_a\,[h_b\,\mathcal{B}^b(\mathbf{u})])\}\,]\ \hat{\mathbf{e}}_n\end{aligned}$$

$$= [\,(1/h_n)\,\partial_n\{(1/\sqrt{g})\,\partial_i\,(\sqrt{g}\,\mathcal{B}^i/h_i)\}\ -\ (h_n/\sqrt{g})\,\varepsilon^{ncd}\,\varepsilon^{dab}\ \partial_c\{(1/\sqrt{g})\,h_d{}^2\,(\partial_a\,[h_b\,\mathcal{B}^b])\}\,]\ \hat{\mathbf{e}}_n\,.$$

$$= [\ A_n - (h_n/\sqrt{g})\,B_n\,]\ \hat{\mathbf{e}}_n \tag{13.2.1}$$

where

$$\begin{aligned}A_n &\equiv (1/h_n)\,\partial_n\{(1/\sqrt{g})\,\partial_i\,(\sqrt{g}\,\mathcal{B}^i/h_i)\} \\ B_n &\equiv \varepsilon^{ncd}\,\varepsilon^{dab}\ \partial_c\{(1/\sqrt{g})\,h_d{}^2\,(\partial_a\,[h_b\,\mathcal{B}^b])\}\,.\end{aligned} \tag{13.2.2}$$

Now define,

$$\begin{aligned}T &\equiv (1/\sqrt{g})\,\partial_i\,(\sqrt{g}\,\mathcal{B}^i/h_i) \\ \Gamma_d &\equiv (1/\sqrt{g})\,h_d{}^2\,\varepsilon^{dab}\{(\partial_a\,[h_b\,\mathcal{B}^b])\}\,)\end{aligned} \tag{13.2.3}$$

so that

$$\begin{aligned}A_n &= (1/h_n)\,\partial_n T \\ B_n &= \varepsilon^{ncd}\,\partial_c\Gamma_d\,.\end{aligned} \tag{13.2.4}$$

Then (13.2.1) says

$$\begin{aligned}&\text{✡}\mathbf{B} = [\ (1/h_n)\,\partial_n T\ -\ (h_n/\sqrt{g})\ \varepsilon^{ncd}\,\partial_c\Gamma_d\,]\ \hat{\mathbf{e}}_n \\ &\text{where } T \equiv (1/\sqrt{g})\,\partial_i\,(\sqrt{g}\,\mathcal{B}^i/h_i)\quad \text{and} \\ &\qquad \Gamma_d \equiv (1/\sqrt{g})\,h_d{}^2\,\varepsilon^{dab}\{(\partial_a\,[h_b\,\mathcal{B}^b])\}\,)\ .\end{aligned} \tag{13.2.5}$$

It is conventional to write this out more specifically as

$$\text{✡}\mathbf{B} = [\ (1/h_1)\,\partial_1 T - (h_1/\sqrt{g})\ \varepsilon^{1cd}\,\partial_c\Gamma_d\ ]\ \hat{\mathbf{e}}_1 + \text{cyclic}$$

But $\varepsilon^{1cd}\,\partial_c\Gamma_d = (\partial_2\Gamma_3 - \partial_3\Gamma_2)$ so

$$\text{✡}\mathbf{B} = [\ (1/h_1)\,\partial_1 T - (h_1/\sqrt{g})\ (\partial_2\Gamma_3 - \partial_3\Gamma_2)]\ \hat{\mathbf{e}}_1 + \text{cyclic}$$

$$\begin{aligned} &= [(1/h_1)\,\partial_1 T + (h_1/\sqrt{g})\,(\partial_3\,\Gamma_2 - \partial_2\,\Gamma_3)\,]\ \hat{\mathbf{e}}_1 \\ &\quad + [(1/h_2)\,\partial_2 T + (h_2/\sqrt{g})\,(\partial_1\,\Gamma_3 - \partial_3\,\Gamma_1)\,]\ \hat{\mathbf{e}}_2 \\ &\quad + [(1/h_3)\,\partial_3 T + (h_3/\sqrt{g})\,(\partial_2\,\Gamma_1 - \partial_1\,\Gamma_2)\,]\ \hat{\mathbf{e}}_3 \quad . \qquad \text{// M\&S 1.11} \qquad (13.2.6) \end{aligned}$$

Writing out the Γ_d from (13.2.3),

$$\begin{aligned} \Gamma_1 &= (1/\sqrt{g})\,h_1^2\,(\,\partial_2\,[h_3\,\mathcal{B}^3] - \partial_3\,[h_2\,\mathcal{B}^2]\,) \\ \Gamma_2 &= (1/\sqrt{g})\,h_2^2\,(\,\partial_3\,[h_1\,\mathcal{B}^1] - \partial_1\,[h_3\,\mathcal{B}^3]\,) \\ \Gamma_3 &= (1/\sqrt{g})\,h_3^2\,(\,\partial_1\,[h_2\,\mathcal{B}^2] - \partial_2\,[h_1\,\mathcal{B}^1]\,)\ . \end{aligned}$$

$$T = (1/\sqrt{g})\,\partial_i\,(\sqrt{g}\,\mathcal{B}^i/h_i) \qquad (13.2.7)$$

With the replacements

$$\mathbf{B} \rightarrow \mathbf{E} \qquad \mathcal{B}^n \rightarrow E_n \qquad \hat{\mathbf{e}}_n \rightarrow \mathbf{a}_n \qquad h_i \rightarrow \sqrt{g_{ii}} \qquad T \rightarrow \Upsilon \qquad (13.2.8)$$

equations (13.2.6) and (13.2.7) agree with Moon and Spencer p 3 (1.11).

13.3 The Vector Laplacian in Cartesian coordinates

First we verify that our general orthogonal coordinates form (13.2.5) reduces to the expected result for the case that u-space is the same Cartesian space as x-space. One then has,

$$\begin{aligned} &h_i = 1 \quad g = 1 \qquad \mathbf{u = F(x) = x} \qquad \Rightarrow\ R = S = 1 \\ &(\hat{\mathbf{e}}_n)^i = S_n{}^i = \delta_n{}^i \qquad \Rightarrow\ \hat{\mathbf{e}}_n = \hat{\mathbf{n}}, \\ &\mathbf{B} = \mathcal{B}^n\ \hat{\mathbf{e}}_n = B^n\,\hat{\mathbf{n}} \qquad \Rightarrow\ \mathcal{B}^n = B^n = B_n \quad . \qquad (13.3.1) \end{aligned}$$

Then (13.2.5) becomes

$$\begin{aligned} &\text{✡}\mathbf{B} = [\ \partial_n T - \varepsilon^{nab}\partial_a\,\Gamma_b\ ]\ \hat{\mathbf{n}} \qquad \text{// implied sum on n} \\ &\text{where}\ \ T = \partial_i B^i \\ &\qquad\quad \Gamma_b = \varepsilon^{bcd}\,(\partial_c B^d) \qquad (13.3.2) \end{aligned}$$

or

$$\begin{aligned} \text{✡}\mathbf{B} &= [\ \partial_n\{\partial_i B^i\} - \varepsilon^{nab}\partial_a\ \{\ \varepsilon^{bcd}\ (\partial_c B^d)\}\ ]\ \hat{\mathbf{n}} \\ &= [\ \partial_n\{\text{div }\mathbf{B}\} - \varepsilon^{nab}\partial_a\ \{\ (\text{curl }\mathbf{B})_b\ \}\ ]\ \hat{\mathbf{n}} \\ &= [\ \partial_n\{\text{div }\mathbf{B}\} - [\text{curl }(\text{curl }\mathbf{B})]_n\ \}\ ]\ \hat{\mathbf{n}} \\ &= \ \text{grad}(\text{div }\mathbf{B}) - \ \text{curl }(\text{curl }\mathbf{B}) \end{aligned} \tag{13.3.3}$$

which agrees with our starting definition of ✡$\mathbf{B}$ (13.1.4).

<u>Second</u>, we verify the claim of (13.1.3) that in Cartesian coordinates

$$[\text{✡}\mathbf{B}]_n = \nabla^2 B_n\ . \tag{13.1.3}$$

We start with the first line of (13.3.3),

$$\begin{aligned} [\text{✡}\mathbf{B}]_n &= \ \partial_n\partial_i B^i - \ \varepsilon^{nab}\ \varepsilon^{bcd}\ \partial_a\partial_c B^d \\ &= \partial_n\partial_i B^i - \ (\delta_{nc}\delta_{ad} - \delta_{nd}\delta_{ac})\ \partial_a\partial_c B^d \qquad // \text{ from (D.10.22)} \\ &= \partial_n\partial_i B^i - \ \partial_n\partial_d B^d\ +\ \partial_c\partial_c B^n \\ &= \partial^2{}_c B^n\ = \nabla^2(B^n) = \nabla^2(B_n)\ . \end{aligned} \tag{13.3.4}$$

<u>Comment</u>: This is a proof of the Cartesian-coordinates vector identity $[\nabla(\nabla\bullet B) - \nabla x(\nabla x\mathbf{B})]_n = \nabla^2(B_n)$.

Thus it has been shown that, in Cartesian coordinates,

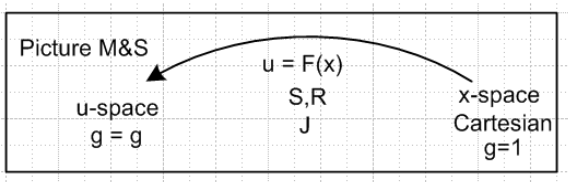
$$[\text{✡}\mathbf{B}]_n\ = \nabla^2\ (B_n)\ =\ [\ \text{grad}(\text{div }\mathbf{B}) - \ \text{curl }(\text{curl }\mathbf{B})\ ]_n\ . \tag{13.3.5}$$

It is this second and rather complicated form which allowed us to obtain the expression (13.2.5) for ✡$\mathbf{B}$ expressed in general curvilinear coordinates.

14. Summary of Differential Operators in curvilinear coordinates

14.1 Summary of Conventions and How To

The results are given in the Picture M&S context, and are copied from Chapters 9-13.



(14.1.1)

The Standard Notation of Chapter 7 is used throughout the tables.

In all the differential operator equations below, an operator acts either on a tensorial vector field **B** or on a tensorial scalar field f. On the right side of the drawing above, objects are said to be in x-space, and f(**x**) and $B_n(\mathbf{x})$ (components of **B**) are x-space tensorial objects. On the left side of the drawing objects are said to be in u-space. The function f is represented in u-space as $f(\mathbf{u})$, while there are three different ways to represent the components of **B**, called $B^n(\mathbf{u})$, $\mathcal{B}^n(\mathbf{u})$ and $B_n(\mathbf{u})$. There is a big distinction between the x-space objects and the u-space objects. For the scalar, $f(\mathbf{u}) = f(\mathbf{x}) = f(\mathbf{x}(\mathbf{u}))$ and so *f* has a different functional form than f. For the vector components, the u-space components are linear combinations of the x-space components, for example $B^n = R^n{}_m B^m$ (or $\boldsymbol{B} = R\mathbf{B}$, contravariant vector transformation).

See (9.3.6) concerning the font used for $\mathcal{B}^n(\mathbf{u})$.

On lines marked "Cartesian", $\partial_n = \partial/\partial x^n$ and f(**x**) and $B_n(\mathbf{x})$ appear (Cartesian components).

On other lines, $\partial_n = \partial/\partial u^n$, and the f and B objects appear in italics and are functions of **u**. The other functions like h_n, g_{ab} and g are also functions of **u**.

The vectors $\mathbf{e}_n$, $\hat{\mathbf{e}}_n$ and $\mathbf{e}^n$ all exist in Cartesian x-space. The $\mathbf{e}_n$ are the tangent base vectors of Chapter 3, and the $\mathbf{e}^n$ are the reciprocal base vectors of Chapter 6. The unit vectors $\hat{\mathbf{e}}_n \equiv \mathbf{e}_n/ |\mathbf{e}_n| = \mathbf{e}_n/h_n$ are used as well.

The dot product **A•B** is the covariant one of Section 5.10, namely, $A_i\overline{B}_i \rightarrow A_iB^i = g_{ij}A^iB^j$.

For each differential operator, the object on the LHS of the equations is always the same: it is a differential operator acting on f(**x**) or **B(x)** *in x-space*. In the Cartesian lines, the RHS expresses that LHS object in terms of Cartesian objects and Cartesian coordinates. On the other lines, the RHS expresses that exact same LHS x-space object *in terms of* Curvilinear (u-space) objects and coordinates. When the LHS is a scalar, the LHS object can be considered to be in either x-space or u-space. When the LHS is a vector, that LHS object is in x-space but can be related to u-space objects by a linear transformation by R.

How to Compute Things for some Arbitrary Coordinate System

For some arbitrary (and possibly non-orthogonal) coordinate system in some arbitrary number of dimensions N, one can start with the defining equations *such as* the following ones for spherical coordinates,

$x = r\sin\theta\cos\varphi$ that is: $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{u})$ $\mathbf{u} = (r,\theta,\varphi)$
$y = r\sin\theta\sin\varphi$
$z = r\cos\theta$.

In the Picture M&S context and in developmental notation, one uses (2.1.5) $S_{ik} \equiv (\partial x_i/\partial u_k)$ to compute the matrix S. Then the covariant metric tensor comes from (5.3.3), $\bar{g} = S^T G S$. Normally one has G = 1 for Cartesian x-space, but see (1.10) for quasi-Cartesian. The scale factors are then $h_i = \sqrt{\bar{g}_{ii}}$ and $g = \det(\bar{g})$, as shown in (5.11.7) and (5.12.20). The contravariant metric tensor is then found as $g = \bar{g}^{-1}$ (matrices). The conversion to Standard Notation is then $\bar{g} \rightarrow g_{ab}$ and $g \rightarrow g^{ab}$. These calculations are easily automated as shown for example in the Maple code of (G.6.1) and (G.6.2).

If the coordinate system is orthogonal, then $\bar{g}$ will be diagonal and one will have $\bar{g}_{nm} = h_n^2\ \delta_{n,m}$ and $g_{nm} = h_n^{-2}\ \delta_{n,m}$ and finally $g = \det(\bar{g}) = \Pi_i h_i$. For the eleven classical orthogonal systems the scale factors h_i are easily found online or in other reference sources, so one can then avoid the above steps (or verify them).

At this point, any formulae appearing in the tables below can be used.

The expressions marked below appear on pages 2 or 3 of Moon & Spencer (M&S).

general: $g \equiv \det(g_{ab})$ $h_n^2 \equiv g_{nn}$ $\partial^i = g^{ij}\partial_j$ $\mathcal{B}^n = h_n B^n$ $\mathbf{e}_n = h_n \hat{\mathbf{e}}_n$

orthogonal: $\sqrt{g} = (\Pi_i h_i) = h_1 h_2 ... h_N$ $g_{nm} = h_n^2\ \delta_{n,m}$ $g^{nm} = h_n^{-2}\ \delta_{n,m}$ (14.1.2)

14.2 divergence

divergence general: *(9.3.3)* (14.2.1)

$[\text{div }\mathbf{B}](\mathbf{x}) = [1/\sqrt{g}\,]\,\partial_n\,[\sqrt{g}\ B^n]$ $\mathbf{B} = B^n\mathbf{e}_n$

$[\text{div }\mathbf{B}](\mathbf{x}) = [1/\sqrt{g}\,]\,\partial_n\,[\sqrt{g}\ \mathcal{B}^n / h_n\,]$ $\mathbf{B} = \mathcal{B}^n\hat{\mathbf{e}}_n$ // M&S 1.06

$[\text{div }\mathbf{B}](\mathbf{x}) = [1/\sqrt{g}\,]\,\partial_n\,[\sqrt{g}\ g^{nm}\,B_m]$ $\mathbf{B} = B_n\mathbf{e}^n$

$[\text{div }\mathbf{B}](\mathbf{x}) = \partial_n B^n$ // B^n = Cartesian components of **B** $\mathbf{B} = B^n\ \hat{\mathbf{n}} = B_n\ \hat{\mathbf{n}}$

$[\text{div }\mathbf{B}](\mathbf{x}) = [\mathit{div}\ \boldsymbol{B}](\mathbf{u})$ // transformation (scalar)

divergence orthogonal: *(9.3.4)* (14.2.2)

$[\text{div }\mathbf{B}](\mathbf{x}) = [1/(\Pi_i h_i)]\,\partial_n\,[(\Pi_i h_i)\ B^n]$ $\mathbf{B} = B^n\mathbf{e}_n$

$[\text{div }\mathbf{B}](\mathbf{x}) = [1/(\Pi_i h_i)]\,\partial_n\,[(\Pi_i h_i)\ \mathcal{B}^n / h_n]$ $\mathbf{B} = \mathcal{B}^n\ \hat{\mathbf{e}}_n$

$[\text{div }\mathbf{B}](\mathbf{x}) = [1/(\Pi_i h_i)]\,\partial_n\,[(\Pi_i h_i)\ B_n / h_n^2]$ $\mathbf{B} = B_n\mathbf{e}^n$

divergence orthogonal N=3: *(9.3.5)* (14.2.3)

$[\text{div }\mathbf{B}](\mathbf{x}) = [1/(h_1h_2h_3)]\ \{\ \partial_1[h_2h_3\ \mathcal{B}^1(\mathbf{u})\ ]\ + \text{cyclic}\ \}$ $\mathbf{B} = \mathcal{B}^n\ \hat{\mathbf{e}}_n$

14.3 gradient and gradient dot vector

gradient general: *(10.1.16)* (14.3.1)

$[\text{grad f}](\mathbf{x}) = (\partial_i f)\ \mathbf{e}^i$

$[\text{grad f}](\mathbf{x}) = (\partial^i f)\ \mathbf{e}_i = g^{ij}\ (\partial_j f)\ \mathbf{e}_i$

$[\text{grad f}](\mathbf{x}) = h_i(\partial^i f)\ \hat{\mathbf{e}}_i = h_i\ g^{ij}\ (\partial_j f)\ \hat{\mathbf{e}}_i$

$[\text{grad f}](\mathbf{x}) = (\partial_i \text{f})\ \hat{\mathbf{n}}$ // Cartesian

$[\mathit{grad}\,f]_i(\mathbf{u}) = R_i{}^j[\text{grad f}]_j(\mathbf{x})$ // transformation (vector) $f(\mathbf{u}) = \text{f}(\mathbf{x}) = \text{f}(\mathbf{x}(\mathbf{u}))$

gradient orthogonal: *(10.1.17)* (14.3.2)

$[\text{grad f}](\mathbf{x}) = (\partial_i f)\ \mathbf{e}^i$

$[\text{grad f}](\mathbf{x}) = (\partial^i f)\ \mathbf{e}_i = (1/h_i^2)\ (\partial_i f)\ \mathbf{e}_i$

$[\text{grad f}](\mathbf{x}) = h_i(\partial^i f)\ \hat{\mathbf{e}}_i = (1/h_i)\ (\partial_i f)\ \hat{\mathbf{e}}_i$ // M&S 1.05

gradient dotted with a vector: *(10.2.5)* (14.3.3)

$[\text{grad } f](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) = (\partial^n f)\, B_n$ $\mathbf{B} = B_n \mathbf{e}^n$
$[\text{grad } f](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) = (\partial_n f)\, B^n$ $\mathbf{B} = B^n \mathbf{e}_n$
$[\text{grad } f](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) = (\partial_n f)\, \mathcal{B}^n / h_n$ $\mathbf{B} = \mathcal{B}^n\, \hat{\mathbf{e}}_n$
$[\text{grad } f](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) = (\partial_n f)\, B^n$ // Cartesian $\mathbf{B} = B^n\, \hat{\mathbf{n}}$
$[\text{grad } f](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) = [grad\, f](\mathbf{u}) \bullet \boldsymbol{B}(\mathbf{u})$ // transformation (scalar)

14.4 Laplacian

Laplacian general: *(11.10)* (14.4.1)

$[\text{lap } f](\mathbf{x}) = [1/\sqrt{g}\,]\ \partial_m[\sqrt{g}\ g^{nm}\, (\partial_n f)\,]$
$[\text{lap } f](\mathbf{x}) = \partial_n{}^2 f(\mathbf{x})$ // Cartesian $f(\mathbf{u}) = f(\mathbf{x}) = f(\mathbf{x}(\mathbf{u}))$
$[\text{lap } f](\mathbf{x}) = [lap\, f](\mathbf{u})$ // transformation (scalar)

Laplacian orthogonal: *(11.12)* (14.4.2)

$[\text{lap } f](\mathbf{x}) = [1/(\Pi_i h_i)]\ \partial_m[\, (\Pi_i h_i)\, (1/h_m{}^2)\, (\partial_m f)\,]$ // orthogonal // M&S 1.09

Laplacian orthogonal N=3: *(11.13)* (14.4.3)

$[\text{lap } f](\mathbf{x}) = 1/(h_1 h_2 h_3)\ \{\ \partial_1\, [\, (h_2 h_3/h_1)\, \partial_1 f\ ]\ \ + \text{cyclic}\ \}$

14.5 curl

curl general: *(12.4.7-11)* // N=3 only (14.5.1)

$$[\mathrm{curl}\ \mathbf{B}](\mathbf{x}) = \varepsilon^{nab}\,[(1/\sqrt{g}\,)\,\partial_a B_b\,]\,\mathbf{e}_n \qquad \mathbf{B} = B_n\mathbf{e}^n$$

$$[\mathrm{curl}\ \mathbf{B}](\mathbf{x}) = \varepsilon^{nab}\,[(1/\sqrt{g}\,)\,\partial_a(g_{bc}B^c\,)]\,\mathbf{e}_n \qquad \mathbf{B} = B^n\mathbf{e}_n$$

$$[\mathrm{curl}\ \mathbf{B}](\mathbf{x}) = \varepsilon^{nab}[(1/\sqrt{g}\,)\,h_n\ \partial_a(g_{bc}\,\mathcal{B}^c/h_c\,)]\,\hat{\mathbf{e}}_n \qquad \mathbf{B} = \mathcal{B}^n\hat{\mathbf{e}}_n$$

$$[\mathrm{curl}\ \mathbf{B}](\mathbf{x}) = \varepsilon^{nab}\,\partial_a \mathrm{B}_b(\mathbf{x})\ \ \hat{\mathbf{n}} \qquad \text{// Cartesian} \qquad \mathbf{B} = \mathrm{B}^n\ \hat{\mathbf{n}}$$

$$[\mathrm{curl}\ \mathbf{B}](\mathbf{x}) = (1/\sqrt{g}\,)\begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ B_1 & B_2 & B_3 \end{vmatrix} \qquad \mathbf{B} = B_n\mathbf{e}^n$$

$$[\mathrm{curl}\ \mathbf{B}](\mathbf{x}) = (1/\sqrt{g}\,)\begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ g_{1c}B^c & g_{2c}B^c & g_{3c}B^c \end{vmatrix} \qquad \mathbf{B} = B^n\mathbf{e}_n$$

$$[\mathrm{curl}\ \mathbf{B}](\mathbf{x}) = (1/\sqrt{g}\,)\begin{vmatrix} h_1\ \hat{\mathbf{e}}_1 & h_2\ \hat{\mathbf{e}}_2 & h_3\ \hat{\mathbf{e}}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ (g_{1c}/h_c)\,\mathcal{B}^c & (g_{2c}/h_c)\,\mathcal{B}^c & (g_{3c}/h_c)\,\mathcal{B}^c \end{vmatrix} \qquad \mathbf{B} = \mathcal{B}^n\hat{\mathbf{e}}_n \qquad \text{// M\&S 1.07}$$

$$[\mathrm{curl}\ \mathbf{B}](\mathbf{x}) = \begin{vmatrix} \hat{\mathbf{1}} & \hat{\mathbf{2}} & \hat{\mathbf{3}} \\ \partial_1 & \partial_2 & \partial_3 \\ \mathrm{B}_1 & \mathrm{B}_2 & \mathrm{B}_3 \end{vmatrix} \qquad \text{// here } \partial_n = \partial/\partial x^n \text{ and } \mathrm{B}_i = \mathrm{B}_i(\mathbf{x}) \qquad \mathbf{B} = \mathrm{B}^n\ \hat{\mathbf{n}}$$

curl orthogonal: *(12.5.2-5)* (14.5.2)

$$[\mathrm{curl}\ \mathbf{B}](\mathbf{x}) = \varepsilon^{nab}\,[(h_1h_2h_3)^{-1}\,\partial_a B_b\,]\,\mathbf{e}_n \qquad \mathbf{B} = B_n\mathbf{e}^n$$

$$[\mathrm{curl}\ \mathbf{B}](\mathbf{x}) = \varepsilon^{nab}\,[(h_1h_2h_3)^{-1}\,\partial_a(h_b^{\,2}B^b\,)]\,\mathbf{e}_n \qquad \mathbf{B} = B^n\mathbf{e}_n$$

$$[\mathrm{curl}\ \mathbf{B}](\mathbf{x}) = \varepsilon^{nab}[(h_1h_2h_3)^{-1}\,h_n\ \partial_a(h_b\,\mathcal{B}^b)]\,\hat{\mathbf{e}}_n \qquad \mathbf{B} = \mathcal{B}^n\hat{\mathbf{e}}_n$$

$$[\mathrm{curl}\ \mathbf{B}](\mathbf{x}) = (h_1h_2h_3)^{-1}\begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ B_1 & B_2 & B_3 \end{vmatrix} \qquad \mathbf{B} = B_n\mathbf{e}^n$$

$$[\mathrm{curl}\ \mathbf{B}](\mathbf{x}) = (h_1h_2h_3)^{-1}\begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ h_1^{\,2}B^1 & h_2^{\,2}B^2 & h_3^{\,2}B^3 \end{vmatrix} \qquad \mathbf{B} = B^n\mathbf{e}_n$$

$$[\mathrm{curl}\ \mathbf{B}](\mathbf{x}) = (h_1h_2h_3)^{-1}\begin{vmatrix} h_1\ \hat{\mathbf{e}}_1 & h_2\ \hat{\mathbf{e}}_2 & h_3\ \hat{\mathbf{e}}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ h_1\,\mathcal{B}^1 & h_2\,\mathcal{B}^2 & h_3\,\mathcal{B}^3 \end{vmatrix} \qquad \mathbf{B} = \mathcal{B}^n\hat{\mathbf{e}}_n \qquad \text{// M\&S 1.07a}$$

14.6 vector Laplacian

vector Laplacian general: *(13.1.22)* // N=3 only (14.6.1)

$$[\text{✡}\,\mathbf{B}](\mathbf{x}) = \mathbf{G} - \mathbf{V}$$

$$\mathbf{G} = \partial^n\{\, (1/\sqrt{g}\,)\, \partial_i\, (\sqrt{g}\;\; B^i)\}\;\; \mathbf{e_n} \qquad\qquad \mathbf{B} = B^n\mathbf{e_n}$$

$$\mathbf{V} = [\ (1/\sqrt{g}\,)\ \varepsilon^{ncd}\, \varepsilon^{eab}\, \partial_c\ \{\ \ (1/\sqrt{g}\,)\ g_{de}(\partial_a[g_{bf}B^f])\ \}]\ \mathbf{e_n}$$

$$\mathbf{G} = \ h_n\ \partial^n\{\, (1/\sqrt{g}\,)\, \partial_i\, (\sqrt{g}\, \mathcal{B}^{\,i}/h_i)\}\ \ \hat{\mathbf{e}}_\mathbf{n} \qquad\qquad \mathbf{B} = \ \mathcal{B}^n\ \ \hat{\mathbf{e}}_\mathbf{n}$$

$$\mathbf{V} = h_n\ [\ (1/\sqrt{g}\,)\ \varepsilon^{ncd}\, \varepsilon^{eab}\, \partial_c\ \{\ \ (1/\sqrt{g}\,)\ g_{de}(\partial_a[g_{bf}\mathcal{B}^f/h_f])\ \}]\ \hat{\mathbf{e}}_\mathbf{n}$$

$$\text{✡}\,\mathbf{B}\ =\ [\nabla^2 B^n\ ]\ \hat{\mathbf{n}} \qquad \text{// Cartesian, } \hat{\mathbf{n}} = \mathbf{u_n} \qquad\qquad \mathbf{B} = B^n\ \hat{\mathbf{n}}$$

vector Laplacian orthogonal: *(13.2.6-7)* (14.6.2)

$$\begin{aligned}
[\text{✡}\,\mathbf{B}](\mathbf{x}) \quad &= [(1/h_1)\ \partial_1 T + (h_1/\sqrt{g}\,)\ (\partial_3\ \Gamma_2 - \partial_2\ \Gamma_3)\ ]\ \hat{\mathbf{e}}_\mathbf{1} \qquad\qquad \text{// M\&S 1.11}\\
&+ [(1/h_2)\ \partial_2 T + (h_2/\sqrt{g}\,)\ (\partial_1\ \Gamma_3 - \partial_3\ \Gamma_1)\ ]\ \hat{\mathbf{e}}_\mathbf{2}\\
&+ [(1/h_3)\ \partial_3 T + (h_3/\sqrt{g}\,)\ (\partial_2\ \Gamma_1 - \partial_1\ \Gamma_2)\ ]\ \hat{\mathbf{e}}_\mathbf{3}
\end{aligned}$$

$$\begin{aligned}
T &= \ (1/\sqrt{g}\,)\ \partial_i\ (\sqrt{g}\, \mathcal{B}^i/h_i)\\
\Gamma_1 &= \ \ (1/\sqrt{g}\,)\ h_1^{\,2}\ (\ \partial_2\ [h_3\, \mathcal{B}^3]\ -\ \partial_3\ [h_2\, \mathcal{B}^2]\ )\\
\Gamma_2 &= \ \ (1/\sqrt{g}\,)\ h_2^{\,2}\ (\ \partial_3\ [h_1\, \mathcal{B}^1]\ -\ \partial_1\ [h_3\, \mathcal{B}^3]\ )\\
\Gamma_3 &= \ \ (1/\sqrt{g}\,)\ h_3^{\,2}\ (\ \partial_1\ [h_2\, \mathcal{B}^2]\ -\ \partial_2\ [h_1\, \mathcal{B}^1]\ )
\end{aligned}$$

or *(13.2.5)* (14.6.3)

$$[\text{✡}\,\mathbf{B}](\mathbf{x}) \ = [\ \ (1/h_n)\ \partial_n T\ -\ (h_n/\sqrt{g}\,)\ \varepsilon^{nab}\partial_a\ \Gamma_b\ ]\ \hat{\mathbf{e}}_\mathbf{n}$$

$$\begin{aligned}
T &= \ (1/\sqrt{g}\,)\ \partial_i\ (\sqrt{g}\, \mathcal{B}^i/h_i)\\
\Gamma_b &= \ \ (1/\sqrt{g}\,)\ h_b^{\,2}\ \varepsilon^{bcd}\ \ (\ \partial_c\ [h_d\, \mathcal{B}^d])
\end{aligned}$$

14.7 Example 1: Polar coordinates: a practical curvilinear notation

From our many visits to this example we know that :

$$\text{general:}\qquad g \equiv \det(g_{ab}) \qquad h_n{}^2 \equiv g_{nn} \qquad \partial^i = g^{ij}\partial_j \qquad \mathcal{B}^n = h_n B^n \qquad \mathbf{e}_n = h_n\hat{\mathbf{e}}_n$$

$$\text{orthogonal:}\qquad \sqrt{g} = (\Pi_i h_i) = h_1 h_2 ... h_N \qquad g_{nm} = h_n{}^2\ \delta_{n,m} \qquad g^{nm} = h_n{}^{-2}\ \delta_{n,m} \qquad \textit{(14.1.2)}$$

$$\mathbf{e}_1 = r(-\sin\theta,\cos\theta) \quad = \mathbf{e}_\theta \quad = r\,\hat{\mathbf{e}}_\theta \qquad // = r\,\hat{\boldsymbol{\theta}}$$
$$\mathbf{e}_2 = (\cos\theta,\sin\theta) \quad = \mathbf{e}_r \quad = \hat{\mathbf{e}}_r \qquad // = \hat{\mathbf{r}} \qquad \textit{(3.4.2)} \qquad (14.7.1)$$

$$g_{ij} = \begin{array}{cc} \begin{array}{cc} \theta & r \end{array} & \\ \begin{pmatrix} r^2 & 0 \\ 0 & 1 \end{pmatrix} & \begin{array}{c} \theta \\ r \end{array} \end{array} \qquad R = \begin{array}{cc} \begin{array}{cc} x & y \end{array} & \\ \begin{pmatrix} -\sin\theta/r & \cos\theta/r \\ \cos(\theta) & \sin\theta \end{pmatrix} & \begin{array}{c} \theta \\ r \end{array} \end{array} \qquad u_1 = \theta \qquad u_2 = r \quad (!!) \qquad (14.7.2)$$
$$\textit{(5.13.11)} \qquad\qquad \textit{(3.4.1)}$$

$$h_1 = h_\theta = \sqrt{g_{\theta\theta}} = r \qquad h_2 = h_r = \sqrt{g_{rr}} = 1\ . \qquad \textit{(5.13.11)} \qquad (14.7.3)$$

Assume one is working with a 2D vector velocity field **v(x)** -- our first encounter with a "lower case" vector field which we have been careful to support with the general notations above. Since one knows the names of the variables 1=θ and 2=r, one might define the following new variables on the first line to be the officially named variables on the second line

$v^\theta \quad v^r$	$v_\theta \quad v_r$	$\mathrm{v}_\theta \quad \mathrm{v}_r$	$\mathrm{v}_x \qquad \mathrm{v}_y$
$v^1 \quad v^2$	$v_1 \quad v_2$	$\boldsymbol{\upsilon}^1 \quad \boldsymbol{\upsilon}^2$	$\mathrm{v}_1 = \mathrm{v}^1 \qquad \mathrm{v}_2 = \mathrm{v}^2$
contravariant	covariant	unit vector	Cartesian
(italic)	(italic)	(non-italic)	(non-italic)

(14.7.4)

One ends up with the comfortable $\mathbf{v} = \mathrm{v}_\theta\hat{\boldsymbol{\theta}} + \mathrm{v}_r\hat{\mathbf{r}}$ notation as shown below.

$$\mathrm{v}_\theta \equiv \boldsymbol{\upsilon}^1 = h_1 v^1 = h_\theta\, v^\theta = r\, v^\theta \qquad \text{// unit vector projection components}$$
$$\mathrm{v}_r \equiv \boldsymbol{\upsilon}^2 = h_2 v^2 = h_r\, v^r = v^r$$

$$v^\theta = R^\theta{}_x \mathrm{v}_x + R^\theta{}_y \mathrm{v}_y = -\sin\theta/r\ \mathrm{v}_x + \cos\theta/r\ \mathrm{v}_y$$
$$v^r = R^r{}_x \mathrm{v}_x + R^r{}_y \mathrm{v}_y = \cos\theta\ \mathrm{v}_x + \sin\theta\ \mathrm{v}_y \qquad (14.7.5)$$

so

$$\mathrm{v}_\theta = r\, v^\theta = -\sin\theta\ \mathrm{v}_x + \cos\theta\ \mathrm{v}_y$$
$$\mathrm{v}_r = v^r = \cos\theta\ \mathrm{v}_x + \sin\theta\ \mathrm{v}_y$$

$$\mathbf{v} = v^n \mathbf{e}_n = v^\theta \mathbf{e}_\theta + v^r \mathbf{e}_r$$
$$\mathbf{v} = \boldsymbol{\upsilon}^n \hat{\mathbf{e}}_n = \mathrm{v}_\theta \hat{\mathbf{e}}_\theta + \mathrm{v}_r \hat{\mathbf{e}}_r = \mathrm{v}_\theta \hat{\boldsymbol{\theta}} + \mathrm{v}_r \hat{\mathbf{r}}$$
$$\mathbf{v} = \mathrm{v}_n \hat{\mathbf{n}} = \mathrm{v}_x \hat{\mathbf{x}} + \mathrm{v}_y \hat{\mathbf{y}}\ . \qquad (14.7.6)$$

As an example of a differential operator, consider the divergence for orthogonal coordinates from (14.2.2)

$$[\mathrm{div}\ \mathbf{B}](\mathbf{x}) = [1/(\Pi_i h_i)]\ \partial_n\ [(\Pi_i h_i)\ \mathscr{B}^n / h_n]\ . \qquad\qquad \mathbf{B} = \mathscr{B}^n\ \hat{\mathbf{e}}_n \tag{14.7.7}$$

Applied to the present situation one gets ($h_1 = h_\theta = r$ and $h_2 = h_r = 1$),

$$\begin{aligned} [\mathrm{div}\ \mathbf{v}](\mathbf{x}) &= [1/(h_1 h_2)]\ \{\ \partial_1\ [h_2\ \mathscr{v}^1] \quad + \partial_2[h_1\ \mathscr{v}^2]\ \} \qquad\qquad \mathbf{V} = \mathscr{v}^n\ \hat{\mathbf{e}}_n \\ &= [1/(h_\theta h_r)]\ \{\ \partial_\theta\ [h_r\ v_\theta\ ] + \partial_r[h_\theta\ v_r]\ \} \\ &= (1/r)\ \{\ \partial_\theta v_\theta + \partial_r(r v_r)\ \}\ . \end{aligned} \tag{14.7.8}$$

Suppose v_x and v_y are constants. Then the Cartesian expression says

$$[\mathrm{div}\ \mathbf{v}](\mathbf{x}) = \partial_n v^n = \partial_x v_x + \partial_y v_y = 0 + 0 = 0\ . \tag{14.7.9}$$

Equation (14.7.9) agrees :

$$\begin{aligned} [\mathrm{div}\ \mathbf{v}](\mathbf{x}) &= (1/r)\ \{\ \partial_\theta v_\theta + \partial_r(r v_r)\ \} \\ &= (1/r)\ \{\ \partial_\theta[-\sin\theta\ v_x + \cos\theta\ v_y] + \partial_r(r\ [\cos\theta\ v_x + \sin\theta\ v_y])\ \} \\ &= (1/r)\ \{\quad [-\cos\theta\ v_x - \sin\theta\ v_y] + [\cos\theta\ v_x + \sin\theta\ v_y]\ \} \\ &= 0\ . \end{aligned} \tag{14.7.10}$$

The notation illustrated here works for any curvilinear coordinates.

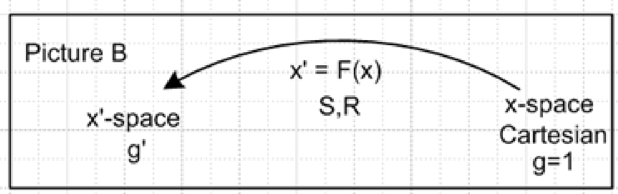
15. Covariant derivation of all curvilinear differential operator expressions

15.1 Review of Chapters 9 through 13

Let us briefly review the discussion of Chapters 9 through 13 concerning the expression of differential operators

div, grad, lap, curl and ✡

in curvilinear coordinates. The underlying framework was provided by Picture B,



(15.1.1)

where the coordinates x'^n of x'-space were the curvilinear coordinates of interest, while the coordinates of x-space were the Cartesian coordinates. The transformation **x'** = **F**(**x**) provided the connection between the curvilinear coordinates and the Cartesian ones.

In **Chapter 9** the object div **B** was treated as the total **B** "flux" ($\int$ **B**•d**A**) emitting from an N-piped in x-space divided by the volume of that N-piped, in the limit the N-piped shrank to a point. Using Appendix B formulas for the N-piped area, and using the key expansion

$\mathbf{B} = B'^n \mathbf{e}_n$, // (7.13.10) line 3

where the $\mathbf{e}_n$ are tangent base vectors in x-space and the B'^n are components of **B** in x'-space, we obtained in (9.1.18) a way to write the Cartesian-space div **B** in terms of curvilinear **x'** coordinates and x'-space objects, namely, $B'^n(\mathbf{x}')$, $g'(\mathbf{x}')$ and the gradient operator $\partial'_i = \partial/\partial x'^i$.

In **Chapter 10** the vector **grad f** was expanded using (7.13.10) line 4 that $\mathbf{V} = V'_n \mathbf{e}^n$ to get

$\mathbf{grad\ f} = [\mathrm{grad\ f}]'_n \mathbf{e}^n = (\partial'_n f')\mathbf{e}^n$

where $\mathbf{e}^n$ are the reciprocal base vectors. This result was expressed in various ways, and **gradf •B** was also treated such that **gradf •B** $= (\partial'_n f')\mathbf{e}^n \bullet B'^m \mathbf{e}_m = (\partial'_n f')\, B'^n$. [Recall that $f'(\mathbf{x}') = f(\mathbf{x})$.]

In **Chapter 11** the Laplacian was written as lap f = div (grad f) and then the results of Chapters 9 and 10 for div and grad were used to obtain (11.5) which expresses lap f in terms of x'-space coordinates and objects $B'^n(\mathbf{x}')$, $g'(\mathbf{x}')$, $g'^{nm}(\mathbf{x}')$ and ∂'_i . In traditional notation, one writes $\nabla^2 f = \nabla \bullet [\ \nabla f\ ] = \mathrm{div\ (grad\ f)}$.

In **Chapter 12** the x'-space curl component $C'^n = [\mathrm{curl}\ \mathbf{B}]'^n$ was treated as the circulation line integral of **B** around near face n of the same N-piped used in Chapter 9 for divergence, divided by the area of that face, again in the limit the N-piped shrank to a point. With the expansions $\mathbf{B} = B'_n \mathbf{e}^n$ and $\mathbf{C} = J^{-1}C'^n \mathbf{e}_n$, this led to an expression (12.3.3) for curl **B** in terms of the curvilinear **x**' coordinates and the x'-space objects $B'_n(\mathbf{x}')$, $g'(\mathbf{x}')$, $g'^{nm}(\mathbf{x}')'$ ε'^{nab} and ∂'_i .

In **Chapter 13** the vector Laplacian was treated using the definition

✡ **B** ≡ grad(div **B**) – curl (curl **B**)

and then the results of Chapters 9, 10 and 12 on div, grad and curl were recruited to produce an expression (13.1.17) for ✡ **B** in terms of x'-space coordinates and various x'-space objects.

15.2 The Covariant Method

We shall now repeat all of the above work using "the covariant method" which, as will be shown, gets most results extremely quickly, but at a cost of requiring knowledge of tensor densities and covariant differentiation, as described in Appendices D and F.

Imagine that one has a Cartesian x-space expression for some tensor object of interest Q,

$$Q\text{----} \ = \ [\text{ Cartesian form with various up and down indices and various derivatives }] \qquad (15.2.1)$$

where Q---- means that Q has some number of up and down indices. The following four steps should then be carried out:

1. Arrange any summed index pairs within [....] so they are "tilted". Then when [...] is "tensorized" those tilted pairs will be true contractions.

Note that doing this does not affect the Cartesian value of this object [...] since up and down indices are the same in Cartesian space.

2. Replace all derivatives with covariant derivatives. Thus is done by first writing all derivatives in comma notation, and then by replacing those commas with semicolons. For example:

$$\partial_\alpha B_a \ = \ B_{a,\alpha} \rightarrow \ B_{a;\alpha} \ . \qquad // \ B_{a;\alpha} = \partial_\alpha B_a - \Gamma^n{}_{a\alpha} B_n \quad \textit{(F.9.2)} \qquad (15.2.2)$$

Objects like $\partial_\alpha B_a$ which are not tensors become objects like $B_{a;\alpha}$ which *are* tensors according to the theorem of (F.7.1).

Note again that doing this does not affect the Cartesian value of this object [...], because in Cartesian x-space $\Gamma^n{}_{a\alpha} = 0$ as per (F.4.16). For a more general case like $B_{abc..x;\alpha}$ shown in (F.7.2), making the replacement $B_{abc..x,\alpha}$ makes no difference as long as $B_{abc..x}$ is a true tensor (W = 0), because *all* those Γ correction terms seen in (F.7.2) vanish in Cartesian space. The same is true regardless of index positions.

3. Insert appropriate powers of g as needed to achieve appropriate tensor density weights so that all terms in [...] have the same weight and this weight equals the weight of Q---- . Tensor densities are discussed in Appendix D, and this weight adjustment idea appears in (D.2.3c).

Note again that doing this does not affect the Cartesian value of this object [...] because g = 1 in Cartesian space ($g = \det(g_{ij})$).

Comment: Since x-space is Cartesian, we know from (5.12.16) that g > 0 and g' > 0, so we don't need to use expressions like |g| and |g'| in our current context.

4. At this point one has

$$Q\text{----} = \{ \text{ tensorized form of the Cartesian expression of interest } \} \quad . \tag{15.2.3}$$

Since this is then a "true tensor equation" (it could be a true tensor density equation), according to the rule of covariance described in Section 7.15 the above equation can be converted to x'-space by simply priming Q and all objects within { }. One then has

$$Q'\text{----} = \{ \text{ tensorized form of the Cartesian expression of interest } \}' \tag{15.2.4}$$

where the notation {...}' means that everything inside {...} is primed.

Perhaps at this point one does some simplifications of the resulting {...}' object. The result then *is* the expression of Q---- in general curvilinear coordinates if the underlying transformation **x**' = **F**(**x**) describes the transformation from Cartesian to curvilinear coordinates. For a vector, recall (7.13.10) that $\mathbf{Q} = Q'^n \mathbf{e_n}$ so the "curvilinear coordinates" expression of **Q** involves the the x'-space components $Q'^a(\mathbf{x}')$ but the tangent base vectors $\mathbf{e_n}$ are in x-space. A similar statement can be made for Q being any tensor, as shown for example in (E.2.9) which says $Q = \Sigma_{ijk} Q'^{ijk} (\mathbf{e_i} \otimes \mathbf{e_j} \otimes \mathbf{e_k})$.

These four steps then comprise "the covariant method". The method of course applies to tensor analysis applications other than "curvilinear coordinates".

Uniqueness. One might obtain two tensorizations of a Cartesian expression that look different. Both these tensorizations must lead to the same Q'---- in x'-space. For example, suppose one finds two tensorizations of a vector Q^a , call them $\{...\}_1{}^a$ and $\{...\}_2{}^a$ which, when evaluated in Cartesian x-space, are equal

$$\{...\}_1{}^a = \{...\}_2{}^a \qquad \text{when both sides are evaluated in Cartesian x-space} \quad . \tag{15.2.5}$$

One can then transform both objects to x'-space in the usual manner,

$$\{...\}'_1{}^a = R^a{}_b \{...\}_1{}^b \qquad \text{and} \qquad \{...\}'_2{}^a = R^a{}_b \{...\}_2{}^b \; . \tag{15.2.6}$$

Therefore, since $\{...\}_1{}^a = \{...\}_2{}^a$ in x-space, one must have $\{...\}'_1{}^a = \{...\}'_2{}^a$ in x'-space.

Example: In (15.7.6) below we will establish the following tensorization of $(✡B)^n$ (the vector Laplacian),

$$(✡B)^n = \{.....\}_1{}^n = (B^j{}_{;j})^{;n} - g^{-1/2} \varepsilon^{nab} (g^{-1/2} g_{bc} \varepsilon^{cde} B_{e;d})_{;a} \quad . \tag{15.2.7}$$

Since in Cartesian x-space one can write $(\text{✡}B)^{n} = \nabla^2(B^n) = \partial_j\partial^j B^n = B^{n,j}{}_{,j}$, another tensorization of $(\text{✡}B)^{n}$ is given by

$$(\text{✡}B)^n = \{.....\}_2{}^n = B^{n;j}{}_{;j} \quad . \tag{15.2.8}$$

Therefore, the following must be true in x'-space

$$B'^{n;j}{}_{;j} = (B'^{j}{}_{;j})^{;n} - g'^{-1/2}\varepsilon'^{nab}(g'^{-1/2}g'_{bc}\varepsilon'^{cde}B'_{e;d})_{;a} \tag{15.2.9}$$

Explicitly verifying (15.2.9) takes some work and we do it below in Section 15.8. Both forms are used in Appendix I to compute **✡B** in spherical and cylindrical coordinates.

In the following Sections we shall in short order derive the curvilinear expressions for the basic differential operators using the covariant method outlined above. This method is used as well (along with parallel brute force methods) to obtain expressions for objects **∇v**, **divT**,∇T in Appendices G, H, J.

15.3 divergence (Chapter 9)

In Cartesian coordinates, [div **B**] = $\partial_i B^i = B^i{}_{,i}$. The true tensorial tensor which matches this is [div **B**] ≡ $B^i{}_{;i}$ since $\Gamma = 0$ for Cartesian coordinates by (F.4.16). That is to say, from (F.9.3),

$$B^a{}_{;\alpha} = \partial_\alpha B^a + \Gamma^a{}_{\alpha n} B^n \tag*{\textit{(F.9.3)}}$$

so

$$B^i{}_{;i} = \partial_i B^i + \Gamma^i{}_{in} B^n = \partial_i B^i = B^i{}_{,i} \qquad\qquad \text{since} \quad \Gamma^c{}_{ab} \equiv 0 \ .$$

Since $B^i{}_{;i}$ is a scalar, it is the same in x-space as in x'-space (Section 7.15 covariance),

$$[\text{div } \mathbf{B}] = [\text{div } \mathbf{B}]' = B'^i{}_{;i} = \partial'_i B'^i + \Gamma'^i{}_{in} B'^n \tag{15.3.1}$$

where B'^n are the contravariant components of vector **B** in x'-space. But recall (F.4.2) in x'-space,

$$\Gamma'^i{}_{in} = (1/\sqrt{g'}\,)\, \partial'_n(\sqrt{g'}\,) \ . \tag*{\textit{(F.4.2)}}$$

Therefore

$$[\text{div } \mathbf{B}] = \partial'_n B'^n + (1/\sqrt{g'}\,)\,\partial'_n(\sqrt{g'}\,)B'^n = (1/\sqrt{g'}\,)\,\partial'_n\,(\sqrt{g'}\,B'^n). \tag{15.3.2}$$

This matches the result (9.2.1) obtained in Chapter 9 by geometric methods.

15.4 gradient and gradient dot vector (Chapter 10)

If f is a scalar, then **G =** grad f transforms as an ordinary vector according to (2.4.2) and (2.5.1), so it is *already* a tensor. According to line 4 of (7.13.10) the expansion of such a vector may be written

$$\mathbf{G} = [\text{grad f}] = G'_i \mathbf{e}^i = (\partial'_i f')\, \mathbf{e}^i \ . \tag{15.4.1}$$

Also, grad f • **B** is a scalar, again already a tensor. Using this fact and line 3 of (7.13.10) that **B** = $B'^n \mathbf{e}_n$,

$$[\text{grad f}]' \bullet \mathbf{B}' = [\text{grad f}] \bullet \mathbf{B} = [(\partial'_i f')\, \mathbf{e}^i] \bullet (B'^n \mathbf{e}_n) = (\partial'_i f')B'^n \delta^i{}_n = (\partial'_i f')B'^i \tag{15.4.2}$$

These two results match (10.1.7) and the second line of (10.2.4) of Chapter 10. As a reminder, since f is a scalar field, f '(**x**') = f(**x**).

15.5 Laplacian (Chapter 11)

In Cartesian x-space the Laplacian of a scalar function f can be written as

$$[\text{lap f}]_{\text{Cart}} = \partial^i \partial_i f = f^{,i}{}_{,i} \ . \tag{15.5.1}$$

The tensorized version of the Laplacian is then the following, using the ,→; rule,

$$\text{lap f} = f^{;i}{}_{;i} \ . \tag{15.5.2}$$

To verify that this is so (we already know it is so), we call upon (F.9.19) with scalar B → f,

$$f^{;a}{}_{;\alpha} = \partial_\alpha f^{;a} + \Gamma^a{}_{\alpha n} f^{;n} \qquad . \qquad\qquad (F.9.19) \tag{15.5.3}$$

Then (F.4.16) says Γ = 0 so $f^{;a}{}_{;\alpha} = \partial_\alpha f^{;a} = \partial_\alpha f^{,a} = f^{,a}{}_{,\alpha}$ so finally $f^{;i}{}_{;i} = f^{,i}{}_{,i}$.

Since (15.5.3) is a true tensor equation, Section 7.15 covariance says it takes this form in x'-space,

$$f'^{;a}{}_{;\alpha} = \partial'_\alpha f'^{;a} + \Gamma'^a{}_{\alpha n} f'^{;n} \ . \tag{15.5.4}$$

But lap f = $f^{;i}{}_{;i}$ is a scalar, so we have

$$[\text{lap f}] = [\text{lap f}]' = f'^{;i}{}_{;i} = \partial'_i f'^{;i} + \Gamma'^i{}_{in} f'^{;n} = \partial'_i f'^{,i} + \Gamma'^i{}_{in} f'^{,n}$$

$$= \partial'_i \partial'^i f' + \Gamma'^i{}_{in} (\partial'^n f') \ . \tag{15.5.5}$$

But identity (F.4.2) says $\Gamma'^i{}_{in} = (1/\sqrt{g'}\,)\, \partial'_n(\sqrt{g'}\,)$ so then

$$\text{lap f} = \partial'_i \partial'^i f' + (1/\sqrt{g'}\,)\, \partial'_n(\sqrt{g'}\,)\, (\partial'^n f')$$

$$= [1/\sqrt{g'}\,]\ \partial'_n [\sqrt{g'}\ (\partial'^n f')] \ . \tag{15.5.6}$$

This matches (11.5) found in Chapter 11.

15.6 curl (Chapter 12)

In Cartesian space we know that **C** = curl **B** and $C^i = \varepsilon^{ijk}\partial_j B_k = \varepsilon^{ijk} B_{k,j}$. The obvious tensorized form is $C^i = \varepsilon^{ijk} B_{k;j}$ which we could verify using (F.9.2) with $\Gamma = 0$, but we know this is correct from the general theory (Step 2) presented in Section 15.2. Since this is a true tensor equation, Section 7.15 covariance tells us that it takes this form in x'-space,

$$C'^i = \varepsilon'^{ijk} B'_{k;j} \quad . \tag{15.6.1}$$

Recall from (D.4.8) that $\varepsilon'^{ijk} = \varepsilon^{ijk}$ = the permutation tensor.

As shown in Section D.8, **C** ≡ curl **B** is a vector density of weight -1 because it contains the ε tensor which has that weight. According to (E.2.26) the expansion of such a vector density is given by $\mathbf{C} = J^{-1}C'^i\mathbf{e}_i$. Therefore,

$$\text{curl } \mathbf{B} = \mathbf{C} = J^{-1}C'^i \ \mathbf{e}_i = J^{-1} [\varepsilon'^{ijk}B'_{k;j}]\mathbf{e}_i = (1/\sqrt{g'}\,)\, \varepsilon'^{ijk} \ B'_{k;j}\mathbf{e}_i \tag{15.6.2}$$

But we now do call upon (F.9.2) in x'-space to show that $\varepsilon'^{ijk} \ B'_{k;j} = \varepsilon'^{ijk} \ B'_{k,j}$:

$$B'_{a;\alpha} = \partial'_\alpha B'_a - \Gamma'^n{}_{a\alpha}B'_n \qquad \text{covariant rank-2 tensor} \quad \text{// 2nd term is sym on } a\leftrightarrow\alpha \qquad \textit{(F.9.2)}$$

so

$$B'_{k;j} = \partial'_j B'_k - \Gamma'^n{}_{kj}B'_n$$

so

$$\varepsilon'^{ijk} \ B'_{k;j} = \varepsilon'^{ijk} \ [\partial'_j B'_k - \Gamma'^n{}_{kj}B'_n] = \varepsilon'^{ijk} \ \partial'_j B'_k - \varepsilon'^{ijk}\Gamma'^n{}_{kj}B'_n$$

$$= \varepsilon'^{ijk} \ \partial'_j B'_k - 0 \qquad \text{// because } \varepsilon' \text{ is antisymmetric on jk, but } \Gamma'^n{}_{kj} \text{ is symmetric on jk } ^\dagger$$

$$= \varepsilon'^{ijk} \ B'_{k,j} \ . \tag{15.6.3}$$

Inserting this result into (15.6.2) gives

$$\text{curl } \mathbf{B} = (1/\sqrt{g'}\,)\, \varepsilon'^{ijk} \ B'_{k,j}\mathbf{e}_i = (1/\sqrt{g'}\,)\, \varepsilon'^{ijk}(\partial'_jB'_k)\mathbf{e}_i \ .$$

This matches the result (12.3.3) found in Chapter 12 based on geometric circulation integrals.

$$^\dagger \ Q \equiv A_{ij}S_{ij} = \underset{i\leftrightarrow j}{A_{ji}S_{ji}} = \underset{\text{symmetries}}{(-A_{ij})(+S_{ij})} = - A_{ij}S_{ij} = -Q \quad \Rightarrow \ Q = 0 \tag{15.6.4}$$

15.7 vector Laplacian (Chapter 13)

The vector Laplacian operator takes a bit more work. Since this is a major example of carrying out the "tensorization process" of Section 15.2, much detail is provided.

In Chapter 13 the vector Laplacian is written ✡ (instead of ∇^2) and is defined by

$$\text{✡}\mathbf{B} \ \equiv\ \text{grad}(\text{div}\ \mathbf{B}) - \text{curl}\ (\text{curl}\ \mathbf{B}) \ . \qquad\qquad \mathit{(13.1.4)} \qquad\qquad (15.7.1)$$

Recall from (13.3.4) that that in Cartesian space,

$$[\text{grad}(\text{div}\ \mathbf{B}) - \text{curl}\ (\text{curl}\ \mathbf{B})]^n \ =\ \nabla^2(B^n) \ =\ (\text{✡}B)^n \qquad // \text{ Cartesian x-space} \qquad (15.7.2)$$

so we are allowed to use the LHS here to find the curvilinear expression of ✡**B .** In components,

$$(\text{✡}B)^n = \partial^n(\partial_j B^j) - \varepsilon^{nab}\partial_a\ [\text{curl } B]_b$$

$$= \partial^n(\partial_j B^j) - \varepsilon^{nab}\partial_a\ (\varepsilon_{bde}\partial^d B^e) \ . \qquad (15.7.3)$$

In Cartesian space up and down index position does not matter, see (5.9.1), and we are just jockeying the indices in search of a tensorized form with contracted indices where possible. Writing (15.7.3) in comma notation,

$$(\text{✡}B)^n = (B^j{}_{,j})^{,n} - \varepsilon^{nab}(\varepsilon_{bde}B^{e,d})_{,a} \qquad (15.7.4)$$

one is then free to replace commas with semicolons since $\Gamma = 0$ in Cartesian coordinates, see (F.4.16),

$$(\text{✡}B)^n = (B^j{}_{;j})^{;n} - \varepsilon^{nab}(\varepsilon_{bde}B^{e;d})_{;a} \ . \qquad (15.7.5)$$

There is a "technical difficulty" visible here. The first term $(B^j{}_{;j})^{;n}$ is a true vector since it is the contraction of a true rank-3 tensor on two tilted indices, but the second term is a rank-3 tensor density of weight -2. In the second term, in order to neutralize the weight of each ε tensor density (from (D.4.9) each ε has weight -1), we shall now add a benign factor of $g^{-1/2}$, as suggested in (D.2.3c), since $g^{-1/2}$ is a scalar density of weight +1. This in turn is so from (5.12.13) that $g' = J^2 g$ which says g has a weight of -2. The factor $g^{-1/2}$ is benign since g = 1 in Cartesian coordinates. So, with this upgrade installed,

$$(\text{✡}B)^n = (B^j{}_{;j})^{;n} - g^{-1/2}\varepsilon^{nab}(g^{-1/2}\varepsilon_{bde}B^{e;d})_{;a} \ . \qquad (15.7.6)$$

Now both objects $T_b \equiv (g^{-1/2}\varepsilon_{bde}B^{e;d})$ and $V^n \equiv g^{-1/2}\varepsilon^{nab}\ T_{b;a}$ are true vectors, so the above equation says that a true vector is the difference between two other true vectors.

Anticipating a few steps ahead when our ship lands in x'-space, we know that $\varepsilon^{bde}B_{e;d} = \varepsilon^{bde}B_{e,d}$ due to the e↔d symmetry of the Γ term in $B_{e;d} = B_{e,d} - \Gamma^s{}_{ed}B_s$ of (F.9.24), and this *motivates us* to perform a "tilt change" in this product. Such a tilt-change also gets the ;d index "down" which is a nicer place for it to be since it will soon become ∂_d. The tilt operation is done as follows ,

$$\varepsilon_{bde}B^{e;d} = \varepsilon_b{}^d{}_e B^e{}_{;d} = \varepsilon_b{}^{de}B_{e;d} = g_{bc}\ \varepsilon^{cde}B_{e;d} \ . \qquad (15.7.7)$$

Recall that tilt changes like this are only allowed with true tensor indices, (7.11.3), and that is why the ε tilt reversal can be done only after the semicolon in $B_{e;d}$ is installed. Eq. (15.7.6) then reads,

$$(\text{✡}B)^n = (B^j{}_{;j})^{;n} - g^{-1/2}\varepsilon^{nab}(g^{-1/2}g_{bc}\varepsilon^{cde}B_{e;d})_{;a} \quad . \qquad (15.7.8)$$

Although $\varepsilon_b{}^{de}$ is a fine tensor, we really prefer ε^{cde} since *this* is the permutation tensor, and that is why the g_{bc} factor is added in (15.7.7).

At this point every index is a tensor index, so (15.7.8) is a "true tensor equation" in the sense of Section 7.15. The above equation agrees exactly with the Cartesian expression of $(\text{✡}B)^n$ since $g = 1$, $g_{bc} = \delta_{b,c}$, and the semicolons are commas since $\Gamma = 0$, a fact verified in the Comment above. The point is that we have found a way to write the Cartesian expression of $(\text{✡}B)^n$ such that both terms are true tensors. The equation is therefore now "covariant" and the equation in x'-space can be obtained simply by priming all objects:

$$(\text{✡}B)'^n = (B'^j{}_{;j})^{;n} - g'^{-1/2}\varepsilon'^{nab}(g'^{-1/2}g'_{bc}\varepsilon'^{cde}B'_{e;d})_{;a} \quad . \qquad (15.7.9)$$

The semicolons were installed to obtain a true tensor equation, but now that we have successfully arrived in x'-space, we want to remove as many semicolons as possible because they imply "extra terms" since now $\Gamma' \neq 0$. Consider for example, this grouping which appears in the above,

$$\varepsilon'^{nab}\,(g'^{-1/2}g'_{bc}\varepsilon'^{cde}B'_{e;d})_{;a} = \varepsilon'^{nab}\;T'_{b;a} \qquad T'_b \equiv (g'^{-1/2}g'_{bc}\varepsilon'^{cde}B'_{e;d}) \quad . \qquad (15.7.10)$$

Since $T_{b;a} = T_{b,a} - \Gamma^n{}_{ba}T_n$ from (F.9.24), $\varepsilon'^{nab}T'_{b;a} = \varepsilon'^{nab}\;T'_{b,a}$ because ε^{nab} is antisymmetric on a,b while $\Gamma^n{}_{ba}$ is symmetric, see (A.4.4). Thus,

$$(\text{✡}B)'^n = (B'^j{}_{;j})^{;n} - g'^{-1/2}\varepsilon'^{nab}(g'^{-1/2}g'_{bc}\varepsilon'^{cde}B'_{e;d})_{\underset{\uparrow}{,}a} \quad . \qquad (15.7.11)$$

But $\varepsilon'^{cde}B'_{e;d} = \varepsilon'^{cde}B'_{e,d}$ for the exact same symmetry reason, so

$$(\text{✡}B)'^n = (B'^j{}_{;j})^{;n} - g'^{-1/2}\varepsilon'^{nab}(g'^{-1/2}g'_{bc}\varepsilon'^{cde}B'_{e\underset{\uparrow}{,}d})_{,a} \quad . \qquad (15.7.12)$$

Meanwhile, (F.9.1) says $D^{;n} = D^{,n}$ so that $(B'^j{}_{;j})^{;n} = (B'^j{}_{;j})^{,n}$, since the object $B'^j{}_{;j} = B^{;j}{}_{;j}$ is a scalar like D. Therefore,

$$(\text{✡}B)'^n = (B'^j{}_{;j})^{\underset{\uparrow}{,}n} - g'^{-1/2}\varepsilon'^{nab}(g'^{-1/2}g'_{bc}\varepsilon'^{cde}B'_{e,d})_{,a} \quad . \qquad (15.7.13)$$

Since ε'^{cde} is a constant (the permutation tensor, see below (D.4.1)), we *now* pull it through the ∂'_a derivative implied by ,a and then showing all derivatives the above becomes

$$(\text{✡}B)'^n = \partial'^n(B'^j{}_{;j}) - g'^{-1/2}\varepsilon'^{nab}\varepsilon'^{cde}\;\partial'_a(g'^{-1/2}g'_{bc}\partial'_dB'_e) \quad . \qquad (15.7.14)$$

From (15.3.2) and (F.9.1) that $B^{;j} = B^{,j}$ we know that

$$[\text{div } \mathbf{B}] = B^{;j}{}_{;j} = B^{,j}{}_{;j} = (1/\sqrt{g'}\,)\,\partial'_j\,(\sqrt{g'}\,B'^j) = (1/\sqrt{g'}\,)\,\partial'_i\,(\sqrt{g'}\,B'^i) \quad . \qquad (15.7.15)$$

Using this in the first term on right of (15.7.14) gives,

$$(\text{✡}B)'^{n} = \partial'^{n}\{(1/\sqrt{g'}\,)\,\partial'_{i}(\sqrt{g'}\,B'^{i})\} \;-\; g'^{-1/2}\varepsilon'^{nab}\varepsilon'^{cde}\;\partial'_{a}(g'^{-1/2}g'_{bc}\partial'_{d}B'_{e}) \quad . \tag{15.7.16}$$

To compare this result to that of Chapter 13, we shuffle the indices in the second term as follows,

$$\begin{array}{ll} g'^{-1/2}\varepsilon'^{nab}\varepsilon'^{cde}\;\partial'_{a}(g'^{-1/2}g'_{bc}\partial'_{d}B'_{e}) & \text{// now take a,b,c,d,e} \rightarrow \text{A,B,C,D,E} \\ g'^{-1/2}\varepsilon'^{nAB}\varepsilon'^{CDE}\;\partial'_{A}(g'^{-1/2}g'_{BC}\partial'_{D}B'_{E}) & \text{// now take A}\rightarrow\text{c, B}\rightarrow\text{d, C}\rightarrow\text{e, D}\rightarrow\text{a, E}\rightarrow\text{b} \\ g'^{-1/2}\varepsilon'^{ncd}\varepsilon'^{eab}\;\partial'_{c}(g'^{-1/2}g'_{de}\partial'_{a}B'_{b})\;. & \end{array} \tag{15.7.17}$$

Then (15.7.16) reads,

$$(\text{✡}B)'^{n} = \partial'^{n}\{(1/\sqrt{g'}\,)\,\partial'_{i}(\sqrt{g'}\,B'^{i})\} - g'^{-1/2}\varepsilon'^{ncd}\varepsilon'^{eab}\;\partial'_{c}(g'^{-1/2}g'_{de}\partial'_{a}B'_{b}) \quad . \tag{15.7.18}$$

This may now be compared with (13.1.17) which we quote :

$$\begin{array}{ll} \text{✡}\,\mathbf{B} \;=\mathbf{G}-\mathbf{V} & \\ \mathbf{G} = \partial'^{n}\{\,(1/\sqrt{g'}\,)\,\partial'_{i}\,(\sqrt{g'}\;\,B'^{i})\}\;\;\mathbf{e_n} & \mathbf{B} = \;B'^{n}\mathbf{e_n} \\ \mathbf{V} = [\;(1/\sqrt{g'}\,)\,\varepsilon'^{ncd}\,\varepsilon'^{eab}\,\partial'_{c}\,\{\;\;(1/\sqrt{g'}\,)\,g'_{de}(\partial'_{a}B'_{b})\,\}]\;\mathbf{e_n} & \mathbf{B} = \;B'_{n}\mathbf{e}^{n} \end{array} \tag{15.7.19}$$

The results agree, and further processing of this result is done in Chapter 13.

This concludes our discussion of how the curvilinear-coordinate differential operator forms of Chapters 9,10,11,12,and 13 may be obtained relatively quickly using the notions of tensor densities (App. D) and covariant derivatives (App. F), without resorting to the use of N-piped geometric constructs.

15.8 Verification that two tensorizations are the same

In Section 15.7 we obtained the x-space **✡B** "tensorization" (15.7.8) which written in x'-space says

$$(\text{✡}B)'^{n} = B'^{j}{}_{;j}{}^{;n} - \;g'^{-1/2}\varepsilon'^{nab}(g'^{-1/2}g'_{bc}\varepsilon'^{cde}B'_{e;d})_{;a} \quad . \qquad \textit{(15.7.9)} \tag{15.8.1}$$

Since in Cartesian coordinates $(\text{✡}B)^{n} \;= \nabla^2(B^n) = \partial_j\partial^j\, B^n = B^{n,j}{}_{,j}$, an obvious *alternate* tensorization is $B^{n,j}{}_{,j} \rightarrow B^{n;j}{}_{;j}$ so that in x'-space

$$(\text{✡}B)'^{n} = B'^{n;j}{}_{;j}\;. \tag{15.8.2}$$

In Section 15.2 we showed that tensorization is unique, so these two tensorizations must be the same. We know they are the same, but in the time-honored tradition of "trust but verify", we would like to actually verify their equality. As many readers know, verifying things that are supposedly obvious often leads to the discovery of problems.

The claim that the above two tensorizations are the same is not general in the abstract. It is specific in that it applies in the "tensor world" built upon a transformation **x**' = **F**(**x**) where x-space is Cartesian. The two forms are not equal if x-space has some metric tensor other than g = 1.

Our task then is to show that the following is true when x-space is Cartesian,

$$B'^{n;j}{}_{;j} = B'^{j}{}_{;j}{}^{;n} - g'^{-1/2}\varepsilon'^{nab}(g'^{-1/2}g'_{bc}\varepsilon'^{cde}B'_{e;d})_{;a} \qquad ? \qquad (15.8.3)$$

The plan is to rewrite the above equation in a sequence of reducing steps until we get to an equation that obviously *is* true. Then one then can reverse the sequence to conclude that (15.8.3) is true. We shall mark each step with a question mark to the right, to indicate that this is "something we want to show is true". When the final step is reached which is true, the reader then goes backwards and erases the question marks.

As a first and rather large step in terms of clutter reduction, we claim that,

$$(g'^{-1/2}g'_{bc}\varepsilon'^{cde}B'_{e;d})_{;a} = g'^{-1/2}g'_{bc}\varepsilon'^{cde}\ (B'_{e;d})_{;a} \ . \qquad (15.8.4)$$

This is an application of several facts derived in Appendix F. The main idea is Theorem (F.10.15) concerning what objects are "extractable" from a covariant derivative expression. Eq. (F.10.17) shows that ε'^{cde} is extractable since $\varepsilon'^{cde}{}_{\alpha;} = 0$, while (F.10.18) shows that g'_{bc} is extractable since $g'_{bc;\alpha}= 0$. Finally, (F.10.20) says that, for any real s, g'^{s} is extractable since $(g'^{s})_{;\alpha} = 0$, recalling that $g' > 0$. We then arrive at

$$B'^{n;j}{}_{;j} = B'^{j}{}_{;j}{}^{;n} - g'^{-1/2}\varepsilon'^{nab}(g'^{-1/2}g'_{bc}\varepsilon'^{cde}B'_{e;d})_{;a}$$

$$= B'^{j}{}_{;j}{}^{;n} - g'^{-1/2}\varepsilon'^{nab}g'^{-1/2}g'_{bc}\varepsilon'^{cde}\ (B'_{e;d})_{;a} \quad \text{// now use } g'_{bc} \text{ and note } \varepsilon'^{nab} = \varepsilon'^{ban}$$

$$= B'^{j}{}_{;j}{}^{;n} - g'^{-1}\varepsilon'^{ban}\ \varepsilon'_{b}{}^{de}\ B'_{e;d;a} \ .$$

Therefore

$$B'^{n;j}{}_{;j} = B'^{j}{}_{;j}{}^{;n} - g'^{-1}\varepsilon'_{b}{}^{de}\varepsilon'^{ban}\ B'_{e;d;a} \qquad ? \qquad (15.8.5)$$

We would next like to have all the B' tensors be multiples of $B'_{e;d;a}$ which appears in the last term, so :

$$B'^{n;j}{}_{;j} = B'^{n;a}{}_{;a} = g'^{ne}\,g'^{ad}B'_{e;d;a}$$

$$B'^{j}{}_{;j}{}^{;n} = B'^{d}{}_{;d}{}^{;n} = g'^{de}\,g'^{na}\,B'_{e;d;a} \ . \qquad (15.8.6)$$

Installing these expressions into (15.8.5) gives

$$g'^{ne}\,g'^{ad}B'_{e;d;a} = g'^{de}\,g'^{na}\,B'_{e;d;a} - g'^{-1}\,\varepsilon'_{b}{}^{de}\varepsilon'^{bna}\ B'_{e;d;a} \qquad ?$$

or

$$[\ g'^{ne}\,g'^{ad} - g'^{de}\,g'^{na} + g'^{-1}\ \varepsilon'_{b}{}^{de}\ \varepsilon'^{bna}\ ]\,B'_{e;d;a} = 0 \qquad ? \qquad (15.8.7)$$

At this point, one might hope that [....] = 0 and we are done. If this were true, our claim (15.8.3) would be true for *any* metric tensor g, but we know that is not the case, so we do not expect to have [...] = 0. The next step is to call upon (D.11.11) expressed in x'-space which says

$$g'^{-1}\ \varepsilon'_{s}{}^{AB}\varepsilon'^{sA'B'} = g'^{AA'}g'^{BB'} - g'^{BA'}g'^{AB'} \tag{D.11.11}$$

or

$$g'^{-1}\ \varepsilon'_{b}{}^{de}\varepsilon'^{bna} = g'^{dn}g'^{ea} - g'^{en}\,g'^{da}\quad . \tag{15.8.8}$$

Putting this into (15.8.7) gives

$$[\ g'^{ne}\,g'^{ad} - g'^{de}\,g'^{na} + g'^{dn}g'^{ea} - g'^{en}g'^{da}\ ]\, B'_{e;d;a} = 0 \qquad ?$$

or

$$[\ -\ g'^{de}\,g'^{na} + g'^{dn}g'^{ea}\ ]\, B'_{e;d;a} = 0 \qquad ?$$

or

$$[\ g'^{dn}g'^{ea} - g'^{de}g'^{na}\ ]\, B'_{e;d;a} = 0 \qquad ? \tag{15.8.9}$$

As expected, we find [...] ≠ 0. We have to use the fact that x-space is Cartesian. To this end write the covariant transformation rule for rank-3 tensor $B_{e;d;a}$ as shown for example in (7.10.1),

$$B'_{e;d;a} = R_e{}^E R_d{}^D R_a{}^A\, B_{E;D;A} = R_e{}^E R_d{}^D R_a{}^A\, B_{E,D,A} = R_e{}^E R_d{}^D R_a{}^A\, \partial_D\partial_A\, B_E \tag{15.8.10}$$

where the critical Cartesian x-space fact will be that $\partial_D\partial_A\, B_A$ is symmetric under A↔D , Inserting the above transformation rule into (15.8.9) gives,

$$[\ g'^{dn}g'^{ea} - g'^{de}g'^{na}\ ]\, R_e{}^E R_d{}^D R_a{}^A \partial_D\partial_A\, B_E = 0 \qquad ?$$

or

$$g'^{dn}g'^{ea}\, R_e{}^E R_d{}^D R_a{}^A\, \partial_D\partial_A\, B_E = g'^{de}g'^{na}\, R_e{}^E R_d{}^D R_a{}^A\, \partial_D\partial_A\, B_E \qquad ? \tag{15.8.11}$$

At this point we recall from (7.5.9) that g'^{**} raises the first index of $R_*{}^*$, whereas the second index of $R_*{}^*$ goes up and down for free since g = 1 in x-space, so we will take second indices all down. Thus,

$$R^{aE}R^{nD}R_a{}^A\partial_D\partial_A\, B_E = R^{dE}R_d{}^D R^{nA}\, \partial_D\partial_A\, B_E \qquad ?$$

or

$$R^a{}_E R^n{}_D R_a{}^A\partial_D\partial_A\, B_E = R^d{}_E R_d{}^D R^n{}_A\partial_D\partial_A\, B_E \qquad ? \tag{15.8.12}$$

Now on the right side do D↔A on these dummy indices, but then restore the $\partial_D\partial_A$ order so

$$R^a{}_E R^n{}_D R_a{}^A\partial_D\partial_A\, B_E = R^d{}_E R_d{}^A R^n{}_D\partial_D\partial_A\, B_E \qquad ? \tag{15.8.13}$$

and changing d→a on the right

$$R^a{}_E R_a{}^A R^n{}_D\ \partial_D\partial_A\, B_E = R^a{}_E R_a{}^A R^n{}_D\partial_D\partial_A\, B_E \qquad ? \tag{15.8.14}$$

The two sides are exactly the same, so we are done and thus (15.8.3) is verified.

Appendix A: Reciprocal Base Vectors the Hard Way

A.1 Introduction

Note: This Appendix is written in the developmental notation, not the Standard Notation, though a few equations are translated to the latter form. The rules for translation to Standard Notation are

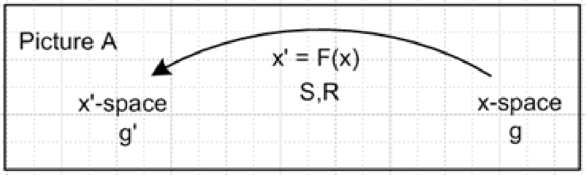
(7.13.1)	*(7.5.2)*	*(7.5.4)*	*(7.4.1)*	*(7.4.1)*
$\mathbf{E}_n \rightarrow \mathbf{e}^n$	$R_{ij} \rightarrow R^i{}_j$	$S_{ij} \rightarrow S^i{}_j$	$\bar{g}'_{nm} \rightarrow g'_{nm}$	$g'_{nm} \rightarrow g'^{nm}$.

The tangent base vectors $\mathbf{e}_n$ are defined in Chapter 3, while the reciprocal base vectors $\mathbf{E}_n$ are defined in Chapter 6. Here are some facts regarding these vectors gathered from those chapters:

(3.2.5) and (6.1.4)	*(6.2.4)*	*(6.2.4) and (6.1.2)*	*(3.2.7) and (6.6.3) and (6.1.2)*
$(\mathbf{e}_n)_k \ = \ S_{kn}$	$\mathbf{e}_n \bullet \mathbf{e}_m \ = \ \bar{g}'_{nm}$	$\lvert\mathbf{e}_n\rvert = \sqrt{\bar{g}'_{nn}} \ = h'_n$	$S = \ [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \ .... \ \mathbf{e}_N \,]$
$(\mathbf{E}_n)_i \ \equiv g_{ia}\, R_{na}$	$\mathbf{E}_n \bullet \mathbf{E}_m \ = g'_{nm}$	$\lvert\mathbf{E}_n\rvert = \sqrt{g'_{nn}}$	$R = \ [\bar{\mathbf{E}}_1, \bar{\mathbf{E}}_2, \bar{\mathbf{E}}_3 \ .... \ \bar{\mathbf{E}}_N \,]^T$
$= g'_{na}\, S_{ia}$	$\mathbf{e}_n \bullet \mathbf{E}_m \ = \delta_{n,m}$	$\mathbf{E}_n \equiv g'_{ni}\, \mathbf{e}_i$	$\mathbf{e}_n \ = \ \bar{g}'_{ni}\, \mathbf{E}_i$.

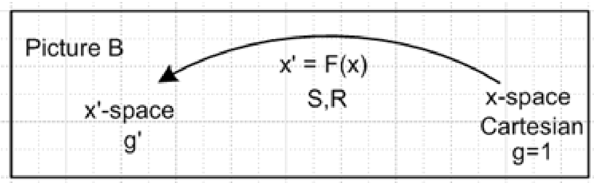
(A.1.1)

These facts are all applicable in the Picture A context with arbitrary metric tensors g' and g,



(A.1.2)

This Appendix begins with a different definition of something called $\mathbf{E}_k$. Although the definition is meaningful in the general Picture A context, the object so defined only agrees with the $\mathbf{E}_k$ of Chapter 6 if x-space is Cartesian (g = 1). The reason can be traced to the fact that the dot product rule $\mathbf{e}_n \bullet \mathbf{E}_m = \delta_{n,m}$ is only valid for the Appendix A definition of $\mathbf{E}_m$ when g = 1 because only then is a cross product orthogonal to all its component vectors. One application of the reciprocal base vectors is in the study of curvilinear coordinates where one always takes g = 1, and g' is then the curvilinear coordinates metric tensor of interest. Therefore, the reader should think of this Appendix in the context of Picture B



(A.1.3)

A.2 Definition of E_n

The reciprocal base vectors are defined in the following very strange looking and clumsy manner,

$$(\mathbf{E_k})_\alpha \equiv \det(R)\,(-1)^{k-1}\,\varepsilon_{\alpha i_1 i_2 i_3 \ldots \cancel{i_k} \ldots i_N}\,(\mathbf{e}_1)_{i_1}\,(\mathbf{e}_2)_{i_2}\ \ldots\ldots\ \cancel{(\mathbf{e_k})_{i_k}} \ldots\ldots\ldots\ (\mathbf{e_N})_{i_N} \qquad \text{(A.2.1)}$$

where N is the number of dimensions of the Cartesian x-space R^N in which the vectors $\mathbf{e_n}$ and $\mathbf{E_n}$ exist. Notice that the ε subscript i_k is "crossed out" and the same for factor $(\mathbf{e_k})_{i_k}$. Crossed out means they are simply missing, they are omitted. Thus, in the above expression there are N-1 implied summation indices (α is fixed) and there are N-1 factors of the form $(\mathbf{e_n})_{i_n}$.

The object ε has N subscripts and is the "totally antisymmetric tensor" in N dimensions: $\varepsilon_{123\ldots N} \equiv +1$, and each time any two indices on ε are swapped, ε negates. For example, $\varepsilon_{1234} = 1$ but $\varepsilon_{1432} = -1$. If two indices are the same, then $\varepsilon = 0$.

A.3 Simpler notation

To avoid dealing with subscripts on subscripts, one can rewrite the above definition in a less precise but simpler notation

$$(\mathbf{E_k})_\alpha \equiv \det(R)\,(-1)^{k-1}\varepsilon_{\alpha abc\ldots x}\,(\mathbf{e}_1)_a(\mathbf{e}_2)_b \ \ldots\ldots\ (\mathbf{e_N})_x \qquad // \ \kappa(k) \text{ and } (\mathbf{e_k})_\kappa \text{ are missing} \qquad \text{(A.3.1)}$$

In this notation, subscript x stands for the N^{th} letter of the alphabet (imagine $N \le 26$). If κ is the k^{th} letter of the alphabet, then κ is missing from the indices on ε, and the factor $(\mathbf{e_k})_\kappa$ is missing from the product of factors. For example, if k = 2, then summation index κ = b is missing from the ε.

Now take the ε subscript α and slide it right to the "hole" where κ is missing, picking up a minus sign for each step of this slide. Moving k-1 positions results in $(-1)^{k-1}$. Thus the above becomes,

$$(\mathbf{E_k})_\alpha \equiv \det(R)\,\varepsilon_{abc..\alpha..x}\,(\mathbf{e}_1)_a(\mathbf{e}_2)_b \ \ldots\ldots\ (\mathbf{e_N})_x \qquad // \ (\mathbf{e_k})_\kappa \text{ is missing, } \alpha \text{ in } \kappa \text{ position (the } k^{th}) \qquad \text{(A.3.2)}$$

<u>Example</u>: For N = 3 the above becomes,

$$\begin{aligned}
(\mathbf{E}_1)_\alpha &\equiv \det(R)\,\varepsilon_{\alpha bc}(\mathbf{e}_2)_b(\mathbf{e}_3)_c \ \Rightarrow\ \mathbf{E}_1 = \det(R)\,\mathbf{e}_2 \text{ x } \mathbf{e}_3 && \text{a is missing} \\
(\mathbf{E}_2)_\alpha &\equiv \det(R)\,\varepsilon_{a\alpha c}(\mathbf{e}_1)_a(\mathbf{e}_3)_c \ \Rightarrow\ \mathbf{E}_2 = \det(R)\,\mathbf{e}_3 \text{ x } \mathbf{e}_1 && \text{b is missing} \\
(\mathbf{E}_3)_\alpha &\equiv \det(R)\,\varepsilon_{ab\alpha}(\mathbf{e}_1)_a(\mathbf{e}_2)_b \ \Rightarrow\ \mathbf{E}_3 = \det(R)\,\mathbf{e}_1 \text{ x } \mathbf{e}_2 && \text{c is missing} \qquad \text{(A.3.3)}
\end{aligned}$$

and the results are cyclic. Here is a detail from the middle line

$$\varepsilon_{a\alpha c}(\mathbf{e}_1)_a(\mathbf{e}_3)_c = -\,\varepsilon_{\alpha ac}(\mathbf{e}_1)_a(\mathbf{e}_3)_c = +\,\varepsilon_{\alpha ca}(\mathbf{e}_1)_a(\mathbf{e}_3)_c = \varepsilon_{\alpha ca}(\mathbf{e}_3)_c\,(\mathbf{e}_1)_a = [\ \mathbf{e}_3 \text{ x } \mathbf{e}_1\ ]_\alpha \qquad \text{(A.3.4)}$$

A.4 Generalized Cross Product of N-1 vectors of dimension N

One can define a generalized "cross product" of N-1 vectors, each of dimension N, in this fashion:

$$Q_a \equiv \varepsilon_{abc\ldots x}\,B_bC_cD_d\ldots\ldots X_x \qquad \text{(A.4.1)}$$

where x and X represent the N^{th} letter of the alphabet. The ε object is again the totally antisymmetric tensor with N indices. In vector notation one writes this symbolically as

$$\mathbf{Q} = \mathbf{B} \times \mathbf{C} \times \mathbf{D} \times \ldots \times \mathbf{X} \qquad \text{/ N-1 factors, N-2 crosses .} \qquad (A.4.2)$$

This vector notation is *defined* by the previous line.

The vector **Q** is orthogonal to all the vectors from which it is constructed! For example (here is the point where $\mathbf{Q} \bullet \mathbf{C} \equiv g_{ab}Q_aC_b$ needs to be Q_aC_a, so g = 1 is required in x-space)

$$\mathbf{Q} \bullet \mathbf{C} = C_aQ_a = C_a\, \varepsilon_{abc...x}\, B_bC_cD_d.....X_x = B_bD_d...X_x \ \{ \varepsilon_{abc...x} \ C_aC_c \} \ . \qquad (A.4.3)$$

But {..} is the contraction of something symmetric under a↔c (C_aC_c) with something antisymmetric under a↔c ($\varepsilon_{\alpha abc...x}$) and therefore {..} = 0. In general,

$$\begin{aligned} S_{ac}A_{ac} &= S_{ca}\, A_{ca} && \text{// relabel both dummy summation indices} \\ &= S_{ac}\,(-A_{ac}) && \text{// S is Symmetric, A is antisymmetric} \\ &= -\, S_{ac}\, A_{ac} && \text{// = the negative of the starting expression} \\ &= 0 \ . && \qquad (A.4.4) \end{aligned}$$

Similarly, $\mathbf{Q}\bullet\mathbf{A} = 0$, $\mathbf{Q}\bullet\mathbf{B} = 0$ and so on.

Swapping the position of any two vectors in the generalized cross product causes **Q** to change sign. For example, swapping **B** and **C**,

$$\begin{aligned} Q_a \equiv \ & \varepsilon_{abc...x}\, C_bB_cD_d.....X_x = \varepsilon_{acb...x}\, C_cB_bD_d.....X_x && \text{// } b \leftrightarrow c \\ & = -\, \varepsilon_{abc...x}\, B_bC_cD_d.....X_x \quad = \ -Q_a \ . && \text{// swap indices on } \varepsilon \qquad (A.4.5) \end{aligned}$$

Thus, the notions of orthogonality and interchange are consistent with the regular $\mathbf{Q} = \mathbf{B} \times \mathbf{C}$ cross product for N=3.

When N=2, one must be a little careful with this notation. The component equation is

$$Q_a \equiv \varepsilon_{ab}\, B_b \qquad \Rightarrow \qquad Q_1 = B_2 \ \text{ and } Q_2 = -B_1 \ . \qquad (A.4.6)$$

One might be tempted to express the vector equation as $\mathbf{Q} = \mathbf{B}$ since there are no "no crosses". This vector equation is *wrong*, while the component equation is correct. One can rescue the vector notation by a simple trick. When N=2 the vector **B** can be represented of course as $\mathbf{B} = B_1\hat{\mathbf{1}} + B_2\ \hat{\mathbf{2}}$. Imagine this 2D space to be embedded in the usual 3D space with a third axis $\hat{\mathbf{3}}$. Then consider this 3D cross product:

$$\mathbf{Q} = \mathbf{B} \times \hat{\mathbf{3}} \qquad \Rightarrow \qquad Q_a \equiv \ \varepsilon_{abc}\, B_b(\hat{\mathbf{3}})_c = \varepsilon_{abc}\, B_b\delta_{3,c} = \varepsilon_{ab3}\, B_b \ = \varepsilon_{ab}B_b \quad . \qquad (A.4.7)$$

Thus, this trick reproduces the correct component equation, and it makes more obvious the fact that **Q** is orthogonal to **B**.

Summary: The generalized cross product **Q** of N-1 vectors each of dimension N can be expressed in both component and vector notation:

$$Q_a \equiv \varepsilon_{abc...x} B_b C_c D_d.....X_x$$
$$\mathbf{Q} = \mathbf{B} \times \mathbf{C} \times \mathbf{D} \times ... \times \mathbf{X} \qquad \text{/ N-1 factors, N-2 crosses} \qquad (A.4.8)$$

Q is orthogonal to all the vectors from which it is composed. Swapping any two vectors negates **Q**. When N=2, one can rescue the otherwise failing vector notation by thinking of it as saying $\mathbf{Q} = \mathbf{B} \times \hat{\mathbf{3}}$.

Comment: Notice that **Q** = **B** x **C** x **D** is defined for 4-vectors only. This is a completely different animal from the object **Q** = **B** x (**C** x **D**) which is defined for 3-vectors only. This latter object contains two ε factors, while the former only one.

A.5 Missing Man Formation

We now make a small variation in the notation. Start with the above equation,

$$Q_a \equiv \varepsilon_{abc...x} B_b C_c D_d.....X_x \ , \qquad (A.5.1)$$

then change a to α, back up all the Latin letters by one (but leave the last as "unknown" x), and assume that some subscript κ and factor K_κ are "missing". The result is,

$$Q_\alpha \equiv \varepsilon_{\alpha ac...x} A_a B_b C_c.....X_x \ . \qquad \text{// } \kappa \text{ and } K_\kappa \text{ are missing} \qquad (A.5.2)$$

There are still N-1 factors, and one can still write this in vector notation

$$\mathbf{Q} = \mathbf{A} \times \mathbf{B} \times \mathbf{C} \times ... \times \mathbf{X} \qquad \text{// } \mathbf{K} \text{ is missing} \qquad (A.5.3)$$

and of course it is still true that **Q•C** = 0, etc. For N=2 the vector notation is rescued as in (A.4.7) above.

A.6 Apply this Notation to E

Compare the above Q_α of (A.5.2) to the {...} part of the (A.3.1) definition of $(\mathbf{E_k})_\alpha$,

$$(\mathbf{E_k})_\alpha \equiv \det(R)\, (-1)^{k-1} \{ \varepsilon_{\alpha abc...x} (\mathbf{e_1})_a (\mathbf{e_2})_b ...... (\mathbf{e_N})_x \} \qquad \text{// } \kappa(k) \text{ and } (\mathbf{e_k})_\kappa \text{ are missing; } N \geq 2$$

(A.3.1) (A.6.1)

Therefore, the definition of $\mathbf{E_k}$ for N > 2 can be written in this vector notation,

$$\mathbf{E_k} \equiv \det(R)\, (-1)^{k-1}\, \mathbf{e_1} \times \mathbf{e_2} \times ...... \times \mathbf{e_N} \qquad \text{// } \mathbf{e_k} \text{ missing; } N > 2 \ . \qquad (A.6.2)$$

The reciprocal base vector $\mathbf{E_k}$ is thus orthogonal to all the tangent base vectors from which it is constructed (remember $\mathbf{e_k}$ is missing)! For example, for N=3 the three **E** vectors are given by

$$\mathbf{E_1} = \det(R)\, (-1)^{1-1}\, \mathbf{e_2} \times \mathbf{e_3} = \det(R)\ \mathbf{e_2} \times \mathbf{e_3}$$
$$\mathbf{E_2} = \det(R)\, (-1)^{2-1}\, \mathbf{e_1} \times \mathbf{e_3} = \det(R)\ \mathbf{e_3} \times \mathbf{e_1}$$
$$\mathbf{E_3} = \det(R)\, (-1)^{3-1}\, \mathbf{e_1} \times \mathbf{e_2} = \det(R)\ \mathbf{e_1} \times \mathbf{e_2} \qquad (A.6.3)$$

which agrees with the results (A.3.3). For N =2 (**E's** label corresponds to the missing **e**'s label),

$$\begin{aligned} \mathbf{E}_1 &= \mathrm{det}(R)\,(-1)^{1-1}\,\mathbf{e}_2 \text{ x } \hat{3} \;=\; \mathrm{det}(R)\,\mathbf{e}_2 \text{ x } \hat{3} &&\text{or}\qquad (\mathbf{E}_1)_k = \mathrm{det}(R)\,\varepsilon_{ka}(\mathbf{e}_2)_a \\ \mathbf{E}_2 &= \mathrm{det}(R)\,(-1)^{2-1}\,\mathbf{e}_1 \text{ x } \hat{3} \;=\; -\,\mathrm{det}(R)\,\mathbf{e}_1 \text{ x } \hat{3} &&\text{or}\qquad (\mathbf{E}_2)_k = -\mathrm{det}(R)\,\varepsilon_{ka}(\mathbf{e}_1)_a \qquad (\text{A.6.4}) \end{aligned}$$

One can combine these two lines into one as follows (eg, k = 1, then 3-1 = 2, etc)

$$\mathbf{E}_k = \mathrm{det}(R)\,(-1)^{k-1}\,\mathbf{e}_{3-k} \text{ x } \hat{3} \;=\; \mathrm{det}(R)\,\mathbf{e}_{3-k} \text{ x } \hat{3} \quad\text{or}\quad (\mathbf{E}_1)_k = \mathrm{det}(R)\,(-1)^{k-1}\varepsilon_{ka}(\mathbf{e}_{3-k})_a\ . \qquad (\text{A.6.5})$$

The vector "trick" notation shows that $\mathbf{E}_1 \bullet \mathbf{e}_2 = 0$ and $\mathbf{E}_2 \bullet \mathbf{e}_1 = 0$,

$$\begin{aligned} \mathbf{E}_1 \bullet \mathbf{e}_2 &= \mathrm{det}(R)\,\mathbf{e}_2 \text{ x } \hat{3} \bullet \mathbf{e}_2 = 0 \\ \mathbf{E}_2 \bullet \mathbf{e}_1 &= -\mathrm{det}(R)\,\mathbf{e}_1 \text{ x } \hat{3} \bullet \mathbf{e}_1 = 0 \qquad (\text{A.6.6}) \end{aligned}$$

and also

$$\begin{aligned} \mathbf{E}_1 \bullet \mathbf{e}_1 &= \mathrm{det}(R)\,\varepsilon_{ka}(\mathbf{e}_2)_a\,(\mathbf{e}_1)_k = \mathrm{det}(R)\mathrm{det}[\mathbf{e}_1, \mathbf{e}_2] = \mathrm{det}(R)\mathrm{det}(S) = 1 \\ \mathbf{E}_2 \bullet \mathbf{e}_2 &= -\mathrm{det}(R)\,\varepsilon_{ka}(\mathbf{e}_1)_a\,(\mathbf{e}_2)_k = -\mathrm{det}(R)\mathrm{det}[\mathbf{e}_2, \mathbf{e}_1] = \mathrm{det}(R)\mathrm{det}(S) = 1\ . \qquad (\text{A.6.7}) \end{aligned}$$

It is shown next that these N=2 results are special cases of a general fact: $\mathbf{E}_m \bullet \mathbf{e}_n = \delta_{m,n}$.

Eq. (5.11.5) showed that $\mathbf{e}_m \bullet \mathbf{e}_n = \bar{g}'_{mn}$. The other two dot products are now considered.

A.7 Compute $\mathbf{E}_m \bullet \mathbf{e}_n$

One can now compute, for general N,

$$\mathbf{E}_k \bullet \mathbf{e}_k = (\mathbf{E}_k)_\alpha(\mathbf{e}_k)_\alpha = \{\ \mathrm{det}(R)\,(-1)^{k-1}\varepsilon_{\alpha abc...x}\,(\mathbf{e}_1)_a(\mathbf{e}_2)_b \ldots\ldots (\mathbf{e}_N)_x\ \}\quad (\mathbf{e}_k)_\alpha\ . \qquad (\text{A.7.1})$$

k is missing

Slide α to the right in the ε subscript field and put it into the hole of the missing subscript κ, picking up $(-1)^{k-1}$. At the same time, move the $(\mathbf{e}_\kappa)_\alpha$ to the left and position it in its proper place in the product of factors,

$$\begin{aligned} \mathbf{E}_k \bullet \mathbf{e}_k &= (\mathbf{E}_k)_\alpha(\mathbf{e}_k)_\alpha = \{\ \mathrm{det}(R)\,\varepsilon_{abc..\alpha..x}\,(\mathbf{e}_1)_a(\mathbf{e}_2)_b \ldots (\mathbf{e}_\kappa)_\alpha \ldots (\mathbf{e}_N)_x\ \} \\ &= \mathrm{det}(R)\,\mathrm{det}\,[\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\ \ldots.\ \mathbf{e}_N\,] = \mathrm{det}(R)\,\mathrm{det}(S) = \mathrm{det}(RS) = \mathrm{det}(1) = 1 \qquad (\text{A.7.2}) \end{aligned}$$

where $S = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\ \ldots.\ \mathbf{e}_N\,]$ from (3.2.7). We already know that $\mathbf{E}_k$ is orthogonal to all the $\mathbf{e}_n$ which form the generalized cross product, therefore

$$\mathbf{E}_m \bullet \mathbf{e}_n = \delta_{m,n} \qquad (\text{A.7.3})$$

which is the "duality relation" discussed more generally in Section 6.2.

A.8 Compute $\mathbf{E}_n \bullet \mathbf{E}_m$

Since the vectors { $\mathbf{e}_n$ } are linearly independent and thus form a basis in R^N, $\mathbf{E}_m$ can be expanded onto the $\mathbf{e}_n$,

$$\mathbf{E}_m = \Sigma_n A_n^{(m)} \mathbf{e}_n \tag{A.8.1}$$

$$\delta_{m,k} = \mathbf{E}_m \bullet \mathbf{e}_k = \Sigma_n A_n^{(m)} \mathbf{e}_n \bullet \mathbf{e}_k = \Sigma_n A_n^{(m)} \bar{g}'_{nk} . \tag{A.8.2}$$

Multiplying both sides of (A.8.2) by g'_{ki} and summing on k gives

$$\text{LHS} = \Sigma_k g'_{ki} \delta_{m,k} = g'_{mi}$$

$$\text{RHS} = \Sigma_n A_n^{(m)} (\Sigma_k \bar{g}'_{nk} g'_{ki}) = \Sigma_n A_n^{(m)} (\bar{g}'g')_{ni} = \Sigma_n A_n^{(m)} \delta_{n,i} = A_i^{(m)} . \tag{A.8.3}$$

Therefore $A_i^{(m)} = g'_{mi}$ so,

$$\mathbf{E}_m = \Sigma_n A_n^{(m)} \mathbf{e}_n = \Sigma_n g'_{mn} \mathbf{e}_n \tag{A.8.4}$$

which is to say $\mathbf{E}_k$ is this linear combination of the $\mathbf{e}_i$, which is the definition used in Chapter 6, (6.1.2),

$$\mathbf{E}_k = \Sigma_i g'_{ki} \mathbf{e}_i = g'_{ki} \mathbf{e}_i \quad \text{// implied sum on i} \quad \text{// Std Notation: } e^k = \Sigma_i g'^{ki} \mathbf{e}_i \tag{A.8.5}$$

This may be compared with the previous result (A.6.2),

$$\mathbf{E}_k \equiv \det(R) (-1)^{k-1} \mathbf{e}_1 \times \mathbf{e}_2 \times \text{......} \times \mathbf{e}_N \qquad \text{// } \mathbf{e}_k \text{ missing;} \qquad \textit{(A.6.2)}$$

It seems rather impressive that these two dissimilar ways of writing **E** are equal. Finally,

$$\mathbf{E}_n \bullet \mathbf{E}_m = \mathbf{E}_n \bullet (g'_{mi} \mathbf{e}_i) = g'_{mi} (\mathbf{E}_n \bullet \mathbf{e}_i) = g'_{mi} \delta_{n,i} = g'_{mn} = g'_{nm} \quad \text{// g' is symmetric} \tag{A.8.6}$$

A.9 Summary of relationship between the tangent and reciprocal base vectors

$$\mathbf{e}_n \bullet \mathbf{e}_m = \bar{g}'_{nm} \qquad \mathbf{E}_n \bullet \mathbf{E}_m = g'_{nm} \qquad \mathbf{e}_n \bullet E_m = \delta_{n,m}$$

$$\mathbf{E}_n = \Sigma_i g'_{ni} \mathbf{e}_i \qquad \mathbf{e}_n = \Sigma_i \bar{g}'_{ni} \mathbf{E}_i \qquad \bar{g}' = g'^{-1} . \tag{A.9.1}$$

Although these results have just been derived in the Picture B context, they are also valid in the more general Picture A context, as shown in Chapter 6 in which the equation $\mathbf{E}_n = \Sigma_i g'_{ni} \mathbf{e}_i$ is used as the definition of $\mathbf{E}_n$. As a reminder, the cross product expression for $\mathbf{E}_n$ is only valid in Picture B where g=1.

In Standard Notation, the summary above can be restated as

$$\mathbf{e}_n \bullet \mathbf{e}_m = g'_{nm} \qquad \mathbf{e}^n \bullet \mathbf{e}^m = g'^{nm} \qquad \mathbf{e}_n \bullet \mathbf{e}^m = \delta_n{}^m$$

$$\mathbf{e}^n = \Sigma_i g'^{ni} \mathbf{e}_i \qquad \mathbf{e}_n = \Sigma_i g'_{ni} \mathbf{e}^i \qquad g'_{ab} = (g'^{ab})^{-1} \tag{A.9.2}$$

A.10 Another Cross Product Notation and another expression for E

Go back to the general cross product of N-1 vectors each of dimension N,

$$\mathbf{Q} = \mathbf{B} \text{ x } \mathbf{C} \text{ x } \mathbf{D} \text{ x } ... \text{ x } \mathbf{X} \quad . \qquad \text{// N-1 factors, N-2 crosses} \qquad \textit{(A.4.2)}$$

Replace **B**,**C**,**D ...** by vectors $\mathbf{A}^{(n)}$,

$$\mathbf{Q} = \mathbf{A}^{(1)} \text{ x } \mathbf{A}^{(2)} \text{ x } \mathbf{A}^{(3)} \text{ x ... x } \mathbf{A}^{(N-1)} \qquad \text{// N-1 factors, N-2 crosses} \qquad (A.10.1)$$

It is convenient to write this using a product symbol Π^x ,

$$\mathbf{Q} = \Pi^x{}_{i=1}{}^{N-1}\, \mathbf{A}^{(i)} = \Pi^x{}_i\, \mathbf{A}^{(i)} \qquad (A.10.2)$$

where in the second form it is understood that i takes on all values i = 1 to N-1. The superscript x means that this is not a regular product, it is our generalized cross product. This Π^x symbol also implies correct handling of the special case N=2 such that

$$\mathbf{Q} = \Pi^x{}_{i=1}{}^{1}\, \mathbf{A}^{(i)} \equiv \mathbf{A}^{(1)} \text{ x } \hat{\mathbf{3}} \qquad \text{// } \neq \mathbf{A}^{(1)} \qquad (A.10.3)$$

as discussed in (A.4.7) above.

This same Π^x notation can be applied to the "missing man formation" of Section A.5 above. Suppose

$$\mathbf{Q} = \mathbf{A}^{(1)} \text{ x } \mathbf{A}^{(2)} \text{ x } \mathbf{A}^{(3)} \text{ x ... x } \mathbf{A}^{(N)} \qquad \text{// } \mathbf{A}^{(n)} \text{ is missing .} \qquad (A.10.4)$$

One can write this as

$$\mathbf{Q} = \Pi^x{}_{i=1..N,i\neq n}\, \mathbf{A}^{(i)} \equiv \Pi^x{}_{i\neq n}\, \mathbf{A}^{(i)} \ . \qquad (A.10.5)$$

And of course this idea can be applied to the expression for $\mathbf{E_k}$

$$\mathbf{E_k} = \det(R)\, (-1)^{k-1}\, \mathbf{e_1} \text{ x } \mathbf{e_2} \text{ x ......x } \mathbf{e_N} \qquad \text{// } \mathbf{e_k} \text{ missing;} \qquad \textit{(A.6.2)}$$

$$\mathbf{E_k} = \det(R)\, (-1)^{k-1}\, \Pi^x{}_{i\neq k}\, \mathbf{e_i} \ . \qquad (A.10.6)$$

Once again, for N=2 the Π^x symbol implies that ($\mathbf{e_k}$ = "missing", $\mathbf{e_{3-k}}$ = the one not missing)

$$\Pi^x{}_{i\neq k}\, \mathbf{e_i} = \Pi^x{}_{i=1..2,i\neq k}\, \mathbf{e_i} = \mathbf{e_{3-k}} \text{ x } \hat{\mathbf{3}} \qquad (A.10.7)$$

$$\mathbf{E_k} = \det(R)\, (-1)^{k-1}\, \mathbf{e_{3-k}} \text{ x } \hat{\mathbf{3}} \qquad (A.10.8)$$

which is the "trick" notation of s(A.4.7) above for the N=2 case.

Appendix B: The Geometry of Parallelepipeds in N dimensions

B.1 Overview

This Appendix presents a simple method for constructing an N dimensional parallelepiped, which name we shorten to "N-piped". It is found that an N-piped has $2^{\mathbf{N}}$ vertices and N pairs of faces for a total of 2N faces, and the locus of points that make up each of these faces is stated. Each face of an N-piped is in fact an (N-1)-piped which has $2^{\mathbf{N-1}}$ vertices and is planar in N dimensions (meaning it lies on an N-1 dimensional flat surface). The two faces which make up each face pair lie on parallel planes in $R^{\mathbf{N}}$.

For example, for N=3 each face is a 2-piped having 2^{3-1} = 4 vertices, and there are N=3 face pairs for a total of 6 faces, and each pair of faces is planar in 3 dimensions.

For N=4 there are 4 pairs of faces for a total of 8 faces. Each face is a 3-piped having 2^{4-1} = 8 vertices. For example, one would say that each face of a 4-cube is a 3-cube. It is not intuitively obvious that two faces each of which is a regular cube can in fact lie on surfaces which are planar and parallel in 4 dimensions, but we show how this works below.

It is then shown that, if the N-piped is spanned by the N tangent base vectors $\mathbf{e_n}$ of Chapter 3, the normal vectors for the pairs of parallel faces are just the reciprocal base vectors $\mathbf{E_n}$ of Chapter 6.

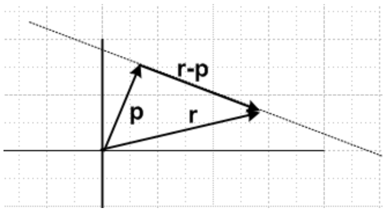
Section B.5 focuses on the area and volume of N-pipeds in various dimensions, and simple expressions for the volume and vector areas of the faces of an N-piped are obtained.

Rather than just state the results in N dimensions, we attempt an inductive approach to provide motivation for the N dimensional results. In this approach, cases N = 2,3.. are treated with nearly identical boilerplate templates to build up the inductive case.

A reader interested in the details here should read Appendix A first since use is made of various Appendix A results including the generalized cross product idea. On the other hand, a reader not interested in details might just read the Appendix B Summary presented below in Section B.6.

B.2 Preliminary: Equation of a plane in N dimensions

Consider an arbitrary plane drawn in N space which does not pass through the origin. There is some point on that plane which lies closer to the origin than all other points on the plane. Let **p** be a vector from the origin to that closest point, and let **r** represent a point lying on the plane,



(B.2.1)

Since **p** is normal to the plane, and since **r-p** is a vector lying *in* the plane, it follows that

$$\mathbf{p}\bullet(\mathbf{r}-\mathbf{p}) = 0 \quad\Rightarrow\quad \mathbf{r}\bullet\mathbf{p} = p^2 \quad\Rightarrow\quad \mathbf{r}\bullet\hat{\mathbf{p}} = p \ . \qquad (B.2.2)$$

Therefore, one way to write the equation of a plane in N dimensions is

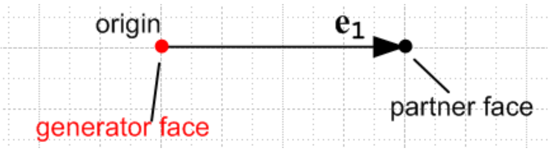
$$\mathbf{r}\bullet\hat{\mathbf{p}} = p \qquad\qquad \mathbf{r} = (x_1, x_2, ...x_N) \ . \tag{B.2.3}$$

where $\hat{\mathbf{p}}$ is the unit vector normal to the plane which points "away from the origin", and where $p > 0$ is the distance of closest approach of the plane to the origin. In the limit $p\rightarrow 0$, the plane passes through the origin and the equation is then $\mathbf{r}\bullet\hat{\mathbf{p}} = 0$ where $\hat{\mathbf{p}}$ is either normal to the plane.

B.3 N-pipeds and their Faces in Various Dimensions

(a) The 1-piped

Start with N = 1 where the piped is some arbitrary line segment $\mathbf{e}_1$ in direction $\hat{\mathbf{e}}_1$ having length e_1, with one end affixed to the origin of the real axis :



(B.3.a.1)

This piped has two vertices located at $\mathbf{v}_1 = 0$ and $\mathbf{v}_2 = \mathbf{e}_1$. These two vertices are also the "faces" of this 1-piped, so there are two faces (one pair of faces). These faces are 0 dimensional and therefore don't point in any direction (they are the endpoints of the vector $\mathbf{e}_1$). The 1-piped is a piece of a plane in 1 dimension (a line). One can think of the vertex at the origin as the "generator 0-piped" and the other vertex as the partner face of the generator, in the sense of the generator idea described below.

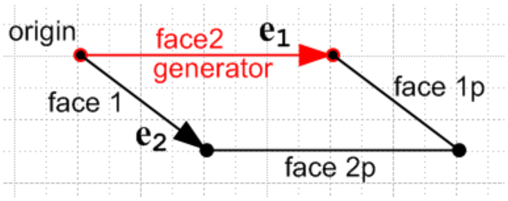
The loci of points in the 1-piped's interior "volume" is given by

$$\mathbf{r}_{\mathtt{volume1}} = \alpha_1 \, \mathbf{e}_1 \qquad\qquad 0 \le \alpha_1 \le 1 \ . \tag{B.3.a.2}$$

The volume of this 1-piped is e_1.

(b) The 2-piped

Now add another dimension, going to N=2. Introduce a unit vector $\hat{\mathbf{e}}_2$ in some arbitrary direction in R^2 other than $\mathbf{e}_1$ so that $\mathbf{e}_1$ and $\mathbf{e}_2$ are linearly independent. Take the 1-piped described above (line segment) and translate it by $\mathbf{e}_2$ to create a new copy of the line segment. The original 1-piped we call the generator piped, and the copy is the partner of the generator piped which, it will be shown, lies on a plane (a 1-plane = line) which is parallel to the plane of the generator piped, but its plane does not pass through the origin. In N=2 dimensions, the generator 1-piped and its partner are now "faces" of a 2-dimensional object, a parallelogram = a 2-piped. Draw line segments from all the vertices of the generator piped to matching vertices of its partner piped (add 2 line segments) to make 2 additional "side" faces. One of these faces necessarily touches the origin, and the other face does not. Faces always occur in parallel pairs, one face of which touches the origin, and one of which does not, the latter we will call the "partner" face. For our 2-piped, each face is a 1-piped. There are now four faces, each is a line segment.



(B.3.b.1)

The loci of the 2-piped's volume and of its four 1-piped faces are given by

$$\mathbf{r}_{\mathtt{volume2}} = \alpha_1\mathbf{e}_1 + \alpha_2\,\mathbf{e}_2 \qquad 0 \le \alpha_1,\alpha_2 \le 1$$

$$\mathbf{r}_{\mathtt{face2}} = \alpha_1\mathbf{e}_1 \qquad 0 \le \alpha_1 \le 1 \qquad \text{// the generator face}$$
$$\mathbf{r}_{\mathtt{face2p}} = \alpha_1\mathbf{e}_1 + \mathbf{e}_2 \qquad 0 \le \alpha_1 \le 1 \qquad \text{// partner of the generator face}$$
$$\mathbf{r}_{\mathtt{face1}} = \alpha_2\mathbf{e}_2 \qquad 0 \le \alpha_2 \le 1 \qquad \text{// side face touching the origin}$$
$$\mathbf{r}_{\mathtt{face1p}} = \alpha_2\mathbf{e}_2 + \mathbf{e}_1 \qquad 0 \le \alpha_2 \le 1 \qquad \text{// partner of the above side face} \qquad \text{(B.3.b.2)}$$

The origin-touching faces are numbered using the index of the $\mathbf{e}_n$ vector that does *not* appear in the locus for the face. This seems strange but for N > 2 it will be clear why this is done. The above results can be summarized as

$$\mathbf{r}_{\mathtt{volume3}} = \Sigma_n\alpha_n\mathbf{e}_n \qquad 0 \le \alpha_n \le 1$$

$$\mathbf{r}_{\mathtt{face(i)}} = \Sigma_{n\ne i}\alpha_n\mathbf{e}_n \qquad 0 \le \alpha_n \le 1 \qquad i = 1,2,3$$
$$\mathbf{r}_{\mathtt{face(ip)}} = \Sigma_{n\ne i}\alpha_n\mathbf{e} + \mathbf{e}_i \qquad 0 \le \alpha_n \le 1 \qquad i = 1,2,3 \qquad \text{(B.3.b.3)}$$

According to Note 1 following (6.2.7), it is possible to construct vectors $\mathbf{E}_1$ and $\mathbf{E}_2$ as linear combinations of $\mathbf{e}_1$ and $\mathbf{e}_2$ such that the following is true:

$$\mathbf{E}_i\bullet\,\mathbf{e}_j = \delta_{i,j} \qquad \text{// } \mathbf{E}_k = \Sigma_{i=1}^{\;2}\, g'_{ki}\,\mathbf{e}_i\text{, see Note 4 following (6.2.7)} \qquad \text{(B.3.b.4)}$$

If one interprets the $\mathbf{e}_n$ vectors as tangent base vectors for some transformation F, then the two vectors $\mathbf{E}_n$ are the corresponding reciprocal base vectors which are discussed in Chapter 6 and Appendix A.

Consider now these dot products:

$$\mathbf{E}_2 \bullet \mathbf{r}_{\mathtt{face2}} = \mathbf{E}_2 \bullet \alpha_1\mathbf{e}_1 = 0 \qquad \Rightarrow \qquad \hat{\mathbf{E}}_2 \bullet \mathbf{r}_{\mathtt{face2}} = 0$$
$$\mathbf{E}_2 \bullet \mathbf{r}_{\mathtt{face2p}} = \mathbf{E}_2 \bullet [\,\alpha_1\mathbf{e}_1 + \mathbf{e}_2] = 1 \qquad \Rightarrow \qquad \hat{\mathbf{E}}_2 \bullet \mathbf{r}_{\mathtt{face2p}} = 1/E_2\ . \qquad \text{(B.3.b.5)}$$

According to (B.2.3), the first line says that face 2 lies on a plane which passes through the origin and which has **normal vector** $\hat{\mathbf{E}}_2$. Also according to (B.2.3), the second line says that face 2p has the same normal and its plane is therefore parallel to face 1 but misses the origin by distance $1/|E_2|$. Similarly,

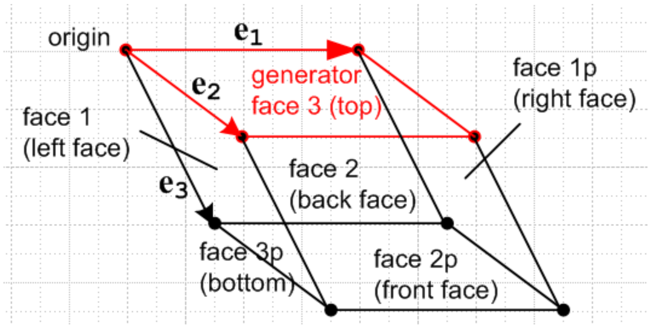
$$\mathbf{E}_1 \bullet \mathbf{r}_{\mathtt{face1}} = \mathbf{E}_1 \bullet \alpha_2\mathbf{e}_2 = 0 \qquad \Rightarrow \qquad \hat{\mathbf{E}}_1 \bullet \mathbf{r}_{\mathtt{face1}} = 0$$
$$\mathbf{E}_1 \bullet \mathbf{r}_{\mathtt{face1p}} = \mathbf{E}_1 \bullet [\,\alpha_2\mathbf{e}_2 + \mathbf{e}_1] = 1 \qquad \Rightarrow \qquad \hat{\mathbf{E}}_1 \bullet \mathbf{r}_{\mathtt{face1p}} = 1/E_1\ . \qquad \text{(B.3.b.6)}$$

These two faces are also parallel, both having normal $\hat{\mathbf{E}}_1$. The first touches the origin while the partner's plane misses the origin by distance $1/|E_1|$.

The conclusions that $\mathbf{E_n}$ is normal to face n, and that the pair of faces n and np are parallel, do not depend on the specific upper endpoints of the ranges of α_1 and α_2 which happen to be given as 1 above. This seems pretty obvious since rescaling the edges of a parallelogram does not affect its normal vector.

(c) The 3-piped

Now add another dimension, going to N=3. Introduce a unit vector $\hat{\mathbf{e}}_3$ in some arbitrary direction in R^3 so that $(\hat{\mathbf{e}}_1,\hat{\mathbf{e}}_2,\hat{\mathbf{e}}_3)$ are linearly independent. Take the 2-piped described above (parallelogram) and translate it by distance e_3 in the $\hat{\mathbf{e}}_3$ direction to create a new copy of the 2-piped. The original 2-piped we call the generator piped, and the copy is the partner of the generator piped which, as will now be shown, lies on a plane which is parallel to that of the generator piped, but which does not pass through the origin. In N=3 dimensions, the generator 2-piped and its partner are now "faces" of a 3-dimensional object, a parallelepiped = a 3-piped. Draw line segments from all 2^2 vertices of the generator piped to the corresponding vertices of its partner piped (add 4 line segments), to get 4 additional side faces. Two of these faces necessarily touch the origin, and the other two do not. For the 3-piped, each face is a 2-piped. There are now 2*3 = 6 faces, each is a 2-piped.



(B.3.c.1)

The loci of the 3-piped's volume and of its six 2-piped faces are given by

$$\mathbf{r}_{\mathbf{volume3}} = \alpha_1\mathbf{e}_1 + \alpha_2\,\mathbf{e}_2 + \alpha_3\,\mathbf{e}_3 \qquad 0 \le \alpha_1,\alpha_2,\alpha_3 \le 1$$

$\mathbf{r}_{\mathbf{face3}} = \alpha_1\mathbf{e}_1 + \alpha_2\,\mathbf{e}_2$ $\quad 0 \le \alpha_1,\alpha_2 \le 1$ // the generator face

$\mathbf{r}_{\mathbf{face3p}} = \alpha_1\mathbf{e}_1 + \alpha_2\,\mathbf{e}_2 + \mathbf{e}_3$ $\quad 0 \le \alpha_1,\alpha_2 \le 1$ // partner face to the above

$\mathbf{r}_{\mathbf{face2}} = \alpha_1\mathbf{e}_1 + \alpha_3\,\mathbf{e}_3$ $\quad 0 \le \alpha_1,\alpha_3 \le 1$ // the generator face

$\mathbf{r}_{\mathbf{face2p}} = \alpha_1\mathbf{e}_1 + \alpha_3\,\mathbf{e}_3 + \mathbf{e}_2$ $\quad 0 \le \alpha_1,\alpha_3 \le 1$ // partner face to the above

$\mathbf{r}_{\mathbf{face1}} = \alpha_2\mathbf{e}_2 + \alpha_3\,\mathbf{e}_3$ $\quad 0 \le \alpha_2,\alpha_3 \le 1$ // the generator face

$\mathbf{r}_{\mathbf{face1p}} = \alpha_2\mathbf{e}_2 + \alpha_3\,\mathbf{e}_3 + \mathbf{e}_1$ $\quad 0 \le \alpha_2,\alpha_3 \le 1$ // partner face to the above (B.3.c.2)

Notice that the partner face locus is created from the non-partner face by adding "the other" base vector. For example, face 3 is "spanned" by base vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ so $\mathbf{e}_3$ is added to get the partner face face3p. A partner is just a copy of the non-partner which is translated by a constant vector. The above results can be summarized in this concise manner:

$$\mathbf{r}_{\mathtt{volume3}} = \Sigma_{\mathbf{n}}\alpha_{\mathbf{n}}\mathbf{e}_{\mathbf{n}} \qquad 0 \le \alpha_{\mathbf{n}} \le 1$$

$$\begin{aligned} &\mathbf{r}_{\mathtt{face(i)}} = \Sigma_{\mathbf{n}\neq\mathbf{i}}\alpha_{\mathbf{n}}\mathbf{e}_{\mathbf{n}} && 0 \le \alpha_{\mathbf{n}} \le 1 && i = 1,2,3 \\ &\mathbf{r}_{\mathtt{face(ip)}} = \Sigma_{\mathbf{n}\neq\mathbf{i}}\alpha_{\mathbf{n}}\mathbf{e} + \mathbf{e}_{\mathbf{i}} && 0 \le \alpha_{\mathbf{n}} \le 1 && i = 1,2,3 \end{aligned} \qquad \text{(B.3.c.3)}$$

According to Note 1 following (6.2.7), it is possible to construct vectors $\mathbf{E}_1$, $\mathbf{E}_2$, $\mathbf{E}_3$ as linear combinations of $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ such that the following is true,

$$\mathbf{E}_{\mathbf{i}}\bullet \mathbf{e}_{\mathbf{j}} = \delta_{\mathbf{i},\mathbf{j}} \qquad // \ \mathbf{E}_{\mathbf{k}} = \Sigma_{\mathbf{i}=1}{}^{3}\, g'_{\mathbf{ki}}\, \mathbf{e}_{\mathbf{i}}, \text{ see Note 4 following (6.2.7)} \qquad \text{(B.3.c.4)}$$

If one interprets the $\mathbf{e}_{\mathbf{n}}$ vectors as tangent base vectors, then the three vectors $\mathbf{E}_{\mathbf{n}}$ are the corresponding reciprocal base vectors which are discussed in Chapter 6 and Appendix A. Consider now these dot products:

$$\begin{aligned} &\mathbf{E}_1\bullet \mathbf{r}_{\mathtt{face1}} = \mathbf{E}_1\bullet [\ \alpha_2\mathbf{e}_2 + \alpha_3\, \mathbf{e}_3] = 0 && \Rightarrow && \hat{\mathbf{E}}_1\bullet \mathbf{r}_{\mathtt{face1}} = 0 \\ &\mathbf{E}_1\bullet \mathbf{r}_{\mathtt{face1p}} = \mathbf{E}_1\bullet [\alpha_2\mathbf{e}_2 + \alpha_3\, \mathbf{e}_3 + \mathbf{e}_1] = 1 && \Rightarrow && \hat{\mathbf{E}}_1\bullet \mathbf{r}_{\mathtt{face1p}} = 1/E_1\ . \end{aligned} \qquad \text{(B.3.c.5)}$$

According to (B.2.3), the first line says that face 1 lies on a plane which passes through the origin and which has normal vector $\hat{\mathbf{E}}_1$. Also according to (B.2.3), the second line says that face 1p has the same normal and its plane is therefore parallel to face 1 but misses the origin by distance $1/|E_1|$.

A similar pair of equations obtains for each of the other face pairs.

(d) The N-piped

Now add another dimension, going from N-1 to N. Introduce a unit vector $\hat{\mathbf{e}}_{\mathbf{N}}$ in some arbitrary direction in $R^{\mathbf{N}}$ so that ($\hat{\mathbf{e}}_1 ... \hat{\mathbf{e}}_{\mathbf{N}}$) are linearly independent. Take the (N-1)-piped described above and translate it by distance $e_{\mathbf{N}}$ in the $\hat{\mathbf{e}}_{\mathbf{N}}$ direction to create a new copy of the (N-1)-piped. The original (N-1)-piped we call the generator piped, and the copy is the partner of the generator piped which, it will be shown, lies on a plane which is parallel to that of the generator piped, but which does not pass through the origin. The generator (N-1)-piped and its partner are now "faces" of a N-dimensional object, an N-piped. Adding this partner piped doubles the total vertex count. Draw line segments from all $2^{\mathbf{N-1}}$ vertices of the generator piped to the corresponding vertices of its partner piped to get 2N-2 additional side faces for a total now of 2N faces. There are N pairs of "faces" because there are N ways to omit a single $\mathbf{e}_{\mathbf{i}}$ from the list of vectors which span a face, so including the partner faces an N-piped has 2N faces in total. Half of these faces necessarily touch the origin, and the other half do not. Each face is an (N-1)-piped.

It is convenient to refer to the partner face of a pair as "the far face" and the other one, which touches the origin, as "the near face".

We leave the N-piped drawing which would be labeled (B.3.d.1) to the reader's imagination, and we omit a detailed list of the face volumes which we could called (B.3.d.2), giving just the summary below.

The loci of the N-piped's volume and of its 2N (N-1)-piped faces are given by:

$$\begin{array}{llll} \mathbf{r}_{\mathtt{volumeN}} = \Sigma_n\alpha_n\mathbf{e}_n & 0 \le \alpha_n \le 1 & & \\ \mathbf{r}_{\mathtt{face(i)}} = \Sigma_{n\neq i}\alpha_n\mathbf{e}_n & 0 \le \alpha_n \le 1 & i = 1,2...N & \\ \mathbf{r}_{\mathtt{face(ip)}} = \Sigma_{n\neq i}\alpha_n\mathbf{e}_n + \mathbf{e}_i & 0 \le \alpha_n \le 1 & i = 1,2...N & \text{(B.3.d.3)} \end{array}$$

Notice that the partner face locus is created from the non-partner face by adding "the other" base vector. For example, face i is "spanned" by base vectors $\mathbf{e}_n$ $n \neq i$, so it is $\mathbf{e}_i$ that one adds to get the partner. A partner is just a copy of the non-partner which is translated by a constant vector.

According to Note 1 following (6.2.7), it is possible to construct vectors $\mathbf{E}_1...\mathbf{E}_N$ as linear combinations of $\mathbf{e}_1.. \mathbf{e}_N$ such that the following is true,

$$\mathbf{E}_i\bullet \mathbf{e}_j = \delta_{i,j} \qquad // \ \mathbf{E}_k = \Sigma_{i=1}{}^2\, g'_{ki}\, \mathbf{e}_i, \text{ see Note 4 following (6.2.7)} \qquad \text{(B.3.d.4)}$$

If one interprets the $\mathbf{e}_n$ vectors as tangent base vectors, then the N vectors $\mathbf{E}_n$ are the corresponding reciprocal base vectors which are discussed in Chapter 6 and Appendix A. Consider now these dot products:

$$\begin{array}{llll} \mathbf{E}_i\bullet \mathbf{r}_{\mathtt{face(i)}} = \mathbf{E}_i\bullet [\Sigma_{n\neq i}\alpha_n\mathbf{e}_n] = 0 & \Rightarrow & \hat{\mathbf{E}}_i\bullet \mathbf{r}_{\mathtt{face(i)}} = 0 & \\ \mathbf{E}_i\bullet \mathbf{r}_{\mathtt{face(ip)}} = \mathbf{E}_i\bullet [\Sigma_{n\neq i}\alpha_n\mathbf{e} + \mathbf{e}_i] = 1 & \Rightarrow & \hat{\mathbf{E}}_i\bullet \mathbf{r}_{\mathtt{face(ip)}} = 1/E_i & \text{(B.3.d.5)} \end{array}$$

The first line says that face i lies on a plane which passes through the origin and which has normal vector $\hat{\mathbf{E}}_i$. The second line says that face ip has the same normal and its plane is therefore parallel to face i but misses the origin by distance $1/|E_i|$.

B.4 The question of inward versus outward facing normal vectors

It has been shown above that, for an N-piped, the pair of faces i and ip has normal vector $\mathbf{E}_i$. For one of these faces, $\mathbf{E}_i$ will be an outward directed normal, while for the other it will be an inward directed normal. One might like to know which is which. Here is one way to find out.

First, construct these three vectors

$$\begin{array}{lll} (\text{piped})_{\mathtt{center}} = \Sigma_n(1/2)\mathbf{e}_n & // \text{ vector from origin to piped center} & \\ (\text{face i})_{\mathtt{center}} = \Sigma_{n\neq i}(1/2)\mathbf{e}_n & // \text{ vector from origin to center of face i} & \\ (\text{face ip})_{\mathtt{center}} = \Sigma_{n\neq i}(1/2)\, \mathbf{e}_n + \mathbf{e}_i & // \text{ vector from origin to center of face ip .} & \text{(B.4.1)} \end{array}$$

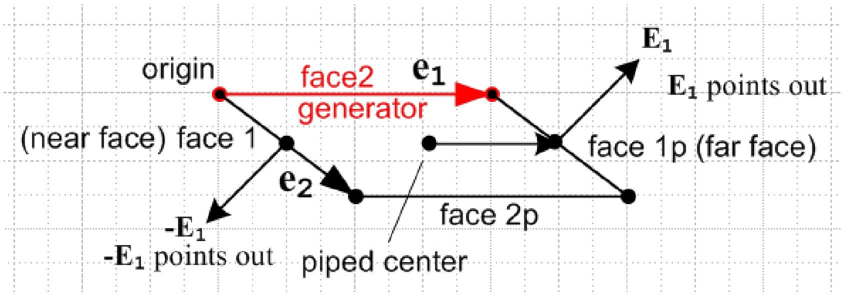
Construct vectors from piped center to face centers (results here are fairly obvious)

$$(\text{face i})_{\text{center}} - (\text{piped})_{\text{center}} = [\Sigma_{n\neq i}(1/2)\mathbf{e}_n] - \Sigma_n(1/2)\mathbf{e}_n = -(1/2)\mathbf{e}_i$$
$$(\text{face ip})_{\text{center}} - (\text{piped})_{\text{center}} = [\Sigma_{n\neq i}(1/2)\mathbf{e}_n + \mathbf{e}_i\ ] - \Sigma_n(1/2)\mathbf{e}_n = +(1/2)\mathbf{e}_i\ . \qquad (B.4.2)$$

Then compute

$$\mathbf{E}_i \bullet \{(\text{face i})_{\text{center}} - (\text{piped})_{\text{center}}\ \} = \mathbf{E}_i \bullet [-(1/2)\mathbf{e}_i\ ] = -(1/2) < 0$$
$$\mathbf{E}_i \bullet \{(\text{face ip})_{\text{center}} - (\text{piped})_{\text{center}}\ \} = \mathbf{E}_i \bullet [+(1/2)\mathbf{e}_i\ ] = +(1/2) > 0\ . \qquad (B.4.3)$$

One may conclude that $\mathbf{E}_i$ is an outward pointing normal for face ip (the far face of the pair). Therefore, $-\mathbf{E}_i$ is an outward pointing normal for face i (the near face of the pair), which recall is the face which touches the origin.. Here is an illustration for the case N = 2:



(B.4.4)

B.5 The Face Area and Volume of N-pipeds in Various Dimensions

We embark now on another long march to inductively arrive at results for the general N case. Tracing the first few cases N = 2,3,4 and then extrapolating to N = N probably gives more insight than a formal induction proof which is not attempted here. Each case below is treated with the same boilerplate template which first treats Face Area and then Volume.

(a) The 2-piped

For this case N = 2, the reader is encouraged to stare at the above Fig (B.4.4) while reading the text.

Face Area. The area of a 2-piped face (a line segment) is just the length of the edge which is the face,

$$A_1 = |\mathbf{e}_2|$$
$$A_2 = |\mathbf{e}_1|\ . \qquad (B.5.a.1)$$

where here we maintain the plan of labeling an area by the index of the spanning vector which is omitted in making the area. The *vector* areas can be written, based on the work above,

$$\mathbf{A}_1 = |\mathbf{e}_2|\ \hat{\mathbf{E}}_1 \qquad // \ \mathbf{E}_k = \Sigma_i\ g'_{ki}\ \mathbf{e}_i\ \text{, see Note 4 following (6.2.7)}$$
$$\mathbf{A}_2 = |\mathbf{e}_1|\ \hat{\mathbf{E}}_2 \qquad (B.5.a.2)$$

and these vectors are out-facing for faces 1p and 2p. We claim that both these results can be expressed in a single formula,

$$\mathbf{A_n} = |\det(S)|\, \mathbf{E_n} \,. \qquad \text{(B.5.a.3)}$$

One can see that the direction is correct for n = 1,2, so it is just a question of verifying the magnitude. One must show that

$$|\det(S)|\ |\mathbf{E_1}| = |\mathbf{e_2}| \qquad \text{and} \qquad |\det(S)|\ |\mathbf{E_2}| = |\mathbf{e_1}| \qquad \text{(B.5.a.4)}$$

or

$$|\mathbf{E_k}| = |\det(R)|\, |\mathbf{e_{3-k}}| \qquad k=1,2 \quad . \qquad // \text{ RS} = 1 \qquad \text{(B.5.a.5)}$$

Using the N=2 trick notation from (A.6.5),

$$\mathbf{E_k} = \det(R)\, (-1)^{k-1}\, \mathbf{e_{3-k}} \times \hat{\mathbf{3}} \qquad \textit{(A.6.5)} \qquad \text{(B.5.a.6)}$$

so that

$$|\,\mathbf{E_k}\,| = |\det(R)|\ |\,\mathbf{e_{3-k}} \times \hat{\mathbf{3}}| = |\det(R)|\ |\,\mathbf{e_{3-k}}| \qquad k = 1,2 \qquad \text{(B.5.a.7)}$$

since $\mathbf{e_{3-k}}$ and $\hat{\mathbf{3}}$ are perpendicular. QED.

We stress the formula $\mathbf{A_n} = |\det(S)|\, \mathbf{E_n}$ because it will turn out that this is valid for all N ≥ 2 .

One can restate $\mathbf{A_n} = |\det(S)|\, \mathbf{E_n}$ using the cross product notation presented in (A.10.6):

$$\mathbf{A_n} = |\det(S)|\, \mathbf{E_n} = |\det(S)| \det(R)\, (-1)^{n-1}\ \Pi^{\mathbf{x}}{}_{i\neq n}\, \mathbf{e_i}$$

$$= \sigma\ (-1)^{n-1}\ \Pi^{\mathbf{x}}{}_{i\neq n}\, \mathbf{e_i} \qquad \sigma \equiv \text{sign}(\det(S)) = \text{sign}(\det(R)) \qquad \text{(B.5.a.8)}$$

with the understanding that for the N=2 case $\Pi^{\mathbf{x}}{}_{i\neq n}\, \mathbf{e_i}$ has the special meaning of (A.10.7), so that in fact

$$\mathbf{A_n} = \sigma\ (-1)^{n-1} [\ \Pi^{\mathbf{x}}{}_{i\neq n}\, \mathbf{e_i}\ ] = \sigma\ (-1)^{n-1} [\ \mathbf{e_{3-n}} \times \hat{\mathbf{3}}\ ] \qquad // \text{ for N = 2, n = 1,2} \qquad \text{(B.5.a.9)}$$

<u>Volume</u>. The volume of a 2-piped is the base times the height of a parallelogram, familiarly given by the cross product of the edges,

$$\text{volume}(2) = |\,\mathbf{e_1} \times \mathbf{e_2}\,| = |\,\varepsilon_{ab}\, (\mathbf{e_1})_a (\mathbf{e_2})_b\,| = |\det[\,\mathbf{e_1}, \mathbf{e_2}\,]\,| = |\det(S)\,| \quad // \text{ see (3.2.7)} \qquad \text{(B.5.a.10)}$$

where S is the linearized transformation matrix for N=2, see Chapter 2. Of course strictly in N=2 the notation $\mathbf{e_1} \times \mathbf{e_2}$ has no meaning, so one has to imagine a $\hat{\mathbf{3}}$ dimension to give it meaning. The second form *does* have a meaning for N=2, and that meaning is $|(\mathbf{e_1})_1 (\mathbf{e_2})_2 - (\mathbf{e_1})_2 (\mathbf{e_2})_1|$.

(b) The 3-piped

<u>Face Area:</u> The faces of a 3-piped are 2-pipeds. For N=2, the 2-piped volume is this from (B.5.a.9),

$$\text{volume}(2) = |\,\varepsilon_{ab}\, (\mathbf{e_1})_a (\mathbf{e_2})_b\,| \,, \qquad \textit{(B.5.a.9)} \qquad \text{(B.5.b.1)}$$

where $\mathbf{e}_1$ and $\mathbf{e}_2$ are 2D vectors. For the 2-piped which is "face 3" of the 3-piped -- a "near" face which touches the origin of the 3D skewed $\mathbf{e}_\mathbf{n}$ coordinate system -- vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ are 3D vectors. This face 3 is the top face shown in Fig (B.3.c.1) above. The first 2 components of each of these 3D vectors are the same as the components of the 2D $\mathbf{e}_\mathbf{i}$ vectors, while the 3rd components are both 0. This is so because face 3 lies in a plane defined by this 3rd component being 0. The volume(2) formula expressed in terms of these new 3D vectors is therefore $|\varepsilon_{\mathbf{ab3}}\ (\mathbf{e}_1)_\mathbf{a}(\mathbf{e}_2)_\mathbf{b}|$, where ε is now a 3D ε tensor. The conclusion is that the area of the top face in Fig (B.3.c.1) is,

$$A_3 = |\varepsilon_{\mathbf{ab3}}\ (\mathbf{e}_1)_\mathbf{a}(\mathbf{e}_2)_\mathbf{b}| \qquad \text{(B.5.b.2)}$$

and this then is the scalar area of both face 3 and its partner face 3p, the far face. Similar arguments would then support these other area expressions

$$A_1 = |\varepsilon_{\mathbf{1ab}}\ (\mathbf{e}_2)_\mathbf{a}(\mathbf{e}_3)_\mathbf{b}|$$
$$A_2 = |\varepsilon_{\mathbf{a2b}}\ (\mathbf{e}_3)_\mathbf{a}(\mathbf{e}_1)_\mathbf{b}| \ . \qquad \text{(B.5.b.3)}$$

Since indices on ε can be swapped for free due to the absolute value, the non-summed index can be put first in all three cases and one may then conclude that

$A_1 = |\mathbf{e}_2 \text{ x } \mathbf{e}_3|$ face 1 and face 1p
$A_2 = |\mathbf{e}_3 \text{ x } \mathbf{e}_1|$ face 2 and face 2p
$A_3 = |\mathbf{e}_1 \text{ x } \mathbf{e}_2|$. face 3 and face 3p (B.5.b.4)

In (A.3.3) it is shown that $\mathbf{E}_1 = \det(R)\ \mathbf{e}_2 \text{ x } \mathbf{e}_3$ so that $\mathbf{e}_2 \text{ x } \mathbf{e}_3$ lines up with $\mathbf{E}_1$ if $\det(R) > 0$. Regardless of the sign of det(R), we define the *vector* areas to point in the $+\hat{\mathbf{E}}_\mathbf{n}$ directions. Thus,

$\mathbf{A}_1 \equiv |\mathbf{e}_2 \text{ x } \mathbf{e}_3|\ \hat{\mathbf{E}}_1$ face 1p out-facing
$\mathbf{A}_2 \equiv |\mathbf{e}_3 \text{ x } \mathbf{e}_1|\ \hat{\mathbf{E}}_2$ face 2p out-facing
$\mathbf{A}_3 \equiv |\mathbf{e}_1 \text{ x } \mathbf{e}_2|\ \hat{\mathbf{E}}_3$. face 3p out-facing (B.5.b.5)

These three equations can be combined into the following single formula

$$\mathbf{A}_\mathbf{n} = |\mathbf{e}_1 \text{ x } \mathbf{...} \text{ x } \mathbf{e}_3|\ \hat{\mathbf{E}}_\mathbf{n} \ . \qquad // \ \mathbf{e}_\mathbf{n} \text{ missing} \qquad \text{(B.5.b.6)}$$

where the $\mathbf{e}_\mathbf{i}$ are reordered for free due to the absolute value signs. But (A.6.2) says

$$\mathbf{E}_\mathbf{n} = \det(R)\ (-1)^{\mathbf{n-1}}\ \mathbf{e}_1 \text{ x ... x } \mathbf{e}_3 \qquad // \ \mathbf{e}_\mathbf{n} \text{ missing} \qquad \textit{(A.6.2)} \qquad \text{(B.5.b.7)}$$

so

$$|\mathbf{E}_\mathbf{n}| = |\det(R)|\ |\ \mathbf{e}_1 \text{ x ... x } \mathbf{e}_3\ | \ . \qquad // \ \mathbf{e}_\mathbf{n} \text{ missing} \qquad \text{(B.5.b.8)}$$

Thus,

$$\begin{aligned} \mathbf{A_n} &= \hat{\mathbf{E}}_\mathbf{n}\ |\mathbf{E_n}| / |\det(R)| = |\det(S)|\ \mathbf{E_n} \\ &= |\det(S)|\ \det(R)\ (-1)^{n-1}\mathbf{e_1} \times ... \times \mathbf{e_3} && // \ \mathbf{e_n} \text{ missing} \\ &= \sigma\ (-1)^{n-1}\mathbf{e_1} \times ... \times \mathbf{e_3} && // \ \mathbf{e_n} \text{ missing} \\ &\text{where} \qquad \sigma \equiv \text{sign}[\det(S)] = \text{sign}[\det(R)] \end{aligned} \qquad \text{(B.5.b.9)}$$

To summarize, for N=3 one has

$$\begin{aligned} \mathbf{A_n} &= |\det(S)|\ \mathbf{E_n} = \sigma\ (-1)^{n-1}\mathbf{e_1} \times ... \times \mathbf{e_3} && // \ \mathbf{e_n} \text{ missing} \\ &= \sigma\ (-1)^{n-1}\ \Pi^{x}{}_{i\neq n}\ \mathbf{e_i} \end{aligned} \qquad \text{(B.5.b.10)}$$

where the last line uses the shorthand notation of (A.10.5). This $\mathbf{A_n}$ result has the same formal form as that of the 2-piped case (B.5.a.8).

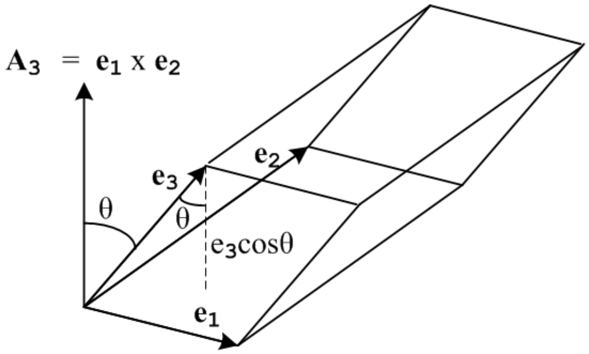
Volume. The volume of a 3-piped (with each of the above $\mathbf{A_n}$ in turn treated as the base) is base times height, so

$$\text{volume}(3) = |\ \mathbf{A_1} \bullet \mathbf{e_1}\ | = |\ \mathbf{A_2} \bullet \mathbf{e_2}\ | = |\ \mathbf{A_3} \bullet \mathbf{e_3}\ |$$

or

$$\text{volume}(3) = |\ \mathbf{e_2} \times \mathbf{e_3} \bullet \mathbf{e_1}\ | = |\ \mathbf{e_3} \times \mathbf{e_1} \bullet \mathbf{e_2}\ | = |\ \mathbf{e_1} \times \mathbf{e_2} \bullet \mathbf{e_3}\ | \quad . \qquad \text{(B.5.b.11)}$$

Here is a drawing showing the last case ($\sigma = +1$), where "base" is $A_3 = |\ \mathbf{e_1} \times \mathbf{e_2}\ |$ and "height" is $e_3 \cos\theta$,



(B.5.b.12)

Using ε notation one can write

$$\mathbf{e_3} \bullet \mathbf{e_1} \times \mathbf{e_2} = (e_3)_i\ \varepsilon_{ijk}(e_1)_j(e_2)_k = \varepsilon_{ijk}\ (e_1)_j(e_2)_k\ (e_3)_i = \varepsilon_{jki}\ (e_1)_j(e_2)_k(e_3)_i$$

$$= \det[\ \mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}] = \det(S) \qquad // \text{ see (3.2.7)} \qquad \text{(B.5.b.13)}$$

so that

$$\text{volume}(3) = |\ \mathbf{e_3} \bullet \mathbf{e_1} \times \mathbf{e_2}\ | = |\ \varepsilon_{abc}\ (\mathbf{e_1})_a(\mathbf{e_2})_b(\mathbf{e_3})_c\ | = |\ \det[\ \mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}]\ | = |\ \det(S)\ | \ . \qquad \text{(B.5.b.14)}$$

This volume expression has the same form as that for the 2-piped, (B.5.a.10).

(c) The 4-piped

Face Area: The faces of a 4-piped are 3-pipeds. For N=3, the 3-piped volume is stated in (B.5.b.14),

$$\text{volume}(3) = |\,\varepsilon_{abc}\,(\mathbf{e}_1)_a(\mathbf{e}_2)_b(\mathbf{e}_3)_c\,| \qquad \text{(B.5.c.1)}$$

where $\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3$ are 3D vectors. For the 3-piped which is "face 4" of the 4-piped -- a "near" face which touches the origin of the 4D skewed $\mathbf{e}_n$ coordinate system -- vectors $\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3$ are 4D vectors. The first 3 components of each of these 4D vectors are the same as the components of the 3D $\mathbf{e}_i$ vectors, while the 4th components are all 0. This is so because face 4 lies in a plane defined by this 4th component being 0. The volume(3) formula expressed in terms of these new 4D vectors is therefore $|\,\varepsilon_{abc4}\,(e_1)_a(e_2)_b(e_3)_c\,|$, where ε is now a 4D ε tensor. The conclusion is that

$$A_4 = |\,\varepsilon_{abc4}\,(e_1)_a(e_2)_b(e_3)_c\,| \qquad \text{(B.5.c.2)}$$

and this then is the scalar area of both face 4 and its partner face 4p, the far face. Similar arguments would then support these other area expressions

$$A_1 = |\,\varepsilon_{1abc}\,(e_2)_a(e_3)_b(e_4)_c\,|$$
$$A_2 = |\,\varepsilon_{a2bc}\,(e_3)_a(e_4)_b(e_1)_c\,|$$
$$A_3 = |\,\varepsilon_{ab3c}\,(e_4)_a(e_1)_b(e_2)_c\,| \quad . \qquad \text{(B.5.c.3)}$$

Since indices on ε can be swapped for free due to the absolute value, the non-summed index can be put first in all three cases and one may then conclude that

$A_1 = |\mathbf{e}_2 \text{ x } \mathbf{e}_3 \text{ x } \mathbf{e}_4|$ face 1 and face 1p
$A_2 = |\mathbf{e}_3 \text{ x } \mathbf{e}_4 \text{ x } \mathbf{e}_1|$ face 2 and face 2p
$A_3 = |\mathbf{e}_4 \text{ x } \mathbf{e}_1 \text{ x } \mathbf{e}_2|$ face 3 and face 3p
$A_4 = |\mathbf{e}_1 \text{ x } \mathbf{e}_2 \text{ x } \mathbf{e}_3|$ face 4 and face 4p . (B.5.c.4)

where, as discussed in (A.4.1) and (A.4.2),

$$\mathbf{Q} = \mathbf{A} \text{ x } \mathbf{B} \text{ x } \mathbf{C} \qquad \text{is defined by} \qquad Q_k = \varepsilon_{kabc}\, A_a B_b C_c \ . \qquad \text{(B.5.c.5)}$$

In (A.6.2) it was shown that $\mathbf{E}_1 = \det(R)\ \mathbf{e}_2 \text{ x } \mathbf{e}_3 \text{ x } \mathbf{e}_4$ so that $\mathbf{e}_2 \text{ x } \mathbf{e}_3 \text{ x } \mathbf{e}_4$ lines up with $\mathbf{E}_1$. Regardless of the sign of det(R), we define the *vector* areas to point in the $+\hat{\mathbf{E}}_n$ directions. Thus

$$\mathbf{A}_n = |\mathbf{e}_1 \text{ x } \mathbf{...} \text{ x } \mathbf{e}_3|\ \hat{\mathbf{E}}_n \qquad // \ \mathbf{e}_n \text{ missing} \qquad n = 1,2,3,4 \qquad \text{(B.5.c.6)}$$

where the $\mathbf{e}_i$ are reordered for free due to the absolute value signs. But (A.6.2) says

$$\mathbf{E}_n = \det(R)\,(-1)^{n-1}\ \mathbf{e}_1 \text{ x ... x } \mathbf{e}_4 \qquad // \ \mathbf{e}_n \text{ missing} \qquad \text{(B.5.c.7)}$$

so

$$|\mathbf{E}_n| = |\det(R)|\,|\,\mathbf{e}_1 \text{ x ... x } \mathbf{e}_4\,| \ . \qquad // \ \mathbf{e}_n \text{ missing} \qquad \text{(B.5.c.8)}$$

Thus,

$$\begin{aligned} \mathbf{A_n} &= |\mathbf{E_n}|\ \hat{\mathbf{E}}_\mathbf{n} / |\det(R)| = \ |\det(S)|\ \mathbf{E_n} \\ &= \ |\det(S)|\ \det(R)\ (-1)^{n-1}\mathbf{e_1} \text{ x } ... \text{ x } \mathbf{e_4} && \text{// } \mathbf{e_n} \text{ missing} \\ &= \sigma\ (-1)^{n-1}\mathbf{e_1} \text{ x } ... \text{ x } \mathbf{e_4} \qquad \text{// } \mathbf{e_n} \text{ missing} && \sigma \equiv \text{sign}[\det(S)] = \text{sign}[\det(R)] \qquad \text{(B.5.c.9)} \end{aligned}$$

To summarize, for N=4 one has

$$\begin{aligned} \mathbf{A_n} &= \ |\det(S)|\ \mathbf{E_n} \ = \ \sigma\ (-1)^{n-1}\mathbf{e_1} \text{ x } ... \text{ x } \mathbf{e_4} && \text{// } \mathbf{e_n} \text{ missing} \\ &= \sigma\ (-1)^{n-1}\ \Pi^{\mathbf{x}}{}_{\mathbf{i\neq n}}\ \mathbf{e_i} && \sigma \equiv \text{sign}[\det(S)] = \text{sign}[\det(R)] \qquad \text{(B.5.c.10)} \end{aligned}$$

This $\mathbf{A_n}$ expression has the same form as those of the 2-piped and the 3-piped found earlier.

Volume. The volume of a 4-piped (with each of the above $\mathbf{A_n}$ in turn treated as the base) is base times height, so

$$\text{volume}(4) = |\ \mathbf{A_1} \bullet \mathbf{e_1}\ | \ = \ |\ \mathbf{A_2} \bullet \mathbf{e_2}\ | = \ |\ \mathbf{A_3} \bullet \mathbf{e_3}\ | = \ |\ \mathbf{A_4} \bullet \mathbf{e_4}\ | \qquad \text{(B.5.c.11)}$$

or

$$\text{volume}(4) = |\ \mathbf{e_2} \text{ x } \mathbf{e_3} \text{ x } \mathbf{e_4} \ \bullet \mathbf{e_1}\ | \ = \ |\ \mathbf{e_3} \text{ x } \mathbf{e_4} \text{ x } \mathbf{e_1} \ \bullet \mathbf{e_2}\ | = \ |\ \mathbf{e_4} \text{ x } \mathbf{e_4} \text{ x } \mathbf{e_1} \ \bullet \mathbf{e_3}\ | = \ |\ \mathbf{e_1} \text{ x } \mathbf{e_4} \text{ x } \mathbf{e_2} \ \bullet \mathbf{e_4}\ | \ .$$

Using ε notation one can write the first case as

$$\begin{aligned} \mathbf{e_2} \text{ x } \mathbf{e_3} \text{ x } \mathbf{e_4} \ \bullet \mathbf{e_1} &= \ (e_1)_a\ \varepsilon_{abcd}(e_2)_b(e_3)_c(e_4)_d = \varepsilon_{abcd}(e_1)_a(e_2)_b(e_3)_c(e_4)_d \\ &= \det[\ \mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}, \mathbf{e_4}] \ = \det(S) \qquad \text{// see (3.2.7)} \qquad \text{(B.5.c.12)} \end{aligned}$$

so that

$$\text{volume}(4) \ = \ |\ \varepsilon_{abcd}(\mathbf{e_1})_a(\mathbf{e_2})_b(\mathbf{e_3})_c(\mathbf{e_4})_d\ | \ = |\ \det[\ \mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}, \mathbf{e_4}]\ | \ = |\ \det(S)\ | \qquad \text{(B.5.c.13)}$$

This volume expression has the same form as those of the 2-piped and the 3-piped.

(d) The N-piped

Face Area: The faces of an N-piped are (N-1)-pipeds. If there had been an (N-1)-piped section prior to this one, the volume formula there would have been

$$\text{volume}(N\text{-}1) = \ |\ \varepsilon_{abc...x}\ (\mathbf{e_1})_a(\mathbf{e_2})_b ...\ (\mathbf{e_{N-1}})_x\ | \qquad \text{(B.5.d.1)}$$

where $\mathbf{e_1},\mathbf{e_2},...\mathbf{e_{N-1}}$ were (N-1)D vectors (that is, 3D vectors if N=4). For the N-piped which is "face N" of the N-piped -- a "near" face which touches the origin of the ND skewed $\mathbf{e_n}$ coordinate system -- vectors $\mathbf{e_1},\mathbf{e_2},...\mathbf{e_{N-1}}$ are ND vectors. The first N-1 components of each of these ND vectors are the same as the components of the (N-1)D $\mathbf{e_i}$ vectors, while the Nth components are all 0. This is so because face N lies in a plane defined by this Nth component being 0. The volume(N-1) formula expressed in terms of these new ND vectors is therefore | $\varepsilon_{abc...xN}\, (\mathbf{e}_1)_a(\mathbf{e}_2)_b ... (\mathbf{e}_{N-1})_x$ |, where ε is now an ND ε tensor. The conclusion is that (compare this to the N=4 expression (B.5.c.2)),

$$A_N = | \varepsilon_{abc...xN}\, (\mathbf{e}_1)_a(\mathbf{e}_2)_b ... (\mathbf{e}_{N-1})_x | \qquad \text{(B.5.d.2)}$$

and this then is the scalar area of both face N and its partner face Np, the far face. Similar arguments would then support similar expressions for the other faces, for example,

$$A_1 = | \varepsilon_{1abc...x}\, (\mathbf{e}_2)_a(\mathbf{e}_3)_b ... (\mathbf{e}_{N-1})_w\, (\mathbf{e}_N)_x |$$

$$A_2 = | \varepsilon_{a2bc...x}\, (\mathbf{e}_3)_a(\mathbf{e}_4)_b ...\ (\mathbf{e}_N)_w\ (\mathbf{e}_1)_x | \ . \qquad \text{(B.5.d.3)}$$

Since indices on ε can be swapped for free due to the absolute value, the non-summed index can be put first in all three cases and one may then conclude that

$A_1 = |\mathbf{e}_2 \text{ x } \mathbf{e}_3 \text{ x } \mathbf{e}_4 \text{ x } \mathbf{e}_5 ... \text{ x } \mathbf{e}_N|$ face 1 and face 1p
$A_2 = |\mathbf{e}_3 \text{ x } \mathbf{e}_4 \text{ x } \mathbf{e}_5 \text{ x } \mathbf{e}_6 ... \text{ x } \mathbf{e}_1|$ face 2 and face 2p
$A_3 = |\mathbf{e}_4 \text{ x } \mathbf{e}_5 \text{ x } \mathbf{e}_6 \text{ x } \mathbf{e}_7 ... \text{ x } \mathbf{e}_2|$ face 3 and face 3p
...
$A_N = |\mathbf{e}_5 \text{ x } \mathbf{e}_6 \text{ x } \mathbf{e}_7 \text{ x } \mathbf{e}_8 .. \text{ x } \mathbf{e}_3|$ face N and face Np (B.5.d.4)

where, as discussed in (A.4.1) and (A.4.2),

$$\mathbf{Q} = \mathbf{A} \text{ x } \mathbf{B} \text{ x } \mathbf{C} .... \mathbf{X} \qquad \text{is defined by} \qquad Q_k = \varepsilon_{kabc...x}\, A_a B_b C_c .....X_x \ . \qquad \text{(B.5.d.5)}$$

In (A.6.2) it was shown that $\mathbf{E}_1 = \det(R)\, \mathbf{e}_2 \text{ x } \mathbf{e}_3 ... \mathbf{e}_N$ so that $\mathbf{e}_2 \text{ x } \mathbf{e}_3 ... \mathbf{e}_N$ lines up with $\mathbf{E}_1$. Regardless of the sign of det(R), we define the *vector* areas to point in the $+\hat{\mathbf{E}}_n$ directions. Thus

$$\mathbf{A}_n = |\mathbf{e}_1 \text{ x } ... \text{ x } \mathbf{e}_N|\, \hat{\mathbf{E}}_n \qquad // \ \mathbf{e}_n \text{ missing} \qquad n = 1,2,3...N \qquad \text{(B.5.d.6)}$$

where the $\mathbf{e}_i$ are reordered for free due to the absolute value signs. But (A.6.2) says

$$\mathbf{E}_n = \det(R)\, (-1)^{n-1}\ \mathbf{e}_1 \text{ x } ... \text{ x } \mathbf{e}_N \qquad // \ \mathbf{e}_n \text{ missing} \qquad \text{(B.5.d.7)}$$

so

$$|\mathbf{E}_n| = |\det(R)|\, |\, \mathbf{e}_1 \text{ x } ... \text{ x } \mathbf{e}_N\, | \qquad // \ \mathbf{e}_n \text{ missing} \qquad \text{(B.5.d.8)}$$

Thus,

$$\begin{aligned} \mathbf{A}_n &= |\mathbf{E}_n|\ \hat{\mathbf{E}}_n / |\det(R)| = \ |\det(S)|\, \mathbf{E}_n \\ &= \ |\det(S)|\, \det(R)\, (-1)^{n-1} \mathbf{e}_1 \text{ x } ... \text{ x } \mathbf{e}_N && // \ \mathbf{e}_n \text{ missing} \\ &= \sigma\, (-1)^{n-1} \mathbf{e}_1 \text{ x } ... \text{ x } \mathbf{e}_N \quad // \ \mathbf{e}_n \text{ missing} && \sigma \equiv \text{sign}[\det(S)] = \text{sign}[\det(R)] \qquad \text{(B.5.d.9)} \end{aligned}$$

To summarize, for N=N one has

$$\begin{aligned} \mathbf{A}_n = \ &|\det(S)|\, \mathbf{E}_n \ = \ \sigma\, (-1)^{n-1} \mathbf{e}_1 \text{ x } ... \text{ x } \mathbf{e}_N && // \ \mathbf{e}_n \text{ missing} \\ &= \sigma\, (-1)^{n-1}\, \Pi^{\mathbf{x}}{}_{i\neq n}\ \mathbf{e}_i && \sigma \equiv \text{sign}[\det(S)] = \text{sign}[\det(R)] \qquad \text{(B.5.d.10)} \end{aligned}$$

This A_n expression has the same form as those of the 2-piped, the 3-piped and the 4-piped.

Volume. The volume of a N-piped (with each of the above $\mathbf{A_n}$ in turn treated as the base) is base times height, so

$$\text{volume(N)} = | \mathbf{A_1} \bullet \mathbf{e_1} | = | \mathbf{A_2} \bullet \mathbf{e_2} | = \ ... \ = | \mathbf{A_N} \bullet \mathbf{e_N} |$$

or

$$\text{volume(N)} = | \mathbf{e_2} \text{ x } \mathbf{e_3} \text{ x } \mathbf{e_4}....\mathbf{e_N} \bullet \mathbf{e_1} | \ = ... \tag{B.5.d.11}$$

Using ε notation one can write the first case as

$$\mathbf{e_2} \text{ x } \mathbf{e_3} \text{ x } \mathbf{e_4}...\mathbf{e_N} \bullet \mathbf{e_1} = \quad (\mathbf{e_1})_a \, \varepsilon_{abc...x}(\mathbf{e_2})_b(\mathbf{e_3})_c.....(\mathbf{e_N})_x = \varepsilon_{abc...x}(\mathbf{e_1})_a(\mathbf{e_2})_b(\mathbf{e_3})_c....(\mathbf{e_N})_x$$

$$= \det [\, \mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}, ...\mathbf{e_N}] = \det(S) \qquad // \text{ see } (3.2.7) \tag{B.5.d.12}$$

where x is the N^{th} letter of the alphabet, so that

$$\text{volume(N)} = | \varepsilon_{abc...x}(\mathbf{e_1})_a(\mathbf{e_2})_b(\mathbf{e_3})_c....(\mathbf{e_N})_x | = | \det [\, \mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}, ...\mathbf{e_N}] | = | \det(S) | \ . \tag{B.5.d.13}$$

This volume expression has the same form as those of the 2-piped, the 3-piped and the 4-piped. This result is also consistent with the volume(N-1) expression stated in (B.5.d.1).

B.6 Summary of Main Results of this Appendix

1. One way to write the equation of a plane in N dimensions is

$$\mathbf{r} \bullet \hat{\mathbf{p}} = p \qquad \mathbf{r} = (x_1, x_2, ...x_n) \tag{B.6.1}$$

where $\hat{\mathbf{p}}$ is the unit vector normal to the plane which points "away from the origin", and where $p > 0$ is the distance of closest approach of the plane to the origin. In the limit $p \to 0$, the plane passes through the origin and the equation is then $\mathbf{r} \bullet \hat{\mathbf{p}} = 0$ where $\hat{\mathbf{p}}$ is either normal to the plane.

2. An N-piped has 2^N vertices as shown by the inductive construction method presented above. (B.6.2)

3. The locus of points making up the (closed) interior of an N-piped spanned by $\mathbf{e_1}...\mathbf{e_N}$ is given by

$$\mathbf{r}_{\text{volumeN}} = \Sigma_{n=1}{}^N \alpha_n \mathbf{e_n} \qquad 0 \le \alpha_n \le 1 \tag{B.6.3}$$

The tails of all the vectors $\mathbf{e_1}...\mathbf{e_N}$ meet at the origin of R^N space.

4. There are N pairs of faces on an N-piped, and each face is an (N-1)-piped having 2^{N-1} vertices. The total face count is 2N. Each face is spanned by a subset of N-1 of the base vectors $\mathbf{e_n}$, so each face is

"missing" one of the $\mathbf{e_n}$ and the face is labeled using the index of this missing base vector. The loci of points making up the faces of an N-piped are given by

$$\mathbf{r}_{\mathtt{face(i)}} = \Sigma_{n\neq i}\alpha_n\mathbf{e}_n \qquad 0 \le \alpha_n \le 1 \qquad i = 1,2...N$$
$$\mathbf{r}_{\mathtt{face(ip)}} = \Sigma_{n\neq i}\alpha_n\mathbf{e}_n + \mathbf{e}_i \qquad 0 \le \alpha_n \le 1 \qquad i = 1,2...N \qquad \text{(B.6.4)}$$

where "face i" has a corner touching the origin of the N-piped (near face), while its parallel partner face "ip" does not touch the origin (far face).

5. If the N-piped spanning vectors $\mathbf{e_n}$ are the tangent base vectors associated with some transformation F, then $\mathbf{E_i}\bullet \mathbf{e_j} = \delta_{i,j}$ where $\mathbf{E_i}$ are the reciprocal base vectors. In this case, the equations of the planes of the faces of the N-piped can be written in the form shown in item 1 above,

$$\hat{\mathbf{E}}_i\bullet \mathbf{r}_{\mathtt{(face\ i)}} = 0 \qquad \hat{\mathbf{E}}_i\bullet \mathbf{r}_{\mathtt{(face\ ip)}} = 1/E_i \qquad i = 1,2...N \qquad \text{(B.6.5)}$$

so that both faces of a pair i are planar (in N dimensional space) and they have the same normal vector $\hat{\mathbf{E}}_i$ so the faces of a pair lie on parallel planes.

6. The vector $\mathbf{E_i}$ is an outward-facing normal for face ip, while $-\mathbf{E_i}$ is an outward-facing normal vector for face i (which touches the origin). (B.6.6)

7. The out-facing vector area of face ip of an N-piped can be expressed as

$$\mathbf{A}_i = |\det(S)|\ \mathbf{E}_i$$

$$\mathbf{A}_i = \sigma\ (-1)^{i-1}\ \Pi^{x}{}_{j\neq i}\ \mathbf{e}_j \qquad \sigma \equiv \text{sign}[\det(S)] = \text{sign}[\det(R)]$$

$$\mathbf{A}_i = \sigma\ (-1)^{i-1}\ \mathbf{e}_1 \times \mathbf{e}_2 \ldots \times \mathbf{e}_N \qquad //\ \mathbf{e}_i \text{ missing} \qquad \text{(B.6.7)}$$

where $\mathbf{e_i}$ is the vector missing from the face's spanning set. The outfacing area for face i is - $\mathbf{A_i}$. The last two lines are shorthands for the following, as discussed in (A.5.2),

$$(\mathbf{A}_i)_\alpha = \sigma\ (-1)^{i-1}\ \varepsilon_{\alpha abc..x}\ (\mathbf{e}_1)_a\ (\mathbf{e}_2)_b \ldots\ (\mathbf{e}_N)_x \qquad //\ \text{where } (\mathbf{e}_i)_i \text{ and i are missing} \qquad \text{(B.6.8)}$$

For N=2 the last two expressions for $\mathbf{A_i}$ are interpreted as shown in (B.5.a.9),

$$\mathbf{A}_i = \sigma\ (-1)^{i-1}\ \mathbf{e}_{3-i} \times \hat{\mathbf{3}} \qquad i = 1,2 \qquad \text{(B.6.9)}$$

8. The volume of an N-piped spanned by $\mathbf{e_1}...\mathbf{e_N}$ is given by

$$\text{volume(N)} = |\ \det(S)\ | = |\ \det\ [\ \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \ldots\mathbf{e}_N]\ | = |\ \varepsilon_{abc...x}(\mathbf{e}_1)_a(\mathbf{e}_2)_b(\mathbf{e}_3)_c\ldots(\mathbf{e}_N)_x\ | \qquad \text{(B.6.10)}$$

where one can regard the tangent base vectors as the columns of the linearized transformation matrix S as shown in (3.2.7).

Appendix C: Elliptical Polar Coords, Views of x'-space, Jacobian Integration Rule

This Appendix is written in the developmental notation of Chapters 1-6.

C.1 Elliptical polar coordinates

The 2D "elliptic" coordinate system has coordinate lines which are orthogonal ellipses and hyperbolas. When rotated about its two symmetry axes, this system generates 3D prolate or oblate spheroidal coordinates. This is *not* the 2D coordinate system described in this Appendix. For "elliptical polar" coordinates, the coordinate lines are taken instead as the ellipses from elliptic coordinates, and the rays from polar coordinates. This non-orthogonal system is perhaps not very useful, but provides a good "sandbox" in which to study general aspects of coordinate systems.

The transformation **x**' = **F**(**x**) is given by

$$\begin{array}{llll} & & \underline{\text{x'-space}} & \underline{\text{x-space (Cartesian)}} \\ \rho^2 = x^2/a^2 + y^2/b^2 & x^2 + y^2 = r^2 \text{ still} & x'_1 = \theta & x_1 = x \\ \tan\theta = y/x & & x'_2 = \rho & x_2 = y \end{array} \qquad (C.1.1)$$

Writing the first equation above as

$$1 = x^2/(\rho a)^2 + y^2/(\rho b)^2 \qquad (C.1.2)$$

it should be clear that ρ serves to label an ellipse of semi-major axis ρa, and semi-minor axis ρb, while θ labels the ray at angle θ, as in polar coordinates. The inverse transform **x** = $\mathbf{F}^{-1}$(**x**') is given by

$$\begin{array}{llll} x = a\rho\cos\theta & x/a = \rho\cos\theta & \Rightarrow & x^2/a^2 + y^2/b^2 = \rho^2 \\ y = b\rho\sin\theta & y/b = \rho\sin\theta & \Rightarrow & \tan\theta = y/x \ . \end{array} \qquad (C.1.3)$$

The matrix S is given by

$$\begin{array}{ll} S_{11} = (\partial x/\partial\theta) = -a\rho\sin\theta & \\ S_{12} = (\partial x/\partial\rho) = a\cos\theta & S_{ik} \equiv (\partial x_i/\partial x'_k) \quad \textit{(2.1.5)} \\ S_{21} = (\partial y/\partial\theta) = b\rho\cos\theta & \\ S_{22} = (\partial y/\partial\rho) = b\sin\theta & \end{array} \qquad (C.1.4)$$

$$S = \begin{pmatrix} -a\rho\sin\theta & a\cos\theta \\ b\rho\cos\theta & b\sin\theta \end{pmatrix} \quad \Rightarrow \quad \det(S) = -ab\rho \quad \text{and} \quad R = S^{-1} = \begin{pmatrix} -\sin\theta/(a\rho) & \cos\theta/(b\rho) \\ \cos\theta/a & \sin\theta/b \end{pmatrix} \ .$$

The tangent base vectors $\mathbf{e_n}$ can be read off as the columns of S, see (3.2.7) :

$$\begin{array}{llll} \mathbf{e_1} = \rho(-a\sin\theta, b\cos\theta) & = \mathbf{e_\theta} & |\mathbf{e_\theta}| = \rho\sqrt{a^2\sin^2\theta + b^2\cos^2\theta} \equiv h_\theta & \mathbf{e_\theta} = |\mathbf{e_\theta}|\ \hat{\mathbf{e}}_\theta \\ \mathbf{e_2} = (a\cos\theta, b\sin\theta) & = \mathbf{e_\rho} & |\mathbf{e_\rho}| = \sqrt{a^2\cos^2\theta + b^2\sin^2\theta} \equiv h_\rho & \mathbf{e_\rho} = |\mathbf{e_\rho}|\ \hat{\mathbf{e}}_\rho \ . \end{array} \qquad (C.1.5)$$

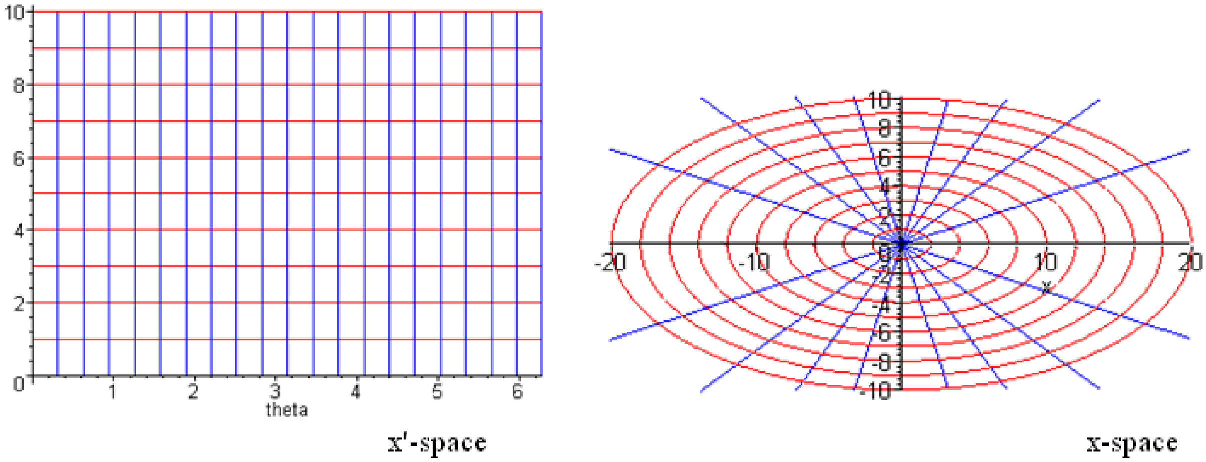
The covariant metric tensor from (5.7.9) or (5.11.3) is,

$$\overline{g}' = S^TS \;= \begin{pmatrix} \rho^2\{a^2\sin^2(\theta) + b^2\cos^2(\theta)\} & [b^2\text{-}a^2]\rho\,\sin(\theta)\cos(\theta) \\ [b^2\text{-}a^2]\rho\,\sin(\theta)\cos(\theta) & a^2\cos^2(\theta) + b^2\sin^2(\theta) \end{pmatrix} = \begin{pmatrix} e_1\bullet e_1 & e_1\bullet e_2 \\ e_2\bullet e_1 & e_2\bullet e_2 \end{pmatrix} \qquad \text{(C.1.6)}$$

which is clearly non-diagonal (but symmetric) as expected. When a = b = 1 it reduces to the polar coordinates system metric tensor where then $\rho = r$. The coordinate system is non-orthogonal because $e_1\bullet e_2 \neq 0$, or equivalently, because $\overline{g}'$ is non-diagonal.

C.2 Forward coordinate lines

Here is a Maple plot of some x-space forward coordinate lines (parameters a = 2 and b = 1)



—

(C.2.1)

where ρ is the x'-space vertical axis. The coordinate lines in x-space are plotted using these equations,

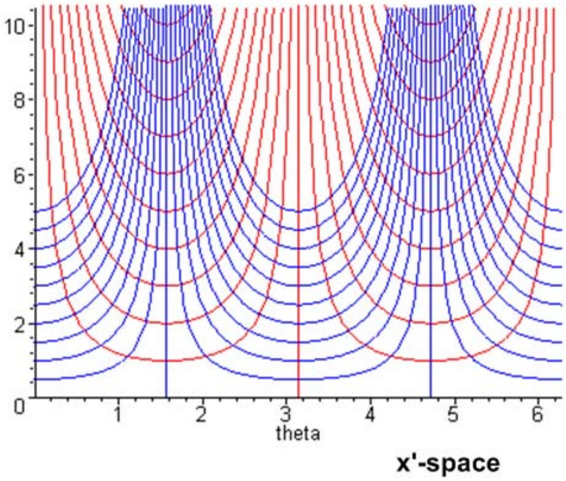
$$\begin{array}{llll} y = b\sqrt{\rho_i^{\,2}\text{-}(x/a)^2} & \text{// ellipses} & \rho_i = 1,2...10 & \text{10 ellipses} \\ y = x\,\tan\theta_i & \text{// rays} & \theta_i = 2\pi\,(i/20)\;,\; i = 1,2...20 & \text{20 rays} \end{array} \qquad \text{(C.2.2)}$$

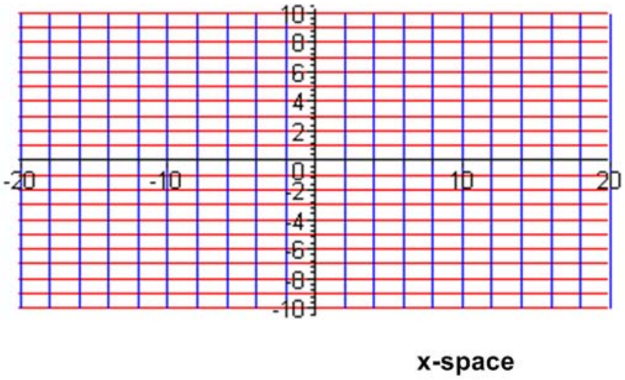
which are obtained from the forward transformation equations

$$\begin{aligned} \rho^2 &= x^2/a^2 + y^2/b^2 \\ \tan\theta &= y/x\;\;. \end{aligned} \qquad \textit{(C.1.1)}$$

C.3 Inverse coordinate lines

Here is a Maple plot of some x'-space inverse coordinate lines (parameters a = 2 and b = 1)





The coordinate lines in x'-space are plotted using these equations (C.3.1)

$\rho = x_i/(a\cos\theta)$ x_i = -10 to +10 20 blue curves
$\rho = y_i/(b\sin\theta)$ y_i = -10 to +10 20 red curves (C.3.2)

which are obtained from the inverse transformation equations

$$x = a\rho\cos\theta$$
$$y = b\rho\sin\theta \qquad \textit{(C.1.3)}$$

The secθ and cscθ curve families appear to "change shape", but that is just what happens when functions are scaled up vertically but not horizontally. If one plots one sine hump at different vertical scalings, the humps have different shapes.

C.4 Drawing a contravariant vector V in x-space: the meaning of V'_n .

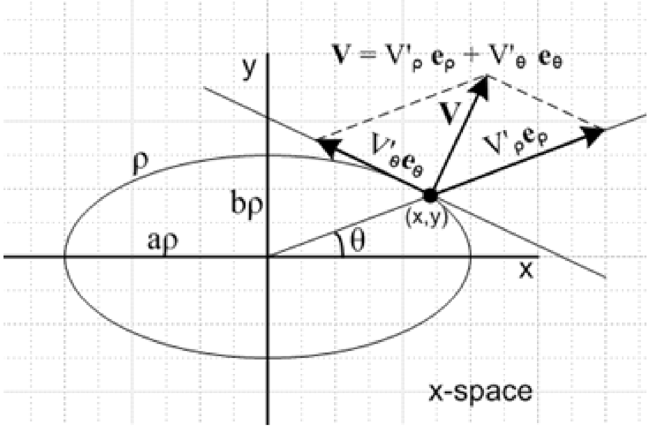
A contravariant vector field **V(x)** can be expanded in these two ways, as shown in (6.6.9),

$$\mathbf{V} = V_1(\mathbf{x})\,\hat{\mathbf{1}} + V_2(\mathbf{x})\,\hat{\mathbf{2}} = V_x(\mathbf{x})\,\hat{\mathbf{x}} + V_y(\mathbf{x})\,\hat{\mathbf{y}} \qquad // \ \mathbf{u}_n = \hat{\mathbf{n}} \text{ for Cartesian}$$

$$\mathbf{V} = V'_1(\mathbf{x}')\,\mathbf{e}_1 + V'_2(\mathbf{x}')\,\mathbf{e}_2 = V'_\theta(\mathbf{x}')\,\mathbf{e}_\theta + V'_\rho(\mathbf{x}')\,\mathbf{e}_\rho \qquad \mathbf{V}'(\mathbf{x}') = R(x)\,\mathbf{V}(\mathbf{x}) \qquad \text{(C.4.1)}$$

where the V'_n are the components of **V** transformed into x'-space where **V** becomes **V**'. The prime is not necessary on V'_ρ but we maintain it as a reminder that it is an x'-space component. R(**x**) is the matrix of (2.1.6) and $\mathbf{e}_n$ are the tangent base vectors of Chapter 3. The fields $V'_n(\mathbf{x}')$ are "the components of vector field **V**' in x'-space", since **V**' = R**V** , or $V'_i = R_{ij}V_j$. Moreover, these $V'_n(\mathbf{x}')$ are "expressed in terms of curvilinear coordinates" **x**'. If one is asked to "express a vector **V** in curvilinear coordinates", one is usually being asked to write **V** as the second expansion above. The vectors $\mathbf{e}_n$ and **V** exist *in x-space*, and in the second expansion it just happens that the coefficients $V'_n(\mathbf{x}')$ are the components of **V**', the transformed vector in x'-space, when it is expanded on the axis-aligned vectors $\mathbf{e}'_n$ in x'-space.

Here is a graphical representation of this vector **V** in x-space:



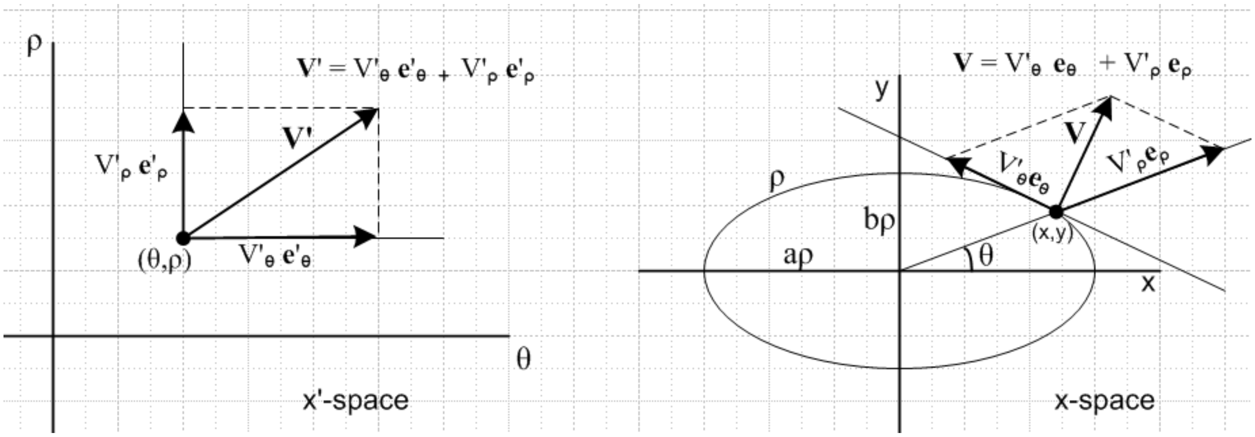
(C.4.2)

As advertised, the tangent base vectors are not at right angles. The **V** parallelogram accurately illustrates the equation $\mathbf{V} = V'_\theta\ \mathbf{e}_\theta + V'_\rho\ \mathbf{e}_\rho$. Since x-space is Cartesian, there is no distinction between Cartesian length (graphical length) and covariant length for vectors in x-space. Graphically, one could find the values for V'_θ and V'_ρ as follows: (1) for the point (x,y), compute the vectors $\mathbf{e}_\theta$ and $\mathbf{e}_\rho$ and compute their lengths $|\mathbf{e}_\theta| = h'_\theta$ and $|\mathbf{e}_\rho| = h'_\rho$; (2) draw the parallelogram shown aligned with these vectors for some given **V** and find the edge lengths. The edges of the parallelogram are $V'_\theta\ h'_\theta$ and $V'_\rho\ h'_\rho$ so then the values of V'_θ and V'_ρ can be found.

The alternative method is to compute R and use $V'_i = R_{ij}V_j$.

C.5 Drawing a contravariant vector V' in x'-space: two "Views"

As with previous examples, the drawing (C.4.2) is drawn to the right of an x'-space drawing as follows:



(C.5.1)

In Chapter 3 the axis-aligned basis vectors $\mathbf{e}'_n$ were introduced as

$\mathbf{e}'_n$, n = 1,2...N // $(\mathbf{e}'_n)_i = \delta_{n,i}$ $\mathbf{e}'_1 = (1,0,0...)$ etc *(3.2.1)*

and $\mathbf{e}_n$ was shown to be a contravariant vector,

$$\mathbf{e}'_n = R(\mathbf{x})\, \mathbf{e}_n \ . \qquad (3.3.2)$$

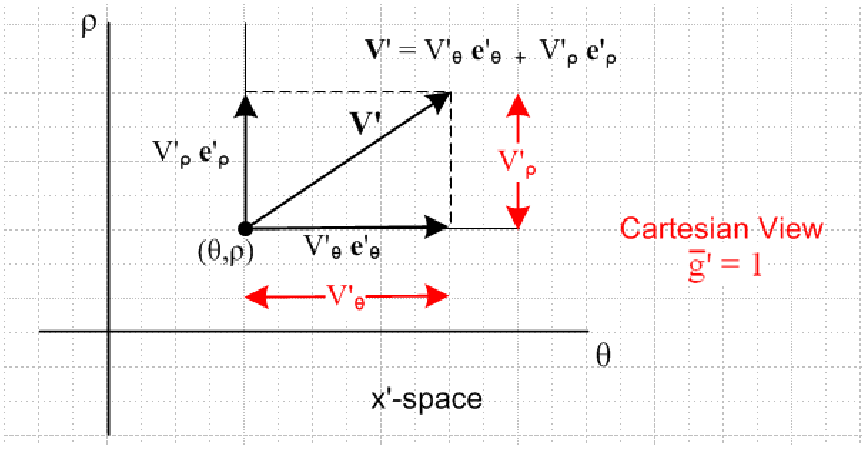
Applying matrix R(x) to the equation $\mathbf{V} = V'_\theta\ \mathbf{e}_\theta + V'_\rho\ \mathbf{e}_\rho$ one gets the expansion shown in Fig (C.5.1),

$$\mathbf{V}' = V'_\theta\ \mathbf{e}'_\theta\ + V'_\rho\ \mathbf{e}'_\rho \qquad (6.6.15) \qquad\qquad (C.5.2)$$

which appears first in the list of expansions of **V**' in (6.6.15). There is no ambiguity concerning this last equation. Ambiguity can arise, however, when one tries to represent this equation graphically in x'-space.

There are two very different "views" one can take of a drawing in x'-space. In the first view, we take x'-space to be "flat" (Cartesian) so that g' = 1. In the second view, we take x'-space to be "curved" with g' ≠ 1. These views are really two *different* x'-spaces since the metric tensors are different.

Cartesian View of x'-space



(C.5.3)

In this view of x'-space the length (norm) of a vector is given by $|\mathbf{A}|^2 = \delta_{ij}A_iA_j = \Sigma A_i^{\ 2}$, so one has, since $(\mathbf{e}'_n)_i = \delta_{n,i}$,

$$\bar{g}' = 1 \qquad |\mathbf{e}'_n| = 1 \qquad \mathbf{e}'_n = \hat{\mathbf{e}}'_n = \hat{\mathbf{n}}' = \text{the usual axis-aligned unit vectors in x'-space}$$

$$\mathbf{V}' = V'_\theta\ \hat{\boldsymbol{\theta}}\ + V'_\rho\ \hat{\boldsymbol{\rho}} \qquad \hat{\boldsymbol{\theta}} = \hat{\mathbf{e}}'_1 \qquad \hat{\boldsymbol{\rho}} = \hat{\mathbf{e}}'_2\ . \qquad (C.5.4)$$

The left-side of (C.5.1) is, in this view, a "normal Cartesian graph" and the vectors add up properly. For example Pythagoras tells us that

$$|\mathbf{V}'|^2 = V'^{\,2}_\theta + V'^{\,2}_\rho$$

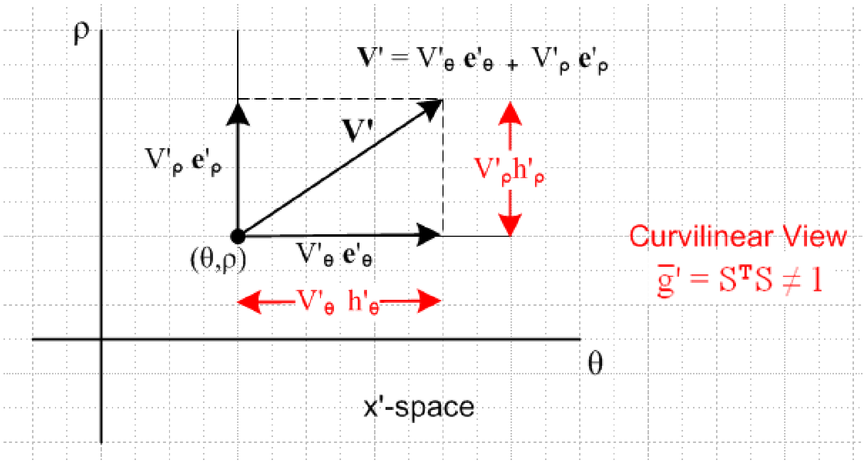
and

$$\hat{\boldsymbol{\theta}} \bullet \hat{\boldsymbol{\theta}}\ = \mathbf{1}$$
$$\hat{\boldsymbol{\rho}} \bullet \hat{\boldsymbol{\rho}}\ = \mathbf{1}$$
$$\hat{\boldsymbol{\theta}} \bullet \hat{\boldsymbol{\rho}} = 0\ . \qquad (C.5.5)$$

This Cartesian View, which is x'-space with $\bar{g}' = 1$, is appropriate in applications in which it is not required or desired that norms and dot products be tensorial scalars, as discussed at the end of Section 5.2. For example, when $\bar{g}'$ is set to 1, one has $|\mathbf{V}'| \neq |\mathbf{V}|$ in (C.5.1).

Another use of this view involves integration as will be seen below.

Curvilinear View of x'-space



(C.5.6)

In this view of x'-space, one assumes that $\bar{g}'$ takes a value which enforces the scalarity of norms and dot products between x-space and x'-space, which is to say, one takes $\bar{g}' = S^T \bar{g}\, S$ from (5.7.6) where $\bar{g}$ is the x-space metric tensor for x-space. Normally $\bar{g}$=1 (Cartesian x-space), so $\bar{g}' = S^T S$. In this Curvilinear View, then, the length $|\mathbf{A}'|$ of a contravariant vector $\mathbf{A}'$ is determined by $|\mathbf{A}'|^2 = \bar{g}'_{ij} A'_i A'_j$ where

$$\bar{g}' = S^T S \neq 1 \qquad |\mathbf{e}'_n| = |\mathbf{e}_n| = h'_n \equiv \sqrt{\bar{g}'_{nn}} \qquad n = 1,2 \text{ for } \theta,\rho$$

(C.5.7)

$$\mathbf{V}' = V'_\theta \mathbf{e}'_\theta + V'_\rho \mathbf{e}'_\rho = (V'_\theta h'_\theta)\, \hat{\mathbf{e}}'_\theta + (V'_\rho h'_\rho)\, \hat{\mathbf{e}}'_\rho \qquad \hat{\mathbf{e}}'_n \equiv \mathbf{e}'_n / |\mathbf{e}_n| = \mathbf{e}'_n / h'_n$$

$$V_x^2 + V_y^2 = |\mathbf{V}|^2 = |\mathbf{V}'|^2 \neq (V'_\theta h'_\theta)^2 + (V'_\rho h'_\rho)^2 \qquad // \text{ unless } x'_i \text{ are orthogonal coordinates}$$

This last inequality says that in the Curvilinear View the Pythagorean Theorem is invalid. In fact

$$|\mathbf{V}'|^2 = \bar{g}'_{ij} V'_i V'_j = \bar{g}'_{\theta\theta} V'^2_\theta + \bar{g}'_{\rho\rho} V'^2_\rho + 2\, \bar{g}'_{\theta\rho} V'_\theta V'_\rho$$

$$= (V'_\theta h'_\theta)^2 + (V'_\rho h'_\rho)^2 + 2\, \bar{g}'_{\theta\rho} V'_\theta V'_\rho \ . \qquad \text{(C.5.8)}$$

In writing $|\mathbf{e}'_n| = |\mathbf{e}_n|$ and $|\mathbf{V}'|^2 = |\mathbf{V}|^2$ above, we use the rule shown in (5.10.4) which says $|\mathbf{A}'|^2 = |\mathbf{A}|^2$ for any contravariant vector $\mathbf{A}$ ($|\mathbf{A}|^2$ is a scalar). Moreover,

$$\hat{\mathbf{e}}'_\theta \bullet \hat{\mathbf{e}}'_\theta = \mathbf{e}'_\theta \bullet \mathbf{e}'_\theta \ / (h'^2_\theta) = \mathbf{e}_\theta \bullet \mathbf{e}_\theta \ / (h'^2_\theta) = \overline{g}'_{\theta\theta} / (h'^2_\theta) \quad = 1$$
$$\hat{\mathbf{e}}'_\rho \bullet \hat{\mathbf{e}}'_\rho = \mathbf{e}'_\rho \bullet \mathbf{e}'_\rho \ / (h'^2_\rho) = \mathbf{e}_\rho \bullet \mathbf{e}_\rho \ / (h'^2_\rho) = \overline{g}'_{\rho\rho} / (h'^2_\rho) \quad = 1$$
$$\hat{\mathbf{e}}'_\theta \bullet \hat{\mathbf{e}}'_\rho = \mathbf{e}'_\theta \bullet \mathbf{e}'_\rho \ / (h'_\theta h'_\rho) = \mathbf{e}_\theta \bullet \mathbf{e}_\rho \ / (h'_\theta h'_\rho) = \overline{g}'_{\theta\rho} / (h'_\theta h'_\rho) \quad \neq 0 \qquad <= \ !! \qquad (C.5.9)$$

so the $\hat{\mathbf{e}}'_n$ are unit vectors having unit covariant length, but $\hat{\mathbf{e}}'_\theta \bullet \hat{\mathbf{e}}'_\rho \neq 0$ despite the fact that these vectors are drawn at right angles in the x'-space graph above, $\mathbf{e}_n = h'_n \hat{\mathbf{e}}'_n$. One might imagine trying to slant the lines in (C.5.6) to cause all intersection points to have angles which match the metric tensor, which is to say, at each intersection point one would need an angle ψ where $\hat{\mathbf{e}}'_\theta \bullet \hat{\mathbf{e}}'_\rho = \cos\psi$. But in general $\hat{\mathbf{e}}'_n \bullet \hat{\mathbf{e}}'_m = \overline{g}'_{nm} / (h'_n h'_m)$ has a different value at every point, so such a graph would be quite complex.

The upshot is that for a *non-orthogonal* system, the axes in x'-space are still drawn at right angles and the purpose of the graph is mainly to "locate" all the points **x**' which correspond to points **x** in x-space according to **x**' = **F**(**x**). The graph does successfully represent the idea that $\mathbf{V}' = V'_\rho \mathbf{e}'_\rho + V'_\theta \mathbf{e}'_\theta$, but one must give up on Euclidean geometry for this vector sum triangle. It might be imagined that the x'-space graph is the projection onto the plane of paper of some vectors drawn on a curved surface emerging from the plane of paper, and that is then why Pythagoras is wrong.

In the case of an *orthogonal* coordinate system (diagonal $\overline{g}'$), the 90 degree angles between the axes in x'-space are accurate representations of the fact that $\hat{\mathbf{e}}'_n \bullet \hat{\mathbf{e}}'_m = 0$ when $n \neq m$. And since scalars are preserved, one has in the Curvilinear View,

$$|\mathbf{V}'|^2 = \Sigma_n (h'_n V'_n)^2 = \Sigma_n \mathcal{V}'^2_n = |\mathbf{V}|^2 = \Sigma_n V_n^2 \qquad // \text{ orthogonal only}$$

where

$$\mathcal{V}'_n \equiv h'_n V'_n \quad \text{and} \quad \mathbf{V}' = \Sigma_n \mathcal{V}'_n \hat{\mathbf{e}}'_n$$
$$\text{and} \quad \mathbf{V} = \Sigma_n \mathcal{V}'_n \hat{\mathbf{e}}_n \qquad (C.5.10)$$

One can then still apply regular Euclidean geometry to the vector addition N-piped in x'-space in the sense that $|\mathbf{V}'|^2 = \Sigma_n (h'_n V'_n)^2$.

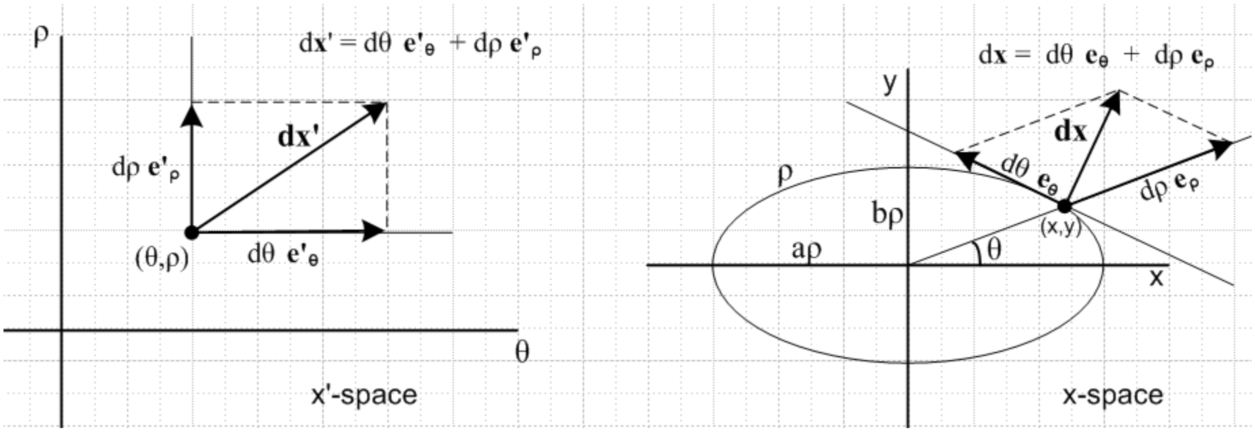
Although the x'-space perpendicular-unit-vectors and Pythagorean paradoxes go away if the x' coordinates are orthogonal (as they would be for polar coordinates where a = b = 1), one still has the two x'-space views to keep in mind: the Cartesian View with $\overline{g}' = 1$ and the Curvilinear View with $\overline{g}' = S^T S$.

C.6 Drawing the specific contravariant vector dx in x-space and x'-space

Since d**x** is the primordial contravariant vector, everything stated in the last two Sections applies with $\mathbf{V} \to d\mathbf{x}$ and $V'_\theta \to dx'_\theta = d\theta$, $V'_\rho \to dx'_\rho = d\rho$, where we finally drop the primes on dθ and dρ. The expansions of d**x** and d**x**' are,

$$d\mathbf{x} = d\theta \ \mathbf{e}_\theta + d\rho \ \mathbf{e}_\rho \qquad // \text{ in x-space}$$
$$d\mathbf{x}' = d\theta \ \mathbf{e}'_\theta + d\rho \ \mathbf{e}'_\rho \qquad // \text{ in x'-space} \qquad (C.6.1)$$

For **V** = d**x** Fig (C.5.1) becomes,



(C.6.2)

It must be understood that now the vector arrows like d**x** are highly magnified and in reality are very small compared to, say, the curvature of the ellipse. From above,

$$\mathbf{e'}_\theta \bullet \mathbf{e'}_\theta = \bar{g}'_{\theta\theta} \qquad \hat{\mathbf{e}}'_\theta \bullet \hat{\mathbf{e}}'_\theta = 1$$
$$\mathbf{e'}_\rho \bullet \mathbf{e'}_\rho = \bar{g}'_{\rho\rho} \qquad \hat{\mathbf{e}}'_\rho \bullet \hat{\mathbf{e}}'_\rho = 1$$
$$\mathbf{e'}_\rho \bullet \mathbf{e'}_\theta = \bar{g}'_{\rho\theta} \qquad \hat{\mathbf{e}}'_\theta \bullet \hat{\mathbf{e}}'_\rho = \bar{g}'_{\theta\rho} / (h'_\theta h'_\rho) \qquad (C.6.3)$$

and once again the "right angle" in the x'-space picture is deceptive.

The x'-space side of (C.6.2) is subject to the two "views" described above:

Cartesian View of x'-space: (elliptical polar coordinates)

$$\bar{g}' = 1 \qquad |\mathbf{e'}_n| = 1$$
$$\mathbf{dx}' = d\theta\, \hat{\boldsymbol{\theta}} + d\rho\, \hat{\boldsymbol{\rho}} \qquad \hat{\boldsymbol{\theta}} = \hat{\mathbf{e}}'_1 \qquad \hat{\boldsymbol{\rho}} = \hat{\mathbf{e}}'_2$$
$$\hat{\boldsymbol{\theta}} \bullet \hat{\boldsymbol{\theta}} = \mathbf{1} \qquad \hat{\boldsymbol{\rho}} \bullet \hat{\boldsymbol{\rho}} = \mathbf{1} \qquad \hat{\boldsymbol{\theta}} \bullet \hat{\boldsymbol{\rho}} = 0 \quad . \qquad (C.6.4)$$

Curvilinear View of x'-space: (elliptical polar coordinates)

$$\bar{g}' = S^T S \qquad |\mathbf{e'}_n| = |\mathbf{e}_n| = h'_n \equiv \sqrt{\bar{g}'_{nn}}$$

$$\mathbf{dx}' = d\theta\, \mathbf{e'}_\theta + d\rho\, \mathbf{e'}_\rho = (d\theta\ h'_\theta)\, \hat{\mathbf{e}}'_\theta + (d\rho\ h'_\rho)\, \hat{\mathbf{e}}'_\rho$$

$$(dx)^2 + (dy)^2 = |\mathbf{dx}|^2 = |\mathbf{dx}'|^2$$

$$= (d\theta\ h'_\theta)^2 + (d\rho h'_\rho)^2 + 2\, \bar{g}'_{\theta\rho} d\theta d\rho$$
$$= (d\theta\ h'_\theta)^2 + (d\rho\ h'_\rho)^2 \quad \text{only if } \bar{g}'_{\theta\rho} = 0 \text{ (orthogonal)} \ . \qquad (C.6.5)$$

Just to have a specific orthogonal coordinates example, we reduce (C.1.3) to polar coordinates by setting a = b = 1. In this case, (C.1.6) reduces to $\overline{g}' = \begin{pmatrix} \rho^2 & 0 \\ 0 & 1 \end{pmatrix}$ as in (5.13.11) so $h_\theta = \rho$, $h_\rho = 1$ and $\overline{g}'_{\theta\rho} = 0$. We can then restate the two Views in this case

Cartesian View of x'-space: (polar coordinates)

$\overline{g}' = 1 \qquad |\mathbf{e}'_n| = 1$

$\mathbf{dx}' = d\theta\,\hat{\boldsymbol{\theta}} + d\rho\,\hat{\boldsymbol{\rho}} \qquad \hat{\boldsymbol{\theta}} = \hat{\mathbf{e}}'_1 \qquad \hat{\boldsymbol{\rho}} = \hat{\mathbf{e}}'_2$

$\hat{\boldsymbol{\theta}} \bullet \hat{\boldsymbol{\theta}} = \mathbf{1} \qquad \hat{\boldsymbol{\rho}} \bullet \hat{\boldsymbol{\rho}} = \mathbf{1} \qquad \hat{\boldsymbol{\theta}} \bullet \hat{\boldsymbol{\rho}} = 0$.

$$(ds)^2 = |\mathbf{dx}|^2 = d\mathbf{x} \bullet d\mathbf{x} = (dx)^2 + (dy)^2 \qquad // \text{ from x-space (which has } \overline{g} = 1\text{)}$$

$$(ds')^2 = |\mathbf{dx'}|^2 = d\mathbf{x'} \bullet d\mathbf{x'} = (d\theta)^2 + (d\rho)^2 \quad \neq \quad |\mathbf{dx}|^2 = (ds)^2 \qquad \text{(C.6.6)}$$

Curvilinear View of x'-space: (polar coordinates)

$\overline{g}' = S^T S = \begin{pmatrix} \rho^2 & 0 \\ 0 & 1 \end{pmatrix} \qquad |\mathbf{e}'_\theta| = |\mathbf{e}_\theta| = h'_\theta = \rho \equiv \sqrt{\overline{g}'_{\theta\theta}}$

$|\mathbf{e}'_\rho| = |\mathbf{e}_\rho| = h'_\rho = 1 \equiv \sqrt{\overline{g}'_{\rho\rho}}$

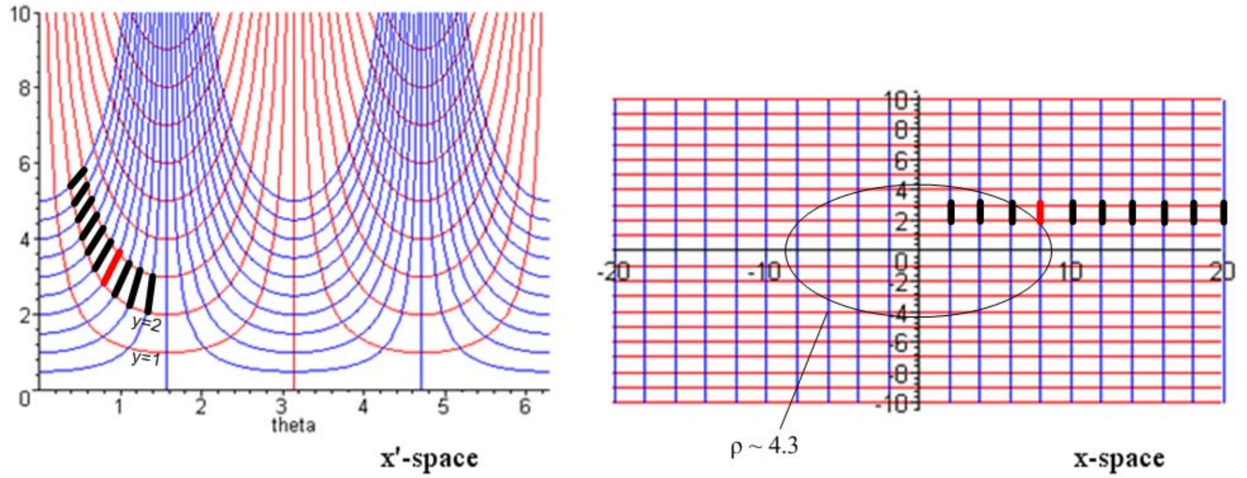
$$\mathbf{dx}' = d\theta\,\mathbf{e}'_\theta + d\rho\,\mathbf{e}'_\rho = (d\theta\,\rho)\,\hat{\mathbf{e}}'_\theta + (d\rho)\,\hat{\mathbf{e}}'_\rho$$

$$(ds)^2 = |\mathbf{dx}|^2 = d\mathbf{x} \bullet d\mathbf{x} = (dx)^2 + (dy)^2 \qquad // \text{ from x-space (which has } \overline{g} = 1\text{)}$$

$$(ds')^2 = |\mathbf{dx'}|^2 = d\mathbf{x'} \bullet d\mathbf{x'} = (d\theta\,\rho)^2 + (d\rho)^2 \quad = \quad |\mathbf{dx}|^2 = (ds)^2 \qquad \text{(C.6.7)}$$

C.7 Study of how dx transforms in the mapping between x-space and x'-space

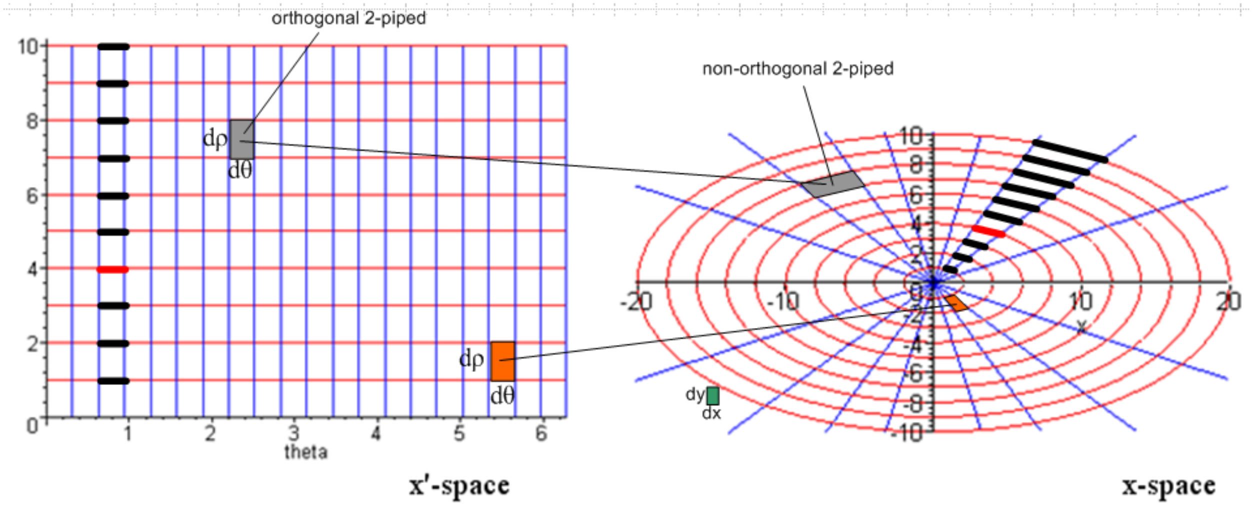
Consider this drawing which shows a representative set of vectors d**x** in x-space (the bars), along with the forward mappings (d**x**' = **F**(d**x**) or d**x**' = Rd**x** **)** of the corresponding vectors d**x**' in x'-space. The vectors on the right all point up, those on the left point generally to the northeast.



(C.7.1)

Now select the red d**x** bar on the right and take it to be the dx of Fig (C.6.2). First determine the tangent base vectors $\mathbf{e}_\theta$ and $\mathbf{e}_\rho$ at the location of the red bar. Then setting d**x** = dθ $\mathbf{e}_\theta$ + dρ $\mathbf{e}_\rho$, consider the value of the two numbers dθ and dρ for this red bar. Graphically, knowing which way $\mathbf{e}_\theta$ and $\mathbf{e}_\rho$ point at the bottom of the red d**x**, one expects dθ > 0 and dρ > 0. The red d**x**' bar on the left has these Cartesian values dθ and dρ, and has a Cartesian-view length of $|d\mathbf{x}|^2 = (d\theta)^2+(d\rho)^2$. One can see from the picture that these Cartesian lengths *vary* for the 10 bars shown, though the lengths are all the same in x-space. The Curvilinear-view lengths of the x'-space bars are all the same, and are equal to the Cartesian length of those bars in x-space since **dx'•dx' = dx•dx.**

Consider now some bar mapping in the other direction:



(C.7.2)

Now the d**x** bars on the right all have different lengths. Those on the left have the same Cartesian length, which is what the drawing shows, but each one's Curvilinear-view length matches that of its corresponding bar on the right. The ratio of the length of a bar on the right to the Cartesian length of the corresponding bar on the left is the scale factor h_θ which recall is a function of location in space:

bar on right = $d\mathbf{x}^{(1)} = \mathbf{e}_1\, dx'^1 = \hat{\mathbf{e}}_1\, h'_1\, dx'^1 = \mathbf{e}_\theta\, d\theta = \hat{\mathbf{e}}_\theta\, h_\theta\, d\theta$ graph length = $h_\theta\, d\theta$

bar on left (Cartesian view) = $d\mathbf{x}'^{(1)} = \mathbf{e}'_1\, dx'^1 = \hat{\mathbf{e}}'_1\, dx'^1 = \hat{\mathbf{e}}_\theta\, d\theta$ graph length = $d\theta$

$$\Rightarrow \text{right bar length / left bar length} = h_\theta = \rho\sqrt{a^2\sin^2\theta + b^2\cos^2\theta} \quad \text{(increases with } \rho) \qquad \text{(C.7.3)}$$

If a=b, then $h_\theta = \rho$ and the bar length on the right is then $\rho d\theta$ as is obvious in polar coordinates.

C.8 A Derivation of the Jacobian Integration Rule

Consider now an integral $\int d\theta d\rho\, f(\theta,\rho)$. The tiny rectangles of area $d\theta d\rho$, like the specific gray and orange ones highlighted on the left above, are regarded for the purposes of integration as being in the Cartesian view of x'-space. One then writes [dA' is called $d\mathcal{V}'$ in Chapter 8]

$$dA' \equiv d\rho d\theta \;=\; \text{the area of a differential patch in Cartesian-view x'-space} \qquad \text{(C.8.1)}$$

This is the graphical area one sees in the picture. There is no need to define or consider any Curvilinear-view area in x'-space because the Cartesian-view area is being used.

In the limiting process which defines the integration, each $d\theta d\rho$ patch on the left has the same area $d\rho d\theta$. The interior of each patch on the left maps into some parallelogram patch on the right. One is not surprised to see that the patch areas on the right are different, though they map into patches on the left of the same Cartesian-view area. As shown in (8.4.d.3) the ratio of the two patch areas (volumes) is the absolute value of the Jacobian $|J(\mathbf{x}')|$,

$$\text{(area of skewed patch on the right at location } \mathbf{x}) \;=\; |J(\mathbf{x}')|\, dA' \;=\; |J(\mathbf{x}')|\, d\rho d\theta \qquad \text{(C.8.2)}$$

This is *not* what we mean by "the Jacobian Integration Rule" in the Section title. That is coming below and it is going to involve the quantity dxdy.

The mapping shown above between patches is an N=2 example of the general N-dimensional discussion in Section 8.2 which describes an orthogonal differential N-piped in (Cartesian-view) x'-space mapping into a non-orthogonal differential N-piped in x-space.

Now back to the integration issue. There are two ways an integration can be done in Cartesian x-space:

$$\text{integral of } f(x) \;=\; \lim \Sigma_i\, dA_1(\mathbf{x}_i)\, f(\mathbf{x}_i) \qquad dA_1(\mathbf{x}_i) \;= \text{patches shown on the right above}$$

$$\text{integral of } f(x) \;=\; \lim \Sigma_i\, dA_2(\mathbf{x}_i)\, f(\mathbf{x}_i) \qquad dA_2(\mathbf{x}_i) \;= dxdy \qquad \text{(C.8.3)}$$

In the <u>first integral</u>, every patch $dA_1(\mathbf{x}_i)$ on the right has a different shape and a different area as the integral is computed in the usual limiting-sum manner. The gray and orange patches on the right are two of these many patches. Despite their non-uniform shape and area, this rag-tag band of patches certainly "covers" the area being integrated over, and does so perfectly in the calculus limit. The area of one of these rag-tag patches is $|J(\mathbf{x}')|dA' = |J(\mathbf{x}')|d\theta d\rho$ and the areas are different because the Jacobian is a function of $\mathbf{x} = \mathbf{x}(\mathbf{x}')$.

In the second integral, every patch $dA_2(\mathbf{x_i})$ has the same area dxdy, so really $dA_2(\mathbf{x_i})$ does not depend on $\mathbf{x_i}$ in this form of the integration. One such dxdy patch is shown in green above. The coverage of the dA_2 patches is of course also "perfect coverage" in the calculus limit.

Since both integrals cover the same area perfectly, they both give the same result in the limiting process that defines the integral. This point is sometimes misunderstood. One is not just "replacing" a parallelogram patch such as the orange one on the right with some dxdy patch that approximates it in area, like the green patch. The statement is about an integration. Thus one has

$$\lim \Sigma_i\, dA_1(\mathbf{x_i})\, f(\mathbf{x_i}) \;=\; \lim \Sigma_i\, dA_2(\mathbf{x_i})\, f(\mathbf{x_i}) \qquad \text{(C.8.4)}$$

or

$$\int [|J(\mathbf{x}')|\, d\theta d\rho]\;\; f(\mathbf{x}(\mathbf{x}')) \;=\; \int [dxdy]\, f(\mathbf{x}) \qquad \text{(C.8.5)}$$

where on the left $f(\mathbf{x}) = f(\mathbf{x}(\mathbf{x}'))$ where $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}') \equiv \mathbf{x}(\mathbf{x}')$. In the sense of distribution theory (Stakgold Chapters 1 and 5), one can then make this *symbolic* statement

$$|J(\theta,\rho)|\, d\rho d\theta \;=\; dxdy \qquad \text{(C.8.6)}$$

where the meaning of this symbolic equality is the integral statement above,

$$\int_D\, dxdy\, f(\mathbf{x}) \;=\; \int_{D'}\, d\theta d\rho\, |J(\mathbf{x}')|\, f(\mathbf{x}(\mathbf{x}'))\;\;, \qquad \text{(C.8.7)}$$

valid for any integrable $f(\mathbf{x})$ and any integration region D (region D' corresponds to D in x'-space.) Either of these last two equations constitute the "Jacobian Integration Rule" of the Section title.

The integral on the left is well defined in 2D calculus, so the expression on the right shows how to "evaluate the integral on the left in curvilinear coordinates".

At this point one may introduce a new but obvious symbol

$$dA \equiv dxdy \qquad \text{(C.8.8)}$$

so the above equality of integrals can be written

$$\int dA\, f(\mathbf{x}) \;=\; \int\, dA'\, |J(\mathbf{x}')|\, f(\mathbf{x}(\mathbf{x}')) \qquad\qquad |J(\mathbf{x}')|\, dA' = dA \qquad \text{(C.8.9)}$$

In N dimensions, dA and dA' are differential "volumes", and the general Jacobian Integration rule takes the form,

$$\int dV\ f(\mathbf{x}) = \int dV'\ |J(\mathbf{x}')|\ f(\mathbf{x}(\mathbf{x}')) \qquad |J(\mathbf{x}')|\ dV' = dV$$

$$dV' \equiv dx'_1 dx'_2 ....dx'_N = \text{the volume of an orthogonal differential N-piped in the Cartesian-view x'-space}$$

$$dV = dx_1 dx_2 ....dx_N = \text{the volume of an orthogonal differential N-piped in x-space.} \tag{C.8.10}$$

Notice that these are *not* the two N-pipeds which "map into each other" as noted above. The N-piped dV has nothing to do that that mapping which involved a non-orthogonal N-piped in x-space.

To finish off our sample N=2 case, recall from earlier that for our polar elliptical coordinate system

$$|J'(x')| = |\det(S)| = ab\rho \tag{C.8.11}$$

and therefore

$$\int dxdy\ f(x,y) = \int d\theta d\rho\ |J(\mathbf{x}')|\ f(a\rho\cos\theta, b\rho\sin\theta) = ab \int d\theta d\rho\ \rho\ f(a\rho\cos\theta, b\rho\sin\theta)\ . \tag{C.8.12}$$

In the limit of regular polar coordinates, one then has a = b = 1 and ρ = r so

$$\int dxdy\ f(x,y) = \int rdrd\theta\ f(r\cos\theta, r\sin\theta) \tag{C.8.13}$$

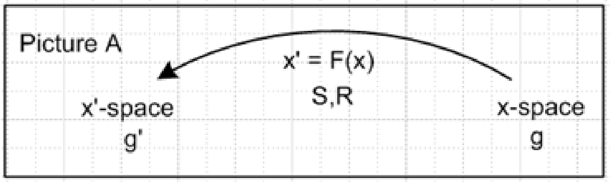
which is the familiar result.

Appendix D: Tensor Densities and the ε tensor

In this Section, DN means Developmental Notation.

D.1 Definition of a tensor density

Picture A is used in this Appendix along with the Standard Notation of Chapter 7.



(D.1.1)

First, equations (5.12.20) summarize facts about the Jacobian J which were converted to Standard Notation in (7.5.18) through (7.5.22). Here we add a symbol $\sigma \equiv \text{sign}(J)$ and restate those facts in Standard Notation as

$$J \equiv \det(S^i{}_j) = \sigma\sqrt{sg'}/\sqrt{sg} = \sigma(sg'/sg)^{1/2} = \sigma(g'/g)^{1/2} \quad \Rightarrow \quad (g'/g)^{1/2} = \sigma J = |J| > 0 \qquad J^2 = (g'/g)$$

$$s = \text{sign}[\det(g_{ij})] = \text{sign}(g) = \text{sign}(g') \qquad g = \det(g_{ij}) \qquad sg = |g| > 0 \qquad S^i{}_j \equiv (\partial x^i/\partial x'^j)$$
$$\sigma = \text{sign}[\det(S^i{}_j)] = \text{sign}(J) \qquad g' = \det(g'_{ij}) \qquad sg' = |g'| > 0 \ . \tag{D.1.2}$$

For proper Lorentz transformations of special relativity, det(S) = 1 so σ = +1. For curvilinear coordinates, one normally selects an ordering of the x_i so that σ = +1, such as r,θ,φ in spherical coordinates. Nevertheless, we allow for the possibility of J < 0.

Next, recall our generic sample tensor transformation (7.10.1), written two ways using (7.5.13) $S^a{}_b = R_b{}^a$:

$$T'^{abc}{}_{de} = R^a{}_{a'} R^b{}_{b'} R^c{}_{c'} S^{d'}{}_d S^{e'}{}_e T^{a'b'c'}{}_{d'e'}$$

$$T'^{abc}{}_{de} = R^a{}_{a'} R^b{}_{b'} R^c{}_{c'} R_d{}^{d'} R_e{}^{e'} T^{a'b'c'}{}_{d'e'} \ . \qquad \textit{(7.10.1)} \tag{D.1.3}$$

T is a mixed rank-5 tensor, meaning it transforms as shown above with respect to the underlying transformation F. T is a regular standard-issue tensorial tensor.

Now suppose instead that the object T were to transform like this, with J being the Jacobian noted above,

$$T'^{abc}{}_{de} = J^{-W} R^a{}_{a'} R^b{}_{b'} R^c{}_{c'} R_d{}^{d'} R_e{}^{e'} T^{a'b'c'}{}_{d'e'} \tag{D.1.4}$$

where the extra factor J^{-W} has been introduced. If T transforms this way, it is called a tensor density of weight W. Thus, an ordinary tensor is a tensor density of weight 0.

The convention for the sign of W used here is that of Weinberg p 99 Eq. (4.4.4), which equation has the following factor on the right side of a sample tensor density transform equation,

$$|\partial x'/\partial x|^{W} \equiv [\det(\partial x'/\partial x)]^{+W} = [\det(\partial x'_i/\partial x_k)]^{+W} = J^{-W} . \qquad (D.1.5)$$

Some authors use -W as the "weight" instead of +W, but we shall stick with Weinberg's convention.

An immediate example of a tensor density is provided by $(g'/g)^{1/2} = |J|$ rewritten as

$$g' = J^2 g = J^{-(-2)} g \Rightarrow \quad \text{weight}(g) = -2 \qquad g \text{ is a scalar density of weight - 2 .} \qquad (D.1.6)$$

so $g \equiv \det(g_{ij})$ is a scalar density of weight -2. Notice that from g one can construct other scalar densities of other weights, for example :

$$g'^{-1} = J^{-(2)} g^{-1} \quad \Rightarrow \quad \text{weight}(g^{-1}) = +2 \qquad g'^{-1} \text{ is a scalar density of weight +2 .} \qquad (D.1.7)$$

D.2 A few facts about tensor densities

1. It is pretty obvious that a **sum** of two index-similar tensor densities of weight W has weight W. (D.2.1)

2. **Contracting** indices within a tensor density does not alter its weight W. (D.2.2)

If indices a and d are contracted in the (D.1.4) example above, one gets

$$T'^{abc}{}_{ae} = J^{-W} R^a{}_{a'} R^b{}_{b'} R^c{}_{c'} R_a{}^{d'} R_e{}^{e'} T^{a'b'c'}{}_{d'e'}$$

$$= J^{-W} (R^a{}_{a'} R_a{}^{d'}) R^b{}_{b'} R^c{}_{c'} R_e{}^{e'} T^{a'b'c'}{}_{d'e'}$$

$$= J^{-W} \delta_{a'}{}^{d'} R^b{}_{b'} R^c{}_{c'} R_e{}^{e'} T^{a'b'c'}{}_{d'e'}$$

$$= J^{-W} R^b{}_{b'} R^c{}_{c'} R_e{}^{e'} T^{a'b'c'}{}_{a'e'} .$$

The factor J^{-W} just sits there, impervious to contraction activities.

3. Going the other direction, when a larger tensor density is formed from two smaller ones, called a direct product or **outer product**, the **weights get added**. (D.2.3)

Example 1: Start with two tensor densities A and B of weights W1 and W2, form the outer product:

$$A'^a = J^{-W1} R^a{}_{a'} A^{a'}$$

$$B'^c{}_d = J^{-W2} R^c{}_{c'} R_d{}^{d'} B^{c'}{}_{d'}$$

$$\Rightarrow \qquad (A'^a B'^c{}_d) = J^{-(W1+W2)} R^a{}_{a'} R^c{}_{c'} R_d{}^{d'} (A^{a'} B^{c'}{}_{d'}) .$$

Example 2: Start with $|g'| = J^2|g|$ from (D.1.2) and raise both sides to the power -W1/2 :

$$|g'|^{-W1/2} = J^{-W1} |g|^{-W1/2} . \qquad // \text{ no R factors since scalar} \qquad (D.2.3a)$$

Thus, the factor ($|g|^{-W1/2}$) is a scalar density of weight W1. Let this be the first factor of an "outer product" where the second factor is the following tensor density of weight W2,

$$B'^{c}{}_{d} = J^{-W2}\, R^{c}{}_{c'} R_{d}{}^{d'}\, B^{c'}{}_{d'} \qquad \text{// same as in Example 1}$$

$$\Rightarrow \qquad (|g'|^{-W1/2} B'^{c}{}_{d}\,) = J^{-(W1+W2)}\, R^{a}{}_{a'}\, R^{c}{}_{c'} R_{d}{}^{d'}\, (|g|^{-W1/2}\, B^{c'}{}_{d'}) \,. \qquad \text{(D.2.3b)}$$

If one selects W1 = –W2, the added factor *neutralizes* the weight of the tensor density to which it is prefixed, generating thereby a regular tensor (weight 0). So if tensor density B has weight W,

$$(|g'|^{W/2}\; B'^{c}{}_{d}) \;=\; R^{a}{}_{a'}\, R^{c}{}_{c'} R_{d}{}^{d'}\, (|g|^{W/2}\, B^{c'}{}_{d'}) \qquad \text{(D.2.3c)}$$

and then ($|g|^{W/2}\; B^{i}{}_{j}$) transforms under F as a regular tensor. (One should always keep in mind the fact that there is an underlying transformation **x**' = **F**(**x**) upon which the House of Tensor is built). This is a standard method of converting a tensor density to a regular tensor.

4. Although sometimes authors take a differing stance for certain tensors like ε below, we shall assume that indices are raised and lowered on a tensor density in *exactly the same way* they are raised and lowered on an ordinary tensor of the same index structure (see Section 7.4). This means the $\mathbf{g^{ab}}$ **raises** an index and $\mathbf{g_{ab}}$ **lowers** an index. (D.2.4)

5. <u>Raising or lowering an index does not change the weight of a tensor density</u>. (D.2.5)

Again, using our generic example (D.1.4) above,

$$T'^{abc}{}_{de} = \; J^{-W}\;\, R^{a}{}_{a'}\, R^{b}{}_{b'}\, R^{c}{}_{c'}\, R_{d}{}^{d'}\; R_{e}{}^{e'}\; T^{a'b'c'}{}_{d'e'} \qquad (1)$$

$$T'^{abc}{}_{d}{}^{e} = g'^{ex}\, T'^{abc}{}_{dx} \qquad (2) \qquad \text{// raise last index e on T' in x'-space}$$

$$T^{a'b'c'}{}_{d'e'} = g_{e'e''}\, T^{a'b'c'}{}_{d'}{}^{e''} \;. \quad (3) \qquad \text{// lower last index on T in x-space}$$

Therefore,

$$T'^{abc}{}_{d}{}^{e} = g'^{ex}\, [\, J^{-W}\;\, R^{a}{}_{a'}\, R^{b}{}_{b'}\, R^{c}{}_{c'}\, R_{d}{}^{d'}\; R_{x}{}^{e'}\; T^{a'b'c'}{}_{d'e'}] \qquad \text{// this is (2) + (1) above}$$

$$= g'^{ex}\, [\, J^{-W}\;\, R^{a}{}_{a'}\, R^{b}{}_{b'}\, R^{c}{}_{c'}\, R_{d}{}^{d'}\; R_{x}{}^{e'}\; (g_{e'e''}\, T^{a'b'c'}{}_{d'}{}^{e''})] \qquad \text{// use (3) above}$$

$$= \;\; J^{-W}\;\, R^{a}{}_{a'}\, R^{b}{}_{b'}\, R^{c}{}_{c'}\, R_{d}{}^{d'}\; (g'^{ex}\, R_{x}{}^{e'}\; g_{e'e''})\, T^{a'b'c'}{}_{d'}{}^{e''} \qquad \text{// regroup}$$

$$= \;\; J^{-W}\;\, R^{a}{}_{a'}\, R^{b}{}_{b'}\, R^{c}{}_{c'}\, R_{d}{}^{d'}\; (R^{e}{}_{e''})\, T^{a'b'c'}{}_{d'}{}^{e''} \qquad \text{// from last line in (7.5.9)}^{\dagger}$$

and again J^{-W} passively watches all the action fly by. The weight of our generic tensor density with its last index raised is still W. † The last step uses (7.5.9) $R_{a}{}^{b} = g'_{aa'} R^{a'}{}_{b'} g^{b'b}$ but we raise a, lower b, and reverse two tilts to get $R^{a}{}_{b} = g'^{aa'} R_{a'}{}^{b'}\, g_{b'b}$.

6. The covariant **dot product** of vector densities **A** and **B** of weights W and w is a scalar density of weight W + w and therefore $\mathbf{A}'\bullet\mathbf{B}' = J^{-(W+w)}\ \mathbf{A}\bullet\mathbf{B}$. (D.2.6)

Proof: First form the rank-2 tensor density A^iB^j which by (D.2.3) has weight W+w. Lower the second index and the mixed rank-2 tensor A^iB_j by item (D.2.5) still has weight W+w. Then contract to get **A•B** = A^iB_i and by (D.2.2), the weight is still W+w.

Corollary: The magnitude of a vector density **A** of weight W is a scalar density of weight W, and therefore $|\mathbf{A}|' = J^{-W}|\mathbf{A}|$. (D.2.7)

Proof: $|\mathbf{A}|^2 = \mathbf{A}\bullet\mathbf{A}$ has weight 2W by (D.2.6), meaning $|\mathbf{A}'|^2 = J^{-2W}|\mathbf{A}|^2$. Therefore $|\mathbf{A}'| = J^{-W}|\mathbf{A}|$.

7. As J→1, tensor densities become true tensors. (D.2.8)

One could imagine some limiting/morphing process on an underlying transformation F such that the linearized transformation matrix R approaches a rotation matrix at all points in space (RR^T= 1 and detR = 1, DN) and then J = detS → 1. In this case $J^{-W} \rightarrow 1^{-W} = 1$ and therefore any tensor density, regardless of its weight W, becomes an ordinary tensor. Perhaps we should restrict this comment to underlying transformations F having J = detS > 0 since passing through detS = 0 is problematical. As was noted below (5.12.16), for an invertible $\mathbf{x}' = \mathbf{F}(\mathbf{x})$, one does not have $\det S(\mathbf{x}) = 0$. at any **x**.

Example: The cross product considered in Section D.7 below of N-1 contravariant vectors becomes in this limit an ordinary covariant vector. If g=1 in x-space, then $g' = RR^T = 1$ in x'-space (DN) and then that resulting vector can be considered either contravariant or covariant since both spaces are then Cartesian. This is the case with **A** = **B** x **C** under rotations in 3D space. On can think of the ε_{abc} as moving in this limit from a tensor density of weight -1 to an ordinary tensor (ε is treated in Section D.4 below).

8. If vector **V** is a vector density of weight W, then the four **expansions** shown in (7.13.10) become

$$\begin{aligned}
\mathbf{V} &= \Sigma_n V^n\, \mathbf{u}_n && \text{with} \quad \mathbf{u}^n \bullet \mathbf{V} = V^n \\
\mathbf{V} &= \Sigma_n V_n\, \mathbf{u}^n && \text{with} \quad \mathbf{u}_n \bullet \mathbf{V} = V_n \\
\mathbf{V} &= J^W\, \Sigma_n V'^n\, \mathbf{e}_n && \text{with} \quad \mathbf{e}^n \bullet \mathbf{V} = J^W\, V'^n \\
\mathbf{V} &= J^W\, \Sigma_n V'_n\, \mathbf{e}^n && \text{with} \quad \mathbf{e}_n \bullet \mathbf{V} = J^W\, V'_n \ .
\end{aligned} \qquad (D.2.9)$$

Proof: If **V** has weight W then, according to item 6 above, the four dot products shown on the right are scalar densities of weight W (the basis vectors all have weight 0). We then evaluate the four dot products :

$$\begin{aligned}
&\mathbf{u}^n \bullet \mathbf{V} = (\mathbf{u}^n)_i\, V^i = \delta^n{}_i V^i = V^n && \text{// see (7.18.3)} \\
&\mathbf{u}_n \bullet \mathbf{V} = (\mathbf{u}_n)^i\, V_i = \delta_n{}^i V_n = V_n && \text{// see (7.18.3)}
\end{aligned}$$

$$\begin{aligned}
&(\mathbf{e}^n \bullet \mathbf{V}\,) = J^W\, (\mathbf{e}'^n \bullet \mathbf{V}'\,)\ \ J^W\ (\mathbf{e}'^n)_i V'^i = J^W\, \delta^n{}_i V'^i = J^W\ V'^n && \text{// see (7.18.1)} \\
&(\mathbf{e}_n \bullet \mathbf{V}\,) = J^W\, (\mathbf{e}'_n \bullet \mathbf{V}'\,)\ \ J^W\ (\mathbf{e}'_n)^i V'_i = J^W\, \delta_n{}^i V'_i = J^W\ V'_n && \text{// see (7.18.1)}
\end{aligned} \qquad (D.2.10)$$

In the last two lines, we use item 6 that $\mathbf{A}\bullet\mathbf{B} = J^{+(W+w)}\ \mathbf{A'}\bullet\mathbf{B'}$ with w=0.

Thus the two $\mathbf{u_n}$ expansions are unaltered, but the $\mathbf{e_n}$ expansions pick up an extra factor of J^W. For example, assume that $\mathbf{V} = \Sigma_m\alpha_m\,\mathbf{e_m}$ with α_m unknown. Then

$$\mathbf{V}\bullet\mathbf{e^n} = (\Sigma_m\alpha_m\,\mathbf{e_m})\bullet\mathbf{e^n} = \Sigma_m\alpha_m(\mathbf{e_m}\bullet\mathbf{e^n}) = \Sigma_m\alpha_m\delta_m{}^n = \alpha_n$$

Therefore $\alpha_n = \mathbf{V}\bullet\mathbf{e^n} = J^W\ V'_n$ so the expansion must be $\mathbf{V} = J^W\ \Sigma_m\ V'^n\ \mathbf{e_m}$.

9. This same idea applies to the more general **tensor expansions** of Appendix E. For example, if A^{ijk} is a tensor density of weight W, one will have in place of (E.2.9),

$$A = J^W\ \Sigma_{ijk}\ A'^{ijk}\ (\mathbf{e_i}\otimes\mathbf{e_j}\otimes\mathbf{e_k})\ . \tag{D.2.11}$$

Proof: According to item 6, the direct product space dot product shown below is a scalar density of weight W, so

$$A\bullet(\mathbf{e^a}\otimes\mathbf{e^b}\otimes\mathbf{e^c}) = J^W\ A'\bullet(\mathbf{e'^a}\otimes\mathbf{e'^b}\otimes\mathbf{e'^c}) = J^W\ A'^{ijk}\ (\mathbf{e'^a})_i\ (\mathbf{e'^b})_j\ (\mathbf{e'^b})_k = J^W\ A'^{abc}$$

If we write $A = \Sigma_{ijk}\ \alpha^{ijk}\ (\mathbf{e_i}\otimes\mathbf{e_j}\otimes\mathbf{e_k})$ with α^{ijk} unknown, we find as in the vector case above that the coefficients are $\alpha^{ijk} = J^W\ A'^{abc}$.

D.3 Theorem about Totally Antisymmetric Tensors: there is really only one: $\varepsilon^{abc...}$

The permutation tensor $\varepsilon^{abc...x}$ has a number of indices equal to the dimension N of the space in which one is working. For N = 3 one has ε^{abc}, and ε^{ab} or ε^{abcd} do not exist. The normal calibration of the tensor is $\varepsilon^{123..N} = 1$. Each index swap causes a minus sign such as $\varepsilon^{abc} = -\varepsilon^{bac} = -\ \varepsilon^{cba}$, and if any pair of indices are the same, the result is zero, such as $\varepsilon^{131} = 0$, Due to this swapping property on any index, this tensor is called "totally antisymmetric".

Theorem: Apart from a scalar factor, there exists only one totally antisymmetric (TA) tensor. (D.3.1)

Proof: Suppose there were two TA tensors called $\varepsilon^{abc...}$ and $r^{abc...}$. If two or more of the indices are equal, both tensors are 0, so for such index sets, one can say $r^{abc..} = f\ \varepsilon^{abc..}$ where f is any finite function whatsoever. Consider now the case where all the indices are distinct and therefore exhaust the set 123...N, and consider abc... to be a permutation of 123...N obtained by doing S pairwise swaps,

$$abc... = P(123...) \qquad\qquad p = (-1)^S\ . \tag{D.3.2}$$

If one were to associate a sign change with each swap, the total sign change would be p, the parity. Since ε and r are both TA tensors, each tensor can be "unwound" back to a standard index order by doing these S swaps, and the swaps will cause a total sign of p relative to that standard order, so

$$r^{abc..} = p\ r^{123...} \qquad\qquad // \text{ for example, } r^{2134..} = (-1)^1\ r^{1234..} \tag{D.3.3}$$
$$\varepsilon^{abc..} = p\ \varepsilon^{123...}$$

Define scalar function $f \equiv r^{123...} / e^{123...}$, whatever it might be. Then

$$r^{abc..} = p(f\,\varepsilon^{123...})$$
$$\varepsilon^{abc..} = p\,\varepsilon^{123...} \tag{D.3.4}$$

and dividing these two equations one finds,

$$r^{abc..} = f\,\varepsilon^{abc..} \tag{D.3.5}$$

which has now been shown valid for all index sets abc.. . Therefore, any "other" totally antisymmetric tensor is just a scalar function times the ε tensor.

Alternate proof assuming (D.11.8) below. Assume $r^{abc..} = f\,\varepsilon^{abc..}$. Apply $\varepsilon_{abc..}$ to both sides and sum on abc.. to get $\varepsilon_{abc..}r^{abc..} = \varepsilon_{abc..}(f\,\varepsilon^{abc..}) = (\varepsilon_{abc..}\,\varepsilon^{abc..})f = N!\ |g|\ f$ by (D.11.8). Therefore the assumed form $r^{abc..} = f\,\varepsilon^{abc..}$ is valid with $f = (N!|g|)^{-1}\varepsilon_{abc..}r^{abc..}$.

Similar Example: For N= 3, if rank-4 tensor M^{abcd} is totally antisymmetric on indices abc, then M^{abcd} can be written in the form $M^{abcd} = \varepsilon^{abc}q^d$ where q^d is a vector. In fact $q^d = (3!|g|)^{-1}\,\varepsilon_{abc}M^{abcd}$.

D.4 The contravariant ε tensor

The "Levi-Civita symbol" or tensor is one that equals the permutation tensor when all indices are "up", as we shall see (a convention). One does not raise and lower indices of the mechanical permutation "tensor", but one does do this on the Levi-Civita tensor, which we shall just call "the ε tensor". As we shall also see, the ε tensor is really a tensor density of weight -1. When writing $\varepsilon_{abc..}$ one must have clearly in mind whether one means the permutation tensor or the Levi-Civita tensor. Unless indices are all up, they are in general not the same.

Knowing nothing to start, assume that the totally antisymmetric $\varepsilon^{abc..}$ tensor transforms under F as a tensor density of some weight W which we hope to determine. Then

$$\varepsilon'^{abc..} = J^{-W}\ R^a{}_{a'}\ R^b{}_{b'} \ldots \varepsilon^{a'b'c'..} \quad . \tag{D.4.1}$$

Assume that $\varepsilon^{abc..}$ is the usual permutation tensor normalized to $\varepsilon^{123...N} = +1$. This is the convention used by Weinberg p 99. This means each index swap changes the sign, and if two or more indices are the same, ε = 0. This is an important starting assumption, and from it most everything follows.

Given this assumption, the RHS of (D.4.1) is totally antisymmetric (TA). The argument is given once here and then used later several times. Consider an a↔b swap. Then

$$\varepsilon'^{bac..} = J^{-W}\ R^b{}_{a'}\ R^a{}_{b'} \ldots \varepsilon^{a'b'c'..} = J^{-W}\ R^b{}_{b'}\ R^a{}_{a'} \ldots \varepsilon^{b'a'c'..} \quad \text{// b'↔a' dummy indices}$$

$$= J^{-W}\ R^a{}_{a'}\ R^b{}_{b'} \ldots (-\varepsilon^{a'b'c'..}) = -\,\varepsilon'^{abc..} \quad . \tag{D.4.2}$$

The same result is true for any swap, thus RHS (D.4.1) is TA. Since according to Section D.3 there is only one TA tensor available, apart from a scalar function factor, it follows that

$$\varepsilon'^{abc..} = K\varepsilon^{abc...} \qquad \text{(D.4.3)}$$

where K is some scalar function, perhaps just a constant. Equation (D.4.1) above then reads

$$K\,\varepsilon^{abc...} = J^{-W}\ R^{a}{}_{a'}\ R^{b}{}_{b'}\ ...\ \varepsilon^{a'b'c'..} \quad . \qquad \text{(D.4.4)}$$

Setting in the standard order, one finds that

$$K\,\varepsilon^{123...} = J^{-W}\ R^{1}{}_{a'}\ R^{2}{}_{b'}\ ...\ \varepsilon^{a'b'c'..} \qquad \text{(D.4.5)}$$

or

$$K = J^{-W}\det(R^{i}{}_{j}) = J^{-W}(J)^{-1} = J^{-(W+1)} \qquad // \ J = \det(S) = 1/\det(R) \qquad \text{(D.4.6)}$$

so now

$$\varepsilon'^{abc..} = K\varepsilon^{abc...} = J^{-(W+1)}\ \varepsilon^{abc...} \qquad \text{(D.4.7)}$$

A second **assumption** is now made: that $\varepsilon^{abc..}$ (contravariant!) has the same value structure in any frame of reference, which is to say it is the *same* in x'-space as it is in x-space,

$$\varepsilon'^{abc..} = \varepsilon^{abc...} \quad . \qquad \text{(D.4.8)}$$

where $\varepsilon^{abc...}$ is the usual permutation tensor with $\varepsilon^{123...} = +1$. This assumption is consistent with taking W = -1 in (D.4.7).

Fact: The Levi-Civita tensor $\varepsilon^{abc...}$ has weight W = -1. From (D.2.5) this is true regardless of the index positions on the ε tensor. (D.4.9)

Again, this follows the convention of Weinberg p 99. Some authors instead arrange for the above equation to be true for the *covariant* ε tensors, and use then $\varepsilon_{123...N} = \varepsilon'_{123...N} = +1$, but we shall follow Weinberg.

To summarize, assuming that $\varepsilon^{abc...}$ is the usual permutation tensor normalized in the usual way, and assuming that $\varepsilon'^{abc..} = \varepsilon^{abc...}$ so this tensor is the same in all frames or spaces, THEN one concludes that $\varepsilon^{abc...}$ must transform as a rank-N tensor density of weight W = -1. That is to say,

$$\varepsilon'^{abc..} = J\,R^{a}{}_{a'}\ R^{b}{}_{b'}\ ...\ \varepsilon^{a'b'c'..} \qquad // \text{ this is (D.4.1) above with W = -1} \qquad \text{(D.4.10)}$$

Viewed in this light, the tensor $\varepsilon^{abc..}$ is known as the Levi-Civita tensor.

Tullio Levi-Civita (1873-1941). Italian, University of Padua 1892, with Ricci published the theory of tensor algebra in 1900 (see Refs.), which work assisted Einstein circa 1915 in formulating the theory of general relativity. The ε tensor bears his name. Sometimes the affine connection (Appendix F) is called the Levi-Civita connection.

D.5 Some facts about the ε tensor

1. Consider, with the assumption of (D.2.4) above applied to the ε tensor,

$$\varepsilon_{abc...} = g_{aa'} g_{bb'} ..... \varepsilon^{a'b'c'...} \qquad \text{// just as with any other tensor} \qquad (D.5.1)$$

This is again in the convention of Weinberg p 99 (4.4.10). Although we show all indices lowered at once, one can lower them one at a time in the usual manner using g_{ij}.

> Added-sign-s convention. Some authors make a special exception for the ε tensor and introduce an extra sign s into the above equation (recall that s = -1 for special relativity)
>
> $$\varepsilon_{abc...} = s\, g_{aa'} g_{bb'} ..... \varepsilon^{a'b'c'...} \qquad s = \text{sign}[\det(g_{ij})] \quad . \qquad (D.5.2)$$
>
> This convention then invalidates the idea (7.4.11) that g_{ab} lowers a tensor index on ε and g^{ab} raises a tensor index on ε. If that idea were valid, the above would read $\varepsilon_{abc..} = s\, \varepsilon_{abc}$ and would force s =1. Furthermore, the meaning of a mixed ε tensor like $\varepsilon_a{}^b$ becomes undefined in this convention because we don't know how to obtain $\varepsilon_a{}^b$ from say ε_{ab}. Since this convention is nevertheless used by some authors, we shall present some results below in this added-sign-s convention.
>
> Up-Equals-Down convention. Recall that we use the Weinberg convention (D.5.1), which says ε's indices work like those of any other tensor, *and* (D.4.8) which says $\varepsilon'^{abc..} = \varepsilon^{abc...}$. Some authors instead replace the "fiat rule" $\varepsilon'^{abc..} = \varepsilon^{abc...}$ by a different fiat rule : $\hat{\varepsilon}_{abc...} = \varepsilon^{abc...}$ where we use a hat to distinguish their low index ε from our low index ε. Below we will show in the Weinberg convention that $\varepsilon_{abc...} = g\, \varepsilon^{abc...}$ so we can then identify $\hat{\varepsilon}_{abc...} = (1/g)\, \varepsilon_{abc...}$. The upper index version is the same, so $\hat{\varepsilon}^{abc..} = \varepsilon^{abc..}$. In this convention one finds that $\hat{\varepsilon}^{abc..}$ is a tensor density of weight - 1 and then $\hat{\varepsilon}_{abc...}$ has weight 2 - 1 = +1. We mention this convention only because it exists, but we shall not express any of our results in this "up equals down" convention.

Continuing with (D.5.1), install the reference sequence to obtain

$$\varepsilon_{123..} = g_{1a'} g_{2b'} ..... \varepsilon^{a'b'c'...} = \det(g_{ij}) = g \quad . \qquad \text{// see (D.2.2)} \qquad (D.5.3)$$

Similarly, $\varepsilon'_{123..} = \det(g'_{ij})$. To summarize,

$$\begin{aligned} &\varepsilon_{123..} = \det(g_{ij}) = g && \text{// } \varepsilon \text{ all-down index reference values} \\ &\varepsilon'_{123..} = \det(g'_{ij}) = g' \,. && \qquad (D.5.4) \end{aligned}$$

In the "added sign s convention", these last two equations would have sg = |g| and sg' = |g'| on the right which means then these two ε values would be always positive.

2. Take the same starting point as above

$$\varepsilon_{abc...} = g_{aa'}\, g_{bb'} .....\, \varepsilon^{a'b'c'...} \qquad \textit{(D.5.1)}$$

The RHS is a totally antisymmetric in indices abc... . This is by argument (D.4.2) repeated in this context,

$$\begin{aligned} \varepsilon_{bac...} &= g_{ba'}\, g_{ab'} .....\, \varepsilon^{a'b'c'...} = g_{bb'}\, g_{aa'} .....\, \varepsilon^{b'a'c'...} \\ &= g_{bb'}\, g_{aa'} .....\, (-\varepsilon^{a'b'c'...}) = -\ g_{aa'} g_{bb'} .....\, \varepsilon^{a'b'c'...} = -\, \varepsilon_{abc...} \end{aligned} \qquad \text{(D.5.5)}$$

Then according to Section D.3, the right side of the above can be written as

$$\text{RHS} = C\, \varepsilon^{abc...} \qquad \text{(D.5.6)}$$

since we there is only one TA tensor apart from scalar C. Therefore

$$\varepsilon_{abc...} = C\, \varepsilon^{abc...} \quad . \qquad \text{(D.5.7)}$$

Insert the reference sequence

$$\varepsilon_{123...} = C\, \varepsilon^{123...} = C \qquad \text{(D.5.8)}$$

But $\varepsilon_{123...} = \det(g_{ij})$ from (D.5.3), so

$$C = \det(g_{ij}) = g \qquad \text{(D.5.9)}$$

and then (D.5.1) says, for the "Weinberg convention",

$$\begin{aligned} \varepsilon_{abc...} &= \det(g_{ij})\, \varepsilon^{abc...} = g\, \varepsilon^{abc...} && \text{// relating all down to all up} \\ \varepsilon'_{abc...} &= \det(g'_{ij})\, \varepsilon'^{abc...} = g'\, \varepsilon'^{abc...} && \text{// = } g'\, \varepsilon^{abc...} \quad . \end{aligned} \qquad \text{(D.5.10)}$$

where the second line follows by the same argument. These equations relate all indices down to all up in the same space. Notice that both $\varepsilon_{abc...}$ and $\varepsilon'_{abc...}$ are totally antisymmetric.

In the "added sign s convention" the above equations are instead

$$\begin{aligned} \varepsilon_{abc...} &= |\det(g_{ij})|\, \varepsilon^{abc...} = |g|\, \varepsilon^{abc...} && \text{// relating all down to all up} \\ \varepsilon'_{abc...} &= |\det(g'_{ij})|\, \varepsilon'^{abc...} = |g'|\, \varepsilon'^{abc...} && \text{// = } |g'|\, \varepsilon^{abc...} \end{aligned} \qquad \text{(D.5.11)}$$

3. Divide the two Weinberg convention equations in (D.5.10) to find that

$$\varepsilon'_{abc...} = [\det(g'_{ij}) / \det(g_{ij})]\ \ \varepsilon_{abc...} = (g'/g)\, \varepsilon_{abc...} \qquad \text{(D.5.12)}$$

From (S.1.2) one has $(g'/g) = J^2$ so the conclusions regarding ε are these:

$$\begin{aligned} \varepsilon'_{abc...} = J^2\, \varepsilon_{abc...} = (g'/g)\, \varepsilon_{abc..} \qquad & \varepsilon'^{abc...} = \varepsilon^{abc...} = \text{permutation tensor} \quad \text{// general} \\ & \varepsilon_{abc...} = \text{permutation tensor if } g=1 \end{aligned} \qquad \text{(D.5.13)}$$

These conclusions are valid for the "added sign s convention" as well since det(g) and det(g') always have the same sign as shown in (D.1.2).

Two comments:

- Although we set $\varepsilon'^{abc...} = \varepsilon^{abc...}$ by fiat, we cannot similarly set $\varepsilon'_{abc...} = \varepsilon_{abc...}$ by fiat. This latter result comes out being $\varepsilon'_{abc...} = J^2\, \varepsilon_{abc...}$ as just shown. (D.5.14)

- The fact that $\varepsilon'_{abc...} = J^2\, \varepsilon_{abc...}$ does *not* say that ε_{abc} is a tensor density of weight -2 because there are no R factors showing (see tensor density definition (D.1.4)) . (D.5.15)

D.6 The covariant ε tensor : repeat Section D.4 as if its weight were not known

According to (D.2.5), lowering indices does not change the weight of a tensor density. Eq (D.4.9) shows that $\varepsilon^{abc..}$ is a tensor density of weight -1, so we know right away that $\varepsilon_{abc..}$ is also a tensor density of weight -1. Nevertheless, it is interesting (and tests our consistency) to see what happens when the same method used in Section D.4 for $\varepsilon^{abc...}$ is applied to $\varepsilon_{abc...}$.

We start by assuming $\varepsilon_{abc..}$ is a tensor density of some unknown weight W,

$$\varepsilon'_{abc..} = J^{-W} [\, R_a{}^{a'} R_b{}^{b'} .... \varepsilon_{a'b'c'..}] \quad . \tag{D.6.1}$$

Eq (D.5.10) shows that $\varepsilon_{a'b'c'..}$ is totally antisymmetric (TA) since we know $\varepsilon^{abc...}$ is TA. The argument of (D.4.2) with up↔down indices then shows from (D.6.1) that $\varepsilon'_{abc...}$ is TA as well. Since according to Section D.3 there is only one TA tensor available, apart from a scalar function factor, it follows that RHS(D.6.1) = K $\varepsilon_{abc..}$ so (D.6.1) then says,

$$K\, \varepsilon_{abc..} = J^{-W} [\, R_a{}^{a'} R_b{}^{b'} .... \varepsilon_{a'b'c'..}] \quad . \tag{D.6.2}$$

Use (D.5.10) to set $\varepsilon_{a'b'c'..} = \det(g_{ij})\, \varepsilon^{a'b'c'..}$ inside the bracket,

$$K\, \varepsilon_{abc..} = J^{-W} [\, R_a{}^{a'} R_b{}^{b'} .... \det(g_{ij})\, \varepsilon^{a'b'c'..}] \ , \tag{D.6.3}$$

and then install the reference sequence on both sides

$$K\, \varepsilon_{123..} = J^{-W} [\, R_1{}^{a'} R_2{}^{b'} .... \det(g_{ij})\, \varepsilon^{a'b'c'..}] \quad . \tag{D.6.4}$$

But (D.5.3) says $\varepsilon_{123..} = \det(g_{ij})$, so cancel $\det(g_{ij})$ on both sides to get [recall $R_i{}^j = S^j{}_i$]

$$K = J^{-W} [\, R_1{}^{a'} R_2{}^{b'} .... \varepsilon^{a'b'c'..}] = J^{-W} \det(R_i{}^j) = J^{-W} \det(S^j{}_i) = J^{-W}\, J = J^{-(W-1)} \ . \tag{D.6.5}$$

So here the result is $K = J^{-(W-1)}$ whereas in (D.4.6) the result was $K = J^{-(W+1)}$. In the current case, since (D.6.1) and (D.6.2) have the same RHS, setting the LHS's equal says

$$\varepsilon'_{abc..} = K\,\varepsilon_{abc..} = J^{-(W-1)}\,\varepsilon_{abc..} \qquad \text{(D.6.6)}$$

But (D.5.13) says that $\varepsilon'_{abc..} = J^2\,\varepsilon_{abc..}$ and therefore W = -1.

Thr conclusion is that $\varepsilon_{abc...}$ transforms with weight -1, the same as $\varepsilon^{abc...}$, so (D.6.1) becomes

$$\varepsilon'_{abc..} = J\ [\,R_a{}^{a'}\,R_b{}^{b'}....\ \varepsilon_{a'b'c'..}]\,. \qquad \text{(D.6.7)}$$

D.7 Generalized cross products

In (A.4.1) and (A.4.2) the following cross product of N-1 vectors is considered (now in Standard Notation, all indices are contracted except a)

$$Q_a \equiv \varepsilon_{abc...x}\,B^bC^cD^d.....X^x \qquad \text{or} \qquad \mathbf{Q} = \mathbf{B}\ \mathrm{x}\ \mathbf{C}\ \mathrm{x}\ \mathbf{D}\ ....\ \mathrm{x}\ \mathbf{X}\ . \qquad \text{(D.7.1)}$$

If the vectors B,C,D..X are all contravariant vectors, then applying the rule (D.2.3), one concludes that, since ε is a tensor density of weight -1 and since all the RHS vectors have weight 0, the object Q_a is a covariant vector density of weight = -1, and thus has this transformation rule

$$Q'_a = J\,R_a{}^bQ_b\,. \qquad \text{(D.7.2)}$$

Similarly, one may consider

$$Q^a \equiv \varepsilon^{abc...x}\ B_bC_cD_d.....X_x\ . \qquad \text{(D.7.3)}$$

If vectors B,C,D...X are covariant vectors, then Q^a is a vector density of weight -1 and

$$Q'^a = J\,R^a{}_bQ^b\,. \qquad \text{(D.7.4)}$$

D.8 The tensorial nature of curl B

It has just been shown that **C** = **A** x **B** [$C_a \equiv \varepsilon_{abc}\,A^bB^c$] is a vector density of weight -1, this being a special case of the generalized cross product discussion above. As noted in the Example below (D.2.8), if $g_{ij} = 1$ and r is a (global) rotation, then $g'_{ij} = 1$ as well. This means from (D.1.2) that g = g' = 1 and J = 1. Then the weight -1 transformation rule $C'_a = J\,R_a{}^bC_b$ becomes $C'_a = R_a{}^bC_b$ which is the same as a weight 0 rule, so then **C** can be regarded as a regular tensor in its transformation properties (in this special case). [Under inversion, if **A**→ **-A** and **B**→ **-B**, one gets **C**→ +**C**, so **C** is sometimes called a pseudotensor.]

One might conjecture that **C** = ∇ x **B** is also a vector density of weight -1, and that conjecture is correct as is now shown. Consider

$$C^n = \varepsilon^{nab}\,\partial_aB_b \qquad \text{(D.8.1)}$$

where **B** is assumed to be an tensorial vector. It is helpful to write this equation in the following manner

$$C^n \;= \varepsilon^{nab}\ [\,\partial_a B_b \,-\, \partial_b B_a\,]/2 \qquad \text{(D.8.2)}$$

where the second term is the same as the first term, since

$$-\,\varepsilon^{nab}\,\partial_b B_a = \,\varepsilon^{nba}\,\partial_b B_a \,= \varepsilon^{nab}\,\partial_a B_b\,. \qquad \text{(D.8.3)}$$

Recall now from (7.16.4) that the covariant derivative of a vector is given by

$$B_{b;a} \;=\; \partial_a B_b - \Gamma^c{}_{ab}\, B_c \qquad \text{(D.8.4)}$$

where the affine connection $\Gamma^c{}_{ab}$ is symmetric under a↔b. Therefore,

$$B_{b;a} - B_{a;b} \;=\; [\partial_a B_b - \Gamma^c{}_{ab}\, B_c] - [\partial_b B_a - \Gamma^c{}_{ba}\, B_c] = \;\partial_a B_b - \partial_b B_a\,. \qquad \text{(D.8.5)}$$

Therefore C^n can be expressed as

$$C^n \;=\; \;\varepsilon^{nab}\;\; [B_{b;a} \;- B_{a;b}\,]/2 \qquad \text{(D.8.6)}$$

so by the same ε anti-symmetry noted above the final result is

$$C^n = \varepsilon^{nab}\, B_{b;a} \quad . \qquad \text{(D.8.7)}$$

The major feature of $B_{b;a}$ -- as discussed in (7.16.5) -- is that it is a rank-2 tensor if **B** is a tensorial vector. The weight addition rule (D.2.3) can then be applied to $\varepsilon^{nab}\, B_{b;a}$. Since ε^{nab} has weight -1 and $B_{b;a}$ has weight 0, the conclusion is that C^n is a vector density of weight -1.

Thus, **C** = ∇ x **B** = curl **B** is a vector density of weight -1.

D.9 Tensor $\mathcal{E}$ as a weight 0 version of ε : three conventions

1. Equations in the "Weinberg Convention"

In this Section it is assumed as in (D.2.4) that g_{ij} and g^{ij} raise and lower indices of the ε tensor just as they do for any other tensor (Weinberg convention). Above it was established that

$$\varepsilon_{abc...} \;= g\, \varepsilon^{abc...} \qquad \varepsilon^{123..} = +1 \qquad \varepsilon_{123...} = g \qquad \textit{(D.5.10)}$$
$$\varepsilon'_{abc...} \;= g'\varepsilon'^{abc...} \qquad \varepsilon'^{123..} = +1 \qquad \varepsilon'_{123...} = g' \qquad \textit{(D.5.10)}$$

$$\varepsilon'_{abc...} \;= J^2\, \varepsilon_{abc...} \;= (g'/g)\, \varepsilon_{abc...} \qquad \textit{(D.5.13)}$$

ε is a rank-N tensor of weight W = -1 *(D.4.9) and (D.6.7)* (D.9.1)

Again, just in passing, notice that $\varepsilon'_{abc...} = J^2\, \varepsilon_{abc...}$ does *not* say ε has weight -2 because the R factors are not present on the right side.
Consider now the following new objects defined by (sg = |g|, s= sign(g) = sign(g') as in (D.1.2)),

$$\begin{aligned} \mathcal{E}_{abc...} &\equiv |g|^{-1/2}\ \varepsilon_{abc...} &&\Rightarrow\quad \mathcal{E}_{123...} = |g|^{-1/2}\ g \ = |g|^{-1/2}\, s\, |g| = \ s\, |g|^{1/2} \\ \mathcal{E}'_{abc...} &\equiv |g'|^{-1/2}\ \varepsilon'_{abc..} &&\Rightarrow\quad \mathcal{E}'_{123...} = |g'|^{-1/2}\ g' \ = |g'|^{-1/2}\, s\, |g'| = \ s\, |g'|^{1/2}\ . \end{aligned} \tag{D.9.2}$$

From (D.1.2) $g' = J^2 g$ so that g transforms as a scalar density of weight -2. Since the sign of g and g' are the same, if follows that (sg) = |g| is also a scalar density of weight -2, and then the quantity $|g|^{-1/2}$ transforms as a scalar density of weight +1, since $|g'|^{-1/2} = J^{-1}\, |g|^{-1/2}$. Looking at the equation $\mathcal{E}_{abc...} \equiv |g|^{-1/2}\ \varepsilon_{abc...}$ above, and using the weight summation rule (D.2.3), one concludes that $\mathcal{E}_{abc...}$ transforms as a tensor of weight (+1) + (-1) = 0, and so $\mathcal{E}_{abc...}$ is an ordinary rank-N tensor.

<u>Comment</u>: It was shown near (7.15.9) that a tensor density equation with matching weights is "covariant", so one is not surprised to see the second line of (D.9.2) having the same form as the first line but everything is primed (s = s').

Raising indices on both sides of (D.9.2) gives these contravariant ordinary rank-N tensors,

$$\begin{aligned} \mathcal{E}^{abc...} &\equiv |g|^{-1/2}\ \varepsilon^{abc...} &&\Rightarrow\quad \mathcal{E}^{123...} = |g|^{-1/2} \\ \mathcal{E}'^{abc...} &\equiv |g'|^{-1/2}\ \varepsilon'^{abc...} &&\Rightarrow\quad \mathcal{E}'^{123...} = |g'|^{-1/2}\ . \end{aligned} \tag{D.9.3}$$

To compare $\mathcal{E}_{abc...}$ and $\mathcal{E}^{abc...}$ we first evaluate $\mathcal{E}_{abc...}$,

$$\begin{aligned} \mathcal{E}_{abc...} &= |g|^{-1/2}\ \varepsilon_{abc...} && \text{// lower indices on both sides of 1st Eq of (D.9.3) since covariant} \\ &= |g|^{-1/2}\, g\, \varepsilon^{abc...} && \text{// from (D.9.1)} \\ &= s\, |g|^{-1/2}\, |g|\, \varepsilon^{abc...} && \text{// since } g = s|g| \text{ by (D.1.2)} \\ &= \ s\, |g|^{1/2}\, \varepsilon^{abc...} \end{aligned} \tag{D.9.4}$$

We then pair this last equation with the first of (D.9.3),

$$\begin{aligned} \mathcal{E}_{abc...} &= s\, |g|^{1/2}\, \varepsilon^{abc...} && \textit{(D.9.4)} \\ \mathcal{E}^{abc...} &= \ |g|^{-1/2}\, \varepsilon^{abc...} && \textit{(D.9.3)} \end{aligned} \tag{D.9.5}$$

Dividing the two equations gives $\mathcal{E}_{abc...} / \mathcal{E}^{abc...} = s\, |g| = g$ so that

$$\mathcal{E}_{abc...} = g\, \mathcal{E}^{abc...} \tag{D.9.6}$$

We can then construct an $\mathcal{E}\,\mathcal{E}$ product as follows

$$\mathcal{E}^{abc...}\ \mathcal{E}_{ABC...} = [\ |g|^{-1/2}\ \varepsilon^{abc...}\ ]\, [\ |g|^{-1/2}\ \varepsilon_{ABC...}] \quad \text{// (D.9.3) and 1st line (D.9.4)}$$

$$= |g|^{-1}\, \varepsilon^{abc...}\ \ \varepsilon_{ABC...}\ \ . \tag{D.9.7}$$

Next, we use standard ordering to get

$$\begin{array}{lll} \mathcal{E}^{123...} = |g|^{-1/2} & \text{// from (D.9.3)} & \\ \mathcal{E}_{123...} = s|g|^{+1/2} \,. & \text{// from (D.9.4)} & \text{(D.9.8)} \end{array}$$

Summarizing (D.9.8) and (D.9.6) and (D.9.7),

$$\begin{array}{lll} \mathcal{E}^{123...} = |g|^{-1/2} & \mathcal{E}_{123...} = s|g|^{+1/2} & \mathcal{E}_{abc...} = g\,\mathcal{E}^{abc...} = s|g|\,\mathcal{E}^{abc...} \\ \mathcal{E}'^{123...} = |g'|^{-1/2} & \mathcal{E}'_{123...} = s|g'|^{+1/2} & \end{array}$$

$$\begin{array}{ll} \mathcal{E}^{abc...}\ \mathcal{E}_{ABC...} = |g|^{-1}\,\varepsilon^{abc...}\ \varepsilon_{ABC...} & \\ \mathcal{E}'^{abc...}\ \mathcal{E}'_{ABC...} = |g'|^{-1}\,\varepsilon'^{abc...}\ \varepsilon'_{ABC...} & \text{(D.9.9)} \end{array}$$

Since $|g|^{-1}$ is a scalar density of weight +2 and each ε has weight -1 and each $\mathcal{E}$ has weight 0, one is happy to see the weights balance of the two sides of this pair of covariant equations.

2. Equations in the "added sign s convention"

The previous section shows how things work out using the "Weinberg convention" noted at the start of Section D.5. Here is the previous section redone in the "added sign s convention" where $\varepsilon_{abc...}$ has an extra sign s as shown in (D.5.2). To get from Weinberg to added-sign-s, take $\varepsilon_{abc...} \rightarrow s\varepsilon_{abc...}$. Doing this change in (D.5.10) for both $\varepsilon_{abc..}$ and $\varepsilon'_{abc..}$ gives

$$\begin{array}{lllll} \varepsilon_{abc...} = sg\varepsilon^{abc...} & \varepsilon^{123..} = +1 & \varepsilon_{123...} = sg = |g| & \text{// } g \rightarrow sg & \\ \varepsilon'_{abc...} = sg'\varepsilon'^{abc...} & \varepsilon'^{123..} = +1 & \varepsilon'_{123...} = sg' = |g'| & \text{// } g' \rightarrow sg' & \text{(D.9.10)} \end{array}$$

$$\varepsilon'_{abc...} = |J|^2\,\varepsilon_{abc...} = (g'/g)\,\varepsilon_{abc...} \qquad \varepsilon \text{ is rank-N tensor of weight } W = -1 \text{ // same}$$

Consider the following *new* objects defined by ($sg = |g|$, $s= \text{sign}(g) = \text{sign}(g')$ as in (D.1.2))

$$\begin{array}{ll} \mathcal{E}_{abc...} \equiv |g|^{-1/2}\,\varepsilon_{abc...} \Rightarrow \quad \mathcal{E}_{123...} = |g|^{-1/2}\,|g| = |g|^{1/2} & \\ \mathcal{E}'_{abc...} \equiv |g'|^{-1/2}\,\varepsilon'_{abc...} \Rightarrow \quad \mathcal{E}'_{123...} = |g'|^{-1/2}\,|g'| = |g'|^{1/2} \,. & \text{(D.9.11)} \end{array}$$

By the same argument given above, $\mathcal{E}_{abc...}$ is an ordinary covariant tensor (ie, weight = 0). However, the indices cannot be raised by g^{ij}. In this convention then one must make independent definitions of the contravariant components as follows,

$$\begin{array}{ll} \mathcal{E}^{abc...} \equiv |g|^{-1/2}\,\varepsilon^{abc...} \Rightarrow \quad \mathcal{E}^{123...} = |g|^{-1/2} & \\ \mathcal{E}'^{abc...} \equiv |g'|^{-1/2}\,\varepsilon'^{abc...} \Rightarrow \quad \mathcal{E}'^{123...} = |g'|^{-1/2} & \text{(D.9.12)} \end{array}$$

To compare $\mathcal{E}_{abc...}$ and $\mathcal{E}^{abc...}$,

$$\begin{aligned}
&\mathcal{E}_{\mathbf{abc}...} = |g|^{-1/2}\ \varepsilon_{\mathbf{abc}...} = |g|^{-1/2}\,|g|\ \varepsilon^{\mathbf{abc}...} = |g|^{-1/2}\,|g|\ \varepsilon^{\mathbf{abc}...} \\
&\mathcal{E}^{\mathbf{abc}...} = |g|^{-1/2}\ \varepsilon^{\mathbf{abc}...}
\end{aligned} \qquad \text{(D.9.13)}$$

so that

$$\mathcal{E}_{\mathbf{abc}...} = |g|\ \mathcal{E}^{\mathbf{abc}...} \qquad \text{(D.9.14)}$$

Summarizing,

$$\begin{aligned}
&\mathcal{E}^{123...} = |g|^{-1/2} \qquad && \mathcal{E}_{123...} = |g|^{+1/2} \qquad && \mathcal{E}_{\mathbf{abc}...} = |g|\ \mathcal{E}^{\mathbf{abc}...} \\
&\mathcal{E}'_{123...} = |g'|^{-1/2} && \mathcal{E}'_{123...} = |g'|^{+1/2} && \mathcal{E}'_{\mathbf{abc}...} = |g'|\ \mathcal{E}'^{\mathbf{abc}...}
\end{aligned}$$

$$\begin{aligned}
&\mathcal{E}^{\mathbf{abc}...}\ \mathcal{E}_{\mathbf{ABC}...} = |g|^{-1}\ \varepsilon^{\mathbf{abc}...}\ \varepsilon_{\mathbf{ABC}...} \\
&\mathcal{E}'^{\mathbf{abc}...}\ \mathcal{E}'_{\mathbf{ABC}...} = |g'|^{-1}\ \varepsilon'^{\mathbf{abc}...}\ \varepsilon'_{\mathbf{ABC}...}
\end{aligned} \qquad \text{(D.9.15)}$$

In this "added sign s" convention, all these summarized results involve only |g| and there are no factors of s floating around. The cost of this benefit is a lack of true covariance (when s = -1), as demonstrated in Section D.11 below.

3. Equations in the "Ricci-Levi-Civita convention"

Ricci and Levi-Civita use the "added s convention" but add a factor σ = sign(det(S)) into their definition of $\mathcal{E}$ (see their paper p 135 or Hermann pp 31-21) so that

$$\begin{aligned}
&\mathcal{E}_{\mathbf{abc}...} \equiv \sigma|g|^{-1/2}\ \varepsilon_{\mathbf{abc}...} && \Rightarrow && \mathcal{E}_{123...} = \sigma|g|^{-1/2}\,|g| = \sigma|g|^{1/2} \\
&\mathcal{E}'_{\mathbf{abc}...} \equiv \sigma|g'|^{-1/2}\ \varepsilon'_{\mathbf{abc}...} && \Rightarrow && \mathcal{E}'_{123...} = \sigma|g'|^{-1/2}\,|g'| = \sigma|g'|^{1/2} \\
\\
&\mathcal{E}^{\mathbf{abc}...} \equiv \sigma|g|^{-1/2}\ \varepsilon^{\mathbf{abc}..} && \Rightarrow && \mathcal{E}^{123...} = \sigma|g|^{-1/2} \\
&\mathcal{E}'^{\mathbf{abc}...} \equiv \sigma|g'|^{-1/2}\ \varepsilon'^{\mathbf{abc}...} && \Rightarrow && \mathcal{E}'^{123...} = \sigma|g'|^{-1/2}
\end{aligned} \qquad \text{(D.9.16)}$$

Summarizing,

$$\begin{aligned}
&\mathcal{E}^{123...} = \sigma|g|^{-1/2} \qquad && \mathcal{E}_{123...} = \sigma|g|^{+1/2} \qquad && \mathcal{E}_{\mathbf{abc}...} = |g|\ \mathcal{E}^{\mathbf{abc}...} \\
&\mathcal{E}'_{123...} = \sigma|g'|^{-1/2} && \mathcal{E}'_{123...} = \sigma|g'|^{+1/2} && \mathcal{E}'_{\mathbf{abc}...} = |g'|\ \mathcal{E}'^{\mathbf{abc}...}
\end{aligned}$$

$$\begin{aligned}
&\mathcal{E}^{\mathbf{abc}...}\ \mathcal{E}_{\mathbf{ABC}...} = |g|^{-1}\ \varepsilon^{\mathbf{abc}...}\ \varepsilon_{\mathbf{ABC}...} \\
&\mathcal{E}'^{\mathbf{abc}...}\ \mathcal{E}'_{\mathbf{ABC}...} = |g'|^{-1}\ \varepsilon'^{\mathbf{abc}...}\ \varepsilon'_{\mathbf{ABC}...}
\end{aligned} \qquad \text{(D.9.17)}$$

Notice that in all three conventions, the last equation pair is the same.

Since Ricci and Levi-Civita did not raise and lower individual indices in their 1900 paper, they were not concerned about their convention being non-covariant in that sense.

D.10 Representation of ε, εε and contracted εε as determinants

This section is in developmental notation and ε is the permutation tensor.

1. Theorem about a certain permutation sum

Consider the following object Q defined as a signed permutation sum of the product of N matrix elements of a matrix M_{ij},

$$Q_{abc..x} \equiv \Sigma_P \ p\ P_2(M_{a1}M_{b2}\ M_{c3}.....M_{xN}) \qquad \text{(D.10.1)}$$

In this equation, P_2 represents a permutation of the set of 2nd indices of the N matrix elements, and the sum is over all N! such permutations.

There are many ways to arrive at a given permutation of 123...N by doing pairwise swaps, but for all these ways, the number of swaps S will be either even or odd. The parity p of a permutation is defined then as $(-1)^S$ and this p appears in the above sum.

If one were to swap indices 2↔3 on the right above, each permutation would have S→ S+1 since an extra swap is needed to undo 2↔3. Thus, all parities p → -p and in fact the whole object negates. But the swap 2↔3 is the same as b↔c since $M_{b3}\ M_{c2} = M_{c2}\ M_{b3}$. Applying this argument to any pair of indices, one concludes that $Q_{abc..x}$ is totally antisymmetric and therefore by (D.3.1) can be written as K $\varepsilon_{abc...x}$:

$$\Sigma_P \ p\ P_2(M_{a1}M_{b2}\ M_{c3}.....M_{xN}) = K\ \varepsilon_{abc...x} \ . \qquad \text{(D.10.2)}$$

where in this Section $\varepsilon_{abc...x}$ is just the permutation tensor (not the Levi-Civita tensor).

Setting abc..x to 123..N, one gets.

$$\Sigma_P \ p\ P_2(M_{11}M_{22}\ M_{33}.....M_{NN}) = K \ . \qquad \text{(D.10.3)}$$

The left side of this last equation can be written as

$$\Sigma_P \ p\ P_2(M_{11}M_{22}\ M_{33}.....M_{NN}) = \Sigma_{abc..x}\ \varepsilon_{abc...x}\ M_{1a}M_{2b}\ M_{3c}.....M_{Nx} \qquad \text{(D.10.4)}$$

because p = $\varepsilon_{abc...x}$ correctly assesses the parity of any given permutation. But this object is simply det(M) so the conclusion is that K = det(M) and then

$$\Sigma_P \ p\ P_2(M_{a1}M_{b2}\ M_{c3}.....M_{xN}) = \det(M)\ \varepsilon_{abc...x} \ . \qquad \text{(D.10.5)}$$

Consider now the following matrix where abc..x is some permutation of 123...x,

$$M^{(abc..)} = \begin{matrix} M_{a1} & M_{a2} & M_{a3} & ... & M_{aN} \\ M_{b1} & M_{b2} & M_{b3} & ... & M_{bN} \\ M_{c1} & M_{c2} & M_{c3} & ... & M_{cN} \\ ... \\ M_{x1} & M_{x2} & M_{x3} & ... & M_{xN} \end{matrix} \ . \qquad \text{(D.10.6)}$$

By rearranging the rows into their normal numerical order, one obtains matrix M, but incurs a sign from the various row swaps which sign is just $\varepsilon_{abc..x}$. Therefore,

$$\det(M^{(abc..)}) = \varepsilon_{abc...x} \det(M) \tag{D.10.7}$$

and therefore we obtain the following "theorem" :

$$\Sigma_P \; p\, P_2(M_{a1}M_{b2}\, M_{c3}.....M_{xN}) = \det(M^{(abc..)}) = \det(M)\, \varepsilon_{abc...x} \;. \tag{D.10.8}$$

The permutation sum is thus just the determinant of matrix $M^{(abc..)}$. The first term in the permutation sum, the term with an identity permutation, corresponds to the product of the diagonals of $M^{(abc..)}$.

2. Application of the theorem to M = δ: a representation of ε

Apply the above theorem to matrix $M = 1 \equiv \delta$, the identity matrix, so $M_{ij} = \delta_{i,j}$. Clearly $\det(\delta) = 1$ and one then has

$$\Sigma_P \; p\, P_2(\delta_{a,1}\delta_{b,2}\, \delta_{c,3}.....\delta_{x,N}) = \det[\delta^{(abc..)}] = \varepsilon_{abc...x} \;. \tag{D.10.9}$$

Thus is obtained a well-known representation of $\varepsilon_{abc..x}$ as a certain determinant of Kronecker deltas,

$$\varepsilon_{abc...x} = \det[\delta^{(abc..)}]$$

where
$$\delta^{(abc..)} = \begin{array}{lllll} \delta_{a,1} & \delta_{a,2} & \delta_{a,3} & ... \; \delta_{a,N} & = \mathbf{R_a} \\ \delta_{b,1} & \delta_{b,2} & \delta_{b,3} & ... \; \delta_{b,N} & = \mathbf{R_b} \\ \delta_{c,1} & \delta_{c,2} & \delta_{c,3} & ... \; \delta_{c,N} & = \mathbf{R_c} \\ ... & & & & \\ \delta_{x,1} & \delta_{x,2} & \delta_{x,3} & ... \; \delta_{x,N} & = \mathbf{R_x} \end{array} \tag{D.10.10}$$

For future use, each row vector has been given a name like $\mathbf{R_a}$ where $(\mathbf{R_a})_i = \delta_{a,i}$.

The conclusion then is that (vertical bars here mean determinant) :

$$\varepsilon_{abc...x} = \begin{vmatrix} \delta_{a,1} & \delta_{a,2} & \delta_{a,3} & ... & \delta_{a,N} \\ \delta_{b,1} & \delta_{b,2} & \delta_{b,3} & ... & \delta_{b,N} \\ \delta_{c,1} & \delta_{c,2} & \delta_{c,3} & ... & \delta_{c,N} \\ ... & & & & \\ \delta_{x,1} & \delta_{x,2} & \delta_{x,3} & ... & \delta_{x,N} \end{vmatrix} = |\delta^{(abc..)}| \tag{D.10.11}$$

which is the same as

$$\varepsilon_{abc...x} = \Sigma_P \; p\, P_2(\delta_{a,1}\delta_{b,2}\, \delta_{c,3}.....\delta_{x,N}) \;. \tag{D.10.12}$$

3. Outer product of two ε tensors.

Consider now

$$\varepsilon_{\mathbf{abc...x}} = \det\begin{pmatrix}\mathbf{R_a}\\ \mathbf{R_b}\\ ...\\ \mathbf{R_x}\end{pmatrix} \qquad \text{and} \qquad \varepsilon_{\mathbf{a'b'c'...x'}} = \det\begin{pmatrix}\mathbf{R_{a'}}\\ \mathbf{R_{b'}}\\ ...\\ \mathbf{R_{x'}}\end{pmatrix} \quad . \tag{D.10.13}$$

Then

$$\varepsilon_{\mathbf{abc...x}}\,\varepsilon_{\mathbf{a'b'c'...x'}} = \det\begin{pmatrix}\mathbf{R_a}\\ \mathbf{R_b}\\ ...\\ \mathbf{R_x}\end{pmatrix} \det\begin{pmatrix}\mathbf{R_{a'}}\\ \mathbf{R_{b'}}\\ ...\\ \mathbf{R_{x'}}\end{pmatrix} = \det\begin{pmatrix}\mathbf{R_a}\\ \mathbf{R_b}\\ ...\\ \mathbf{R_x}\end{pmatrix} \det\,(\,\mathbf{R_{a'}}\ \mathbf{R_{b'}}\ ...\ \mathbf{R_{x'}})$$

$$= \det\left\{\begin{pmatrix}\mathbf{R_a}\\ \mathbf{R_b}\\ ...\\ \mathbf{R_x}\end{pmatrix}(\,\mathbf{R_{a'}}\ \mathbf{R_{b'}}\ ...\ \mathbf{R_{x'}})\right\} \tag{D.10.14}$$

which is the determinant of this matrix

$$\begin{array}{llll}
\mathbf{R_a}\bullet\mathbf{R_{a'}} & \mathbf{R_a}\bullet\mathbf{R_{b'}} & \mathbf{R_a}\bullet\mathbf{R_{c'}} & ...\ \mathbf{R_a}\bullet\mathbf{R_{x'}}\\
\mathbf{R_b}\bullet\mathbf{R_{a'}} & \mathbf{R_b}\bullet\mathbf{R_{b'}} & \mathbf{R_b}\bullet\mathbf{R_{c'}} & ...\ \mathbf{R_b}\bullet\mathbf{R_{x'}}\\
\mathbf{R_c}\bullet\mathbf{R_{a'}} & \mathbf{R_c}\bullet\mathbf{R_{b'}} & \mathbf{R_c}\bullet\mathbf{R_{c'}} & ...\ \mathbf{R_c}\bullet\mathbf{R_{x'}}\\
... & & & \\
\mathbf{R_x}\bullet\mathbf{R_{a'}} & \mathbf{R_x}\bullet\mathbf{R_{b'}} & \mathbf{R_x}\bullet\mathbf{R_{c'}} & ...\ \mathbf{R_x}\bullet\mathbf{R_{x'}}\ .
\end{array} \tag{D.10.15}$$

A typical element of this matrix is given by

$$\mathbf{R_c}\bullet\mathbf{R_{b'}} = (\mathbf{R_c})_\mathbf{i}(\mathbf{R_{b'}})_\mathbf{i} = \delta_{\mathbf{c,i}}\,\delta_{\mathbf{b',i}} = \delta_{\mathbf{c,b'}} \tag{D.10.16}$$

so that matrix can be written as

$$\begin{array}{llll}
\delta_{\mathbf{a,a'}} & \delta_{\mathbf{a,b'}} & \delta_{\mathbf{a,c'}} & ....\ \delta_{\mathbf{a,x'}}\\
\delta_{\mathbf{b,a'}} & \delta_{\mathbf{b,b'}} & \delta_{\mathbf{b,c'}} & ....\ \delta_{\mathbf{b,x'}}\\
\delta_{\mathbf{c,a'}} & \delta_{\mathbf{c,b'}} & \delta_{\mathbf{c,c'}} & ....\ \delta_{\mathbf{c,x'}}\\
.... & & & \\
\delta_{\mathbf{x,a'}} & \delta_{\mathbf{x,b'}} & \delta_{\mathbf{x,c'}} & ....\ \delta_{\mathbf{x,x'}}
\end{array} \qquad \equiv\ \delta^{(\mathbf{abc..x;\ a'b'c'..x'})} \tag{D.10.17}$$

where we have made up a name for this matrix as shown.

The conclusion then is that

$$\varepsilon_{abc...x}\,\varepsilon_{a'b'c'...x'} = \begin{vmatrix} \delta_{a,a'} & \delta_{a,b'} & \delta_{a,c'} & .... & \delta_{a,x'} \\ \delta_{b,a'} & \delta_{b,b'} & \delta_{b,c'} & .... & \delta_{b,x'} \\ \delta_{c,a'} & \delta_{c,b'} & \delta_{c,c'} & .... & \delta_{c,x'} \\ .... & & & & \\ \delta_{x,a'} & \delta_{x,b'} & \delta_{x,c'} & .... & \delta_{x,x'} \end{vmatrix} \equiv \left|\delta^{(abc..x;\ a'b'c'..x')}\right|$$

(D.10.18)

which is the same as

$$\varepsilon_{abc...x}\,\varepsilon_{a'b'c'...x'} = \Sigma_P\ p\ P_2(\delta_{a,a'}\delta_{b,b'}\delta_{c,c'}.....\delta_{x,x'}) \qquad \text{(D.10.19)}$$

because the right side of (D.10.19) is precisely the determinant appearing in (D.10.18). As usual, the argument of P_2 is the product of the diagonal elements of the matrix of interest.

4. Contracting the first index of the outer product of two ε tensors.

Consider what happens if one sums on the first index of the εε product:

$$\Sigma_a\,\varepsilon_{abc...x}\,\varepsilon_{ab'c'...x'} \quad . \qquad \text{(D.10.20)}$$

For fixed given values of bc..x and b'c'...x' , there is only one way this sum can be non-zero. In that one way, bc..x and b'c'...x' must each be permutations of the set {12..N exclude A} where A is the "hit value" of a in the sum on a. In the sum, only a = A contributes. Then using this hit value A and (D.10.19),

$$\Sigma_a\,\varepsilon_{abc...x}\,\varepsilon_{ab'c'...x'} = \varepsilon_{Abc...x}\,\varepsilon_{Ab'c'...x'}$$

$$= \Sigma_P\ p\ P_2(\delta_{A,A}\delta_{b,b'}\delta_{c,c'}.....\delta_{x,x'}) = \Sigma_P\ p\ P_2(\delta_{b,b'}\delta_{c,c'}.....\delta_{x,x'}) \qquad \text{(D.10.21)}$$

where in this last expression the sum can be regarded as being over permutations where b'c'...x' is a permutation of b,c..x. Each of these lists of integers is in turn a permutation of {12..N exclude A}. *Now*, parity p = $(-1)^S$ where S is a number of swaps it takes to connect b'c'...x' with b,c..x, since a = a' = A. One might wonder if the overall sign of the RHS of the last equation is correct. A check of the first term in this sum which is just $\delta_{b,b'}\delta_{c,c'}.....\delta_{x,x'}$ shows that this overall sign is indeed correct. This first term must be positive because the product of two ε's is either +1 or 0. As an example,

$$\Sigma_a\,\varepsilon_{abc}\,\varepsilon_{ab'c'} = \Sigma_P\ p\ P_2(\delta_{b,b'}\delta_{c,c'}) = \delta_{b,b'}\delta_{c,c'} - \delta_{b,c'}\delta_{c,b'}\ . \qquad \text{(D.10.22)}$$

The permutation sum shown on the right side of (D.10.21) is the determinant of $\delta^{(abc..x;\ a'b'c'..x')}$ but with the first row and column crossed out. It can then be thought of as either the minor or cofactor of the element aa of this big δ matrix. Therefore,

$$\Sigma_a\,\varepsilon_{abc...x}\,\varepsilon_{ab'c'...x'} = [\mathrm{cof}\ \delta^{(abc..x;\ a'b'c'..x')}]_{aa} \qquad \text{(D.10.23)}$$

where the notation cofM refers to a matrix of cofactors with elements $[\mathrm{cofM}]_{ij}$. Don't confuse the a on the right side with the local dummy summation index a on the left side.

The conclusion then is that (implied summation on a on the LHS)

$$\varepsilon_{abc...x}\,\varepsilon_{ab'c'...x'} = \begin{vmatrix} \delta_{b,b'} & \delta_{b,c'} & \cdots\ \delta_{b,x'} \\ \delta_{c,b'} & \delta_{c,c'} & \cdots\ \delta_{c,x'} \\ & & \\ \delta_{x,b'} & \delta_{x,c'} & \cdots\ \delta_{x,x'} \end{vmatrix} = [\mathrm{cof}\,\delta^{(abc..x;\ a'b'c'..x')}]_{aa}$$

(D.10.24)

5. Contracting two or more indices of the outer product of two ε tensors.

Consider what happens if one sums on the first two indices of the εε product:

$$\Sigma_{a,b}\,\varepsilon_{abcd...x}\,\varepsilon_{abc'd'...x'} \quad . \tag{D.10.25}$$

For fixed given values of c,d..x and c'd'...x' , in order for this double sum to be non-zero, the index sets cd..x and c'd'...x' must each be permutations of the set {12..N exclude A,B} where A,B are a pair of hit values for the a and b sums. If a=A and b=B is a hit value, then so is a=B and a=A, so there are 2! contributing terms in the sum, and each term is +1. Therefore

$$\Sigma_{a,b}\,\varepsilon_{abcd...x}\,\varepsilon_{abc'd'...x'} = 2!\ \varepsilon_{ABc...x}\,\varepsilon_{ABc'...x'}$$

$$= 2!\ \Sigma_P\ p\ P_2(\delta_{A,A}\delta_{B,B'}\delta_{c,c'}\delta_{d,d'}.....\delta_{x,x'}) = 2!\Sigma_P\ p\ P_2(\delta_{c,c'}\ \delta_{d,d'}.....\delta_{x,x'}) \tag{D.10.26}$$

where in this last expression the sum is over permutations where c'd'...x' is a permutation of c,d....x. Each of these lists of integers is in turn a permutation of {12..N exclude A,B}. Now parity p = $(-1)^S$ where S is a number of swaps it takes to connect c'd'...x' with c,d....x. Since the product of two ε's is either +1 or 0, the overall sign of the right side shown must be correct. As an example,

$$\Sigma_{a,b}\,\varepsilon_{abc}\,\varepsilon_{abc'} = 2!\Sigma_P\ p\ P_2(\delta_{c,c'}) = 2\ \delta_{c,c'} \quad . \tag{D.10.27}$$

If c = c' = 2, then this says

$$\Sigma_{a,b}\,\varepsilon_{ab2}\,\varepsilon_{ab2} = \varepsilon_{132}\,\varepsilon_{132} + \varepsilon_{312}\,\varepsilon_{312} = 1 + 1 = 2 \quad . \tag{D.10.28}$$

The permutation sum shown on the right side of (D.10.26) is the determinant of $\delta^{(abc..x;\ a'b'c'..x')}$ but with the first 2 rows and columns crossed out. Therefore,

$$\Sigma_{a,b}\,\varepsilon_{abcd...x}\,\varepsilon_{abc'd'...x'} = 2!\ \{\ [\mathrm{cof}\,\delta^{(abc..x;\ a'b'c'..x')}]_{aa}\}_{bb} \quad . \tag{D.10.29}$$

The conclusion then is that (implied summation on a,b on the LHS)

$$\varepsilon_{\mathtt{abcd...x}}\,\varepsilon_{\mathtt{abc'd'...x'}} = 2! \begin{vmatrix} \delta_{\mathtt{c,c'}} & .... & \delta_{\mathtt{c,x'}} \\ .... & & \\ \delta_{\mathtt{x,c'}} & .... & \delta_{\mathtt{x,x'}} \end{vmatrix} = 2!\ \{\ [\mathrm{cof}\,\delta^{(\mathtt{abc..x;\ a'b'c'..x'})}]_{\mathtt{aa}}\}_{\mathtt{bb}}$$

(D.10.30)

Here a and b on the left are just dummy summation indices, whereas a and b on the right indicate that the right side is 2! {...} where {...} is the cofactor of a matrix which is the full $\delta^{(\mathtt{abc..x;\ a'b'c'..x'})}$ matrix but with the first two rows and columns crossed out. One can imagine doing this crossing out in the picture of the full δ matrix shown in (D.10.18). The diagonal element of the first row and column is $\delta_{\mathtt{a,a'}}$ and of the second row and column $\delta_{\mathtt{b,b'}}$ so we use the notation aa and bb to denote these rows and columns which are crossed out.

This pattern continues as more indices are contracted. If three indices a,b,c are contracted, there will then be 3! hit values which are A,B,C and its permutations, and one just repeats the above discussion. The result will then be

$$\Sigma_{\mathtt{a,b,c}}\,\varepsilon_{\mathtt{abcd...x}}\,\varepsilon_{\mathtt{abcd'...x'}} = 3!\ \{\{\ [\mathrm{cof}\,\delta^{(\mathtt{abc..x;\ a'b'c'..x'})}]_{\mathtt{aa}}\}_{\mathtt{bb}}\}_{\mathtt{cc}}\ . \qquad \text{(D.10.31)}$$

The conclusion then is that (implied summation on a,b,c on the LHS)

$$\varepsilon_{\mathtt{abcd...x}}\,\varepsilon_{\mathtt{abcd'...x'}} = 3! \begin{vmatrix} \delta_{\mathtt{d,d'}} & .... & \delta_{\mathtt{d,x'}} \\ .... & & \\ \delta_{\mathtt{x,d'}} & .... & \delta_{\mathtt{x,x'}} \end{vmatrix} = 3!\ \{\{\ [\mathrm{cof}\,\delta^{(\mathtt{abc..x;\ a'b'c'..x'})}]_{\mathtt{aa}}\}_{\mathtt{bb}}\}_{\mathtt{cc}}$$

(D.10.32)

Eventually one arrives at a point where all but one of the indices are summed, so that

$$\varepsilon_{\mathtt{abcd...x}}\,\varepsilon_{\mathtt{abcd...x'}} = (N\text{-}1)!\ |\delta_{\mathtt{x,x'}}| = (N\text{-}1)!\ \delta_{\mathtt{x,x'}} \qquad \text{(D.10.33)}$$

an example being

$$\varepsilon_{\mathtt{abc2}}\,\varepsilon_{\mathtt{abc2}} = 3!\ \delta_{\mathtt{22}} = 3! \quad = \varepsilon_{\mathtt{1342}}\,\varepsilon_{\mathtt{1342}} + \varepsilon_{\mathtt{1432}}\,\varepsilon_{\mathtt{1432}} \ + 4 \text{ more terms} = 1+1+4 = 6\ . \qquad \text{(D.10.34)}$$

The final point is that at which *all* indices are summed, with result

$$\varepsilon_{\mathtt{abcd...x}}\,\varepsilon_{\mathtt{abcd...x}} = N! \qquad \text{(D.10.35)}$$

and example of which is

$$\varepsilon_{\mathtt{abc}}\varepsilon_{\mathtt{abc}} = \varepsilon_{\mathtt{123}}^{\mathtt{2}} + \varepsilon_{\mathtt{213}}^{\mathtt{2}} + 4 \text{ more terms } = 1 + 1 + 4 = 6\ . \qquad \text{(D.10.36)}$$

6. Summary of Results

$$\varepsilon_{\mathbf{abc...x}}\,\varepsilon_{\mathbf{a'b'c'...x'}} = \begin{vmatrix} \delta_{\mathbf{a,a'}} & \delta_{\mathbf{a,b'}} & \delta_{\mathbf{a,c'}} & \ldots\ \delta_{\mathbf{a,x'}} \\ \delta_{\mathbf{b,a'}} & \delta_{\mathbf{b,b'}} & \delta_{\mathbf{b,c'}} & \ldots\ \delta_{\mathbf{b,x'}} \\ \delta_{\mathbf{c,a'}} & \delta_{\mathbf{c,b'}} & \delta_{\mathbf{c,c'}} & \ldots\ \delta_{\mathbf{c,x'}} \\ \ldots & & & \\ \delta_{\mathbf{x,a'}} & \delta_{\mathbf{x,b'}} & \delta_{\mathbf{x,c'}} & \ldots\ \delta_{\mathbf{x,x'}} \end{vmatrix} \equiv \left|\delta^{\mathbf{(abc..x;\ a'b'c'..x')}}\right|$$

$$\varepsilon_{\mathbf{abc...x}}\,\varepsilon_{\mathbf{ab'c'...x'}} = \begin{vmatrix} \delta_{\mathbf{b,b'}} & \delta_{\mathbf{b,c'}} & \ldots\ \delta_{\mathbf{b,x'}} \\ \delta_{\mathbf{c,b'}} & \delta_{\mathbf{c,c'}} & \ldots\ \delta_{\mathbf{c,x'}} \\ & & \\ \delta_{\mathbf{x,b'}} & \delta_{\mathbf{x,c'}} & \ldots\ \delta_{\mathbf{x,x'}} \end{vmatrix} = [\mathrm{cof}\,\delta^{\mathbf{(abc..x;\ a'b'c'..x')}}]_{\mathbf{aa}}$$

$$\varepsilon_{\mathbf{abcd...x}}\,\varepsilon_{\mathbf{abc'd'...x'}} = 2!\begin{vmatrix} \delta_{\mathbf{c,c'}} & \ldots\ \delta_{\mathbf{c,x'}} \\ \ldots\ldots & \\ \delta_{\mathbf{x,c'}} & \ldots\ \delta_{\mathbf{x,x'}} \end{vmatrix} = 2!\ \{\,[\mathrm{cof}\,\delta^{\mathbf{(abc..x;\ a'b'c'..x')}}]_{\mathbf{aa}}\}_{\mathbf{bb}}$$

$$\varepsilon_{\mathbf{abcd...x}}\,\varepsilon_{\mathbf{abcd'...x'}} = 3!\begin{vmatrix} \delta_{\mathbf{d,d'}} & \ldots\ \delta_{\mathbf{d,x'}} \\ \ldots\ldots & \\ \delta_{\mathbf{x,d'}} & \ldots\ \delta_{\mathbf{x,x'}} \end{vmatrix} = 3!\ \{\{\,[\mathrm{cof}\,\delta^{\mathbf{(abc..x;\ a'b'c'..x')}}]_{\mathbf{aa}}\}_{\mathbf{bb}}\}_{\mathbf{cc}}$$

• • • •

$$\varepsilon_{\mathbf{abcd...x}}\,\varepsilon_{\mathbf{abcd...x'}} = (N\text{-}1)!\ \delta_{\mathbf{x,x'}}$$

$$\varepsilon_{\mathbf{abcd...x}}\,\varepsilon_{\mathbf{abcd...x}} = N! \qquad \text{(D.10.37)}$$

The determinants appearing above are sometimes called **generalized Kronecker deltas**. For example, here is a sample determinant followed by our (easy to type) notation and then the official Kronecker delta notation:

$$\begin{vmatrix} \delta_{\mathbf{c,c'}} & \ldots\ \delta_{\mathbf{c,x'}} \\ \ldots\ldots & \\ \delta_{\mathbf{x,c'}} & \ldots\ \delta_{\mathbf{x,x'}} \end{vmatrix} \equiv \delta^{\mathbf{(cd...x;\ c',d'...x')}} \equiv \delta^{\mathbf{c\,d}\ \ \ \ \mathbf{x}}_{\mathbf{c'\,d'}\cdots\,\mathbf{x'}} \qquad \text{(D.10.38)}$$

Often the deltas are written in the form $\delta_{\mathbf{c,c'}} = \delta^{\mathbf{c}}{}_{\mathbf{c'}}$ or $\delta^{\mathbf{c}}_{\mathbf{c'}}$

Using both our notation and the generalized Kronecker delta notation, we rewrite (D.10.37) this way

$$\varepsilon_{abc..x}\varepsilon_{a'b'c'..x'} = \delta^{(abc...x;\ a'b'c'...x')} = \delta^{a\ b\ c\ \dots\ x}_{a'\ b'\ c'\ \dots\ x'}$$

$$\varepsilon_{abc..x}\varepsilon_{ab'c'..x'} = \delta^{(bcd...x;\ b'c'd'...x')} = \delta^{b\ c\ \dots\ x}_{b'\ c'\ \dots\ x'}$$

$$\varepsilon_{abcd..x}\varepsilon_{abc'd..x'} = 2!\ \delta^{(cd...x;\ c'd'...x')} = 2!\ \delta^{c\ d\ \dots\ x}_{c'\ d'\dots\ x'}$$

$$\varepsilon_{abcde..x}\varepsilon_{abcd'e'..x'} = 3!\ \delta^{(de...x;\ d'e'...x')} = 3!\ \delta^{d\ e\ \dots\ x}_{d'\ e'\dots\ x'}$$

•••••

$$\varepsilon_{abcd...x}\ \varepsilon_{abcd...x'} = (N-1)!\delta^{(x,x')} = (N-1)!\delta^{x}_{x'} = (N-1)!\ \delta_{x,x'}$$

$$\varepsilon_{abcd...x}\ \varepsilon_{abcd...x} = N! \qquad (N = \text{dimension of x-space}) \qquad (D.10.39)$$

Keep in mind that in all of Section D.10, ε has represented the permutation tensor.

D.11 Covariant forms of the previous Section results

The Section D.11 results were in effect all developed in Cartesian x-space where up and down indices on the ε's did not matter and we regarded ε as the permutation tensor. The rules for converting any of these results to covariant form are as follows: [after conversion, ε becomes the Levi-Civita tensor where $\varepsilon^{ab...}$ is the only index position where it is the same as the permutation tensor]

Weinberg convention: (D.11.1)

• Make the replacement $\varepsilon_{*****}\ \varepsilon_{*****} \to |g|^{-1}\ \varepsilon^{*****}\ \varepsilon_{*****} = \mathcal{E}^{*****}\ \mathcal{E}_{*****}$ from (D.9.9). In this way we "continue off" the Cartesian equation (where g = 1) to obtain a covariant form. The objects $\mathcal{E}^{*****}$ and $\mathcal{E}_{*****}$ are true contravariant and covariant (weight = 0) rank-N tensor. This is vaguely reminiscent of "analytically continuing" a real function f(x) "off" the real axis to get a function of a complex variable f(z). When evaluated on the real axis, f(z) and f(x) are the same. See the tensorization discussion in Section 15.2 for more detail.

• Write the right side of the expression of interest replacing every $\delta_{a,b} = \delta^a{}_b = g^a{}_b$ as shown in (7.4.19). Then the right side will also be a true tensor.

Example 1: For g = 1, consider the εε product with no summed indices [first line of (D.10.37)] for N = 2:

$$\varepsilon_{ab}\varepsilon_{a'b'} = \begin{vmatrix} \delta_{aa'} & \delta_{ab'} \\ \delta_{ba'} & \delta_{bb'} \end{vmatrix} = \delta_{a,a'}\ \delta_{b,b'} - \delta_{a,b'}\ \delta_{b,a'}\ . \qquad // \ \varepsilon = \text{permutation tensor} \qquad (D.11.2)$$

The covariant form is as follows, where now g is some arbitrary metric tensor for x-space,

$$\mathcal{E}^{ab}\mathcal{E}_{a'b'} = |g|^{-1}\,\varepsilon^{ab}\varepsilon_{a'b'} = \begin{vmatrix} g^{a}{}_{a'} & g^{a}{}_{b'} \\ g^{b}{}_{a'} & g^{b}{}_{b'} \end{vmatrix} = g^{a}{}_{a'}\,g^{b}{}_{b'} - g^{a}{}_{b'}\,g^{b}{}_{a'} \quad . \tag{D.11.3}$$

The equation in x'-space would then be

$$\mathcal{E}'^{ab}\mathcal{E}'_{a'b'} = |g'|^{-1}\,\varepsilon'^{ab}\varepsilon'_{a'b'} = \begin{vmatrix} g'^{a}{}_{b} & g'^{a}{}_{b'} \\ g'^{a'}{}_{b} & g'^{a'}{}_{b'} \end{vmatrix} = g'^{a}{}_{b}\,g'^{a}{}_{b'} - g'^{a'}{}_{b}\,g'^{a'}{}_{b'} \tag{D.11.4}$$

because true tensor equations are "covariant" as discussed in Section 7.15. One can raise and lower individual indices to get for example these valid tensor equations which are 3 members of the family of 4! = 24 tensor equations obtained by raising and lowering indices :

$$\mathcal{E}^{ab}\mathcal{E}_{a'b'} = |g|^{-1}\,\varepsilon^{ab}\varepsilon_{a'b'} = \begin{vmatrix} g^{a}{}_{a'} & g^{a}{}_{b'} \\ g^{b}{}_{a'} & g^{b}{}_{b'} \end{vmatrix} = g^{a}{}_{a'}\,g^{b}{}_{b'} - g^{a}{}_{b'}\,g^{b}{}_{a'} \tag{D.11.5}$$

$$\mathcal{E}^{a}{}_{b}\mathcal{E}_{a'b'} = |g|^{-1}\,\varepsilon^{a}{}_{b}\varepsilon_{a'b'} = \begin{vmatrix} g^{a}{}_{a'} & g^{a}{}_{b'} \\ g_{ba'} & g_{bb'} \end{vmatrix} = g^{a}{}_{a'}\,g_{bb'} - g^{a}{}_{b'}\,g_{ba'} \tag{D.11.6}$$

$$\mathcal{E}_{ab}\mathcal{E}_{a'b'} = |g|^{-1}\,\varepsilon_{ab}\varepsilon_{a'b'} = \begin{vmatrix} g_{aa'} & g_{ab'} \\ g_{ba'} & g_{bb'} \end{vmatrix} = g_{aa'}\,g_{bb'} - g_{ab'}\,g_{ba'} \tag{D.11.7}$$

and of course in x'-space the equations are the same but everything is primed.

For the Levi-Civita ε tensor, we then have this covariant version of (D.10.39) :

$$\begin{array}{lll}
\mathcal{E}^{abc..x}\mathcal{E}_{a'b'c'..x'} & = |g|^{-1}\varepsilon^{abc..x}\varepsilon_{a'b'c'..x'} & = g(^{abc...x;}\ {}_{a'b'c'...x'}) \\
\mathcal{E}^{abc..x}\mathcal{E}_{ab'c'..x'} & = |g|^{-1}\varepsilon^{abc..x}\varepsilon_{ab'c'..x'} & = \ g(^{bcd...x;}\ {}_{b'c'd'...x'}) \\
\mathcal{E}^{abcd..x}\mathcal{E}_{abc'd..x'} & = |g|^{-1}\varepsilon^{abcd..x}\varepsilon_{abc'd..x'} & = \ 2!\ g(^{cd...x;}\ {}_{c'd'...x'}) \\
\mathcal{E}^{abcde..x}\mathcal{E}_{abcd'e'..x'} & = |g|^{-1}\varepsilon^{abcde..x}\varepsilon_{abcd'e'..x'} & = \ 3!\ g(^{de...x;}\ {}_{d'e'...x'}) \\
\bullet\bullet\bullet\bullet\bullet & & \\
\mathcal{E}^{abcd...x}\ \mathcal{E}_{abcd...x'} & = |g|^{-1}\varepsilon^{abcd...x}\ \varepsilon_{abcd...x'} & = \ (N\text{-}1)!g(^{x}{}_{x'}) \ = \ (N\text{-}1)!g^{x}{}_{x} \\
\mathcal{E}^{bcd...x}\ \mathcal{E}_{abcd...x} & = |g|^{-1}\varepsilon^{abcd...x}\ \varepsilon_{abcd...x} & = N!
\end{array} \tag{D.11.8}$$

where g(...) is the determinant of a matrix of $g^{\alpha}{}_{\beta}$ elements. Eq. (D.11.5) provides an example of the first line of (D.11.8) for the N=2 dimensions. Each equation in (D.11.8) is a true tensor equation allowing individual non-summed indices to be taken up and down on both sides. And of course the contraction tilts can be individually reversed as well. If *all* index positions are reversed, one gets a version of (D.11.8) where all matrix entries have the form $g_{\alpha}{}^{\beta}$ instead of $g^{\alpha}{}_{\beta}$, For example, the second last line would be,

$$\mathcal{E}_{abcd...x}\mathcal{E}^{abcd...x'} = |g|^{-1}\varepsilon_{abcd...x}\varepsilon^{abcd...x'} = (N\text{-}1)!g(_{x}{}^{x'}) = (N\text{-}1)!g_{x}{}^{x'} \tag{D.11.9}$$

As in (7.4.19) one always has $g_{\alpha}{}^{\beta} = g^{\alpha}{}_{\beta} = \delta^{\alpha}{}_{\beta} = \delta_{\alpha}{}^{\beta} = \delta_{\alpha,\beta}$ but of course $g_{\alpha\beta} \neq \delta_{\alpha,\beta}$.

Since the above equations are true tensor equations, Section 7.15 says the equations are also valid in x'-space where every object in the equation is primed.

Example 2: For N = 3 the second line of (D.11.8) reads,

$$\mathcal{E}^{abc}\mathcal{E}_{ab'c'} = |g|^{-1}\varepsilon^{abc}\varepsilon_{ab'c'} = g(^{bc;}{}_{b'c'}) = \begin{vmatrix} g^{b}{}_{b'} & g^{b}{}_{c'} \\ g^{c}{}_{b'} & g^{c}{}_{c'} \end{vmatrix} = g^{b}{}_{b'}g^{c}{}_{c'} - g^{c}{}_{b'}g^{b}{}_{c'} \ . \tag{D.11.10}$$

Raising b' and c' gives

$$\mathcal{E}^{abc}\mathcal{E}_{a}{}^{b'c'} = |g|^{-1}\varepsilon^{abc}\varepsilon_{a}{}^{b'c'} = g(^{bc;\ b'c'}) = \begin{vmatrix} g^{bb'} & g^{bc'} \\ g^{cb'} & g^{cc'} \end{vmatrix} = g^{bb'}g^{cc'} - g^{cb'}g^{bc'} \ .$$

Shuffling indices, this can be written

$$\mathcal{E}^{sAB}\mathcal{E}_{s}{}^{A'B'} = |g|^{-1}\varepsilon^{sAB}\varepsilon_{s}{}^{A'B'} = g(^{AB;\ A'B'}) = \begin{vmatrix} g^{AA'} & g^{AB'} \\ g^{BA'} & g^{BB'} \end{vmatrix} = g^{AA'}g^{BB'} - g^{BA'}g^{AB'}$$

$$= \mathcal{E}_{s}{}^{AB}\mathcal{E}^{sA'B'} = |g|^{-1}\varepsilon_{s}{}^{AB}\varepsilon^{sA'B'} \ . \tag{D.11.11}$$

Added-sign-s and Ricci-Levi-Civita conventions:

Do the above two bullet items, then add an overall sign s to the right side, because $\varepsilon_{abc..} = s\, g\, \varepsilon^{abc...}$ in these conventions instead of $\varepsilon_{abc..} = g\, \varepsilon^{abc...}$ so that $\varepsilon_{abc..}$(Weinberg) = $s\varepsilon_{abc..}$(added-sign).

Example: The example above becomes ($\varepsilon_{a'b'} \rightarrow s\, \varepsilon_{a'b'}$)

$$\mathcal{E}^{ab}\mathcal{E}_{a'b'} = |g|^{-1}\, \varepsilon^{ab}\varepsilon_{a'b'} = s \begin{vmatrix} g^{a}{}_{a'} & g^{a}{}_{b'} \\ g^{b}{}_{a'} & g^{b}{}_{b'} \end{vmatrix} = s\, (\, g^{a}{}_{a'}\, g^{b}{}_{b'} - g^{a}{}_{b'}\, g^{b}{}_{a'}) \ . \tag{D.11.12}$$

The second equation is *undefined* (when s=-1), because we don't know how to lower just one index on ε, hence on $\mathcal{E}$, as discussed near (D.5.2). The third equation is

$$\mathcal{E}_{ab}\mathcal{E}_{a'b'} = |g|^{-1}\, \varepsilon_{ab}\varepsilon_{a'b'} = \begin{vmatrix} g_{aa'} & g_{ab'} \\ g_{ba'} & g_{bb'} \end{vmatrix} = g_{aa'}\, g_{bb'} - g_{ab'}\, g_{ba'} \ . \tag{D.11.13}$$

The first and third equations are true tensor equations, except individual indices cannot be raised and lowered. If one were doing some significant work involving covariance and s=-1, it would certainly seem advisable to use the Weinberg convention since it is completely "covariant" for either sign of s.

D.12 How determinants of rank-2 tensors transform

In this document we have encountered only a few determinants of tensors (like g_{ij}) and tensor-like objects (like $R^i{}_j$ and $S^i{}_j$). Nevertheless, we would like to know how the determinant of a rank-2 tensor transforms under $x = \mathbf{F}(\mathbf{x})$. To this end, we first rewrite the traditional determinant formula in a covariant form. Once this is done, the conclusions come quickly.

Comment: The objects below like $\det(M^i{}_j)$ and $\det(M_{ij})$ of course have no dependence on the "indices" i and j other than on the up/down position of these indices. Elsewhere we write these objects as $\det(M^*{}_*)$ and $\det(M_{**})$ where the "wildcards" just show the nature of the matrix elements (down-tilt or all-down). Here, for technical reasons to be seen below, we maintain the notations $\det(M^i{}_j)$ and $\det(M_{ij})$.

We start with this mechanical statement of the determinant of a matrix $M^i{}_j$

$$\det(M^i{}_j) = \varepsilon_{ab..x}\, M^1{}_a M^2{}_b....M^N{}_x \qquad \text{(D.12.1)}$$

where ε is the permutation tensor. This form is "mechanical" just in the sense that if you drew a picture of the matrix $M^i{}_j$ and mechanically evaluated $\det(M^i{}_j)$, you obtain the above result. We can now restate this determinant, switching from the permutation tensor $\varepsilon_{abc...x}$ to the index-all-up Levi-Civita tensor. As stated below (D.4.1), we take $\varepsilon^{abc..x}$ to be equal to the permutation tensor. We then have

$$\det(M^i{}_j) = \varepsilon^{ab..x}\, M^1{}_a M^2{}_b....M^N{}_x \qquad \text{(D.12.2)}$$

which certainly looks more "covariant" since all summed indices appear to be contracted. But with fixed upper indices 1,2...N "hanging out", it is not quite clear what is going on here. Since a through x are all contracted, and since weight(ε) = - 1, one might be tempted to say $\det(M^i{}_j)$ is a scalar density of weight -1, but that is incorrect.

To get a better view of things, it is useful to rewrite the above determinant as follows (proof follows)

$$\det(M^i{}_j) = (1/N!)\ \varepsilon^{AB...X}\, \varepsilon^{ab..x}\, M^A{}_a\, M^B{}_b\, ...M^X{}_x\ . \qquad \text{(D.12.3)}$$

Again, both ε's shown here are in effect permutation tensors.

We shall now show that the right sides of the previous two equations are exactly the same. First write the right hand side of (D.12.3) as

$$\text{RHS (D.12.3)} = (1/N!)\, \varepsilon^{AB...X} \{\ \varepsilon^{ab..x} M^A{}_a\, M^B{}_b\, ...M^X{}_x\ \}\ . \qquad \text{(D.12.4)}$$

The bracketed quantity can be expanded as,

$$Q^{AB..X} \equiv \{\varepsilon^{ab..x} M^A{}_a\, M^B{}_b\, ...M^X{}_x\} = M^A{}_1\, M^B{}_2\, ...M^X{}_N + \text{all signed permutations}\ . \qquad \text{(D.12.5)}$$

meaning all signed permutations of the upper indices. Notice that

$$Q^{12..N} \equiv \{\varepsilon^{ab..x} M^1{}_a M^2{}_b ...M^N{}_x\} = M^1{}_1 M^2{}_2 ...M^N{}_N + \text{all signed permutations} . \qquad (D.12.6)$$

This is just a mechanical evaluation of $\det(M^i{}_j)$ and we conclude that

$$Q^{12..N} = \det(M^i{}_j) . \qquad (D.12.7)$$

The tensor $Q^{AB..X}$ is totally antisymmetric as this example demonstrates,

$$Q^{BA..X} = \varepsilon^{abc..x} M^B{}_a M^A{}_b M^C{}_c ...M^X{}_x = [- \varepsilon^{bac..x}] M^A{}_b M^B{}_a M^C{}_c ...M^X{}_x$$

$$= - [\varepsilon^{abc..x}] M^A{}_a M^B{}_b M^C{}_c ...M^X{}_x \qquad // a\leftrightarrow b$$

$$= - Q^{AB..X} . \qquad (D.12.8)$$

Then using theorem (D.3.1) we can write $Q^{AB..X}$ in this form.

$$Q^{AB..X} = K \varepsilon^{AB..X} . \qquad (D.12.9)$$

To evaluate K, use the standard ordering 123..N to find from the above and (D.12.7),

$$Q^{12..N} = K \varepsilon^{12..N} = K = \det(M^i{}_j) . \qquad (D.12.10)$$

Then, since $K = \det(M^i{}_j)$ one finds that

$$Q^{AB..X} = \det(M^i{}_j) \varepsilon^{AB..X} . \qquad (D.12.11)$$

Therefore (D.12.4) can be written

$$\text{RHS (D.12.3)} = (1/N!)\varepsilon^{AB...X} \{ \varepsilon^{ab..x} M^A{}_a M^B{}_b ...M^X{}_x \}$$

$$= (1/N!)\varepsilon^{AB...X} \{Q^{AB..X} \}$$

$$= (1/N!)\varepsilon^{AB...X} \{\det(M^i{}_j) \varepsilon^{AB..X}\}$$

$$= \det(M^i{}_j) (1/N!) \Sigma_{AB..X} (\varepsilon^{AB...X})^2 . \qquad (D.12.12)$$

The sum $\Sigma_{AB..X} (\varepsilon^{AB...X})^2$ is a "sum of ones" and there is a one for each permutation of AB..X. All other terms are zero. There are N! total permutations including the first, so

$$\Sigma_{AB..X} (\varepsilon^{AB...X})^2 = N! \qquad (D.12.13)$$

which agrees with the last equation of (D.10.37). Therefore,

$$\text{RHS (D.12.3)} = \det(M^i{}_j) (1/N!) \Sigma_{AB..X} (\varepsilon^{AB...X})^2 = \det(M^i{}_j) . \qquad (D.12.14)$$

This concludes our too-lengthy proof that the right sides of (D.12.2) and (D.12.3) are identical.

Now we start with the proven result,

$$\det(M^i{}_j) = (1/N!)\ \varepsilon^{ab..x}\ \varepsilon^{AB...X}\ M^A{}_a\, M^B{}_b \ldots M^X{}_x\,. \qquad \textit{(D.12.3)}$$

Eq (D.5.10) says that $\varepsilon_{ABC...} = g\, \varepsilon^{ABC...}$ so the above can be written

$$\det(M^i{}_j) = (1/g)\,(1/N!)\ \varepsilon^{ab..x}\ \varepsilon_{AB...X}\, M^A{}_a\, M^B{}_b \ldots M^X{}_x\,. \qquad \text{(D.12.15)}$$

Now, finally, we have a form in which all tensor indices are contracted with no loose ends. We can then use theorem (D.2.3) about the additivity of weights. Recall from (D.4.9) and (D.6.7) that both ε tensors shown in (D.12.15) have weight -1, and that the object 1/g has weight +2 from (D.1.7). Adding, we find that the object $\det(M^i{}_j)$ has weight 0.

We have therefore proven: (scalar = scalar density of weight 0)

Theorem: The determinant $\det(M^i{}_j)$ of a mixed rank-2 tensor $M^i{}_j$ transforms as a scalar under the transformation $\mathbf{x} = \mathbf{F}(\mathbf{x})$. (D.12.16)

<u>Comment</u>: Since our Chapter 2 S matrix $S^i{}_j$ is not a tensor, $J = \det(S^i{}_j)$ is not a scalar, and is fact not a tensor of any kind since it bridges x-space and x'-space.

It is now straightforward to determine the transformation nature of the other three rank-2 tensor types, and in fact to find simple relations between the four determinants. In all these cases, we use of the up-down altering property of the g tensor of (7.4.11), the "up-tilt" or "down-tilt" version of matrix multiplication of (7.8.11), the matrix rule det(AB) = det(A)det(B), facts (7.5.20) and (7.5.21), and the weight of g and 1/g as determined in (D.1.6) and (D.1.7) :

$$\underset{-2}{\det(M_{ij})} = \det(g_{ia}M^a{}_j) = \det(g_{ij})\,\det(M^i{}_j) = \underset{-2}{g}\ \underset{0}{\det(M^i{}_j)}.$$

$$\underset{0}{\det(M_i{}^j)} = \det(M_{ia}g^{aj}) = \det(M_{ij})\det(g^{ij}) = \{\, \underset{-2}{g}\ \underset{0}{\det(M^i{}_j)}\}\{\underset{+2}{g^{-1}}\} = \underset{0}{\det(M^i{}_j)}$$

$$\underset{+2}{\det(M^{ij})} = \det(g^{ia}M_a{}^j) = \det(g^{ij})\,\det(M_i{}^j) = \underset{+2}{g^{-1}}\ \underset{0}{\det(M_i{}^j)} = \underset{+2}{g^{-1}}\ \underset{0}{\det(M^i{}_j)}\ . \qquad \text{(D.12.17)}$$

The weights are shown under each line of equations. The conclusions are these:

$\det(M_i{}^j) = \det(M^i{}_j)$ // scalar densities of weight 0 (= scalar)

$\det(M_{ij}) = g\,\det(M^i{}_j)$ // scalar density of weight -2 $g = \det(g_{ij})$

$\det(M^{ij}) = g^{-1}\det(M^i{}_j)$ // scalar density of weight +2 $g^{-1} = \det(g^{ij})$. (D.12.18)

Here we have related each of the three partner determinants to the down-tilt determinant, and have shown the weight of each type of determinant.

Example: We know that g_{ij} is a rank-2 tensor from Section 5.7. Thus, we expect to find from (D.12.18) that,

$$
\begin{array}{llll}
\det(g_i{}^j) = \det(g^i{}_j) & \Rightarrow & \det(\delta_i{}^j) = \det(\delta^i{}_j) \quad \text{or} \quad 1 = 1 & \text{ok, weight } 0 \\
\det(g_{ij}) = g \det(g^i{}_j) & \Rightarrow & g = g \det(\delta^i{}_j) = g*1 = g & \text{ok, weight } -2 \\
\det(g^{ij}) = g^{-1} \det(g^i{}_j) & \Rightarrow & g^{-1} = g^{-1}\det(\delta^i{}_j) = g^{-1}* 1 = g^{-1} & \text{ok, weight } 2\ .
\end{array}
\qquad (D.12.19)
$$

Equations (D.12.18) are valid only for rank-2 tensors. They are not valid for R and S.

Go back now to the first line of (D.12.18),

$$\det(M_i{}^j) = \det(M^i{}_j) = \det([M^{T}]_j{}^i)$$

where M^{T} is the "covariant transpose" of M as shown in (7.9.3). This shows that

$$\det(M_*{}^*) = \det([M^{T}]_*{}^*)$$

or

$$\det(M_{\mathtt{ut}}) = \det(M^{T}{}_{\mathtt{ut}})\ . \qquad \text{// ut = up-tilt}$$

Since $\det(M) = \det(M^{\mathrm{T}})$ is valid for any matrix regardless of index position (M^{T} is the "matrix transpose" of M), we find that

$$\det(M^{T}{}_{\mathtt{ut}}) = \det(M^{\mathrm{T}}{}_{\mathtt{ut}}) = \det(M_{\mathtt{ut}})\ .$$

A similar argument beginning with

$$\det(M^i{}_j) = \det(M_i{}^j) = \det([M^{T}]^j{}_i) \qquad \Rightarrow \qquad \det(M_{\mathtt{dt}}) = \det(M^{T}{}_{\mathtt{dt}})$$

shows that

$$\det(M^{T}{}_{\mathtt{dt}}) = \det(M^{\mathrm{T}}{}_{\mathtt{dt}}) = \det(M_{\mathtt{dt}})\ . \qquad \text{// dt = down-tilt}$$

Since $M^{\mathrm{T}} = M^{T}$ for both-up or both-down index positions, we arrive at this interesting generalization of the traditional determinant theorem $\det(M) = \det(M^{\mathrm{T}})$:

Fact: $\det(M) = \det(M^{\mathrm{T}}) = \det(M^{T})$ for all four possible index positions (D.12.20)

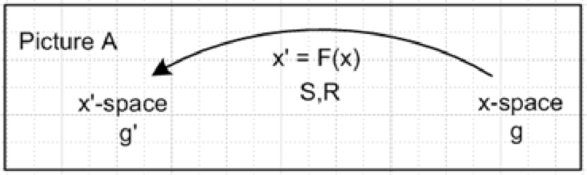
where T is the matrix transpose and *T* is the covariant transpose (7.9.3).

The scalar density weights of these determinants are given in (D.12.18) and do depend on the index positions.

Appendix E: Tensor Expansions: direct product, polyadic and operator notation

E.1 Direct Product Notation

This entire Section uses the general Picture A context where x-space need not be Cartesian,



(E.1.1)

The Standard Notation of Chapter 7 is used throughout. Useful forms can be found in Section 7.18.

The key tool required for the expression of tensor expansions is the notion of a direct product (tensor product, outer product) of n tensorial vectors defined in this simple way,

$$(\mathbf{A}\otimes\mathbf{B}\otimes\mathbf{C}\ \mathbf{...})^{abc\cdots} \equiv A^aB^bC^c.....$$
$$(\mathbf{A}\otimes\mathbf{B}\otimes\mathbf{C}\ \mathbf{...})^{a}{}_{b}{}^{c\cdots} \equiv A^aB_bC^c..... \quad \text{etc} \quad . \qquad \text{(E.1.2)}$$

The tensor **A⊗B⊗C ...** is nothing more than the outer product of vectors A,B,C as in (7.1.1) for contravariant vectors, but later extended to any mixture of vector types. One can, however, construct outer products from more complicated tensors, see Lucht *Tensor Products*. Suppose one has

A = vector F,G = rank-2 tensors M,N = rank-3 tensors . (E.1.3)

Then here is a small sampling of outer products of tensors that can be formed,

$$(F\otimes A)^{abc} = F^{ab}A^c \qquad (F\otimes A\otimes G)^{abcde} = F^{ab}A^cG^{de} \qquad (F\otimes G)^{abcd} = F^{ab}G^{cd}$$

$$(F\otimes M\otimes A)^{abcdef} = F^{ab}M^{cde}A^f \qquad (M\otimes N)^{abcdef} = M^{abc}N^{def} \ . \qquad \text{(E.1.4)}$$

In any of these equations, any one or more of the indices can be lowered on both sides to provide a valid equation since all objects in the list (E.1.3) are assumed to be true tensors. All the left-hand-side objects above are tensors because the right sides show that they transform as tensors.

In what follows, only the (E.1.2) direct product of vectors shall be considered. One can define the dot product of two direct-product-space vectors in this obvious manner,

$$(\mathbf{A}\otimes\mathbf{B}\otimes\mathbf{C}\ \mathbf{...}) \bullet (\mathbf{A'}\otimes\mathbf{B'}\otimes\mathbf{C'}\ \mathbf{...}) \equiv (\mathbf{A}\otimes\mathbf{B}\otimes\mathbf{C}\ \mathbf{...})^{abc\cdots}\ (\mathbf{A'}\otimes\mathbf{B'}\otimes\mathbf{C'}\ \mathbf{...})_{abc}$$

$$= A^aB^bC^c.....\ A'_aB'_bC'_c..... = \mathbf{A}\bullet\mathbf{A'}\ \mathbf{B}\bullet\mathbf{B'}\ \mathbf{C}\bullet\mathbf{C'}\ ... \qquad \text{(E.1.5)}$$

where of course the indices abc can be "tilted" in any way desired according to (7.11.3).

E.2 Tensor Expansions and Bases

Preamble: A seeming paradox and how it is resolved

Before attacking tensor expansions below, we wish to head off a possible confusion between a quantity transforming as a scalar under some transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ versus transforming as a component of a tensor. The "paradox" is presented in a few simple examples.

Example 1: Consider the vector $\mathbf{r} = (x,y,z)$ and a transformation $\mathbf{r}' = \mathbf{F}(\mathbf{r}) = R\mathbf{r}$ which is a rotation. Assume x-space is Cartesian so $g = 1$. Consider the dot product $s = \mathbf{r} \bullet \hat{\mathbf{x}}$. As noted in (5.10.2), this dot product transforms as a scalar under $\mathbf{F}$. In x'-space we still find that $\mathbf{r}' \bullet \hat{\mathbf{x}}' = s$. After all, the projection of one vector onto another cannot change when the two vectors are rotated together.

On the other hand, direct calculation shows that $\mathbf{r} \bullet \hat{\mathbf{x}} = x$, and x is a component of the vector $\mathbf{r} = (x,y,z)$. So how can one say that quantity $\mathbf{r} \bullet \hat{\mathbf{x}}$ is a scalar when $\mathbf{r} \bullet \hat{\mathbf{x}} = x$ which is a component of a vector? This *type* of question arises frequently in the study of tensor analysis: there seems to be a paradox that needs resolving.

The physicist considers several different "thought experiments".

In Experiment 1, two measurements are made. The first is made in Frame S and one finds that $s = x$. The second measurement is made of rotated vector $\mathbf{r}' = R\mathbf{r}$ in rotated Frame S' and there one finds that $s = x'$. The two measurements give the same number, so in Experiment 1 one finds that $x' = x$ and the quantity so measured is a scalar. In this experiment $x = \mathbf{r} \bullet \hat{\mathbf{x}}$ and $x' = \mathbf{r}' \bullet \hat{\mathbf{x}}'$.

In Experiment 2, everything is done in Frame S. One draws a vector $\mathbf{r}$ and measures x. One then rotates this vector within Frame S to get a new vector $\mathbf{r}' = R\mathbf{r}$ where $\mathbf{r}' = (x',y',z')$. One then measures x' and finds $x' \neq x$. In *this* experiment, $x = \mathbf{r} \bullet \hat{\mathbf{x}}$ and $x' = \mathbf{r}' \bullet \hat{\mathbf{x}}$ where $\hat{\mathbf{x}}$ has no prime.

In Experiment 3 one observes unrotated $\mathbf{r}$ from rotated Frame S' so $x = \mathbf{r} \bullet \hat{\mathbf{x}}$ and $x' = \mathbf{r} \bullet \hat{\mathbf{x}}'$ and this again results in $x' \neq x$

Thus, in these three experiments, x has the same definition, but x' has three different definitions. That is why one can have $x' = x$ in Experiment 1 and $x' \neq x$ in Experiments 2 and 3. The concept of $\mathbf{r} \bullet \hat{\mathbf{x}}$ as a scalar applies to Experiment 1 where $s = \mathbf{r} \bullet \hat{\mathbf{x}} = x = \mathbf{r}' \bullet \hat{\mathbf{x}}' = x'$.

Example 2. Consider the vector expansion (7.13.10) which says $\mathbf{V} = \Sigma_n V'^n \mathbf{e}_n$ with $V'^n = \mathbf{V} \bullet \mathbf{e}^n$. Here we have the same paradoxical situation: we know that $s = \mathbf{V} \bullet \mathbf{e}^n$ must be a scalar with respect to $\mathbf{F}$, yet it is equal to the component of a contravariant vector V'^n under $\mathbf{F}$. If we write $\mathbf{V} = \Sigma_n \mathcal{V}^n \mathbf{e}_n$, then in Frame S we find that $\mathbf{V} \bullet \mathbf{e}^n = \mathcal{V}^n$ which is like $\mathbf{r} \bullet \hat{\mathbf{x}} = x$ of the previous Example. In Frame S' we find instead that $\mathbf{V}' \bullet \mathbf{e}'^n = V'^i(\mathbf{e}'^n)_i = V'^i\delta^n{}_i = V'^n$ which is like $\mathbf{r}' \bullet \hat{\mathbf{x}}' = x'$ in Example 1. In Experiment 1 outlined above, it is not a paradox to have $s = \mathbf{e}^n \bullet \mathbf{V} = \mathcal{V}^n = V'^n$ be a scalar.

Example 3. Below we expand a rank-3 tensor $A = \Sigma_{ijk}\, \alpha^{ijk}\, (\mathbf{b}_i \otimes \mathbf{b}_j \otimes \mathbf{b}_k)$ and we find that $\alpha^{ijk} = A \bullet (\mathbf{b}^i \otimes \mathbf{b}^j \otimes \mathbf{b}^k)$ where $\bullet$ is a the dot product (E.1.5). We know that this quantity α^{ijk}, analogous to s in the two Examples above, must be a scalar under $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. Yet when we take $\mathbf{b}_n = \mathbf{e}_n$ we find in (E.2.8) that $\alpha^{ijk} = A'^{ijk}$ so our scalar α^{ijk} is equal to the component of a rank-3 tensor in x'-space. Once again, in Experiment 1 this is not a contradiction. It just happens that in x'-space the scalar α^{ijk}

appears as the value of a tensor component A'^{ijk} in x'-space, just the way s in Example 1 appears as the value x' of vector **r'** in Frame S'.

Tensor Expansions

In what follows, we use the example of a rank-3 tensor and the reader can easily see how this applies to a rank-n tensor.

Let $\mathbf{b}_i$ be an arbitrary complete set of basis vectors in x-space. As shown in the notes following (6.2.8) there exists a unique set of *dual* ("reciprocal") basis vectors $\mathbf{b}^i$ (also in x-space) such that $\mathbf{b}^i \bullet \mathbf{b}_j = \delta^i{}_j$, where we now use the Standard Notation $\mathbf{b}_n$ equations of (7.18.6). Consider then the following expansion of rank-3 tensor A with contravariant components A^{abc},

$$A = \Sigma_{ijk}\, \alpha^{ijk}\, (\mathbf{b}_i \otimes \mathbf{b}_j \otimes \mathbf{b}_k) \quad \text{where } (\mathbf{b}_i \otimes \mathbf{b}_j \otimes \mathbf{b}_k \ldots)^{abc} = (\mathbf{b}_i)^a\, (\mathbf{b}_j)^b\, (\mathbf{b}_k)^c \quad \textit{(E.1.2)}$$

$$A^{abc} = \Sigma_{ijk}\, \alpha^{ijk}\, (\mathbf{b}_i \otimes \mathbf{b}_j \otimes \mathbf{b}_k \ldots)^{abc} = \Sigma_{ijk}\, \alpha^{ijk}\, (\mathbf{b}_i)^a\, (\mathbf{b}_j)^b\, (\mathbf{b}_k)^c \ . \qquad \text{(E.2.1)}$$

As we did in Example 3 (2.9.5) for $\mathbf{e}_n$, we declare the $\mathbf{b}_n$ vectors to be "contravariant by definition" by writing the rule $(\mathbf{b'}_n)^i \equiv R^i{}_j(\mathbf{b}^n)_j$. So then the vectors $\mathbf{b}_n$ are true rank-1 tensors under $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. According to Section 7.1, the outer product $(\mathbf{b}_i)^a\, (\mathbf{b}_j)^b\, (\mathbf{b}_k)^c$ is a true rank-3 contravariant tensor under **F**. Then (E.2.1) says that A^{abc} is a linear combination of these tensors, which we know is a tensor because the sum of two tensors of some type is a tensor of the same type.

The expansion coefficients α^{ijk} can be obtained by dotting both sides of (E.2.1) with $(\mathbf{b}^{i'} \otimes \mathbf{b}^{j'} \otimes \mathbf{b}^{k'})$ and using (E.1.5),

$$(\mathbf{b}^{i'} \otimes \mathbf{b}^{j'} \otimes \mathbf{b}^{k'}) \bullet (\mathbf{b}_i \otimes \mathbf{b}_j \otimes \mathbf{b}_k) = \mathbf{b}^{i'} \bullet \mathbf{b}_i\ \ \mathbf{b}^{j'} \bullet \mathbf{b}_j\ \mathbf{b}^{k'} \bullet \mathbf{b}_k = \delta^{i'}{}_i \delta^{j'}{}_j \delta^{k'}{}_k \ . \qquad \text{(E.2.2)}$$

The result is then

$$A \bullet (\mathbf{b}^{i'} \otimes \mathbf{b}^{j'} \otimes \mathbf{b}^{k'}) = \{\Sigma_{ijk} \alpha^{ijk}\, (\mathbf{b}_i \otimes \mathbf{b}_j \otimes \mathbf{b}_k)\} \bullet (\mathbf{b}^{i'} \otimes \mathbf{b}^{j'} \otimes \mathbf{b}^{k'}) = \Sigma_{ijk}\, \alpha^{ijk}\, \delta^{i'}{}_i \delta^{j'}{}_j \delta^{k'}{}_k = \alpha^{i'j'k'}$$

or unpriming indices,

$$\alpha^{ijk} = A \bullet (\mathbf{b}^i \otimes \mathbf{b}^j \otimes \mathbf{b}^k) \ .$$

But since both A and $(\mathbf{b}^i \otimes \mathbf{b}^j \otimes \mathbf{b}^k)$ are elements of the triple direct-product space spanned by the vectors $(\mathbf{b}^i \otimes \mathbf{b}^j \otimes \mathbf{b}^k)$, we use the dot product of (E.1.5) to claim that

$$A \bullet (\mathbf{b}^i \otimes \mathbf{b}^j \otimes \mathbf{b}^k) = A^{abc}\, (\mathbf{b}^i \otimes \mathbf{b}^j \otimes \mathbf{b}^k)_{abc} = A^{abc}\, (\mathbf{b}^i)_a\, (\mathbf{b}^j)_b\, (\mathbf{b}^k)_c \ .$$

The coefficients α^{ijk} may then be written in all these ways :

$$\alpha^{ijk} = A \bullet (\mathbf{b}^i \otimes \mathbf{b}^j \otimes \mathbf{b}^k) = A^{abc}\, (\mathbf{b}^i \otimes \mathbf{b}^j \otimes \mathbf{b}^k)_{abc} = A^{abc}\, (\mathbf{b}^i)_a\, (\mathbf{b}^j)_b\, (\mathbf{b}^k)_c \qquad \text{(E.2.3)}$$

where A^{abc} are the contravariant components of tensor A in x-space, and $(b^i)_a$ are the covariant components of vector $\mathbf{b}^i$ in x-space. As noted in Example 3 of the previous subsection, the object α^{ijk} transforms as a *scalar*, since it is the dot product of two direct-product-space vectors. This fact is especially obvious from the last expression in (E.2.3) where all tensor indices are contracted so the result must be a scalar according to the neutralization rule (7.12.1). So α^{ijk} is a set of 3^3 =27 scalars.

Two special cases attract our attention.

The $(\mathbf{u}_i)^a$ are axis-aligned basis vectors in x-space, as shown in (7.13.9) or (7.18.3). For these basis vectors, one has $(\mathbf{u}_i)^a = \delta_i{}^a$ and $(\mathbf{u}^i)_a = \delta^i{}_a$. If one considers expansion (E.2.1) with $\mathbf{b}_n = \mathbf{u}_n$, then

$$A = \Sigma_{ijk}\, \alpha^{ijk}\, (\mathbf{u}_i \otimes \mathbf{u}_j \otimes \mathbf{u}_k)$$

$$\text{where} \qquad (\mathbf{u}_i \otimes \mathbf{u}_j \otimes \mathbf{u}_k \ldots)^{abc} = (\mathbf{u}_i)^a\, (\mathbf{u}_j)^b\, (\mathbf{u}_k)^c = \delta_i{}^a\, \delta_j{}^b\, \delta_k{}^c \qquad \text{(E.2.4)}$$

and the coefficients are found to be

$$\alpha^{ijk} = A \bullet (\mathbf{u}^i \otimes \mathbf{u}^j \otimes \mathbf{u}^k) = A^{abc}\ \delta^i{}_a\, \delta^j{}_b\, \delta^k{}_c = A^{ijk}\,, \qquad \text{(E.2.5)}$$

so the scalar coefficients α^{ijk} are exactly the x-space contravariant components of the tensor A. If this seems paradoxical, see Example 1 above where we found that scalar s = $\mathbf{r} \bullet \hat{\mathbf{x}} =$ x. Thus,

$$A = \Sigma_{ijk}\, A^{ijk}\, (\mathbf{u}_i \otimes \mathbf{u}_j \otimes \mathbf{u}_k)\ . \qquad \text{(E.2.6)}$$

On the other hand, if $\mathbf{e}_i$ are the tangent base vectors in x-space (see Chapters 3), the dual vectors are the $\mathbf{e}^i$ and from (7.13.1), $(\mathbf{e}_i)^a = S^a{}_i = R_i{}^a$ and $(\mathbf{e}^i)_a = S_a{}^i = R^i{}_a$. If one considers the expansion

$$A = \Sigma_{ijk}\, \alpha^{ijk}\, (\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k)$$

$$\text{where} \qquad (\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \ldots)^{abc} = (\mathbf{e}_i)^a\, (\mathbf{e}_j)^b\, (\mathbf{e}_k)^c = R_i{}^a\, R_j{}^b\, R_k{}^c \qquad \text{(E.2.7)}$$

then the coefficients are found to be

$$\alpha^{ijk} = A \bullet (\mathbf{e}^i \otimes \mathbf{e}^j \otimes \mathbf{e}^k) = A^{abc}\ (\mathbf{e}^i)_a (\mathbf{e}^j)_b (\mathbf{e}^k)_c = A^{abc}\, R^i{}_a\, R^j{}_b\, R^k{}_c$$

$$= R^i{}_a\, R^j{}_b\, R^k{}_c\, A^{abc} = A'^{ijk} \qquad \text{(E.2.8)}$$

and thus the scalar coefficients α^{ijk} in this case are exactly the x'-space contravariant components of tensor A, as shown for M in (7.10.9). Thus,

$$A = \Sigma_{ijk}\, A'^{ijk}\, (\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k)\ . \qquad \text{(E.2.9)}$$

Expansions like the above are the generalizations (to tensors of any rank) of these vector expansions stated in (7.13.12),

$\mathbf{A} = \Sigma_i \alpha^i\, \mathbf{b}_i$ $\qquad \alpha^i = \mathbf{b}^i \bullet \mathbf{A}$ // arbitrary basis

$\mathbf{A} = \Sigma_i A^i\, \mathbf{u}_i$ // axis aligned unit vectors

$\mathbf{A} = \Sigma_i A'^i\, \mathbf{e}_i$ // tangent base vectors (E.2.10)

where we continue to write rank-1 tensors (vectors) in bold font: **A**.

To summarize, here is the general rank-n tensor expansion for an arbitrary basis, and then for the two specific bases just discussed:

$A = \Sigma_{ijk...}\, \alpha^{ijk...}\, (\mathbf{b}_i\otimes\mathbf{b}_j\otimes\mathbf{b}_k...)$ $\qquad \alpha^{ijk...} = A^{abc...}\, (b^i)_a\, (b^j)_b\, (b^k)_c...$

$A = \Sigma_{ijk...}\, A^{ijk...}\, (\mathbf{u}_i\otimes\mathbf{u}_j\otimes\mathbf{u}_k...)$ $\qquad A^{ijk...}$ = contravariant components of A in x-space

$A = \Sigma_{ijk...}\, A'^{ijk...}\, (\mathbf{e}_i\otimes\mathbf{e}_j\otimes\mathbf{e}_k...)$ $\qquad A'^{ijk...}$ = contravariant components of A in x'-space

Tilting labels (E.2.11)

In the above discussion, suppose we make the swap $\mathbf{b}_i \rightarrow \mathbf{b}^i$. We quickly trace the steps:

$A = \Sigma_{ijk}\, \alpha_i{}^{jk}\, (\mathbf{b}^i\otimes\mathbf{b}_j\otimes\mathbf{b}_k)$ where $(\mathbf{b}^i\otimes\mathbf{b}_j\otimes\mathbf{b}_k\, ...)_a{}^{bc} = (\mathbf{b}^i)_a\, (\mathbf{b}_j)^b\, (\mathbf{b}_k)^c$ *(E.2.1)*

$A_a{}^{bc} = \Sigma_{ijk}\, \alpha_i{}^{jk}\, (\mathbf{b}^i\otimes\mathbf{b}_j\otimes\mathbf{b}_k\, ...)_a{}^{bc} = \Sigma_{ijk}\, \alpha_i{}^{jk}\, (\mathbf{b}^i)_a\, (\mathbf{b}_j)^b\, (\mathbf{b}_k)^c$ = outer product sum *(E.2.1)*

$\alpha_i{}^{jk} = A \bullet (\mathbf{b}_i\otimes\mathbf{b}^j\otimes\mathbf{b}^k) = A_a{}^{bc}\, (\mathbf{b}_i\otimes\mathbf{b}^j\otimes\mathbf{b}^k)^a{}_{bc} = A_a{}^{bc}\, (\mathbf{b}_i)^a\, (\mathbf{b}^j)_b\, (\mathbf{b}^k)_c$ = scalar *(E.2.3)*

$A = \Sigma_{ijk}\, \alpha_i{}^{jk}\, (\mathbf{u}^i\otimes\mathbf{u}_j\otimes\mathbf{u}_k)$ // first special case *(E.2.4)*

$\alpha_i{}^{jk} = A \bullet (\mathbf{u}_i\otimes\mathbf{u}^j\otimes\mathbf{u}^k) = A_a{}^{bc}\, (\mathbf{u}_i)^a\, (\mathbf{u}^j)_b\, (\mathbf{u}^k)^c{}_c = A_a{}^{bc}\, \delta_i{}^a\, \delta^j{}_b\, \delta^k{}_c = A_i{}^{jk}$ *(E.2.5)*

$A = \Sigma_{ijk}\, \alpha_i{}^{jk}\, (\mathbf{e}^i\otimes\mathbf{e}_j\otimes\mathbf{e}_k)$ // second special case *(E.2.7)*

$\alpha_i{}^{jk} = A \bullet (\mathbf{e}_i\otimes\mathbf{e}^j\otimes\mathbf{e}^k) = A_a{}^{bc}\, (\mathbf{e}_i)^a(\mathbf{e}^j)_b(\mathbf{e}^k)_c = A_a{}^{bc}\, R_i{}^a\, R^j{}_b\, R^k{}_c = A'_i{}^{jk}$ *(E.2.8)*

Here $\alpha_i{}^{jk} = A_a{}^{bc}\, (\mathbf{b}_i)^a\, (\mathbf{b}^j)_b\, (\mathbf{b}^k)_c = A^{abc}\, (\mathbf{b}_i)_a\, (\mathbf{b}^j)_b\, (\mathbf{b}^k)_c$ is a different set of scalars compared with the original $\alpha^{ijk} = A^{abc}\, (\mathbf{b}^i)_a\, (\mathbf{b}^j)_b\, (\mathbf{b}^k)_c$ because in general $\mathbf{b}_i \neq \mathbf{b}^i$. The sequence above shows that everything goes through as before, and so we may expand tensor A in either of these two ways,

$$A = \Sigma_{ijk}\, \alpha^{ijk}\, (\mathbf{b}_i \otimes \mathbf{b}_j \otimes \mathbf{b}_k)$$

$$A = \Sigma_{ijk}\, \alpha_i{}^{jk}\, (\mathbf{b}^i \otimes \mathbf{b}_j \otimes \mathbf{b}_k)\ . \qquad \text{(E.2.12)}$$

The second expansion is really a mixed-basis expansion as discussed in Section E.10 below. The main point of this exercise is to show that one may "reverse the tilt" of the i summation index and still have a viable tensor expansion. This is *not* an example of the "contraction tilt-reversal rule" (7.11.1) because the object α^{ijk} is not a rank-3 tensor, it is a set of scalars. Moreover, index i is not a tensor index, it is a label on the basis vector $\mathbf{b}_i$. Nevertheless, in this case tilt-reversal is allowed and in fact just serves to define a new set of coefficients $\alpha_i{}^{jk}$

It turns out that the indices on α^{ijk} can be lowered by the matrix $w'_{nm} = \mathbf{b}_n \bullet \mathbf{b}_m$ which appears in (7.18.6). To show this, consider

$$\alpha^{ijk} = A^{abc}\, (\mathbf{b}^i)_a\, (\mathbf{b}^j)_b\, (\mathbf{b}^k)_c = A_a{}^{bc}\, (\mathbf{b}^i)^a\, (\mathbf{b}^j)_b\, (\mathbf{b}^k)_c \qquad \text{// (E.2.3) then tensor tilt rule}$$

so

$$w'_{si}\alpha^{ijk} = A_a{}^{bc}\, [w'_{si}\mathbf{b}^i]^a\, (\mathbf{b}^j)_b\, (\mathbf{b}^k)_c = A_a{}^{bc}\, [\mathbf{b}_s]^a\, (\mathbf{b}^j)_b\, (\mathbf{b}^k)_c \quad \text{// } \mathbf{b}_n = w'_{ni}\ \mathbf{b}^i \text{ in (7.18.6)}$$

$$= \alpha_s{}^{jk} \qquad \text{// as shown in } \textit{(E.2.3)} \text{ above} \qquad \text{(E.2.13)}$$

which shows that w'_{**} lowers an index on α^{ijk}. Despite this ability of w'_{**} and w'^{**} to lower and raise indices on the object family α^{ijk}, each element of this family (such as $\alpha_i{}^{jk}$) is a scalar. When $\mathbf{b}_n = \mathbf{e}_n$, we find that g' then raises and lowers indices on α^{ijk}, but this does not make α^{ijk} a rank-3 tensor, because α^{ijk} transforms as a scalar under $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. We apologize for constantly hammering on this point, but it can be a source of confusion.

Obviously one can tilt any or all of the indices in expansion (E.2.1). In particular, tilting all indices gives this alternate version of (E.2.11)

$A = \Sigma_{ijk...}\, \alpha_{ijk}{}^{...}\, (\mathbf{b}^i \otimes \mathbf{b}^j \otimes \mathbf{b}^k ...)$ $\qquad \alpha_{ijk..}{}^{.} = A_{abc...}\, (b_i)^a\, (b_j)^b\, (b_k)^c ...$

$A = \Sigma_{ijk...}\, A_{ijk}{}^{...}\, (\mathbf{u}^i \otimes \mathbf{u}^j \otimes \mathbf{u}^k ...)$ $\qquad A_{ijk...}$ = covariant components of A in x-space

$A = \Sigma_{ijk...}\, A'_{ijk}{}^{...}\, (\mathbf{e}^i \otimes \mathbf{e}^j \otimes \mathbf{e}^k ...)$ $\qquad A'_{ijk...}$ = covariant components of A in x'-space

Orthonormal basis (E.2.14)

If the basis vectors $\mathbf{b}_i$ happen to be orthonormal, as defined by $\mathbf{b}_i \bullet \mathbf{b}_j = \delta_{i,j}$ then $\mathbf{b}^i = \mathbf{b}_i$ because the dual basis is unique. In this case $w'_{ij} = \mathbf{b}_i \bullet \mathbf{b}_j = \delta_{ij}$ so the coefficient α^{ijk} is unchanged if any or all indices are lowered, see (E.2.13).

An example of orthonormal basis vectors arises if $\mathbf{b}_i = \hat{\mathbf{e}}_i \equiv \mathbf{e}_i/|\mathbf{e}_i| = \mathbf{e}_i/h'_i$ and $\hat{\mathbf{e}}_n \bullet \hat{\mathbf{e}}_m = \delta_{n,m}$. In this case we find from (7.18.1) that

$$g'_{nm} = \mathbf{e}_n \bullet \mathbf{e}_m = \hat{\mathbf{e}}_n \bullet \hat{\mathbf{e}}_m\, h'_m h'_n = \delta_{n,m}\ h'_m h'_n = h'^2_n \delta_{n,m} \qquad \text{(E.2.15)}$$

so x'-space has the diagonal metric tensor $g'_{nm} = h'_n{}^2\delta_{n,m}$. A specific example arises with polar coordinates as shown in Fig (3.4.3) where $\hat{\mathbf{e}}_1 = \hat{\boldsymbol{\theta}}$ and $\hat{\mathbf{e}}_2 = \hat{\mathbf{r}}$.

Since the dual basis is unique, one has $\mathbf{b}_i = \mathbf{b}^i = \hat{\mathbf{e}}^i = \hat{\mathbf{e}}_i$ †. In this case (E.2.1) and (E.2.3) state,

$$A = \Sigma_{ijk}\, \alpha^{ijk}\, (\hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j \otimes \hat{\mathbf{e}}_k) \qquad \alpha^{ijk}(\text{any up/down}) = A^{abc}\, (\hat{\mathbf{e}}^i)_a\, (\hat{\mathbf{e}}^j)_b\, (\hat{\mathbf{e}}^k)_c\,. \tag{E.2.16}$$

On the other hand, one can expand A as in (E.2.9), using $\mathbf{e}_i = h'_i \hat{\mathbf{e}}_i$,

$$A = \Sigma_{ijk}\, A'^{ijk}\, (\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k) = \Sigma_{ijk}\, [A'^{ijk}\, h'_i\, h'_j\, h'_k\,](\hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j \otimes \hat{\mathbf{e}}_k)\,. \tag{E.2.17}$$

Comparison of the last two equations shows that

$$\alpha^{ijk}(\text{any up/down}) = A^{abc}\, (\hat{\mathbf{e}}^i)_a\, (\hat{\mathbf{e}}^j)_b\, (\hat{\mathbf{e}}^k)_c = A'^{ijk}\, h'_i\, h'_j\, h'_k \quad \text{// no implied sums} \tag{E.2.18}$$

Expansions on the unit versions of the tangent base vectors $\hat{\mathbf{e}}_i$ are discussed more in Section E.8 below.

---

† Is it really true that $\mathbf{b}_i = \mathbf{b}^i \Rightarrow \hat{\mathbf{e}}^i = \hat{\mathbf{e}}_i$ where the latter two vectors are defined as unit vectors? This will serve as a little check on our notation. Since $g'_{ab} = h'_a{}^2\delta_{a,b}$ we know that $g'^{ab} = h'_a{}^{-2}\delta_{a,b}$. Then $g'_{nn} = h'_n{}^2$ and $g'^{nn} = h'_n{}^{-2}$. From (7.18.1) one has

$$|\mathbf{e}_n| = \sqrt{g'_{nn}} = h'_n \qquad \mathbf{e}^n = g'^{ni}\, \mathbf{e}_i = g'^{nn}\, \mathbf{e}_n = h'_n{}^{-2}\, \mathbf{e}_n$$
$$|\mathbf{e}^n| = \sqrt{g'^{nn}} = (1/\, h'_n) \qquad \mathbf{e}_n = h'_n{}^2\, \mathbf{e}^n$$

so

$$\hat{\mathbf{e}}_n \equiv \mathbf{e}_n/|\mathbf{e}_n| = \mathbf{e}_n\, h'_n{}^{-1} = [h'_n{}^2\mathbf{e}^n]\, h'_n{}^{-1} = \mathbf{e}^n\, h'_n = \mathbf{e}^n / |\mathbf{e}^n| \equiv \hat{\mathbf{e}}^n \quad \Rightarrow \hat{\mathbf{e}}_n = \hat{\mathbf{e}}^n \tag{E.2.19}$$

Thus one is free to move the $\hat{\mathbf{e}}_n$ label up or down at will.

Tensor density expansions

If A is a tensor density of weight W, the general rule is to make this replacement:

$$A'^{ijk\ldots} \rightarrow J^W A'^{ijk\ldots} \tag{E.2.20}$$

As justification for this rule, start with a regular tensor transformation for A,

$$A'^{ijk\ldots} = R^i{}_{i'}\, R^j{}_{j'}\, R^k{}_{k'} \ldots\ldots A^{i'j'k'\ldots} \tag{E.2.21}$$

The rule then gives

$$J^W A'^{ijk\ldots} = R^i{}_{i'}\, R^j{}_{j'}\, R^k{}_{k'} \ldots\ldots A^{i'j'k'\ldots} \tag{E.2.22}$$

or

$$A'^{ijk\ldots} = J^{-W}\, R^i{}_{i'}\, R^j{}_{j'}\, R^k{}_{k'} \ldots\ldots A^{i'j'k'\ldots} \tag{E.2.23}$$

which is the correct form for the transformation of a tensor density of weight W as in the example (D.1.4).

Using this rule, the expansion (E.2.9) of tensor density A would be written,

$$A = J^W \Sigma_{ijk...} A'^{ijk...} (\mathbf{e_i}\otimes\mathbf{e_j}\otimes\mathbf{e_k}...) \qquad A'^{ijk...} = \text{contravariant components of A in x'-space} \tag{E.2.24}$$

a result that is verified in (D.2.11). Taking components,

$$A^{abc...} = J^W \Sigma_{ijk...} A'^{ijk...} (\mathbf{e_i}\otimes\mathbf{e_j}\otimes\mathbf{e_k}...)^{abc..} = J^W \Sigma_{ijk...} A'^{ijk...} (\mathbf{e_i})^a (\mathbf{e_j})^b (\mathbf{e_k})^c ...$$

$$= J^W \Sigma_{ijk...} A'^{ijk...} R_i{}^a R_j{}^b R_k{}^c... \qquad // \text{ see (7.18.1)}$$

$$= J^W R_i{}^a R_j{}^b R_k{}^c ...A'^{ijk} \tag{E.2.25}$$

which is just the inverse of (E.2.23). Applied to a vector **A** of weight W, the expansion (E.2.24) becomes

$$A^a = J^W A'^i \mathbf{e_i} . \tag{E.2.26}$$

E.3 Polyadic Notation

Some fields of study historically use "polyadic notation" as follows

$$(\mathbf{ABC}...) \equiv \mathbf{A}\otimes\mathbf{B}\otimes\mathbf{C} ... \tag{E.3.1}$$

where the direct product notation was discussed above in Section E.1. It is sometimes a bit disturbing to modern readers to see bolded vectors stacked directly against each other, but the direct product makes the meaning clear. For arbitrary basis vectors, one would then have, for example,

$$(\mathbf{b_i b_j b_k}...) \equiv \mathbf{b_i}\otimes\mathbf{b_j}\otimes\mathbf{b_k} ... \quad . \tag{E.3.2}$$

Sometimes this basis vector notation is compressed even more, to wit,

$$\mathbf{i\ j\ k\ ...} \equiv (\mathbf{b_i b_j b_k}...) \equiv \mathbf{b_i}\otimes\mathbf{b_j}\otimes\mathbf{b_k} \mathbf{...} \tag{E.3.3}$$

although *this* notation seems to be mostly used when the $\mathbf{b_i}$ are the unit vectors $\mathbf{u_i}$.

In all these notations, one must be aware that the symbols generally do not "commute". For example

$$\mathbf{i\ j} = (\mathbf{b_i b_j}) = \mathbf{b_i}\otimes\mathbf{b_j} \qquad \Rightarrow \qquad \begin{array}{l} (\mathbf{i\ j})^{nm} = (\mathbf{b_i b_j})^{nm} = (\mathbf{b_i}\otimes\mathbf{b_j})^{nm} = (b_i)^n (b_j)^m \\ (\mathbf{j\ i})^{nm} = (\mathbf{b_j b_i})^{nm} = (\mathbf{b_j}\otimes\mathbf{b_i})^{nm} = (b_j)^n (b_i)^m \neq (\mathbf{i\ j})^{nm} \end{array} \tag{E.3.4}$$

and therefore one cannot in general write **i j** = **j i**.

The general expansion (E.2.1) now appears as

$$\begin{aligned} A &= \Sigma_{ijk\ldots}\ \alpha^{ijk\ldots}\ (\mathbf{b}_i\otimes\mathbf{b}_j\otimes\mathbf{b}_{k\ldots}) \\ &= \Sigma_{ijk\ldots}\ \alpha^{ijk\ldots}\ (\mathbf{b}_i\mathbf{b}_j\mathbf{b}_{k\ldots}) \\ &= \Sigma_{ijk\ldots}\ \alpha^{ijk\ldots}\ (\mathbf{i}\ \mathbf{j}\ \mathbf{k}_{\ldots}) \end{aligned} \qquad \text{(E.3.5)}$$

where

$$\begin{aligned} \alpha^{ijk\ldots} &= A \bullet (\mathbf{b}^i\otimes\mathbf{b}^j\otimes\mathbf{b}^k\ldots) \\ &= A \bullet (\mathbf{b}^i\mathbf{b}^j\mathbf{b}^k\ldots) \\ &= A \bullet\ (\mathbf{i}^d\,\mathbf{j}^d\,\mathbf{k}^d\ \ldots\ ) \\ &= A^{abc\ldots}\ (b^i)_a\,(b^j)_b\,(b^k)_c\ldots \end{aligned} \qquad \text{(E.3.6)}$$

where we have just made up a notation $\mathbf{i}^d$ to stand for the dual vector $\mathbf{b}^i$.

One can find further discussion of polyadic notation for example in Backus.

E.4 Dyadic Products

When two vectors **A** and **B** are combined in polyadic notation, the result is called a dyadic product (**AB**) [also known as a dyad or just a dyadic]

$$(\mathbf{AB})^{ij} \equiv A^iB^j \qquad // = (\mathbf{A}\otimes\mathbf{B})^{ij} = \text{component of the matrix } \mathbf{A}\otimes\mathbf{B} = \mathbf{AB}\ . \qquad \text{(E.4.1)}$$

In this notation, the expansion (E.3.5) for a rank-2 tensor becomes

$$A = \Sigma_{ij}\ \alpha^{ij}\ (\mathbf{b}_i\mathbf{b}_j) \qquad \alpha^{ij} = A^{ab}(b^i)_a\,(b^j)_b\ =\ A^{ab}\,(\mathbf{b}^i\mathbf{b}^j)_{ab}\ . \qquad \text{(E.4.2)}$$

Notice from (7.1.1) that the dyadic product (**AB**) *is* a rank-2 tensor if we assume that the underlying A^i and B^i are the x-space contravariant components of tensorial vectors **A** and **B** (which we normally assume). As a reminder, x-space need not be Cartesian. In Section E.7 it will be shown that the *matrix* $(\mathbf{AB})^{ij}$ can be associated with an *operator* (**AB**) in the $\mathbf{u}_n$ basis so $(\mathbf{AB})^{ij} = \langle u^i\ |(\mathbf{AB})|\ u^j\ \rangle$, but this interpretation is not necessary for what follows.

If vector **A** is replaced by the gradient operator ∇, one gets

$$(\nabla\mathbf{B})^{ij} \equiv \nabla^iB^j = \partial^iB^j\ , \qquad // \text{ dyadic notation} \qquad \text{(E.4.3)}$$

but it is more common to define the operator (∇**B**) using a reverse dyadic notation

$$(\nabla\mathbf{B})^{ij} \equiv \nabla^jB^i = \partial^jB^i\ = \partial B^i/\partial x^j \qquad // \text{ reverse dyadic notation} \qquad \text{(E.4.4)}$$

since the index order i,j on the far right matches the order on the left.

E.5 Matrix notation for dyadics (Cartesian Space)

In Cartesian space g = 1 so for any vector **a**, we have $a^i = a_i$.

If we think of **b** as a column vector with components b_j, then the corresponding row vector $\mathbf{b}^T$ contains those same elements b_j, so we set $(b^T)_j = b^j$. Therefore one can express the dyadic product in this more down-to-earth manner,

$$(\mathbf{ab})_{ij} \equiv a_i b_j = a_i (b^T)_j = (\mathbf{ab}^T)_{ij} \tag{E.5.1}$$

or

$$\mathbf{ab} = \mathbf{ab}^T \quad . \tag{E.5.2}$$

Here one knows that **ab** is a "dyadic" because there is no other meaning for two bolded column vectors abutting each other with no intervening operator, so no special notation like [**ab**] is needed to indicate that **ab** is a dyadic. The object $\mathbf{ab}^T$ on the other hand has a well-defined meaning in matrix algebra,

$$\mathbf{ab}^T = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} (b_1\ b_2) = \begin{pmatrix} a_1b_1 & a_1b_2 \\ a_2b_1 & a_2b_2 \end{pmatrix} = \text{a matrix} \qquad // \text{ same as matrix } \mathbf{ab} = \mathbf{a}\otimes\mathbf{b} \tag{E.5.3}$$

and one sees that in fact

$$(\mathbf{ab})_{ij} = (\mathbf{ab}^T)_{ij} = a_i b^T{}_j = a_i b_j \quad . \tag{E.5.4}$$

Meanwhile, the object $\mathbf{a}^T\mathbf{b}$ is just a number,

$$\mathbf{a}^T\mathbf{b} = (a_1\ a_2)\begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = a_1b_1 + a_2b_2 = \mathbf{a} \bullet \mathbf{b} \ . \qquad // \ g = 1 \qquad . \tag{E.5.5}$$

This transpose notation can then be applied to the dyadic expansion of a 2x2 matrix A,

$$A = \Sigma_{ij}\, \alpha_{ij}\, \mathbf{b}_i\mathbf{b}_j = \Sigma_{ij}\, \alpha_{ij}\, \mathbf{b}_i\mathbf{b}_j^T = \alpha_{11}\, \mathbf{b}_1\, \mathbf{b}_1^T + \alpha_{12}\, \mathbf{b}_1\, \mathbf{b}_2^T \ \ldots \quad . \tag{E.5.6}$$

In the special case that the $\mathbf{b}_i$ are the unit vectors $\mathbf{u}_i$, and assuming N = 2 dimensions, one has

$$A = \Sigma_{nm} A_{nm}\, \mathbf{u}_n\mathbf{u}_m = \Sigma_{nm} A_{nm}\, \mathbf{u}_n\mathbf{u}_m^T = A_{11}\, \mathbf{u}_1\, \mathbf{u}_1^T + A_{12}\, \mathbf{u}_1\, \mathbf{u}_2^T + A_{21}\, \mathbf{u}_2\, \mathbf{u}_1^T + A_{22}\, \mathbf{u}_2\, \mathbf{u}_2^T$$
$$= \text{a matrix with } A_{12} \text{ in the upper right corner} \tag{E.5.7}$$

where $\mathbf{u}_n$ is a column unit vector and $\mathbf{u}_n^T$ is the corresponding row unit vector. For example,

$$\mathbf{u}_1\mathbf{u}_2^T = \begin{pmatrix} 1 \\ 0 \end{pmatrix} (\ 0\ 1) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \ . \tag{E.5.8}$$

Obviously this matrix visualization is valid for any dimension N, not just N=2. For rank n > 2, however, this transpose-of-vector concept does not conveniently generalize. For n=3 the object $\mathbf{u}_a\mathbf{u}_b\mathbf{u}_c$ would be a cube of zeros with a single 1 located at coordinates a,b,c, and so on for n > 3. One cannot write this as $\mathbf{u_a u_b u_c}^T$ for example.

In associating of M = $\mathbf{ab}^T$ with a rank-2 tensor, we must limit our interest to Cartesian space where g = 1. That is because $\mathbf{ab}^T$ is the single matrix shown above in (E.5.5). If g ≠ 1, then there are four different matrices M^{ij}, $M^i{}_j$, $M_i{}^j$ and M_{ij} and the notation M = $\mathbf{ab}^T$ cannot support this fact. When g = 1 all four matrices are the same matrix, *and* the x-space dot product is as shown in (E.5.5). g' ≠ 1 is still allowed.

Ambiguity of **ab**:

There is a certain ambiguity in (E.5.2) that **ab** = $\mathbf{ab}^T$. Writing A = **ab** = $\mathbf{ab}^T$, we imply that "**ab**" is the *name* of a matrix, an alternate to the name A. This matrix A = **ab** has matrix elements $A_{ij} = (\mathbf{ab})_{ij} = a_i b_j$. On the other hand, in (E.3.1) we might say that B = **ab** = **a**⊗**b** which is a vector in a tensor product space. So this version of the object **ab** is not a matrix, it is a tensor product space vector. So A and B are different types of objects both indicated by the notation **ab**. However, the tensor components of the rank-2 tensor B are the same as the matrix elements of the matrix A. That is because $(B)_{ij} = [\mathbf{a}\otimes\mathbf{b}]_{ij} = a_i b_j$, being an outer product of two vectors. In Dirac notation described below in (E.7.4) we would write the objects A and B in this manner which stresses the distinction,

$$A = |\mathbf{a}\rangle\langle\mathbf{b}| \quad = \text{the name of a matrix } \mathbf{ab}^T \text{ (or the corresponding operator in the Dirac Hilbert space)}$$

$$B = |\mathbf{a}\rangle \otimes |\mathbf{b}\rangle = \text{a vector in a tensor product space.} \qquad \text{(E.5.9)}$$

The matrix elements and tensor components (in the $\mathbf{u}_i$ basis) would be

$$A_{ij} = \langle\mathbf{u}_i|\, A\, |\,\mathbf{u}_j\rangle = \langle\mathbf{u}_i|\,\mathbf{a}\rangle\langle\mathbf{b}|\,\mathbf{u}_j\rangle = a_i b_j$$

$$B_{ij} = [\langle\mathbf{u}_i| \otimes \langle\mathbf{u}_j|\,]\ [\,|\mathbf{a}\rangle \otimes |\mathbf{b}\rangle\,] = \langle\mathbf{u}_i|\mathbf{a}\rangle\langle\mathbf{u}_j|\mathbf{b}\rangle = a_i b_j \ . \qquad \text{(E.5.10)}$$

E.6 Large and small dots used with dyadics (Cartesian Space)

Sometimes a small-size dot • is used to indicate the action of a dyadic (matrix) on a vector. If A is a dyadic (same symbol for matrix), and if **c** and **d** are vectors, then one defines:

$$\begin{aligned} A \cdot \mathbf{c} &\equiv A\mathbf{c} = \text{a column vector} &&\Rightarrow (A \cdot \mathbf{c})_i = (A\mathbf{c})_i = A_{ij}c_j &&\Rightarrow A \cdot \mathbf{c} = \Sigma_{ij}A_{ij}c_j\,\mathbf{u}_i \\ \mathbf{c} \cdot A &\equiv \mathbf{c}^T A = \text{a row vector} &&\Rightarrow (\mathbf{c} \cdot A)_i = (\mathbf{c}^T A)_i = c_j A_{ji} &&\Rightarrow \mathbf{c} \cdot A = \Sigma_{ij}c_j A_{ji}\,\mathbf{u}_i \\ \mathbf{d} \cdot A \cdot \mathbf{c} &= \mathbf{d}^T A\mathbf{c} = \text{a number} &&= \Sigma_{ij}d_i A_{ij}c_j \ . && \qquad \text{(E.6.1)} \end{aligned}$$

It then follows that, for the particular dyadic A = **ab** ,

$$\begin{aligned} (\mathbf{ab}) \cdot \mathbf{c} &\equiv (\mathbf{ab})\,\mathbf{c} = (\mathbf{ab}^T)\mathbf{c} = \mathbf{a}(\mathbf{b}^T\mathbf{c}) = \mathbf{a}\,(\mathbf{b} \bullet \mathbf{c}) = (\mathbf{b} \bullet \mathbf{c})\,\mathbf{a} = \text{a column vector} \\ \mathbf{c} \cdot (\mathbf{ab}) &\equiv \mathbf{c}^T(\mathbf{ab}) = \mathbf{c}^T(\mathbf{ab}^T) = (\mathbf{c}^T\mathbf{a})\,\mathbf{b}^T = (\mathbf{c} \bullet \mathbf{a})\,\mathbf{b}^T = \text{a row vector} \\ \mathbf{d} \cdot (\mathbf{ab}) \cdot \mathbf{c} &= \mathbf{d}^T(\mathbf{ab})\,\mathbf{c} = \mathbf{d}^T\mathbf{ab}^T\,\mathbf{c} = (\mathbf{d}^T\mathbf{a})(\,\mathbf{b}^T\,\mathbf{c}) = (\mathbf{d} \bullet \mathbf{a})(\mathbf{b} \bullet \mathbf{c}) = \text{a number} \ . \qquad \text{(E.6.2)} \end{aligned}$$

Here is more detail on the first line of the above group showing a skeletal matrix structure,

$$\begin{array}{ccccccccc} (\mathbf{ab})\,\mathbf{c} & = & (\mathbf{a}\,\mathbf{b}^{\mathrm{T}})\mathbf{c} & = & \mathbf{ab}^{\mathrm{T}}\mathbf{c} & = & \mathbf{a}(\mathbf{b}^{\mathrm{T}}\mathbf{c}) & = & \mathbf{a}(\mathbf{b}\bullet\mathbf{c}) \\ \{\begin{pmatrix} x & x \\ x & x \end{pmatrix}\}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} & = & \{\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}(b_1\ b_2)\}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} & = & \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}(b_1\ b_2)\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} & = & \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}\{(b_1\ b_2)\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}\} & = & \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}\mathbf{b}\bullet\mathbf{c} \end{array} \tag{E.6.3}$$

The same small dot is used to indicate the product of two dyadics, which is to say, matrix multiplication

$$A\bullet B \equiv AB\,. \tag{E.6.4}$$

Regarding this small size dot • : (1) from a matrix algebra point of view, it is completely superfluous except in the case $\mathbf{c} \bullet A \equiv \mathbf{c}^{\mathrm{T}}A$; (2) it is completely different from the dot • used in $\mathbf{b}^{\mathrm{T}}\mathbf{c} = \mathbf{b}\bullet\mathbf{c}$. The next Section provides an explanation of the small dot as part of an operator interpretation for dyadics.

E.7 Operators and Matrices for Rank-2 tensors: the bra-ket notation (Cartesian Space)

We continue with the assumption that x-space is Cartesian, g = 1. The discussion below applies only to rank-2 tensors. One of our goals is to make contact with the dyadic discussion of Morse and Feshbach, one of the few places dyadic notation appears in a general mathematical physics book.

Operator concept. As discussed in Section 5.10, x-space and x'-space of Picture A are both N-dimensional real Hilbert Spaces, with a scalar product indicated by the large dot •, and one can regard **V** as a vector in either space. The first line of (E.2.10) with **A** replaced by **V** states this expansion,

$$\mathbf{V} = \Sigma_i \alpha^i\, \mathbf{b}_i \qquad \alpha^i = \mathbf{b}^i \bullet \mathbf{V} \qquad // \text{ arbitrary basis} \tag*{\textit{(E.2.10)}}$$

which we now rewrite, renaming the coefficients,

$$\mathbf{V} = \Sigma_i [V^{(b)}]^i\, \mathbf{b}_i \qquad [V^{(b)}]^i = \mathbf{b}^i \bullet \mathbf{V} \tag{E.7.1}$$

Moreover, one can regard a rank-2 tensor A as an "operator" in this Hilbert space, using notation (E.5.6),

$$A = \Sigma_{ij}\, [A^{(b)}]^{ij}\, \mathbf{b}_i\mathbf{b}_j^{\mathrm{T}}\,. \tag{E.7.2}$$

In (E.2.10) and (E.5.6) the coefficients in these two expansions were called α^i and α^{ij}, so here we are providing more descriptive names for these coefficients.

Application of $(\mathbf{b}^n)^{\mathrm{T}}$ on the left of (E.7.2) and $\mathbf{b}^m$ on the right, and then a double use of $(\mathbf{b}^n)^{\mathrm{T}}\mathbf{b}_i = \mathbf{b}^n \bullet \mathbf{b}_i = \delta_{n,i}$ (see 7.18.6) gives,

$$[A^{(b)}]^{nm} = (\mathbf{b}^n)^{\mathrm{T}}\, A\, \mathbf{b}^m\,. \tag{E.7.3}$$

Here, one regards A as an operator in the x Hilbert space, whereas $[A^{(b)}]^{nm}$ is a "matrix" which is *associated* with the operator A in the particular $\mathbf{b}^n$ basis. If $\mathbf{b}^n = \mathbf{u}^n$, then $[A^{(u)}]^{nm} = A^{nm}$ which is a particular matrix for this particular $\mathbf{u}^n$ basis. More generally,

$$[A^{(b)}]^{nm} = (\mathbf{b}^T)^n A\, \mathbf{b}^m = [(\mathbf{b}^T)^n]_i A^{ij} [\mathbf{b}^m]_j = [\mathbf{b}^m]_i A^{ij} [\mathbf{b}^m]_j$$

which is in general a completely different matrix formed by linearly combining elements A^{ij}.

Bra-ket Notation. For the author of this document, the bra-ket notation commonly used in quantum mechanics (Paul Dirac 1939) provides a useful way to look at a rank-2 tensor A as an operator. It is true that in quantum mechanics one usually deals with infinite-dimensional Hilbert spaces and complex numbers, but the formalism applies just as well to real Hilbert spaces with finite dimensions (used for example to study "spin"). Here is how the bra-ket notation works:

$\mathbf{b}_i$ → $|b_i>$ // vector
$\mathbf{b}_i^T$ → $<b_i|$ // transpose vector

$\mathbf{b}^i$ → $|b^i>$ // vector
$(\mathbf{b}^i)^T$ → $<b^i|$ // transpose vector

$\mathbf{a}\mathbf{b}^T$ *(E.5.3)* → $|a><b|$ // a matrix (outer product)
$\mathbf{a}^T\mathbf{b} = \mathbf{a} \bullet \mathbf{b}$ *(E.5.5)* → $<a\,|\,b>$ // a number (inner product)

$\mathbf{a} \bullet \mathbf{b} = \mathbf{b} \bullet \mathbf{a}$ *(7.4.14)* → $<a\,|\,b> = <b\,|\,a>$ for any a,b // *real* Hilbert Space

$\mathbf{V} = \Sigma_i\ [V^{(b)}]^i\, \mathbf{b}_i$ *(E.7.1)* → $|V> = \Sigma_i\ [V^{(b)}]^i\, |b_i>$ // vector expansion ...

$[V^{(b)}]^i = \mathbf{b}^i \bullet \mathbf{V}$ *(E.7.1)* → $[V^{(b)}]^i = <b^i|V>$ // and coefficients

$\mathbf{b}^i \bullet \mathbf{b}_j = \delta^i{}_j = \delta_{i,j}$ *(7.18.3)* → $<b^i|b_j> = \delta_{i,j}$ // orthogonality
$= <b_j|b^i> = <b_i|b^j> = <b_j|b^i>$

$\Sigma_i\, \mathbf{b}^i\, \mathbf{b}_i^T = 1$ *(7.18.6)* † → $\Sigma_i\, |b^i><b_i| = 1 = \Sigma_i\, |b_i><b^i|$ // completeness

$A = \Sigma_{ij}\ [A^{(b)}]^{ij}\, \mathbf{b}_i \mathbf{b}_j^T$ *(E.7.2)* → $A = \Sigma_{ij}\, [A^{(b)}]^{ij}\, |\, b_i><b_j|$ // tensor expansion ...

$[A^{(b)}]^{ij} = (\mathbf{b}^i)^T A\, \mathbf{b}^j$ → $[A^{(b)}]^{ij} = <b^i\,|\,A\,|\,b^j>$ // and coefficients

$\mathbf{b}^i \bullet (A\mathbf{b}^j) = \mathbf{b}^j \bullet (A^T\, \mathbf{b}^i)$ *(7.9.17)* → $<b^i\,|\,A\,|\,b^j> = <b^j\,|\,A^T\,|\,b^i>$ // covariant transpose

(E.7.4)

---

† What appears in (7.18.6) for completeness is $(\mathbf{b}^n)_i(\mathbf{b}_n)^j = \delta_i{}^j$, but since g = 1 we can write this with indices down as $(\mathbf{b}^n)_i(\mathbf{b}_n)_j = \delta_{i,j}$ which in matrix notation says $\Sigma_i\, \mathbf{b}^i\, \mathbf{b}_i^T = 1$.

In this bra-ket notation, the N $|b_i>$ are a set of basis vectors which span an N-dimensional real Hilbert Space, while $<b_i|$ span the so-called adjoint (or transpose in our case) Hilbert Space. One then refers to $[A^{(b)}]^{ij} = <b^i \,|\, A \,|\, b^j>$ as "the matrix element of the operator A in the $\mathbf{b}_i$ basis ", since the expansion (E.7.2) was $A = \Sigma_{ij}\, [A^{(b)}]^{ij}\, \mathbf{b}_i\mathbf{b}_j = \Sigma_{ij}\, [A^{(b)}]^{ij}\, \mathbf{b}_i\otimes\mathbf{b}_j$ which is an expansion in the $\mathbf{b}_i$ basis. In general, $|b_i>$ and $|b^i>$ are different vectors because $\mathbf{b}_i$ and $\mathbf{b}^i$ are different. For example $\mathbf{e}^n = g'^{nm}\,\mathbf{e}_m$ and our restriction to $g = 1$ certainly allows $g' \neq 1$. [Do not confuse with $(\mathbf{b}_i)^a = (\mathbf{b}_i)_a$ being the same; here a is a tensor index raised and lowered by $g = 1$, whereas i is a basis vector *label* and these are raised and lowered by w' as shown in (7.18.6).]

In this notation, based on what was presented earlier, one can write,

$A^{nm} = <u^n \,|\, A \,|\, u^m>$ = the x-space components of tensor A (basis $\mathbf{u}_n$) raise/lower with g

$A'^{nm} = <e^n \,|\, A \,|\, e^m>$ = the x'-space components of tensor A (basis $\mathbf{e}_n$) raise/lower with g'

$[A^{(b)}]^{nm} = <b^n \,|\, A \,|\, b^m>$ = the matrix of A in the $\mathbf{b}_n$ basis raise/lower with w (E.7.5)

In the first of these three lines, one can raise and lower indices with g^{ab} and g_{ab} on both sides of the equation, but we are taking $g = 1$ so up and down here do not matter. On the second line this can be done with g'^{ab} and g'_{ab}. Eq (7.18.6) shows that $\mathbf{b}^n = w'^{nm}\,\mathbf{b}_m$ and conversely $\mathbf{b}_n = w'_{nm}\mathbf{b}^m$ where w'_{nm} is the metric tensor g'_{nm} one *would get* for some underlying transformation F_b which causes $\mathbf{b}_n$ to be its tangent base vectors $\mathbf{e}_n$. So, on the third line above we can raise and lower indices on each side with w^{ab} and w_{ab} where $w'_{nm} \equiv \mathbf{b}_n \bullet \mathbf{b}_m$ as in (7.18.6).

Notice in the three equations of (E.7.5) that the operator A between the vertical bars is *the exact same operator* in each case. The matrices are different not because the operator has changed, but because the basis vectors are different.

In a more consistent notation one might write $A^{nm} = [A^{(u)}]^{nm}$ and $A'^{nm} = [A^{(e)}]^{nm}$.

Bases are related by a transformation. Consider again,

$$[A^{(b)}]^{nm} = <b^n \,|\, A \,|\, b^m> = (\mathbf{b}^n)^T A\, \mathbf{b}^m = [\mathbf{b}^n]_i\, A^{ij}\, [\mathbf{b}^m]_j = <b^n|u_i><u^i|A|u^j><u_j|b^m> . \qquad (E.7.6)$$

We lower index m on both sides (using w'_{ab} as noted above) and reverse the j tilt to get

$$[A^{(b)}]^n{}_m = <b^n \,|\, A \,|\, b_m> = (\mathbf{b}^n)^T A\, \mathbf{b}_m = [\mathbf{b}^n]_i\, A^i{}_j\, [\mathbf{b}_m]^j = <b^n|u_i><u^i|A|u_j><u^j|b_m> . \qquad (E.7.7)$$

One could then define the following tensor-like object,

$$B^n{}_i \equiv [\mathbf{b}^n]_i \quad . \qquad (E.7.8)$$

The first index on B is raised and lowered by w', while the second is raised and lowered by g, so this object is a bit like R and S in its non-tensor nature. Lowering n and raising i then gives

$$B_n{}^i = [\mathbf{b}_n]^i = (B^T)^i{}_n \ , \tag{E.7.9}$$

where we use the covariant transpose of a tilted matrix described in (7.9.3). One then has

$$[A^{(b)}]^n{}_m = B^n{}_i \, A^i{}_j (B^T)^j{}_m \, . \tag{E.7.10}$$

Since all the matrices are tilted the same way and summed indices are contractions, this is one of the "legal" Standard Notation matrix multiplication forms like (7.8.6) and we then write,

$$A^{(b)} = BAB^T \qquad \text{or more precisely} \qquad [A^{(b)} = BAB^T]_{SN,dt} \tag{E.7.11}$$

where SN,dt means Standard Notation, down-tilt, as described below (7.8.6). The matrix equation $A^{(b)} = BAB^T$ shows that the $[A^{(b)}]^n{}_m$ are related to the $A^i{}_j$ by a "congruence transformation" with a matrix $B_n{}^i = [\mathbf{b}_n]^i$ whose rows are the basis vectors $\mathbf{b}_n$. When $\mathbf{b}_m = \mathbf{u}_m$, matrix B is the identity matrix, and when $\mathbf{b}_m = \mathbf{e}_m$ one has $B_n{}^i = [\mathbf{e}_n]^i = R_n{}^i$, so that B = R in this case. In Standard Notation the general R matrix is "covariant real orthogonal", $[RR^T = 1]_{SN,dt}$ and $[R^T = R^{-1}]_{SN,dt}$ (see (7.9.3) and following text) so in fact one has for the $\mathbf{b}_m = \mathbf{e}_m$ basis,

$$A^{(e)} = B\,A\,B^T = R\,A\,R^T = R\,A\,R^{-1} = R\,A\,S \, . \tag{E.7.12}$$

Specifically in this case,

$$[A^{(e)}]^n{}_m = R^n{}_i A^i{}_j S^j{}_m = R^n{}_i R_m{}^j A^i{}_j = A'^n{}_m \, . \tag{E.7.13}$$

which is the expected result looking at the second line of (7.5.8).

<u>Comment</u>: Recall that in Developmental Notation one has $RR^T = 1$ and $R^T = R^{-1}$ only when R is a rotation, whereas in Standard Notation one has $RR^T = 1$ and $R^T = R^{-1}$ for any R matrix, see (7.9.3). We used this fact above to show that a basis change from basis $\mathbf{u}_n$ to $\mathbf{b}_n$ on operator A can be thought of as a congruence transformation by a matrix B whose rows are the vectors $\mathbf{b}_n$. This is a standard concept in linear algebra where one operates in Cartesian x-space.

<u>More on bra-ket notation and its relation to the small dyadic dot.</u>

Consider the following facts, where A^T is the covariant transpose as in (7.9.3) and (7.9.17),

$$<d\,|\,A\,|\,c> \ = \ \mathbf{d}^T A\,\mathbf{c} \ = \mathbf{d}^T [A\,\mathbf{c}\,] \qquad = \ <d\,|Ac>$$

$$<d\,|\,A\,|\,c> \ = \ \mathbf{d}^T A\,\mathbf{c} \ = [\,\mathbf{d}^T A]\,\mathbf{c} \ = \ [A^T \mathbf{d}]^T \mathbf{c} \qquad = <A^T d\,|\,c> \tag{E.7.14}$$

where

|(Ac) > = a new Hilbert space vector which results when operator A is applied to |c> = A|c>

$<(A^T d)\,|$ = a new transpose Hilbert space vector which results when A is applied to <d| = <d|A .

(E.7.15)

Detail: In covariant notation $[\mathbf{d}^T A] = [A^T \mathbf{d}]^T$ because $[\mathbf{d}^T A]^i = ([A^T \mathbf{d}]^T)^i$. To verify this,

$$[\mathbf{d}^T A]^i = (\mathbf{d}^T)^j A_j{}^i = d^j A_j{}^i = A_j{}^i d^j$$

$$([A^T \mathbf{d}]^T)^i = [A^T \mathbf{d}]^i = (A^T)^i{}_j d^j = A_j{}^i d^j$$

So one has this general idea that

$$\langle d \,|\, A \,|\, c \rangle = \langle d \,|Ac \rangle = \langle A^T d \,|\, c\rangle$$

$$A \,|\, c \rangle = |(Ac)\rangle , \quad \langle d \,|\, A = \langle (A^T d) \,| \quad . \tag{E.7.16}$$

In this last line, the isolated A's are the same operator A sitting in the Hilbert space. This operator can "act" either to the right or to the left as shown. The object $|(Ac)\rangle \equiv |e\rangle$ is some different vector in the Hilbert space (different from $|c\rangle$), call it $|e\rangle$, and the grouping (Ac) labels this vector. Similarly, $\langle (A^T d) \,|$ is some vector $\langle f|$ in the transpose Hilbert space. The distinction between A as an abstract operator in the Hilbert space, and the A in (Ac) and $(A^T d) = (d^T A)^T$ as vectors in the Hilbert space is a subtle one. It is just this distinction that is implied by the small dot in the dyadic notation discussed in the previous section, and here is the correspondence between the dyadic notation and the bra-ket notation:

$$\begin{array}{llll} A \bullet \mathbf{c} = A\mathbf{c} & \mathbf{d} \bullet A = (A^T \mathbf{d})^T = \mathbf{d}^T A & \mathbf{d} \bullet A \bullet \mathbf{c} = \mathbf{d}^T A\mathbf{c} & A \bullet B\, \mathbf{c} \\ A \,|\, c \rangle = |Ac\rangle & \langle d \,|\, A = \langle (A^T d)| & \langle d \,|\, A \,|\, c \rangle = \langle d \,|\, Ac \rangle & AB|\, c \rangle \ . \end{array} \tag{E.7.17}$$

In the rightmost column operator B is applied first to $|c\rangle$ to get vector $|(Bc)\rangle$, and then operator A is applied to $|(Bc)\rangle$ to give yet another vector $|\,(ABc)\rangle$. In bra-ket notation the product of two abstract operators is given just as AB, but in dyadic notation it is written A • B.

Dyadics as operators. According to the above discussion, one can regard a dyadic (**AB**), being a rank-2 tensor, as an *operator* and not as a matrix. The matrix $T^{nm} = (\mathbf{AB})^{nm} = A^n B^m$ is specific to the $\mathbf{u_n}$ basis in x-space,

$$T^{nm} = (\mathbf{AB})^{nm} = \langle u^n \,|(\mathbf{AB})|\, u^m \rangle = (\mathbf{u}^n)^T \mathbf{A}\, \mathbf{B}^T \mathbf{u}^m$$

$$= [(\mathbf{u}^n)^T]_a \mathbf{A}^a (\mathbf{B}^T)^b [\mathbf{u}^m]_b = \delta^n{}_a A^a B^b \delta^m{}_b = A^n B^m \,. \tag{E.7.18}$$

In the generic $\mathbf{b_n}$ basis one has

$$[(\mathbf{AB})^{(b)}]^{nm} = \langle b^n \,|(\mathbf{AB})|\, b^m \rangle \,. \tag{E.7.19}$$

It is to emphasize this operator view of a dyadic that Morse and Feshbach use fancy letters like $\mathcal{U}$ to represent dyadics. Then their small-dot notation $\mathcal{U} \bullet \mathbf{B}$ emphasizes the idea of an operator acting on a vector, equivalent to $\mathcal{U}|\, B\rangle$. Here are a few samples from their Section 1.6 on dyadics. Under each clip we have tried to relate the stated equation(s) to our notation above. Like us, Morse and Feshbach are working in a Cartesian x-space, so up and down indices are the same for x-space objects. Note that $\mathbf{u}^n = \mathbf{u}_n$ when g = 1. The M&F symbols $\boldsymbol{\mathcal{U}}$ and $\mathbf{a}_n$ are our A and $\mathbf{u}_n$:

$$\mathfrak{A} \cdot \mathbf{B} = \sum_{mn} \mathbf{a}_m A_{mn} B_n; \quad \mathbf{B} \cdot \mathfrak{A} = \sum_{mn} B_m A_{mn} \mathbf{a}_n \qquad (1.6.5)$$

p 55

$A \bullet \mathbf{b} \ = \ \Sigma_{mn}\ \mathbf{u}_m\ A_{mn}\ b_n \qquad \mathbf{b} \bullet A \ = \ \Sigma_{mn}\ b_m\ A_{mn}\ \mathbf{u}_n$ *(E.6.1)*

or $A\ |b> \ = \ \Sigma_{mn}\ |u_m><u_m|A|u_n><u_n|b> \qquad <b|A \ = \ \Sigma_{mn}<b|u_m><u_m|A|u_n><u_n|$ (E.7.20)

$$\mathfrak{A} = \mathbf{a}_1 A_{11} \mathbf{a}_1 + \mathbf{a}_1 A_{12} \mathbf{a}_2 + \mathbf{a}_1 A_{13} \mathbf{a}_3 + \mathbf{a}_2 A_{21} \mathbf{a}_1 + \mathbf{a}_2 A_{22} \mathbf{a}_2 + \mathbf{a}_2 A_{23} \mathbf{a}_3 + \mathbf{a}_3 A_{31} \mathbf{a}_1 + \mathbf{a}_3 A_{32} \mathbf{a}_2 + \mathbf{a}_3 A_{33} \mathbf{a}_3 \quad (1.6.6)$$

p 55

$A \ = \Sigma_{nm}\ \mathbf{u}_n A_{nm}\ \mathbf{u}_m$ *(E.5.7)* (E.7.21)

$$(\mathfrak{A}^{-1}) \cdot \mathfrak{A} = \mathfrak{A} \cdot (\mathfrak{A}^{-1}) = \mathfrak{J}$$

$A^{-1} \bullet A \ = \ A \bullet A^{-1} \ = \ 1$ // A^{-1} defined if matrix $[A^{(u)}]_{mn} = A_{mn}$ is invertible (E.7.22)

There is, of course, a zero dyadic $\mathfrak{O}$ and a unity dyadic $\mathfrak{J}$ called the *idemfactor:*

$$\mathfrak{O} \cdot \mathbf{F} = 0; \quad \mathfrak{J} \cdot \mathbf{F} = \mathbf{F}; \quad \mathfrak{J} = \mathbf{a}_1\mathbf{a}_1 + \mathbf{a}_2\mathbf{a}_2 + \mathbf{a}_3\mathbf{a}_3$$

where **F** is any vector.

p 57

$0 \bullet \mathbf{f} = \mathbf{0} \qquad 1 \bullet \mathbf{f} = \mathbf{f}; \qquad 1 = \Sigma_n\ \mathbf{u}_n\mathbf{u}_n^{\mathbf{T}}$ *(E.7.4)* with $\mathbf{b}_n = \mathbf{u}_n$

or

$0\,|\,f> \ = \ |0>; \qquad 1\,|\,f> \ = \ |\,f>; \qquad 1 = \ \Sigma_n\ |\,u_n><u_n|$ *(E.7.4)* (E.7.23)

$$\mathfrak{A}_s = A_x\mathbf{ii} + B_z\mathbf{ij} + B_y\mathbf{ik} + B_z\mathbf{ji} + A_y\mathbf{jj} + B_x\mathbf{jk} + B_y\mathbf{ki} + B_x\mathbf{kj} + A_z\mathbf{kk}$$

p 59

$U_s = \Sigma_{nm}\,[U_s]_{nm}\,\mathbf{u}_n\mathbf{u}_m \ = \ [U_s]_{11}\mathbf{u}_1\mathbf{u}_1 + [U_s]_{12}\mathbf{u}_1\mathbf{u}_2\ + \ldots \ = \ [U_s]_{11}\,\mathbf{i}\,\mathbf{i}\ +\ [U_s]_{12}\mathbf{i}\,\mathbf{j}\ + ..$

where $\mathbf{i} = \mathbf{u}_1$, $\mathbf{j} = \mathbf{u}_2$, $\mathbf{k} = \mathbf{u}_3$ and where $[U_s]_{ij} = \begin{pmatrix} A_x & B_z & B_y \\ B_z & A_y & B_x \\ B_y & B_x & A_z \end{pmatrix}$ = a symmetric matrix (E.7.24)

Notice the impressive name "idemfactor" for the identity operator $1 = \Sigma_i\ |\ a_i><a_i|\ = \Sigma_i\ \mathbf{a}_i\mathbf{a}_i^{\mathbf{T}} = \Sigma_i\mathbf{a}_i\mathbf{a}_i$ where $\Sigma_i\mathbf{a}_i\mathbf{a}_i$ uses the dyadic notation (E.5.2) .

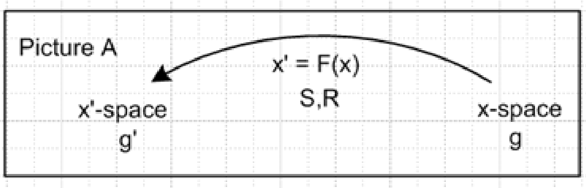
Comment: A favorite bra-ket notation trick is to obtain useful results by inserting $1 = \Sigma_n\ |\ u_n><u_n\ |$ in opportune places, as in (E.7.20) above :

$$A\,|b> = 1\,A\,1\;|b> = [\Sigma_m\;|\,u_m><u_m\,|\,]\;A\;\;[\Sigma_n\;|\,u_n><u_n\,|\,]\;\;|b>\;\;= \Sigma_n\Sigma_m\,|\,u_m><u_m\,|A|\,u_n><u_n\,|b>$$
$$= \Sigma_n\Sigma_m\,|\,u_m>\,A_{mn}\,b_n\;\;= \Sigma_{nm}\,A_{mn}\,b_n\;\;|u_m> \qquad \text{(E.7.25)}$$

The vector $A\,|b> = |Ab>$ has thus been expanded on the basis $|\,u_m>$.

E.8 Expansions of tensors on unit tangent base vectors: M and N

We start with the general Picture A (and later specialize to orthogonal coordinates),



(E.8.1)

In (E.2.9) it was established that one can expand a tensor A on the tangent base vectors $\mathbf{e}_n$ as

$$A = \Sigma_{ijk\ldots}\;A'^{\,ijk\ldots}\;(\mathbf{e}_i\otimes\mathbf{e}_j\otimes\mathbf{e}_k\ldots) \qquad A'^{\,ijk\ldots} = \text{contravariant components of A in x'-space}$$

$$A'^{\,ijk\ldots}\;\;= R^i{}_{i'}R^j{}_{j'}R^k{}_{k'}\ldots\ldots\;\;A^{\,i'j'k'\ldots}\;\;. \qquad \text{// tensor transformation rule} \qquad \text{(E.8.2)}$$

We know from (7.13.6) that $\mathbf{e}_n = |\mathbf{e}_n|\;\hat{\mathbf{e}}_n = h'_n\;\hat{\mathbf{e}}_n$ so the above expansion for tensor A can be written

$$A = \Sigma_{ijk\ldots}\;\{\,h'_i h'_j h'_k\ldots\ldots A'^{\,ijk\ldots}\}\;\;(\hat{\mathbf{e}}_i\otimes\hat{\mathbf{e}}_j\otimes\hat{\mathbf{e}}_k\ldots)$$

$$\equiv\;\Sigma_{ijk\ldots}\;[A^{(\hat{e})}]^{ijk\ldots}\;\;(\hat{\mathbf{e}}_i\otimes\hat{\mathbf{e}}_j\otimes\hat{\mathbf{e}}_k\ldots)\,, \qquad \text{(E.8.3)}$$

where the unit-vector expansion coefficients are given by

$$[A^{(\hat{e})}]^{ijk\ldots}\;\;= \{\,h'_i h'_j h'_k\ldots\ldots A'^{\,ijk\ldots}\} \qquad \text{// } \equiv\;\mathcal{a}'^{\,ijk\ldots}\;\;\text{see (E.8.5)}$$

$$= h'_i h'_j h'_k\ldots\ldots\;R^i{}_{i'}R^j{}_{j'}R^k{}_{k'}\ldots\ldots\;\;A^{\,i'j'k'\ldots}$$

$$= (h'_i\,R^i{}_{i'})(\,h'_j\,R^j{}_{j'})(\,h'_k\,R^k{}_{k'})\;\ldots\ldots\;A^{\,i'j'k'\ldots} \qquad \text{(E.8.4)}$$

Note that we have not at this point assumed that the $\{\hat{\mathbf{e}}_i\}$ unit vectors are orthonormal, so g' can be non-diagonal.

Coefficient notation

In a curvilinear coordinates application of these expansions, the expansion coefficients are usually written in the following manner,

$$[A^{(\hat{e})}]^{ijk\ldots} = \mathcal{A}'^{\,ijk\ldots} = A^{x'_ix'_jx'_k} \cdots\cdots \tag{E.8.5}$$

where the x'_n are the *names* of the coordinates. For example, for a rank-4 tensor in spherical coordinates with coordinates $x'_1 = r$, $x'_2 = x'_3 = \theta$ and $x'_3 = \varphi$ one might write

$$[A^{(\hat{e})}]^{2213} = \mathcal{A}'^{2213} = A^{\theta\theta r\varphi} \quad . \tag{E.8.6}$$

The script notation is the same shorthand we used in (7.13.12) where $\mathbf{V} = V'^n \mathbf{e}_n = \mathcal{V}'^n \hat{\mathbf{e}}_n$. Note that there is a prime on $\mathcal{V}'^n$ and a prime on $\mathcal{A}'^{\,ijk\ldots}$ above. *The reason* for this prime is that these quantities are associated with the curvilinear coordinates (variables) which in Picture A and B are called x'^i . For example we have $\mathcal{V}'^n = h'_n V'^n$. In the Moon & Spencer Picture of (14.11) with u-space (metric tensor g) on the left, the curvilinear coordinates are called u^i. Since they then have no prime, we would use unprimed scripted variables. For example, $\mathcal{V}^n = h_n V^n$. This is the convention of Chapter 14, but in these Appendices we are using Pictures A and B.

As shown below, one may interpret $[A^{(\hat{e})}]^{ijk\ldots} = \mathcal{A}'^{ijk\ldots}$ as the contravariant components of tensor A with respect to a certain transformation F_M, and in this sense the up and down index position is significant. However, when we later assume the $\hat{\mathbf{e}}_i$ are orthonormal, that forces g'_{ij} to be diagonal and forces the condition $\hat{\mathbf{e}}_i = \hat{\mathbf{e}}^i$ as shown below (E.2.14). Then the up or down index position on $[A^{(\hat{e})}]^{ijk\ldots}$ makes no difference in expansion (E.8.3). In the spherical coordinates example where the $\hat{\mathbf{e}}_i$ *are* orthonormal, we would normally write the above coefficient as $A_{\theta\theta r\varphi}$.

Matrices M and N

It is convenient now to define some modified R matrices as

$$M^a{}_b \equiv h'_a R^a{}_b \tag{E.8.7}$$

so that (E.8.4) becomes

$$[A^{(\hat{e})}]^{ijk\ldots} = M^i{}_{i'} M^j{}_{j'} M^k{}_{k'} \ldots\ldots \; A^{i'j'k'\ldots} \quad . \tag{E.8.8}$$

As shown below, these matrices M are the R-matrices of a transformation called F_H that takes x-space to a new x"-space in which the tangent base vectors are the unit vectors $\hat{\mathbf{e}}_n$, but we defer that interpretation to first get some facts laid out.

Defining $N^a{}_b$ to be the **inverse** of $M^a{}_b$, one has (verified just below),

$$N^a{}_b \equiv h'_b{}^{-1}S^a{}_b \ = h'_b{}^{-1}R_b{}^a \tag{E.8.9}$$

so the inversion of (E.8.8) is given by

$$A^{ijk...} = N^i{}_{i'}N^j{}_{j'}N^k{}_{k'}......\ [A^{(\hat{e})}]^{i'j'k'...} \quad . \tag{E.8.10}$$

To verify that this N is the correct inverse of M, use (E.8.7) and (E.8.9) to show

$$M^a{}_kN^k{}_c = (h'_a R^a{}_k)(h'_c{}^{-1}R_c{}^k) \ = (h'_a/ h'_c) R^a{}_k R_c{}^k \ = \ (h'_a/ h'_c)\delta^a{}_c \ = \delta^a{}_c \tag{E.8.11}$$

making use of orthogonality rule #3 of (7.6.4), $R^a{}_k R_c{}^k = \delta^a{}_c$. M and N can be written in terms of the $\mathbf{e}^n$ and $\mathbf{e}_n$ vectors as follows,

$$M^n{}_i \equiv h'_n R^n{}_i \ = h'_n(\mathbf{e}^n)_i \qquad // (7.13.1)$$

$$N^i{}_n = h'_n{}^{-1}R_n{}^i \ = h'_n{}^{-1}(\mathbf{e}_n)^i \ = \ (\hat{\mathbf{e}}_n)^i \ , \qquad // (7.13.1) \tag{E.8.12}$$

which says that

$$N^i{}_n \ = \{ \ \hat{\mathbf{e}}_1 \ , \ \hat{\mathbf{e}}_2, \ .... \} \ . \tag{E.8.13}$$

Thus the columns of $N^i{}_n$ are the unit tangent base vectors.

Matrix H and x"-space

In the discussion above one has x-space with basis vectors $\mathbf{u}_n$ and x'-space with basis vectors $\mathbf{e}_n$ (the tangent base vectors). It is useful then to define x"-space as the space whose basis vectors are the $\hat{\mathbf{e}}_n$ unit vectors, which are generally not orthogonal. The relation (E.8.7) $M^a{}_b \equiv h'_a R^a{}_b$ can be written as a down-tilt matrix equation,

$$M = HR \qquad \text{where} \qquad H^i{}_j \equiv \text{diag}(h'_1, h'_2.....) \quad . \tag{E.8.14}$$

It then follow that

$$N = M^{-1} = R^{-1}H^{-1} = SH^{-1} \ . \tag{E.8.15}$$

Finally, note that

$$\mathbf{x} = x^i\mathbf{u}_i = x'^i\mathbf{e}_i = x'^i(h'_i\hat{\mathbf{e}}_i) \ = x''^i\hat{\mathbf{e}}_i \quad \Rightarrow \quad x''^i = h'_i \ x'^i$$

so

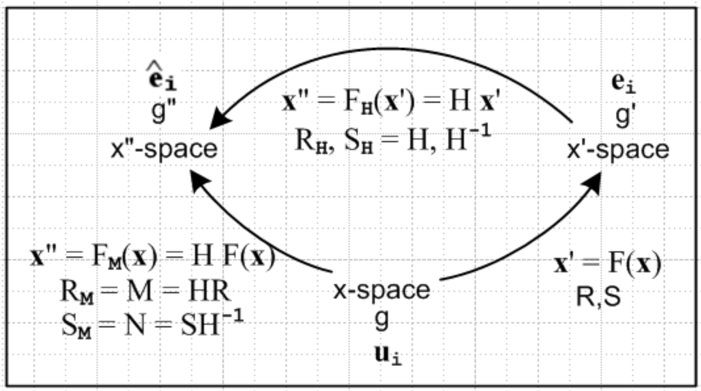
$$\mathbf{x}'' = H\ \mathbf{x}' \ . \tag{E.8.16}$$

This shows that the transformation $\mathbf{F_H}$ from x'-space to x"-space is linear with matrix H,

$$\mathbf{x}'' = \mathbf{F_H}(\mathbf{x}') = H\,\mathbf{x}' \ . \tag{E.8.17}$$

Eq. (7.5.2) then says that the corresponding R-matrix is $(R_H)^i{}_k \equiv (\partial x''^i/\partial x'^k) = H^i{}_k$ so $R_H = H$ and then $S_H = R_H^{-1} = H^{-1}$.

We can now show all three spaces in the same Picture as follows,



(E.8.18)

This picture shows that the transformation directly from x-space to x"-space is

$$\mathbf{F_M}(\mathbf{x}) = H\,\mathbf{F}(\mathbf{x}) \ . \tag{E.8.19}$$

Here $\mathbf{F}(\mathbf{x})$ is a (generally) non-linear transformation assumed to connect x'-space to x-space. This is then concatenated with linear transformation H to get non-linear transformation $\mathbf{F_M}(\mathbf{x})$.

We can now write $A^{(\hat{e})}$ as A" and restate (E.8.3) and (E.8.8) above as

$A = \Sigma_{ijk...}\, A''^{ijk...}\, (\hat{\mathbf{e}}_i\otimes\hat{\mathbf{e}}_j\otimes\hat{\mathbf{e}}_k...)$ // rank-n tensor expanded on $\hat{\mathbf{e}}_i\otimes\hat{\mathbf{e}}_j\otimes\hat{\mathbf{e}}_k...$

$A''^{ijk...} = M^i{}_{i'} M^j{}_{j'} M^k{}_{k'} ......\ A^{i'j'k'...}$ // rank-n tensor transformation

$A^{ijk...} = N^i{}_{i'} N^j{}_{j'} N^k{}_{k'} ......\ A''^{i'j'k'...}$ // inverse of the above

$A''^{i} = M^i{}_j\, A^j$ // rank-1 tensor

$A''^{ij} = M^i{}_{i'} M^j{}_{j'} A^{i'j'}$ // rank-2 tensor

where $A''^{ijk...} \equiv [A^{(\hat{e})}]^{ijk...} \equiv \mathcal{a}'^{ijk...}$ (E.8.20)

The word "tensor" suddenly has a new meaning in the above equations. The equations indicate objects being tensors with respect to this non-linear transformation $\mathbf{F_M}(\mathbf{x})$ whose linearized-at-a-point matrix is $R_M = M = HR$, where R is the linearized-at-a-point matrix version of $\mathbf{F}(\mathbf{x})$. H is the diagonal matrix of scale factors h'_i which are associated with g'_{ij} in x'-space. The $A^{ijk\cdots}$ are contravariant components of tensor A in x-space, while $A''^{ijk\cdots}$ are the corresponding contravariant components of A in x"-space, all with respect to $\mathbf{F_M}(\mathbf{x})$ and its matrix $R_M = M$ where, for example, $d\mathbf{x}'' = M\, d\mathbf{x}$.

One can compare the last line of (E.8.20) to the first line of the generic (7.5.8) showing how a rank-2 tensor transforms. The second line of (7.5.8) then tells us the manner in which the mixed tensor $A^a{}_b$ transforms under F_M :

$$A''^a{}_b = M^a{}_{a'}\ M_b{}^{b'}\ A^{a'}{}_{b'}\ =\ M^a{}_{a'}A^{a'}{}_{b'}M_b{}^{b'}\ . \tag{E.8.21}$$

If we use the Standard Notation covariant transpose shown in (7.9.3) [$R_M = M$], then $M_b{}^{b'} = (M^T)^{b'}{}_b$ and the above becomes

$$A''^a{}_b =\ M^a{}_{a'}A^{a'}{}_{b'}\ (M^T)^{b'}{}_b = [\ MAM^T\ ]^a{}_b$$

from which we conclude that

$$A'' = MAM^T \qquad\qquad // \text{ that is to say, } [\ A'' = MAM^T]_{SN,dt} \tag{E.8.22}$$

Comment:

This matrix result is a special case of the generic result (E.7.11) that $A^{(b)} = BAB^T$ when $\mathbf{b_n} = \hat{\mathbf{e}}_n$, but it is not obvious how this works out so here are some details. In order to compute $B^n{}_i \equiv [\mathbf{b}^n]_i$ in (E.7.8) one has to know $\mathbf{b}^n$. One sees from (7.18.6) that $\mathbf{b}^n = w'^{ni}\ \mathbf{b}_i$ where

$$w'_{nm} = \mathbf{b}_n \bullet \mathbf{b}_m = \hat{\mathbf{e}}_n \bullet \hat{\mathbf{e}}_m = (h'_n h'_m)^{-1}\ \mathbf{e}_n \bullet \mathbf{e}_m = (h'_n h'_m)^{-1} g'_{nm} \quad \Rightarrow$$

$$w'^{nm} = (h'_n h'_m) g'^{nm}\ . \qquad\qquad // \text{ since } w_{nm} \text{ and } w^{nm} \text{ are inverses} \tag{E.8.23}$$

Therefore,

$$\mathbf{b}^n = w'^{nm}\ \mathbf{b}_m = (h'_n h'_m) g'^{nm}\ \hat{\mathbf{e}}_m = h'_n g'^{nm}\ \mathbf{e}_m = h'_n\ \mathbf{e}^n \qquad \Rightarrow$$

$$B^n{}_i \equiv [\mathbf{b}^n]_i = h'_n\ (\mathbf{e}^n)_i \tag{E.8.24}$$

But (E.8.12) says $M^n{}_i = h'_n(\mathbf{e}^n)_i$ so we have shown that $B^n{}_i = M^n{}_i$ so $BAB^T = MAM^T$.

The metric tensor g" in x"-space appearing in Fig (E.8.18) is this,

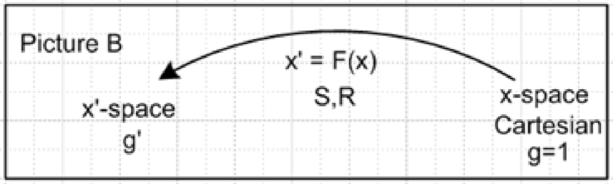
$$g''_{ab} = \mathbf{e}''_a \bullet \mathbf{e}''_b = \hat{\mathbf{e}}_a \bullet \hat{\mathbf{e}}_b = (h'_a h'_b)^{-1}\ \mathbf{e}_a \bullet e_b\ =\ (h'_a h'_b)^{-1}\ g'_{ab}\ =\ w'_{ab}\ , \tag{E.8.25}$$

so g" = w' in the Comment above.

At this point the x and x' spaces are completely general, and none of R, $R_H = H$, $R_M = M$ is a rotation matrix. In the next Section, we shall specialize the above picture so that x-space is Cartesian with g = 1, and x'-space has *orthogonal* curvilinear coordinates **x**' which means g' is a diagonal matrix. In this scenario, **F**(**x**) is in general non-linear and so then is $\mathbf{F_M}(\mathbf{x}) = H\,\mathbf{F}(\mathbf{x})$. Since the $\hat{\mathbf{e}}_i$ now form a frame of *orthonormal* vectors, and since $\mathbf{u}_i$ also form such a frame, one will not be surprised to find that M is now a rotation which relates these two frame sets.

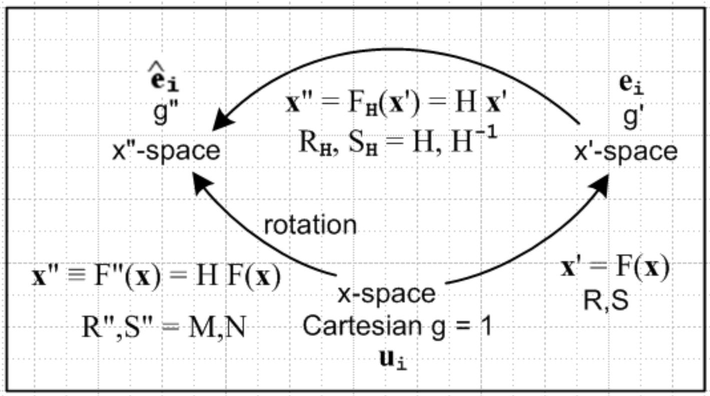
E.9 Application of Section E.8 to Orthogonal Curvilinear Coordinates

We now switch to Picture B (g=1) *and* assume that the x'_i are orthogonal coordinates,



(E.9.1)

and our three-frame picture (E.8.18) above then becomes



(E.9.2)

In this situation, x-space is Cartesian with

$$g_{ab} = g^{ab} = \delta_{a,b} \qquad g'_{ab} = h'^{2}_{a}\,\delta_{a,b} \qquad g'^{ab} = h'^{-2}_{a}\,\delta_{a,b}.$$

$$g''_{ab} = (h'_a h'_b)^{-1}\, g'_{ab} = (h'_a h'_b)^{-1} h'^{2}_{a}\,\delta_{a,b} = \delta_{a,b} \qquad \text{// from (E.8.25)} \quad . \qquad \text{(E.9.3)}$$

Notice that $g''_{ab} = g''(\mathbf{x})_{ab} = \hat{\mathbf{e}}_a(\mathbf{x}) \bullet \hat{\mathbf{e}}_b(\mathbf{x}) = \delta_{ab}$ so x"-space is locally Cartesian at any point **x**. For example, in polar coordinates one has $\hat{\mathbf{r}} \bullet \hat{\boldsymbol{\theta}} = 0$ at any location **x**. Since g = 1 and g" = 1, the up or down index position on a tensor object is not significant.

Now, regardless of what R and S are, M is a "rotation", where we include in this term possible axis reflections. To show this fact, first note from (E.8.7) and (7.5.9) that

$$M^j{}_k \;= (h'_j\, R^j{}_k) \;= (h'_j\, g'^{ja}R_{ak}) \;=\; (h'_j\;[\delta^{ja}\, h'_j{}^{-2}]R_{ak}) \;=\; (h'_j{}^{-1}R_{jk}) \;= (h'_j{}^{-1}R_j{}^k)\,. \qquad \text{(E.9.4)}$$

Using orthogonality rule #3 that $R^a{}_b\, R_c{}^b = \delta^a{}_c = \delta_{a,c}$ it then follows that,

$$M^i{}_k M^j{}_k \;=\; (h'_i\, R^i{}_k)\,(h'_j{}^{-1}R_j{}^k) \;=\; h'_i h'_j{}^{-1}\,(\,R^i{}_k R_j{}^k) = h'_i h'_j{}^{-1}\,\delta_{i,j} \;= \delta_{i,j}\,. \qquad \text{(E.9.5)}$$

But $M^i{}_k M^j{}_k = \delta_{i,j}$ is the Standard Notation statement that M is a "rotation", as was shown in (7.9.11). In developmental notation this appears as $[MM^T = 1]_{DN}$, as shown there. Since M is a rotation, N must be the inverse rotation since $N = M^{-1}$. From (7.9.11) we then also know that

$$M^i{}_k = M_i{}^k \qquad \text{(E.9.6)}$$

which is the alternate form for $M^i{}_k$ being a rotation, Again, M is the R-matrix R_M = M for F_M.

Relation between M and N. Looking at (E.9.5),

$$M^a{}_b\, M^c{}_b \;= \delta_{a,c} \qquad \textit{(E.9.5)}$$

and knowing that

$$M^a{}_b\,(M^{-1})^b{}_c \;= \delta^a{}_c = \;\delta_{a,c} \qquad \text{(E.9.7)}$$

one concludes that $(M^{-1})^b{}_c \;= M^c{}_b$. But $(M^{-1})^b{}_c = N^b{}_c$ so

$N^b{}_c = M^c{}_b$ // note: this does *not* say that $N = M^T$ in standard notation, see (7.9.3) (E.9.8)

which can be verified from the above expressions for N and M. Since (E.9.6) must also be true for rotation N, we find that

$M_i{}^k = M^i{}_k = N^k{}_i = N_k{}^i$ // g = 1, g' = diagonal (E.9.9)

Interpretation of N and M. Since $\mathbf{e_n} = S\,\mathbf{u_n}$ (this is (3.2.4) and (3.2.1) with $\mathbf{e'_n} = \mathbf{u_n}$) and since $\mathbf{e_n} = h'_n \hat{\mathbf{e}}_n$, it follows that

$$\hat{\mathbf{e}}_n = h'_n{}^{-1}\mathbf{e_n} \;=\; h'_n{}^{-1}\, S\, \mathbf{u_n}$$

or

$$(\hat{\mathbf{e}}_n)^a = h'_n{}^{-1}\, S^a{}_b\,(\mathbf{u_n})^b = h'_n{}^{-1}\, S^a{}_b\,\delta_n{}^b \;=\; h'_n{}^{-1}\, S^a{}_n \;= N^a{}_n \text{ (E.8.9)} \;= N^a{}_b\delta^b{}_n = N^a{}_b(\mathbf{u_n})^b$$

or

$\hat{\mathbf{e}}_n = N\,\mathbf{u_n}$ and $(\hat{\mathbf{e}}_n)^a = N^a{}_n$. // $\Rightarrow$ $\mathbf{u_n} = M\,\hat{\mathbf{e}}_n$ (E.9.10)

Since the $\mathbf{u_n}$ are the Cartesian unit vectors, it seems intuitively obvious that the transformation that moves this frame of orthonormal unit vectors $\{\mathbf{u_n}\}$ into the orthonormal frame $\{\hat{\mathbf{e}}_n\}$ must be a "rotation".

It was shown in (E.9.8) that $N^a{}_n = M^n{}_a$, therefore

$$M^n{}_a = N^a{}_n = (\hat{e}_n)^a \ . \qquad (E.9.11)$$

The rotation matrix $N^a{}_n = (\hat{e}_n)^a$ has the orthonormal basis vectors $\hat{e}_n$ as its columns, while the rotation matrix $M^n{}_a = (\hat{e}_n)^a$ has the $\hat{e}_n$ as its rows. Recall from (E.2.15) and (E.2.19) that orthonormal $\hat{e}_n$ implies both $\hat{e}_n = \hat{e}^n$ and $g'_{nm} = h'_n{}^2\delta_{n,m}$ = diagonal.

The relation $\mathbf{u_n} = M\ \hat{e}_n$ can be written $\mathbf{u_n} = M(\mathbf{x})\ \hat{e}_n(\mathbf{x})$ to emphasize that the rotation $M(\mathbf{x}) = R_M(\mathbf{x})$ is really a different rotation at every point $\mathbf{x}$, since the $\hat{e}_n(\mathbf{x})$ vary with $\mathbf{x}$. This is very different from a global rotation which is the same at all points. For a global rotation R, F = R = linear and $\partial_i u_j$ is a tensor. For R_M being a rotation which varies from point to point, F_M is non-linear just as F defining the curvilinear coordinates is non-linear, and $\partial_i u_j$ fails to be a tensor under either F or F_M.

Matrix form of rank-2 tensor transformation:

One implication of the above picture relates to tensor equations being covariant, as discussed in Section 7.15. If one has a tensor field equation in x-space,

$$Q_a{}^d{}_c(\mathbf{x}) = H_{ab}(\mathbf{x})T^b{}_c(\mathbf{x})\ B^d(\mathbf{x})\ , \qquad (E.9.12a)$$

in which all the objects transform as tensors with respect to the underlying $\mathbf{x}'' = F_M(\mathbf{x})$ (and its linear approximation $M(\mathbf{x})$ as in $d\mathbf{x}'' = M(x)\ d\mathbf{x}$), then the equation is covariant and takes the same form in x"-space,

$$Q''_a{}^d{}_c(\mathbf{x''}) = H''_{ab}(\mathbf{x''})T''^b{}_c(\mathbf{x''})\ B''^d(\mathbf{x''})\ . \qquad \mathbf{x}'' = \mathbf{F_M}(\mathbf{x}) \qquad (E.9.12b)$$

Using our script notation explained in (E.8.5) and (E.8.20), we rewrite the x"-space equation above as the third line below (primed script tensor names). The fourth line is then appropriate to the Moon and Spencer Picture (14.1.1) where the curvilinear coordinates u^i have no primes :

$Q_a{}^d{}_c = H_{ab}T^b{}_c\ B^d$ x-space (general), Picture A or M&S (1.11)
$Q'_a{}^d{}_c = H'_{ab}T'^b{}_c\ B'^d$ x'-space (general), Picture A or B (1.11)
$\mathcal{Q}'_a{}^d{}_c = \mathcal{H}'_{ab}\mathcal{T}'^b{}_c\ \mathcal{B}'^d$ x"-space (orthogonal) , Picture A or B (1.11)
$\mathcal{Q}_a{}^d{}_c = \mathcal{H}_{ab}\mathcal{T}^b{}_c\ \mathcal{B}^d$ x"-space (orthogonal) , Picture M&S (14.1.1) (E.9.13)

Comment: An x-space observer has axes $\mathbf{u_n}$ (Frame S) while an x"-space observer has axes $\hat{e}_n(x)$ (Frame S"), and these two sets of observation axes are related by $\mathbf{u_n} = M(\mathbf{x})\ \hat{e}_n(\mathbf{x})$ where $M(\mathbf{x})$ is a rotation. If the first equation describes something at location $\mathbf{x}$ in the realm of Newtonian mechanics, we expect the equation to have the same form in both Frame S and Frame S" which are related by this local rotation $M(\mathbf{x})$. In other words, rotations are an invariance of Newtonian mechanics, and this means equations are covariant with respect to rotations. The above example, which might apply to fluid dynamics, has this covariance at each point **x** in the fluid, and it may happen that the rotation is a different rotation at different points **x**, but it is always a rotation.

A simple interpretation of (E.9.13) is that the mutually orthogonal unit base vectors $\hat{\mathbf{e}}_n$ form a rotated Cartesian frame of reference, so the x-space equations must have the same form in this rotated frame. This is why the first and last lines of (E.9.13) have the same form.

Example: Polar Coordinates. In polar coordinates now with ordering r,θ = 1,2 one has

$$
\begin{array}{ll}
S^1{}_1 = (\partial x/\partial r) = \cos\theta & x = r\cos\theta \\
S^1{}_2 = (\partial x/\partial\theta) = -r\sin\theta & y = r\sin\theta \\
S^2{}_1 = (\partial y/\partial r) = \sin\theta & \\
S^2{}_2 = (\partial y/\partial\theta) = r\cos\theta &
\end{array}
\qquad (E.9.14)
$$

$$
S^i{}_j = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} \qquad R^i{}_j = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta/r & \cos\theta/r \end{pmatrix} \qquad R = S^{-1}
$$

$$
[g' = RR^T]_{DN} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta/r & \cos\theta/r \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta/r \\ \sin\theta & \cos\theta/r \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1/r^2 \end{pmatrix} \quad \rightarrow \quad g'^{ab} = \begin{pmatrix} h'_r{}^{-2} & 0 \\ 0 & h'_\theta{}^{-2} \end{pmatrix}
$$

so $h_r = 1$ and $h_\theta = r$.

The N and M matrices may be computed as follows: [recall $M = HR$ and $N = SH^{-1}$ from (E.8.14,15)]

$$
M^a{}_b \equiv h'_a R^a{}_b = \begin{pmatrix} 1 & 0 \\ 0 & r \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta/r & \cos\theta/r \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = R_z(-\theta) \qquad (E.9.15)
$$

$$
N^a{}_b \equiv S^a{}_b h'_b{}^{-1} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1/r \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = R_z(\theta) \qquad (E.9.16)
$$

Therefore, the relation between a rank-2 tensor's $\hat{\mathbf{e}}_n$-expanded-form components $[A^{(\hat{e})}]^i{}_j$ and the Cartesian form components $A^i{}_j$ is given (for polar coordinates) by (E.8.22) which says

$$
A'' = A^{(\hat{e})} = \mathcal{A}' = M\,A\,M^{\mathcal{T}} = M\,A\,M^T \qquad (E.9.17)
$$

or

$$
\begin{pmatrix} A_{rr} & A_{r\theta} \\ A_{\theta r} & A_{\theta\theta} \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \qquad \text{// verified in Lai p 316 Problem 5.71}
$$

where recall that up or down index position has no significance in either x-space or x"-space and $M^{\mathcal{T}} = M^T$ as in (7.9.5). That is to say, one has $A_{12} = A^1{}_2$ as well as $A_{r\theta} = A^r{}_\theta$, so one can think of the above as a down-tilt matrix equation.

Applications in Continuum Mechanics

1. In isotropic elastic stress analysis for states of plane stress and plane strain (see Lai p 251), the 3D Cartesian stress tensor T_{ij} has a simple form in which the upper left four components can be represented as derivatives of a potential-like function called an Airy function φ, so that $T_{11} = \partial_2{}^2\varphi$, $T_{22} = \partial_1{}^2\varphi$, and $T_{12} = T_{21} = -\partial_1\partial_2\varphi$. In this case, equation (E.9.17) becomes

$$\begin{pmatrix} T_{rr} & T_{r\theta} \\ T_{\theta r} & T_{\theta\theta} \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \partial_2{}^2\varphi & -\partial_1\partial_2\varphi \\ -\partial_1\partial_2\varphi & \partial_1{}^2\varphi \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \qquad // \; B = MAM^T \qquad \text{(E.9.18)}$$

where

$$\partial_1 = \partial/\partial x_1 = \cos\theta\, \partial_r - (\sin\theta/r)\partial_\theta$$
$$\partial_2 = \partial/\partial x_2 = \sin\theta\, \partial_r + (\cos\theta/r)\partial_\theta \;.$$

Here is a Maple computation of (E.9.18):

```
A :=
matrix(2,2,[D2(D2(phi(r,theta))),-D1(D2(phi(r,theta))),-D1(D2(phi(r,theta))),
D1(D1(phi(r,theta)))]);
```

$$A := \begin{bmatrix} \mathrm{D2}(\mathrm{D2}(\phi(r,\theta))) & -\mathrm{D1}(\mathrm{D2}(\phi(r,\theta))) \\ -\mathrm{D1}(\mathrm{D2}(\phi(r,\theta))) & \mathrm{D1}(\mathrm{D1}(\phi(r,\theta))) \end{bmatrix}$$

```
D1 := f -> cos(theta)*Diff(f,r) - (sin(theta)/r)*Diff(f,theta);
```

$$\mathrm{D1} := f \to \cos(\theta)\left(\frac{\partial}{\partial r} f\right) - \frac{\sin(\theta)\left(\frac{\partial}{\partial \theta} f\right)}{r}$$

```
D2 := f -> sin(theta)*Diff(f,r) + (cos(theta)/r)*Diff(f,theta);
```

$$\mathrm{D2} := f \to \sin(\theta)\left(\frac{\partial}{\partial r} f\right) + \frac{\cos(\theta)\left(\frac{\partial}{\partial \theta} f\right)}{r}$$

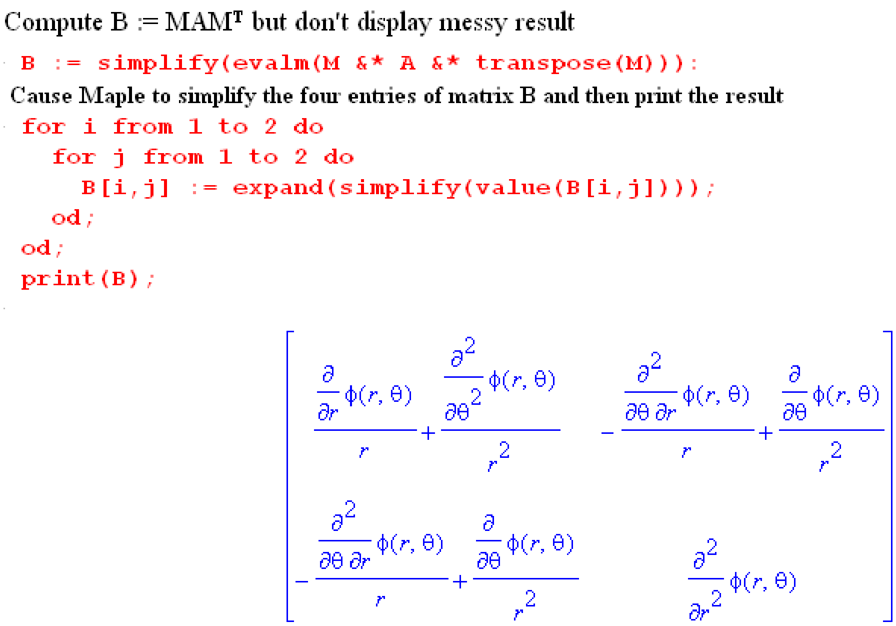
A = A * 1 forces Maple to install the D1 and D2 operators inside matrix A, the result is messy so not showing it

```
A := evalm(A &* &*()):
```

Construct the M matrix

```
M := matrix(2,2,[cos(theta),sin(theta),-sin(theta),cos(theta)]);
```

$$M := \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$$



Compute B := MAM^T but don't display messy result

```
B := simplify(evalm(M &* A &* transpose(M))):
```

Cause Maple to simplify the four entries of matrix B and then print the result

```
for i from 1 to 2 do
   for j from 1 to 2 do
     B[i,j] := expand(simplify(value(B[i,j])));
   od;
od;
print(B);
```

$$\begin{bmatrix} \dfrac{\frac{\partial}{\partial r}\phi(r,\theta)}{r} + \dfrac{\frac{\partial^2}{\partial \theta^2}\phi(r,\theta)}{r^2} & -\dfrac{\frac{\partial^2}{\partial \theta\,\partial r}\phi(r,\theta)}{r} + \dfrac{\frac{\partial}{\partial \theta}\phi(r,\theta)}{r^2} \\ -\dfrac{\frac{\partial^2}{\partial \theta\,\partial r}\phi(r,\theta)}{r} + \dfrac{\frac{\partial}{\partial \theta}\phi(r,\theta)}{r^2} & \dfrac{\partial^2}{\partial r^2}\phi(r,\theta) \end{bmatrix}$$

// verified in Lai p 264 (5.27.3) (E.9.19)

This then is a real-world example of using a rank-2 tensor in curvilinear coordinates expanded on the unit tangent base vectors. The mentioned plane of strain or stress has Cartesian coordinates x_1,x_2 which are converted to polar coordinates r,θ. The third Cartesian coordinate x_3 is more or less ignored.

2. The relation between the stress tensor T_{ij} and the infinitesimal strain tensor E_{ij} for an isotropic material is stated in Cartesian coordinate x-space as (a form of Hooke's Law generalizing F = -kx),

$$T_{ij} = \lambda\, \mathrm{tr}(E) g_{ij} + 2\mu E_{ij} \qquad \text{or the same thing:} \qquad T^{ij} = \lambda\, \mathrm{tr}(E) g^{ij} + 2\mu E^{ij}$$

$$\text{where } \mathrm{tr}(E) = \Sigma_i E_{ii} = E_{ii} \text{ with implied sum on i .} \tag{E.9.20}$$

Here λ and μ are called Lamé's constants and $g_{ij} = g^{ij} = \delta_{i,j}$ (see Lai p 208 (5.3.8)). Some components are (in the case that $E_{33} = 0$),

$$T_{11} = \lambda(E_{11}+ E_{22}) + 2\mu E_{11} \qquad // \ 1,2,3 = x,y,z$$
$$T_{22} = \lambda(E_{11}+ E_{22}) + 2\mu E_{22}$$
$$T_{12} = \ 2\mu E_{12}\ . \tag{E.9.21}$$

With respect to our transformation F_M, (E.9.20) is a "true tensor equation" ($\mathrm{tr}(E) = E^k{}_k = E_{kk}$ is scalar under rotations), so according to Section 7.15 it is "covariant" and in x"-space may be written

$$T''^{ij} = \lambda\, \mathrm{tr}(E'')g''^{ij} + 2\mu E''^{ij}$$

or

$$T''^{ij} == \lambda\, \mathrm{tr}(E'')\delta_{i,j}+ 2\mu E''^{ij} \qquad // \ g''^{ij} = \delta_{i,j} \text{ as in (E.9.3)}$$

or

$$[T^{(\hat{e})}]^{ij} = \lambda\, [T^{(\hat{e})}]^{kk}\ \delta_{i,j} + 2\mu[E^{(\hat{e})}]^{ij}\ .$$

or

$$\mathcal{T}'^{ij} = \lambda\, \mathrm{tr}(\mathcal{E}')\delta_{i,j} + 2\mu\mathcal{E}'^{j} \qquad \mathrm{tr}(\mathcal{E}') = \Sigma^{i}\ \mathcal{E}'_{ii}\ . \qquad // \ \mathcal{E} = \text{script E} \tag{E.9.22}$$

This notion of equation covariance was shown above in example (E.9.13). Recall the notation example from (E.8.6) that, for spherical coordinates 1,2,3 = r,θ,φ,

$$[A^{(\hat{e})}]^{2213} = \mathcal{A}'^{2213} = A^{\theta\theta r\varphi} = A_{\theta\theta r\varphi}\ . \tag*{(E.8.6)}$$

Correspondingly, we would set $\mathcal{T}'^{11} = T_{rr}$ and so on to write the components of (E.9.21) this way in cylindrical coordinates, where 1,2,3 = r,θ,z,

$$\begin{aligned} T_{rr} &= \lambda\, [\ E_{rr}+ E_{\theta\theta}]\ + 2\mu\ E_{rr} \qquad // \ E_{zz} = 0 \\ T_{\theta\theta} &= \lambda\, [\ E_{rr}+ E_{\theta\theta}]) + 2\mu E_{\theta\theta} \\ T_{r\theta} &= 2\mu\ E_{r\theta}\quad . \end{aligned} \tag{E.9.23}$$

This has the same form as (E.9.21) with 1→r and 2→θ. These equations are consistent with Lai p 264 (5.27.7) with $\lambda = -E_Y v/[(v+1)(2v-1)]$ and $2\mu = E_Y/(1+v)$. Statements of (E.9.23) are rare on the web, but here is one instance (Victor Saouma draft Lecture Notes on Continuum Mechanics, 1998),

$$\begin{aligned} T_{rr} &= \lambda e + 2\mu\varepsilon_{rr} \\ T_{\theta\theta} &= \lambda e + 2\mu\varepsilon_{\theta\theta} \\ T_{r\theta} &= 2\mu\varepsilon_{r\theta} \qquad\qquad e = \varepsilon_{rr} + \varepsilon_{\theta\theta}. \end{aligned} \tag{E.9.24}$$

3. By way of contrast, the Cartesian-coordinates equation $E_{ij} = (\partial_i u_j + \partial_j u_i)/2$, which relates strain tensor E_{ij} to the vector displacement **u** of a continuum particle, is not a "true tensor equation", so $E_{r\theta} \neq (\partial_r u_\theta + \partial_\theta u_r)/2$. In fact, this relation is $E = [(\nabla\mathbf{u})^T + (\nabla\mathbf{u})]\ /\ 2$ and $(\nabla\mathbf{u})$ for polar coordinates is computed in (G.6.6) and one ends up with $E_{r\theta} = (\partial_r u_\theta + (1/r)\ \partial_\theta u_r\ - u_\theta/r)\ /\ 2$.

E.10 Tensor expansions in a mixed basis

Recall the expansion (E.2.1) and (E.2.3) written for a general rank-n tensor A,

$$A = \Sigma_{ijk\ldots}\ \alpha^{ijk\ldots}\ (\mathbf{b}_i\otimes\mathbf{b}_j\otimes\mathbf{b}_k\ldots) \qquad\qquad \alpha^{ijk\ldots} = A \bullet (\mathbf{b}^i\otimes\mathbf{b}^j\otimes\mathbf{b}^k\ldots) \tag{E.10.1}$$

where $\alpha^{ijk\ldots}$ are the coefficients of the expansion of A on the direct product basis shown. To make explicit the fact that the coefficients depend on the choice of basis, one might write (one b for each index, number of b's is the rank of the tensor) ,

$$\alpha^{\mathtt{ijk\ldots}} = [\,A^{(\mathtt{b,b,b\ldots})}]^{\mathtt{ijk\ldots}} \quad . \tag{E.10.2}$$

The fact that the indices ijk... are "up" indicates that the b label stands for the $\mathbf{b}_\mathtt{i}$ basis and not $\mathbf{b}^\mathtt{i}$. So here is an example showing the generic expansion of a rank-3 tensor,

$$A = \Sigma_{\mathtt{ijk}}\ [\,A^{(\mathtt{b,b,b})}]^{\mathtt{ijk}}\,(\mathbf{b}_\mathtt{i}\otimes\mathbf{b}_\mathtt{j}\otimes\mathbf{b}_\mathtt{k}) \qquad [\,A^{(\mathtt{b,b,b})}]^{\mathtt{ijk}} = A \bullet (\mathbf{b}^\mathtt{i}\otimes\mathbf{b}^\mathtt{j}\otimes\mathbf{b}^\mathtt{k})\ . \tag{E.10.3}$$

Earlier we used the simpler notation $[A^{(\mathtt{b})}]^{\mathtt{ijk}}$ for the above coefficient, but now we want to show all the basis elements because now we want to consider a "mixed basis expansion" such as

$$A = \Sigma_{\mathtt{ijk}}\ [\,A^{(\mathtt{b,e,u})}]^{\mathtt{ijk}}\,(\mathbf{b}_\mathtt{i}\otimes\mathbf{e}_\mathtt{j}\otimes\mathbf{u}_\mathtt{k}) \qquad [\,A^{(\mathtt{b,e,u})}]^{\mathtt{ijk}} = A \bullet (\mathbf{b}^\mathtt{i}\otimes\mathbf{e}^\mathtt{j}\otimes\mathbf{u}^\mathtt{k})\ . \tag{E.10.4}$$

This is a completely viable expansion since the b, e and u basis vectors are each a complete set within their part of the direct-product space. To verify the validity of this expansion, consider :

$$[\,A^{(\mathtt{b,e,u})}]^{\mathtt{ijk}} = \{A\} \bullet (\mathbf{b}^\mathtt{i}\otimes\mathbf{e}^\mathtt{j}\otimes\mathbf{u}^\mathtt{k})$$

$$= \{\,\Sigma_{\mathtt{i'j'k'}}\ [\,A^{(\mathtt{b,e,u})}]^{\mathtt{i'j'k'}}\,(\mathbf{b}_{\mathtt{i'}}\otimes\mathbf{e}_{\mathtt{j'}}\otimes\mathbf{u}_{\mathtt{k'}})\}\bullet (\mathbf{b}^\mathtt{i}\otimes\mathbf{e}^\mathtt{j}\otimes\mathbf{u}^\mathtt{k})$$

$$= \{\,\Sigma_{\mathtt{i'j'k'}}\ [\,A^{(\mathtt{b,e,u})}]^{\mathtt{i'j'k'}}\,(\mathbf{b}^{\mathtt{i'}}\bullet\mathbf{b}_\mathtt{i})\ (\mathbf{e}^{\mathtt{j'}}\bullet\mathbf{e}_\mathtt{j})\,(\mathbf{u}^{\mathtt{k'}}\bullet\mathbf{u}_\mathtt{k})$$

$$= \{\,\Sigma_{\mathtt{i'j'k'}}\ [\,A^{(\mathtt{b,e,u})}]^{\mathtt{i'j'k'}}\ \delta^{\mathtt{i'}}{}_{\mathtt{i}}\delta^{\mathtt{j'}}{}_{\mathtt{j}}\delta^{\mathtt{k'}}{}_{\mathtt{k}}$$

$$= [\,A^{(\mathtt{b,e,u})}]^{\mathtt{ijk}}\,. \tag{E.10.5}$$

In the case of a rank-2 tensor, one has the option of using the other notations discussed above,

$$A = \Sigma_{\mathtt{ij}}\ [\,A^{(\mathtt{b,u})}]^{\mathtt{ij}}\,(\mathbf{b}_\mathtt{i}\otimes\mathbf{u}_\mathtt{j}) = \Sigma_{\mathtt{ij}}\ [\,A^{(\mathtt{b,u})}]^{\mathtt{ij}}\,(\mathbf{b}_\mathtt{i}\mathbf{u}_\mathtt{j})$$

direct product dyadic

$$= \Sigma_{\mathtt{ij}}\ [\,A^{(\mathtt{b,u})}]^{\mathtt{ij}}\,(\mathbf{b}_\mathtt{i}\mathbf{u}_\mathtt{j}{}^{\mathtt{T}}) = \Sigma_{\mathtt{ij}}\ [\,A^{(\mathtt{b,u})}]^{\mathtt{ij}}\ |b_\mathtt{i}\rangle\langle u_\mathtt{j}| \tag{E.10.6}$$

matrix bra-ket

where

$$[\,A^{(\mathtt{b,u})}]^{\mathtt{ij}} = A \bullet (\mathbf{b}^\mathtt{i}\otimes\mathbf{u}^\mathtt{j}) = A^{\mathtt{ab}}\ (\mathbf{b}^\mathtt{i})_\mathtt{a}\ (\mathbf{u}^\mathtt{j})_\mathtt{b} = (\mathbf{b}^\mathtt{i})^{\mathtt{T}}\,A\,(\mathbf{u}^\mathtt{j}) = \langle\, b^\mathtt{i}|\,A\,|\,u^\mathtt{j}\,\rangle\,. \tag{E.10.7}$$

One can of course use unit versions of the $\mathbf{e}_\mathbf{n}$ basis vectors, $\hat{\mathbf{e}}_\mathbf{n}$, and then one might write for example

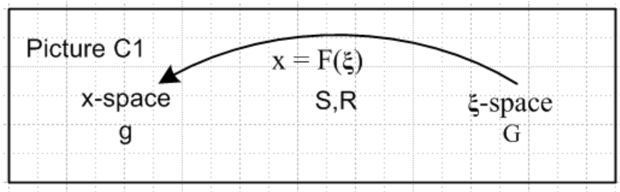
$$A = \Sigma_{\mathtt{ijk}}\ [\,A^{(\hat{\mathtt{e}},\hat{\mathtt{e}},\mathtt{u})}]^{\mathtt{ijk}}\,(\hat{\mathbf{e}}_\mathtt{i}\otimes\hat{\mathbf{e}}_\mathtt{j}\otimes\mathbf{u}_\mathtt{k}) \tag{E.10.8}$$

where the hats are replicated into the superscript tensor label.

Appendix F: The Affine Connection $\Gamma^{c}{}_{ab}$ and Covariant Derivatives

F.1 Definition and Interpretation of Γ : $\Gamma^{c}{}_{ab} = \mathbf{q}^{c} \bullet (\partial_{a}\mathbf{q}_{b}) \;\;=\;\; R^{c}{}_{i}(\partial_{a}R_{b}{}^{i})$

In this Section we shall use a modified Picture C in which the quasi-Cartesian space on the right is called ξ-space instead of $x^{(0)}$-space as in Picture C. The notation ξ^{i} for the coordinates of ξ-space seems traditional in general relativity work where the Γ object appears frequently.



(F.1.1)

Recall from the discussion near (1.10) that the metric tensor G is a diagonal matrix whose elements are independently +1 or -1. Since the metric tensor transforms as a rank-2 tensor, we know from the last line of (7.5.8) (adapted from Picture A to Picture C1) that $g_{ab} = R_{a}{}^{i}R_{b}{}^{j}G_{ij}$. Since G is diagonal, we write this as

$$g_{ab} = R_{a}{}^{i}R_{b}{}^{i}G_{ii} \qquad \text{and} \qquad g^{ab} = R^{a}{}_{i}R^{b}{}_{i}G^{ii} \tag{F.1.2}$$

with a single implied sum on i. If G = 1, then the x^{i} coordinates of x-space are "the curvilinear coordinates" and the ξ^{i} are "the Cartesian coordinates".

In Picture C1 the tangent base vectors exist in ξ-space and we will call them $\mathbf{q}_{n}$. From (7.18.1) (adapted from Picture A to Picture C1) we know both that $(\mathbf{q}_{n})^{i} = R_{n}{}^{i}$ and that the dot product $\mathbf{q}_{n} \bullet \mathbf{q}_{m} = g_{nm}$ where g is the metric tensor in x-space. In general $\mathbf{q}_{n} = \mathbf{q}_{n}(\xi)$. However, since $\mathbf{x} = \mathbf{F}(\xi)$, we are free *instead* to regard $\mathbf{q}_{n} = \mathbf{q}_{n}(\mathbf{x})$, and that is what we shall do for Picture C1.

If one moves a small amount d**x** in x-space, $\mathbf{q}_{n}(\mathbf{x})$ (a vector in ξ-space) will change by some small amount. We are used to this idea in the Picture A context of Fig (3.4.3) where the x-space tangent base vectors $\mathbf{e}_{\theta} = r\,\hat{\boldsymbol{\theta}}$ and $\mathbf{e}_{r} = \hat{\mathbf{r}}$ both change if one takes θ→θ+dθ and r→r + dr in x'-space.

So, for a small change dx^{j} in x-space, the change in $\mathbf{q}_{n}(\mathbf{x})$ is given by,

$$d(\mathbf{q}_{n})^{i} = \partial_{j}(\mathbf{q}_{n})^{i}\, dx^{j} \qquad\qquad // \;\partial_{j} \equiv \partial/\partial x^{j}$$

or

$$(d\mathbf{q}_{n})^{i} = (\partial_{j}\mathbf{q}_{n})^{i}\, dx^{j}\,. \qquad\qquad // \;\partial_{\mathbf{x}}(f^{i}) = \frac{\partial(f^{i})}{dx} = \frac{\partial f^{i}}{dx} = (\partial_{\mathbf{x}}f^{i}) \tag{F.1.3}$$

Since $\mathbf{q}_{n}$ and $(\partial_{j}\mathbf{q}_{n})$ are both vectors in ξ-space, and since the $\mathbf{q}_{k}$ are known to form a complete basis in ξ-space, it must be possible to expand $(\partial_{j}\mathbf{q}_{n})$ on the $\mathbf{q}_{k}$ with some appropriate coefficients, call them $\Gamma^{k}{}_{jn}$:

$$(\partial_j \mathbf{q_n}) = \Gamma^k{}_{jn}\, \mathbf{q_k} \qquad \Rightarrow \qquad (d\mathbf{q_n})^i = (\partial_j \mathbf{q_n})dx^a = \Gamma^k{}_{jn}\, (\mathbf{q_k})^i\, dx^a \ . \tag{F.1.4}$$

The coefficients $\Gamma^k{}_{jn}$ are known as the **affine connection**. They measure how the tangent basis vectors change in ξ-space as a function of **x** in x-space. One regards $\Gamma^k{}_{jn}(\mathbf{x})$ as an object associated with x-space, though it is not a tensor object as (F.6.3) below shows.

Dotting the left equation into $\mathbf{q}^m$, using $\mathbf{q}^m \bullet \mathbf{q_k} = \delta^m{}_k$ as in (7.18.1), and then doing m→ k gives

$$\Gamma^k{}_{jn} = \mathbf{q}^k \bullet (\partial_j \mathbf{q_n}) \ = \ (\mathbf{q}^k)_i(\partial_j \mathbf{q_n})^i \ = \ R^k{}_i(\partial_j R_n{}^i)\ , \tag{F.1.5}$$

where from (7.18.1) and (7.5.16) (adjusted from Picture A to Picture C1) we have used

$$(\mathbf{q}^k)_i = R^k{}_i = \frac{\partial x^k}{\partial \xi^i} \ \text{ and } \ (\mathbf{q_n})^i = R_n{}^i = \frac{\partial \xi^i}{\partial x^n}\ . \tag{F.1.6}$$

Inserting the partial derivatives from (F.1.6) into (F.1.5) gives

$$\Gamma^k{}_{jn} = R^k{}_i(\partial_j R_n{}^i) \ = \frac{\partial x^k}{\partial \xi^i}\,(\,\partial_j \frac{\partial \xi^i}{\partial x^n}\,) \ = \frac{\partial x^k}{\partial \xi^i}\,\frac{\partial^2 \xi^i}{\partial x^j \partial x^n} \tag{F.1.7}$$

Notice that

$$(\partial_j R_n{}^i) \ = (\partial_n R_j{}^i) \qquad // \text{ since } \frac{\partial^2 \xi^i}{\partial x^j \partial x^n} \ = \frac{\partial^2 \xi^i}{\partial x^n \partial x^j} \tag{F.1.8}$$

Form (F.1.7) shows that:

Fact: $\Gamma^k{}_{jn}$ is symmetric on the lower two indices, so $\Gamma^k{}_{jn} = \Gamma^k{}_{nj}$ (F.1.9)

If we take k→λ, i→α, j→μ and n→ν, then (F.1.9) becomes

or

$$\Gamma^\lambda{}_{\mu\nu} = \frac{\partial x^\lambda}{\partial \xi^\alpha}\,\frac{\partial^2 \xi^\alpha}{\partial x^\mu \partial x^\nu}$$

and this equation appears as Weinberg p 100 (4.5.1), with traditional Greek indices used for relativity.

We now restate some of the results above with more commonly used indices :

$$(\partial_a \mathbf{q_n}) = \Gamma^k{}_{an}\, \mathbf{q_k} \quad \rightarrow \quad (\partial_a \mathbf{q_b}) = \Gamma^c{}_{ab}\, \mathbf{q_c}$$
$$\Gamma^k{}_{jn} = \mathbf{q}^k \bullet (\partial_j \mathbf{q_n}) \quad \rightarrow \quad \Gamma^c{}_{ab} = \mathbf{q}^c \bullet (\partial_a \mathbf{q_b})$$

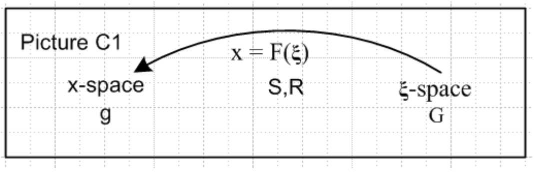
so that

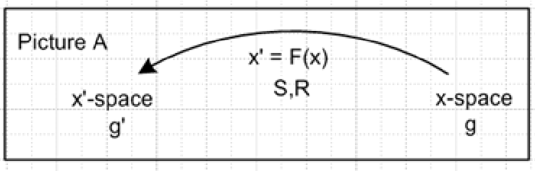
$$(\partial_a \mathbf{q_b}) = \Gamma^c{}_{ab}\, \mathbf{q_c} \qquad \textit{(F.1.4)}$$
$$\Gamma^c{}_{ab} = \mathbf{q^c} \bullet (\partial_a \mathbf{q_b}) \qquad \textit{(F.1.5)}$$
$$\Gamma^c{}_{ab} = \Gamma^c{}_{ba} \qquad \textit{(F.1.9)}$$

$$\Gamma^c{}_{ab} = \quad R^c{}_i(\partial_a R_b{}^i) \qquad \textit{(F.1.5)}$$
$$\Gamma^c{}_{ab} = -\, R_b{}^i\ (\partial_a R^c{}_i) \qquad \text{(F.1.10)}$$

The last line will be derived in the next Section.

One can convert the Γ-related equations from Picture C1 to Picture A,





For example, since the tangent base vectors in x-space of Picture A are called $\mathbf{e_n}$,

$$(\partial_a \mathbf{q_b}) = \Gamma^c{}_{ab}\, \mathbf{q_c} \qquad \rightarrow \qquad (\partial'_a \mathbf{e_b}) = \Gamma'^c{}_{ab}\, \mathbf{e_c}$$

$$\Gamma^c{}_{ab} = \mathbf{q^c} \bullet (\partial_a \mathbf{q_b}) \qquad \rightarrow \qquad \Gamma'^c{}_{ab} = \mathbf{e^c} \bullet (\partial'_a \mathbf{e_b}) \qquad \text{(F.1.11)}$$

where $\Gamma' = \Gamma'(\mathbf{x'})$ and $\mathbf{e_n} = \mathbf{e_n}(\mathbf{x'})$.

F.2 Identities of the form $(\partial_a R^d{}_n) = -\, R^e{}_n\, R^d{}_m\ (\partial_a R_e{}^m)$

The identities (to be proven below) are :

$$R^d{}_m\ (\partial_a R_e{}^m) = -\ R_e{}^m\ (\partial_a R^d{}_m) \qquad 1$$

$$(\partial_a R^d{}_n) = -\, R^e{}_n\, R^d{}_m\ (\partial_a R_e{}^m) \qquad 2 \qquad \partial_a \equiv \partial/\partial x^a$$

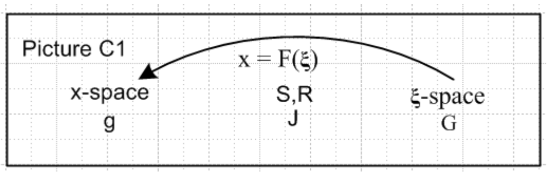
$$(\partial_a R_d{}^n) = -\, R_e{}^n\, R_d{}^m\ (\partial_a R^e{}_m) \qquad 3\ . \qquad \text{(F.2.1)}$$

Corollary: The first identity above allows the alternate form for the affine connection in terms of R, as quoted above in (F.1.10). Here we take d→c, m→i and e→b in identity #1,

$$R^d{}_m\ (\partial_a R_e{}^m) = -\ R_e{}^m\ (\partial_a R^d{}_m) \qquad \text{// identity 1}$$
$$R^c{}_i\ (\partial_a R_b{}^i) = -\ R_b{}^i\ (\partial_a R^c{}_i) \qquad \text{// same with new indices}$$

which then verifies (F.1.10).

Our context is:



. (F.2.2)

Proofs of identities 1,2,3: These identities are a simple consequence of the fact that RS = 1 which in standard notation is written $\delta^c{}_b = R^c{}_\alpha R_b{}^\alpha$ (orthogonality rule #3 in (7.6.4)). So,

$$0 = \partial_a(\delta^d{}_e) = \partial_a(R^d{}_m R_e{}^m) = R^d{}_m\, (\partial_a R_e{}^m) + R_e{}^m\, (\partial_a R^d{}_m)\ . \qquad \text{QED 1} \qquad (F.2.3)$$

Apply $\Sigma_e R^e{}_n$ to both ends of (F.2.3) to get

$$\begin{aligned} 0 &= R^e{}_n\, R^d{}_m\, (\partial_a R_e{}^m) + (R^e{}_n\, R_e{}^m)\, (\partial_a R^d{}_m) = R^e{}_n\, R^d{}_m\, (\partial_a R_e{}^m) + \delta_n{}^m\, (\partial_a R^d{}_m) \\ &= R^e{}_n\, R^d{}_m\, (\partial_a R_e{}^m) + (\partial_a R^d{}_n) \end{aligned}$$

$$\Rightarrow\ (\partial_a R^d{}_n) = -\, R^e{}_n\, R^d{}_m\, (\partial_a R_e{}^m)\ . \qquad \text{QED 2} \qquad (F.2.4)$$

Alternatively, apply $\Sigma_d R_d{}^n$ to both ends of (F.2.3) to get

$$\begin{aligned} 0 &= (R_d{}^n\, R^d{}_m)\, (\partial_a R_e{}^m) + R_d{}^n\, R_e{}^m\, (\partial_a R^d{}_m) = \delta^n{}_m\, (\partial_a R_e{}^m) + R_d{}^n\, R_e{}^m\, (\partial_a R^d{}_m) \\ &= (\partial_a R_e{}^n) + R_d{}^n\, R_e{}^m\, (\partial_a R^d{}_m) \end{aligned}$$

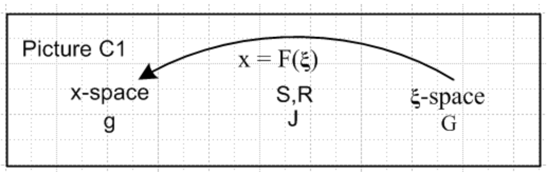
$$\Rightarrow\ (\partial_a R_e{}^n) = -\, R_d{}^n\, R_e{}^m\, (\partial_a R^d{}_m) \qquad \text{now swap d and e:}$$
$$\Rightarrow\ (\partial_a R_d{}^n) = -\, R_e{}^n\, R_d{}^m\, (\partial_a R^e{}_m)\ . \qquad \text{QED 3} \qquad (F.2.5)$$

F.3 Identities of the form $(\partial_c g^{ab}) = -\,[g^{an}\, \Gamma^b{}_{cn} + g^{bn}\, \Gamma^a{}_{cn}]$

The derivatives of the metric tensor are given by ($\partial_c = \partial/\partial x^c$)

$$(\partial_c g^{ab}) = -\,[g^{an}\, \Gamma^b{}_{cn} + g^{bn}\, \Gamma^a{}_{cn}] \qquad 1 \qquad (F.3.1)$$

$$(\partial_c g_{ab}) = +\,[g_{an}\, \Gamma^n{}_{cb} + g_{bn}\, \Gamma^n{}_{ca}] \qquad 2 \qquad (F.3.2)$$



(F.3.3)

Proof of 1: $(\partial_c g^{ab}) = -[g^{an}\,\Gamma^b{}_{cn} + g^{bn}\,\Gamma^a{}_{cn}]$ *(F.3.1)*

Using (F.1.2) that $g^{ab} = R^a{}_i R^b{}_i G^{ii}$, the LHS of (F.3.1) becomes,

$$\text{LHS} = (\partial_c g^{ab}) = \partial_c(R^a{}_i R^b{}_i)G^{ii} = R^a{}_i(\partial_c R^b{}_i)G^{ii} + R^b{}_i(\partial_c R^a{}_i)G^{ii} \ . \tag{F.3.4}$$

For the RHS, the Γ objects can be replaced by their alternate definitions from (F.1.10),

$\Gamma^c{}_{ab} = -R_b{}^i\ (\partial_a R^c{}_i)$ // from (F.1.10)
$\Gamma^b{}_{cn} = -R_n{}^i\ (\partial_c R^b{}_i)$ // b→n then c→b then a→c
$\Gamma^a{}_{cn} = -R_n{}^i\ (\partial_c R^a{}_i)$. // b→a (F.3.5)

The RHS of the claimed identity (F.3.1) may then be written

$$\text{RHS} = -g^{an}\,\Gamma^b{}_{cn} - g^{bn}\,\Gamma^a{}_{cn}$$

$$= \{R^a{}_k R^n{}_k G^{kk}\}\{R_n{}^i\ (\partial_c R^b{}_i)\} + \{R^b{}_k R^n{}_k G^{kk}\}\{R_n{}^i(\partial_c R^a{}_i)\}$$

$$= R^a{}_k(R^n{}_k\, R_n{}^i)\ (\partial_c R^b{}_i)\ G^{kk} + R^b{}_k(R^n{}_k\, R_n{}^i)\ (\partial_c R^a{}_i)\ G^{kk}$$

$$= R^a{}_k\delta_k{}^i\ (\partial_c R^b{}_i)\ G^{kk} + R^b{}_k\delta_k{}^i\ (\partial_c R^a{}_i)\ G^{kk} \qquad \text{// orthog rule \#2 of (7.6.4)}$$

$$= R^a{}_i\ (\partial_c R^b{}_i)\ G^{ii} + R^b{}_i\ (\partial_c R^a{}_i)\ G^{ii}$$

$$= (\partial_c g^{ab}) \qquad \text{// using (F.3.4)} \qquad \text{QED 1} \tag{F.3.6}$$

Proof of 2: $(\partial_c g_{ab}) = +[g_{an}\,\Gamma^n{}_{cb} + g_{bn}\,\Gamma^n{}_{ca}]$ *(F.3.2)*

Using (F.1.2) that $g_{ab} = R_a{}^i R_b{}^i G_{ii}$, the LHS of (F.3.2) becomes,

$$\text{LHS} = (\partial_c g_{ab}) = \partial_c(R_a{}^i R_b{}^i)G_{ii} = R_a{}^i\ (\partial_c R_b{}^i)G_{ii} + R_b{}^i\ (\partial_c R_a{}^i)G_{ii} \ . \tag{F.3.7}$$

For the RHS, the Γ objects can be replaced by their primary definitions in (F.1.10),

$\Gamma^c{}_{ab} = R^c{}_i(\partial_a R_b{}^i)$
$\Gamma^n{}_{cb} = R^n{}_i(\partial_c R_b{}^i)$ // c→n then a→c then i→k
$\Gamma^n{}_{ca} = R^n{}_i(\partial_c R_a{}^i)$. // b→a (F.3.8)

The RHS of the claimed identity (F.3.2) may then be written

$$\text{RHS} = g_{an}\,\Gamma^n{}_{cb} + g_{bn}\,\Gamma^n{}_{ca}$$

$$= \{R_a{}^k R_n{}^k G_{kk}\}\{R^n{}_i(\partial_c R_b{}^i)\} + \{R_b{}^k R_n{}^k G_{kk}\}\{R^n{}_i(\partial_c R_a{}^i)\}$$

$= R_a{}^k\ (R_n{}^k R^n{}_i)(\partial_c R_b{}^i)G_{kk}\ +\ R_b{}^k(R_n{}^k R^n{}_i)(\partial_c R_a{}^i)G_{kk}$

$= R_a{}^k\ \delta^k{}_i(\partial_c R_b{}^i)G_{kk}\ +\ R_b{}^k\partial^k{}_i(\partial_c R_a{}^i)G_{kk}$ // orthog rule #1 of (7.6.4)

$= R_a{}^i\ (\partial_c R_b{}^i)G_{ii}\ +\ R_b{}^i\ (\partial_c R_a{}^i)G_{ii}$

$= (\partial_c g_{ab})$ // using (F.3.7) QED 2 (F.3.9)

F.4 Identity: $\Gamma^d{}_{ab} = (1/2)\ g^{dc}\ [\ \partial_a g_{bc} + \partial_b g_{ca} - \partial_c g_{ab}]$

The **identity** states that $\Gamma^d{}_{ab}$ may be expressed entirely in terms of the metric tensor,

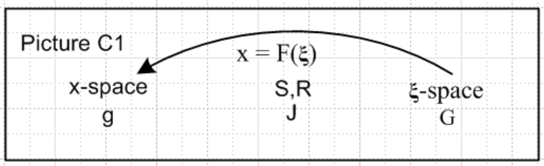
$$\Gamma^d{}_{ab} = (1/2)\ g^{dc}\ [\ \partial_a g_{bc} + \partial_b g_{ca} - \partial_c g_{ab}]\ . \tag{F.4.1}$$

Recall in our definition above, $\Gamma^d{}_{ab} = R^d{}_k(\partial_a R_b{}^k)$, that Γ was given in terms of R matrices.

The following **corollary** concerns summing of the upper Γ index with a lower one,

$$\Gamma^a{}_{an} = (1/2)\ g^{ad}\ \partial_n g_{ad} = (1/2)(1/g)\partial_n g = (1/\sqrt{|g|}\ )\ \partial_n(\sqrt{|g|}\ )\ . \tag{F.4.2}$$

We shall prove the identity (F.4.1) first, and the corollary (F.4.2) after that.



Proof of Identity: $\Gamma^d{}_{ab} = (1/2)\ g^{dc}\ [\ \partial_a g_{bc} + \partial_b g_{ca} - \partial_c g_{ab}]$ (F.4.3)

We first rewrite (F.1.2) with new index names,

$g_{ab} = R_a{}^e R_b{}^e\ G_{ee}$ // sum on e
$g^{dc} = R^d{}_i R^c{}_i\ G^{ii}$. // sum on i (F.4.4)

The first line below is computed from the first line in the pair above, then the next two lines below are obtained by doing forward cyclic index permutations of the first line :

$\partial_c g_{ab} = [R_b{}^e\ (\partial_c R_a{}^e) + R_a{}^e(\partial_c R_b{}^e)\ ]G_{ee}$
$\partial_a g_{bc} = [R_c{}^e\ (\partial_a R_b{}^e) + R_b{}^e(\partial_a R_c{}^e)\ ]G_{ee}$
$\partial_b g_{ca} = [R_a{}^e\ (\partial_b R_c{}^e) + R_c{}^e(\partial_b R_a{}^e)\ ]G_{ee}$. (F.4.5)

The last four lines can be appropriately inserted into the RHS of (F.4.3) to obtain

$$(\mathrm{RHS})^{\mathbf{d}}{}_{\mathbf{ab}} = (1/2)\, g^{\mathbf{dc}} \left[\, \partial_{\mathbf{a}} g_{\mathbf{bc}} + \partial_{\mathbf{b}} g_{\mathbf{ca}} - \partial_{\mathbf{c}} g_{\mathbf{ab}} \right]$$

$$= (1/2)\, (R^{\mathbf{d}}{}_{\mathbf{i}} R^{\mathbf{c}}{}_{\mathbf{i}}\, G^{\mathbf{ii}})\, G_{\mathbf{ee}} \, *$$

$$[\underset{1}{R_{\mathbf{c}}{}^{\mathbf{e}} (\partial_{\mathbf{a}} R_{\mathbf{b}}{}^{\mathbf{e}})} + \underset{2}{R_{\mathbf{b}}{}^{\mathbf{e}}(\partial_{\mathbf{a}} R_{\mathbf{c}}{}^{\mathbf{e}})} + \underset{3}{R_{\mathbf{a}}{}^{\mathbf{e}} (\partial_{\mathbf{b}} R_{\mathbf{c}}{}^{\mathbf{e}})} + \underset{4}{R_{\mathbf{c}}{}^{\mathbf{e}}(\partial_{\mathbf{b}} R_{\mathbf{a}}{}^{\mathbf{e}})} - \underset{5}{R_{\mathbf{b}}{}^{\mathbf{e}} (\partial_{\mathbf{c}} R_{\mathbf{a}}{}^{\mathbf{e}})} - \underset{6}{R_{\mathbf{a}}{}^{\mathbf{e}}(\partial_{\mathbf{c}} R_{\mathbf{b}}{}^{\mathbf{e}})}\,] \;. \qquad (\mathrm{F.4.6})$$

Due to the symmetry $(\partial_{\mathbf{i}} R_{\mathbf{j}}{}^{\mathbf{e}}) = (\partial_{\mathbf{j}} R_{\mathbf{i}}{}^{\mathbf{e}})$ noted in (F.1.8), terms 2 and 5 cancel as do terms 3 and 6, while terms 1 and 4 are equal. Therefore,

$$(\mathrm{RHS})^{\mathbf{d}}{}_{\mathbf{ab}} = (1/2)\, R^{\mathbf{d}}{}_{\mathbf{i}} R^{\mathbf{c}}{}_{\mathbf{i}}\, G^{\mathbf{ii}}\, G_{\mathbf{ee}} * 2\, R_{\mathbf{c}}{}^{\mathbf{e}} (\partial_{\mathbf{a}} R_{\mathbf{b}}{}^{\mathbf{e}})$$

$$= R^{\mathbf{d}}{}_{\mathbf{i}}\, (R_{\mathbf{c}}{}^{\mathbf{e}}\, R^{\mathbf{c}}{}_{\mathbf{i}})\, G^{\mathbf{ii}}\, G_{\mathbf{ee}}\, (\partial_{\mathbf{a}} R_{\mathbf{b}}{}^{\mathbf{e}})$$

$$= R^{\mathbf{d}}{}_{\mathbf{i}}\, \delta^{\mathbf{e}}{}_{\mathbf{i}}\, G^{\mathbf{ii}}\, G_{\mathbf{ee}}\, (\partial_{\mathbf{a}} R_{\mathbf{b}}{}^{\mathbf{e}}) \qquad \text{// orthog rule \#1 of (7.6.4)}$$

$$= R^{\mathbf{d}}{}_{\mathbf{e}}\, G^{\mathbf{ee}}\, G_{\mathbf{ee}}\, (\partial_{\mathbf{a}} R_{\mathbf{b}}{}^{\mathbf{e}})$$

$$= R^{\mathbf{d}}{}_{\mathbf{e}} (\partial_{\mathbf{a}} R_{\mathbf{b}}{}^{\mathbf{e}}) \qquad \text{// } G^{\mathbf{ee}} = G_{\mathbf{ee}} = \pm 1$$

$$= \Gamma^{\mathbf{d}}{}_{\mathbf{ab}} \qquad \text{// from (F.1.10)}$$

$$= \text{LHS of (F.4.3)} \qquad \text{QED} \qquad (\mathrm{F.4.7})$$

Proof of Corollary: The corollary is this ($g \equiv \det(g_{\mathbf{ab}})$)

$$\underset{\text{1st}}{\Gamma^{\mathbf{a}}{}_{\mathbf{an}}} = \underset{\text{2nd}}{(1/2)\, g^{\mathbf{ad}}(\, \partial_{\mathbf{a}} g_{\mathbf{nd}} + \partial_{\mathbf{n}} g_{\mathbf{ad}} - \partial_{\mathbf{d}} g_{\mathbf{an}}\,)} = \underset{\text{3rd}}{(1/2)\, g^{\mathbf{ad}}\, \partial_{\mathbf{n}} g_{\mathbf{ad}}} = \underset{\text{4th}}{(1/2)(1/g)\partial_{\mathbf{n}} g} = \underset{\text{5th}}{(1/\sqrt{|g|}\,)\, \partial_{\mathbf{n}}(\sqrt{|g|}\,)} \;. \qquad (\mathrm{F.4.8})$$

The **2nd** expression is of course just (F.4.1) with $c = b = n$ and implied sum on n.

The first and third terms of the 2nd expression cancel due to symmetry of g:

$$\text{first term} = g^{\mathbf{ad}}\, \partial_{\mathbf{a}} g_{\mathbf{nd}} = \underset{a \leftrightarrow d}{g^{\mathbf{da}}\, \partial_{\mathbf{d}} g_{\mathbf{an}}} = \underset{g^{\mathbf{da}} = g^{\mathbf{ad}}}{g^{\mathbf{ad}}\, \partial_{\mathbf{d}} g_{\mathbf{an}}} = -\text{ third term} \qquad (\mathrm{F.4.9})$$

so we then have just

$$\Gamma^{\mathbf{a}}{}_{\mathbf{an}} = (1/2)\, g^{\mathbf{ad}}(\partial_{\mathbf{n}} g_{\mathbf{ad}})$$

which is the **3rd** expression shown in (F.4.8).

To get the **4th** expression, we must show that $g^{\mathbf{ab}}(\partial_{\mathbf{n}} g_{\mathbf{ab}}) = (1/g)\partial_{\mathbf{n}} g$. Here are the steps to prove this fact, where $g \equiv \det(g_{\mathbf{ab}})$:

(1) $g^{ab} = (g^{-1})_{ab} = \mathrm{cof}(g_{ab})^T/\det(g_{ab}) = \mathrm{cof}(g_{ab})/g \quad \Rightarrow \quad \mathrm{cof}(g_{ab}) = g\, g^{ab}$. (F.4.10)

(2) Apply g_{ba} and sum on b to get: $g_{ba}\mathrm{cof}(g_{ab}) = g\, g_{ba}g^{ab} = g\, \delta_a{}^a = g$. (F.4.11)

(3) $g = \det(g_{ab}) = g_{ab}\mathrm{cof}(g_{ab}) \quad \Rightarrow \quad \partial g/\partial g_{ab} = \mathrm{cof}(g_{ab}) = g\, g^{ab}$ * (F.4.12)

(4) $\partial_n g = \partial g/\partial x^n = (\partial g/\partial g_{ab})(\partial g_{ab}/\partial x^n) = g\, g^{ab}\ (\partial_n g_{ab}) \Rightarrow (1/g)\, \partial_n g = g^{ab}(\partial_n g_{ab})$ (F.4.13)

* In Step 3, notice that $\partial\, \mathrm{cof}(g_{ab})/\partial g_{ab} = 0$ because $\mathrm{cof}(g_{ab})$ is not a function of variable g_{ab}. The element g_{ab} is one of those "crossed out" in obtaining the cofactor of g_{ab}. Despite this fact, $\mathrm{cof}(g_{ab})$ is still an object having indices a,b. In contrast, $g = \det(g_{ab}) = \det(g_{**})$ has no indices!

It remains only to obtain the **5th** expression in (F.4.8), and that is easy to show:

$$g^{-1/2}\, \partial_n(g^{1/2}) = g^{-1/2}\, (1/2)\, g^{-1/2}\, \partial_n(g) = (1/2)\, (1/g)\, (\partial_n g) . \qquad g > 0$$

$$[-g]^{-1/2}\partial_n([-g]^{1/2}) = [-g]^{-1/2}\, (1/2)[-g]^{-1/2}\partial_n([-g]) = (1/2)[-g]^{-1}\, \partial_n([-g]) \qquad g < 0$$

which we summarize as

$$|g|^{-1/2}\, \partial_n(|g|^{1/2}) = (1/2)\, (1/g)\, (\partial_n g) \qquad \text{(F.4.14)}$$

We close this Section with the following observations:

Theorem: If x-space of Picture C1 of (F.1.1) has a metric tensor $g_{ij}(\mathbf{x})$ whose elements are constants independent of **x**, then $\Gamma^c{}_{ab} \equiv 0$. (F.4.15)

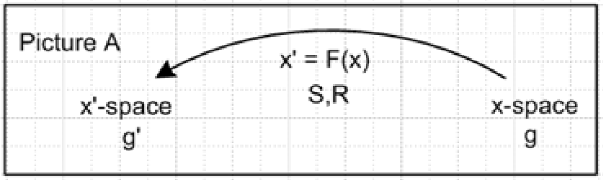
Corollary: If the x-space in Figure C1 of (F.1.1) is Cartesian (g =1) or quasi-Cartesian (g = G), then $\Gamma^c{}_{ab} \equiv 0$. (F.4.16)

Proof: The identity (F.4.1) says $\Gamma^d{}_{ab} = (1/2)\, g^{dc}\, [\, \partial_a g_{bc} + \partial_b g_{ca} - \partial_c g_{ab}]$,
so if g_{ij} are constants, one must have $\Gamma^d{}_{ab} = 0$.

We shall often quote this Corollary with the phrase "Γ = 0 for a Cartesian space".

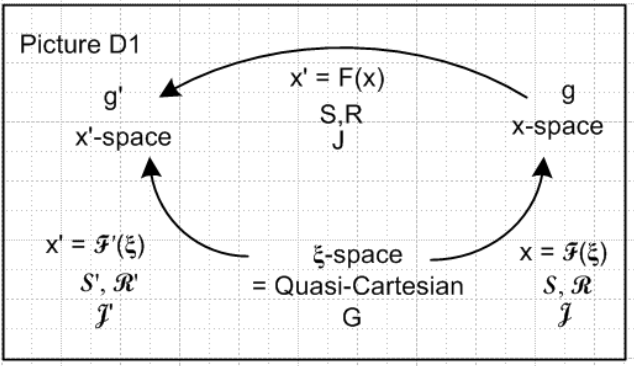
F.5 Picture D1 Context

Our main context of interest is Picture A,



. (F.5.1)

For the proof to be presented in Section F.6, it is useful to think of Picture A as the top part of this Picture D1,



(F.5.2)

The relationship between the R's and S's are these

$$R = \mathcal{R}' \mathcal{R}^{-1} = \mathcal{R}' \mathcal{S} \qquad \Rightarrow \qquad \mathcal{R}' = R\mathcal{R}$$
$$S = \mathcal{R} \mathcal{R}'^{-1} = \mathcal{R} \mathcal{S}' \ . \qquad (F.5.3)$$

The tangent and reciprocal base vectors in ξ-space associated with transformations $\mathcal{F}$ and $\mathcal{F}'$ are these, based on (F.1.6),

$$(\boldsymbol{q}_n)^i = \mathcal{R}_n{}^i \qquad (\boldsymbol{q}^n)_i = \mathcal{R}^n{}_i \qquad \text{x-space}$$
$$(\boldsymbol{q}'_n)^i = \mathcal{R}'_n{}^i \qquad (\boldsymbol{q}'^n)_i = \mathcal{R}'^n{}_i \qquad \text{x'-space} \ . \qquad (F.5.4)$$

Now there are *two* affine connections, one for x-space and the other for x'-space,

$$\Gamma^c{}_{ab} \equiv (\partial x^c/\partial\xi^i)\ (\partial^2\xi^i/\partial x^a\partial x^b) = \mathcal{R}^c{}_i\, \partial_a\, (\partial\xi^i/\partial x^b) = \mathcal{R}^c{}_i(\partial_a\mathcal{R}_b{}^i) \qquad \partial_a = \partial/\partial x^a$$
$$= [\boldsymbol{q}^c]_i\, (\partial_a[\boldsymbol{q}_b]^i) = \boldsymbol{q}^c \bullet (\partial_a\boldsymbol{q}_b) \qquad (F.5.5)$$

$$\Gamma'^c{}_{ab} \equiv (\partial x'^c/\partial\xi^i)\ (\partial^2\xi^i/\partial x'^a\partial x'^b) = \mathcal{R}'^c{}_i\, \partial'_a\, (\partial\xi^i/\partial x'^b) = \mathcal{R}'^c{}_i(\partial'_a\mathcal{R}'_b{}^i) \qquad \partial'_a = \partial/\partial x'^a$$
$$= [\boldsymbol{q}'^c]_i\, (\partial'_a[\boldsymbol{q}'_b]^i) = \boldsymbol{q}'^c \bullet (\partial'_a\boldsymbol{q}'_b) \ . \qquad (F.5.6)$$

The expressions appearing here in (F.5.5) are just converted versions of (F.1.7) and (F.1.5).

F.6 Relations between Γ and Γ'

The claimed relations are the following, in the context of Picture A shown above,

$$\Gamma'^{c}{}_{ab} = R^{c}{}_{d}\, R_{a}{}^{\alpha}\, R_{b}{}^{\beta}\, \Gamma^{d}{}_{\alpha\beta} + R^{c}{}_{\alpha}\, (\partial'_{a} R_{b}{}^{\alpha}) \qquad \text{// Weinberg p 100 (4.5.2)} \qquad (F.6.1)$$

$$\Gamma'^{c}{}_{ab} = R^{c}{}_{d}\, R_{a}{}^{\alpha}\, R_{b}{}^{\beta}\, \Gamma^{d}{}_{\alpha\beta} - R_{b}{}^{\beta}(\partial'_{a} R^{c}{}_{\beta}) \qquad (F.6.2)$$

$$\Gamma'^{c}{}_{ab} = R^{c}{}_{d}\, R_{a}{}^{\alpha}\, R_{b}{}^{\beta}\, \Gamma^{d}{}_{\alpha\beta} - R_{a}{}^{\alpha}\, R_{b}{}^{\beta}\, (\partial_{\alpha} R^{c}{}_{\beta}) \qquad \text{// Weinberg p 102 (4.5.8)} \qquad (F.6.3)$$

The first terms are all the same, only the second terms vary. If the second term were not present, the relation would state that $\Gamma^{d}{}_{\alpha\beta}$ transforms as a mixed rank-3 tensor in the usual manner, as in example (7.10.1). Since the second term *is* present, $\Gamma^{d}{}_{\alpha\beta}$ is *not* a tensor (unless $R_{b}{}^{\alpha}(\mathbf{x})$ = constant, in which case the transformation F is linear so $\mathbf{x}' = \mathbf{F}(\mathbf{x}) = R\mathbf{x}$ as mentioned in Section 2.8) .

Proof of the first relation : We now make use of Picture D1 shown above. Start with the Γ' definition expression shown in (F.5.6),

$$\Gamma'^{c}{}_{ab} \equiv \mathcal{R}'^{c}{}_{n}(\partial'_{a}\mathcal{R}'_{b}{}^{n}) = (R\mathcal{R})^{c}{}_{n}\, \partial'_{a}(R\mathcal{R})_{b}{}^{n} = R^{c}{}_{d}\mathcal{R}^{d}{}_{n}\, \partial'_{a}(R_{b}{}^{\beta}\mathcal{R}_{\beta}{}^{n}) \qquad \text{// } \mathcal{R}' = R\mathcal{R} \ \textit{(F.5.3)}$$

$$= R^{c}{}_{d}\mathcal{R}^{d}{}_{n}R_{b}{}^{\beta}(\partial'_{a}\mathcal{R}_{\beta}{}^{n}) + R^{c}{}_{d}(\mathcal{R}^{d}{}_{n}\mathcal{R}_{\beta}{}^{n})(\partial'_{a}R_{b}{}^{\beta}) \qquad \text{// } \mathcal{R}^{d}{}_{n}\mathcal{R}_{\beta}{}^{n} = \delta^{d}{}_{\beta} \ \textit{(7.6.4)}$$

$$= R^{c}{}_{d}\mathcal{R}^{d}{}_{n}R_{b}{}^{\beta}([R_{a}{}^{\alpha}\partial_{\alpha}]\mathcal{R}_{\beta}{}^{n}) + R^{c}{}_{d}(\delta^{d}{}_{\beta})(\partial'_{a}R_{b}{}^{\beta}) \qquad \text{// } \partial'_{a} = R_{a}{}^{\alpha}\partial_{\alpha} \text{ in first term only}$$

$$= R^{c}{}_{d}R_{a}{}^{\alpha}R_{b}{}^{\beta}\mathcal{R}^{d}{}_{n}(\partial_{\alpha}\mathcal{R}_{\beta}{}^{n}) + R^{c}{}_{\beta}\,(\partial'_{a}R_{b}{}^{\beta}) \qquad \text{// next use (F.5.5) for } \mathcal{R}^{d}{}_{n}(\partial_{\alpha}\mathcal{R}_{\beta}{}^{n})$$

$$= R^{c}{}_{d}R_{a}{}^{\alpha}R_{b}{}^{\beta}\,\Gamma^{d}{}_{\alpha\beta} + R^{c}{}_{\alpha}\,(\partial'_{a}R_{b}{}^{\alpha}) \qquad \text{// } \Gamma^{d}{}_{\alpha\beta} \equiv \mathcal{R}^{d}{}_{n}(\partial_{\alpha}\mathcal{R}_{\beta}{}^{n}) \qquad \text{QED} \qquad (F.6.4)$$

Magically, all the $\mathcal{R}$'s have gone away.

The second term in the above relation can be written a different manner as follows. Consider,

$$0 = \partial'_{a}(\delta^{c}{}_{b}) = \partial'_{a}(R^{c}{}_{\alpha}R_{b}{}^{\alpha}) = R^{c}{}_{\alpha}\ (\partial'_{a}R_{b}{}^{\alpha}) + R_{b}{}^{\alpha}\ (\partial'_{a}R^{c}{}_{\alpha})$$

$$\Rightarrow\ R^{c}{}_{\alpha}\,(\partial'_{a}R_{b}{}^{\alpha}) = -R_{b}{}^{\alpha}(\partial'_{a}R^{c}{}_{\alpha}) = -R_{b}{}^{\beta}(\partial'_{a}R^{c}{}_{\beta})$$

$$= -R_{b}{}^{\beta}\, R_{a}{}^{\alpha}(\partial_{\alpha}R^{c}{}_{\beta}) \qquad (F.6.5)$$

and this gives the other two relations stated above.

F.7 Statement and Proof of the Covariant Derivative Theorem

Many examples of this theorem will be given later. In this proof it is assumed that the tensor density of interest is purely covariant (all indices "down"). In the next Section it will then be shown how to adjust the theorem if one or more of the tensor density indices is "up".

Covariant Derivative Theorem: The covariant derivative $B_{abc..x;\alpha}$ (as defined below) of a covariant tensor $B_{abc..x}$ of rank n and weight W transforms as a covariant tensor density of rank n+1 and weight W. (F.7.1)

The implication is that all the indices including α of $B_{abc..x;\alpha}$ can be treated as ordinary tensor indices with respect to raising, lowering, contraction, and so on. The first term below in $B_{abc..x;\alpha}$ is the regular derivative $\partial_\alpha B_{abc..x}$, often written as $B_{abc..x,\alpha}$ (comma, not semicolon), and this first term is *not* a tensor. Only when all the "correction terms" are included does the object become a tensor.

The covariant derivative in x-space and then in x'-space is defined as follows: (Weinberg p 104 4.6.12)

$$B_{abc..x;\alpha} \equiv \underset{\text{del}}{\partial_\alpha B_{abc..x}} - \underset{\text{a-term}}{\Gamma^n{}_{a\alpha}B_{nbc..x}} - \underset{\text{b-term}}{\Gamma^n{}_{b\alpha}B_{anc..x}} - \ .... \ - \underset{\text{x-term}}{\Gamma^n{}_{x\alpha}B_{abc..n}} \qquad // \text{ x-space}$$

$$+ \ [W/(2g)]\, (\partial_\alpha g)\, B_{abc..x} \qquad \text{(F.7.2)}$$

$$B'_{abc..x;\alpha} \equiv \underset{\text{del}}{\partial'_\alpha B'_{abc..x}} - \underset{\text{a-term}}{\Gamma'^n{}_{a\alpha}B'_{nbc..x}} - \underset{\text{b-term}}{\Gamma'^n{}_{b\alpha}B'_{anc..x}} - \ .... \ - \underset{\text{x-term}}{\Gamma'^n{}_{x\alpha}B'_{abc..n}} \qquad // \text{ x'-space}$$

$$+ \ [W/(2g')]\, (\partial'_\alpha g')\, B'_{abc..x} \qquad \text{(F.7.2)'}$$

The two definitions are the same except everything is primed in x'-space (except constant weight W). This is as one would expect if the x-space equation were a "true tensor equation" as discussed in Section 7.15 and were therefore "covariant". Although the Γ objects are not tensors themselves, the combination of terms shown in the definition of $B_{abc..x;\alpha}$ is a rank n+1 tensor (as will be demonstrated).

As was shown in (F.4.2), $(1/2)(1/g)\partial_\alpha g = \Gamma^\kappa{}_{\kappa\alpha}$ so the W terms could be written as $W\,\Gamma^\kappa{}_{\kappa\alpha}\,B_{abc..x}$ and $W\,\Gamma'^\kappa{}_{\kappa\alpha}\,B'_{abc..x}$, and this form is commonly seen in the literature on this subject.

A proof of the theorem must then show that $B_{abc..x;\alpha}$ as defined above in fact transforms as a tensor density of rank n+1 and weight W, which is to say, one must show that

$$B'_{abc..x;\alpha} = J^{-W}\, R_a{}^{a'} R_b{}^{b'} .....\, R_x{}^{x'}\, R_\alpha{}^{\alpha'}\ B_{a'b'c'..x';\alpha'} \qquad \text{(F.7.3)}$$

or, changing abc..x→ABC..X and then a'b'c'.. → abc.. ,

$$\begin{aligned} &B'_{ABC..X;\alpha} \\ &\ = J^{-W} R_\alpha{}^{\alpha'} \{R_A{}^a R_B{}^b .....\, R_X{}^x\} \ * \\ &\qquad\qquad B_{abc..x';\alpha'} \end{aligned} \qquad \text{(F.7.4)}$$

or, in gory detail,

$$\underbrace{\partial'_\alpha B'_{ABC..x}}_{\text{del'}} - \underbrace{\Gamma'^{n}{}_{A\alpha} B'_{nBC..x}}_{\text{a'-term}} - \underbrace{\Gamma'^{n}{}_{B\alpha} B'_{AnC..x}}_{\text{b'-term}} - \ \ .... \ \ \underbrace{- \Gamma'^{n}{}_{X\alpha} B'_{ABC..n}}_{\text{x'-term}} \qquad // \text{ LHS}$$

$$+ \ [W/(2g')]\ (\partial'_\alpha g')\ B'_{ABC..x}$$

$$= J^{-W}\ R_\alpha{}^{\alpha'} \{R_A{}^a R_B{}^b ..... R_X{}^x\} \ * \qquad // \text{ RHS}$$

$$\{\underbrace{\partial_{\alpha'} B_{abc..x}}_{\text{del}} - \underbrace{\Gamma^n{}_{a\alpha'} B_{nbc..x}}_{\text{a-term}} - \underbrace{\Gamma^n{}_{b\alpha'} B_{anc..x}}_{\text{b-term}} - \ \ .... \ \ \underbrace{- \Gamma^n{}_{x\alpha'} B_{abc..n}}_{\text{x-term}}$$

$$+ (\ [W/(2g)]\ (\partial_{\alpha'} g)\ B_{abc..x}\ \} \qquad \text{(F.7.5)}$$

Proof:

There is probably a faster way to prove that (F.7.5) is valid, but we shall be content with a brute force proof where we in essence evaluate both sides of (F.7.5) and show they are the same. In doing so, we make use of identities obtained in previous Sections. The proof is carried out in five steps.

<u>1. Expand the LHS del' term and show del'-del matches the RHS del term.</u>

$$\partial'_\alpha B'_{ABC..x} = (R_\alpha{}^\beta \partial_\beta)(\ J^{-W} R_A{}^a R_B{}^b ..... R_X{}^x\ B_{abc..x}) \qquad \text{del'}$$

$$= R_\alpha{}^\beta (\partial_\beta\ J^{-W})\ R_A{}^a R_B{}^b ..... R_X{}^x\ B_{abc..x} \qquad \text{del'-J}$$

$$+ R_\alpha{}^\beta\ J^{-W}\ (\partial_\beta R_A{}^a) R_B{}^b ..... R_X{}^x\ B_{abc..x} \qquad \text{del'-a}$$

$$+ R_\alpha{}^\beta\ J^{-W}\ R_A{}^a (\partial_\beta R_B{}^b) ..... R_X{}^x\ B_{abc..x} \qquad \text{del'-b}$$

$$...$$

$$+ R_\alpha{}^\beta\ J^{-W}\ R_A{}^a R_B{}^b. \quad ...\ (\partial_\beta R_X{}^x)\ B_{abc..x} \qquad \text{del'-x}$$

$$+ R_\alpha{}^\beta\ J^{-W}\ R_A{}^a R_B{}^b ....R_X{}^x\ (\partial_\beta\ B_{abc..x}) \qquad \text{del'-del} \qquad \text{(F.7.6)}$$

We claim that this last del'-del term of the LHS matches the del term on the RHS:

$$\text{del'-del LHS} \ = \ R_\alpha{}^\beta\ J^{-W}\ \{R_A{}^a R_B{}^b ....R_X{}^x\}\ (\partial_\beta\ B_{abc..x}) \qquad \text{(F.7.7)}$$

$$\text{del RHS} \ = J^{-W}\ R_\alpha{}^{\alpha'} \{R_A{}^a R_B{}^b ..... R_X{}^x\}\ \{\partial_{\alpha'}\ B_{abc..x}\} \qquad \text{(F.7.8)}$$

Setting α' = β shows that these two terms indeed match.

<u>2. Show that the a-related terms balance.</u> The a'-term on the LHS of (F.7.5) is

$$- \Gamma'^{n}{}_{A\alpha} B'_{nBC..x}\ . \qquad // \text{ a'-term of LHS} \qquad \text{(F.7.9)}$$

The connection (F.6.3) between Γ' and Γ reads (then shuffle indices as shown),

$$\Gamma'^{c}{}_{ab} = R^c{}_d\ R_a{}^\alpha\ R_b{}^\beta\ \Gamma^d{}_{\alpha\beta} \ - \ R_a{}^\alpha\ R_b{}^\beta\ (\partial_\alpha R^c{}_\beta) \quad // \ \alpha \to \kappa, \ \ \beta \to \sigma \qquad \textit{(F.6.3)}$$

$$\Gamma'^{c}{}_{ab} = R^c{}_d\ R_a{}^\kappa\ R_b{}^\sigma\ \Gamma^d{}_{\kappa\sigma} \ - \ R_a{}^\kappa\ R_b{}^\sigma\ (\partial_\kappa R^c{}_\sigma) \quad // \ c \to n, \ \ a \to A, \ b \to \alpha$$

$$\Gamma'^{n}{}_{A\alpha} = R^n{}_d\ R_A{}^\kappa\ R_\alpha{}^\sigma\ \Gamma^d{}_{\kappa\sigma} \ - \ R_A{}^\kappa\ R_\alpha{}^\sigma\ (\partial_\kappa R^n{}_\sigma) \quad . \qquad \text{(F.7.10)}$$

Now insert into (F.7.9) the expression (F.7.10) for $\Gamma'^{n}{}_{A\alpha}$ and the tensor transformation rule for $B'_{nBC..x}$,

$$-\{ R^n{}_d R_A{}^\kappa R_\alpha{}^\sigma \Gamma^d{}_{\kappa\sigma} - R_A{}^\kappa R_\alpha{}^\sigma (\partial_\kappa R^n{}_\sigma)\} \{ J^{-W} R_n{}^{n'} R_B{}^b R_C{}^c....R_x{}^x B_{n'bc..x} \} \quad // \text{ a'-term} \tag{F.7.11}$$

and to this we must add the contribution del'-a shown in (F.7.6).

Meanwhile, the RHS a-term in (F.7.5) is this

$$J^{-W} R_\alpha{}^{\alpha'} \{R_A{}^a R_B{}^b..... R_X{}^x\}\{-\Gamma^n{}_{a\alpha'} B_{nbc..x}\} \qquad // \text{ now do } n\to n'$$

$$= -\{ J^{-W} R_\alpha{}^{\alpha'} \{R_A{}^a R_B{}^b..... R_X{}^x\}\{\Gamma^{n'}{}_{a\alpha'} B_{n'bc..x}\} \quad . \qquad // \text{ a-term of RHS} \tag{F.7.12}$$

We want to show that

a'-term + del'-a = a-term

or

LHS a'-term (F.7.11) + LHS del'-a term (F.7.6) = RHS a-term (F.7.12)

or

$$\begin{aligned} &-\{ R^n{}_d R_A{}^\kappa R_\alpha{}^\sigma \Gamma^d{}_{\kappa\sigma} - R_A{}^\kappa R_\alpha{}^\sigma (\partial_\kappa R^n{}_\sigma)\} \{ J^{-W} R_n{}^{n'} R_B{}^b R_C{}^c....R_x{}^x B_{n'bc..x} \} \\ &\quad + R_\alpha{}^\beta J^{-W} (\partial_\beta R_A{}^{n'}) R_B{}^b..... R_X{}^x B_{n'bc..x} \qquad // \text{ a}\to\text{n' , this is the del'-a term} \\ &= -\ J^{-W} R_\alpha{}^{\alpha'} \{R_A{}^a R_B{}^b..... R_X{}^x\}\{\Gamma^{n'}{}_{a\alpha'} B_{n'bc..x}\} \qquad ? \end{aligned} \tag{F.7.13}$$

We can see that the factors $J^{-W} R_B{}^b R_C{}^c....R_x{}^x B_{n'bc..x}$ are common to both sides so they can be removed to give a simpler relation which we must show to be valid:

$$\begin{aligned} &-\{ R^n{}_d R_A{}^\kappa R_\alpha{}^\sigma \Gamma^d{}_{\kappa\sigma} - R_A{}^\kappa R_\alpha{}^\sigma (\partial_\kappa R^n{}_\sigma)\} \{ R_n{}^{n'} \} \\ &\quad + R_\alpha{}^\beta (\partial_\beta R_A{}^{n'}) \\ &= -\ R_\alpha{}^{\alpha'} \{R_A{}^a \}\{\Gamma^{n'}{}_{a\alpha'}\} \ . \end{aligned} \tag{F.7.14}$$

In the first term orthogonality (7.6.4) says $R^n{}_d R_n{}^{n'} = \delta_d{}^{n'}$ which pins d to n' in that term only, so the above becomes

$$- R_A{}^\kappa R_\alpha{}^\sigma \Gamma^{n'}{}_{\kappa\sigma} + R_n{}^{n'} R_A{}^\kappa R_\alpha{}^\sigma (\partial_\kappa R^n{}_\sigma) + R_\alpha{}^\beta (\partial_\beta R_A{}^{n'}) = - R_\alpha{}^\sigma R_A{}^\kappa \Gamma^{n'}{}_{\kappa\sigma} \ . \tag{F.7.15}$$

The first term on the left cancels the only term on the right. Then do β→σ in the third term to get,

$$R_A{}^\kappa R_\alpha{}^\sigma \{R_n{}^{n'} (\partial_\kappa R^n{}_\sigma)\} + R_\alpha{}^\sigma (\partial_\sigma R_A{}^{n'}) = 0 \ . \tag{F.7.16}$$

Now cancel the common $R_\alpha{}^\sigma$ factor and use the symmetry $(\partial_\kappa R^n{}_\sigma) = (\partial_\sigma R^n{}_\kappa)$ to get

$$(\partial_\sigma R_A{}^{n'}) = - R_n{}^{n'} R_A{}^\kappa (\partial_\sigma R^n{}_\kappa) \ . \tag{F.7.17}$$

Now in this order do σ→a, A→d, n→e, n'→n, κ→m to get

$$(\partial_a R_d{}^n) = -R_e{}^n R_d{}^m (\partial_a R^e{}_m) \;. \qquad \text{(F.7.18)}$$

But this is seen to be the third identity of (F.2.1)! By reversing the above sequence of steps, one shows that the three a-related terms in the above LHS = RHS equation balance:

a'-term + del'-a = a-term.

or

$$-\{ R^n{}_d R_A{}^\kappa R_\alpha{}^\sigma \Gamma^d{}_{\kappa\sigma} - R_A{}^\kappa R_\alpha{}^\sigma (\partial_\kappa R^n{}_\sigma)\} \{ J^{-W} R_n{}^{n'} R_B{}^b R_C{}^c....R_x{}^x B_{n'bc..x} \}$$
$$+ R_\alpha{}^\beta J^{-W} (\partial_\beta R_A{}^{n'}) R_B{}^b..... R_X{}^x B_{n'bc..x} \qquad \text{// a→n' , this is the del'-a term}$$

$$= -\; J^{-W} R_\alpha{}^{\alpha'} \{R_A{}^a R_B{}^b..... R_X{}^x\}\{\Gamma^{n'}{}_{a\alpha'} B_{n'bc..x}\} \qquad ? \qquad \textit{(F.7.13)}$$

In reversing the sequence, one of course adds back the deleted common factors as well as the various implied index sums. It seems clearer to state the proof this way rather than start with the last equality with no justification and artificially thread backwards to the desired equation. This method is used as well in the next section.

3. Show that the b-related terms balance. We have proven that *(F.7.13)* is valid, and it is true for any tensor $B_{n'bc..x}$, since this is a common factor on both sides. Eq (F.7.13) is therefore still valid if we replace the tensor $B_{n'bc..x}$ by the different tensor $Q_{n'bc..x} \equiv B_{bn'c..x}$. Tensor Q is tensor B with the first two indices swapped. Make this replacement, and after doing so, make the free index swap A↔B and summation index swap a↔b. One ends up then with this *known-valid* equation:

$$-\{ R^n{}_d R_B{}^\kappa R_\alpha{}^\sigma \Gamma^d{}_{\kappa\sigma} - R_B{}^\kappa R_\alpha{}^\sigma (\partial_\kappa R^n{}_\sigma)\} \{ J^{-W} R_n{}^{n'} R_A{}^a R_C{}^c....R_x{}^x B_{an'c..x} \}$$
$$+ R_\alpha{}^\beta J^{-W} (\partial_\beta R_B{}^{n'}) R_A{}^a..... R_X{}^x B_{an'c..x}$$

$$= -\; J^{-W} R_\alpha{}^{\alpha'} \{R_B{}^b R_A{}^a..... R_X{}^x\}\{\Gamma^{n'}{}_{b\alpha'} B_{an'c..x} \} \;. \qquad \text{(F.7.19)}$$

As the reader may suspect, the three terms in this last equation will have this interpretation with respect to equation (F.7.5) :

b'-term + del'-b = b-term .

To verify this claim, we treat the terms one at a time.

b'-term: First state the equation (F.7.9) = (F.7.11) on the first line below, then get the second line doing A↔B and dummy b↔ a :

$$- \Gamma'^n{}_{A\alpha} B'_{nBC..x} \;=\; -\{ R^n{}_d R_A{}^\kappa R_\alpha{}^\sigma \Gamma^d{}_{\kappa\sigma} - R_A{}^\kappa R_\alpha{}^\sigma (\partial_\kappa R^n{}_\sigma)\} \{ J^{-W} R_n{}^{n'} R_B{}^b R_C{}^c....R_x{}^x B_{n'bc..x} \}$$
$$- \Gamma'^n{}_{B\alpha} B'_{nAC..x} \;=\; -\{ R^n{}_d R_B{}^\kappa R_\alpha{}^\sigma \Gamma^d{}_{\kappa\sigma} - R_B{}^\kappa R_\alpha{}^\sigma (\partial_\kappa R^n{}_\sigma)\} \{ J^{-W} R_n{}^{n'} R_A{}^a R_C{}^c....R_x{}^x B_{n'ac..x} \}$$

Since this second equation is valid for any tensor $B_{n'ac..x}$, it is true for $Q_{n'ac..x} \equiv B_{an'c..x}$ where the first two indices are swapped. This gives,

$$- \Gamma'^n{}_{B\alpha} B'_{AnC..x} \;=\; -\{ R^n{}_d R_B{}^\kappa R_\alpha{}^\sigma \Gamma^d{}_{\kappa\sigma} - R_B{}^\kappa R_\alpha{}^\sigma (\partial_\kappa R^n{}_\sigma)\} \{ J^{-W} R_n{}^{n'} R_A{}^a R_C{}^c....R_x{}^x B_{an'c..x} \}$$

The left side is the b'-term on the LHS of (F.7.5) and the right side is the first term in (F.7.19).

del'-b: The del'-b term in (F.7.6) is the first line below,

$$\begin{aligned}&R_\alpha{}^\beta\, J^{-W} R_A{}^a(\partial_\beta R_B{}^b\ )..... R_X{}^x\, B_{abc..x} && \text{// b→n}\\ &R_\alpha{}^\beta\, J^{-W} R_A{}^a(\partial_\beta R_B{}^{n'})..... R_X{}^x\, B_{an'c..x}\end{aligned}$$

The second line matches the second term in (F.7.19).

b-term: The b-term on the RHS of (F.7.5) is

$$\begin{aligned}&-J^{-W}\ R_\alpha{}^{\alpha'}\{R_A{}^a R_B{}^b..... R_X{}^x\}[\ \Gamma^n{}_{b\alpha'} B_{anc..x}] && \text{// n→n'}\\ &-J^{-W}\ R_\alpha{}^{\alpha'}\{R_A{}^a R_B{}^b..... R_X{}^x\}[\ \Gamma^{n'}{}_{b\alpha'} B_{an'c..x}]\end{aligned}$$

The second line matches the right side of (F.7.19).

We have thus shown that the b terms balance in equation (F.7.5), just the way the a terms balance.

4. Similarly, the c,d.....x terms match. Just repeat step 3 above for each extra index.

5. It remains to show that the three terms so far neglected in (F.7.5) match as well. These terms are

$$\begin{aligned}\text{LHS:}\quad & R_\alpha{}^\beta(\partial_\beta\, J^{-W})\, R_A{}^a R_B{}^b..... R_X{}^x\, B_{abc..x} && \text{// del'-J in (F.7.6)}\\ & +\ \ [W/(2g')]\ (\partial'_\alpha g')\ B'_{ABC..x} && \text{// the LHS W term}\qquad (F.7.20)\\[1em] \text{RHS:}\quad & J^{-W}\ R_\alpha{}^{\alpha'}\{R_A{}^a R_B{}^b..... R_X{}^x\}\ [W/(2g)]\ (\partial_{\alpha'} g)\ B_{abc..x} && \text{// the RHS W term}\end{aligned}$$

That is, one must show that

$$\begin{aligned}&R_\alpha{}^\beta(\partial_\beta\, J^{-W})\, R_A{}^a R_B{}^b..... R_X{}^x\, B_{abc..x} + [W/(2g')]\ (\partial'_\alpha g')\ B'_{ABC..x}\\ &\quad = J^{-W}\ R_\alpha{}^{\alpha'}\{R_A{}^a R_B{}^b..... R_X{}^x\}\ [W/(2g)]\ (\partial_{\alpha'} g)\ B_{abc..x}\quad .\end{aligned}\tag{F.7.21}$$

Inserting the tensor transformation (F.7.4) of $B'_{ABC..x}$ in the second term gives

$$\begin{aligned}&R_\alpha{}^\beta(\partial_\beta\, J^{-W})\, R_A{}^a R_B{}^b..... R_X{}^x\, B_{abc..x} +\ [W/(2g')]\ (\partial'_\alpha g')\ J^{-W}\{R_A{}^a R_B{}^b..... R_X{}^x\}\ *\ \ B_{abc..x}\\ &\quad = J^{-W}\ R_\alpha{}^{\alpha'}\{R_A{}^a R_B{}^b..... R_X{}^x\}\ [W/(2g)]\ (\partial_{\alpha'} g)\ B_{abc..x}\quad .\end{aligned}\tag{F.7.22}$$

Set β = α' in the first term to get

$$\begin{aligned}&R_\alpha{}^{\alpha'}(\partial_{\alpha'}\, J^{-W})\, R_A{}^a R_B{}^b..... R_X{}^x\, B_{abc..x} +\ [W/(2g')]\ (\partial'_\alpha g')\ J^{-W}\{R_A{}^a R_B{}^b..... R_X{}^x\}\ *\ \ B_{abc..x}\\ &\quad = J^{-W}\ R_\alpha{}^{\alpha'}\{R_A{}^a R_B{}^b..... R_X{}^x\}\ [W/(2g)]\ (\partial_{\alpha'} g)\ B_{abc..x}\quad .\end{aligned}\tag{F.7.23}$$

Next, remove the common factor $R_A{}^a R_B{}^b$..... $R_X{}^x B_{abc..x}$ (and associated implied sums) to get

$$R_\alpha{}^{\alpha'} (\partial_{\alpha'} J^{-W}) + [W/(2g')] (\partial'_\alpha g') J^{-W} = J^{-W} R_\alpha{}^{\alpha'} [W/(2g)] (\partial_{\alpha'} g) \ . \tag{F.7.24}$$

Since $\partial'_\alpha = R_\alpha{}^{\alpha'} \partial_{\alpha'}$ (covariant vector transformation) this becomes

$$R_\alpha{}^{\alpha'} (\partial_{\alpha'} J^{-W}) + [W/(2g')] (R_\alpha{}^{\alpha'} \partial_{\alpha'} g') J^{-W} = J^{-W} R_\alpha{}^{\alpha'} [W/(2g)] (\partial_{\alpha'} g)$$

or

$$(\partial'_\alpha J^{-W}) + [W/(2g')] (\partial'_\alpha g') J^{-W} = J^{-W} [W/(2g)] (\partial'_\alpha g) \ . \tag{F.7.25}$$

Move the second term to the RHS and do the left side derivative to get

$$(-W) J^{-W-1}(\partial'_\alpha J) = J^{-W} (W/2) [(\partial'_\alpha g)/g – (\partial'_\alpha g')/g'] \tag{F.7.26}$$

so it remains then to show that

$$J^{-1}(\partial'_\alpha J) = (1/2)[(\partial'_\alpha g')/g' - (\partial'_\alpha g)/g] \ . \tag{F.7.27}$$

From (5.12.14) one has $J^2 = g'/g$ and $J = (g'/g)^{1/2}$ so we need to show that

$$\partial'_\alpha (g'/g)^{1/2} = (1/2) (g'/g)^{1/2}[(\partial'_\alpha g')/g' - (\partial'_\alpha g)/g]$$

or

$$(1/2) (g'/g)^{-1/2} \partial'_\alpha (g'/g) = (1/2) (g'/g)^{1/2}[(\partial'_\alpha g')/g' - (\partial'_\alpha g)/g]$$

or

$$\partial'_\alpha (g'/g) = (g'/g) [(\partial'_\alpha g')/g' - (\partial'_\alpha g)/g] \tag{F.7.28}$$

Evaluation of the left side of (F.7.28) gives

$$\partial'_\alpha (g'/g) = [g(\partial'_\alpha g') - g'(\partial'_\alpha g)]/g^2 = (1/g) (\partial'_\alpha g') - (g'/g^2)(\partial'_\alpha g) \ .$$

Evaluation of right side of (F.7.28) gives

$$(g'/g) [(\partial'_\alpha g')/g' - (\partial'_\alpha g)/g] = (1/g)(\partial'_\alpha g') - (g'/g^2) (\partial'_\alpha g) \ ,$$

Since these agree, (F.7.27) is valid and therefore working backwards we find that (F.7.20) is valid.

This concludes our lengthy 5-step proof of the Covariant Derivative Theorem stated in (F.7.1).

F.8 Rules for raising any index on a covariant derivative of a covariant tensor density.

The Rules are stated in (F.8.9) and (F.8.11), but will only make sense after looking at the examples presented here.

Consider the general form given in (F.7.2) for a covariant derivative

$$B_{abc..x;\alpha} \equiv \underbrace{\partial_\alpha B_{abc..x}}_{\text{del}} - \underbrace{\Gamma^n{}_{a\alpha}B_{nbc..x}}_{\text{a-term}} - \underbrace{\Gamma^n{}_{b\alpha}B_{anc..x}}_{\text{b-term}} - \;....\; - \underbrace{\Gamma^n{}_{x\alpha}B_{abc..n}}_{\text{x-term}} \qquad \text{// x-space}$$
$$+ (W/2g)\,(\partial_\alpha g)\,B_{abc..x} \;. \qquad \text{(F.8.1)}$$

Notice that there are n indices on $B_{abc..x}$ and there are n corresponding terms on the RHS in addition to the del and W terms. Each index of $B_{abc..x}$ thus has its own "correction term". What happens if one of the indices (say b) on $B_{abc..x}$ is raised? To find out, apply $g^{\beta b}$ to both sides. The effect of doing this is trivial for all terms *except* the del term and the b-term, since b is a regular tensor index on all such terms, so we get

$$B_a{}^\beta{}_{c..x;\alpha} \equiv \underbrace{g^{\beta b}\,\partial_\alpha B_{abc..x}}_{\text{del}} - \underbrace{\Gamma^n{}_{a\alpha}B_n{}^\beta{}_{c..x}}_{\text{a-term}} - \underbrace{g^{\beta b}\,\Gamma^n{}_{b\alpha}B_{anc..x}}_{\text{b-term}} - \;....\; - \underbrace{\Gamma^n{}_{x\alpha}B_a{}^\beta{}_{c..n}}_{\text{x-term}}$$
$$+ (W/2g)\,(\partial_\alpha g)\,B_a{}^\beta{}_{c..x} \;. \qquad \text{(F.8.2)}$$

The del term can be written

$$\begin{aligned} g^{\beta b}\,(\partial_\alpha B_{abc..x}) &= \partial_\alpha\,(g^{\beta b} B_{abc..x}) - (\partial_\alpha g^{\beta b})\,B_{abc..x} \\ &= \partial_\alpha B_a{}^\beta{}_{c..x} \qquad - (\partial_\alpha g^{\beta b})\,B_{abc..x} \;. \end{aligned} \qquad \text{(F.8.3)}$$

The second term del term here can be combined with the b-term to give

$$\begin{aligned} \text{del-2nd + b-term} &= -(\partial_\alpha g^{\beta b})\,B_{abc..x} \;-\; g^{\beta b}\,\Gamma^n{}_{b\alpha}B_{anc..x} \\ &= -(\partial_\alpha g^{\beta n})\,B_{anc..x} \;-\; g^{\beta b}\,\Gamma^n{}_{b\alpha}B_{anc..x} && \text{// b}\rightarrow\text{n in first term only} \\ &= [-(\partial_\alpha g^{\beta n}) - g^{\beta b}\,\Gamma^n{}_{b\alpha}]\,B_{anc..x} \;. \end{aligned} \qquad \text{(F.8.4)}$$

The identity (F.3.1) with n→ i reads,

$$(\partial_c g^{ab}) \;=\; -\,[g^{ai}\,\Gamma^b{}_{ci} + g^{bi}\,\Gamma^a{}_{ci}]$$

or

$$\begin{aligned} [-(\partial_c g^{ab}) - g^{ai}\,\Gamma^b{}_{ci}] &= g^{bi}\,\Gamma^a{}_{ci} && \text{// do c}\rightarrow\alpha\text{, b}\rightarrow\text{n, a}\rightarrow\beta \\ [-(\partial_\alpha g^{\beta n}) - g^{\beta i}\,\Gamma^n{}_{\alpha i}] &= g^{ni}\,\Gamma^\beta{}_{\alpha i} && \text{// then i}\rightarrow\text{b on the left} \\ [-(\partial_\alpha g^{\beta n}) - g^{\beta b}\,\Gamma^n{}_{\alpha b}] &= g^{ni}\,\Gamma^\beta{}_{\alpha i} \;. \end{aligned} \qquad \text{(F.8.5)}$$

Replacing [...] in (F.8.4) by the last result gives,

$$\text{del-2nd + b-term} \;= g^{ni}\,\Gamma^\beta{}_{\alpha i}\,B_{anc..x} \;=\; \Gamma^\beta{}_{\alpha i}\,B_a{}^i{}_{c..x} \;=\; \Gamma^\beta{}_{\alpha n}\,B_a{}^n{}_{c..x} \qquad \text{(F.8.6)}$$

so it has been shown that

$$B_a{}^\beta{}_{c..x;\alpha} \equiv \partial_\alpha B_a{}^\beta{}_{c..x} - \Gamma^n{}_{a\alpha}B_n{}^\beta{}_{c..x} + \underbrace{\Gamma^\beta{}_{\alpha n}\,B_a{}^n{}_{c..x}}_{\text{new b term}} - ... - \Gamma^n{}_{x\alpha}B_a{}^\beta{}_{c..n} + W\,\Gamma^\kappa{}_{\kappa\alpha}\,B_a{}^b{}_{c..x} \qquad \text{(F.8.7)}$$

Here then is a comparison where β→b in the second line:

$$\begin{aligned} B_{abc..x;\alpha} &\equiv \partial_\alpha B_{abc..x} - \Gamma^n{}_{a\alpha}B_{nbc..x} - \Gamma^n{}_{b\alpha}B_{anc..x} - .... - \Gamma^n{}_{x\alpha}B_{abc..n} + W\,\Gamma^\kappa{}_{\kappa\alpha} B_{abc..x} \\ B_a{}^b{}_{c..x;\alpha} &\equiv \partial_\alpha B_a{}^b{}_{c..x} - \Gamma^n{}_{a\alpha}B_n{}^b{}_{c..x} + \Gamma^b{}_{\alpha n} B_a{}^n{}_{c..x} - .... - \Gamma^n{}_{x\alpha}B_a{}^b{}_{c..n} + W\,\Gamma^\kappa{}_{\kappa\alpha} B_a{}^b{}_{c..x} \end{aligned}$$

del a-term b-term x-term W-term

(F.8.8)

This demonstrates the following rule (stated for q = b or any other index):

Rule for raising some non-last index q:

(1) In all terms, raise the q-position B index (in the q-term that index is called n)
(2) in the q correction term, make the replacement $-\Gamma^n{}_{q\alpha} \rightarrow +\Gamma^q{}_{n\alpha}$ $(= \Gamma^q{}_{\alpha n})$ (F.8.9)

We refer to a correction term like the a-term and x-term in (F.8.8) as a "covariant correction term", and a correction term like the b-term as a "contravariant correction term". Thus, for an arbitrary up/down placement of the abc..x indices, there will be a covariant correction term for each down index, and a contravariant correction term for each up index. Here is another example where x is also taken up:

$$B_a{}^b{}_{c..}{}^x{}_{;\alpha} \equiv \partial_\alpha B_a{}^b{}_{c..}{}^x - \Gamma^n{}_{a\alpha}B_n{}^b{}_{c..}{}^x + \Gamma^b{}_{\alpha n} B_a{}^n{}_{c..}{}^x - .... + \Gamma^x{}_{n\alpha}B_a{}^b{}_{c..}{}^n + W\,\Gamma^\kappa{}_{\kappa\alpha} B_a{}^b{}_{c..}{}^x$$

del a-term b-term x-term W-term

(F.8.10)

where the contravariant correction terms are marked in red.

Rule for raising the last index α: Raising the ;α index must be done "manually" so the first term will have $g^{\alpha\alpha'}\partial_\alpha = \partial^{\alpha'}$ and all remaining terms will have explicit $g^{\alpha\alpha'}$ factors. (F.8.11)

F.9 Examples of covariant derivative expressions

Example J = 0 (covariant derivative of a *scalar* B) // J is the rank of the B tensor

$B_{;\alpha} = \partial_\alpha B$ covariant vector $// = B_{,\alpha}$

$B^{;\alpha} = \partial^\alpha B$ contravariant vector $// = B^{,\alpha}$ (F.9.1)

Example J=1: (covariant derivative of a vector **B**)

$B_{a;\alpha} = \partial_\alpha B_a - \Gamma^n{}_{a\alpha}B_n$ covariant rank-2 tensor // 2nd term is sym on a↔α (F.9.2)

$B^a{}_{;\alpha} = \partial_\alpha B^a + \Gamma^a{}_{\alpha n} B^n$ mixed rank-2 tensor (F.9.3)

To obtain $B_a{}^{;\alpha}$, write (F.9.2) as $B_{a;\beta} = \partial_\beta B_a - \Gamma^n{}_{a\beta}B_n$ and apply $g^{\alpha\beta}$ to both sides. The result is then $B_{a;}{}^\alpha = \partial^\alpha B_a - g^{\alpha\beta}\Gamma^n{}_{a\beta}B_n$ .Finlly, raise a and alter the correction term to get $B^{a;\alpha} = \partial^\alpha B^a + g^{\alpha\beta}\Gamma^a{}_{n\beta}B^n$.

$B_a{}^{;\alpha} = \partial^\alpha B_a - g^{\alpha\beta}\Gamma^n{}_{a\beta}B_n$ mixed rank-2 tensor (F.9.4)

$B^{a;\alpha} = \partial^\alpha B^a + g^{\alpha\beta}\Gamma^a{}_{\beta n} B^n$ contravariant rank-2 tensor (F.9.5)

Example J=2: (covariant derivative of a rank-2 tensor)

$$B_{ab;\alpha} \equiv \partial_\alpha B_{ab} - \Gamma^n{}_{a\alpha}B_{nb} - \Gamma^n{}_{b\alpha}B_{an} \qquad \text{covariant rank-3 tensor} \qquad (F.9.6)$$

$$B^a{}_{b;\alpha} \equiv \partial_\alpha B^a{}_b + \Gamma^a{}_{\alpha n} B^n{}_b - \Gamma^n{}_{b\alpha}B^a{}_n \qquad \text{etc.} \qquad (F.9.7)$$

$$B_a{}^b{}_{;\alpha} \equiv \partial_\alpha B_a{}^b - \Gamma^n{}_{a\alpha}B_n{}^b + \Gamma^b{}_{\alpha n} B_a{}^n \qquad (F.9.8)$$

$$B^{ab}{}_{;\alpha} \equiv \partial_\alpha B^{ab} + \Gamma^a{}_{\alpha n} B^{nb} + \Gamma^b{}_{\alpha n} B^{an} \quad . \qquad (F.9.9)$$

Again, application of $g^{\alpha\beta}$ would give expressions for the other four possibilities with ;α being "up". These "other possibilities" are always present, but we shall no longer mention them in the following examples.

Example J=3: (covariant derivative of a rank-3 tensor)

$$B_{abc;\alpha} \equiv \partial_\alpha B_{abc} - \Gamma^n{}_{a\alpha}B_{nbc} - \Gamma^n{}_{b\alpha}B_{anc} - \Gamma^n{}_{c\alpha}B_{abn} \qquad (F.9.10)$$

$$B^a{}_{bc;\alpha} \equiv \partial_\alpha B^a{}_{bc} + \Gamma^a{}_{\alpha n} B^n{}_{bc} - \Gamma^n{}_{b\alpha}B^a{}_{nc} - \Gamma^n{}_{c\alpha}B^a{}_{bn} \qquad (F.9.11)$$

...

$$B^{abc}{}_{;\alpha} \equiv \partial_\alpha B^{abc} + \Gamma^a{}_{n\alpha}B^{nbc} + \Gamma^b{}_{n\alpha}B^{anc} + \Gamma^c{}_{n\alpha}B^{abn} \qquad (F.9.12)$$

Special J=2 application to the metric tensor:

$$g_{ab;\alpha} \equiv \partial_\alpha g_{ab} - \Gamma^n{}_{a\alpha}g_{nb} - \Gamma^n{}_{b\alpha}g_{an} \quad = 0 \qquad \text{// by identity (F.3.2)} \qquad (F.9.13)$$

$$g^a{}_{b;\alpha} \equiv \partial_\alpha g^a{}_b + \Gamma^a{}_{\alpha n} g^n{}_b - \Gamma^n{}_{b\alpha}g^a{}_n \quad = \; 0 + \Gamma^a{}_{\alpha b} - \Gamma^a{}_{b\alpha} = 0 \qquad \text{// } g^a{}_b = \delta^a{}_b \qquad (F.9.14)$$

$$g_a{}^b{}_{;\alpha} \equiv \partial_\alpha g_a{}^b - \Gamma^n{}_{a\alpha}g_n{}^b + \Gamma^b{}_{\alpha n} g_a{}^n \quad = \; 0 - \Gamma^b{}_{a\alpha} + \Gamma^b{}_{\alpha a} = 0 \qquad (F.9.15)$$

$$g^{ab}{}_{;\alpha} \equiv \partial_\alpha g^{ab} + \Gamma^a{}_{\alpha n} g^{nb} + \Gamma^b{}_{\alpha n} g^{an} \quad = 0 \qquad \text{// by identity (F.3.1)} \qquad (F.9.16)$$

The middle lines use the fact (7.4.19) that $g^a{}_b = \delta^a{}_b$. Since $g_{ab;\alpha}$ is a tensor, knowing that any one of the above vanishes implies that all four lines vanish! The net result is

$$g_{ab;\alpha} = g^a{}_{b;\alpha} = g_a{}^b{}_{;\alpha} = g^{ab}{}_{;\alpha} = 0 \quad . \qquad \text{// Weinberg p 105 (4.6.16,17,18)} \qquad (F.9.17)$$

The covariant derivative of any form of the metric tensor vanishes. As Weinberg points one, one knows that in a quasi-Cartesian x-space $g_{ab;\alpha} = 0$ since $g_{ab} = G_{aa}\delta_{a,b}$ = constant and $\Gamma = 0$. Then in any x'-space $g'_{ab;\alpha} = 0$ as well since $g'_{ab;\alpha} = R_a{}^{a'} R_b{}^{b'} R_\alpha{}^{\alpha'} g_{a'b';\alpha'}$.

Example J=2: (double covariant derivatives)

Consider again the J=1 examples from above

$$\begin{aligned} B_{a;\alpha} &= \partial_\alpha B_a - \Gamma^n{}_{a\alpha} B_n && \textit{(F.9.2)} \\ B^a{}_{;\alpha} &= \partial_\alpha B^a + \Gamma^a{}_{\alpha n} B^n \quad . && \textit{(F.9.3)} \end{aligned}$$

This applies to *any* vector B_a. As the J=0 example shows, $B_{;a}$ and $B^{;a}$ are bona-fide vectors (covariant and contravariant components of the same vector) and therefore

$$\begin{aligned} B_{;a;\alpha} &= \partial_\alpha B_{;a} - \Gamma^n{}_{a\alpha} B_{;n} && \text{(F.9.18)} \\ B^{;a}{}_{;\alpha} &= \partial_\alpha B^{;a} + \Gamma^a{}_{\alpha n} B^{;n} && \text{(F.9.19)} \end{aligned}$$

where we have simply inserted a semicolon in each term. In these equations B is a scalar.

Consider again the J=2 examples from above,

$$\begin{aligned} B_{ab;\alpha} &= \partial_\alpha B_{ab} - \Gamma^n{}_{a\alpha} B_{nb} - \Gamma^n{}_{b\alpha} B_{an} && \textit{(F.9.6)} \\ B^a{}_{b;\alpha} &= \partial_\alpha B^a{}_b + \Gamma^a{}_{\alpha n} B^n{}_b - \Gamma^n{}_{b\alpha} B^a{}_n \quad . && \textit{(F.9.7)} \end{aligned}$$

This applies to any rank-2 tensors B_{ab} or $B^a{}_b$. But $B_{a;b}$ and $B^a{}_{;b}$ are bona-fide rank-2 tensors, and therefore we can restate the above as,

$$\begin{aligned} B_{a;b;\alpha} &= \partial_\alpha B_{a;b} - \Gamma^n{}_{a\alpha} B_{n;b} - \Gamma^n{}_{b\alpha} B_{a;n} && \text{(F.9.20)} \\ B^a{}_{;b;\alpha} &= \partial_\alpha B^a{}_{;b} + \Gamma^a{}_{\alpha n} B^n{}_{;b} - \Gamma^n{}_{b\alpha} B^a{}_{;n} && \text{(F.9.21)} \end{aligned}$$

In a similar manner one can derive expressions for triple covariant derivatives and beyond. For example

$$B_{a;b;c;\alpha} = \partial_\alpha B_{a;b;c} - \Gamma^n{}_{a\alpha} B_{n;b;c} - \Gamma^n{}_{b\alpha} B_{a;n;c} - \Gamma^n{}_{c\alpha} B_{a;b;n} \quad . \qquad \text{(F.9.22)}$$

The next examples are for **tensor densities** so there will be a W term as shown in (F.8.8).

Example J = 0 (scalar density of weight W) // recall $\Gamma^\kappa{}_{\kappa\alpha} = (2g)^{-1}\partial_\alpha g$ *from (F.4.2)*

$$B_{;\alpha} = \partial_\alpha B + W\, \Gamma^\kappa{}_{\kappa\alpha} B \qquad \text{(F.9.23)}$$

Example J = 1 (vector density of weight W)

$$B_{a;\alpha} = \partial_\alpha B_a - \Gamma^n{}_{a\alpha} B_n + W\, \Gamma^\kappa{}_{\kappa\alpha} B_a$$

covariant rank-2 tensor density

$$B^a{}_{;\alpha} = \partial_\alpha B^a + \Gamma^a{}_{\alpha n} B^n + W\, \Gamma^\kappa{}_{\kappa\alpha} B^a \qquad \text{(F.9.24)}$$

Example J = 2 (rank-2 tensor density of weight W)

$$B_{ab;\alpha} = \partial_\alpha B_{ab} - \Gamma^n{}_{a\alpha}B_{nb} - \Gamma^n{}_{b\alpha}B_{an} + W\,\Gamma^\kappa{}_{\kappa\alpha}\,B_{ab} \qquad \text{covariant rank-3 tensor density} \tag{F.9.25}$$

$$B^a{}_{b;\alpha} = \partial_\alpha B^a{}_b + \Gamma^a{}_{\alpha n}\,B^n{}_b - \Gamma^n{}_{b\alpha}B^a{}_n + W\,\Gamma^\kappa{}_{\kappa\alpha}\,B^a{}_b \tag{F.9.26}$$

$$B_a{}^b{}_{;\alpha} = \partial_\alpha B_a{}^b - \Gamma^n{}_{a\alpha}B_n{}^b + \Gamma^b{}_{\alpha n}\,B_a{}^n + W\,\Gamma^\kappa{}_{\kappa\alpha}\,B_a{}^b \tag{F.9.27}$$

$$B^{ab}{}_{;\alpha} = \partial_\alpha B^{ab} + \Gamma^a{}_{\alpha n}\,B^{nb} + \Gamma^b{}_{\alpha n}\,B^{an} + W\,\Gamma^\kappa{}_{\kappa\alpha}\,B^{ab} \quad . \tag{F.9.28}$$

As noted in Example 2 below (D.2.3), adding a factor $g^{W/2}$ to a tensor density of weight W neutralizes the weight, and the result is a regular tensor. Here then are a few examples in which this is done. Since the product is a tensor, there are no W correction terms.

Example J = 0 (covariant derivative of a scalar density B of weight W)

$$(g^{W/2}B)_{;\alpha} = \partial_\alpha(g^{W/2}B) \qquad \text{covariant vector} \tag{F.9.29}$$

$$(g^{W/2}B)^{;\alpha} = \partial^\alpha(g^{W/2}B) \qquad \text{contravariant vector} \tag{F.9.30}$$

Example J=1: (covariant derivative of a vector density **B** of weight W)

$$(g^{W/2}B_a)_{;\alpha} = \partial_\alpha\,(g^{W/2}B_a) - g^{W/2}\,\Gamma^n{}_{a\alpha}\,B_n \qquad \text{covariant rank-2 tensor} \tag{F.9.31}$$

$$(g^{W/2}B^a)_{;\alpha} = \partial_\alpha\,(g^{W/2}B^a) + g^{W/2}\,\Gamma^a{}_{\alpha n}\,B^n \qquad \text{mixed rank-2 tensor} \tag{F.9.32}$$

Example J=2: (covariant derivative of a tensor density B of weight W)

$$(g^{W/2}B_{ab})_{;\alpha} = \partial_\alpha\,(g^{W/2}B_{ab}) - \Gamma^n{}_{a\alpha}(g^{W/2}B_{nb}) - \Gamma^n{}_{b\alpha}(g^{W/2}B_{an}) \quad \text{covariant rank-3 tensor} \tag{F.9.33}$$

$$(g^{W/2}B^a{}_b)_{;\alpha} = \partial_\alpha\,(g^{W/2}B^a{}_b) + \Gamma^a{}_{\alpha n}\,(g^{W/2}B^n{}_b) - \Gamma^n{}_{b\alpha}(g^{W/2}B^a{}_n) \qquad \text{etc.} \tag{F.9.34}$$

$$(g^{W/2}B_a{}^b)_{;\alpha} = \partial_\alpha\,(g^{W/2}B_a{}^b) - \Gamma^n{}_{a\alpha}(g^{W/2}B_n{}^b) + \Gamma^b{}_{\alpha n}\,(g^{W/2}B_a{}^n) \tag{F.9.35}$$

$$(g^{W/2}B^{ab})_{;\alpha} = \partial_\alpha\,(g^{W/2}B^{ab}) + \Gamma^a{}_{\alpha n}\,(g^{W/2}B^{nb}) + \Gamma^b{}_{\alpha n}\,(g^{W/2}B^{an}) \tag{F.9.36}$$

F.10 The Leibniz rule for the covariant derivative of the product of two tensor densities

If A and B are arbitrary tensor densities each with an arbitrary set of up and down indices and arbitrary weight, then the *claim* of the Leibniz or product rule is this:

$$(A\text{----}B\text{----})_{;\alpha} \equiv A\text{----}{}_{;\alpha}\,B\text{----} + A\text{----}\;\;B\text{----}{}_{;\alpha} \qquad \text{// Weinberg p 105 (4.6.14)} \tag{F.10.1}$$

Recall from the Covariant Derivative Theorem (F.7.1) that A---- and $A\text{----}{}_{;\alpha}$ have the same weight, call it W_A. Similarly, B---- and $B\text{----}{}_{;\alpha}$ have the same weight W_B. According to the outer product rule (D.2.3), both terms on the RHS above have weight W_A+W_B and therefore this sum is the weight of the LHS $(A\text{----}B\text{----})_{;\alpha}$ as well.

Proof of (F.10.1): Start with the covariant derivative of a generic tensor A in the form of (F.8.8),

$$A\text{----}_{;\alpha} = A\text{----}_{,\alpha} + (\text{A index correction terms}) + W_A\, \Gamma^{\kappa}{}_{\kappa\alpha}\, A\text{----}$$

$$B\text{----}_{;\alpha} = B\text{----}_{,\alpha} + (\text{B index correction terms}) + W_B\, \Gamma^{\kappa}{}_{\kappa\alpha}\, B\text{----} \tag{F.10.2}$$

where --- is any number of up/down indices. The "index correction terms" are those Γ terms discussed below (F.8.9). One can then write out the two terms on the RHS of (F.10.1) above as:

$$A\text{----}_{;\alpha}\, B\text{----} = A\text{----}_{,\alpha}\, B\text{----} + (\text{A index correction terms})\, B\text{----} + [W_A\, \Gamma^{\kappa}{}_{\kappa\alpha}\, A\text{----}\,]\, B\text{----}$$

$$A\text{----}\; B\text{----}_{;\alpha} = A\text{----}\; B\text{----}_{,\alpha} + A\text{----}\, (\text{B index correction terms}) + A\text{----}\, [W_B\, \Gamma^{\kappa}{}_{\kappa\alpha}\, B\text{----}\,]\ . \tag{F.10.3}$$

Meanwhile, the LHS of (F.10.1) can be written as

$$(A\text{----}B\text{----})_{;\alpha} = (A\text{----}B\text{----})_{,\alpha} + (\text{ all index correction terms}) + (W_A + W_B)\, \Gamma^{\kappa}{}_{\kappa\alpha}\, (A\text{----}B\text{----})\ . \tag{F.10.4}$$

Momentarily ignoring the index correction terms, it is clear that the other terms match between LHS and RHS. The W terms match by visual inspection, while the regular derivative terms match due to the "regular" Leibniz rule for the derivative of a product

$$(A\text{----}B\text{----})_{,\alpha} = A\text{----}_{,\alpha}\, B\text{----} + A\text{----}\; B\text{----}_{,\alpha} \tag{F.10.5}$$

which is to say

$$\partial_\alpha(A\text{----}B\text{----}) = (\partial_\alpha A\text{----})B\text{----} + A\text{----}\; (\partial_\alpha B\text{----})\ . \tag{F.10.6}$$

Consider now the LHS terms called (all index correction terms) above. This set of terms can be partitioned into two groups,

$$(\text{all index correction terms}) = (\text{terms involving A indices}) + (\text{terms involving B indices})\ . \tag{F.10.7}$$

Let us pause to look at a simple example where ICT means we just show the index correction terms,

$$(A_{ab}B^{cd})_{;\alpha}|^{ICT} = -\,\Gamma^{n}{}_{a\alpha}(A_{nb}B^{cd}) - \Gamma^{n}{}_{b\alpha}(A_{an}B^{cd}) + \Gamma^{c}{}_{\alpha n}\,(A_{ab}B^{nd}) + \Gamma^{d}{}_{\alpha n}\,(A_{ab}B^{cn})\ . \tag{F.10.8}$$

The correction terms can be reordered in this way

$$= \{ -\Gamma^{n}{}_{a\alpha}(A_{nb}) - \Gamma^{n}{}_{b\alpha}(A_{an}) \}\, B^{cd} + A_{ab}\, \{ \Gamma^{c}{}_{\alpha n}B^{nd} + \Gamma^{d}{}_{\alpha n}B^{cn} \}$$

$$= \{ \text{index correction terms for } A_{ab} \}\, B^{cd} + A_{ab}\, \{ \text{index correction terms for } B^{cd} \}\ . \tag{F.10.9}$$

Just so, in the general case one has

$$(A\text{----}B\text{----})_{;\alpha}|^{ICT} = (\text{ all index correction terms})$$

$$= \{ \text{index correction terms for } A\text{----} \}\, B\text{----} + A\text{----}\, \{ \text{index correction terms for } B\text{----} \} \tag{F.10.10}$$

and this then shows that the index correction terms on the two sides of (F.10.1) do in fact match. **QED**

Once the above product rule is verified, it is then easy to generalize just as for regular derivatives:

$$(A\text{----}B\text{----}C\text{----})_{;\alpha} \equiv A\text{----}_{;\alpha}\ B\text{----}\ C\text{----} + A\text{----}\ B\text{----}_{;\alpha} C\text{----} + A\text{----}\ B\text{----}\ C\text{----}_{;\alpha} \qquad \text{(F.10.11)}$$

Examples with two vectors:

$$\begin{aligned}
(A_aB_b)_{;\alpha} &= A_{a;\alpha}B_b + A_aB_{b;\alpha} \\
(A^aB_b)_{;\alpha} &= A^a{}_{;\alpha}B_b + A^aB_{b;\alpha} \\
(A^aB^b)_{;\alpha} &= A^a{}_{;\alpha}B^b + A^aB^b{}_{;\alpha} \\
(A_aB^b)_{;\alpha} &= A_{a;\alpha}B^b + A_aB^b{}_{;\alpha} \qquad \text{(F.10.12)}
\end{aligned}$$

Example with a scalar function A and a vector B:

$$(AB_b)_{;\alpha} = A_{;\alpha}B_b + AB_{b;\alpha} = A_{,\alpha}B_b + AB_{b;\alpha} \quad // \ A_{;\alpha} = A_{,\alpha},\ \textit{(F.9.1)} \qquad \text{(F.10.13)}$$

A more general example:

$$(A_{ab}{}^cB_{de})_{;\alpha} \equiv A_{ab}{}^c{}_{;\alpha}\ B_{de} + A_{ab}{}^c\ B_{de;\alpha} \qquad \text{(F.10.14)}$$

Theorem: Any object X--- for which $(X\text{---})_{;\alpha} = 0$ can be "extracted" from a covariant derivative group which contains X---. That is to say (F.10.15)

$$(A\text{----}\ X\text{---}\ B\text{-----})_{;\alpha} = X\text{---}\ (A\text{----}\ B\text{-----})_{;\alpha}$$

Proof: By the Leibniz Rule (F.10.1),

$$(A\text{----}\ X\text{---}\ B\text{-----})_{;\alpha} = (A\text{----}\ B\text{-----})_{;\alpha}\ X\text{---} + (A\text{----}\ B\text{-----})\ (X\text{---})_{;\alpha} = X\text{---}\ (A\text{----}\ B\text{-----})_{;}$$

We shall now apply this theorem in a few examples.

Example 1: Any constant can be extracted from a group.

$$(\pi\ A\text{----}\ B\text{-----})_{;\alpha} = \pi\ (A\text{----}\ B\text{-----})_{;\alpha} \qquad // \text{ extract a constant} \qquad \text{(F.10.16)}$$

$$(\varepsilon^{ijk}\ A\text{----}\ B\text{-----})_{;\alpha} = \varepsilon^{ijk}\ (A\text{----}\ B\text{-----})_{;\alpha}\ // \ \varepsilon^{ijk} = \text{element of perm. tensor} = \pm 1 \text{ or } 0 \qquad \text{(F.10.17)}$$

Example 2: Since by (F.7.19) $g_{ab;\alpha} = 0$ one can always extract g_{ab} from a group. This applies to g^{ab} as well. And of course it applies to $g_a{}^b = \delta_a{}^b$ since $\delta_a{}^b$ is a constant.

$$(A\text{----}\ B\text{--}_a\text{--})_{;\alpha} = (g_{ab}\ A\text{----}\ B\text{--}^b\text{--})_{;\alpha} = g_{ab}(A\text{----}\ B\text{--}^b\text{--})_{;\alpha} \qquad \text{(F.10.18)}$$

This says that lowering (or raising) an index "commutes" with covariant differentiation -- one can lower an index ignoring the fact that :α is sitting there. But we already know this must be true because we know that the object $(A{-}{-}{-}{-}B{-}{-}{}^{b}{-}{-})_{;\alpha}$ is a true tensor, and g_{ab} can lower any index on a true tensor.

Fact : $(|g|^s)_{;\alpha} = 0$, s = real This is |g| raised to any real power s (s is not an index) (F.10.19)

<u>Proof</u>: Recall from (D.2.3a) that

$$|g'|^{-W/2} = J^{-W}\,|g|^{-W/2}\,. \qquad \text{// no R factors since scalar} \qquad \textit{(D.2.3a)}$$

Setting s = -W/2 gives

$$|g'|^s = J^{2s}\,|g|^s$$

so $|g'|^s$ transforms as a scalar density of weight W = -2s. Recall next the derivative rule for a scalar density of weight W,

$$B_{;\alpha} \;=\; \partial_\alpha B + W\,\Gamma^\kappa{}_{\kappa\alpha}\,B \qquad \textit{(F.9.23)}$$

which we apply with B = $|g|^s$ to find,

$$\begin{aligned}(|g|^s)_{;\alpha} &= \partial_\alpha|g|^s \;-2s\,\Gamma^\kappa{}_{\kappa\alpha}\,|g|^s \\ &= \partial_\alpha|g|^s \;-2s\,\{\,(1/2)(1/|g|)\partial_\alpha|g|\,\}\;|g|^s \qquad \text{// (F.4.2) for } \Gamma^\kappa{}_{\kappa\alpha} \\ &= s|g|^{s-1}\,\partial_\alpha|g| \;-s\,|g|^{s-1}\partial_\alpha|g| \\ &= 0\,. \qquad \text{QED}\end{aligned}$$

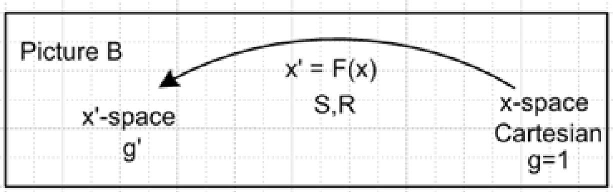
<u>Example 3:</u> Since by (F.10.19) $(|g|^s)_{;\alpha} = 0$, $|g|^s$ may be extracted from a covariant derivative group,

$$(\;|g|^s\;A{-}{-}{-}{-}\;B{-}{-}{-}{-}{-})_{;\alpha} = \;|g|^s\,(A{-}{-}{-}{-}\;B{-}{-}{-}{-}{-})_{;\alpha} \qquad \text{(F.10.20)}$$

Appendix G: Expansion of (∇v) in curvilinear coordinates (v = vector)

G.1 Continuum Mechanics motivation

This Appendix assumes the usual curvilinear coordinates context, Picture B



(G.1.1)

Although the polyadic notation is regarded as archaic by some writers (eg, Wolfram), it is well embedded into the literature of continuum mechanics, a field awash in tensors.

In the literature one sometimes sees, in Cartesian coordinates,

$$(\nabla \mathbf{A})_{ij} \equiv \partial_j A_i \ \equiv A_{i,j} \qquad \text{(G.1.2)}$$

where the indices are the reverse of the normal dyadic definition of (E.4.1),

$$(\mathbf{BA})_{ij} \equiv B_i A_j \ . \qquad \text{(G.1.3)}$$

For example, in continuum mechanics one encounters the so-called convective or material derivative of an arbitrary vector field **A(x**,t**)** in the Eulerian or spatial "view" of the motion of a blob of continuous matter,

$$DA_i/Dt \ = \partial_t A_i + \nabla A_i \bullet \mathbf{v} \ \ = \partial_t A_i + (\partial_j A_i)\, v_j = \ \partial_t A_i + (\nabla\mathbf{A})_{ij}\, v_j \ = \ \partial_t A_i + [(\nabla\mathbf{A})\ \mathbf{v}]_i$$

$$\Rightarrow \ \ D\mathbf{A}/Dt = \partial_t \mathbf{A} + (\nabla\mathbf{A})\ \mathbf{v} \qquad // = \partial_t \mathbf{A} + (\mathbf{v}\bullet\nabla)\mathbf{A} \ = (\partial_t + \mathbf{v}\bullet\nabla)\ \mathbf{A} \ \ = D/Dt\ (\mathbf{A}). \qquad \text{(G.1.4)}$$

DA_i/Dt is a historical notation for the total derivative $dA_i(\mathbf{x},t)/dt$. Here **v**(**x**,t) is the velocity field of the moving matter blob. An example is acceleration **a**, where **A** = **v**,

$$\mathbf{a} = D\mathbf{v}/Dt = \partial_t \mathbf{v} + (\nabla\mathbf{v})\ \mathbf{v} \qquad // = \partial_t \mathbf{v} + (\mathbf{v}\bullet\nabla)v \ = (\partial_t + \mathbf{v}\bullet\nabla)\ \mathbf{v} \ \ = D/Dt\ (\mathbf{v}). \qquad \text{(G.1.5)}$$

see Lai p 76 (3.4.3) and p 78 (3.4.8). The object (∇**v**), called the velocity gradient, is of great interest in fluid mechanics. Correspondingly, the object (∇**u**) is of great interest in the theory of elastic solids, where **u** is the displacement field.

The equation (G.1.4) written with proper index positions,

$$DA_i/Dt = \ \partial_t A_i + (\nabla\mathbf{A})_i{}^j\, v_j \qquad \textit{(G.1.4)}$$

becomes a true tensor equation if we make the assumption shown in the next Section that $(\nabla\mathbf{A})_i{}^j \equiv A_i{}^{;j}$. In this case, we know from the covariance theorem exemplified in (E.9.13) that, for orthogonal coordinates, this equation will appear in terms of $\hat{e}_n$-expanded coefficients (like $\mathscr{v}_j = h'_j v_j$) as follows:

$$D\mathcal{A}'_i/Dt = \partial_t\mathcal{A}'_i + (\nabla\mathcal{A})'_{ij}\,\mathscr{v}'_j \qquad \text{(G.1.6)}$$

where $(\nabla\mathcal{A})'_{ij} = [(\nabla\mathbf{A})^{(\hat{e})}]^{ij}$. To actually *use* equation (G.1.6) in some system of orthogonal curvilinear coordinates, one has to compute the coefficients $[(\nabla\mathbf{A})^{(\hat{e})}]^{ij}$. This is done below in Section G.5 generally and Section G.6 for a few specific coordinate systems (where **A** is called **v**) .

G.2 Expansion of ∇v on $e^i\otimes e^j$ by Method 1: Use the fact that $v_{b;a}$ is a tensor.

The covariant derivative $v_{b;a}$ is discussed in Sections F.7, F.8 and F.9. We shall define a true tensor object (∇**v**) as $v_{b;a}$ so that, according to (F.9.2),

$$\begin{array}{llll} (\nabla\mathbf{v})_{ba} \equiv v_{b;a} = [v_{b,a} - \Gamma^c{}_{ab}\,v_c] & // \; v_{b,a} \equiv \partial_a v_b & \text{x-space} & \\ (\nabla\mathbf{v})'_{ba} \equiv v'_{b;a} = [v'_{b,a} - \Gamma'^c{}_{ab}\,v'_c]\;. & // \; v'_{b,a} \equiv \partial'_a v'_b & \text{x'-space} & \text{(G.2.1)} \end{array}$$

If x-space is Cartesian, then Γ = 0 as in (F.4.16), in which case

$$(\nabla\mathbf{v})_{ba} = v_{b,a} = \partial_a v_b \qquad \text{(G.2.2)}$$

so $(\nabla\mathbf{v})_{ba} = v_{b;a}$ aligns with our dyadic (∇**v**) object (E.4.4) in Cartesian space.

As shown in (E.2.14) one can expand the rank-2 tensor $v_{b;a}$ in either of these ways,

$$\begin{array}{lll} \nabla\mathbf{v} = \Sigma_{ij}\,v_{i;j}\,\mathbf{u}^i\otimes\mathbf{u}^j & v_{i;j} = [v_{i,j} - \Gamma^c{}_{ij}\,v_c] = v_{i,j} = \partial_j v_i & \\ \nabla\mathbf{v} = \Sigma_{ij}\,v'_{i;j}\,\mathbf{e}^i\otimes\mathbf{e}^j & v'_{i;j} = [v'_{i,j} - \Gamma'^c{}_{ij}\,v'_c] & \text{(G.2.3)} \end{array}$$

where the $\mathbf{u}^i$ are the Cartesian basis vectors in x-space, while $\mathbf{e}^i$ are the reciprocal base vectors.

According to (F.5.6) with Fig (F.5.2), the affine connection Γ ' in x'-space is given by $\Gamma'^c{}_{ab} = \mathcal{R}'^c{}_n(\partial'_a\mathcal{R}'_b{}^n)$. When x-space is Cartesian, $\mathcal{R} = 1$ and then $\mathcal{R}' = R$ (F.5.3), so

$$\begin{array}{ll} \Gamma'^c{}_{ab} = R^c{}_i(\partial'_a R_b{}^i) & \\ \Gamma'^c{}_{ij} = R^c{}_k(\partial'_j R_i{}^k)\;. & \text{(G.2.4)} \end{array}$$

Inserting this into the second expansion of (G.2.3) gives

$$\nabla\mathbf{v} = \Sigma_{ij}\,v'_{i;j}\,\mathbf{e}^i\otimes\mathbf{e}^j \qquad v'_{i;j} = [(\partial'_j v'_i) - R^c{}_k(\partial'_j R_i{}^k)\,v'_c]$$

or

$$\nabla\mathbf{v} = \Sigma_{ij}\,[(\nabla\mathbf{v})^{(e)}]_{ij}\,\mathbf{e}^i\otimes\mathbf{e}^j \qquad [(\nabla\mathbf{v})^{(e)}]_{ij} = [(\partial'_j v'_i) - R^c{}_k(\partial'_j R_i{}^k)\,v'_c]\;. \qquad \text{(G.2.5)}$$

Note that the expression shown contains only x'-space coordinates and objects. The (e) superscript tells us that this matrix element of operator (∇v) is taken in the $\mathbf{e_n}$ basis, see Section E.7:

$$[(\nabla\mathbf{v})^{(e)}]_{ij} = \langle e_i|\nabla\mathbf{v}|\ e_j\rangle = \ (\mathbf{e_i})^T\ (\nabla\mathbf{v})\ \mathbf{e_j} = \mathbf{e_i} \bullet (\nabla\mathbf{v})\ \mathbf{e_j}\ = e_i \bullet (\nabla\mathbf{v}) \bullet \mathbf{e_j} \qquad \text{(G.2.6)}$$

bra-ket matrix dot of vectors dyadic

Recall now the third identity from (F.2.1),

$$(\partial_a R_d{}^n) = -\, R_e{}^n\, R_d{}^m\ (\partial_a R^e{}_m) \qquad 3\ . \qquad \textit{(F.2.1)}$$

This was derived for Picture C1 of (F.1.1) with x on the left and ξ on the right. Adjusting this to Picture B of (G.1.1) with x' on the left and x on the right gives the first line below, and then an index shuffle gives the second line,

$$(\partial'_a R_d{}^n) = -\, R_e{}^n\, R_d{}^m\ (\partial'_a R^e{}_m) \qquad // \ a\to j,\ d\to i,\ n\to k$$
$$(\partial'_j R_i{}^k) = -\, R_e{}^k\, R_i{}^m\ (\partial'_j R^e{}_m)\ . \qquad \text{(G.2.7)}$$

Then,

$$R^c{}_k(\partial'_j R_i{}^k) = -\, R^c{}_k\, R_e{}^k\, R_i{}^m\ (\partial'_j R^e{}_m)\ = -\, \delta^c{}_e\, R_i{}^m\ (\partial'_j R^e{}_m)\ = -R_i{}^m(\partial'_j R^c{}_m) \qquad \text{(G.2.8)}$$

so that (G.2.5) can be written in this way,

$$\nabla\mathbf{v} = \Sigma_{ij}\ [(\nabla\mathbf{v})^{(e)}]_{ij}\ \mathbf{e^i}\otimes\mathbf{e^j} \qquad [(\nabla\mathbf{v})^{(e)}]_{ij} = v'_{i;j} =\ [(\partial'_j v'_i) + R_i{}^m(\partial'_j R^c{}_m)\ v'_c]\ . \qquad \text{(G.2.9)}$$

We shall use this second form for $[(\nabla\mathbf{v})^{(e)}]_{ij}$ below.

Footnote: A variation of the above development would be to start this way, using the expansion of (E.2.12) on the left, and using (F.9.3) on the right,

$$\nabla\mathbf{v} = \Sigma_{ij}\ v^i{}_{;j}\ \mathbf{u_i}\otimes\mathbf{u^j} \qquad v^i{}_{;j}\ =\ [v^i{}_{,j} + \Gamma^i{}_{jc}\, v_c]\ = v^i{}_{,j} = \partial_j v^i$$
$$\nabla\mathbf{v} = \Sigma_{ij}\ v'^i{}_{;j}\ \mathbf{e_i}\otimes\mathbf{e^j} \qquad v'^i{}_{;j}\ = [v'^i{}_{,j} + \Gamma'^i{}_{jc}\, v'^c] \qquad \text{(G.2.10)}$$

Eq (G.2.4) says $\Gamma'^c{}_{ab}\ = R^c{}_i(\partial'_a R_b{}^i)$, but (F.1.10) gives the alternate form shown on the first line below, and then an index shuffle gives the second line,

$$\Gamma'^c{}_{ab}\ =\ -\, R_b{}^i\ (\partial'_a R^c{}_i) \qquad // \ i\to m,\ \text{then } c\to i,\ a\to j,\ b\to c$$
$$\Gamma'^i{}_{jc}\ =\ -\, R_c{}^m\ (\partial'_j R^i{}_m)\ .$$

Then (G.2.10) says

$$\nabla\mathbf{v} = \Sigma_{ij}\ [(\nabla\mathbf{v})^{(e)}]^i{}_j\ \mathbf{e_i}\otimes\mathbf{e^j} \qquad [(\nabla\mathbf{v})^{(e)}]^i{}_j = v'^i{}_{;j}\ =\ [(\partial'_j v'^i) -\ R_c{}^m\ (\partial'_j R^i{}_m)\ v'^c] \qquad \text{(G.2.11)}$$

which is similar to (G.2.5) but applies to the down-tilt index configuration.

G.3 Expansion of ∇v on $\mathbf{e}^i \otimes \mathbf{e}^j$ by Method 2: Use brute force.

Method 1 is perhaps elegant in that it makes use of tensor transformations and the affine connection Γ. But a simple brute force method is really just as simple and does not require knowledge of Γ and covariant differentiation. Instead of using the $\mathbf{e}^i \otimes \mathbf{e}^j$ notation for basis vectors, here we use the alternate notation $\mathbf{e}^i(\mathbf{e}^j)^T$ which works for rank-2 tensor expansions. Recall that $\mathbf{e}^i(\mathbf{e}^j)^T$ is an NxN matrix as illustrated in (E.5.3) for N = 2.

In this brute force method, start with the Cartesian-space expansion,

$$(\nabla \mathbf{v}) = \Sigma_{cd} (\nabla \mathbf{v})_{dc}\, \mathbf{u}^d \mathbf{u}^{cT} = \Sigma_{cd} (\partial_c v_d)\, \mathbf{u}^d \mathbf{u}^{cT} \,. \qquad \text{// note that } \mathbf{u}_d = \mathbf{u}^d \text{ in Cartesian space} \qquad (G.3.1)$$

This expansion is the second line of (E.2.14) using the matrix notation (E.5.3), $\mathbf{u}^d \otimes \mathbf{u}^c = \mathbf{u}^d \mathbf{u}^{cT}$.

To express things in x'-coordinates, first write

$$\partial_c v_d = (R^i{}_c \partial'_i)(R^j{}_d v'_j) = R^i{}_c \left[(\partial'_i R^j{}_d)\, v'_j + R^j{}_d\, (\partial'_i v'_j)\right] \,. \qquad (G.3.2)$$

The next step is to express the matrix $\mathbf{u}^d \mathbf{u}^{cT}$ as a linear combination of $\mathbf{e}^e \mathbf{e}^{fT}$. We start with the orthogonality rule # 1 of (7.6.4) .

$$\delta^n{}_i = R_b{}^n R^b{}_i \qquad \text{// implied sum on b, as usual} \qquad \textit{(7.6.4)}$$

We know that

$$(\mathbf{e}^n)_i = R^n{}_i \qquad \textit{(7.18.1)}$$
$$(\mathbf{u}^n)_i = \delta^n{}_i \qquad \textit{(7.18.3)}$$

so the above orthogonality rule says:

$$(\mathbf{u}^n)_i = R_b{}^n\, (\mathbf{e}^b)_i \qquad \Rightarrow \qquad \mathbf{u}^n = R_b{}^n\, \mathbf{e}^b$$
$$\text{or} \qquad \mathbf{u}^d = R_e{}^d\, \mathbf{e}^e \,. \qquad (G.3.3)$$

Taking the transpose of both sides (sum of column vectors to sum of row vectors),

$$\mathbf{u}^{Td} = R_e{}^d\, \mathbf{e}^{Te} \qquad \Rightarrow \qquad \mathbf{u}^{cT} = R_f{}^c\, \mathbf{e}^{fT} \,. \qquad (G.3.4)$$

Combining these last two results gives

$$\mathbf{u}^d \mathbf{u}^{cT} = R_e{}^d R_f{}^c\, \mathbf{e}^e\, \mathbf{e}^{fT} = R_e{}^d R_f{}^c\, \mathbf{e}^e\, \mathbf{e}^{fT} \,. \qquad (G.3.5)$$

The expansion (G.3.1) of (∇**v**) may then be written,

$$(\nabla \mathbf{v}) = (\nabla \mathbf{v})_{dc}\, \mathbf{u}_d \mathbf{u}_c^T \ = \ (\partial_c v_d)\, \mathbf{u}_d \mathbf{u}_c^T$$

$$= \ R^i{}_c\, [(\partial'_i R^j{}_d)\, v'_j + R^j{}_d\, (\partial'_i v'_j)] \ \{ R_e{}^d\, R_f{}^c\, \mathbf{e}^e\, \mathbf{e}^{fT} \} \qquad \text{// (G.3.2) and (G.3.5)}$$

$$= \ (R_f{}^c\, R^i{}_c)[R_e{}^d\, (\partial'_i R^j{}_d)\, v'_j + (R_e{}^d\, R^j{}_d)(\partial'_i v'_j)] \ \mathbf{e}^e\, \mathbf{e}^{fT}$$

$$= \ \delta_f{}^i\, [R_e{}^d\, (\partial'_i R^j{}_d)\, v'_j + \delta_e{}^j\, (\partial'_i v'_j)]\, \mathbf{e}^e\, \mathbf{e}^{fT} \qquad \text{// orthogonality twice}$$

$$= \ [R_e{}^d\, (\partial'_f R^j{}_d)\, v'_j + (\partial'_f v'_e)]\, \mathbf{e}^e\, \mathbf{e}^{fT}$$

$$= \ \Sigma_{ef}\, [(\nabla \mathbf{v})^{(e)}]_{ef}\ \mathbf{e}^e\, \mathbf{e}^{fT} \tag{G.3.6}$$

where

$$[(\nabla \mathbf{v})^{(e)}]_{ef} \ = \ (\partial'_f v'_e) + R_e{}^d\, (\partial'_f R^j{}_d)\, v'_j \tag{G.3.7}$$

or

$$[(\nabla \mathbf{v})^{(e)}]_{ij} \ = \ (\partial'_j v'_i) + R_i{}^m\, (\partial'_j R^j{}_m)\, v'_j \tag{G.3.8}$$

This agrees with (G.2.9) obtained by the previous method.

G.4 Expansion on $\mathbf{e}_i \otimes \mathbf{e}_j$ and $\hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j$

Reversing the index tilts in (G.2.5) [as justified in (E.2.12)] one gets

$$\nabla \mathbf{v} = \Sigma_{ij}\, v'^{i;j}\, \mathbf{e}_i \otimes \mathbf{e}_j$$

$$\text{where}\ \ v'^{i;j} = g'^{ia} g'^{jb}\, v'_{a;b} \ = g'^{ia} g'^{jb}\, [(\partial'_b v'_a) + R_a{}^m (\partial'_b R^c{}_m)\, v'_c] \quad . \ \ \text{// (G.2.9)} \tag{G.4.1}$$

Then since $\mathbf{e}_i = h'_i \hat{\mathbf{e}}_i$ one gets yet another expansion (more generally see Section E.8),

$$\nabla \mathbf{v} = \Sigma_{ij}\, (v'^{i;j}\, h'_i\, h'_j)\, \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j \ = \Sigma_{ij}\, [(\nabla \mathbf{v})^{(\hat{e})}]_{ij}\, \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j$$

$$\text{where} \quad [(\nabla \mathbf{v})^{(\hat{e})}]_{ij} = h'_i\, h'_j\, v'^{i;j} = \ h'_i\, h'_j\, g'^{ia} g'^{jb}\, [(\partial'_b v'_a) + R_a{}^m (\partial'_b R^c{}_m)\, v'_c]\ . \tag{G.4.2}$$

Here is a summary of results so far:

$$\nabla \mathbf{v} = \Sigma_{ij}\, [(\nabla \mathbf{v})^{(e)}]_{ij}\, \mathbf{e}^i \otimes \mathbf{e}^j \qquad [(\nabla \mathbf{v})^{(e)}]_{ij} = \ [(\partial'_j v'_i) \ + \ R_i{}^m (\partial'_j R^c{}_m)\, v'_c] \qquad \textit{(G.2.9)}$$

$$\nabla \mathbf{v} = \Sigma_{ij}\, [(\nabla \mathbf{v})^{(e)}]^i{}_j\, \mathbf{e}_i \otimes \mathbf{e}^j \qquad [(\nabla \mathbf{v})^{(e)}]^i{}_j = \ [(\partial'_j v'^i) - \ R_c{}^m\ (\partial'_j R^i{}_m)\, v'^c]\ . \qquad \textit{(G.2.11)}$$

$$\nabla \mathbf{v} = \Sigma_{ij}\, [(\nabla \mathbf{v})^{(e)}]^{ij}\, \mathbf{e}_i \otimes \mathbf{e}_j \qquad [(\nabla \mathbf{v})^{(e)}]^{ij} = g'^{ia} g'^{jb}\, [(\partial'_b v'_a) + R_a{}^m (\partial'_b R^c{}_m)\, v'_c] \qquad \textit{(G.4.1)}$$

$$\nabla \mathbf{v} = \Sigma_{ij}\, [(\nabla \mathbf{v})^{(\hat{e})}]^{ij}\, \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j \qquad [(\nabla \mathbf{v})^{(\hat{e})}]^{ij} \ = \ h'\, h'_j\ \ g'^{ia} g'^{jb}\, [(\partial'_b v'_a) + R_a{}^m (\partial'_b R^c{}_m)\, v'_c] \quad \textit{(G.4.2)}$$

(G.4.3)

One could replace $v'_a = g'_{ad}v'^d$ in any of the above results. For example, the last object becomes

$$[(\nabla\mathbf{v})^{(\hat{e})}]^{ij} = h'_i\, h'_j\, g'^{ia}g'^{jb}\, [(\partial'_b[g'_{ad}v'^d]) + R_a{}^m(\partial'_b R^c{}_m)\,(g'_{cd}v'^d)] \ . \tag{G.4.4}$$

Another choice is to use the $\mathcal{v}'^n$ components of $\mathbf{v}$ obtained when expanding $\mathbf{v}$ on the $\hat{\mathbf{e}}_n$,

$$\mathbf{v} = \Sigma_n v'^n\, \mathbf{e}_n = \Sigma_n v'^n\, (h'_n\hat{\mathbf{e}}_n) = \Sigma_n (h'_n\, v'^n)\, \hat{\mathbf{e}}_n = \Sigma_n \mathcal{v}'^n\ \hat{\mathbf{e}}_n \qquad \Rightarrow \qquad \mathcal{v}'^n \equiv h'_n\, v'^n \tag{G.4.5}$$

so one then replaces $v'^n = h'_n{}^{-1}\mathcal{v}'^n$. The same object above then becomes

$$[(\nabla\mathbf{v})^{(\hat{e})}]^{ij} = h'_i\, h'_j\, g'^{ia}g'^{jb}\, [(\partial'_b[g'_{ad}\, h'_d{}^{-1}\mathcal{v}'^d]) + R_a{}^m(\partial'_b R^c{}_m)\,(g'_{cd}\, h'_d{}^{-1}\mathcal{v}'^d)] \ . \tag{G.4.6}$$

As discussed in Section 14.7, the components $\mathcal{v}'^n$ are convenient since they all have the same dimensions. Moreover, when a specific curvilinear system is selected, one can dispense with the unpleasant font used in $\mathcal{v}'^n$ and just write $\mathcal{v}'^n = v_{x'(n)}$. For example, in spherical coordinates r,θ,φ :

$$\mathcal{v}'^1 = v_r \qquad \mathcal{v}'^2 = v_\theta \qquad \mathcal{v}'^3 = v_\varphi \qquad\qquad \mathbf{v} = \Sigma_n \mathcal{v}'^n\ \hat{\mathbf{e}}_n = v_r\hat{\mathbf{r}} + v_\theta\hat{\boldsymbol{\theta}} + v_\varphi\ \hat{\boldsymbol{\varphi}} \ . \tag{G.4.7}$$

In Chapter 14 the scripted variables have no primes because Picture (14.1.1) is being used instead of Picture B.

G.5 Orthogonal coordinate systems

In this case $g'_{ab} = h'_a{}^2\delta_{a,b}$ and $g'^{ab} = h'_a{}^{-2}\delta^{a,b}$ and things simplify. The four expansions (G.4.3) become (there is no change in the first two expansions)

$$\begin{aligned}
&\nabla\mathbf{v} = \Sigma_{ij}\, [(\nabla\mathbf{v})^{(e)}]_{ij}\ \mathbf{e}^i\otimes\mathbf{e}^j && [(\nabla\mathbf{v})^{(e)}]_{ij} = [(\partial'_j v'_i) + R_i{}^m(\partial'_j R^c{}_m)\, v'_c] \\
&\nabla\mathbf{v} = \Sigma_{ij}\, [(\nabla\mathbf{v})^{(e)}]^i{}_j\ \mathbf{e}_i\otimes\mathbf{e}^j && [(\nabla\mathbf{v})^{(e)}]^i{}_j = [(\partial'_j v'^i) - R_c{}^m\, (\partial'_j R^i{}_m)\, v'^c] \ . \\
&\nabla\mathbf{v} = \Sigma_{ij}\, [(\nabla\mathbf{v})^{(e)}]^{ij}\ \mathbf{e}_i\otimes\mathbf{e}_j && [(\nabla\mathbf{v})^{(e)}]^{ij} = h'_i{}^{-2}\, h'_j{}^{-2}\, [(\partial'_j v'_i) + R_i{}^m(\partial'_j R^c{}_m)\, v'_c] \\
&\nabla\mathbf{v} = \Sigma_{ij}\, [(\nabla\mathbf{v})^{(\hat{e})}]^{ij}\ \hat{\mathbf{e}}_i\otimes\hat{\mathbf{e}}_j && [(\nabla\mathbf{v})^{(\hat{e})}]^{ij} = h'_i{}^{-1}\, h'_j{}^{-1}\, [(\partial'_j v'_i) + R_i{}^m(\partial'_j R^c{}_m)\, v'_c]
\end{aligned} \tag{G.5.1}$$

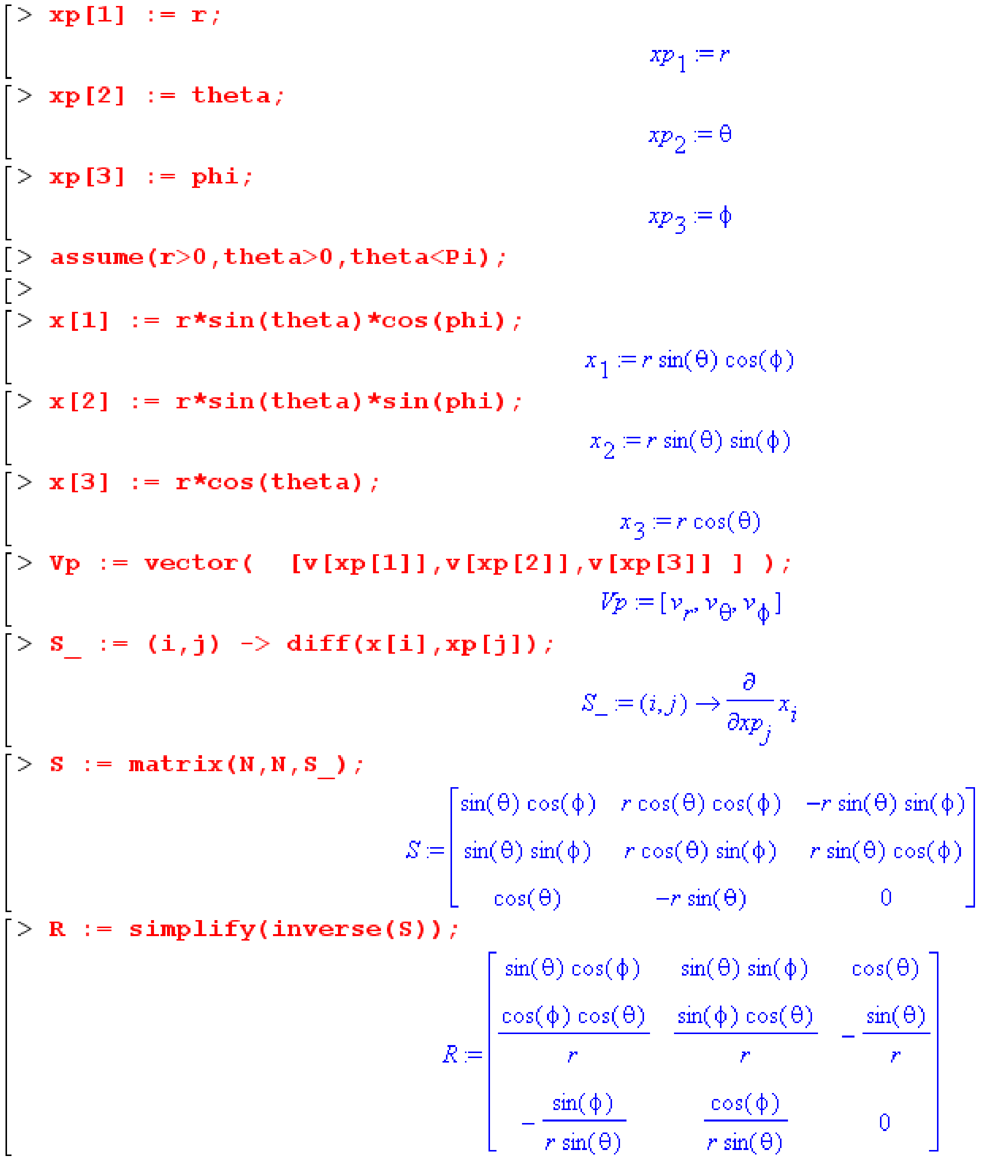
The (G.4.6) version of $[(\nabla\mathbf{v})^{(\hat{e})}]^{ij}$ becomes,

$$\begin{aligned}
P^{ij} \equiv [(\nabla\mathbf{v})^{(\hat{e})}]^{ij} &= h'_i{}^{-1}\, h'_j{}^{-1}\, [(\partial'_j[h'_i\mathcal{v}'^i]) + R_i{}^m(\partial'_j R^c{}_m)\, (h'_c\mathcal{v}'^c)] \\
&= h'_i{}^{-1}\, h'_j{}^{-1}\, [(\partial'_j[h'_i\mathcal{v}'^i]) + R_i{}^d(\partial'_j R^b{}_d)\, (h'_b\mathcal{v}'^b)] \qquad // \ m\to d, c\to b \\
&= h'_i{}^{-1}\, h'_j{}^{-1}\, [\ \underbrace{(\partial'_j h'_i)\ \mathcal{v}'^i}_{T2} + \underbrace{h'_i(\partial'_j\mathcal{v}'^i)}_{T3} + \underbrace{h'_i{}^2 R^i{}_d(\partial'_j R^b{}_d)\, (h'_b\mathcal{v}'^b)}_{T1}]
\end{aligned} \tag{G.5.2}$$

where $R_i{}^d = g'_{ia}R^a{}_b\, g^{bd} = h'_i{}^2 R^i{}_d$ is used to put all R's into their down-tilt form.

G.6 Maple evaluation of (∇v) in several coordinate systems

This object P^{ij} shown above can be computed in Maple by a simple program which is easily modified for other orthogonal curvilinear systems. The first task is to obtain the R matrix from the inverse transformation equations $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$, where we assume spherical coordinates as an example. In this code, N = 3 and p means prime, as in hp = h'.



```
> xp[1] := r;
```

$$xp_1 := r$$

```
> xp[2] := theta;
```

$$xp_2 := \theta$$

```
> xp[3] := phi;
```

$$xp_3 := \phi$$

```
> assume(r>0,theta>0,theta<Pi);
>
> x[1]  := r*sin(theta)*cos(phi);
```

$$x_1 := r\sin(\theta)\cos(\phi)$$

```
> x[2]  := r*sin(theta)*sin(phi);
```

$$x_2 := r\sin(\theta)\sin(\phi)$$

```
> x[3]  := r*cos(theta);
```

$$x_3 := r\cos(\theta)$$

```
> Vp  := vector(   [v[xp[1]],v[xp[2]],v[xp[3]] ] );
```

$$Vp := [v_r, v_\theta, v_\phi]$$

```
> S_  := (i,j) -> diff(x[i],xp[j]);
```

$$S\_ := (i,j) \to \frac{\partial}{\partial xp_j} x_i$$

```
> S  := matrix(N,N,S_);
```

$$S := \begin{bmatrix} \sin(\theta)\cos(\phi) & r\cos(\theta)\cos(\phi) & -r\sin(\theta)\sin(\phi) \\ \sin(\theta)\sin(\phi) & r\cos(\theta)\sin(\phi) & r\sin(\theta)\cos(\phi) \\ \cos(\theta) & -r\sin(\theta) & 0 \end{bmatrix}$$

```
> R  := simplify(inverse(S));
```

$$R := \begin{bmatrix} \sin(\theta)\cos(\phi) & \sin(\theta)\sin(\phi) & \cos(\theta) \\ \frac{\cos(\phi)\cos(\theta)}{r} & \frac{\sin(\phi)\cos(\theta)}{r} & -\frac{\sin(\theta)}{r} \\ -\frac{\sin(\phi)}{r\sin(\theta)} & \frac{\cos(\phi)}{r\sin(\theta)} & 0 \end{bmatrix}$$

(G.6.1)

Then one needs the scale factors $\mathrm{h_n}'$ from the metric tensor $\overline{\mathrm{g}}' = \mathrm{S}^{\mathrm{T}}\mathrm{S}$ in (5.7.9) [developmental notation],

```
gcov := simplify(evalm( transpose(S) &* S));
```

$$gcov := \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 - r^2\cos(\theta)^2 \end{bmatrix}$$

```
for n from 1 to N do hp[n] := simplify(sqrt(gcov[n,n])) od;
```

$$hp_1 := 1$$

$$hp_2 := r$$

$$hp_3 := r\sin(\theta) \qquad \text{(G.6.2)}$$

The terms T1,T2 and T3 shown in (G.5.2) are then entered,

$$P^{ij} \equiv [(\nabla\mathbf{v})^{(\hat{\mathbf{e}})}]^{ij} = {}_{\mathrm{pol}}h'^{-1}_{i}\, h'^{-1}_{j}\, [\ \underbrace{(\partial'_j h'_i)\, \boldsymbol{u}'^{i}}_{T2} + \underbrace{h'_i(\partial'_j \boldsymbol{u}'^{i})}_{T3} + \underbrace{h'^{2}_{i} R^{i}{}_{d}(\partial'_j R^{b}{}_{d})\,(h'_b \boldsymbol{u}'^{b})}_{T1}] \qquad (G.5.2)$$

```
> T1_ := (i,j) -> (hp[i])^2 * sum(
  sum(R[i,d]*Diff(R[b,d],xp[j])*hp[b]*Vp[b],d=1..N),b=1..N) ;
```

$$T1\_ := (i,j) \to hp_i^{\,2}\left(\sum_{b=1}^{N}\left(\sum_{d=1}^{N} R_{i,d}\left(\frac{\partial}{\partial xp_j} R_{b,d}\right) hp_b\, Vp_b\right)\right)$$

```
> T2_ := (i,j) -> Vp[i] *diff(hp[i],xp[j]) ;
```

$$T2\_ := (i,j) \to Vp_i\left(\frac{\partial}{\partial xp_j} hp_i\right)$$

```
> T3_ := (i,j) ->hp[i] * Diff(Vp[i],xp[j]) ;
```

$$T3\_ := (i,j) \to hp_i\left(\frac{\partial}{\partial xp_j} Vp_i\right) \qquad \text{(G.6.3)}$$

The terms are then added and simplified and out pops the result,

```
> P_ := (i,j) -> (1/hp[i])*(1/hp[j])*(value(T1_(i,j)) + (T2_(i,j)) + T3_(i,j));
```

$$P\_ := (i,j) \to \frac{\mathrm{value}(T1\_(i,j)) + T2\_(i,j) + T3\_(i,j)}{hp_i\, hp_j}$$

```
> P :=matrix(N,N):
> for n from 1 to N do
     for m from 1 to N do
      P[n,m] := expand(simplify(P_(n,m)));
     od;
  od;
> evalm(P);
>
```

$$\begin{bmatrix} \frac{\partial}{\partial r} v_r & -\frac{v_\theta}{r} + \frac{\frac{\partial}{\partial \theta} v_r}{r} & -\frac{v_\phi}{r} + \frac{\frac{\partial}{\partial \phi} v_r}{r\sin(\theta)} \\ \frac{\partial}{\partial r} v_\theta & \frac{v_r}{r} + \frac{\frac{\partial}{\partial \theta} v_\theta}{r} & -\frac{\cos(\theta)\, v_\phi}{r\sin(\theta)} + \frac{\frac{\partial}{\partial \phi} v_\theta}{r\sin(\theta)} \\ \frac{\partial}{\partial r} v_\phi & \frac{\frac{\partial}{\partial \theta} v_\phi}{r} & \frac{v_r}{r} + \frac{\cos(\theta)\, v_\theta}{r\sin(\theta)} + \frac{\frac{\partial}{\partial \phi} v_\phi}{r\sin(\theta)} \end{bmatrix} \qquad \text{(G.6.4)}$$

Below are some sample results (including the above):

$$(\nabla\mathbf{v}) = \Sigma_{ij}\ P^{ij}\ \hat{\mathbf{e}}_i\ \hat{\mathbf{e}}_j^{T} \qquad\qquad P^{ij} = [(\nabla\mathbf{v})^{(\hat{\mathbf{e}})}]^{ij} \tag{G.6.5}$$

- P^{ij} in **polar** coordinates (where 1,2 = r,θ) :

$$\begin{bmatrix} \frac{\partial}{\partial r} v_r & -\frac{v_\theta}{r} + \frac{\frac{\partial}{\partial \theta} v_r}{r} \\ \frac{\partial}{\partial r} v_\theta & \frac{v_r}{r} + \frac{\frac{\partial}{\partial \theta} v_\theta}{r} \end{bmatrix} \qquad \text{// agrees with Lai p 57 (2.33.23)} \tag{G.6.6}$$

- P^{ij} in **cylindrical** coordinates (where 1,2,3 = r,θ,z) :

$$\begin{bmatrix} \frac{\partial}{\partial r} v_r & -\frac{v_\theta}{r} + \frac{\frac{\partial}{\partial \theta} v_r}{r} & \frac{\partial}{\partial z} v_r \\ \frac{\partial}{\partial r} v_\theta & \frac{v_r}{r} + \frac{\frac{\partial}{\partial \theta} v_\theta}{r} & \frac{\partial}{\partial z} v_\theta \\ \frac{\partial}{\partial r} v_z & \frac{\frac{\partial}{\partial \theta} v_z}{r} & \frac{\partial}{\partial z} v_z \end{bmatrix} \qquad \text{// agrees with Lai p 60 (2.34.5)} \tag{G.6.7}$$

The polar coordinates results are seen to be the upper-left 2x2 piece of the cylindrical results.

- P^{ij} in **spherical** coordinates (where 1,2,3 = r,θ,φ) :

$$\begin{bmatrix} \frac{\partial}{\partial r} v_r & -\frac{v_\theta}{r} + \frac{\frac{\partial}{\partial \theta} v_r}{r} & -\frac{v_\phi}{r} + \frac{\frac{\partial}{\partial \phi} v_r}{r\sin(\theta)} \\ \frac{\partial}{\partial r} v_\theta & \frac{v_r}{r} + \frac{\frac{\partial}{\partial \theta} v_\theta}{r} & -\frac{\cos(\theta)\, v_\phi}{r\sin(\theta)} + \frac{\frac{\partial}{\partial \phi} v_\theta}{r\sin(\theta)} \\ \frac{\partial}{\partial r} v_\phi & \frac{\frac{\partial}{\partial \theta} v_\phi}{r} & \frac{v_r}{r} + \frac{\cos(\theta)\, v_\theta}{r\sin(\theta)} + \frac{\frac{\partial}{\partial \phi} v_\phi}{r\sin(\theta)} \end{bmatrix} \qquad \text{// agrees with Lai p 64 (2.35.25)} \tag{G.6.8}$$

By entering the usual inverse transformation equations $\mathbf{x}' = F^{-1}(\mathbf{x})$, and with suitable small alterations, the above Maple code can compute $\nabla\mathbf{v}$ in any orthogonal curvilinear coordinate system in any number of dimensions N.

Appendix H: Expansion of div(T) in curvilinear coordinates (T = rank-2 tensor)

H.1 Introduction

The object of attention in this Appendix, **divT**, is expressed this way in Cartesian coordinates,

$$(\mathrm{divT})^i = \partial_j T^{ij} \ , \tag{H.1.1}$$

where T^{ij} is a rank-2 tensor. One can regard the above equation as describing the normal divergence of the "vector" which forms the i^{th} row of the matrix T_{ij}. In general (that is to say, under general transformations F), the rows of T_{ij} are not tensorial vectors, and that is why $(\mathrm{divT})^i$ is not a tensorial scalar. Nor, in fact, are the $(\mathrm{divT})^i$ as defined above the components of a tensorial vector! As shown in Section H.3 below, **divT** will be redefined as a true tensorial vector which equals $\partial_j T^{ij}$ in Cartesian coordinates.

H.2 Continuum Mechanics motivation

The vector **divT** arises for example when Newton's 2nd Law **F** = m**a** is applied to a particle of continuous matter, in which case this law is known as Cauchy's Equation of Motion,

$$\mathbf{divT} + \rho\mathbf{B} = \rho\mathbf{a} \qquad // \text{ Lai p 169 (4.7.4)} \tag{H.2.1}$$

In this equation T is known as the Cauchy stress tensor, ρ is the mass density of the particle, **B** is any action-at-a-distance force (body force) per unit mass (such as gravity), and of course **a** is the acceleration of the particle.

Under rotations and translations the quantity $(\mathrm{divT})^i = \partial_j T^{ij}$ *is* a tensorial vector, ρ is a tensorial scalar, and **a** and **B** are tensorial vectors. As one would expect, **divT** + ρ**B** = ρ**a** is covariant in the sense of Section 7.15 under these kinds of transformations.

More generally, if $(\mathrm{divT})^a$ is taken to mean $T^{ab}{}_{;b}$ as shown in the next Section. then (H.2.1) is a true tensor equation under any transformation **x**' = **F**(**x**). Then as shown in (E.9.13), equation (H.2.1) stated in terms of $\hat{e}_n$-expanded coefficients takes the form

$$(\boldsymbol{div\mathcal{T}})' + \rho\boldsymbol{\mathcal{B}}' = \rho\boldsymbol{a}' \qquad (\boldsymbol{div\mathcal{T}})'^i = [(\mathrm{divT})^{(e)}]^i \tag{H.2.2}$$

where $\mathbf{e}_n$ are the tangent base vectors for some system of orthogonal curvilinear coordinates. One again, in order to use this equation, one must compute the coefficients $[(\mathrm{divT})^{(e)}]^i$ and this task is done below. The result is rather complicated, as shown in (H.5.7). For spherical coordinates, the three components of $(\boldsymbol{div\mathcal{T}})'^i = [(\mathrm{divT})^{(e)}]^i$ are displayed in (H.6.3), where each component is seen to have seven terms.

H.3 Expansion of divT on $\mathbf{e_n}$ by Method 1: Use fact that $T^{ab}{}_{;\alpha}$ is a tensor.

As shown in (F.9.9), the following object transforms as a rank-3 tensor,

$$\begin{aligned} &T^{ab}{}_{;\alpha} \equiv \partial_\alpha T^{ab} + \Gamma^a{}_{\alpha n} T^{nb} + \Gamma^b{}_{\alpha n} T^{an} = \partial_\alpha T^{ab} && \text{// x-space [g=1, } \Gamma = 0 \text{ (F.4.16)]} \\ &T'^{ab}{}_{;\alpha} \equiv \partial'_\alpha T'^{ab} + \Gamma'^a{}_{\alpha n} T'^{nb} + \Gamma'^b{}_{\alpha n} T'^{an} && \text{// x'-space} \qquad \text{(H.3.1)} \end{aligned}$$

Contracting b with α yields the following tensorial vector equations,

$$\begin{aligned} &(\mathrm{div}T)^a = T^{ab}{}_{;b} = \partial_b T^{ab} && \text{// x-space , } \Gamma = 0 \\ &(\mathrm{div}T)'^a = T'^{ab}{}_{;b} \equiv \partial'_b T'^{ab} + \Gamma'^a{}_{bn} T'^{nb} + \Gamma'^b{}_{bn} T'^{an} && \text{// x'-space} \qquad \text{(H.3.2)} \end{aligned}$$

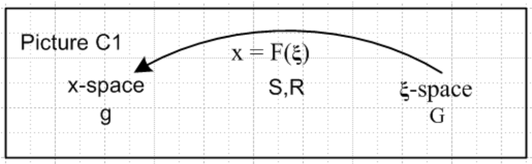
Note that the Cartesian space statement $(\mathrm{div}T)^a = \partial_b T^{ab}$ is obtained, as noted in the opening comments above. As in (E.2.10), the vector **divT** can be expanded as

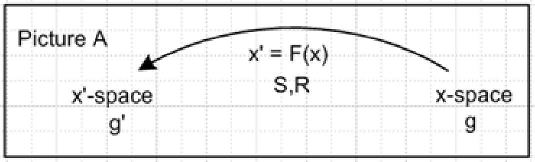
$$\mathbf{divT} = \Sigma_a (\mathrm{div}T)'^a \mathbf{e_a} \qquad (\mathrm{div}T)'^a = [(\mathrm{div}T)^{(e)}]^a \qquad \text{// } \mathbf{divT} \text{ in the } \mathbf{e_n} \text{ basis} \qquad \text{(H.3.3)}$$

Recall from (F.1.10) that

$$\begin{aligned} &\Gamma^c{}_{ab} = R^c{}_i \, (\partial_a R_b{}^i) && \text{// for Picture C1 in (F.1.1)} \\ &\Gamma^c{}_{ab} = - R_b{}^i \, (\partial_a R^c{}_i) \ . && \qquad \textit{(F.1.10)} \end{aligned}$$

Adjusting from Picture C1 to Picture A,





the last equation above becomes the first one below. We then shuffle indices to get the second line, and contract a with b to get the third:

$$\begin{aligned} &\Gamma'^c{}_{ab} = - R_b{}^i \, (\partial'_a R^c{}_i) && \text{// do b→n then a→b then c→a} \\ &\Gamma'^a{}_{bn} = - R_n{}^i \, (\partial'_b R^a{}_i) && \text{// next, contract indices a and b} \\ &\Gamma'^b{}_{bn} = - R_n{}^i \, (\partial'_b R^b{}_i) \ . && \qquad \text{(H.3.4)} \end{aligned}$$

Installing the last two lines into the second line of (H.3.2) then gives

$$\begin{aligned} (\mathrm{div}T)'^a &= \partial'_b T'^{ab} + \Gamma'^a{}_{bn} T'^{nb} + \Gamma'^b{}_{bn} T'^{an} \\ &= \partial'_b T'^{ab} - R_n{}^i \, (\partial'_b R^a{}_i) \, T'^{nb} - R_n{}^i \, (\partial'_b R^b{}_i) T'^{an} \ . \qquad \text{(H.3.5)} \end{aligned}$$

The conclusion is that

$$\mathbf{divT} = \Sigma_a[(\mathrm{divT})^{(e)}]^a \, \mathbf{e}_a \qquad \text{(H.3.6)}$$

$$[(\mathrm{divT})^{(e)}]^a = \partial'_b \, T'^{ab} - R_n{}^i(\partial'_b R^a{}_i) \, T'^{nb} - R_n{}^i \, (\partial'_b R^b{}_i) T'^{an} \qquad \text{// sum on n and b}$$

The vector $\mathbf{divT} = \Sigma_a(\partial_b T^{ab})\mathbf{u}_a$ has thus been expressed in terms of x'-space coordinates and objects, and as usual the $\mathbf{e}_n$ are the tangent base vectors in x-space.

H.4 Expansion of divT on $\mathbf{e}_n$ by Method 2: Use brute force.

Start with the known expansion of **divT** in Cartesian x-space, as in (E.2.10),

$$\mathbf{divT} = \Sigma_i \, [\mathrm{divT}]^i \, \mathbf{u}_i = \Sigma_i \, (\partial_j T^{ij}) \, \mathbf{u}_i \, . \qquad \text{(H.4.1)}$$

Compute

$$(\partial_j T^{ij}) = (R^a{}_j \partial'_a)(R_{b'}{}^i \, R_{c'}{}^j \, T'^{b'c'})$$

$$= R^a{}_j \, R_{b'}{}^i \, R_{c'}{}^j \, (\partial'_a \, T'^{b'c'}) + R^a{}_j \, R_{b'}{}^i \, (\partial'_a \, R_{c'}{}^j) \, T'^{b'c'} + R^a{}_j \, R_{c'}{}^j \, (\partial'_a \, R_{b'}{}^i) \, T'^{b'c'}$$

$$= (R^a{}_j \, R_{c'}{}^j) \, R_{b'}{}^i \, (\partial'_a \, T'^{b'c'}) + R^a{}_j \, R_{b'}{}^i \, (\partial'_a \, R_{c'}{}^j) \, T'^{b'c'} + (R^a{}_j \, R_{c'}{}^j) \, (\partial'_a \, R_{b'}{}^i) \, T'^{b'c'}$$

$$= \delta^a{}_{c'} \, R_{b'}{}^i \, (\partial'_a \, T'^{b'c'}) + R^a{}_j \, R_{b'}{}^i \, (\partial'_a \, R_{c'}{}^j) \, T'^{b'c'} + \delta^a{}_{c'} \, (\partial'_a \, R_{b'}{}^i) \, T'^{b'c'}$$

$$= R_{b'}{}^i \, (\partial'_a \, T'^{b'a}) + R^a{}_j \, R_{b'}{}^i \, (\partial'_a \, R_{c'}{}^j) \, T'^{b'c'} + (\partial'_a \, R_{b'}{}^i) \, T'^{b'a} \, . \qquad \text{(H.4.2)}$$

In (G.3.3) it was shown that $\mathbf{u}^n = \Sigma_b R_b{}^n \, \mathbf{e}^b$. This is valid with indices tilted the other way, as the reader can show mimicking the derivation of (G.3.3) or by raising and lowering labels using (7.18.1) $\mathbf{e}^n = g'^{ni} \, \mathbf{e}_i$ and (7.18.3) $\mathbf{u}^n = g^{ni} \, \mathbf{u}_i$ and (7.5.9) $R_a{}^b = g'_{aa'} R^{a'}{}_{b'} \, g^{b'b}$. The result is

$$\mathbf{u}_i = \Sigma_e \, R^n{}_i \, \mathbf{e}_n \, . \qquad \text{(H.4.3)}$$

Inserting (H.4.2) and (H.4.3) into (H.4.1) gives, twice using orthogonality rules (7.6.4),

$$\mathbf{divT} = \Sigma_i \, (\partial_j T^{ij}) \, \mathbf{u}_i$$

$$= \{ (R^n{}_i \, R_{b'}{}^i) \, (\partial'_a \, T'^{b'a}) + R^a{}_j \, (R^n{}_i \, R_{b'}{}^i) \, (\partial'_a \, R_{c'}{}^j) \, T'^{b'c'} + R^n{}_i \, (\partial'_a \, R_{b'}{}^i) \, T'^{b'a} \} \, \mathbf{e}_n$$

$$= \{ \delta^n{}_{b'} \, (\partial'_a \, T'^{b'a}) + R^a{}_j \, \delta^n{}_{b'} \, (\partial'_a \, R_{c'}{}^j) \, T'^{b'c'} + R^n{}_i \, (\partial'_a \, R_{b'}{}^i) \, T'^{b'a} \} \, \mathbf{e}_n$$

$$= \{ (\partial'_a \, T'^{na}) + R^a{}_j \, (\partial'_a \, R_{c'}{}^j) \, T'^{nc'} + R^n{}_i \, (\partial'_a \, R_{b'}{}^i) \, T'^{b'a} \} \, \mathbf{e}_n$$

$$= \Sigma_n[(\mathrm{divT})^{(e)}]^n \, \mathbf{e}_n \qquad \text{(H.4.4)}$$

where

$$[(\mathrm{div}T)^{(e)}]^n = (\partial'_a T'^{na}) \ + \ R^a{}_j\,(\partial'_a R_{c'}{}^j)\, T'^{nc'} \ + \ R^n{}_i\,(\partial'_a R_{b'}{}^i)\, T'^{b'a} \qquad . \qquad \text{(H.4.5)}$$

The third identity in (F.2.1) says

$$(\partial_a R_d{}^n) = -\, R_e{}^n\, R_d{}^m\, (\partial_a R^e{}_m) \qquad 3\ . \qquad \text{// for Picture C1 of (F.1.1)} \qquad \textit{(F.2.1)}$$

Adjusted to Picture C1' of (F.1.4) this becomes the first line below, then the next two lines are obtained by index shuffles,

$$\begin{aligned}
&(\partial'_a R_d{}^n) \ = -\, R_e{}^n\, R_d{}^m \ \ (\partial'_a R^e{}_m) && \text{// do d→c', n→j} \\
&(\partial'_a R_{c'}{}^j) = -\, R_e{}^j\, R_{c'}{}^m\, (\partial'_a R^e{}_m) && \text{// do j → i} \\
&(\partial'_a R_{b'}{}^i) = -\, R_e{}^i\, R_{b'}{}^m\, (\partial'_a R^e{}_m)\ . && \qquad\qquad\qquad \text{(H.4.6)}
\end{aligned}$$

Inserting these last two lines into (H.4.5) gives,

$$\begin{aligned}
[(\mathrm{div}T)^{(e)}]^n &= (\partial'_a T'^{na}) \ + \ R^a{}_j\,(\partial'_a R_{c'}{}^j)\, T'^{nc'} \ + \ R^n{}_i\,(\partial'_a R_{b'}{}^i)\, T'^{b'a} \qquad \textit{(H.4.5)} \\
&= (\partial'_a T'^{na}) \ - \ R^a{}_j\, R_e{}^j\, R_{c'}{}^m\, (\partial'_a R^e{}_m)\, T'^{nc'} \ - \ R^n{}_i\, R_e{}^i\, R_{b'}{}^m\, (\partial'_a R^e{}_m)\, T'^{b'a} \\
&= (\partial'_a T'^{na}) \ - \ \delta^a{}_e R_{c'}{}^m\, (\partial'_a R^e{}_m)\, T'^{nc'} \ - \ \delta^n{}_e\, R_{b'}{}^m\, (\partial'_a R^e{}_m)\, T'^{b'a} \qquad \text{// orthog.} \\
&= (\partial'_a T'^{na}) \ - \ R_{c'}{}^m\, (\partial'_a R^a{}_m)\, T'^{nc'} \ - \ R_{b'}{}^m\, (\partial'_a R^n{}_m)\, T'^{b'a}\ . \qquad \text{(H.4.7)}
\end{aligned}$$

Reverse the order of the last two terms,

$$[(\mathrm{div}T)^{(e)}]^n = (\partial'_a T'^{na}) \ - \ R_{b'}{}^m\, (\partial'_a R^n{}_m)\, T'^{b'a} \ - \ R_{c'}{}^m\, (\partial'_a R^a{}_m)\, T'^{nc'} \ . \qquad \text{(H.4.8)}$$

Replace summation indices m→i, a→b,

$$= (\partial'_b T'^{nb}) \ - \ R_{b'}{}^i\, (\partial'_b R^n{}_i)\, T'^{b'b} \ - \ R_{c'}{}^i\, (\partial'_b R^b{}_i)\, T'^{nc'}\ . \qquad \text{(H.4.9)}$$

Finally, change n→a and then c'→ n and b'→n,

$$[(\mathrm{div}T)^{(e)}]^a \ = (\partial'_b T'^{ab}) \ - \ R_n{}^i\ (\partial'_b R^a{}_i)\, T'^{nb} \ - \ R_n{}^i\ (\partial'_b R^b{}_i)\, T'^{an}\ . \quad \text{// sum on n and b} \quad \text{(H.4.10)}$$

This is seen to match the result (H.3.6) of the previous Section. The brute force method is in fact not too bad and requires no explicit use of the affine connection Γ.

<u>Technical Note:</u> In the above expression one can write for example $R_n{}^i(\partial'_b R^a{}_i) = g'_{nm} R^m{}_i(\partial'_b R^a{}_i)$. Then using the fact that $R^m{}_i R^a{}_i = g'^{ma}$ one finds $R^m{}_i(\partial'_b R^a{}_i) + R^a{}_i(\partial'_b R^m{}_i) = (\partial'_b g'^{ma})$ which then allows the replacement $R_n{}^i(\partial'_b R^a{}_i) = g'_{nm}\,(\partial'_b\, g'^{ma}) - g'_{nm}\, R^a{}_i(\partial'_b R^m{}_i)$. This kind of transformation leads to an alternate form for **divT**, still with all down-tilt R matrices, but the index structure is different. This other form appears when one does the "brute force" method starting with $(\partial_j T_{ij})$ instead of $(\partial_j T^{ij})$. Of course in Cartesian space these two objects must be the same.

H.5 Adjustment for T expanded on $(\hat{e}_i\otimes\hat{e}_j)$ and divT expanded on $\hat{e}_a$

In the above, it has been assumed that T^{ij} is a rank-2 tensor so that, in the notation of (E.2.11),

$$T = \Sigma_{ij}T^{ij}(\mathbf{u}_i\otimes\mathbf{u}_j) = \Sigma_{ij}T'^{ij}(\mathbf{e}_i\otimes\mathbf{e}_j) \quad . \tag{H.5.1}$$

If one is interested in an expansion of T on the unit vectors $\hat{\mathbf{e}}_i$ where $\mathbf{e}_i = h'_i\,\hat{\mathbf{e}}_i$ this becomes

$$T = \Sigma_{ij}[T'^{ij}h'_ih'_j]\,(\hat{\mathbf{e}}_i\otimes\hat{\mathbf{e}}_j) = \Sigma_{ij}[T^{(\hat{e})}]^{ij}\,(\hat{\mathbf{e}}_i\otimes\hat{\mathbf{e}}_j) \tag{H.5.2}$$

One then has

$$[T^{(\hat{e})}]^{ij} = h'_ih'_jT'^{ij} \qquad \Rightarrow\ T'^{ij} = h'_i{}^{-1}\,h'_j{}^{-1}\,[T^{(e\hat{})}]^{ij} \quad . \tag{H.5.3}$$

If one is interested in this form of the T matrix elements, then one is likely also interested in this expansion for **divT**,

$$\mathbf{divT} = \Sigma_a[(\mathrm{div}T)^{(e)}]^a\,\mathbf{e}_a = \Sigma_n\ \{\,[(\mathrm{div}T)^{(e)}]^a\,h'_a\,\}\ \ \hat{\mathbf{e}}_a\ \equiv\ [(\mathrm{div}T)^{(\hat{e})}]^a\,\hat{\mathbf{e}}_a \tag{H.5.4}$$

where then

$$[(\mathrm{div}T)^{(\hat{e})}]^a = h'_a[(\mathrm{div}T)^{(e)}]^a$$

$$= h'_a\,[(\partial'_b\,T'^{ab}) - R_n{}^i\,(\partial'_bR^a{}_i)\,T'^{nb} - R_n{}^i\,(\partial'_bR^b{}_i)\,T'^{an}] \qquad // \text{ from (H.4.10)}$$

$$= h'_a * \ \{\,\partial'_b\,(h'_a{}^{-1}\,h'_b{}^{-1}\,[T^{(\hat{e})}]^{ab}) - R_n{}^i\,(\partial'_bR^a{}_i)\,h'_n{}^{-1}\,h'_b{}^{-1}\,[T^{(\hat{e})}]^{nb}$$
$$- R_n{}^i\,(\partial'_bR^b{}_i)\,h'_a{}^{-1}\,h'_n{}^{-1}\,[T^{(\hat{e})}]^{an}\ \} \quad // \text{ from (H.5.3)}$$

$$= h'_a * \ \{\,\partial'_b\,(h'_a{}^{-1}\,h'_b{}^{-1})\,[T^{(\hat{e})}]^{ab} - R_n{}^i\,(\partial'_bR^a{}_i)\,h'_n{}^{-1}\,h'_b{}^{-1}\,[T^{(\hat{e})}]^{nb} \qquad // \ \partial(xy) = dx\ y + x\ dy$$
$$+ h'_a{}^{-1}\,h'_b{}^{-1}\,(\partial'_b\,[T^{(\hat{e})}]^{ab}) - R_n{}^i\,(\partial'_bR^b{}_i)\,h'_a{}^{-1}\,h'_n{}^{-1}\,[T^{(\hat{e})}]^{an}\ \} \quad . \tag{H.5.5}$$

Now

$$\partial'_b\,(h'_a{}^{-1}\,h'_b{}^{-1}) = \partial'_b(h'_a\,h'_b)^{-1} = -\,(h'_a\,h'_b)^{-2}\,\partial'_b(h'_a\,h'_b) = -\,h'_a{}^{-2}\,h'_b{}^{-2}\,\partial'_b(h'_a\,h'_b)$$

so the above sequence for $[(\mathrm{div}T)^{(\hat{e})}]^a$ continues,

$$= h'_a * \ \{-\,h'_a{}^{-2}\,h'_b{}^{-2}\,\partial'_b(h'_a\,h'_b)\,[T^{(\hat{e})}]^{ab} - R_n{}^i\,(\partial'_bR^a{}_i)\,h'_n{}^{-1}\,h'_b{}^{-1}\,[T^{(\hat{e})}]^{nb}$$
$$+ h'_a{}^{-1}\,h'_b{}^{-1}\,(\partial'_b\,[T^{(\hat{e})}]^{ab}) \qquad - R_n{}^i\,(\partial'_bR^b{}_i)\,h'_a{}^{-1}\,h'_n{}^{-1}\,[T^{(\hat{e})}]^{an}\ \}$$

$$= \ \{-\,h'_a{}^{-1}\,h'_b{}^{-2}\,\partial'_b(h'_a\,h'_b)\,[T^{(\hat{e})}]^{ab} - h'_a\,h'_n{}^{-1}\,h'_b{}^{-1}\,g'_{nm}\,R^m{}_i\,(\partial'_bR^a{}_i)\,[T^{(\hat{e})}]^{nb}$$
$$+ h'_b{}^{-1}\,(\partial'_b\,[T^{(\hat{e})}]^{ab}) \qquad - h'_n{}^{-1}\,g'_{nm}\,R^m{}_i\,(\partial'_bR^b{}_i)\,[T^{(\hat{e})}]^{an}\ \} \tag{H.5.6}$$

where recall that $\mathrm{R^n{}_i} = \mathrm{R^{ni}}$ since x-space is Cartesian. In the last form only the down-tilt R matrix appears which simplifies calculation with Maple. This and all other results above are valid for general curvilinear coordinates, orthogonal as well as non-orthogonal.

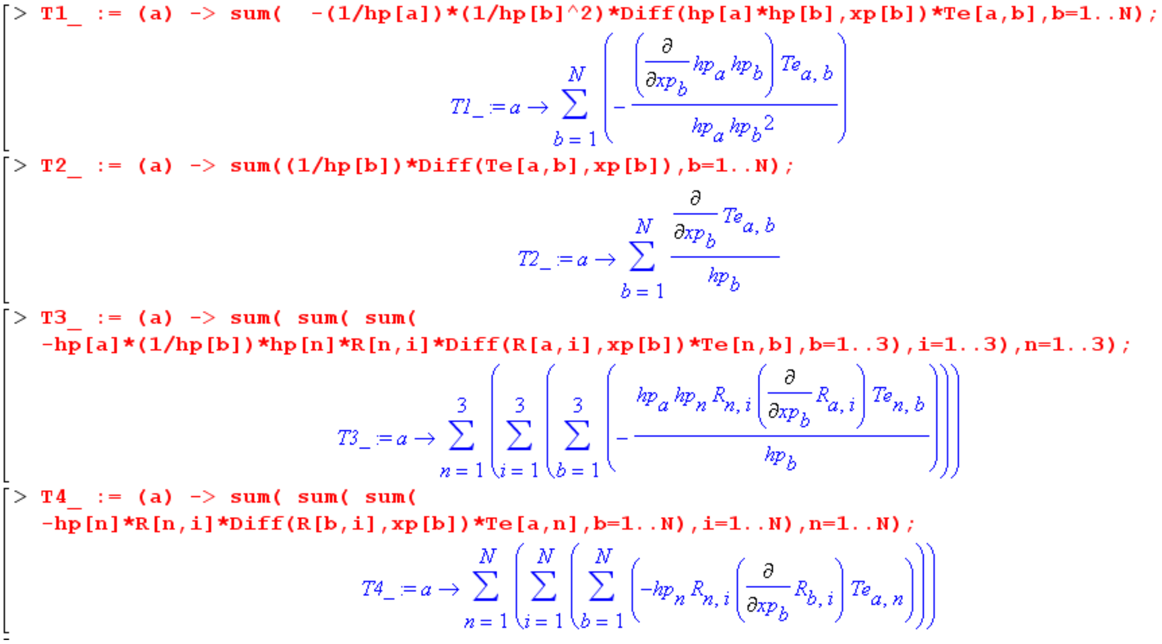
At this point we specialize to **orthogonal** systems, so (H.5.6) becomes

$$\begin{aligned}[(\mathrm{divT})^{(\hat{\mathrm{e}})}]^{\mathrm{a}} = \{ & \overset{\mathrm{T1}}{-\,h'_{\mathrm{a}}{}^{-1}\, h'_{\mathrm{b}}{}^{-2}\, \partial'_{\mathrm{b}}(h'_{\mathrm{a}}\, h'_{\mathrm{b}})\, [\mathrm{T}^{(\hat{\mathrm{e}})}]^{\mathrm{ab}}} \quad \overset{\mathrm{T3}}{-\, h'_{\mathrm{a}}\, h'_{\mathrm{b}}{}^{-1}\, h'_{\mathrm{n}}\, \mathrm{R^n{}_i}\, (\partial'_{\mathrm{b}} \mathrm{R^a{}_i})\, [\mathrm{T}^{(\hat{\mathrm{e}})}]^{\mathrm{nb}}} \\ & \underset{\mathrm{T2}}{+\, h'_{\mathrm{b}}{}^{-1}\, (\partial'_{\mathrm{b}}\, [\mathrm{T}^{(\hat{\mathrm{e}})}]^{\mathrm{ab}})} \quad \underset{\mathrm{T4}}{-\, h'_{\mathrm{n}}\, \mathrm{R^n{}_i}\, (\partial'_{\mathrm{b}} \mathrm{R^b{}_i})\, [\mathrm{T}^{(\hat{\mathrm{e}})}]^{\mathrm{an}}} \ \} \end{aligned} \qquad \text{(H.5.7)}$$

Even for an orthogonal curvilinear coordinate system, the form of this tensor divergence is amazingly complicated. Here it is expressed in terms of the down-tilt R matrix and the scale factors, and as usual repeated indices are summed.

H.6 Maple: divT in cylindrical and spherical coordinates

The above object $[(\mathrm{divT})^{(\hat{\mathrm{e}})}]^{\mathrm{a}}$ can be evaluated by Maple code very similar to that shown in Appendix G for (**∇v**). The main difference is the set of entry lines for the terms,



```
> T1_ := (a) -> sum(  -(1/hp[a])*(1/hp[b]^2)*Diff(hp[a]*hp[b],xp[b])*Te[a,b],b=1..N);
```

$$T1\_ := a \to \sum_{b=1}^{N} \left(- \frac{\left(\frac{\partial}{\partial xp_b} hp_a\, hp_b \right) Te_{a,b}}{hp_a\, hp_b{}^2} \right)$$

```
> T2_ := (a) -> sum((1/hp[b])*Diff(Te[a,b],xp[b]),b=1..N);
```

$$T2\_ := a \to \sum_{b=1}^{N} \frac{\frac{\partial}{\partial xp_b} Te_{a,b}}{hp_b}$$

```
> T3_ := (a) -> sum( sum( sum(
  -hp[a]*(1/hp[b])*hp[n]*R[n,i]*Diff(R[a,i],xp[b])*Te[n,b],b=1..3),i=1..3),n=1..3);
```

$$T3\_ := a \to \sum_{n=1}^{3} \left(\sum_{i=1}^{3} \left(\sum_{b=1}^{3} \left(- \frac{hp_a\, hp_n\, R_{n,i} \left(\frac{\partial}{\partial xp_b} R_{a,i} \right) Te_{n,b}}{hp_b} \right) \right) \right)$$

```
> T4_ := (a) -> sum( sum( sum(
  -hp[n]*R[n,i]*Diff(R[b,i],xp[b])*Te[a,n],b=1..N),i=1..N),n=1..N);
```

$$T4\_ := a \to \sum_{n=1}^{N} \left(\sum_{i=1}^{N} \left(\sum_{b=1}^{N} \left(-hp_n\, R_{n,i} \left(\frac{\partial}{\partial xp_b} R_{b,i} \right) Te_{a,n} \right) \right) \right)$$

(H.6.1)

Here are some results for $[(\mathrm{divT})^{(\hat{\mathbf{e}})}]^{\mathbf{a}}$ = "$\mathrm{divT}_{\mathbf{a}}$" :

- **cylindrical** coordinates (where 1,2,3 = r,θ,z) :

$$divT_1 := \left(\frac{\partial}{\partial r} T_{r,r}\right) + \frac{\frac{\partial}{\partial \theta} T_{r,\theta}}{r} + \left(\frac{\partial}{\partial z} T_{r,z}\right) - \frac{T_{\theta,\theta}}{r} + \frac{T_{r,r}}{r}$$

$$divT_2 := \frac{T_{\theta,r}}{r} + \left(\frac{\partial}{\partial r} T_{\theta,r}\right) + \frac{\frac{\partial}{\partial \theta} T_{\theta,\theta}}{r} + \left(\frac{\partial}{\partial z} T_{\theta,z}\right) + \frac{T_{r,\theta}}{r}$$

$$divT_3 := \left(\frac{\partial}{\partial r} T_{z,r}\right) + \frac{\frac{\partial}{\partial \theta} T_{z,\theta}}{r} + \left(\frac{\partial}{\partial z} T_{z,z}\right) + \frac{T_{z,r}}{r}$$

// agrees with Lai p 60 (2.34.8,9,10)

(H.6.2)

- For **polar** coordinates $(\mathrm{divT})_{\mathbf{r},\boldsymbol{\theta}}$ and are given by the first two lines above with the $\partial_{\mathbf{z}}$ terms set to 0. These polar results then agree with Lai p 58 (2.33.32,33).

- **spherical** coordinates (where 1,2,3 = r,θ,φ) :

$$divT_1 := \left(\frac{\partial}{\partial r} T_{r,r}\right) + \frac{\frac{\partial}{\partial \theta} T_{r,\theta}}{r} + \frac{\frac{\partial}{\partial \phi} T_{r,\phi}}{r\sin(\theta)} - \frac{T_{\theta,\theta}}{r} - \frac{T_{\phi,\phi}}{r} + 2\frac{T_{r,r}}{r} + \frac{\cos(\theta)\, T_{r,\theta}}{r\sin(\theta)}$$

$$divT_2 := \frac{\frac{\partial}{\partial \phi} T_{\theta,\phi}}{r\sin(\theta)} + \frac{\frac{\partial}{\partial \theta} T_{\theta,\theta}}{r} + 2\frac{T_{\theta,r}}{r} + \frac{\cos(\theta)\, T_{\theta,\theta}}{r\sin(\theta)} + \left(\frac{\partial}{\partial r} T_{\theta,r}\right) - \frac{\cos(\theta)\, T_{\phi,\phi}}{r\sin(\theta)} + \frac{T_{r,\theta}}{r}$$

$$divT_3 := 2\frac{T_{\phi,r}}{r} + \frac{\cos(\theta)\, T_{\theta,\phi}}{r\sin(\theta)} + \left(\frac{\partial}{\partial r} T_{\phi,r}\right) + \frac{\frac{\partial}{\partial \phi} T_{\phi,\phi}}{r\sin(\theta)} + \frac{\frac{\partial}{\partial \theta} T_{\phi,\theta}}{r} + \frac{\cos(\theta)\, T_{\phi,\theta}}{r\sin(\theta)} + \frac{T_{r,\phi}}{r}$$

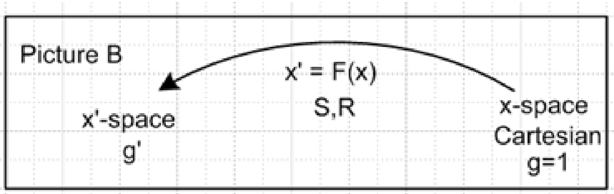
(H.6.3)

The expressions above agree with Lai p 65 (2.35.33,34,35).

Appendix I : The Vector Laplacian in Spherical and Cylindrical Coordinates

I.1 Introduction

This appendix assumes the usual curvilinear coordinates context, Picture B,



. (I.1.1)

In this Section the vector Laplacian is computed in two different ways, each associated with a particular "tensorization" of its Cartesian form. The second method, though less pleasant than the first, gives insight into why the vector Laplacian always includes the scalar Laplacian of the field components. It is in this inclusive form that results are usually stated in the literature.

In passing, it should be noted that the vector Laplacian is not just an idle mathematical curiosity. It shows up for example in the wave equations for electric and magnetic fields in a vacuum,

$$(\nabla^2 + k^2)\mathbf{E(x)} = \mathbf{0} \qquad (\nabla^2 + k^2)\mathbf{B(x)} = \mathbf{0} \qquad k = \omega/c \qquad \mathbf{E,B}(\mathbf{x},t) = \mathbf{E,B}(\mathbf{x})e^{-i\omega t} \qquad \text{(I.1.2)}$$

In continuum mechanics, it appears for example in the Navier/Cauchy equation which describes the small vector displacement field **u** in an isotropic elastic solid,

$$\rho_o\partial_t^2\mathbf{u} = \rho_0\mathbf{B} + (\lambda+\mu)\nabla e + \mu\,\nabla^2\mathbf{u} \qquad e = \text{div}\,\mathbf{u} = \text{dilatation} \qquad \text{// Lai p 216 (5.6.9)} \qquad \text{(I.1.3)}$$

or

$$\rho_o\partial_t^2\mathbf{u} = \rho_0\mathbf{B} + (\lambda+\mu)\,(\textbf{grad div u}) + \mu\,\nabla^2\mathbf{u}$$

Here ρ_0 is the unperturbed mass density, **B** the body force, and λ and μ are the two Lamé constants which describe an isotropic elastic medium. The vector Laplacian makes another appearance in the better-known Navier-Stokes equation which describes the vector velocity field **v** in an incompressible Newtonian fluid,

$$\rho\,[\,\partial_t\mathbf{v} + (\nabla\mathbf{v})\mathbf{v}] = \rho\mathbf{B} - \nabla p + \mu\nabla^2\mathbf{v} \qquad \text{// Lai p 361 (6.7.6)} \qquad \text{(I.1.4)}$$

where ρ is the mass density and p is pressure.

As in previous Appendices, for orthogonal curvilinear coordinates and based on (E.9.13) we can express the above equations in primed scripted $\hat{\mathbf{e}}_n$-expanded coefficients,

$$(\nabla^2\boldsymbol{\mathcal{B}})' + k^2\boldsymbol{\mathcal{B}}' = 0 \qquad \text{// vector equation (Helmholtz)}$$
$$(\nabla^2\boldsymbol{\mathcal{B}})'^i + k^2\boldsymbol{\mathcal{B}}'^i = 0 \qquad \text{// components} \qquad \text{(I.1.5)}$$

$$\rho_o\partial_t^2\boldsymbol{u}' = \rho_0\boldsymbol{\mathcal{B}}' + (\lambda+\mu)\,(\textit{grad div u})' + \mu\,(\nabla^2\boldsymbol{u})' \qquad \text{// vector equation (Navier/Cauchy)}$$
$$\rho_o\partial_t^2\boldsymbol{u}'^i = \rho_0\boldsymbol{\mathcal{B}}'^i + (\lambda+\mu)\,(\textit{grad div u})'^i + \mu\,(\nabla^2\boldsymbol{u})'^i \qquad \text{// components} \qquad \text{(I.1.6)}$$

$\rho\, [\, \partial_t \boldsymbol{u}' + (\nabla \boldsymbol{u})' \boldsymbol{u}'] = \rho \boldsymbol{\mathcal{B}}' - (\nabla \boldsymbol{p})' + \mu(\nabla^2 \boldsymbol{u})'$ // vector equation (Navier-Stokes)

$\rho\, [\, \partial_t \boldsymbol{u}'^{i} + (\nabla \boldsymbol{u})'^{i}{}_{j} \boldsymbol{u}'^{j}] = \rho \boldsymbol{\mathcal{B}}'^{i} - (\nabla \boldsymbol{p})'^{i} + \mu(\nabla^2 \boldsymbol{u})'^{i}$ // components (I.1.7)

In our non-script $^{(\hat{e})}$ notation the objects appearing in the above equations may be written

$\boldsymbol{u}'^{i} = [u^{(\hat{e})}]^{i} = h'_{i} u'^{i}$ // similarly for other vectors

$(\nabla \boldsymbol{p})'^{i} = [(\nabla p)^{(\hat{e})}]^{i} = h'_{i}\, \partial'^{i} p'(x')$ where $p'(\mathbf{x}') = p(\mathbf{x})$ = scalar field, see (10.1.13)

$(\mathit{grad\ div\ u})'^{i} = h'_{n}\, \partial'^{n} \{ (1/\sqrt{g'}\,)\, \partial'_{i} (\sqrt{g'}\, \boldsymbol{u}'^{i}/h'_{i}) \}$ // see **G** in (13.1.17)

$(\nabla^2 \boldsymbol{\mathcal{B}})'^{i} = [(\nabla^2 \mathbf{B})^{(\hat{e})}]^{i} = [(\text{✡}B)^{(\hat{e})}]^{i} = h'_{i}\, [(\text{✡}B)^{(e)}]^{i} = h'_{i}\, (\text{✡}B)'^{i} = (\text{✡}\boldsymbol{\mathcal{B}})'^{i}$

= treated in this Appendix. See (I.2.10) for $(\text{✡}B)'^{i}$. (I.1.8)

As noted in (E.8.5), the primes on scripted variables correspond to the primes on the names of the curvilinear coordinates which are x'^{i} for Picture B. In Picture M&S of (14.1.1) the primes on scripted variables go away, but here we are using Picture B.

I.2 Method 1 : a review

In (13.3.4) it is shown that, in Cartesian coordinates,

$\nabla^2(B_n) = [\text{ grad(div } \mathbf{B}) - \text{ curl(curl } \mathbf{B})]_n$ $\qquad \nabla^2 = \partial_j \partial^j$

or

$\nabla \bullet \nabla\, (B_n) = [\nabla(\nabla \bullet \mathbf{B}) - \nabla \times (\nabla \times \mathbf{B})]_n$. (I.2.1)

When all components are considered in a single equation, one could write

$\nabla^2(\mathbf{B}) = \text{grad(div } \mathbf{B}) - \text{ curl(curl } \mathbf{B})$

or

$\nabla \bullet \nabla\, (\mathbf{B}) = \nabla(\nabla \bullet \mathbf{B}) - \nabla \times (\nabla \times \mathbf{B})$ (I.2.2)

and this is frequently done (see examples cited in the previous Section). Since $\nabla^2(\mathbf{B}) \neq \partial_j \partial^j\ \mathbf{B}$ in curvilinear coordinates, it seems notionally safer in our current context to use a different symbol for the vector Laplacian operator, and following Moon and Spencer we use ✡ so the second last equation above then says:

$\text{✡}\mathbf{B} = \text{grad(div } \mathbf{B}) - \text{ curl (curl } \mathbf{B})$. (I.2.3)

Since $[\text{✡}\mathbf{B}]_n$ agrees with $\nabla^2(B_n)$ in Cartesian coordinates, and since we know how to write div, grad and curl in curvilinear coordinates (Chapters 9,10,12 or Chapter 15), the right hand side of the above equation provides our "first method" of writing ✡**B** in curvilinear coordinates. In (15.7.8) it is shown that the proper "tensorization" of the above equation is given by

$$(\text{✡}B)^{n} = (B^{j}{}_{;j})^{;n} - g^{-1/2}\varepsilon^{nab}(g^{-1/2}\varepsilon_{bde}B^{e;d})_{;a} \qquad \text{✡}\mathbf{B} = (\text{✡}B)^{n}\,\mathbf{u}_{n} \tag{I.2.4}$$

and therefore, in x'-space (**x**' are the curvilinear coordinates of interest),

$$(\text{✡}B)'^{n} = (B'^{j}{}_{;j})^{;n} - g'^{-1/2}\varepsilon'^{nab}(g'^{-1/2}\varepsilon'_{bde}B'^{e;d})_{;a} \qquad \text{✡}\mathbf{B} = (\text{✡}B)'^{n}\,\mathbf{e}_{n} \tag{I.2.5}$$

where $(B'^{j}{}_{;j})$ = [div **B**]' = [div **B**] (a scalar). The object ✡**B** is a normal vector (weight 0), assuming **B** is a normal vector. Doing various simplifying steps, we then arrive at the following expression (15.7.18) which lends itself to calculation:

$$\begin{aligned}(\text{✡}B)'^{n} &= \partial'^{n}\{(1/\sqrt{g'}\,)\,\partial'_{i}(\sqrt{g'}\,B'^{i})\} \qquad\qquad \text{✡}\mathbf{B} = (\text{✡}B)'^{n}\,\mathbf{e}_{n}\\ &\quad - (1/\sqrt{g'}\,)\,\varepsilon'^{ncd}\,\varepsilon'^{eab}\,\partial'_{c}\{\,(1/\sqrt{g'}\,)\,g'_{de}(\partial'_{a}[g'_{bf}B'^{f}])\,\}\end{aligned} \tag{I.2.6}$$

where ε'^{ncd} is the usual permutation tensor ($\varepsilon'^{ncd} = \varepsilon^{ncd}$). In this notation, B'^{i} is an official x'-space contravariant vector component.

We are often interested in working with vectors which are expanded on unit vector versions of the tangent base vectors. In such a unit vector expansion of a vector, a script font has been used for the components. One has,

$$B'^{n} = \mathcal{B}'^{n}/h'_{n} \qquad \mathbf{e}_{n} = h'_{n}\,\hat{\mathbf{e}}_{n} \qquad \mathbf{B} = B'_{n}\mathbf{e}_{n} = \mathcal{B}'^{n}\hat{\mathbf{e}}_{n} \qquad h'_{n} = \text{scale factor} \tag{I.2.7}$$

which leads to this rewrite of (I.2.6),

$$\begin{aligned}(\text{✡}B)'^{n} &= \partial'^{n}\{(1/\sqrt{g'}\,)\,\partial'_{i}(\sqrt{g'}\,\mathcal{B}'^{i}/h'_{i})\} \qquad\qquad \text{✡}\mathbf{B} = (\text{✡}B)'^{n}\,\mathbf{e}_{n}\\ &\quad - (1/\sqrt{g'}\,)\,\varepsilon'^{ncd}\,\varepsilon'^{eab}\,\partial'_{c}\{\,(1/\sqrt{g'}\,)\,g'_{de}(\partial'_{a}[g'_{bf}\mathcal{B}'^{f}/h'_{f}])\,\}\end{aligned} \tag{I.2.8}$$

or, as we will use it with unit vectors,

$$\begin{aligned}(\text{✡}\mathcal{B})'^{n} &= h'_{n}\,\partial'^{n}\{(1/\sqrt{g'}\,)\,\partial'_{i}(\sqrt{g'}\,\mathcal{B}'^{i}/h'_{i})\} \qquad\qquad \text{✡}\mathbf{B} = (\text{✡}\mathcal{B})'^{n}\,\hat{\mathbf{e}}_{n}\\ &\quad - h'_{n}\,(1/\sqrt{g'}\,)\,\varepsilon'^{ncd}\,\varepsilon'^{eab}\,\partial'_{c}\{\,(1/\sqrt{g'}\,)\,g'_{de}(\partial'_{a}[g'_{bf}\mathcal{B}'^{f}/h'_{f}])\,\} \quad .\end{aligned} \tag{I.2.9}$$

Specializing to orthogonal coordinates yields the form we shall use below for computation in Maple,

$$\begin{aligned}(\text{✡}\mathcal{B})'^{n} &= (1/h'_{n})\,\partial'_{n}\overset{T1}{\{(1/\sqrt{g'}\,)}\,\partial'_{i}(\sqrt{g'}\,\mathcal{B}'^{i}/h'_{i})\} \qquad\qquad \text{✡}\mathbf{B} = (\text{✡}\mathcal{B})'^{n}\,\hat{\mathbf{e}}_{n}\\ &\quad - (h'_{n}/\sqrt{g'}\,)\,\underset{T2}{\varepsilon'^{ncd}\,\varepsilon'^{dab}\,\partial'_{c}}\{\,(1/\sqrt{g'}\,)\,h'^{2}_{d}(\partial'_{a}[h'_{b}\mathcal{B}'^{b}])\,\} \quad .\end{aligned} \tag{I.2.10}$$

The second term could be written as two terms by reducing the product $\varepsilon'^{ncd}\,\varepsilon'^{dab}$ into δδ - δδ in the usual manner, but Maple is happy to just "do it" as stated. And for non-orthogonal coordinates, the reduction of $\varepsilon'^{ncd}\,\varepsilon'^{eab}$ is much less friendly (see (D.10.18)) and then one would want to use the εε product as is.

I.3 Method 1 for spherical coordinates: Maple speaks

In this Section, the following notation is used:

$$\mathcal{B}'^1 = B_r \qquad \mathcal{B}'^2 = B_\theta \qquad \mathcal{B}'^3 = B_\varphi$$

$$\mathbf{B} = \mathcal{B}'^n \hat{\mathbf{e}}_n = B_r \hat{\mathbf{e}}_r + B_\theta \hat{\mathbf{e}}_\theta + B_\varphi \hat{\mathbf{e}}_\varphi = B_r \hat{\mathbf{r}} + B_\theta \hat{\boldsymbol{\theta}} + B_\varphi \hat{\boldsymbol{\varphi}} \ . \tag{I.3.1}$$

Maple begins in the same manner as shown earlier in (G.6.1), the idea being that this code could be modified for any coordinate system,

```
N := 3;
```

$$N := 3$$

```
xp[1] := r;
```

$$xp_1 := r$$

```
xp[2] := theta;
```

$$xp_2 := \theta$$

```
xp[3] := phi;
```

$$xp_3 := \phi$$

```
assume(r>0,theta>0,theta<Pi);
x[1] := r*sin(theta)*cos(phi);
```

$$x_1 := r \sin(\theta) \cos(\phi)$$

```
x[2] := r*sin(theta)*sin(phi);
```

$$x_2 := r \sin(\theta) \sin(\phi)$$

```
x[3] := r*cos(theta);
```

$$x_3 := r \cos(\theta)$$

```
Bp := vector(  [B[xp[1]],B[xp[2]],B[xp[3]] ] );
```

$$Bp := [B_r, B_\theta, B_\phi]$$

```
S_ := (i,j) -> diff(x[i],xp[j]);
```

$$S\_ := (i, j) \rightarrow \frac{\partial}{\partial xp_j} x_i$$

```
S := matrix(N,N,S_);
```

$$S := \begin{bmatrix} \sin(\theta)\cos(\phi) & r\cos(\theta)\cos(\phi) & -r\sin(\theta)\sin(\phi) \\ \sin(\theta)\sin(\phi) & r\cos(\theta)\sin(\phi) & r\sin(\theta)\cos(\phi) \\ \cos(\theta) & -r\sin(\theta) & 0 \end{bmatrix}$$

```
gcov := simplify(evalm( transpose(S) &* S));
```

$$gcov := \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 - r^2 \cos(\theta)^2 \end{bmatrix}$$

```
for m from 1 to N do hp[m] := simplify(sqrt(gcov[m,m])) od;
```

$$hp_1 := 1$$

$$hp_2 := r$$

$$hp_3 := r \sin(\theta)$$

```
sq_gp := hp[1]*hp[2]*hp[3];
```

$$sq\_gp := r^2 \sin(\theta) \tag{I.3.2}$$

The last object is $\sqrt{g'}$, where hp[n] $= h'_n$. The next chunk of code computes the *scalar* Laplacian of an unspecified function f, and this expression will be used below in parsing the vector Laplacian results. The Laplacian expression comes from (11.6) with $g'^{nm} = (1/h'^2_n)\ \delta_{n,m}$:

Here is the formula to be used for computing the regular Laplacian:

$$[\text{lap f}](x) = [1/\sqrt{g'(x')}\ ]\ \partial'_n [\ \sqrt{g'(x')}\, h'_n{}^{-2}(x')\ (\partial'_n f\,'(x'))\ ]$$

```
lapf := (1/sq_gp)*sum(Diff(sq_gp *
hp[n]^(-2)*Diff(f(xp[1],xp[2],xp[3]),xp[n]),xp[n]),n=1..N);
```

$$lapf := \frac{\left(\frac{\partial}{\partial r} r^2 \sin(\theta) \left(\frac{\partial}{\partial r} f(r, \theta, \phi)\right)\right) + \left(\frac{\partial}{\partial \theta} \sin(\theta) \left(\frac{\partial}{\partial \theta} f(r, \theta, \phi)\right)\right) + \left(\frac{\partial}{\partial \phi} \frac{\frac{\partial}{\partial \phi} f(r, \theta, \phi)}{\sin(\theta)}\right)}{r^2 \sin(\theta)}$$

```
subs(f(xp[1],xp[2],xp[3])=op(0,f(xp[1],xp[2],xp[3])),value(expand(%)));
```

$$2 \frac{\frac{\partial}{\partial r} f}{r} + \left(\frac{\partial^2}{\partial r^2} f\right) + \frac{\cos(\theta) \left(\frac{\partial}{\partial \theta} f\right)}{r^2 \sin(\theta)} + \frac{\frac{\partial^2}{\partial \theta^2} f}{r^2} + \frac{\frac{\partial^2}{\partial \phi^2} f}{r^2 \sin(\theta)^2} \tag{I.3.3}$$

The subs command used here (and more intensely below) removes the arguments of the function f for purely cosmetic reasons (see our Maple User's Guide for details, operands section). The arguments are added in the first place to prevent Maple from thinking the unspecified function f is a constant.

Next comes a low-budget implementation of the permutation tensor eps(a,b,c) = ε^{abc},

```
eps := proc(a,b,c)
if type(a,numeric) and type(b,numeric) and type(c,numeric) then
    if   (a=1 and b=2 and c=3) then 1;
    elif (a=1 and b=3 and c=2) then -1;
    elif (a=2 and b=1 and c=3) then -1;
    elif (a=2 and b=3 and c=1) then 1;
    elif (a=3 and b=1 and c=2) then 1;
    elif (a=3 and b=2 and c=1) then -1;
    else 0;
    fi;
  else 'eps(a,b,c)';
  fi;
end:
```

(I.3.4)

The two terms of $(✡\mathcal{B})'^{n}$ in (I.2.10) are then duly entered, along with their sum "starB" ,

$$(✡\mathcal{B})'^{n} = T1^{n} + T2^{n}$$

$$T1^{n} = (1/h'_{n})\, \partial'_{n}\{ (1/\sqrt{g'})\, \partial'_{i}\, (\sqrt{g'}\, \mathcal{B}'^{i}/h'_{i})\}$$

$$T2^{n} = -\ (h'_{n}/\sqrt{g'})\, \varepsilon'^{ncd}\, \varepsilon'^{dab}\, \partial'_{c}\{ (1/\sqrt{g'})\, h'_{d}{}^{2}\, (\partial'_{a}\, [h'_{b}\, \mathcal{B}'^{b}]) \}$$

```
T1 := (n) -> (1/hp[n])*sum(Diff((1/sq_gp)*
Diff(sq_gp*Bp[i](xp[1],xp[2],xp[3])/hp[i],xp[i]),xp[n]),i=1..N);
```

$$T1 := n \to \frac{\displaystyle\sum_{i=1}^{N} \left(\frac{\partial}{\partial xp_{n}} \frac{\left(\dfrac{\frac{\partial}{\partial xp_{i}}\, sq\_gp\; Bp_{i}(xp_{1}, xp_{2}, xp_{3})}{hp_{i}}\right)}{sq\_gp} \right)}{hp_{n}}$$

```
T2 := (n) -> -(hp[n]/sq_gp)*sum(sum(sum(sum(eps(n,c,d)*eps(d,a,b) *Diff((1/sq_gp) *
hp[d]^2*Diff(hp[b]*Bp[b](xp[1],xp[2],xp[3]),xp[a]),xp[c]),a=1..N),b=1..N),c=1..N),d=1..N);
```

$$T2 := n \to -\frac{hp_{n} \left(\displaystyle\sum_{d=1}^{N} \left(\sum_{c=1}^{N} \left(\sum_{b=1}^{N} \left(\sum_{a=1}^{N} \operatorname{eps}(n, c, d)\, \operatorname{eps}(d, a, b) \left(\frac{\partial}{\partial xp_{c}} \frac{hp_{d}{}^{2} \left(\frac{\partial}{\partial xp_{a}}\, hp_{b}\, Bp_{b}(xp_{1}, xp_{2}, xp_{3}) \right)}{sq\_gp} \right) \right) \right) \right) \right)}{sq\_gp}$$

```
starB:= (n) -> T1(n) + T2(n):
```

(I.3.5)

In the following code, we use this notation, similar to (I.3.1) for vector **B**,

$$q1\_ \ = \ (✡\mathcal{B})'^{1} \ = \ (✡B)_{r}$$
$$q2\_ \ = \ (✡\mathcal{B})'^{2} \ = \ (✡B)_{\theta}$$
$$q3\_ \ = \ (✡\mathcal{B})'^{3} \ = \ (✡B)_{\varphi} \qquad (I.3.6)$$

and in this notation Maple computes the components of the vector Laplacian :

```
q1 := expand(value(starB(1))):
'subs(Bp[1](xp[1],xp[2],xp[3]) = op(0,Bp[1](xp[1],xp[2],xp[3])),q1)':
'subs(Bp[2](xp[1],xp[2],xp[3]) = op(0,Bp[2](xp[1],xp[2],xp[3])),%)':
 q1_ := 'subs(Bp[3](xp[1],xp[2],xp[3]) = op(0,Bp[3](xp[1],xp[2],xp[3])),%)':
q1_;
```

$$-2\frac{B_r}{r^2}+2\frac{\frac{\partial}{\partial r}B_r}{r}+\left(\frac{\partial^2}{\partial r^2}B_r\right)-2\frac{\cos(\theta)B_\theta}{r^2\sin(\theta)}-2\frac{\frac{\partial}{\partial\theta}B_\theta}{r^2}-2\frac{\frac{\partial}{\partial\phi}B_\phi}{r^2\sin(\theta)}+\frac{\frac{\partial^2}{\partial\phi^2}B_r}{r^2\sin(\theta)^2}+\frac{\cos(\theta)\left(\frac{\partial}{\partial\theta}B_r\right)}{r^2\sin(\theta)}+\frac{\frac{\partial^2}{\partial\theta^2}B_r}{r^2}$$

```
q2 := expand(value(starB(2))):
'subs(Bp[1](xp[1],xp[2],xp[3]) = op(0,Bp[1](xp[1],xp[2],xp[3])),q2)':
'subs(Bp[2](xp[1],xp[2],xp[3]) = op(0,Bp[2](xp[1],xp[2],xp[3])),%)':
 q2_ := 'subs(Bp[3](xp[1],xp[2],xp[3]) = op(0,Bp[3](xp[1],xp[2],xp[3])),%)':
q2_;
```

$$2\frac{\frac{\partial}{\partial\theta}B_r}{r^2}-\frac{\cos(\theta)^2B_\theta}{r^2\sin(\theta)^2}+\frac{\cos(\theta)\left(\frac{\partial}{\partial\theta}B_\theta\right)}{r^2\sin(\theta)}-\frac{B_\theta}{r^2}+\frac{\frac{\partial^2}{\partial\theta^2}B_\theta}{r^2}-2\frac{\cos(\theta)\left(\frac{\partial}{\partial\phi}B_\phi\right)}{r^2\sin(\theta)^2}+\frac{\frac{\partial^2}{\partial\phi^2}B_\theta}{r^2\sin(\theta)^2}+2\frac{\frac{\partial}{\partial r}B_\theta}{r}+\left(\frac{\partial^2}{\partial r^2}B_\theta\right)$$

```
q3 := expand(value(starB(3))):
'subs(Bp[1](xp[1],xp[2],xp[3]) = op(0,Bp[1](xp[1],xp[2],xp[3])),q3)':
'subs(Bp[2](xp[1],xp[2],xp[3]) = op(0,Bp[2](xp[1],xp[2],xp[3])),%)':
 q3_ := 'subs(Bp[3](xp[1],xp[2],xp[3]) = op(0,Bp[3](xp[1],xp[2],xp[3])),%)':
q3_;
```

$$2\frac{\frac{\partial}{\partial\phi}B_r}{\sin(\theta)r^2}+2\frac{\cos(\theta)\left(\frac{\partial}{\partial\phi}B_\theta\right)}{\sin(\theta)^2r^2}+\frac{\frac{\partial^2}{\partial\phi^2}B_\phi}{r^2\sin(\theta)^2}-\frac{\cos(\theta)^2B_\phi}{r^2\sin(\theta)^2}+\frac{\cos(\theta)\left(\frac{\partial}{\partial\theta}B_\phi\right)}{r^2\sin(\theta)}-\frac{B_\phi}{r^2}+\frac{\frac{\partial^2}{\partial\theta^2}B_\phi}{r^2}+2\frac{\frac{\partial}{\partial r}B_\phi}{r}+\left(\frac{\partial^2}{\partial r^2}B_\phi\right) \qquad (I.3.7)$$

See comment below (I.3.3) about the Maple subs commands (purely cosmetic).

I.4 Method 1 for spherical coordinates: putting results in traditional form

It turns out that in each component of the vector Laplacian stated above, 5 of the 9 terms can be represented as if they were the scalar Laplacian acting on the component in question. Here is how it works, where we now use the well-known spherical coordinates Laplacian (which Maple has computed using (11.12) with 1,2,3 = r,θ.φ and $h_1 = 1$, $h_2 = r$, $h_3 = r\sin\theta$ as shown above in (I.3.1)),

$$\nabla^2 f = \underset{1+2}{(1/r^2)\partial_r(r^2\partial_r f)} + \underset{3+4}{(1/r^2\sin\theta)\partial_\theta(\sin\theta\partial_\theta f)} + \underset{5}{(1/r^2\sin^2\theta)\partial_\varphi{}^2 f} =$$

$$\underset{1}{2\frac{\frac{\partial}{\partial r}f}{r}}+\underset{2}{\left(\frac{\partial^2}{\partial r^2}f\right)}+\underset{3}{\frac{\cos(\theta)\left(\frac{\partial}{\partial\theta}f\right)}{r^2\sin(\theta)}}+\underset{4}{\frac{\frac{\partial^2}{\partial\theta^2}f}{r^2}}+\underset{5}{\frac{\frac{\partial^2}{\partial\phi^2}f}{r^2\sin(\theta)^2}} \qquad (I.4.1)$$

`q1_;`

$$-2\frac{B_r}{r^2}+\underset{1}{2\frac{\frac{\partial}{\partial r}B_r}{r}}+\underset{2}{\left(\frac{\partial^2}{\partial r^2}B_r\right)}-2\frac{\cos(\theta)B_\theta}{r^2\sin(\theta)}-2\frac{\frac{\partial}{\partial\theta}B_\theta}{r^2}-2\frac{\frac{\partial}{\partial\phi}B_\phi}{r^2\sin(\theta)}+\underset{5}{\frac{\frac{\partial^2}{\partial\phi^2}B_r}{r^2\sin(\theta)^2}}+\underset{3}{\frac{\cos(\theta)\left(\frac{\partial}{\partial\theta}B_r\right)}{r^2\sin(\theta)}}+\underset{4}{\frac{\frac{\partial^2}{\partial\theta^2}B_r}{r^2}} \qquad (I.4.2)$$

Therefore

$$(\text{✡}B)_r = \nabla^2(B_r) - (2/r^2)B_r - (2/r^2)\cot(\theta)B_\theta - (2/r^2)\partial_\theta B_\theta - (2/r^2\sin\theta)\,\partial_\varphi B_\varphi$$

$$= \nabla^2(B_r) - (2/r^2)\,[\,B_r + \cot\theta\, B_\theta + \partial_\theta B_\theta + \csc\theta\;\partial_\varphi B_\varphi\,] \qquad (I.4.3)$$

`q2_;`

$$2\frac{\frac{\partial}{\partial\theta}B_r}{r^2}-\frac{\cos(\theta)^2B_\theta}{r^2\sin(\theta)^2}+\underset{3}{\frac{\cos(\theta)\left(\frac{\partial}{\partial\theta}B_\theta\right)}{r^2\sin(\theta)}}-\frac{B_\theta}{r^2}+\underset{4}{\frac{\frac{\partial^2}{\partial\theta^2}B_\theta}{r^2}}-2\frac{\left(\frac{\partial}{\partial\phi}B_\phi\right)\cos(\theta)}{r^2\sin(\theta)^2}+\underset{5}{\frac{\frac{\partial^2}{\partial\phi^2}B_\theta}{r^2\sin(\theta)^2}}+\underset{1}{2\frac{\frac{\partial}{\partial r}B_\theta}{r}}+\underset{2}{\left(\frac{\partial^2}{\partial r^2}B_\theta\right)} \qquad (I.4.4)$$

Therefore

$$(\text{✡}B)_\theta = \nabla^2(B_\theta) - (1/r^2)\,[\,-2\,\partial_\theta B_r + \cot^2\theta\, B_\theta + B_\theta + 2\cot\theta\csc\theta\partial_\varphi B_\varphi\,]$$

$$= \nabla^2(B_\theta) - (1/r^2)\,[\csc^2\theta\, B_\theta - 2\,\partial_\theta B_r + 2\cot\theta\csc\theta\partial_\varphi B_\varphi\,] \qquad (I.4.5)$$

`q3_;`

$$2\frac{\frac{\partial}{\partial\phi}B_r}{r^2\sin(\theta)}+2\frac{\cos(\theta)\left(\frac{\partial}{\partial\phi}B_\theta\right)}{r^2\sin(\theta)^2}+\underset{5}{\frac{\frac{\partial^2}{\partial\phi^2}B_\phi}{r^2\sin(\theta)^2}}-\frac{\cos(\theta)^2B_\phi}{r^2\sin(\theta)^2}+\underset{3}{\frac{\cos(\theta)\left(\frac{\partial}{\partial\theta}B_\phi\right)}{r^2\sin(\theta)}}-\frac{B_\phi}{r^2}+\underset{4}{\frac{\frac{\partial^2}{\partial\theta^2}B_\phi}{r^2}}+\underset{1}{2\frac{\frac{\partial}{\partial r}B_\phi}{r}}+\underset{2}{\left(\frac{\partial^2}{\partial r^2}B_\phi\right)} \qquad (I.4.6)$$

$$(\text{✡}B)_\varphi = \nabla^2(B_\varphi) - (1/r^2)\,[\csc^2\theta B_\varphi - 2\csc\theta\partial_\varphi B_r - 2\cot\theta\csc\theta\partial_\varphi B_\theta\,] \qquad (I.4.7)$$

Each component is seen to have nine terms.

In this manner, we end up with the components of the vector Laplacian expressed in the traditional manner,

$$\begin{aligned}(\text{✡}B)_r &= \nabla^2(B_r) - (2/r^2)\,[\,B_r + \cot\theta\, B_\theta + \partial_\theta B_\theta + \csc\theta\;\partial_\varphi B_\varphi\,]\\ (\text{✡}B)_\theta &= \nabla^2(B_\theta) - (1/r^2)\,[\csc^2\theta\, B_\theta - 2\,\partial_\theta B_r + 2\cot\theta\csc\theta\partial_\varphi B_\varphi\,]\\ (\text{✡}B)_\varphi &= \nabla^2(B_\varphi) - (1/r^2)\,[\csc^2\theta\, B_\varphi - 2\csc\theta\partial_\varphi B_r - 2\cot\theta\csc\theta\partial_\varphi B_\theta\,]\end{aligned}$$

where $$\nabla^2 f = (1/r^2)\partial_r(r^2\partial_r f) + (1/r^2\sin\theta)\partial_\theta(\sin\theta\partial_\theta f) + (1/r^2\sin^2\theta)\partial_\varphi{}^2 f$$

$$= (2/r)\partial_r f + \partial_r{}^2 f + (\cot\theta/r^2)\partial_\theta f + (1/r^2)\,\partial_\theta{}^2 f + (1/r^2\sin^2\theta)\partial_\varphi{}^2 f \qquad \text{(I.4.8)}$$

The curious reader might wonder *why*, in each case, five of the nine terms of the vector Laplacian components can be represented by the scalar Laplacian acting on the component. This question is answered in the following Section.

I.5 Method 2, Part A

In the Method 1, described in Section I.2 above, we used this tensorization of the vector Laplacian,

$$(\text{✡}B)'^n = (B'^j{}_{;j})^{;n} - g'^{-1/2}\varepsilon'^{nab}(g'^{-1/2}g'_{bc}\varepsilon'^{cde}B'_{e;d})_{;a} \quad . \qquad \textit{(I.2.5)} \qquad \text{(I.5.1)}$$

An alternative tensorization is stated in (15.8.2),

$$(\text{✡}B)'^n = B'^{n;j}{}_{;j} \quad . \qquad \text{(I.5.2)}$$

These two tensors must be the same since the tensorization of a Cartesian form equation is unique and this fact is verified in Section 7.8. Our Method 2 is to use this $B'^{n;j}{}_{;j}$ tensorization to compute once again the components of the vector Laplacian.

Recall covariant derivative example (F.9.9),

$$B^{ab}{}_{;\alpha} \equiv \partial_\alpha B^{ab} + \Gamma^a{}_{\alpha n} B^{nb} + \Gamma^b{}_{\alpha n} B^{an} \qquad \textit{(F.9.9)}$$

Since $B^{a;b}$ is a rank-2 tensor one can apply the above example to $B^{a;b}$ in place of B^{ab} to get a result of identical form. The second line below is the desired index contraction (I.5.2) with a→n and b=α → j :

$$B^{a;b}{}_{;\alpha} \equiv \partial_\alpha B^{a;b} + \Gamma^a{}_{\alpha k} B^{k;b} + \Gamma^b{}_{\alpha k} B^{a;k} \qquad \text{(I.5.3)}$$
$$B^{n;j}{}_{;j} \equiv \partial_j B^{n;j} + \Gamma^n{}_{jk} B^{k;j} + \Gamma^j{}_{jk} B^{n;k} \qquad // \ \Gamma'^j{}_{jk} = (1/\sqrt{g}\,)\,\partial_k(\sqrt{g}\,) \text{ as in (F.4.2)} \quad .$$

Next, recall covariant derivative example (F.9.5), with a whole sequence of index shuffles following

$$B^{a;\alpha} = \partial^\alpha B^a + g^{\alpha\beta}\Gamma^a{}_{\beta n} B^n \qquad // \text{ do } \beta\to b \text{ and } n\to s \qquad \textit{(F.9.5)}$$

$$B^{a;\alpha} = \partial^\alpha B^a + g^{\alpha b}\Gamma^a{}_{bs} B^s \qquad // \text{ do } n\to a \text{ and } \alpha\to j$$
$$B^{n;j} = \partial^j B^n + g^{jb}\Gamma^n{}_{bs} B^s \qquad // \text{ do } n\to k$$
$$B^{k;j} = \partial^j B^k + g^{jb}\Gamma^k{}_{bs} B^s \qquad // \text{ do } k\to n \text{ and } j\to k$$
$$B^{n;k} = \partial^k B^n + g^{kb}\Gamma^n{}_{bs} B^s \qquad \text{(I.5.4)}$$

Insert the last two of these expressions into the second line of (I.5.3) to get,

$$B^{n;j}{}_{;j} = \partial_j[\partial^j B^n + g^{jb}\Gamma^n{}_{bs} B^s] + \Gamma^n{}_{jk}[\partial^j B^k + g^{jb}\Gamma^k{}_{bs} B^s] + \Gamma^j{}_{jk}[\partial^k B^n + g^{kb}\Gamma^n{}_{bs} B^s] \ . \quad \text{(I.5.5)}$$

Combine the second last term with the first,

$$B^{n;j}{}_{;j} = [\partial_j \partial^j B^n + \Gamma^j{}_{jk}(\partial^k B^n)] \\ + \partial_j(g^{jb}\Gamma^n{}_{bs} B^s) + \Gamma^n{}_{jk}[\partial^j B^k + g^{jb}\Gamma^k{}_{bs} B^s] + \Gamma^j{}_{jk}[g^{kb}\Gamma^n{}_{bs} B^s] \ . \qquad \text{(I.5.6)}$$

Since no terms have been dropped, we are still "covariant" and in x'-space everything gets primed,

$$B'^{n;j}{}_{;j} = [\partial'_j \partial'^j B'^n + \Gamma'^j{}_{jk}(\partial'^k B'^n)] \\ + \partial'_j(g'^{jb}\Gamma'^n{}_{bs} B'^s) + \Gamma'^n{}_{jk}[\partial'^j B'^k + g'^{jb}\Gamma'^k{}_{bs} B'^s] + \Gamma'^j{}_{jk}[g'^{kb}\Gamma'^n{}_{bs} B'^s] \qquad \text{(I.5.7)}$$

Using $\Gamma'^j{}_{jk} = (1/\sqrt{g'})\, \partial'_k(\sqrt{g'})$ from (F.4.2), the first two terms of (I.5.7) can be written this way,

$$[\partial'_j \partial'^j B'^n + \Gamma'^j{}_{jk}(\partial'^k B'^n)] = [\partial'_j \partial'^j B'^n + (1/\sqrt{g'})\, \partial'_k(\sqrt{g'})\, (\partial'^k B'^n)] = \text{lap}(B'^n) \qquad \text{(I.5.8)}$$

and we recognize this as being lap (B'^n) based on (15.5.6),

$$\text{lap } f = \partial'_n \partial'^n f' + (1/\sqrt{g'})\, \partial'_n(\sqrt{g'})\, (\partial'^n f') = [1/\sqrt{g'}]\ \partial'_n [\sqrt{g'}\ (\partial'^n f')] \ . \qquad \textit{(15.5.6)} \qquad \text{(I.5.9)}$$

Therefore (I.5.7) can be written as,

$$B'^{n;j}{}_{;j} = \text{lap}(B'^n) + \partial'_j(g'^{jb}\Gamma'^n{}_{bs} B'^s) + \Gamma'^n{}_{jk}[\partial'^j B'^k + g'^{jb}\Gamma'^k{}_{bs} B'^s] + \Gamma'^j{}_{jk}[g'^{kb}\Gamma'^n{}_{bs} B'^s]$$

$$= \text{lap}(B'^n) \ + \ \text{Extra Terms} \qquad \text{(I.5.10)}$$

So we see why the scalar Laplacian of B'^n appears in $(✡B)^n$. Basically lap(B'^n) is the $B'^{n,j}{}_{,j}$ term within $B'^{n;j}{}_{;j}$, the only term that survives if $\Gamma = 0$.

I.6 Method 2, Part B

Unfortunately, we don't want to see lap(B'^n) in (I.5.10), we want to see lap$(\mathcal{B}'^n)$! Consider then this rewrite of (I.5.9), using (I.2.7) that $B'^n = \mathcal{B}'^n/h'_n$,

$$\text{lap }(B'^n) = \text{lap }(\mathcal{B}'^n/h'_n) = [\partial'_j \partial'^j (\mathcal{B}''^n/h'_n) + (1/\sqrt{g'})\, \partial'_k(\sqrt{g'})\, (\partial'^k (\mathcal{B}''^n/h'_n))] \ . \qquad \text{(I.6.1)}$$

One computes the pieces as follows:

$$\partial'_j \partial'^j (\mathcal{B}'^n/h'_n) = \partial'_j \partial'^j (\mathcal{B}'^n h'^{-1}{}_n) = \partial'_j [(\partial'^j \mathcal{B}'^n)\, h'^{-1}{}_n + \mathcal{B}'^n(\partial'^j h'^{-1}{}_n)] \\ = (\partial'_j \partial'^j \mathcal{B}'^n) h'^{-1}{}_n + (\partial'^j \mathcal{B}'^n)(\partial'_j h'^{-1}{}_n) + (\partial'_j \mathcal{B}'^n)(\partial'^j h'^{-1}{}_n) + \mathcal{B}'^n (\partial'_j \partial'^j h'^{-1}{}_n) \\ = (\partial'_j \partial'^j \mathcal{B}'^n)\, h'^{-1}{}_n + 2(\partial'^j \mathcal{B}'^n)(\partial'_j h'^{-1}{}_n) + \mathcal{B}'^n (\partial'_j \partial'^j h'^{-1}{}_n) \qquad \text{(I.6.2)}$$

$$\partial'^k (\mathcal{B}'^n/h'_n) = [(\partial'^k \mathcal{B}'^n)\, h'^{-1}{}_n + \mathcal{B}'^n(\partial'^k h'^{-1}{}_n)] \qquad \text{(I.6.3)}$$

so that (I.6.1) becomes,

$$\begin{aligned} \mathrm{lap}\,(B'^{n}) \;&=\; (\partial'_j\partial'^j\mathcal{B}'^{n})\, h'^{-1}{}_{n} + 2(\partial'^j\mathcal{B}'^{n})\,(\partial'_j\, h'^{-1}{}_{n}) \;+ \mathcal{B}'^{n}\,(\partial'_j\,\partial'^j\, h'^{-1}{}_{n}) \\ &\quad + (1/\sqrt{g'}\,)\,\partial'_k(\sqrt{g'}\,)[(\partial'^k\mathcal{B}'^{n})\, h'^{-1}{}_{n} \;+ \mathcal{B}'^{n}(\partial'^k\, h'^{-1}{}_{n})\,] \\[1ex] &=\; h'^{-1}{}_{n}\,\{(\partial'_j\partial'^j\mathcal{B}'^{n}) + (1/\sqrt{g'}\,)\,\partial'_k(\sqrt{g'}\,)(\partial'^k\mathcal{B}'^{n})\,\} \\ &\quad +\; 2(\partial'^j\mathcal{B}'^{n})\,(\partial'_j\, h'^{-1}{}_{n}) \;+ \mathcal{B}'^{n}\,(\partial'_j\,\partial'^j\, h'^{-1}{}_{n}) \;+ (1/\sqrt{g'}\,)\,\partial'_k(\sqrt{g'}\,)\mathcal{B}'^{n}(\partial'^k\, h'^{-1}{}_{n}) \\[1ex] &= h'^{-1}{}_{n}\,\mathrm{lap}\,(\mathcal{B}'^{n}) \qquad\qquad // \text{ (I.5.9) for } \{..\} \\ &\quad +\; 2(\partial'^j\mathcal{B}'^{n})\,(\partial'_j\, h'^{-1}{}_{n}) \;+ \mathcal{B}'^{n}\,(\partial'_j\,\partial'^j\, h'^{-1}{}_{n}) \;+ (1/\sqrt{g'}\,)\,\partial'_k(\sqrt{g'}\,)\mathcal{B}'^{n}(\partial'^k\, h'^{-1}{}_{n}) \\[1ex] &= h'^{-1}{}_{n}\,\mathrm{lap}\,(\mathcal{B}'^{n}) \;+ \text{Other Terms} \qquad . \end{aligned} \tag{I.6.4}$$

The conclusion so far from (I.5.2) and (I.5.10) and (I.6.4),

$$(\text{✡}B)'^{n} \;=\; B'^{n;j}{}_{;j} =\; \mathrm{lap}(B'^{n}) \;\; + \text{Extra Terms} \qquad\qquad \textit{(I.5.10)}$$

$$= [\;\, h'^{-1}{}_{n}\,\mathrm{lap}(\mathcal{B}'^{n}) \;+ \text{Other Terms}] \;\; + \text{Extra Terms}\;\; . \qquad\qquad \textit{(I.6.4)} \tag{I.6.5}$$

As before, our interest is with the expansion ✡**B** = $(\text{✡}\mathcal{B})'^{n}\;\hat{\mathbf{e}}_{n}$. Since (✡**B)** is a vector like any other vector, we write $(\text{✡}B)'^{n} = \; (\text{✡}\mathcal{B})'^{n}/h'_{n}$ as in (I.2.7) so that (I.6.5) becomes,

$$(\text{✡}\mathcal{B})'^{n} = \mathrm{lap}(\mathcal{B}'^{n}) + h'_{n}\;\; [\text{Other Terms} \;\; + \text{Extra Terms}]\;\; . \tag{I.6.6}$$

This result applies to any x'-space curvilinear coordinate system, orthogonal or otherwise. Thus we have demonstrated why it is that $\mathrm{lap}(\mathcal{B}'^{n})$ always appears as part of the vector Laplacian component $(\text{✡}\mathcal{B})'^{n}$.

We copy the Extras and Other Terms from (I.5.10) and (I.6.4),

$$\text{Extra Terms} \;\; = \partial'_j(g'^{jb}\Gamma'^{n}{}_{bs}\, B'^{s}) \;\; + \Gamma'^{n}{}_{jk}[\partial'^j\, B'^{k} + g'^{jb}\Gamma'^{k}{}_{bs}\, B'^{s}] + \Gamma'^{j}{}_{jk}[g'^{kb}\Gamma'^{n}{}_{bs}\, B'^{s}] \tag{I.6.7}$$

$$\text{Other Terms} = \;\; 2(\partial'^j\mathcal{B}'^{n})\,(\partial'_j\, h^{-1}{}_{n}) \;\; + \mathcal{B}'^{n}\,(\partial'^j\,\partial'_j\, h^{-1}{}_{n}) \;\; + (1/\sqrt{g'}\,)\,\partial'_k(\sqrt{g'}\,)\mathcal{B}'^{n}(\partial'^k\, h^{-1}{}_{n})\, . \tag{I.6.8}$$

Rather than attempt algebraic simplification of the above terms, we will just throw them into Maple as is. Replacing $B'^{s} = \mathcal{B}'^{s}/h'_{s}$ in Extra Terms and "lowering" the differential operators appropriately, one gets

$$\begin{aligned} &\text{Extra Terms} \;\; = \\ &\underset{\text{ET1}}{\partial'_j(g'^{jb}\,\Gamma'^{n}{}_{bs}\,\mathcal{B}'^{s}/h'_{s})} + \underset{\text{ET2}}{\Gamma'^{n}{}_{jk}g'^{js}\partial'_s(\mathcal{B}'^{k}/h'_{k})} \;\; + \underset{\text{ET3}}{\Gamma'^{n}{}_{jk}g'^{jb}\,\Gamma'^{k}{}_{bs}\,\mathcal{B}'^{s}/h'_{s}} + \underset{\text{ET4}}{\Gamma'^{j}{}_{jk}g'^{kb}\,\Gamma'^{n}{}_{bs}\,\mathcal{B}'^{s}/h'_{s}} \end{aligned} \tag{I.6.9}$$

$$\begin{aligned} &\text{Other Terms} \;\; = \\ &\quad \underset{\text{OT1}}{2(g'^{js}\,\partial'_s\mathcal{B}'^{n})\,(\partial'_j h'_{n}{}^{-1})} \;\; + \underset{\text{OT2}}{\mathcal{B}'^{n}\,(\partial'_j\,\partial'^j h'_{n}{}^{-1})} \;\;\; + \underset{\text{OT3}}{(1/\sqrt{g'}\,)\,\partial'_k(\sqrt{g'}\,)\mathcal{B}'^{n}(\partial'^k h'_{n}{}^{-1})} \;\; . \end{aligned} \tag{I.6.10}$$

I.7 Method 2 for spherical coordinates: Maple speaks again

In Method 1, there was no need in the Maple calculation for the affine connection object of (F.4.1),

$$\Gamma'^{\mathbf{d}}{}_{\mathbf{ab}} = (1/2)\, g'^{\mathbf{dc}}\ [\, \partial'_{\mathbf{a}} g'_{\mathbf{bc}} + \partial'_{\mathbf{b}} g'_{\mathbf{ca}} - \partial'_{\mathbf{c}} g'_{\mathbf{ab}}] \qquad \text{(I.7.1)}$$

which in orthogonal coordinates simplifies to

$$\Gamma'^{\mathbf{d}}{}_{\mathbf{ab}} = (1/2)\, h'_{\mathbf{d}}{}^{-2}\, [\delta^{\mathbf{d}}{}_{\mathbf{b}}\, \partial'_{\mathbf{a}}(h'_{\mathbf{b}}{}^2) + \delta^{\mathbf{d}}{}_{\mathbf{a}} \partial'_{\mathbf{b}}(h'_{\mathbf{a}}{}^2) - \delta_{\mathbf{ab}}\, \partial'_{\mathbf{d}}(h'_{\mathbf{a}}{}^2)] \ . \qquad \text{(I.7.2)}$$

This last form shows that $\Gamma'^{\mathbf{d}}{}_{\mathbf{ab}} = 0$ unless two indices match, in which case it *might* not vanish. For coordinate systems like spherical and cylindrical coordinates, $\Gamma'^{\mathbf{d}}{}_{\mathbf{ab}}$ is "sparsely populated", and this explains why won't be swamped by all those Extra Terms shown above.

For spherical coordinates, only 9 of the 27 elements of the object $\Gamma^{\mathbf{d}}{}_{\mathbf{ab}}$ are non-zero :

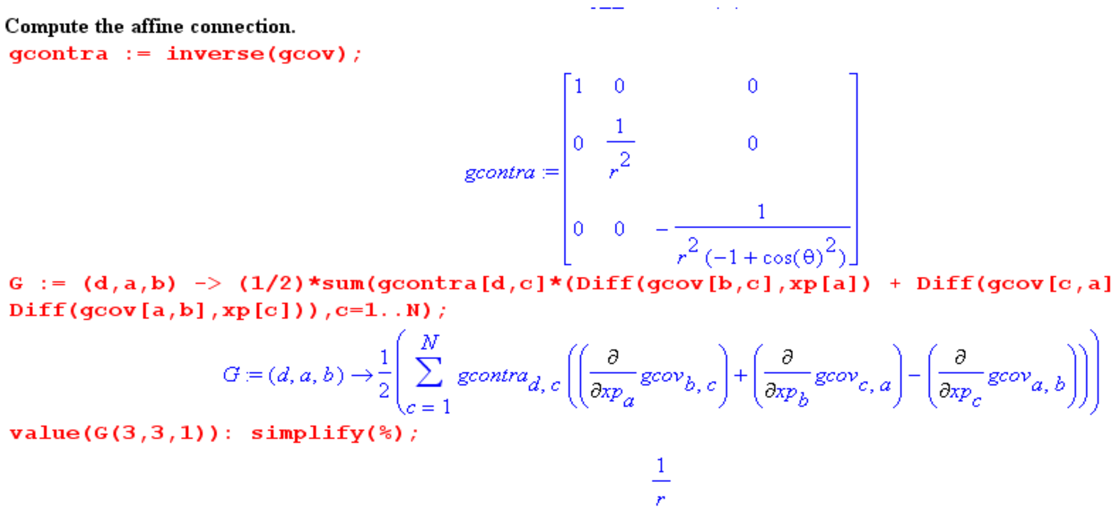
$\Gamma'^1{}_{22} = -r$ $\qquad$ $\Gamma'^2{}_{12} = \Gamma'^2{}_{21} = 1/r$ $\qquad$ // notation: $\quad \Gamma'^1{}_{22} = \Gamma^{\mathbf{r}}{}_{\theta\theta}$
$\Gamma'^1{}_{33} = -r\sin^2\theta$ $\qquad$ $\Gamma'^3{}_{13} = \Gamma'^3{}_{31} = 1/r$
$\Gamma'^2{}_{33} = -\cos\theta\sin\theta$ $\qquad$ $\Gamma'^3{}_{23} = \Gamma'^3{}_{32} = \cot\theta$ $\qquad$ // 1,2,3 = r,θ,φ = radius, polar, azimuthal

$$\Gamma^{\mathbf{r}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -r & 0 \\ 0 & 0 & -r\sin^2\theta \end{bmatrix} \qquad \Gamma^{\theta} = \begin{bmatrix} 0 & 1/r & 0 \\ 1/r & 0 & 0 \\ 0 & 0 & -\sin\theta\cos\theta \end{bmatrix} \qquad \Gamma^{\varphi} = \begin{bmatrix} 0 & 0 & 1/r \\ 0 & 0 & \cot\theta \\ 1/r & \cot\theta & 0 \end{bmatrix} \qquad \text{(I.7.3)}$$

Note that each matrix is symmetric since $\Gamma^{\mathbf{a}}{}_{\mathbf{bc}} = \Gamma^{\mathbf{a}}{}_{\mathbf{cb}}$.

To maintain generality, however, we let Maple compute $\Gamma^{\mathbf{d}}{}_{\mathbf{ab}} = G(d,a,b)$ from (I.7.1), so

$$\Gamma'^{\mathbf{d}}{}_{\mathbf{ab}} = (1/2)\, g'^{\mathbf{dc}}\ [\, \partial'_{\mathbf{a}} g'_{\mathbf{bc}} + \partial'_{\mathbf{b}} g'_{\mathbf{ca}} - \partial'_{\mathbf{c}} g'_{\mathbf{ab}}] \qquad \text{(I.7.4)}$$



Compute the affine connection.

```
gcontra := inverse(gcov);
```

$$gcontra := \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{r^2} & 0 \\ 0 & 0 & -\frac{1}{r^2\,(-1+\cos(\theta)^2)} \end{bmatrix}$$

```
G := (d,a,b) -> (1/2)*sum(gcontra[d,c]*(Diff(gcov[b,c],xp[a]) + Diff(gcov[c,a],xp[b])-
Diff(gcov[a,b],xp[c])),c=1..N);
```

$$G := (d,a,b) \rightarrow \frac{1}{2}\left(\sum_{c=1}^{N} gcontra_{d,c}\left(\left(\frac{\partial}{\partial xp_a} gcov_{b,c}\right) + \left(\frac{\partial}{\partial xp_b} gcov_{c,a}\right) - \left(\frac{\partial}{\partial xp_c} gcov_{a,b}\right)\right)\right)$$

```
value(G(3,3,1)): simplify(%);
```

$$\frac{1}{r}$$

The Extra Terms are then entered, as shown above : (I.7.5)

Extra Terms =

$$\partial'_j(g'^{jb}\,\Gamma'^n{}_{bs}\,\mathcal{B}'^s/h'_s) + \Gamma'^n{}_{jk}g'^{js}\partial'_s(\mathcal{B}'^k/h'_k) + \Gamma'^n{}_{jk}g'^{jb}\,\Gamma'^k{}_{bs}\,\mathcal{B}'^s/h'_s + \Gamma'^j{}_{jk}g'^{kb}\,\Gamma'^n{}_{bs}\,\mathcal{B}'^s/h'_s$$

ET1 ET2 ET3 ET4

Compute ET = "the extra terms"

```
ET1 := (n) ->
sum(sum(sum(Diff(gcontra[j,b]*G(n,b,s)*Bp[s](xp[1],xp[2],xp[3])/hp[s],xp[j]),b=1..N),j=1..N),
s=1..N);
```

$$ET1 := n \to \sum_{s=1}^{N}\left(\sum_{j=1}^{N}\left(\sum_{b=1}^{N}\left(\frac{\partial}{\partial xp_j}\,\frac{gcontra_{j,b}\,\mathrm{G}(n,b,s)\,Bp_s(xp_1,xp_2,xp_3)}{hp_s}\right)\right)\right)$$

```
ET2 := (n) ->
sum(sum(sum(G(n,j,k)*gcontra[j,s]*Diff(Bp[k](xp[1],xp[2],xp[3])/hp[k],xp[s]),j=1..N),k=1..N),
s=1..N);
```

$$ET2 := n \to \sum_{s=1}^{N}\left(\sum_{k=1}^{N}\left(\sum_{j=1}^{N}\mathrm{G}(n,j,k)\,gcontra_{j,s}\left(\frac{\partial}{\partial xp_s}\,\frac{Bp_k(xp_1,xp_2,xp_3)}{hp_k}\right)\right)\right)$$

```
ET3 := (n) -> sum(sum(sum(sum(
G(n,j,k)*gcontra[j,b]*G(k,b,s)*Bp[s](xp[1],xp[2],xp[3])/hp[s],b=1..N),j=1..N),s=1..N),k=1..N)
;
```

$$ET3 := n \to \sum_{k=1}^{N}\left(\sum_{s=1}^{N}\left(\sum_{j=1}^{N}\left(\sum_{b=1}^{N}\frac{\mathrm{G}(n,j,k)\,gcontra_{j,b}\,\mathrm{G}(k,b,s)\,Bp_s(xp_1,xp_2,xp_3)}{hp_s}\right)\right)\right)$$

```
ET4 := (n) -> sum(sum(sum(sum(
G(j,j,k)*gcontra[k,b]*G(n,b,s)*Bp[s](xp[1],xp[2],xp[3])/hp[s],b=1..N),j=1..N),s=1..N),k=1..N)
;
```

$$ET4 := n \to \sum_{k=1}^{N}\left(\sum_{s=1}^{N}\left(\sum_{j=1}^{N}\left(\sum_{b=1}^{N}\frac{\mathrm{G}(j,j,k)\,gcontra_{k,b}\,\mathrm{G}(n,b,s)\,Bp_s(xp_1,xp_2,xp_3)}{hp_s}\right)\right)\right)$$

(I.7.6)

And then the Other Terms are entered as well,

Other Terms =

$$2(g'^{js}\,\partial'_s\mathcal{B}'^n)\,(\partial'_j h'_n{}^{-1}) + \mathcal{B}'^n\,(\partial'_j\,\partial'^j h'_n{}^{-1}) + (1/\sqrt{g'}\,)\,\partial'_k(\sqrt{g'}\,)\mathcal{B}'^n(\partial'^k h'_n{}^{-1}) \qquad \text{(I.7.7)}$$

OT1 OT2 OT3

```
OT1 := (n) ->
sum(sum(2*Diff(1/hp[n],xp[j])*gcontra[j,s]*Diff(Bp[n](xp[1],xp[2],xp[3]),xp[s]),j=1..N),s=1.
.N);
```

$$OT1 := n \to \sum_{s=1}^{N}\left(\sum_{j=1}^{N}\left(2\left(\frac{\partial}{\partial xp_j}\,\frac{1}{hp_n}\right)gcontra_{j,s}\left(\frac{\partial}{\partial xp_s}Bp_n(xp_1,xp_2,xp_3)\right)\right)\right)$$

```
OT2 := (n) -> sum(sum(gcontra[j,s]*
Diff(Diff(1/hp[n],xp[j]),xp[s])*Bp[n](xp[1],xp[2],xp[3]),j=1..N),s=1..N);
```

$$OT2 := n \to \sum_{s=1}^{N}\left(\sum_{j=1}^{N}gcontra_{j,s}\left(\frac{\partial}{\partial xp_s}\left(\frac{\partial}{\partial xp_j}\,\frac{1}{hp_n}\right)\right)Bp_n(xp_1,xp_2,xp_3)\right)$$

```
OT3 := (n) ->
sum(sum((1/sq_gp)*Diff(sq_gp,xp[j])*gcontra[j,s]*Diff(1/hp[n],xp[s])*Bp[n](xp[1],xp[2],xp[3]
),j=1..N),s=1..N);
```

$$OT3 := n \to \sum_{s=1}^{N}\left(\sum_{j=1}^{N}\frac{\left(\frac{\partial}{\partial xp_j}sq\_gp\right)gcontra_{j,s}\left(\frac{\partial}{\partial xp_s}\,\frac{1}{hp_n}\right)Bp_n(xp_1,xp_2,xp_3)}{sq\_gp}\right)$$

(I.7.8)

Maple now computes the non- lap($\mathcal{B}'^n$) terms shown in (I.6.6), called starBx(n) in the code,

$$(✡\mathcal{B})'^n = \text{lap}(\mathcal{B}'^n) + h'_n\ [\text{Other Terms } + \text{Extra Terms}]\ . \tag{I.6.6}$$

The final results are then generated :

```
starBx := (n) -> hp[n]*(ET1(n)+ET2(n)+ET3(n)+ET4(n)+OT1(n)+OT2(n)+OT3(n));
```

$$starBx := n \to hp_n\,(\mathrm{ET1}(n)+\mathrm{ET2}(n)+\mathrm{ET3}(n)+\mathrm{ET4}(n)+\mathrm{OT1}(n)+\mathrm{OT2}(n)+\mathrm{OT3}(n))$$

```
value(starBx(1)):simplify(%):expand(%);
```

$$-2\,\frac{\cos(\theta)\,B_\theta(r,\theta,\phi)}{\sin(\theta)\,r^2} - 2\,\frac{\frac{\partial}{\partial\theta}B_\theta(r,\theta,\phi)}{r^2} - 2\,\frac{\frac{\partial}{\partial\phi}B_\phi(r,\theta,\phi)}{\sin(\theta)\,r^2} - 2\,\frac{B_r(r,\theta,\phi)}{r^2}$$

```
value(starBx(2)):expand(%):simplify(%):subs(cos(theta)^2=1-sin(theta)^2, %):expand(%);
```

$$2\,\frac{\frac{\partial}{\partial\theta}B_r(r,\theta,\phi)}{r^2} - 2\,\frac{\cos(\theta)\left(\frac{\partial}{\partial\phi}B_\phi(r,\theta,\phi)\right)}{r^2\,\sin(\theta)^2} - \frac{B_\theta(r,\theta,\phi)}{r^2\,\sin(\theta)^2}$$

```
value(starBx(3)):expand(%):simplify(%):subs(cos(theta)^2=1-sin(theta)^2, %):expand(%);
```

$$2\,\frac{\cos(\theta)\left(\frac{\partial}{\partial\phi}B_\theta(r,\theta,\phi)\right)}{\sin(\theta)^2\,r^2} - \frac{B_\phi(r,\theta,\phi)}{\sin(\theta)^2\,r^2} + 2\,\frac{\frac{\partial}{\partial\phi}B_r(r,\theta,\phi)}{\sin(\theta)\,r^2}$$

(I.7.9)

Recall that results of Method 1 were,

$$\begin{aligned}
(✡B)_r &= \nabla^2(B_r) - (2/r^2)\,[\ B_r + \cot\theta\ B_\theta + \partial_\theta B_\theta + \csc\theta\ \ \partial_\varphi B_\varphi\ ] \\
(✡B)_\theta &= \nabla^2(B_\theta) - (1/r^2)\,[\csc^2\theta\ B_\theta - 2\ \partial_\theta B_r \quad + 2\cot\theta\csc\theta\partial_\varphi B_\varphi\ ] \\
(✡B)_\varphi &= \nabla^2(B_\varphi) - (1/r^2)\,[\csc^2\theta\ B_\varphi - 2\csc\theta\partial_\varphi B_r - 2\cot\theta\csc\theta\partial_\varphi B_\theta\ \ ]
\end{aligned} \tag{I.4.8}$$

One then sees that the non-scalar-Laplacian terms just computed as (I.7.9) by Method 2 exactly match the non-∇^2 terms in (I.4.8) as found by Method 1.

I.8 Results for Cylindrical Coordinates from both methods

The Maple program was easily modified for this system. Here are the results:

$$\mathcal{B}'^1 = B_r \qquad \mathcal{B}'^2 = B_\varphi \qquad \mathcal{B}'^3 = B_z$$

$$\mathbf{B} = \mathcal{B}'^n \hat{\mathbf{e}}_n = B_r\hat{\mathbf{e}}_r + B_\varphi\hat{\mathbf{e}}_\varphi + B_\varphi\hat{\mathbf{e}}_z = B_r\hat{\mathbf{r}} + B_\theta\hat{\boldsymbol{\varphi}} + B_\varphi\hat{\mathbf{z}}\ . \tag{I.8.1}$$

The scalar Laplacian of an unspecified function f :

$$\nabla^2 f = \underset{1+2}{(1/r)\partial_r(rf)} + \underset{3}{(1/r^2)\partial_\varphi{}^2 f} + \underset{4}{\partial_z{}^2 f} =$$

$$\underset{1}{\frac{\frac{\partial}{\partial r}f}{r}} + \underset{2}{\left(\frac{\partial^2}{\partial r^2}f\right)} + \underset{3}{\frac{\frac{\partial^2}{\partial \phi^2}f}{r^2}} + \underset{4}{\left(\frac{\partial^2}{\partial z^2}f\right)} \tag{I.8.2}$$

The components of the vector Laplacian are found by **Method 1** to be

`q1_;`

$$-\frac{B_r}{r^2} + \underset{1}{\frac{\frac{\partial}{\partial r}B_r}{r}} + \underset{2}{\left(\frac{\partial^2}{\partial r^2}B_r\right)} - 2\,\frac{\frac{\partial}{\partial \phi}B_\phi}{r^2} + \underset{4}{\left(\frac{\partial^2}{\partial z^2}B_r\right)} + \underset{3}{\frac{\frac{\partial^2}{\partial \phi^2}B_r}{r^2}} \tag{I.8.3}$$

Therefore,

$$(\text{✡}B)_r = \nabla^2(B_r) - B_r/r^2 - 2\partial_\varphi B_\varphi/r^2 \ . \tag{I.8.4}$$

`q2_;`

$$2\,\frac{\frac{\partial}{\partial \phi}B_r}{r^2} + \underset{3}{\frac{\frac{\partial^2}{\partial \phi^2}B_\phi}{r^2}} + \underset{4}{\left(\frac{\partial^2}{\partial z^2}B_\phi\right)} - \frac{B_\phi}{r^2} + \underset{1}{\frac{\frac{\partial}{\partial r}B_\phi}{r}} + \underset{2}{\left(\frac{\partial^2}{\partial r^2}B_\phi\right)} \tag{I.8.5}$$

Therefore,

$$(\text{✡}B)_\varphi = \nabla^2(B_\varphi) - B_\varphi/r^2 + 2\partial_\varphi B_r/r^2 \ . \tag{I.8.6}$$

`q3_;`

$$\underset{4}{\left(\frac{\partial^2}{\partial z^2}B_z\right)} + \underset{3}{\frac{\frac{\partial^2}{\partial \phi^2}B_z}{r^2}} + \underset{1}{\frac{\frac{\partial}{\partial r}B_z}{r}} + \underset{2}{\left(\frac{\partial^2}{\partial r^2}B_z\right)} \tag{I.8.7}$$

Therefore,

$$(\text{✡}B)_z = \nabla^2(B_z) \tag{I.8.8}$$

as befits a component which is Cartesian. In this manner, Method 1 produces the components of the vector Laplacian expressed in the traditional manner,

$$\begin{aligned}(\text{✡}B)_r &= \nabla^2(B_r) \ \ - (1/r^2)[B_r + 2\partial_\varphi B_\varphi]\\ (\text{✡}B)_\varphi &= \nabla^2(B_\varphi) \ \ - (1/r^2)[B_\varphi - 2\partial_\varphi B_r]\\ (\text{✡}B)_z &= \nabla^2(B_z)\end{aligned}$$

$$\text{where } \nabla^2 f = (1/r)\partial_r(rf) + (1/r^2)\partial_\varphi{}^2 f + \partial_z{}^2 f \ = \partial^2{}_r f + (1/r)\partial_r f + (1/r^2)\partial_\varphi{}^2 f + \partial_z{}^2 f \qquad \text{(I.8.9)}$$

Method 2 produces the following results for the terms which are added to the scalar Laplacian,

```
starBx := (n) ->
hp[n]*(ET1(n)+ET2(n)+ET3(n)+ET4(n)+OT1(n)+OT2(n)+OT3(n));
```

$$starBx := n \to hp_n\,(\mathrm{ET1}(n) + \mathrm{ET2}(n) + \mathrm{ET3}(n) + \mathrm{ET4}(n) + \mathrm{OT1}(n) + \mathrm{OT2}(n) + \mathrm{OT3}(n))$$

```
value(starBx(1)):simplify(%):expand(%);
```

$$-2\,\frac{\frac{\partial}{\partial\phi}B_\phi(r,\phi,z)}{r^2} - \frac{B_r(r,\phi,z)}{r^2}$$

```
value(starBx(2)):expand(%):simplify(%):subs(cos(phi)^2=1-sin(phi)^2
, %):expand(%);
```

$$2\,\frac{\frac{\partial}{\partial\phi}B_r(r,\phi,z)}{r^2} - \frac{B_\phi(r,\phi,z)}{r^2}$$

```
value(starBx(3)):expand(%):simplify(%):subs(cos(phi)^2=1-sin(phi)^2
, %):expand(%);
```

$$0 \qquad \text{(I.8.10)}$$

and these are seen to agree with non-∇^2 terms in (I.8.9),

$$\begin{aligned}(\text{✡}B)_r &= \nabla^2(B_r) \ \ - (1/r^2)[B_r + 2\partial_\varphi B_\varphi]\\ (\text{✡}B)_\varphi &= \nabla^2(B_\varphi) \ \ - (1/r^2)[B_\varphi - 2\partial_\varphi B_r]\\ (\text{✡}B)_z &= \nabla^2(B_z)\end{aligned} \qquad \textit{(I.8.9)}$$

In cylindrical coordinates the Γ object is even sparser than in spherical coordinates. One has

$$\Gamma'^d{}_{ab} = (1/2)\, h'_d{}^{-2}\, [\delta^d{}_b\, \partial'_a(h'_b{}^2) \ + \ \ \delta^d{}_a \partial'_b(h'_a{}^2) \ \ - \ \delta_{ab}\, \partial'_d(h'_a{}^2)] \ .$$

$\Gamma'^1{}_{22} = -r$ // notation: $\Gamma'^1{}_{22} = \Gamma^r{}_{\varphi\varphi}$

$\Gamma'^2{}_{12} = 1/r$

$\Gamma'^2{}_{21} = 1/r$ // 1,2,3 = r,φ,z

$$\Gamma^r = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -r & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \Gamma^\varphi = \begin{array}{cc} \begin{array}{ccc} r & \varphi & z \end{array} & \\ \begin{bmatrix} 0 & 1/r & 0 \\ 1/r & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} & \begin{array}{c} r \\ \varphi \\ z \end{array} \end{array} \qquad \Gamma^z = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \qquad \text{(I.8.11)}$$

so only 3 of 27 components are non-vanishing.

<u>Hand Calculation of Γ' for cylindrical coordinates</u>

From a simple polar coordinates picture one knows that

$$d\hat{\mathbf{r}} = d\varphi\, \hat{\boldsymbol{\varphi}} \qquad\qquad d\hat{\boldsymbol{\varphi}} = -d\varphi\, \hat{\mathbf{r}}\ . \qquad\qquad (\text{I.8.12})$$

Then from (3.4.2) with θ → φ, we have this situation for cylindrical coordinates,

$$\begin{array}{lll} \mathbf{e_r} = \hat{\mathbf{e}}_\mathbf{r} = \hat{\mathbf{r}} & \Rightarrow & d\mathbf{e_r} = d\hat{\mathbf{r}} = d\varphi\, \hat{\boldsymbol{\varphi}} = d\varphi\, (\mathbf{e}_\varphi/r) = (1/r)d\varphi\, \mathbf{e}_\varphi \\ \mathbf{e}_\varphi = r\, \hat{\boldsymbol{\varphi}} & \Rightarrow & d\mathbf{e}_\varphi = d(r\, \hat{\boldsymbol{\varphi}}) = r\, d\hat{\boldsymbol{\varphi}} + dr\, \hat{\boldsymbol{\varphi}} = r\, (\, -d\varphi\, \hat{\mathbf{r}}) + dr\, \hat{\boldsymbol{\varphi}} = -rd\varphi\, \mathbf{e_r} + (dr/r)\mathbf{e}_\varphi \\ \mathbf{e_z} = \hat{\mathbf{z}} & \Rightarrow & d\mathbf{e_z} = 0\ . \qquad\qquad (\text{I.8.13}) \end{array}$$

Therefore only 3 of 9 partial derivatives are non-zero,

$$\begin{array}{lll} \partial_\mathbf{r}\mathbf{e_r} = 0 & \partial_\mathbf{r}\mathbf{e}_\varphi = (1/r)\mathbf{e}_\varphi & \partial_\mathbf{z}\mathbf{e_r} = 0 \\ \partial_\varphi\mathbf{e_r} = (1/r)\mathbf{e}_\varphi & \partial_\varphi\mathbf{e}_\varphi = -r\mathbf{e_r} & \partial_\mathbf{z}\mathbf{e}_\varphi = 0 \\ \partial_\mathbf{z}\mathbf{e_r} = 0 & \partial_\mathbf{z}\mathbf{e}_\varphi = 0 & \partial_\mathbf{z}\mathbf{e_z} = 0\ . \qquad\qquad (\text{I.8.14}) \end{array}$$

Since we are in the Picture A context of Fig (3.4.3), we use (F.1.11) which says

$$\Gamma'^{\mathbf{c}}{}_{\mathbf{ab}} = \mathbf{e^c} \bullet (\partial'_\mathbf{a}\mathbf{e_b})\ . \qquad\qquad \textit{(F.1.11)}$$

Then

$$\Gamma'^{\mathbf{c}}{}_{\mathbf{ab}} = g'^{\mathbf{cd}}\ \mathbf{e_d} \bullet (\partial'_\mathbf{a}\mathbf{e_b}) \ = h'_\mathbf{c}{}^{-2}\ \mathbf{e_c} \bullet (\partial'_\mathbf{a}\mathbf{e_b})\ . \qquad // \text{ no sum on c} \qquad\qquad (\text{I.8.15})$$

For our situation with **x**' = (r,φ,z), we find that for $\Gamma'^{\mathbf{c}}{}_{\mathbf{ab}}$ to *not* vanish, one must have ab = φr, rφ and φφ since all other $(\partial'_\mathbf{j}\mathbf{e_b}) = 0$ in (I.8.14). In each of these cases only one Γ component survives due the orthogonality of the base vectors :

$$\Gamma'^{\varphi}{}_{\varphi\varphi} = h_\varphi{}^{-2}\ \mathbf{e}_\varphi \bullet (\partial_\varphi\mathbf{e}_\varphi) = r^{-2}\ \mathbf{e}_\varphi \bullet [\ -r\mathbf{e_r}] = 0$$

$$\Gamma'^{\mathbf{z}}{}_{\varphi\varphi} = h_\mathbf{z}{}^{-2}\ \mathbf{e_z} \bullet (\partial_\varphi\mathbf{e}_\varphi) \ = 1 * \mathbf{e_z} \bullet \ [\ -r\mathbf{e_r}] \ = 0$$

$$\Gamma'^{\mathbf{r}}{}_{\varphi\varphi} = h_\mathbf{r}{}^{-2}\ \mathbf{e_r} \bullet (\partial_\varphi\mathbf{e}_\varphi) \ = 1 * \mathbf{e_r} \bullet [\ -r\mathbf{e_r}] \ = -r\ \hat{\mathbf{r}} \bullet \hat{\mathbf{r}} \ = -\ r$$

$$\Gamma'^{\varphi}{}_{\mathbf{r}\varphi} = h_\varphi{}^{-2}\ \mathbf{e}_\varphi \bullet (\partial_\mathbf{r}\mathbf{e}_\varphi) = r^{-2}\ \mathbf{e}_\varphi \bullet [(1/r)\mathbf{e}_\varphi] \ = r^{-2}\ \ r\ \hat{\boldsymbol{\varphi}} \bullet \hat{\boldsymbol{\varphi}} \ = (1/r)$$

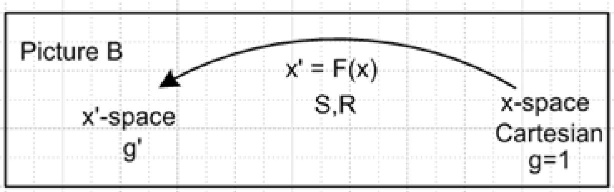
$$\Gamma'^{\varphi}{}_{\varphi\mathbf{r}} = h_\varphi{}^{-2}\ \mathbf{e}_\varphi \bullet (\partial_\varphi\mathbf{e_r}) \ = r^{-2}\ \mathbf{e}_\varphi \bullet [\ (1/r)\mathbf{e}_\varphi] \ = \ r^{-2}\ \ r\ \hat{\boldsymbol{\varphi}} \bullet \hat{\boldsymbol{\varphi}} \ = (1/r) \qquad\qquad (\text{I.8.16})$$

These results agree with (I.8.11).

Appendix J: Expansion of (∇T) in curvilinear coordinates (T = rank-2 tensor)

J.1 Total time derivative as prototype equation

This appendix assumes the usual curvilinear coordinates context, Picture B :



(J.1.1)

The total time derivative of a contravariant rank-2 tensor field $T^{ij}(\mathbf{x},t)$ can be written as

$$dT^{ij}(\mathbf{x},t)/dt = \partial T^{ij}/\partial t + (\partial T^{ij}/\partial x^k)\ (\partial x^k/dt) = \partial T^{ij}/\partial t + (\partial T^{ij}/\partial x^k)\ v^k$$

or

$$d_t T^{ij} = \partial_t T^{ij} + (\partial_k T^{ij})\ v^k \quad . \qquad // \ d_t \equiv d/dt,\ \ \partial_t \equiv \partial/\partial t,\ \ \partial_k \equiv \partial/\partial x^k \qquad (J.1.2)$$

We take this as a useful prototype equation for two reasons. First, it contains our object of interest, which is the gradient of a rank-2 tensor, $\partial_k T^{ij}$. Second, this equation plays a role in the continuum mechanics of non-Newtonian fluids as discussed in Section J.6 below.

One can define, in Cartesian coordinates, a (∇T) object :

$$(\nabla T)^{ij}{}_k \equiv \partial_k T^{ij} \qquad (J.1.3)$$

so that, from (J.1.2),

$$d_t T^{ij} = \partial_t T^{ij} + (\nabla T)^{ij}{}_k\ v^k \ . \qquad (J.1.4)$$

Comment: Our convention has been to bold vectors and not to bold other tensors. In line with this idea, we shall write **∇**T where the grad is bolded and the T is not bolded.

In order to express the above equation in curvilinear coordinates, it must be "tensorized" in the sense of Section 15.2 so that the equation is covariant. Thus, $(\nabla T)^{ij}{}_k$ must be regarded as components of a (mixed) rank-3 tensor which, in Cartesian coordinates, are equal to $\partial_k T^{ij}$. Since v^k are the components of a tensorial vector, $(\nabla T)^{ij}{}_k v^k$ transforms as a rank-2 tensor, and then all terms in the above equation are rank-2 tensors and the equation is then covariant and therefore appears this way in x'-space,

$$d_t T'^{ij} = \partial_t T'^{ij} + (\nabla T)'^{ij}{}_k\ v'^k \ . \qquad (J.1.5)$$

In our usual formalism, x'-space is the space of some generic curvilinear coordinates x'^n (not necessarily orthogonal) and then (J.1.5) tells us the form taken by the time derivative equation in curvilinear coordinates, and it remains only to compute the objects $(\nabla T')^{ij}{}_k$.

For convenience, we can lower the tensorial ij indices on the above tensor equations to get

$$d_t T_{ij} = \partial_t T_{ij} + (\nabla T)_{ijk}\, v^k \qquad (\nabla T)_{ijk} \equiv \partial_k T_{ij} \qquad \text{// Cartesian coordinates}$$

$$d_t T'_{ij} = \partial_t T'_{ij} + (\nabla T)'_{ijk}\, v'^k \qquad \text{// curvilinear coordinates} \qquad (J.1.6)$$

and then we can deal with the pure covariant tensor components $(\nabla T)'_{ijk}$.

As in previous appendices, we shall be interested in (J.1.6) stated in $\hat{\mathbf{e}}_n$-expanded coefficients according to the covariance idea of (E.9.13),

$$d_t \mathcal{T}'_{ij} = \partial_t \mathcal{T}'_{ij} + (\nabla\mathcal{T})'_{ijk}\, \mathcal{v}'^k \qquad (J.1.7)$$

where the components $(\nabla\mathcal{T})'_{ijk} = (\nabla\mathcal{T})'^{ijk} = [(\nabla T)^{(\hat{e})}]^{ijk}$ are computed in (J.7.8) below and evaluated for specific orthogonal coordinate systems in subsequent equations.

J.2 Computation of components $(\nabla T)'_{ijk}$

The covariant derivative is discussed in Sections F.7- F.9. We shall define a true tensor object $(\nabla T)_{abc}$ as $T_{ab;\alpha}$ which is defined in (F.9.6),

$$(\nabla T)_{abc} \equiv T_{ab;c} \equiv T_{ab,c} - \Gamma^n{}_{ac} T_{nb} - \Gamma^n{}_{bc} T_{an} = T_{ab,c} = \partial_c T_{ab} \qquad \text{// x-space}$$
$$(\nabla T)'_{abc} \equiv T'_{ab;c} \equiv T'_{ab,c} - \Gamma'^n{}_{ac} T'_{nb} - \Gamma'^n{}_{bc} T'_{an} \ . \qquad \text{// x'-space} \qquad (J.2.1)$$

Recall that $T_{ab,c}$ is a shorthand for $\partial_c T_{ab}$ and that $\Gamma^c{}_{ab}$ is the affine connection which tells how the basis vectors change as one moves around in space. In Cartesian x-space (first line above, see also Picture C1 in Fig (F.1.1)) the basis vectors are the fixed $\mathbf{u}_n$ which don't change, so $\Gamma^c{}_{ab} \equiv 0$ as in (F.4.16). In curvilinear x'-space, $\Gamma'^c{}_{ab} \neq 0$. Since $T_{ab;c}$ transforms as a true rank-3 tensor, its defining equation is "covariant" (Section 7.15) so that in x'-space the equation has exactly the same form but everything is primed.

The above two lines characterize the process of "tensorization": we find a true "tensorial tensor" $T_{ab;c}$ which agrees with $(\nabla T)_{abc} = \partial_c T_{ab}$ in Cartesian space. The tensorized version of $T_{ab,c}$ is unique (as shown in Section 15.2), and it is $T_{ab;c}$. Since this tensor is given by $T'_{ab;c}$ in x'-space, we use the second line above to compute the components of the tensor (∇T) object in x'-space, which is to say, in curvilinear coordinates.

It is a simple matter to have Maple compute $\Gamma'^c{}_{ab}$ for any curvilinear coordinate system, and then the second line in (J.2.1) reports out the components $(\nabla T)'_{abc}$,

$$(\nabla T)'_{abc} = \partial'_c T'_{ab} - \Gamma'^n{}_{ac} T'_{nb} - \Gamma'^n{}_{bc} T'_{an}$$

$$\Gamma'^d{}_{ab} = (1/2)\, g'^{dc}\, [\, \partial'_a g'_{bc} + \partial'_b g'_{ca} - \partial'_c g'_{ab}] \qquad \text{// (F.4.1)} \qquad (J.2.2)$$

where g'_{ab} is the metric tensor in x'-space. Then we know from (J.1.6) how to write the time derivative equation in any curvilinear coordinates,

$$d_tT'_{ij} = \partial_tT'_{ij} + (\nabla T)'_{ijk}\, v'^k$$

or

$$d_tT'^{ij} = \partial_tT'^{ij} + (\nabla T)'^{ij}{}_k\, v'^k \qquad\qquad (\nabla T)'^{ij}{}_k = g'^{ii'}g'^{jj'}(\nabla T)'_{i'j'k}\ . \qquad (J.2.3)$$

In (I.3.2) we show Maple code to compute the covariant metric tensor g'_{ij} from the curvilinear coordinates' defining equations (such as x = rsinθcosφ for sphericals). Then (I.7.5) shows the extra code for computing g'^{ij} and $\Gamma'^d{}_{ab}$. The reader can then add a few extra lines to have Maple compute the desired $(\nabla T)'_{abc}$ using the equations (J.2.2) above. Later in this Appendix we shall use Maple to compute certain related quantities which we can then verify against a known source.

J.3 Tensor expansions of ∇T on the $\mathbf{u}_n$ and $\mathbf{e}_n$ base vectors

It is useful at this point to write out the tensor expansions for the rank-3 tensor (∇T) to show exactly where the components $(\nabla T)'_{abc}$ appear. As shown in (E.2.11) and (E.2.14) one can expand the rank-3 tensor (∇T) in various ways. The first line below shows expansions on x-space basis vectors, while the second line shows expansion on the reciprocal and tangent base vectors (which also exist in x-space) :

$$\nabla T = \Sigma_{ijk}\, T_{ij;k}\, \mathbf{u}^i\otimes\mathbf{u}^j\otimes\mathbf{u}^k \;=\; \Sigma_{ijk}\, [(\nabla T)^{(u)}]_{ijk}\, \mathbf{u}^i\otimes\mathbf{u}^j\otimes\mathbf{u}^k \;=\Sigma_{ijk}\, [(\nabla T)^{(u)}]^{ijk}\, \mathbf{u}_i\otimes\mathbf{u}_j\otimes\mathbf{u}_k$$
$$\nabla T = \Sigma_{ijk}\, T'_{ij;k}\, \mathbf{e}^i\otimes\mathbf{e}^j\otimes\mathbf{e}^k \;=\; \Sigma_{ijk}\, [(\nabla T)^{(e)}]_{ijk}\, \mathbf{e}^i\otimes\mathbf{e}^j\otimes\mathbf{e}^k \;=\Sigma_{ijk}\, [(\nabla T)^{(e)}]^{ijk}\, \mathbf{e}_i\otimes\mathbf{e}_j\otimes\mathbf{e}_k\ . \qquad (J.3.1)$$

As usual, up and down index positions on T are the same for the first line in Cartesian space, but are significant on the second line (for non-orthogonal coordinates). The superscripts on the $(\nabla T)^{(u)}$ and $(\nabla T)^{(e)}$ components indicate which basis vectors are being expanded upon. One normally just writes $(\nabla T)^{(u)} = (\nabla T)$, and $(\nabla T)^{(e)} = (\nabla T)'$ where the prime indicates the curvilinear x'-space. So, we now have three different notations for the expansion components:

$$T_{ij;k} \;=\; [(\nabla T)^{(u)}]_{ijk} = (\nabla T)_{ijk}$$

$$T'_{ij;k} \;=\; [(\nabla T)^{(e)}]_{ijk} = (\nabla T)'_{ijk}\ . \qquad (J.3.2)$$

Just to fill things out, here are the corresponding expansions for rank-2 tensor T, again from (E.2.11) and (E.2.14),

$$T = \Sigma_{ij}\, T_{ij}\, \mathbf{u}^i\otimes\mathbf{u}^j \;=\; \Sigma_{ij}\, [T^{(u)}]_{ij}\, \mathbf{u}^i\otimes\mathbf{u}^j \;=\Sigma_{ij}\, [T^{(u)}]^{ij}\, \mathbf{u}_i\otimes\mathbf{u}_j \qquad T_{ij} = [T^{(u)}]_{ij}$$
$$T = \Sigma_{ij}\, T'_{ij}\, \mathbf{e}^i\otimes\mathbf{e}^j \;=\; \Sigma_{ij}\, [T^{(e)}]_{ij}\, e^i\otimes e^j \;=\Sigma_{ij}\, [T^{(e)}]^{ij}\, \mathbf{e}_i\otimes\mathbf{e}_j \qquad T'_{ij} = [T^{(e)}]_{ij} \qquad (J.3.3)$$

and then for a rank-1 tensor, as in (7.13.10),

$$\mathbf{v} = \Sigma_i\, v_i\, \mathbf{u}^i \;=\; \Sigma_i\, [v^{(u)}]_i\, \mathbf{u}^i \;=\Sigma_i\, [v^{(u)}]^i\, \mathbf{u}_i \;=\Sigma_i\, v^i\mathbf{u}_i \qquad v_i = [v^{(u)}]_i$$
$$\mathbf{v} = \Sigma_i\, v'_i\, \mathbf{e}^i \;=\; \Sigma_i\, [v^{(e)}]_i\, \mathbf{e}^i \;=\Sigma_i\, [v^{(e)}]^i\, \mathbf{e}_i \;=\Sigma_i\, v'^i\mathbf{e}_i \qquad v'_i = [v^{(e)}]_i\ . \qquad (J.3.4)$$

J.4 Tensor expansions of ∇T on the $\hat{e}_n$ base vectors

In practical applications, it is sometimes useful to deal with components of tensors which are expanded on the unit versions of the tangent base vectors, $\hat{\mathbf{e}}_n \equiv \mathbf{e}_n/|\mathbf{e}_n| = \mathbf{e}_n/h'_n$. On the one hand, this introduces major complications (see below) since such components are non-covariant (neither contravariant nor covariant nor mixed). On the other hand, since the unit vectors are all dimensionless, all tensor components have the same dimensions, which is very useful in any practical engineering work. We shall refer to such tensor components as "unit-base-vector components".

This subject is addressed below (E.2.14) and in Section E.8. We start with one of the (J.3.1) expansions,

$$\nabla T = \Sigma_{ijk}\ [(\nabla T)^{(e)}]^{ijk}\ \mathbf{e}_i\otimes\mathbf{e}_j\otimes\mathbf{e}_k\ \ = \Sigma_{ijk}\ (\nabla T)'^{ijk}\ \mathbf{e}_i\otimes\mathbf{e}_j\otimes\mathbf{e}_k \qquad (J.4.1)$$

and process it in this manner ,

$$= \Sigma_{ijk}\ [(\nabla T)^{(e)}]^{ijk}\ (h'_i\hat{\mathbf{e}}_i)\otimes(h'_j\hat{\mathbf{e}}_j)\otimes(h'_k\hat{\mathbf{e}}_k)$$

$$= \Sigma_{ijk}\ \{\,h'_i\,h'_j\,h'_k\,[(\nabla T)^{(e)}]^{ijk}\,\}\ \ \hat{\mathbf{e}}_i\otimes\hat{\mathbf{e}}_j\otimes\hat{\mathbf{e}}_k$$

$$= \Sigma_{ijk}\ \ [(\nabla T)^{(\hat{e})}]^{ijk}\ \ \hat{\mathbf{e}}_i\otimes\hat{\mathbf{e}}_j\otimes\hat{\mathbf{e}}_k \quad \text{where} \quad [(\nabla T)^{(\hat{e})}]^{ijk} = h'_i\,h'_j\,h'_k\,[(\nabla T)^{(e)}]^{ijk}\ . \qquad (J.4.2)$$

If our x'-space coordinate system happens NOT to be orthogonal, then the up/down position of the indices on $[(\nabla T)^{(\hat{e})}]^{ijk}$ has significance. This is so because, as shown near (E.2.11), these indices are lowered by $w'_{nm} = \hat{\mathbf{e}}_n \bullet \hat{\mathbf{e}}_m$. For an orthogonal system, one has $w'_{nm} = \hat{\mathbf{e}}_n \bullet \hat{\mathbf{e}}_m = \delta_{n,m}$ and then the up/down positions of the indices on $[(\nabla T)^{(\hat{e})}]^{ijk}$ all indicate the same number. One should not confuse this with the fact that the up/down position of an index on any true x'-space tensor like $(\nabla T)'^{ijk}$ is always significant, whether or not a coordinate system is orthogonal, because in general $g'_{ab} \neq \delta_{a,b}$.

Now for convenience later, we make one more definition,

$$(\nabla\mathcal{T})'^{ijk}\ \ \equiv [(\nabla T)^{(\hat{e})}]^{ijk} = h'_i\,h'_j\,h'_k\,[(\nabla T)^{(e)}]^{ijk}\ \ =\ h'_i\,h'_j\,h'_k\,(\nabla T)'^{ijk} \qquad (J.4.3)$$

where we follow our convention that the components of unit-base-vector tensors are written in script, see (E.8.5). (The symbol $\mathcal{T}$ is a script T , not a "tau" τ.)

Similarly, we can write the unit-base-vector expansion for a rank-2 tensor T,

$$T = \Sigma_{ij}\ \ [T^{(\hat{e})}]^{ij}\ \ \hat{\mathbf{e}}_i\otimes\hat{\mathbf{e}}_j \quad \text{where} \quad [T^{(\hat{e})}]^{ij} = h'_i\,h'_j\ \ [T^{(e)}]^{ij} \qquad (J.4.4)$$

with the definition

$$\mathcal{T}'^{ij} \equiv [T^{(\hat{e})}]^{ij} = h'_i\,h'_j\ \ [T^{(e)}]^{ij}\ \ = h'_i\,h'_j\,T'^{ij}\ \ . \qquad (J.4.5)$$

Finally, for a vector **v**,

$$\mathbf{v} = \Sigma_i\ [v^{(\hat{e})}]^i\ \hat{\mathbf{e}}_i \quad \text{where} \quad [v^{(\hat{e})}]^i = h'_i\ [v^{(e)}]^i \tag{J.4.6}$$

with the definition

$$\mathcal{v}'^i \equiv [v^{(\hat{e})}]^i = h'_i\ [v^{(e)}]^i = h'_i\ v'^i\ . \tag{J.4.7}$$

In terms of the scripted unit-base-vector components, our three expansions are:

$$\begin{aligned} \nabla T &= \Sigma_{ijk}\ (\nabla\mathcal{T})'^{ijk}\ \hat{\mathbf{e}}_i\otimes\hat{\mathbf{e}}_j\otimes\hat{\mathbf{e}}_k & (\nabla\mathcal{T})'^{ijk} &= h'_i\ h'_j\ h'_k\ (\nabla T)'^{ijk} \\ T &= \Sigma_{ij}\ \mathcal{T}'^{ij}\ \hat{\mathbf{e}}_i\otimes\hat{\mathbf{e}}_j & \mathcal{T}'^{ij} &= h'_i\ h'_j\ T'^{ij} \\ \mathbf{v} &= \Sigma_i\ \mathcal{v}'^i\ \hat{\mathbf{e}}_i & \mathcal{v}'^i &= h'_i\ v'^i\ . \end{aligned} \tag{J.4.8}$$

The scripted tensor components have indices which are integers, i = 1,2...N. Once we actually *select* a particular curvilinear coordinate system, one can replace the indices with curvilinear coordinate names and then, since those names *indicate* that one is talking about x'-space components, and since we just assume we are dealing with unit-base-vector components, both the script and the prime position can be dropped. For example, in spherical coordinates with 1,2,3 = r,θ,φ one can write

$$\begin{aligned} &(\nabla\mathcal{T})'^{123} = (\nabla T)^{r\theta\varphi} \\ &\mathcal{T}'^{12} = T^{r\theta} \qquad \mathcal{T}'^{11} = T^{rr} \\ &\mathcal{v}'^3 = v^\varphi \qquad \mathcal{v}'^1 = v^r \end{aligned} \tag{J.4.9}$$

As noted below (J.4.2), if a curvilinear coordinate system is orthogonal (and this is generally the case, and is certainly the case for spherical coordinates), then the up/down position does not matter and one writes

$$\begin{aligned} &(\nabla\mathcal{T})'^{123} = (\nabla\mathcal{T})'_{123} = (\nabla T)_{r\theta\varphi} \\ &\mathcal{T}'^{12} = T_{r\theta} \qquad \mathcal{T}'^{11} = T_{rr} \\ &\mathcal{v}'^3 = v_\varphi \qquad \mathcal{v}'^1 = v_r\ . \end{aligned} \tag{J.4.10}$$

Notice that the scripted forms are always necessary when summations like Σ_{ijk} are involved, unless one is willing to write out all the terms in the sum, which is a bit clumsy. The continuum mechanics book of Lai *et al.*, which we shall refer to below, avoids summations in curvilinear coordinates and thus has no need for our scripted components.

J.5 Total time derivative equation written in unit-base-vector curvilinear components

Recall from (J.2.3) our prototype time derivative equation of interest in x'-space,

$$d_t T'^{ij} = \partial_t T'^{ij} + (\nabla T)'^{ij}{}_k\ v'^k\ . \tag{J.5.1}$$

How does one write this equation in terms of unit-base-vector tensor components? We know that for orthogonal coordinates the answer is (J.1.7) based on covariance. Here we shall answer the question for general coordinates, and shall show that (J.1.7) "obtains" for orthogonal coordinates.

From (J.4.5) and (J.4.7) we had (no implied sums here)

$$\mathcal{T}'^{ij} = h'_i\, h'_j\, T'^{ij} \qquad\qquad \mathscr{v}'^k = h'_k\, v'^k \qquad\qquad (J.5.2)$$

so (J.5.1) can be processed as follows:

$$d_tT'^{ij} = \partial_tT'^{ij} + (\nabla T)'^{ij}{}_k\, v'^k \qquad\qquad // \text{ (J.5.1)}$$

$$h'_i\, h'_j\, d_tT'^{ij} = h'_i\, h'_j\, \partial_tT'^{ij} + h'_i\, h'_j\, (\nabla T)'^{ij}{}_k\, h'_k{}^{-1}\, h'_k\, v'^k \qquad\qquad // \text{ mult thru by } h'_i\, h'_j$$

$$d_t(h'_ih'_j\, T'^{ij}) = \partial_t(h'_ih'_j\, T'^{ij}) \;+ h'_i\, h'_j\, h'_k{}^{-1}\, (\nabla T)'^{ij}{}_k\, (h'_k\, v'^k) \qquad\qquad // \; h'_n = h'_n(\mathbf{x}'), \text{ no } t$$

$$d_t\, \mathcal{T}'^{ij} = \partial_t\, \mathcal{T}'^{ij} \;+ [\, h'_i\, h'_j\, h'_k{}^{-1}\, (\nabla T)'^{ij}{}_k\,]\, \mathscr{v}'^k \qquad\qquad // \text{ (J.5.2)}$$

$$d_t\, \mathcal{T}'^{ij} = \partial_t\, \mathcal{T}'^{ij} \;+ Q'^{ij}{}_k\, \mathscr{v}'^k \qquad\qquad Q'^{ij}{}_k \;\equiv\; h'_i\, h'_j\, h'_k{}^{-1}\, (\nabla T)'^{ij}{}_k \;. \qquad\qquad (J.5.3)$$

<u>Comment:</u> Note that $d_th'_n(\mathbf{x}) = \partial_th'_n(\mathbf{x}) = 0$ and not $d_th'_n(\mathbf{x}) = \partial_th'_n(\mathbf{x}) + (\nabla h'_n)\, \mathbf{v}$. The reason is that the field $h'_n(\mathbf{x})$ is not an Eulerian fluid property like A_1 in (J.6.9) below, it is a property of space at point $\mathbf{x}$.

Recall that no assumption has been made that the curvilinear coordinates are orthogonal. Section J.2 showed how the $(\nabla T)'^{ij}{}_k$ can be computed for any curvilinear coordinate system, and thus one can compute the Q'_{ijk} shown above. Our question is now answered: (J.5.3) shows how one writes the total time derivative equation in arbitrary curvilinear coordinates.

If we now *assume* the coordinates are orthogonal, then

$$(\nabla T)'^{ij}{}_k \quad = g_{kk'}(\nabla T)'^{ijk'} = \; h_k{}^2\delta_{k,k'}\, (\nabla T)'^{ijk'} = h_k{}^2(\nabla T)'^{ijk} \qquad\qquad (J.5.4)$$

and then

$$\begin{aligned} Q'^{ij}{}_k \;\equiv\; & h'_i\, h'_j\, h'_k{}^{-1}\, (\nabla T)'^{ij}{}_k \; = h'_i\, h'_j\, h'_k{}^{-1}\, h_k{}^2(\nabla T)'^{ijk} \\ & = h'_i\, h'_j\, h'_k(\nabla T)'^{ijk} \; = (\, \nabla\mathcal{T})^{ijk}\, . \qquad\qquad // \text{ (J.4.3)} \end{aligned} \qquad\qquad (J.5.5)$$

As noted below (J.4.2), for orthogonal coordinates the up/down index position on $(\,\nabla\mathcal{T})^{ijk}$ makes no difference, so for example, $(\,\nabla\mathcal{T})^{ijk} = \;(\,\nabla\mathcal{T})^{ij}{}_k$. Installing $Q'^{ij}{}_k = \;(\,\nabla\mathcal{T})^{ij}{}_k$ into (J.5.3) then gives,

$$d_t\, \mathcal{T}'^{ij} = \partial_t\, \mathcal{T}'^{ij} \;+ (\,\nabla\mathcal{T})'^{ij}{}_k\, \mathscr{v}'^k \qquad\qquad (\,\nabla\mathcal{T})'^{ij}{}_k \;\equiv\; h'_i\, h'_j\, h'_k\, (\nabla T)'^{ijk} \quad . \quad // \text{ orthogonal} \qquad (J.5.6)$$

which is (J.1.7) which was deduced by covariance. Once again, indices go up and down for free on $\mathcal{T}'^{ij}$ and $(\boldsymbol{\nabla}\mathcal{T})^{ijk}$, but not on T'^{ij} and $(\nabla T)'^{ijk}$. We shall assume from now on that the curvilinear coordinates are orthogonal.

As a reminder, here is what the above equation says if i = j = 1,

$$d_t\, \mathcal{T}'^{11} = \partial_t\, \mathcal{T}'^{11} + (\boldsymbol{\nabla}\mathcal{T})'^{11}{}_k\, \boldsymbol{\mathit{v}}'^k \quad . \tag{J.5.7}$$

Using the notation scheme just described above in (J.4.9) and (J.4.10), in spherical coordinates one would write the above equation as,

$$d_t\, T_{rr} = \partial_t\, T_{rr} + (\nabla T)_{rrr}\, v_r + (\nabla T)_{rr\theta}\, v_\theta + (\nabla T)_{rr\varphi}\, v_\varphi \quad . \tag{J.5.8}$$

J.6 Shorthand notations and a continuum mechanics application

Consider our original Cartesian time derivative equation (J.1.2),

$$d_t T^{ij} = \partial_t T^{ij} + (\nabla T)^{ij}{}_k\, v^k \quad . \tag{J.6.1}$$

One could *regard* $(\nabla T)^{ij}{}_k$ as the kth component of a vector $(\nabla T)^{ij}$ labeled by fixed values i and j, so that

$$[(\nabla \mathbf{T})^{ij}]_k = (\nabla T)^{ij}{}_k \quad . \tag{J.6.2}$$

Then one can write (J.6.1) as

$$d_t T^{ij} = \partial_t T^{ij} + (\nabla \mathbf{T})^{ij} \bullet \mathbf{v} \quad . \tag{J.6.3}$$

The next step is to suppress the ij labels, since the equation above is true for any i and j,

$$d_t T = \partial_t T + (\nabla \mathbf{T}) \bullet \mathbf{v} \quad . \tag{J.6.4}$$

This is *only a shorthand notation*, no precision justification is required. We can do the same thing to the equation written for orthogonal unit-base-vector curvilinear coordinates :

$$d_t\, \mathcal{T}'^{ij} = \partial_t\, \mathcal{T}'^{ij} + (\boldsymbol{\nabla}\mathcal{T})'^{ij}{}_k\, \boldsymbol{\mathit{v}}'^k \qquad \textit{(J.5.6)}$$

$$[(\boldsymbol{\nabla}\mathcal{T})'^{ij}]_k = (\boldsymbol{\nabla}\mathcal{T})'^{ij}{}_k \qquad \text{// vector with index k}$$

$$d_t\, \mathcal{T}'^{ij} = \partial_t\, \mathcal{T}'^{ij} + (\boldsymbol{\nabla}\mathcal{T})'_{ij} \bullet \boldsymbol{\mathit{v}}'_k \qquad \text{// dot notation}$$

$$d_t\, \mathcal{T}' = \partial_t\, \mathcal{T}' + (\boldsymbol{\nabla}\mathcal{T})' \bullet \boldsymbol{\mathit{v}}' \qquad \text{// indices stripped} \tag{J.6.5}$$

and then we have

$$d_t T \;=\; \partial_t T + \;(\nabla \mathbf{T})\; \bullet \mathbf{v}$$
$$d_t\, \mathcal{T}' = \;\partial_t \mathcal{T}' + (\,\boldsymbol{\nabla}\mathcal{T})' \bullet\; \boldsymbol{\upsilon}'\;\; . \qquad\qquad (J.6.6)$$

In the book of Lai *et al*., the • is omitted and the shorthand notations are written,

$$d_t T \;=\; \partial_t T \;+\; (\nabla T)\, \mathbf{v}$$
$$d_t\, \mathcal{T}' = \;\partial_t \mathcal{T}' \;\; + (\,\boldsymbol{\nabla}\mathcal{T})'\, \boldsymbol{\upsilon}' \qquad\qquad (J.6.7)$$

where one imagines that (∇T) and ($\boldsymbol{\nabla}\mathcal{T}$)' are 3-index operators which act on a 1-index object to generate a 2-index object. Again, it is just a shorthand. The real meaning of these equations is

$$d_t T^{ij} \;\;= \partial_t T^{ij} \;\;+\;\; (\nabla T)^{ij}{}_{k}\, v^k$$
$$d_t\, \mathcal{T}'^{ij} = \partial_t \mathcal{T}'^{ij} \;\; + (\,\boldsymbol{\nabla}\mathcal{T})'^{ij}{}_{k}\, \boldsymbol{\upsilon}'^{k}\; . \qquad\qquad (J.6.8)$$

Example: An application of our total time derivative equation appears in the first line of Lai p 470 ,

$$DA_1/Dt \;\;= \partial A_1/\partial_t + (\nabla A_1)\; \mathbf{v} \qquad\qquad (J.6.9)$$

where D/Dt is the way a total time derivative is expressed in continuum mechanics (D/Dt = d/dt). It is the convective or material derivative for Eulerian-picture functions (like $A_1(\mathbf{x},t)$), meaning that the second term registers a time change in a fluid property at the fixed point **x** due to "new fluid" with velocity **v** passing through that point (a point being a tiny differential volume of fluid). In the notation of (J.6.7), this would be written

$$d_t \mathcal{A}'_1 \;\;= \partial_t \mathcal{A}'_1 + (\,\boldsymbol{\nabla}\mathcal{A}_1)'\; \boldsymbol{\upsilon}'\;\; . \qquad\qquad (J.6.10)$$

The detailed meaning is

$$d_t\, (\mathcal{A}'_1)^{ij} = \partial_t\, (\mathcal{A}'_1)^{ij} \;\; + (\,\boldsymbol{\nabla}\mathcal{A}_1)'^{ij}{}_{k}\; \boldsymbol{\upsilon}'^{k} \qquad\qquad (J.6.11)$$

and for i = j = 1 this says (spherical coordinates) ,

$$d_t[(A_1)_{rr}] = \partial_t[(A_1)_{rr}] \;\; + (\nabla A_1)_{rrr}\, v_r + (\nabla A_1)_{rr\theta}\, v_\theta + (\nabla A_1)_{rr\varphi}\, v_\varphi \;\; . \qquad\qquad (J.6.12)$$

The tensor A_1 is the "**first Rivlin-Ericksen tensor**" associated with the flow of non-Newtonian fluids (rheology). The A_i tensors, briefly mentioned in Section K.4, are derivatives of a certain deformation tensor called C_t, and computation of the A_i is done in the following iterative manner,

$$A_{i+1} = D_t A_i + A_i(\nabla \mathbf{v}) + (\nabla \mathbf{v})^T A_i \;\; , \qquad\qquad // \text{ Lai p 468 (8.11.3)} \qquad\qquad (J.6.13)$$

where (∇**v**) is the gradient-of-vector object treated in our Appendix G, **v** being the fluid velocity field. In particular, $A_2 = D_t A_1 + A_1(\nabla \mathbf{v}) + (\nabla \mathbf{v})^T A_1$ which involves our object of interest $D_t A_1 = d_t A_1$. So, in order to compute A_2 in curvilinear coordinates, one needs $d_t(\mathcal{A}'_1)_{ij}$ which involves the $(\,\boldsymbol{\nabla}\mathcal{A}_1)'_{ijk}$. In order to use the above equation in practice, one has to *know* for example that $(\nabla A_1)_{rr\theta} \;= [\; \partial_\theta(A_1)_{rr} - (A_1)_{\theta r} - (A_1)_{r\theta}]/r$, a fact that is certainly not immediately obvious (but see the rrθ entry in (J.7.9) below).

So, our next task is to *compute* the $(\nabla\mathcal{T})'_{ijk}$ which appear in our prototype equation above,

$$d_t\,\mathcal{T}' = \partial_t\,\mathcal{T}' + (\nabla\mathcal{T})'\,\boldsymbol{u}' \qquad \text{// shorthand notation (J.6.7)}$$

$$d_t\,\mathcal{T}'^{ij} = \partial_t\,\mathcal{T}'^{ij} + (\nabla\mathcal{T})'^{ij}{}_k\,\boldsymbol{u}'^k \qquad \textit{(J.5.6)} \qquad \text{(J.6.14)}$$

J.7 Maple computation of the $(\nabla\mathcal{T})'_{ijk}$ components for spherical coordinates

Recall from (J.4.3) that, for arbitrary curvilinear coordinates,

$$(\nabla\mathcal{T})'^{ijk} \equiv h'_i\,h'_j\,h'_k\,(\nabla T)'^{ijk} \qquad \text{(J.7.1)}$$

and from (J.2.2),

$$(\nabla T)'_{abc} = \partial'_c T'_{ab} - \Gamma'^n{}_{ac}T'_{nb} - \Gamma'^n{}_{bc}T'_{an} \qquad \text{(J.7.2)}$$

$$\Gamma'^d{}_{ab} = (1/2)\,g'^{dc}\,[\,\partial'_a g'_{bc} + \partial'_b g'_{ca} - \partial'_c g'_{ab}] \quad . \qquad g' = \text{metric tensor for x'-space}$$

But we are now assuming only <u>orthogonal</u> coordinates, so

$$g'_{ab} = \delta_{a,b}h_a{}^2 \qquad\qquad g'^{ab} = \delta_{a,b}h_a{}^{-2} \qquad \text{(J.7.3)}$$

$$\Rightarrow \quad (\nabla T)'^{ijk} = g'^{ii'}g'^{jj'}g'^{kk'}(\nabla T)'_{i'j'k'} = (h'_i\,h'_j\,h'_k)^{-2}\,(\nabla T)'_{ijk} \qquad \text{(J.7.4)}$$

and then from (J.7.1),

$$(\nabla\mathcal{T})'^{ijk} = (\nabla\mathcal{T})'_{ijk} = (h'_i\,h'_j\,h'_k)^{-1}\,(\nabla T)'_{ijk} \ . \qquad \text{(J.7.5)}$$

Remember that up/down index position does not matter on $(\nabla\mathcal{T})'^{ijk}$ for orthogonal coordinates.

We want the result expressed in terms of the $\mathcal{T}'_{ij}$ and not the T'_{ij}, so from (J.4.8),

$$\mathcal{T}'^{ij} = (h'_i\,h'_j)\,T'^{ij} = (h'_i\,h'_j)\,g'^{ii'}\,g'^{jj'}T'_{i'j'} = (h'_i\,h'_j)^{-1}\,T'_{ij}$$

$$\Rightarrow \quad T'_{ij} = \ (h'_i\,h'_j\mathcal{T}'^{ij}) \ . \qquad \text{(J.7.6)}$$

Then (J.7.2) reads,

$$(\nabla T)'_{abc} = \partial'_c T'_{ab} - \Gamma'^n{}_{ac}T'_{nb} - \Gamma'^n{}_{bc}T'_{an} \qquad \text{// do abc→ijk rename}$$

$$(\nabla T)'_{ijk} = \partial'_k T'_{ij} - \Gamma'^n{}_{ik}T'_{nj} - \Gamma'^n{}_{jk}T'_{in}$$

$$(\nabla T)'_{ijk} = \partial'_k(h'_i\,h'_j\mathcal{T}'_{ij}) - \Gamma'^n{}_{ik}(h'_n\,h'_j\mathcal{T}'_{nj}) - \Gamma'^n{}_{jk}(h'_i\,h'_n\mathcal{T}'_{in}) \qquad \text{// (J.7.6)}$$

$$= h'_i h'_j(\partial'_k\mathcal{T}'_{ij}) + \partial'_k(h'_i h'_j)\,\mathcal{T}'_{ij} - \Gamma'^n{}_{ik}(h'_n h'_j\mathcal{T}'_{nj}) - \Gamma'^n{}_{jk}(h'_i h'_n\mathcal{T}'_{in}) \qquad \text{(J.7.7)}$$

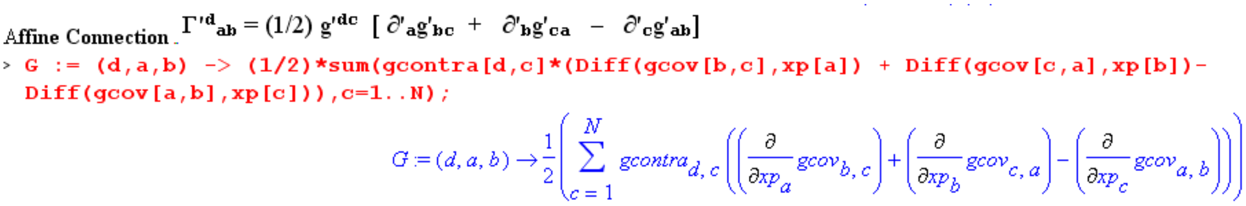
where we break the ∂'_k term in two pieces for Maple technical reasons.

So the Maple program will compute from (J.7.5) and (J.7.7),

$$(\nabla\mathcal{T})'_{ijk} = (h'_i h'_j h'_k)^{-1}\,(\nabla T)'_{ijk} \qquad \text{(J.7.8)}$$

$$\text{where } (\nabla T)'_{ijk} = \underset{T1}{h'_i h'_j(\partial'_k\mathcal{T}'_{ij})} + \underset{T2}{\partial'_k(h'_i h'_j)\,\mathcal{T}'_{ij}} - \underset{T3}{\Gamma'^n{}_{ik}(h'_n h'_j\mathcal{T}'_{nj})} - \underset{T4}{\Gamma'^n{}_{jk}(h'_i h'_n\mathcal{T}'_{in})}\ .$$

For now we assume spherical coordinates. Maple first computes g_{**} = gcov using (I.3.2) and g^{**} = gcontra using (I.7.5). Then the affine connection is computed from these metric tensors, just as in (I.7.5),



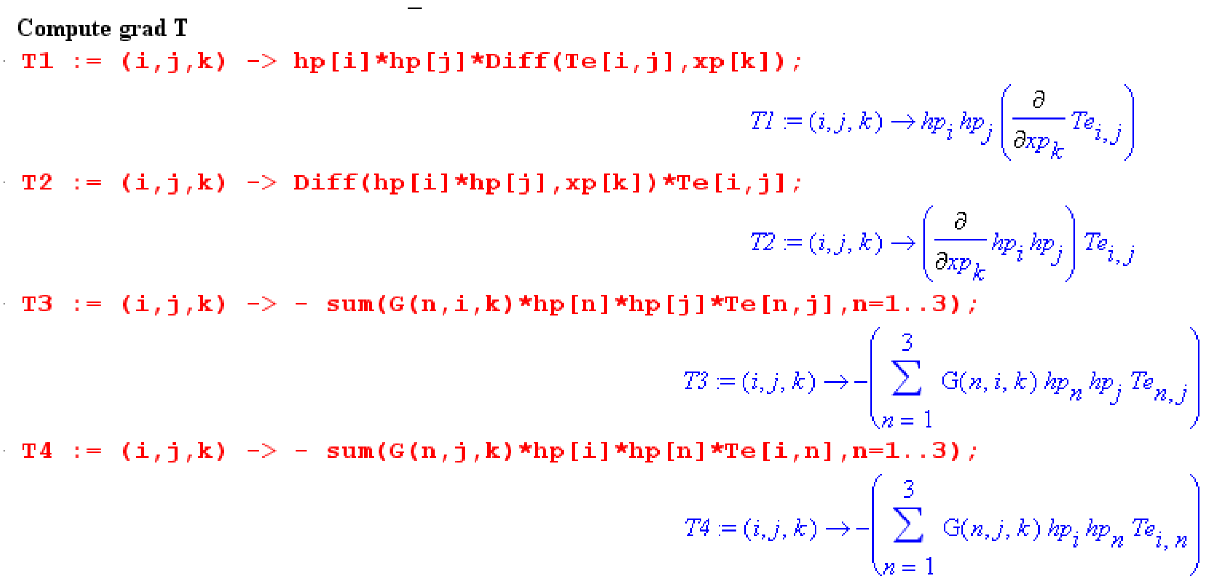
Affine Connection $\Gamma'^d{}_{ab} = (1/2)\, g'^{dc}\,[\,\partial'_a g'_{bc} + \partial'_b g'_{ca} - \partial'_c g'_{ab}]$

```
> G := (d,a,b) -> (1/2)*sum(gcontra[d,c]*(Diff(gcov[b,c],xp[a]) + Diff(gcov[c,a],xp[b])-
  Diff(gcov[a,b],xp[c])),c=1..N);
```

$$G := (d,a,b) \rightarrow \frac{1}{2}\left(\sum_{c=1}^{N} gcontra_{d,c}\left(\left(\frac{\partial}{\partial xp_a} gcov_{b,c}\right) + \left(\frac{\partial}{\partial xp_b} gcov_{c,a}\right) - \left(\frac{\partial}{\partial xp_c} gcov_{a,b}\right)\right)\right)$$

(J.7.9)

Then the terms of (J.7.8) are entered ($\mathcal{T}'_{ij}$ = Te[i,j], h'_i = hp[i], $\Gamma'^d{}_{ab}$ = G(d,a,b), etc.)



Compute grad T

```
T1 := (i,j,k) -> hp[i]*hp[j]*Diff(Te[i,j],xp[k]);
```

$$T1 := (i,j,k) \rightarrow hp_i\, hp_j \left(\frac{\partial}{\partial xp_k} Te_{i,j}\right)$$

```
T2 := (i,j,k) -> Diff(hp[i]*hp[j],xp[k])*Te[i,j];
```

$$T2 := (i,j,k) \rightarrow \left(\frac{\partial}{\partial xp_k} hp_i\, hp_j\right) Te_{i,j}$$

```
T3 := (i,j,k) -> - sum(G(n,i,k)*hp[n]*hp[j]*Te[n,j],n=1..3);
```

$$T3 := (i,j,k) \rightarrow -\left(\sum_{n=1}^{3} \mathrm{G}(n,i,k)\, hp_n\, hp_j\, Te_{n,j}\right)$$

```
T4 := (i,j,k) -> - sum(G(n,j,k)*hp[i]*hp[n]*Te[i,n],n=1..3);
```

$$T4 := (i,j,k) \rightarrow -\left(\sum_{n=1}^{3} \mathrm{G}(n,j,k)\, hp_i\, hp_n\, Te_{i,n}\right)$$

(J.7.10)

The terms are then added, multiplied by $(h'_i\ h'_j\ h'_k)^{-1}$ according to (J.7.8), and then displayed. In this code then, DT(i,j,k) is $(\nabla\mathcal{T})'_{ijk}$ as computed on the first line. In the triple loop, the complicated resulting expressions are cleaned up in a series of cosmetic trigonometric filtering steps: DT(i,j,k)→ f → g → h[k] and it is then the simplified h[k] expressions which are printed out.

```
> DT  := (i,j,k) -> (hp[i]*hp[j]*hp[k])^(-1)*( T1(i,j,k) + value(T2(i,j,k)) +
  value(T3(i,j,k))+ value(T4(i,j,k)) ;
```

$$DT := (i, j, k) \rightarrow \frac{\mathrm{T1}(i, j, k) + \mathrm{value}(\mathrm{T2}(i, j, k)) + \mathrm{value}(\mathrm{T3}(i, j, k)) + \mathrm{value}(\mathrm{T4}(i, j, k))}{hp_i\, hp_j\, hp_k}$$

```
> for i from 1 to 3 do
    for j from 1 to 3 do
      for k from 1 to 3 do
        f := DT(i,j,k);
        g := subs(cos(theta)^2= 1-sin(theta)^2,f);
        h[k] := subs(cos(theta) = cot(theta)*sin(theta), expand(g));
      od;
      print(xp[i],xp[j],xp[1]," = ",h[1],"     ",xp[i],xp[j],xp[2]," = ",h[2],"
  ",xp[i],xp[j],xp[3]," = ",h[3]);
    od
  od;
```

(J.7.11)

and here are the resulting values for $(\nabla\mathcal{T})'_{ijk}$ (for example, $(\nabla\mathcal{T})'_{112} = (\nabla T)_{rr\theta}$)

$$r, r, r, \text{" = "}, \frac{\partial}{\partial r} T_{r,r}, \text{"   "}, r, r, \theta, \text{" = "}, \frac{\frac{\partial}{\partial\theta} T_{r,r}}{r} - \frac{T_{\theta,r}}{r} - \frac{T_{r,\theta}}{r}, \text{"   "}, r, r, \phi, \text{" = "}, \frac{\frac{\partial}{\partial\phi} T_{r,r}}{r \sin(\theta)} - \frac{T_{\phi,r}}{r} - \frac{T_{r,\phi}}{r}$$

$$r, \theta, r, \text{" = "}, \frac{\partial}{\partial r} T_{r,\theta}, \text{"   "}, r, \theta, \theta, \text{" = "}, \frac{\frac{\partial}{\partial\theta} T_{r,\theta}}{r} - \frac{T_{\theta,\theta}}{r} + \frac{T_{r,r}}{r}, \text{"   "}, r, \theta, \phi, \text{" = "}, \frac{\frac{\partial}{\partial\phi} T_{r,\theta}}{r \sin(\theta)} - \frac{T_{\phi,\theta}}{r} - \frac{\cot(\theta)\, T_{r,\phi}}{r}$$

$$r, \phi, r, \text{" = "}, \frac{\partial}{\partial r} T_{r,\phi}, \text{"   "}, r, \phi, \theta, \text{" = "}, \frac{\frac{\partial}{\partial\theta} T_{r,\phi}}{r} - \frac{T_{\theta,\phi}}{r}, \text{"   "}, r, \phi, \phi, \text{" = "}, \frac{\frac{\partial}{\partial\phi} T_{r,\phi}}{r \sin(\theta)} - \frac{T_{\phi,\phi}}{r} + \frac{T_{r,r}}{r} + \frac{\cot(\theta)\, T_{r,\theta}}{r}$$

$$\theta, r, r, \text{" = "}, \frac{\partial}{\partial r} T_{\theta,r}, \text{"   "}, \theta, r, \theta, \text{" = "}, \frac{\frac{\partial}{\partial\theta} T_{\theta,r}}{r} + \frac{T_{r,r}}{r} - \frac{T_{\theta,\theta}}{r}, \text{"   "}, \theta, r, \phi, \text{" = "}, \frac{\frac{\partial}{\partial\phi} T_{\theta,r}}{r \sin(\theta)} - \frac{\cot(\theta)\, T_{\phi,r}}{r} - \frac{T_{\theta,\phi}}{r}$$

$$\theta, \theta, r, \text{" = "}, \frac{\partial}{\partial r} T_{\theta,\theta}, \text{"   "}, \theta, \theta, \theta, \text{" = "}, \frac{\frac{\partial}{\partial\theta} T_{\theta,\theta}}{r} + \frac{T_{r,\theta}}{r} + \frac{T_{\theta,r}}{r}, \text{"   "}, \theta, \theta, \phi, \text{" = "}, \frac{\frac{\partial}{\partial\phi} T_{\theta,\theta}}{r \sin(\theta)} - \frac{\cot(\theta)\, T_{\phi,\theta}}{r} - \frac{\cot(\theta)\, T_{\theta,\phi}}{r}$$

$$\theta, \phi, r, \text{" = "}, \frac{\partial}{\partial r} T_{\theta,\phi}, \text{"   "}, \theta, \phi, \theta, \text{" = "}, \frac{\frac{\partial}{\partial\theta} T_{\theta,\phi}}{r} + \frac{T_{r,\phi}}{r}, \text{"   "}, \theta, \phi, \phi, \text{" = "}, \frac{\frac{\partial}{\partial\phi} T_{\theta,\phi}}{r \sin(\theta)} - \frac{\cot(\theta)\, T_{\phi,\phi}}{r} + \frac{T_{\theta,r}}{r} + \frac{\cot(\theta)\, T_{\theta,\theta}}{r}$$

$$\phi, r, r, \text{" = "}, \frac{\partial}{\partial r} T_{\phi,r}, \text{"   "}, \phi, r, \theta, \text{" = "}, \frac{\frac{\partial}{\partial\theta} T_{\phi,r}}{r} - \frac{T_{\phi,\theta}}{r}, \text{"   "}, \phi, r, \phi, \text{" = "}, \frac{\frac{\partial}{\partial\phi} T_{\phi,r}}{r \sin(\theta)} + \frac{T_{r,r}}{r} + \frac{\cot(\theta)\, T_{\theta,r}}{r} - \frac{T_{\phi,\phi}}{r}$$

$$\phi, \theta, r, \text{" = "}, \frac{\partial}{\partial r} T_{\phi,\theta}, \text{"   "}, \phi, \theta, \theta, \text{" = "}, \frac{\frac{\partial}{\partial\theta} T_{\phi,\theta}}{r} + \frac{T_{\phi,r}}{r}, \text{"   "}, \phi, \theta, \phi, \text{" = "}, \frac{\frac{\partial}{\partial\phi} T_{\phi,\theta}}{r \sin(\theta)} + \frac{T_{r,\theta}}{r} + \frac{\cot(\theta)\, T_{\theta,\theta}}{r} - \frac{\cot(\theta)\, T_{\phi,\phi}}{r}$$

$$\phi, \phi, r, \text{" = "}, \frac{\partial}{\partial r} T_{\phi,\phi}, \text{"   "}, \phi, \phi, \theta, \text{" = "}, \frac{\frac{\partial}{\partial\theta} T_{\phi,\phi}}{r}, \text{"   "}, \phi, \phi, \phi, \text{" = "}, \frac{\frac{\partial}{\partial\phi} T_{\phi,\phi}}{r \sin(\theta)} + \frac{T_{r,\phi}}{r} + \frac{\cot(\theta)\, T_{\theta,\phi}}{r} + \frac{T_{\phi,r}}{r} + \frac{\cot(\theta)\, T_{\phi,\theta}}{r}$$

$(\nabla\mathcal{T})'_{ijk}$ for Spherical Coordinates (J.7.12)

The strange " symbols in the above table should be ignored, just a Maple print command artifact. These results agree with the spherical coordinates table given in Lai p 505. The divT results of Appendix H can be verified from the above table using

$$(\mathbf{div}\mathcal{T})'_{\mathtt{i}} = \partial'_{\mathtt{j}}\mathcal{T}'_{\mathtt{ij}} = (\nabla\mathcal{T})'_{\mathtt{ijj}} = \quad \Sigma_{\mathtt{j}}(\nabla\mathcal{T})'_{\mathtt{ijj}} \ . \qquad \text{(J.7.13)}$$

For example,

$$(\mathrm{divT})_{\mathbf{r}} = (\nabla\mathrm{T})_{\mathbf{rrr}} + (\nabla\mathrm{T})_{\mathbf{r\theta\theta}} + (\nabla\mathrm{T})_{\mathbf{r\varphi\varphi}}$$

$$= \partial_{\mathbf{r}}\mathrm{T}_{\mathbf{rr}} + (1/r)[\partial_{\theta}\mathrm{T}_{\mathbf{r\theta}} - \mathrm{T}_{\theta\theta} + \mathrm{T}_{\mathbf{rr}}] + (1/r)[\ \csc\theta\ \partial_{\varphi}\mathrm{T}_{\mathbf{r\varphi}} - \mathrm{T}_{\varphi\varphi} + \mathrm{T}_{\mathbf{rr}} + \cot\theta\ \mathrm{T}_{\mathbf{r\theta}}]$$

$$= \partial_{\mathbf{r}}\mathrm{T}_{\mathbf{rr}} + (1/r)\partial_{\theta}\mathrm{T}_{\mathbf{r\theta}} - \mathrm{T}_{\theta\theta}/r + (2/r)\mathrm{T}_{\mathbf{rr}} + (1/r\sin\theta)\ \partial_{\varphi}\mathrm{T}_{\mathbf{r\varphi}} - \mathrm{T}_{\varphi\varphi}/r + \cot\theta\ \mathrm{T}_{\mathbf{r\theta}}\ /r \qquad \text{(J.7.14)}$$

and we quote from (H.6.3),

$$divT_1 := \left(\frac{\partial}{\partial r}T_{r,\,r}\right) + \frac{\frac{\partial}{\partial\theta}T_{r,\,\theta}}{r} + \frac{\frac{\partial}{\partial\phi}T_{r,\,\phi}}{r\sin(\theta)} - \frac{T_{\theta,\,\theta}}{r} - \frac{T_{\phi,\,\phi}}{r} + 2\,\frac{T_{r,\,r}}{r} + \frac{\cos(\theta)\,T_{r,\,\theta}}{r\sin(\theta)} \ . \qquad \textit{(H.6.3)}$$

J.8 Maple computation of the $(\nabla\mathcal{T})'_{ijk}$ components for cylindrical coordinates

Making four small edits to the spherical coordinates Maple program converts it to a cylindrical coordinates program. Here are the resulting values for $(\nabla\mathcal{T})'_{ijk}$ (for example, $(\nabla\mathcal{T})'_{123} = (\nabla T)_{r\theta z}$)

$$r, r, r, \text{"="}, \frac{\partial}{\partial r} T_{r, r}, \text{"} \quad \text{"}, r, r, \theta, \text{"="}, \frac{\frac{\partial}{\partial \theta} T_{r, r}}{r} - \frac{T_{\theta, r}}{r} - \frac{T_{r, \theta}}{r}, \text{"} \quad \text{"}, r, r, z, \text{"="}, \frac{\partial}{\partial z} T_{r, r}$$

$$r, \theta, r, \text{"="}, \frac{\partial}{\partial r} T_{r, \theta}, \text{"} \quad \text{"}, r, \theta, \theta, \text{"="}, \frac{\frac{\partial}{\partial \theta} T_{r, \theta}}{r} - \frac{T_{\theta, \theta}}{r} + \frac{T_{r, r}}{r}, \text{"} \quad \text{"}, r, \theta, z, \text{"="}, \frac{\partial}{\partial z} T_{r, \theta}$$

$$r, z, r, \text{"="}, \frac{\partial}{\partial r} T_{r, z}, \text{"} \quad \text{"}, r, z, \theta, \text{"="}, \frac{\frac{\partial}{\partial \theta} T_{r, z}}{r} - \frac{T_{\theta, z}}{r}, \text{"} \quad \text{"}, r, z, z, \text{"="}, \frac{\partial}{\partial z} T_{r, z}$$

$$\theta, r, r, \text{"="}, \frac{\partial}{\partial r} T_{\theta, r}, \text{"} \quad \text{"}, \theta, r, \theta, \text{"="}, \frac{\frac{\partial}{\partial \theta} T_{\theta, r}}{r} + \frac{T_{r, r}}{r} - \frac{T_{\theta, \theta}}{r}, \text{"} \quad \text{"}, \theta, r, z, \text{"="}, \frac{\partial}{\partial z} T_{\theta, r}$$

$$\theta, \theta, r, \text{"="}, \frac{\partial}{\partial r} T_{\theta, \theta}, \text{"} \quad \text{"}, \theta, \theta, \theta, \text{"="}, \frac{\frac{\partial}{\partial \theta} T_{\theta, \theta}}{r} + \frac{T_{r, \theta}}{r} + \frac{T_{\theta, r}}{r}, \text{"} \quad \text{"}, \theta, \theta, z, \text{"="}, \frac{\partial}{\partial z} T_{\theta, \theta}$$

$$\theta, z, r, \text{"="}, \frac{\partial}{\partial r} T_{\theta, z}, \text{"} \quad \text{"}, \theta, z, \theta, \text{"="}, \frac{\frac{\partial}{\partial \theta} T_{\theta, z}}{r} + \frac{T_{r, z}}{r}, \text{"} \quad \text{"}, \theta, z, z, \text{"="}, \frac{\partial}{\partial z} T_{\theta, z}$$

$$z, r, r, \text{"="}, \frac{\partial}{\partial r} T_{z, r}, \text{"} \quad \text{"}, z, r, \theta, \text{"="}, \frac{\frac{\partial}{\partial \theta} T_{z, r}}{r} - \frac{T_{z, \theta}}{r}, \text{"} \quad \text{"}, z, r, z, \text{"="}, \frac{\partial}{\partial z} T_{z, r}$$

$$z, \theta, r, \text{"="}, \frac{\partial}{\partial r} T_{z, \theta}, \text{"} \quad \text{"}, z, \theta, \theta, \text{"="}, \frac{\frac{\partial}{\partial \theta} T_{z, \theta}}{r} + \frac{T_{z, r}}{r}, \text{"} \quad \text{"}, z, \theta, z, \text{"="}, \frac{\partial}{\partial z} T_{z, \theta}$$

$$z, z, r, \text{"="}, \frac{\partial}{\partial r} T_{z, z}, \text{"} \quad \text{"}, z, z, \theta, \text{"="}, \frac{\frac{\partial}{\partial \theta} T_{z, z}}{r}, \text{"} \quad \text{"}, z, z, z, \text{"="}, \frac{\partial}{\partial z} T_{z, z}$$

$(\nabla\mathcal{T})'_{ijk}$ for Cylindrical Coordinates (J.8.1)

and these expressions agree with those in the table on page 504 of Lai. The same comment made about divT at the end of the previous Section applies here as well. The object $\Gamma'^{d}{}_{ab}$ has 3 of 27 components non zero as shown in (I.8.11), only 2 of which are distinct.

It should be emphasized that this same simple Maple code can be used to compute the $(\nabla\mathcal{T})'_{ijk}$ for *any* system of orthogonal curvilinear coordinates in any number of dimensions N. For non-orthogonal coordinates the results would be obtained from these more general equations,

$$(\nabla\mathcal{T})'^{ijk} = h'_i \, h'_j \, h'_k \, (\nabla T)'^{ijk} \tag*{(J.7.1)}$$

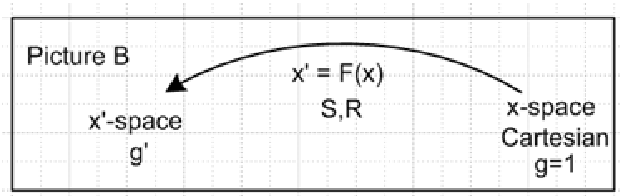
$$(\nabla T)'^{ijk} = g'^{ai} \, g'^{bj} \, g'^{ck} \, (\nabla T)'_{abc} \qquad \text{// raise indices}$$

$$(\nabla T)'_{abc} = \partial'_c T'_{ab} - \Gamma'^n{}_{ac} T'_{nb} - \Gamma'^n{}_{bc} T'_{an} \tag*{(J.7.2)}$$

$$\Gamma'^d{}_{ab} = (1/2) \, g'^{dc} \, [\, \partial'_a g'_{bc} + \partial'_b g'_{ca} - \partial'_c g'_{ab}] \quad . \qquad g' = \text{metric tensor for x'-space} \tag*{(J.8.2)}$$

J.9 The Lai Method of computing $(\nabla\mathcal{T})'_{ijk}$ for orthogonal coordinates

This method appears in Lai pp 501-505. It differs from the method given above in Section J.8 mainly because it uses a different affine connection, but of course it gives the same results. We shall present Lai's computation of the components $(\nabla\mathcal{T})'_{ijk}$ in the context of Picture B,



(J.9.1)

At the very end we will translate the result to Picture C1 of (F.1.1) where **x** instead of **x'** are the curvilinear coordinates.

For Picture B, one has from (F.1.11),

$$(\partial'_j \mathbf{e}_n) = \Gamma'^k{}_{jn} \, \mathbf{e}_k \qquad \Rightarrow \qquad (d\mathbf{e}_n)^i = (\partial'_j \mathbf{e}_n)^i \, dx'^j = \Gamma'^k{}_{jn} \, (\mathbf{e}_k)^i \, dx'^j \tag*{(J.9.3)}$$

$$\Gamma'^k{}_{jn} = \mathbf{e}^k \bullet (\partial'_j \mathbf{e}_n) \tag*{(J.9.4)}$$

where the $\mathbf{e}_n$ are our usual Chapter 3 tangent base vectors and we regard $\mathbf{e}_n = \mathbf{e}_n(\mathbf{x}')$.

It is possible to define a *different* affine connection $\hat{\Gamma}'^k{}_{jn}$ in this manner,

$$(\partial'_j \hat{\mathbf{e}}_n) = \hat{\Gamma}'^k{}_{jn} \, \hat{\mathbf{e}}_k \qquad \Rightarrow \qquad (d\hat{\mathbf{e}}_n)^i = \hat{\Gamma}'^k{}_{jn} \, (\hat{\mathbf{e}}_k)^i \, dx'^j \tag*{(J.9.4)}$$

$$\hat{\Gamma}'^k{}_{jn} = \hat{\mathbf{e}}^k \bullet (\partial'_j \hat{\mathbf{e}}_n) \, . \tag*{(J.9.5)}$$

This new affine connection $\hat{\Gamma}'^k{}_{jn}$ describes how the *unit vectors* $\hat{\mathbf{e}}_n$ vary with **x'**.

Now consider the tensor expansion from (J.4.8), where all objects are functions of **x'**,

$$T = \Sigma_{ij}\ \mathcal{T}'^{ij}\ \hat{\mathbf{e}}_i\otimes\hat{\mathbf{e}}_j\ = \mathcal{T}'^{ij}\ \hat{\mathbf{e}}_i\otimes\hat{\mathbf{e}}_j\ . \tag{J.9.6}$$

The object $T = T(\mathbf{x}')$ is a vector in a double direct-product space spanned by $\hat{\mathbf{e}}_i\otimes\hat{\mathbf{e}}_j$. Apply ∂'_k to get

$$\partial'_k T = \ (\partial'_k\mathcal{T}'^{ij})\ \hat{\mathbf{e}}_i\otimes\hat{\mathbf{e}}_j\ + \mathcal{T}'^{ij}(\partial'_k\hat{\mathbf{e}}_i)\otimes\hat{\mathbf{e}}_j + \mathcal{T}'^{ij}\ \hat{\mathbf{e}}_i\otimes(\partial'_k\hat{\mathbf{e}}_j)\ \ . \tag{J.9.7}$$

Now do index shuffles on (J.9.4),

$$(\partial'_j\hat{\mathbf{e}}_n) = \hat{\Gamma}'^k{}_{jn}\ \hat{\mathbf{e}}_k \qquad\qquad // \text{ do } k\rightarrow s \text{ then } j \rightarrow k \text{ and } n \rightarrow i$$
$$(\partial'_k\hat{\mathbf{e}}_i) = \hat{\Gamma}'^s{}_{ki}\ \hat{\mathbf{e}}_s \qquad\qquad // \text{ do } i\rightarrow j$$
$$(\partial'_k\hat{\mathbf{e}}_j) = \hat{\Gamma}'^s{}_{kj}\ \hat{\mathbf{e}}_s\ . \tag{J.9.8}$$

Insert the last two lines into (J.9.7) to get

$$\partial'_k T = \ (\partial'_k\mathcal{T}'^{ij})\ \hat{\mathbf{e}}_i\otimes\hat{\mathbf{e}}_j\ + \mathcal{T}'^{ij}\hat{\Gamma}'^s{}_{ki}\ \hat{\mathbf{e}}_s\otimes\hat{\mathbf{e}}_j + \mathcal{T}'^{ij}\ \hat{\mathbf{e}}_i\otimes\ \hat{\Gamma}'^s{}_{kj}\ \hat{\mathbf{e}}_s$$

In the second term swap s↔ i and in the third term swap s↔ j (dummy summation indices) to get,

$$\partial'_k T = \ (\partial'_k\mathcal{T}'^{ij})\ \hat{\mathbf{e}}_i\otimes\hat{\mathbf{e}}_j\ + \mathcal{T}'^{sj}\hat{\Gamma}'^i{}_{ks}\ \hat{\mathbf{e}}_i\otimes\hat{\mathbf{e}}_j + \mathcal{T}'^{is}\ \hat{\mathbf{e}}_i\otimes\ \hat{\Gamma}'^j{}_{ks}\ \hat{\mathbf{e}}_j$$

$$= [\partial'_k\mathcal{T}'^{ij}\ +\ \mathcal{T}'^{sj}\hat{\Gamma}'^i{}_{ks} +\ \mathcal{T}'^{is}\hat{\Gamma}'^j{}_{ks}\ ]\ \ \hat{\mathbf{e}}_i\otimes\hat{\mathbf{e}}_j, \tag{J.9.9}$$

so $\partial'_k T$ is another vector in the same double direct-product space. To economize below, we write the above as

$$\partial'_k T = Q^{ij}{}_k\ \hat{\mathbf{e}}_i\otimes\hat{\mathbf{e}}_j$$

$$\text{where } Q^{ij}{}_k \equiv\ [\partial'_k\mathcal{T}'^{ij}\ +\ \mathcal{T}'^{sj}\hat{\Gamma}'^i{}_{ks} +\ \mathcal{T}'^{is}\hat{\Gamma}'^j{}_{ks}\ ]\ \ . \tag{J.9.10}$$

Next, dot both sides of (J.9.10) with the vector $\mathbf{u}_a \otimes \mathbf{u}_b$ where the $\mathbf{u}_n$ are our usual Cartesian space axis-aligned unit vectors $(\mathbf{u}_n)^i = (\mathbf{u}_n)_i = \delta_{i,n}$. On the left side one gets

$$(\partial'_k T) \bullet \mathbf{u}_a \otimes \mathbf{u}_b\ = (\partial'_k T)^{ab} = \partial'_k T^{ab} \tag{J.9.11}$$

where T^{ab} are the contravariant components of the T tensor in x-space. On the right side one gets

$$\hat{\mathbf{e}}_i\otimes\hat{\mathbf{e}}_j \bullet\ \mathbf{u}_a \otimes \mathbf{u}_b\ = (\hat{\mathbf{e}}_i\bullet\mathbf{u}_a)\ (\hat{\mathbf{e}}_j\bullet\mathbf{u}_b)\ = (\hat{\mathbf{e}}_i)^a\ (\hat{\mathbf{e}}_j)^b\ . \tag{J.9.12}$$

Thus (J.9.10) becomes

$$\partial'_k T^{ab} = Q^{ij}{}_k\ (\hat{e}_i)^a\, (\hat{e}_j)^b\ . \tag{J.9.13}$$

We know from (7.6.8) that $\partial_c = R^k{}_c\ \partial'_k$, so applying $R^k{}_c$ to both sides of (J.9.13) gives

$$\partial_c T^{ab} = R^k{}_c\ (\hat{e}_i)^a\, (\hat{e}_j)^b\ Q^{ij}{}_k\ . \tag{J.9.14}$$

Since x-space is Cartesian, this is the same as

$$\partial^c T^{ab} = R^k{}_c\ (\hat{e}_i)^a\, (\hat{e}_j)^b\ Q^{ij}{}_k \tag{J.9.14a}$$

or, using (J.1.3) that $(\nabla T)^{ijk} \equiv \partial^k T^{ij}$,

$$(\nabla T)^{abc} = R^k{}_c\, (\hat{e}_i)^a\, (\hat{e}_j)^b\ Q^{ij}{}_k\ . \tag{J.9.15}$$

Our coefficients of interest are the $(\nabla\mathcal{T})'^{abc}$ which are related to $(\nabla T)^{abc}$ as in (E.8.20),

$$\begin{aligned}(\nabla\mathcal{T})'^{abc} &= M^a{}_A M^b{}_B M^c{}_C\, (\nabla T)^{ABC} \\ &= M^a{}_A M^b{}_B M^c{}_C\ \{\, R^k{}_C\, (\hat{e}_i)^A\, (\hat{e}_j)^B\ Q^{ij}{}_k \,\}\ .\end{aligned} \tag{J.9.16}$$

Things will simplify. First, from (E.8.12) we know that $(\hat{e}_i)^A = N^A{}_i$ and $(\hat{e}_j)^B = N^B{}_j$, so we continue the above

$$\begin{aligned}(\nabla\mathcal{T})'^{abc} &= M^a{}_A M^b{}_B M^c{}_C R^k{}_C N^A{}_i N^B{}_j\ Q^{ij}{}_k \\ &= (M^a{}_A N^A{}_i)(M^b{}_B N^B{}_j) M^c{}_C R^k{}_C\, Q^{ij}{}_k && \text{// reorder} \\ &= (\delta^a{}_i)(\delta^b{}_j) M^c{}_C R^k{}_C\, Q^{ij}{}_k = M^c{}_C R^k{}_C\, Q^{ab}{}_k && \text{// M and N are inverses as in (E.8.9)} \\ &= (h'_c\, R^c{}_C) R^k{}_C\, Q^{ab}{}_k = h'_c (R^c{}_C R^k{}_C) Q^{ab}{}_k && \text{// (E.8.7) then reorder} \\ &= h'_c (g'^{cd} R_d{}^C R^k{}_C) Q^{ab}{}_k = h'_c g'^{cd} (R_d{}^C R^k{}_C) Q^{ab}{}_k && \text{// (7.5.9) with g = 1 then reorder} \\ &= h'_c (g'^{cd} \delta_d{}^k) Q^{ab}{}_k = h'_c (g'^{ck}) Q^{ab}{}_k && \text{// (7.6.4) orthog rule \#4} \\ &= h'_c (h'_k{}^{-2}\, \delta_{c,k}) Q^{ab}{}_k = h'_c (h'_c{}^{-2}) Q^{ab}{}_c && \text{// (5.11.9) for orthog coords} \\ &= h'_c{}^{-1} Q^{ab}{}_c \\ &= h'_c{}^{-1}\, [\partial'_c \mathcal{T}'^{ab} + \mathcal{T}'^{sb}\hat{\Gamma}'^a{}_{cs} + \mathcal{T}'^{as}\hat{\Gamma}'^b{}_{cs}\,]\ . && \text{// (J.9.10)}\end{aligned} \tag{J.9.17}$$

The final result with the Lai affine connection is then,

$$(\nabla\mathcal{T})'^{abc} = (\nabla\mathcal{T})'_{abc} = h'_c{}^{-1}\,[\partial'_c\mathcal{T}'^{ab} + \mathcal{T}'^{sb}\hat{\Gamma}'^{a}{}_{cs} + \mathcal{T}'^{as}\hat{\Gamma}'^{b}{}_{cs}\,] \ . \tag{J.9.18}$$

Now do an index shuffle abc→ ijm to get,

$$(\nabla\mathcal{T})'^{ijm} = h'_m{}^{-1}\,[\partial'_m\mathcal{T}'^{ij} + \mathcal{T}'^{sj}\hat{\Gamma}'^{i}{}_{ms} + \mathcal{T}'^{is}\hat{\Gamma}'^{j}{}_{ms}\,] \,. \tag{J.9.19}$$

Expressing this result for Picture C1 of (F.1.1) removes the primes to give

$$(\nabla\mathcal{T})^{ijm} = (\nabla\mathcal{T})_{ijm} = h_m{}^{-1}\,[\partial_m\mathcal{T}^{ij} + \mathcal{T}^{sj}\hat{\Gamma}^{i}{}_{ms} + \mathcal{T}^{is}\hat{\Gamma}^{j}{}_{ms}\,] \ . \tag{J.9.20}$$

The corresponding equations (J.9.4) and (J.9.5) become in Picture C1 (see (F.1.1) and following text)

$$\begin{aligned}
&(\partial_j\hat{\mathbf{q}}_n) = \hat{\Gamma}^{k}{}_{jn}\,\hat{\mathbf{q}}_k && \Rightarrow \quad (d\hat{\mathbf{q}}_n)^i = \hat{\Gamma}^{k}{}_{jn}\,(\hat{\mathbf{q}}_k)^i\,dx^j \qquad \textit{(J.9.4)}\\
& && \qquad\ d\hat{\mathbf{q}}_n = \hat{\Gamma}^{k}{}_{jn}\,\hat{\mathbf{q}}_k\ dx^j\\
&\hat{\Gamma}^{k}{}_{jn} = \hat{\mathbf{q}}^k \bullet (\partial_j\hat{\mathbf{q}}_n) \quad \textit{(J.9.5)} && \qquad\ d\hat{\mathbf{q}}_i = \hat{\Gamma}^{k}{}_{ji}\,\hat{\mathbf{q}}_k\ dx^j \ .
\end{aligned} \tag{J.9.21}$$

where $\mathbf{q}_n(\mathbf{x})$ are the tangent base vectors in the ξ-space of Picture C1 and $\hat{\mathbf{q}}_n = \mathbf{q}_n(\mathbf{x})/h_n$.

We now translate these results to the Lai *et al.* notation as follows:

Us	Lai	
$\hat{\mathbf{q}}_n$	$\mathbf{e}_n$	// unit vectors in ξ-space
$\mathcal{T}^{ij}$	T_{ij}	
$\hat{\Gamma}^{i}{}_{ms}$	Γ_{smi}	// the Lai affine connection, note reverse index order
$(\nabla\mathcal{T})^{ijm}$	M_{ijm}	
∇T	M .	(J.9.22)

The translation of (J.9.20) and the right side of (J.9.21) is then

$$\begin{aligned}
(\nabla\mathcal{T})^{ijm}h_m &= [\partial_m\mathcal{T}^{ij} + \mathcal{T}^{sj}\hat{\Gamma}^{i}{}_{ms} + \mathcal{T}^{is}\hat{\Gamma}^{j}{}_{ms}\,] && \text{// us, now take s→q}\\
M_{ijm}\,h_m &= [\partial_m T_{ij} + T_{qj}\Gamma_{qmi} + T_{iq}\Gamma_{qmj}\,] && \text{// Lai p 504 (8A.26)}
\end{aligned} \tag{J.9.23}$$

$$\begin{aligned}
d\hat{\mathbf{q}}_i &= \hat{\Gamma}^{k}{}_{ji}\ dx^j\ \hat{\mathbf{q}}_k && \text{// us}\\
d\mathbf{e}_i &= \Gamma_{ijk}\ dx_j\ \mathbf{e}_k \ . && \text{// Lai p 502 (8A.12)}
\end{aligned} \tag{J.9.24}$$

Lai *et al.* use (J.9.23) to compute the $(\nabla\mathcal{T})^{ijm}$ for cylindrical and spherical coordinates on p 504 and 505. Their results are in agreement with our calculations (J.8.1) and (J.7.9).

Appendix K: Deformation Tensors in Continuum Mechanics

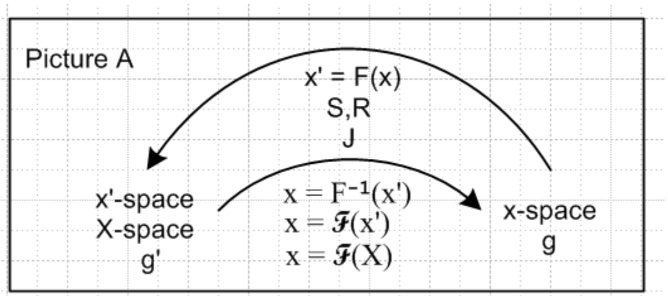
Tensor-like objects appear everywhere in continuum mechanics. As was noted below (2.12.3), continuum mechanics texts generally refer to all objects having indices as being "tensors", whether or not these objects actually transform as tensors with respect to some underlying transformation. The most commonly appearing tensors have two indices and are just 3x3 matrices associated with 3D space. In this category, there are several kinds of stress tensors, and many kinds of strain and deformation tensors which describe how a tiny volume of continuous matter (perhaps a tiny cube near some point **x**) changes shape in response to some applied stress. For a fluid, a fixed stress pattern can cause a continuous ongoing change of shape which is measured then by a "rate of deformation tensor" often called D.

An equation relating stress to strain/deformation is called a constitutive equation and describes the "response" of some physical system to "stimulus". The constitutive equation for a spring is $\mathbf{F} = -k\Delta\mathbf{x}$, for example, which is distinct from the equation of motion for a mass on a spring which is $\mathbf{F} = m\mathbf{a}$. For the spring, the stimulus is the force **F** ("stress"), and the response is the spring stretch $\Delta\mathbf{x}$ ("strain").

In this Appendix we shall study the tensor aspects of several kinds of deformation tensors appearing in continuum mechanics. In the book of Lai *et al.* this material is spread over several chapters, but here it will all be put in one place with Lai references provided. Section K.3 below considers the form of a candidate constitutive equation for a continuous solid whose form is determined by the requirement that the equation be "covariant" as discussed in Section 7.15. Similarly, Sections K.4 and K.5 consider covariant constitutive equations for fluids

K.1 A Preliminary Deformation Flow Picture

Our flow version of Picture A was shown as (5.16.1), where $\mathcal{F} \equiv \mathbf{F}^{-1}$:

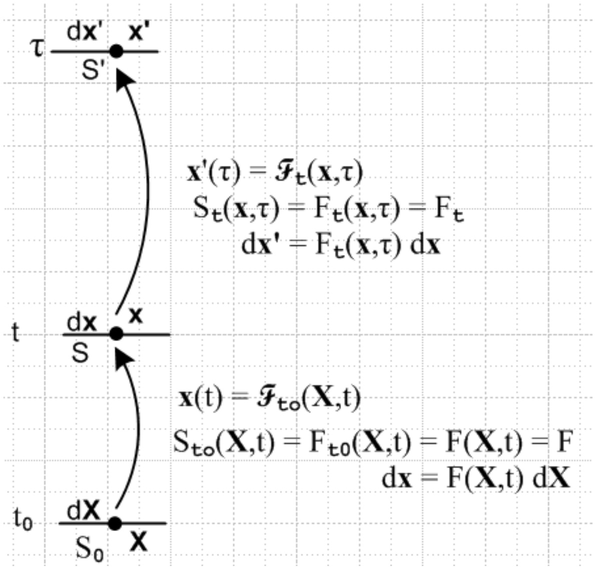


(K.1.1)

The upper arrow represents an underlying generally non-linear transformation between x-space and x'-space given by $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ while matrix S is the linearized-at-point-**x** version of the transformation which defines the notion of a vector with $d\mathbf{x} = S\, d\mathbf{x'}$ as in (2.1.6). Since F will have another meaning below, we change the transformation name so that $\mathbf{x} = \mathcal{F}(\mathbf{X})$ where $\mathcal{F} \equiv \mathbf{F}^{-1}$.

Warning: The x' in Fig (K.1.2) below is unrelated to the x' appearing in (K.1.1) which is really X. Also, overloaded symbol S will be used both for reference frames and for linearized transformation matrices.

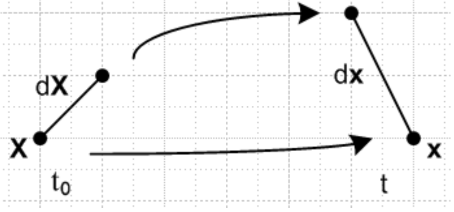
Consider now the picture below in which appear two sequential transformations $\mathcal{F}_{to}$ and $\mathcal{F}_t$. There are three spaces called X-space on the bottom, x-space in the middle, and x'-space on the top. The spaces are associated with frames of reference S_0, S and S'. Each frame has some set of basis vectors to be discussed below. The linearized S matrices associated with the two transformations are shown to the right and are given the names $S_{t0} = F$ on the bottom and $S_t = F_t$ on the top. The transformation $\mathbf{x}' = \mathcal{F}_t(\mathbf{x},\tau)$ is a spatial coordinate transformation only, the time coordinate τ is a parameter. In the picture below, time increases in the upward direction, so $\tau > t > t_0$.



(K.1.2)

Consider for the moment just the bottom transformation. The transformation $\mathcal{F}_{t0} \equiv \mathcal{F}$ describes the deformation of a particle of continuous matter which starts at position **X** and time t_0 and ends up at position **x** at time t. If we look at a large cube of continuous matter, we might find that it deforms in a very complicated manner as determined by the non-linear transformation $\mathcal{F}$ applied to all the particles within this large cube. The cube gets stirred up and is probably no longer recognizable. But if instead we consider a differentially small starting cube at **X** and t_0, we shall find that at time t that cube is at location **x** but has been transformed into a tiny rotated parallelepiped whose axes are no longer orthogonal. It is assumed that the flow is reasonable and smooth, we are not considering some kind of singular "explosion" here. We use the words flow and fluid, but the deformation concept applies to elastic solids as well as fluids since these deform in some way when they are stressed (think jello or even steel).

Rather than think of the flow in terms of the edges of this tiny cube, one can instead consider two very closely spaced points in the fluid close to **X** which are separated by spacing d**X** at time t_0, which we think of as "a little dumbbell". At time t, if one carefully tracks the "pathlines" of the ends of the dumbbell, one finds that the dumbbell tumbles and stretches and ends up as d**x** at time t and location **x,**



(K.1.3)

This differential dumbbell can be regarded as a mathematical "probe" embedded in the continuous medium. The relationship between d**x** and d**X** is given by

$d\mathbf{x} = F(\mathbf{X},t)\, d\mathbf{X} \qquad dx_i = F_{ij}dX_j$ // Lai p 105 (3.18.3) (K.1.4)

where the matrix F_{ij} is called "the deformation gradient". Matrix F is also known as "the deformation gradient tensor" even though it is not a "tensorial tensor" with respect to any identifiable transformation. In Section 5.16 the above equation d**x** = Fd**X** was identified with d**x** = Sd**x'** with S being F. Fig (K.1.3) is just Fig (2.1.2) with x'-space identified with X-space. Thus, the deformation gradient F is the linearized version (at point **x**) of some fancy non-linear (and unknown) "flow transformation" $\mathcal{F}$. Since one can write

$dx_i = (\partial x_i/\partial X_j)dX_j$, (K.1.5)

one finds from (K.1.4) that

$F_{ij} = (\partial x_i/\partial X_j) \equiv \partial_j X_i \qquad$ or $\qquad \mathbf{F} = (\nabla \mathbf{x})$ // Lai p 105 (3.18.4) (K.1.6)

where the gradient ∇ is with respect to **X**, so it is really $\nabla = \nabla^{(\mathbf{X})}$. Thus the name "deformation gradient" for F. [Notice that (∇**x**) is a matrix. In Appendix G the form of (∇**v**) for arbitrary vector **v** is found in arbitrary curvilinear coordinates. The index order reversal $F_{ij} = \partial_j X_i$ is mentioned there as well.]

The deformation gradient F(X,t) depends implicitly on the time t_0. At $t = t_0+\varepsilon$ (with very small ε) no flow has yet taken place, so $dx_i = dX_i$ and then $F(X,t_0) = 1$. Time t_0 is called the reference time and one could display it by writing $F(X,t) = F_{t0}(X,t)$, but normally this t_0 label is suppressed.

Now consider the upper flow in Fig (K.1.2) above. It is entirely analogous to the lower flow, but names are changed. One gets from the lower flow to the upper flow by making these replacements :

$(t,\ \mathbf{x},\ d\mathbf{x},\ S) \rightarrow (\tau, \mathbf{x}', d\mathbf{x}', S')$
then
$(t_0, \mathbf{X}, d\mathbf{X}, S_0) \rightarrow (t, \mathbf{x},\ d\mathbf{x},\ S)$. (K.1.7)

The equations corresponding to (K.1.4) are therefore (as shown in (K.1.2)),

$d\mathbf{x'} = F_t(\mathbf{x},\tau)\, d\mathbf{x} \qquad dx'_i = (F_t)_{ij}dx_j$. // Lai p 457 (8.7.2) (K.1.8)

Since one can write

$dx'_i = (\partial x'_i/\partial x_j)dx_j$,

one finds that

$(F_t)_{ij} = (\partial x'_i/\partial x_j) \qquad$ or $\qquad \mathbf{F_t} = (\nabla \mathbf{x'})$ // Lai p 457 (8.7.3) (K.1.9)

where the gradient ∇ is with respect to **x**, so it is really $\nabla = \nabla^{(\mathbf{x})}$.

In the lower flow, t_0 is the reference time, and t is the "current time". In the upper flow, the current time t is the reference time, and τ is some time τ > t. The upper flow is relative to current time t as reference, and for that reason the word **relative** is pre-pended to the names of all related tensors. Thus, F_t is called the "relative deformation gradient ", whereas F is just the "deformation gradient ".

Why are relative tensors useful?

The main motivation for use of the relative tensors concerns differentiation with respect to time in the vicinity of the current time t. One can write

$$d_tF_t(\mathbf{x},t) \equiv [\partial_\tau F_t(\mathbf{x},\tau)]^{\tau=t} \qquad // \ \mathbf{x} \text{ fixed (for example, } \mathbf{x} = \mathbf{x}_1)$$

where

$$\partial_\tau F_t(\mathbf{x},\tau) \approx [\ F_t(\mathbf{x},\tau+d\tau) - F_t(\mathbf{x},\tau)\ ] / d\tau\ . \qquad (K.1.10)$$

Here $d_t = D_t$ = d/dt = D/Dt = the total time derivative, and $\partial_t = \partial/\partial t$ = the partial time derivative. Since **x** is fixed, d**x** = 0 so $d_t = \partial_t$. We want to know the *rate* of deformation at some fixed current time t, and it is the τ argument of the function $F_t(\mathbf{x},\tau)$ that lets this derivative be computed.

One could in theory carry out this same differentiation using the "non-relative" tensors by doing d/dt_0 with t_0 near t:

$$d_tF_t(\mathbf{X},t) \equiv [\partial_{t0}F_{t0}(\mathbf{X},t)]^{t0=t} \qquad // \ \mathbf{X} \text{ fixed}$$

where

$$\partial_{t0}F_{t0}(\mathbf{X},t) \approx [\ F_{t0+dt0}(\mathbf{X},t) - F_{t0}(\mathbf{X},t)\ ] / dt_0\ , \qquad (K.1.11)$$

but this goes against the grain of the idea that **X** is a material coordinate at constant, fixed initial time t_0 which is earlier than t. And in the above, we end up with a statement about Lagrangian functions f(**X**,t) rather than Eulerian functions f(**x**,t), though one might argue that as t_0→ t, one has **X** → **x**. The relative tensor approach just makes the differentiation process clearer, and will be used below for that purpose.

[In the Lagrangian or material picture, one tracks a blob that started at position X -- one's instrumentation so to speak flows with the blob. In the Eulerian picture the instrumentation is fixed in space and observes the flow passing by.]

The frames of reference and the cameraman

Each of the three frames of reference S_0, S and S' in Fig (K.1.2) has its own set of orthonormal basis vectors which we are completely free to set in any manner. As a construct it is helpful to imagine that, as the flow proceeds, it is observed by a cameraman who flies around on a camera platform which translates and rotates in some arbitrary manner. Since our main concern will be with the dumbbells like d**X**, d**x** and d**x**', the translational part of the camera platform motion is irrelevant since d**X** is invariant under translations. We allow the cameraman's arbitrary orientation at times t_0, t and τ to determine the axes of the three frames S_0, S and S'. The cameraman is an "observer".

The *values* of the deformation gradient matrix elements F_{ij} depend on the choice of basis vectors in frames S_0 and S, so they in fact are dependent on how the cameraman flies his platform. Consider,

$$d\mathbf{x} = dx_1\hat{\mathbf{e}}_1^{(S)} + dx_2\, \hat{\mathbf{e}}_2^{(S)} + dx_3\, \hat{\mathbf{e}}_3^{(S)} \tag{K.1.12}$$

$$d\mathbf{X} = dX_1\hat{\mathbf{e}}_1^{(S0)} + dX_2\, \hat{\mathbf{e}}_2^{(S0)} + dX_3\, \hat{\mathbf{e}}_3^{(S0)}\ . \tag{K.1.13}$$

Once the axes are chosen, the value of the F_{ij} are determined, for example,

$$F_{12} \approx\ (dx_1)/(dX_2)\ . \tag{K.1.14}$$

If we were to rotate the basis vectors in frame S, for example, dx_1 would change, dX_2 would stay the same, and F_{12} would change.

Tensor expansions of the deformation gradients

These two expansions won't be used below, they are just inserted here as an application of the work of Appendix E on tensor expansions. These expansions use the cameraman basis vectors just defined above which are $\hat{\mathbf{e}}_n^{(S0)}$ for frame S_0 and $\hat{\mathbf{e}}_n^{(S)}$ for frame S.

- Consider this candidate expansion for the deformation gradient F,

$$F\ = \Sigma_{ij}\, F_{ij}\, \hat{\mathbf{e}}_i^{(S)} \otimes\ \hat{\mathbf{e}}_j^{(S0)}\quad =\ \Sigma_{ij}\, F_{ij}\ [\hat{\mathbf{e}}_i^{(S)}]\, [\hat{\mathbf{e}}_j^{(S0)}]^T$$

$$\text{where}\qquad F_{ij} = [F^{(S,S0)}]_{ij}\ =\ [\hat{\mathbf{e}}_i^{(S)}]^T\, F\, [\hat{\mathbf{e}}_j^{(S0)}] = (\partial x_i/\partial X_j)\ , \tag{K.1.15}$$

where we write the expansion in both direct product and matrix forms of Appendix E. This is a "mixed basis expansion" as discussed in Section E.10. Consider the application of this expansion to $d\mathbf{X}$:

$$\{\ \Sigma_{ij}\, F_{ij}\ [\hat{\mathbf{e}}_i^{(S)}]\, [\hat{\mathbf{e}}_j^{(S0)}]^T\ \}\ \ d\mathbf{X} \qquad\qquad // \equiv \{\text{expansion}\}\ dX$$

$$= \Sigma_{ij}\, F_{ij}\ [\hat{\mathbf{e}}_i^{(S)}]\, [\hat{\mathbf{e}}_j^{(S0)}]^T\ \{\ \Sigma_k dX_k\hat{\mathbf{e}}_k^{(S0)}\}$$

$$= \Sigma_{ij}\, F_{ij}\ \ \Sigma_k dX_k\ \ [\hat{\mathbf{e}}_i^{(S)}]\, [\hat{\mathbf{e}}_j^{(S0)}]^T\, [\ \hat{\mathbf{e}}_k^{(S0)}]$$

$$= \Sigma_{ij}\, F_{ij}\ \ \Sigma_k dX_k\ \ [\hat{\mathbf{e}}_i^{(S)}]\, \delta_{j,k}$$

$$= [\ \Sigma_{ij}\, F_{ij}\ \ dX_j]\ \ \hat{\mathbf{e}}_i^{(S)}$$

$$= [\ dx_i\ ]\, \hat{\mathbf{e}}_i^{(S)} \qquad\qquad // \text{ using the fact that F } d\mathbf{X}\ = d\mathbf{x}$$

$$= d\mathbf{x}\quad . \tag{K.1.16}$$

Since {expansion}$d\mathbf{X} = d\mathbf{x}$ and since $(\nabla^{(X)}\mathbf{x})\ d\mathbf{X} = d\mathbf{x}$ by the chain rule, it seems reasonable to conclude that {expansion} = $(\nabla^{(X)}\mathbf{x}) = F$.

• If we agree to use the $\hat{\mathbf{e}}_i^{(S)}$ for both d**x** and d**x'**, so that d**x** = $dx_i\hat{\mathbf{e}}_i^{(S)}$ and d**x'** = $dx'_i\hat{\mathbf{e}}_i^{(S)}$, then the following is a viable expansion for the relative deformation gradient F_t:

$$F_t = \Sigma_{ij} (F_t)_{ij}\ \hat{\mathbf{e}}_i^{(S)} \otimes\ \hat{\mathbf{e}}_j^{(S)} \quad = \quad \Sigma_{ij}\ (F_t)_{ij}\ [\hat{\mathbf{e}}_i^{(S)}]\,[\hat{\mathbf{e}}_j^{(S)}]^T$$

$$\text{where} \qquad (F_t)_{ij} = [F_t^{(S,S)}]_{ij} \ = \ [\hat{\mathbf{e}}_i^{(S)}]^T\, F_t\, [\hat{\mathbf{e}}_j^{(S)}] = (\partial x'_i/\partial x_j) \qquad\qquad (K.1.17)$$

The verification is similar to the above,

$$\{\ \Sigma_{ij} (F_t)_{ij}\ [\hat{\mathbf{e}}_i^{(S)}]\,[\hat{\mathbf{e}}_j^{(S)}]^T \}\ \ d\mathbf{x} \qquad\qquad // \equiv \{\text{expansion}\}\ dX$$

$$= \Sigma_{ij} (F_t)_{ij}\ [\hat{\mathbf{e}}_i^{(S)}]\,[\hat{\mathbf{e}}_j^{(S)}]^T \{\ \Sigma_k dx_k \hat{\mathbf{e}}_k^{(S)}\}$$

$$= \Sigma_{ij} (F_t)_{ij}\ \Sigma_k dx_k\ [\hat{\mathbf{e}}_i^{(S)}]\,[\hat{\mathbf{e}}_j^{(S)}]^T\,[\ \hat{\mathbf{e}}_k^{(S)}]$$

$$= \Sigma_{ij} (F_t)_{ij}\ \Sigma_k dx_k\ [\hat{\mathbf{e}}_i^{(S)}]\ \delta_{j,k}$$

$$= [\ \Sigma_{ij} (F_t)_{ij}\ \ dx_j]\ \ \hat{\mathbf{e}}_i^{(S)}$$

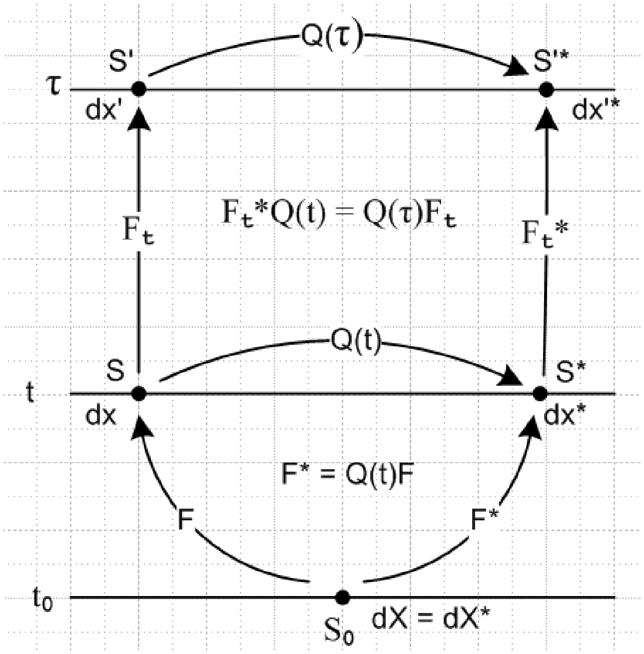
$$= [\ dx'_i\ ]\ \hat{\mathbf{e}}_i^{(S)} \qquad\qquad // \text{ using the fact that } F_t\, d\mathbf{x}\ = d\mathbf{x'}$$

$$= d\mathbf{x'}\quad . \qquad\qquad (K.1.18)$$

Since {expansion}d**x** = d**x'** and since ($\nabla^{(x)}$**x'**) d**x** = d**x'** by the chain rule, it seems reasonable to conclude that {expansion} = ($\nabla^{(x)}$**x'**) = F_t.

K.2 A More Complicated Deformation Flow Picture

Consider now this flow picture,



(K.2.1)

There is much to be said about this drawing.

The left side is the same as in Fig (K.1.2) shown above.

The picture is simplified in that it shows only the linearized S-type transformations like F and not the full transformations like $\mathcal{F}$, and these S-type matrices are now labeled right on the transformation arrows. We only care about these F matrices because we only care about the dumbbells like d**x**. For example, on the lower left we have d**x** = F d**X** .

The two sides of the picture represent observations of *the same flow* by two independent flying cameraman observers, call them C and C*. On each side the basis vectors of the various frames are set by the motions of these cameramen. The two cameramen have agreed to start off at time t_0 with their camera platforms in exact alignment, so there is no need for a frame S_0*.

The two frames of reference S and S* are related by some Galilean transformation (rotation plus translation) which brings the two independent camera platforms into alignment at time t :

$$\mathbf{x}^* = Q(t)\,(\mathbf{x}-\mathbf{x}_0) + \mathbf{c(t)} \qquad \Rightarrow \qquad d\mathbf{x}^* = Q(t)\,d\mathbf{x} \quad . \qquad (K.2.2)$$

Here $\mathbf{x}_0$ is a randomly selected origin for the rotation Q(t), and **c(t)** the corresponding translation. As above, we only care about the Q(t) part of this transformation, so in terms of d**x** objects, the frames S and S* are in effect related by the rotation Q(t).

The arrows in Fig (K.2.1) correctly describe the transformations of the d**x** type objects in moving between frames. In the lower part, for example, we have

$$d\mathbf{x}^* = Q(t)\,d\mathbf{x} \qquad d\mathbf{x} = F\,d\mathbf{X} \qquad d\mathbf{x}^* = F^*\,d\mathbf{X} \quad . \qquad (K.2.3)$$

Comparing the left and right equations one has $Q(t)\, d\mathbf{x} = F^*\, d\mathbf{X}$ and then the center equation can be used on the left side to get $Q(t)\, F\, d\mathbf{X} = F^* d\mathbf{X}$. Since this has to be true for any $d\mathbf{X}$, we have $Q(t)\, F = F^*$ as shown in the drawing. This result is trivially obtained just by looking at the alternate arrow paths from frame S_0 to frame S^*. So:

$$F^* = Q(t)\, F \qquad \text{or} \qquad F^*(\mathbf{x}^*,t^*) = Q(t)\, F(\mathbf{x},t) \ . \qquad t^* = t \qquad \text{(K.2.4)}$$

At time τ we have a similar situation, but there are four arrows instead of three. Comparing the arrow paths from frame S to frame S'* one finds

$$F_t^*Q(t) = Q(\tau)F_t \qquad \Rightarrow$$

$$F_t^* = Q(\tau)F_tQ(t)^T \qquad \text{or} \qquad F_t^*(\mathbf{x}^*,\tau^*) = Q(\tau)\, F_t(\mathbf{x},\tau)Q(t)^T \qquad t^* = t \qquad \text{(K.2.5)}$$

The two Q's are rotations (reflections included) and are therefore orthogonal so $Q^{-1} = Q^T$.

Does F transform as a tensor with respect to rotation Q(t) ?

We can think of $F(\mathbf{x},t)$ as a property of the continuous material at location $\mathbf{x}$ and current time t. F describes the "state of deformation". If F transformed as a tensor with respect to Q(t), one would need this to be true,

$$F^* = Q(t)\, F\, Q(t)^T \ , \qquad // \text{ not true!} \qquad \text{(K.2.6)}$$

which is the matrix form for the transformation of a rank-2 tensor as shown in (5.7.3). But we have just seen that $F^* = Q(t)\, F$ so the required $Q(t)^T$ on the right is missing. We conclude therefore that in fact F, although it is called a tensor, does not transform as a tensor under Q(t). One then says that F is a non-objective tensor with respect to Q(t). Equation $F^* = Q(t)\, F$ in fact says that the columns of matrix F transform as vectors under Q(t), which is very different from saying F transforms as a rank-2 tensor under Q(t).

If one is trying to construct a phenomenological equation modeling a continuous material at point $\mathbf{x}$ and time t, one must make sure that equation is "covariant" (frame-indifferent) with respect to rotation Q(t). The observers (cameramen) in frame S and frame S* must see equations which have exactly the same form, which means the elements in the equations must be objective with respect to Q(t). See Section 7.15 for a general discussion of "covariance". Since F is non-objective, it is not directly useful in the construction of covariant model equations.

Do any of the usual "derived tensors" transform as tensors with respect to Q(t) ?

By "the usual derived tensors" we mean B, C, U, V and associated R all defined as follows:

$$B = FF^T \ = \text{the left Cauchy-Green deformation tensor} \ = \text{the Piola deformation tensor} \qquad \text{(K.2.7)}$$

$$C = F^TF \ = \text{the right Cauchy-Green deformation tensor} \ = \text{the Finger deformation tensor} \qquad \text{(K.2.8)}$$

$$F = RU = VR \qquad R = \text{rotation} \qquad U,V = \text{symmetric positive definite} \qquad \text{(K.2.9)}$$

Tensors B and C are defined as shown, and both are therefore symmetric tensors. The last line is a statement of the **polar decomposition** theorem which says that any (real) non-singular matrix (det ≠ 0) can be *uniquely* written in these two ways (we apply this theorem to the deformation tensor F)

$$F = RU = VR \qquad \Rightarrow \ U = R^T VR \quad \text{and} \ \ V = RUR^T \qquad // \text{ Lai p 110 (3.21.1,2,4)} \qquad (K.2.10)$$

where R is a rotation matrix and V and U are *symmetric* positive definite matrices (meaning the eigenvalues are all positive) known as the left and right stretch tensors. Note that R is the same matrix in both the RU and VR forms. The idea is that the R matrix takes into account the rotational part of the deformation F, while U or V take into account the stretch component of the deformation. If the deformation is a pure rotation, U = V = 1, whereas if the deformation is a pure stretch then R = 1. A general deformation is a rotation/stretch/shear affair and one will find that none of R, U, V are unity.

One can combine the three equations above to find that

$$B = FF^T = (VR)(VR)^T = VRR^TV^T = VV^T = VV = V^2 \qquad // \text{ Lai p 121 (3.25.1)} \qquad (K.2.11)$$

$$C = F^TF = (RU)^T(RU) = U^TR^TRU = U^TU = U^2 \ . \qquad // \text{ Lai p 115 (3.23.1,2)} \qquad (K.2.12)$$

So our task is to discover whether any of these derived tensors transform as a tensor relative to Q(t). If they do transform as tensors (if they are objective), then they are candidates for use in constructing model equations for the continuous material.

We start with B and C:

$$B^* = F^*F^{*T} = (QF)(QF)^T = QF\,F^TQ^T = QBQ^T \qquad \Rightarrow \qquad B^* = Q(t)BQ(t)^T \qquad (K.2.13)$$

$$C^* = F^{*T}F^* = (QF)^T(QF) = F^TQ^TQF = F^TF = C \qquad \Rightarrow \qquad C^* = C \ . \qquad (K.2.14)$$

Thus, the left Cauchy-Green deformation tensor B actually does transform as a rank-2 tensor with respect to Q(t), so it is a tensorial tensor, it is "objective". In contrast, since C* = C, the right Cauchy-Green deformation tensor does not transform as a rank-2 tensor. In fact each element of matrix C transforms as a tensorial scalar with respect to Q(t).

What about V and U as defined above, the left and right stretch tensors?

$$F = RU = VR \qquad\qquad F^* = R^*U^* = V^*R^* \ . \qquad (K.2.15)$$

Consider,

$$F^* = QF = Q(RU) = (QR)\,(U) = R^*U^* \ . \qquad (K.2.16)$$

Since U is positive definite symmetric, and since QR is a rotation, and since the polar decomposition is unique, it must be that

$$R^* = QR \quad \text{and} \quad U^* = U \ . \qquad (K.2.17)$$

Next write

$$F^* = QF = Q(VR) = (QVQ^T)(QR) = V^* R^* \ . \qquad (K.2.18)$$

Since the eigenvalues of symmetric V are determined by det(V-λI) = 0, and since this is the same as the equation $\det(QVQ^T-\lambda I) = 0$, QVQ^T has the same eigenvalues as V and so (QVQ^T) is symmetric and positive definite. Due to this fact and the fact that QR is a rotation, and the fact that the polar decomposition is unique, it must be that

$$V^* = QVQ^T \quad \text{and} \quad QR = R^* \ . \qquad (K.2.19)$$

and thus V transforms as a true tensor. So here is a summary for our tensors of interest. Only two of the five deformation tensors actually transform as tensors. The references are to Lai page 336-337 :

$F^* = Q(t)F$ // Lai (5.56.21)
$B^* = Q(t)BQ(t)^T$ // rank-2 tensor with respect to Q(t) so objective // Lai (5.56.31)
$C^* = C$ // Lai (5.56.28)
$U^* = U$
$V^* = Q(t)VQ(t)^T$ // rank-2 tensor with respect to Q(t) so objective
$R^* = Q(t)R$ (K.2.20)

Comment: Recall that F = F(**x**,t) has a hidden parameter t_0 so in fact $F = F_{t0}(\mathbf{x},t)$. Similarly, all derived tensors have this same hidden parameter. Thus, for example, one could write the transformation of B as

$$B_{t0}^*(\mathbf{x}^*,t^*) = Q(t)\, B_{t0}(\mathbf{x},t)Q(t)^T \qquad t^* = t \qquad \mathbf{x}^* = Q(t)\,(\mathbf{x}-\mathbf{x}_0) + \mathbf{c(t)} \qquad d\mathbf{x}^* = Q(t)\, d\mathbf{x} \ . \quad (K.2.21)$$

The parameter t_0 is treated as a fixed constant here and plays no role in the question of whether or not B transforms as a rank-2 tensor. The important time argument of B is the current time t, and the main idea is that $B^*(t) = Q(t)\, B(t)Q(t)^T$ so that B(t) is objective with respect to the rotation Q(t). The transformation is valid for any value of t_0. In the limit that $t_0 \to t$, the equation says $1 = Q(t)\ 1\ Q(t)^T$ which of course is true since rotation Q(t) is orthogonal.

Do any of the usual *relative* derived tensors transform as tensors with respect to Q(t) ?

Again we think of a relative tensor W_t as being a property of the continuous material at current time t, a measure of the state of deformation. Such a tensor is objective only if $W_t^* = Q(t)W_tQ(t)^T$. With regard to the above Comment, in this new situation it is the t of W_t which is the time variable of interest (the current time), and time τ is regarded as a fixed parameter, as was t_0 in the Comment. It just happens that the notational positions of the current time t and the parameter time τ are swapped in this case relative to the last, so now we have

$$W_t^*(\mathbf{x}^*,\tau^*) = Q(t)W_t(\mathbf{x},\tau)\, Q(t)^T \qquad \tau^* = \tau \qquad \mathbf{x}^* = Q(t)\,(\mathbf{x}-\mathbf{x}_0) + \mathbf{c(t)} \qquad d\mathbf{x}^* = Q(t)\, d\mathbf{x} \ . \quad (K.2.22)$$

A tensor W_t which transforms as a rank-2 tensor (is objective) with respect to rotation Q(t) must satisfy the relation above, where the arguments of both Q rotations are t. As in the Comment above, this transformation is valid for any value of parameter τ, and as τ→t, the equation says $1 = Q(t)\ 1\ Q(t)^T$.

Our study of the transformation properties of the relative tensors proceeds in a manner similar to that used for the regular tensors above. We start with $B_t \equiv F_tF_t^T$:

$$B_t^* = F_t^*F_t^{*T} = [Q(\tau)\ F_t\ Q^T(t)]\ [Q(\tau)\ F_t\ Q^T(t)]^T = Q(\tau)\ F_t\ Q^T(t)\ Q(t)\ F_t^T\ Q(\tau)^T$$

$$= Q(\tau)\ F_t\ F_t^T\ Q(\tau)^T = Q(\tau)\ B_t\ Q(\tau)^T \quad \text{// not a rank-2 tensor since } t \neq \tau \tag{K.2.23}$$

Next comes $C_t \equiv F_t^TF_t$:

$$C_t^* = F_t^{*T}F_t^* = [Q(\tau)\ F_t\ Q^T(t)]^T\ [Q(\tau)\ F_t\ Q^T(t)] = Q(t)\ F_t^T\ Q(\tau)^T\ Q(\tau)\ F_t\ Q^T(t)$$

$$= Q(t)\ F_t^T\ F_t\ Q^T(t) = Q(t)\ C_tQ^T(t) \quad \text{// yes a rank-2 tensor with respect to } Q(t) \tag{K.2.24}$$

What about the left and right relative stretch tensors V_t and U_t?

$$F_t = R_tU_t = V_tR_t \qquad F_t^* = R_t^*U_t^* = V_t^*R_t^* \tag{K.2.25}$$

Consider,

$$F_t^* = Q(\tau)\ F_t\ Q^T(t) = Q(\tau)\ R_tU_tQ^T(t) = [Q(\tau)\ R_tQ^T(t)]\ [Q(t)U_tQ^T(t)] = R_t^*U_t^* \quad . \tag{K.2.26}$$

Since $[Q(\tau)\ R_tQ^T(t)]$ is a rotation and since $[Q(t)U_tQ^T(t)]$ is a symmetric positive definite matrix by the argument given in the previous section, and since the polar decomposition is unique, it must be that

$$R_t^* = Q(\tau)\ R_tQ^T(t) \qquad \text{and} \qquad U_t^* = Q(t)U_tQ^T(t) \quad \text{// } U_t \text{ is a rank-2 tensor} \tag{K.2.27}$$

Finally, write

$$F_t^* = Q(\tau)\ F_tQ^T(t) = Q(\tau)V_tR_tQ^T(t) = [Q(\tau)V_tQ^T(\tau)]\ [Q(\tau)\ R_tQ^T(t)] = V_t^*\ R_t^* \tag{K.2.28}$$

By the same argument used several times above, we conclude that

$$R_t^* = Q(\tau)R_tQ^T(t) \quad \text{and} \quad V_t^* = Q(\tau)V_tQ^T(\tau) \quad \text{// } V_t \text{ is not a rank-2 tensor, } \tau \neq t \tag{K.2.29}$$

The rule for transforming R_t is the same as found a few lines above.

Here then are the conclusions, with references to Lai page 472:

$F_t^* = Q(\tau)F_tQ^T(t)$ // Lai (8.13.6)
$B_t^* = Q(\tau)B_tQ(\tau)^T$ // Lai (8.13.12)
$C_t^* = Q(t)C_tQ^T(t)$ // rank-2 tensor with respect to Q(t) so objective // Lai (8.13.10)
$U_t^* = Q(t)U_tQ^T(t)$ // rank-2 tensor with respect to Q(t) so objective // Lai (8.13.9)
$V_t^* = Q(\tau)V_tQ^T(\tau)$ // Lai (8.13.12)
$R_t^* = Q(\tau)R_tQ^T(t)$ // Lai (8.13.8) (K.2.30)

Notice that among the "normal" tensors, B and V are objective, whereas among the "relative tensors" it is C_t and U_t that are objective. All the other tensors are "non-objective".

K.3 Covariant form of a solid constitutive equation involving the deformation tensor

For a solid continuous material in frame S one can *consider* a stress/deformation relationship of the form T = f(B), where T is the Cauchy stress tensor, B is the left Cauchy-Green deformation tensor mentioned in (K.2.7) above, and f is "some function".

In frame S*, there will be some covariant version of the equation T* = f*(B*). If the medium is isotropic (rotationally invariant in its properties), then f* = f and one will have T* = f(B*) in Frame S*. Two observers of the same system in frames related by a rotation cannot observe different functions $f \neq f^*$ if the material is isotropic. Notice that there are two separate issues here: (1) equation must be covariant under rotations to be viable; (2) isotropic implies f = f*.

If f is a polynomial, or a function which can be approximated by one (f is smooth), then T = f(B) with polynomial coefficients which are rotational scalars (with respect to Q) is a viable equation form for the following reason: since B is a rank-2 tensor by (K.2.20), so is any power of B,

$$B^{*2} = [QBQ^T][QBQ^T] = Q\,B^2Q^T \qquad \text{etc.} \qquad (K.3.1)$$

and if the polynomial coefficients are scalars, then f(B) is a rank-2 tensor.

Just as a particle force **F** transforms as a rank-1 tensor under rotations, the Cauchy stress tensor T transforms as a rank-2 tensor under rotations, and then both sides of T = f(B) transform in the same way -- as rank-2 tensors. Any candidate equation between T and a deformation tensor which did not have both sides transforming the same way would be invalid from the get-go (except perhaps as an approximation).

The scalar coefficients must be functions of the B_{ij} and there are three such scalars known as the principal scalar invariants of B (Lai p 40), one of which is det(B), so the scalar coefficients can be any functions of these three scalar invariants. Furthermore, one can use the fact that $B = FF^T$ is symmetric along with the Cayley-Hamilton theorem (symmetric matrix B satisfies its own secular equation, whose coefficients by the way are those scalar invariants) to show that any powers of B in polynomial f(B) larger than degree 2 can be expressed as a linear combination of I, B and B^2. One ends up then with $T = aI + bB + cB^2$ where a,b,c are functions of the three scalar invariants of tensor B.

Since both sides of T = f(B) transform in the same way (rank-2 tensors), the equation T = f(B) is "covariant" as discussed in Section 7.15, meaning it has the same form in frame S* as it has in S.

The equation T = f(B) is a relation between stress and strain in the form of deformation, and as such is called a constitutive equation for the continuous material. One wants such equations to be covariant between frames of reference related by any Galilean transformation (rotation + translation), even if one or both of these frames are non-inertial. This is an extension of Hooke's Law for a spring, $\mathbf{F} = -k\, \Delta\mathbf{x}$, which is covariant under rotations and translations.

In contrast, equations of motion are covariant only if both frame S and S* are inertial frames.

Notice that this entire discussion falls apart completely if one tries T = f(F) or T = f(C) as a candidate constitutive relation, since then the two sides of the equation don't transform the same way.

This subject is discussed in Lai pp 334-342 and p 40 for the scalar invariants. The requirement of covariance for an isotropic material and the fact that B is symmetric and transforms as a tensor puts a severe restriction on the form of the constitutive equation and we end up with $T = aI + bB + cB^2$. Since one can replace $B^3 = \alpha B^2 + \beta B + \gamma I$, if B is invertible ($\det B \neq 0$) one has $B^2 = \alpha B + \beta I + \gamma B^{-1}$ and this allows the alternate form $T = a'I + b'B + c'B^{-1}$. This last equation is used to model large deformations of an isotropic elastic material. An example is the Mooney-Rivlin theory for rubber.

K.4 Some fluid constitutive equations

It was noted just above (K.1.10) that the *relative* deformation tensors are appropriate when one is interested in time derivatives of the tensors. It was also noted in (K.2.30) that the relative deformation tensor C_t is objective. One can expand $C_t(\mathbf{x},\tau)$ in a Taylor series about current time t in this manner ($\partial_\tau \equiv \partial/\partial\tau$) ,

$$C_t(\mathbf{x},\tau) = \Sigma_{n=0}^{\infty} [\,\partial_\tau{}^n C_t(\mathbf{x},\tau)]^{\tau=t}\, (\tau-t)^n/n! \ = \Sigma_{n=0}^{\infty} A_n(\mathbf{x},t)\ (\tau-t)^n/n! \qquad \text{// Lai p 463 (8.10.1)}$$

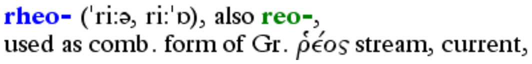
$$A_n(\mathbf{x},t) \equiv [\,\partial_\tau{}^n C_t(\mathbf{x},\tau)]^{\tau=t}\ , \tag{K.4.1}$$

where the coefficient derivatives are given the names $A_n(\mathbf{x},t)$ called Rivlin-Ericksen tensors. Each of these coefficient tensors is in fact objective, just as is C_t, since (as usual, t = t*, τ = τ*)

$$Q(t)\ [\,\partial_\tau{}^n C_t(\mathbf{x},\tau)]^{\tau=t}\, Q^T(t) = \{\, \partial_\tau{}^n\, [Q(t)\, C_t(\mathbf{x},\tau)\, Q^T(t)]\}^{\tau=t} \ = \{\, \partial_{\tau^*}{}^n\, C_t{}^*(\mathbf{x}^*,\tau^*)\}^{\tau^*=t}$$

$$\Rightarrow\ Q(t)\, A_n(\mathbf{x},t)\, Q^T(t) = A^*{}_n(\mathbf{x}^*,t) \tag{K.4.2}$$

These $A_n(x,t)$ tensors appear in various covariant models of "non-Newtonian" fluid behavior, the general study of which is called rheology, based on the Greek word for a current flow (a rheostat controls electric current),



rheo- (ˈriːə, riːˈɒ), also **reo-**,
used as comb. form of Gr. ῥέος stream, current,

// OED2

Here are a few covariant constitutive equations and the names assigned to them (Lai p 481). Note that for any normal fluid, there is always a -pI tensor term in the expression for stress T, where p is the fluid pressure and I is the identity matrix. The diagonal elements of matrix -pI are the equal normal stresses of the surroundings of a tiny cube of fluid pulling out on the cube faces, hence the -p (p > 0) since we know the fluid actually pushes in on the cube.

$T = -pI + \text{functional of } C_t(\tau),\ \tau \le t$ // "simple" fluid, since $\nabla^n F_t$ not involved ($C_t = F_t^T F_t$)

$T = -pI + \int_{-\infty}^{t} d\tau\ f_1(\tau)\ C_t(\tau)$ // single-integral simple fluid. $f_1(\tau)$ = a memory weight function

$T = -pI + f(A_1, A_2....A_N)$ // Rivlin-Ericksen incompressible fluid of complexity N
$T = -pI + f(A_1,A_2)$ // viscometric flow fluid (there are conditions on A_1 and A_2)
$T = -pI + \mu_1 A_1 + \mu_2 A_1^2 + \mu_3 A_2$ // second order fluid (paint, blood, polymers)

$T = -pI + \mu A_1$ // incompressible Newtonian fluid (fluids like water) (K.4.3)

It turns out that $A_1 = 2D$ where $D = [(\nabla \mathbf{v}) + (\nabla \mathbf{v})^T]/2 \equiv (\nabla \mathbf{v})^{sym}$, so A_1 is twice the rate of deformation tensor D. The other A_n can then be found from this recursion relation,

$A_{n+1} = d_t A_n + A_n(\nabla \mathbf{v}) + (\nabla \mathbf{v})^T A_n$ // Lai p 468 (8.11.2) (K.4.4)

Here **v** is the fluid velocity vector and $d_t = d/dt = D/Dt$. Again, $(\nabla \mathbf{v})$ is the subject of Appendix G.

K.5 Corotational and other objective time derivatives of the Cauchy stress tensor

The Cauchy stress tensor T transforms as a tensor under Q(t); it is objective. One can write therefore,

$T^*(\mathbf{x}^*,t^*) = Q(t)\ T(\mathbf{x},t)\ Q(t)^T \qquad t^* = t \qquad \mathbf{x}^* = Q(t)\ (\mathbf{x}-\mathbf{x}_0) + \mathbf{c(t)} \qquad d\mathbf{x}^* = Q(t)\ d\mathbf{x}$. (K.5.1)

Clarification of the above equation

One can think of the above equation $T^* = QTQ^T$ as involving operators in Hilbert Space, as outlined in Section E.7. In the upper part of Fig (K.2.1) above we show four different frames of reference called S, S*, S' and S'* each of which has its own set of basis vectors we might call $\mathbf{u}_n$, $\mathbf{u}^*_n$, $\mathbf{u}'_n$ and $\mathbf{u}^{*\prime}_n$. It happens that the picture refers to S and S* at time t, and S' and S'* at time τ, but any basis vectors can be "used" at any time one wants. For example, here are four expansions of the operator T(**x**,t)

$$T(\mathbf{x},t) = \Sigma_{ab}\ T_{ab}(\mathbf{x},t)\ \mathbf{u}_a \otimes \mathbf{u}_b = \Sigma_{ab}\ T^*_{ab}(\mathbf{x}^*,t)\ \mathbf{u}^*_a \otimes \mathbf{u}^*_b$$

$$= \Sigma_{ab}\ T'_{ab}(\mathbf{x}',t)\ \mathbf{u}'_a \otimes \mathbf{u}'_b = \Sigma_{ab}\ T'^*_{ab}(\mathbf{x}'^*,t)\ \mathbf{u}'^*_a \otimes \mathbf{u}'^*_b \qquad \text{(K.5.2)}$$

in which we see four different kinds of components T_{ab}, T^*_{ab}, T'_{ab}, T'^*_{ab} . The spatial arguments of each component are written as appropriate for that frame of reference and of course all "correspond" to each other (for example, $\mathbf{x}' = \mathcal{F}_t(\mathbf{x},\tau)$). Recall from (2.5.1) the notion of the transformation of a contravariant vector field in developmental notation

$\mathbf{V}'(\mathbf{x}') = R\ \mathbf{V}(\mathbf{x})$ contravariant $R_{ik}(\mathbf{x}) \equiv (\partial x'_i/\partial x_k)$ $R = S^{-1}$ *(2.5.1)*

where the argument is appropriate to the space of interest.

The time argument t in the above four expansions of T can be set to any arbitrary value. The stress tensor at a point **x** is in general a function of time t. One could for example set $t = \tau$ in all the expansions.

Having said this, we now decide that only the frame S basis vectors $\mathbf{u_n}$ shall be used in our expansions and components. Then for example (these $\mathbf{u_n}$ were called $\hat{\mathbf{e}}_n^{(S)}$ earlier)

$$T(\mathbf{x},t) = \Sigma_{ab}\ T_{ab}(\mathbf{x},t)\ \mathbf{u_a} \otimes \mathbf{u_b}$$

$$T^*(\mathbf{x}^*,t) = \Sigma_{ab}\ T^*_{ab}(\mathbf{x}^*,t)\ \mathbf{u_a} \otimes \mathbf{u_b}$$

$$Q(t) = \Sigma_{ab}\ Q_{ab}(t)\ \mathbf{u_a} \otimes \mathbf{u_b}\ . \qquad \text{(K.5.3)}$$

Our operator statement of objectivity then becomes the following when expressed in components,

$$T^*(\mathbf{x}^*,t^*)_{ij} = Q(t)_{ia}\, T(\mathbf{x},t)_{ab}\, Q(t)^T{}_{bj} \qquad t^* = t\quad . \qquad \text{(K.5.4)}$$

Thus, there should be no confusion about the following two equations which we express back in operator notation with the position arguments suppressed (but shown on the right)

$$T^*(t) = Q(t)\ T(t)\ Q(t)^T \qquad // \ T^*(\mathbf{x}^*, t) = Q(t)\ T(\mathbf{x},t)\ Q(t)^T$$

$$T^*(\tau) = Q(\tau)\ T(\tau)\ Q(\tau)^T \qquad // \ T^*(\mathbf{x}^*, \tau) = Q(\tau)\ T(\mathbf{x},\tau)\ Q(\tau)^T\ . \qquad \text{(K.5.5)}$$

<u>Problem: The tensor dT/dt fails to transform as a rank-2 tensor, even though T does so transform.</u>

If one tries to construct covariant constitutive equations involving dT/dt, a problem arises because dT/dt is non-objective,

$$T^*(t) = Q(t)\ T(t)\ Q(t)^T$$

$$(dT^*/dt) = Q\ (dT/dt)\ Q^T + [\ (dQ/dt)\ T\ Q^T + Q\ T\ (dQ/dt)^T\ ]\ , \qquad \text{(K.5.6)}$$

so there are two extra unwanted terms. Just as $B = FF^T$ is constructed to provide an objective derived tensor from non-objective F, one can construct a derived version of (dT/dt) which is objective. In the next three sections, three different derived versions are described.

<u>The corotational/Jaumann derivatives</u>

The first step is to define an adjusted stress tensor $J_t(\tau)$ at time τ according to (see Lai p 483 (8.19.3). Lai does not have a t subscript on J).

$$J_t(\tau)\ \equiv R_t^T(\tau)\ T(\tau)\ R_t(\tau) \qquad // \ J_t(\mathbf{x},\tau)\ \equiv R_t^T(\mathbf{x},\tau)\ T(\mathbf{x},\tau)\ R_t(\mathbf{x},\tau) \qquad \text{(K.5.7)}$$

where $R_t(\tau)$ is the rotation which appears above in (K.2.25), where we had (showing τ arguments),

$$R_t^*(\tau) = Q(\tau)\, R_t(\tau) Q^T(t) \ . \qquad\qquad \textit{(K.2.27)} \qquad\qquad (K.5.8)$$

The tensor $R_t(\tau)$ is non-objective due to appearance of $Q(\tau)$ instead of $Q(t)$ on the left (see comments near (K.2.22) on W_t). Recall that this rotation $R_t(\tau)$ is unique and is determined from the deformation tensor by the polar decomposition $F_t(\tau) = R_t(\tau)U_t(\tau) = V_t(\tau)\, R_t(\tau)$. Thus, in some sense $J_t(\tau)$ knows about the stress tensor $T(\tau)$, *and* it knows something about the deformation tensor through $R_t(\tau)$. [The meaning of the term "corotational" is explained far below.]

The **claim** now is that the time derivative of this corotating stress tensor J_t is objective, meaning that tensor d_tJ_t transforms as a rank-2 tensor under the rotation $Q(t)$. Here is a proof :

We first assemble the following facts,

$$\begin{array}{ll} T^*(\tau) = Q(\tau)\, T(\tau)\, Q(\tau)^T & \text{// transformation of stress tensor T at time } \tau, \textit{(K.5.5)} \\ R_t^*(\tau) = Q(\tau)\, R_t(\tau)Q^T(t) & \text{// how } R_t(\tau) \text{ transforms, where } F_t(\tau) = R_t(\tau)U_t(\tau), \textit{(K.5.8)} \\ J_t(\tau) \ \equiv R_t^T(\tau)\, T(\tau)\, R_t(\tau) & \text{// definition of } J_t(\tau) \text{ in frame S, } \textit{(K.5.7)} \\ J_t^*(\tau) \ \equiv R_t^{*T}(\tau)\, T^*(\tau)\, R^*_t(\tau) & \text{// corresponding } J_t^* \text{ in frame S*} \end{array} \qquad (K.5.9)$$

and then we combine these ingredients to obtain a transformation rule for J_t :

$$J_t^*(\tau) \ \equiv R_t^{*T}(\tau)\, T^*(\tau)\, R^*_t(\tau) \ = \ [Q(\tau)\, R_t(\tau)Q^T(t)]^T\, [Q(\tau)\, T(\tau)\, Q(\tau)^T]\, [Q(\tau)\, R_t(\tau)Q^T(t)]$$

$$= [Q(t)\, R_t^T(\tau)Q^T(\tau)]\, [Q(\tau)\, T(\tau)\, Q(\tau)^T]\, [Q(\tau)\, R_t(\tau)Q^T(t)]$$

$$= Q(t)\, R_t^T(\tau)\, [Q^T(\tau)Q(\tau)]\, T(\tau)\, [Q(\tau)^TQ(\tau)]\, R_t(\tau)Q^T(t)$$

$$= Q(t)\, [\, R_t^T(\tau)T(\tau)\, R_t(\tau)]\, Q^T(t)$$

$$= Q(t)\, J_t(\tau)\, Q^T(t) \ . \qquad (K.5.10)$$

Since this equation $J_t^*(\tau) = Q(t)\, J_t(\tau)\, Q^T(t)$ fulfills the condition described earlier for W_t to be objective, we conclude that the corotating stress transforms as a rank-2 tensor, where τ is treated as a parameter.

Consider now the limit of (K.5.10) as $\tau \rightarrow t$. One finds,

$$\begin{array}{ll} J_t(\tau) \ \ \equiv R_t^T(\tau)\, T(\tau)\, R_t(\tau) & \textit{// (K.5.7)} \\ J_t(t) \ \ \equiv R_t^T(t)\, T(t)\, R_t(t) \ = \ 1\, T(t)\, 1 = T(t) & \end{array} \qquad (K.5.11)$$

and

$$\begin{array}{ll} J^*_t(\tau) \ \ \equiv R^*{}_t^T(\tau)\, T^*(\tau)\, R^*_t(\tau) & \textit{// (K.5.9)} \\ J^*_t(t) \ \ \equiv R^*{}_t^T(t)\, T^*(t)\, R^*_t(t) \ = \ 1\, T^*(t)\, 1 = T^*(t)\ . & \end{array} \qquad (K.5.12)$$

In this limit, the corotation $R_t^{-1}(\tau)$ has come to a halt, and (K.5.10) becomes a statement that T is objective.

More interestingly, we can apply $\partial_\tau{}^n = \partial^n/\partial\tau^n$ to both sides of(K.5.10) to get

$$d_\tau{}^n\, J_t{}^*(\tau) \;\; = Q(t)[\, d_\tau{}^n\, J_t(\tau)\,]\, Q^T(t)\,. \qquad\qquad (K.5.13)$$

Taking the limit τ→t then gives

$$[d_t{}^n\, J_t{}^*](t) \;\; = Q(t)\;\; [d_t{}^n\, J_t](t)\, Q^T(t) \qquad\qquad (K.5.14)$$

which says that $d_t{}^n J_t$ are all objective tensors. And in particular, for n = 1,

$$(d_t J_t)^* \;\; = Q(t)\, (d_t J_t)\, Q^T(t)\;, \qquad\qquad (K.5.15)$$

and this concludes our proof that dJ_t/dt is objective, whereas dT/dt is not objective.

The above objective tensor time derivatives are sometimes written using the following strange notation

$$\overset{o}{T}_n \;\equiv\;\; [d^n J_t(t)/dt^n],\ n = 1,2,3... \qquad \overset{o}{T} \equiv \overset{o}{T}_1 \qquad J_t(\tau) \;\equiv R_t{}^T(\tau)\, T(\tau)\, R_t(\tau) \qquad\qquad (K.5.16)$$

and these are called corotational or Jaumann derivatives (Lai p 484) [Jaumann-Zaremba]. It can be shown that

$$\overset{o}{T} \;= d_t T + TW\text{-}WT \;\;\text{ where }\;\; W = [(\nabla\mathbf{v}) - (\nabla\mathbf{v})^T]/2 \;\; = \text{"the spin tensor"} \qquad // \text{ Lai p 484 (8.19.10)} \qquad (K.5.17)$$

The Oldroyd Lower convected derivatives

An alternative solution to the same problem uses a different adjusted stress tensor,

$$J_L(\tau) \;\equiv F_t{}^T(\tau)\, T(\tau)\, F_t(\tau) \qquad // \text{ Lai p 484 (8.19.12)} \quad (K.5.18)$$

We suppress the t subscript on J_L just to avoid having to write $(J_L)_t(\tau)$. In the table (K.2.30) one sees that F_t transforms the same way R_t does, so one can repeat the above analysis to conclude that the derivatives $[d^n J_L(t)/dt^n]$ are all objective tensors (just replace $R_t \to F_t$ everywhere), so

$$\overset{\Delta}{T}_n \;\equiv\;\; [d^n J_L(t)/dt^n]\ ,\ n = 1,2,3... \qquad \overset{\Delta}{T} \equiv \overset{\Delta}{T}_1, \qquad J_L(\tau) \;\equiv F_t{}^T(\tau)\, T(\tau)\, F_t(\tau) \quad // = \;\overset{\cup}{T}_n \qquad\qquad (K.5.19)$$

and these are the "Oldroyd *lower* convected derivatives" (Lai p 485 uses $\overset{\cup}{T}_n$) . $\overset{\Delta}{T}$ is sometimes called the Cotter-Rivlin stress rate. It can be shown that

$$\overset{\Delta}{T} = \overset{\cup}{T} = \;\; d_t T + T(\nabla\mathbf{v}) + (\nabla\mathbf{v})^T T \quad . \qquad // \text{ Lai p 485 (8.19.21)} \quad (K.5.20)$$

The Oldroyd Upper convected derivatives

Finally, consider again from (K.2.30) the non-objective way that F_t transforms,

$F_t{}^*(\tau) = Q(\tau)F_t(\tau)Q^T(t)$

$\Rightarrow\ (F_t{}^{-1})^*(\tau) = Q(t)\,(F_t{}^{-1}(\tau))\,Q^T(\tau)$ // inverted

$\Rightarrow\ (F_t{}^{-1})^{T*}(\tau) = Q(\tau)\,(F_t{}^{-1})^T(\tau)\,Q^T(t)$. // then transposed (K.5.21)

This object $F_t{}^{-1,T}$ therefore transforms the same way R_t and F_t transform, so we obtain a *third* set of objective time derivatives called the Oldroyd *upper* convected derivatives (Lai p 486 uses $\hat{T}_n$)

$\overset{\nabla}{T}_n \equiv [d^n J_U(t)/dt^n]$, n = 1,2,3... $\overset{\nabla}{T} \equiv \overset{\nabla}{T}_1$, $J_U(\tau) \equiv F_t{}^{-1}(\tau)\,T(\tau)\,F_t{}^{-1,T}(\tau)$ // = $\hat{T}_n$ (K.5.22)

The meaning of the term "convected" is explained below. It can be shown that

$\overset{\nabla}{T} = \hat{T} = d_tT - (\nabla\mathbf{v})T - T(\nabla\mathbf{v})^T$ // Lai p 486 (8.19.26) (K.5.23)

Covariant constitutive equations

Constitutive equations involving an objective time derivative of the stress tensor are called "rate type constitutive equations". Here are some models for incompressible fluids :

$T = -pI + S$ where $S + \lambda\overset{o}{S} = 2\mu D$ // a convected Maxwell fluid

$T = -pI + S$ where $S + \lambda(\partial S/\partial t) = 2\mu D$ // linear Maxwell fluid, see below (non-covariant)

$T = -pI + S$ where $S = 2\mu D$ // Newtonian fluid

$T = -pI + S$ where $S + \lambda_1\overset{o}{S} = 2\mu(D + \lambda_2\overset{o}{D})$ // a corotational Jeffrey fluid

$T = -pI + S$ where $S + \lambda_1\overset{\nabla}{T} = 2\mu(D + \lambda_2\overset{\nabla}{T})$ // Oldroyd fluid A (K.5.24)

Fluids with stress time derivatives in their constitutive equations exhibit both elastic and viscous behavior at the same time. Pull on a chunk of such a fluid and the pull is initially resisted by an elastic force, but after a while the internal stress field damps out (molasses, honey) and that elastic force goes away, as if the fluid were microscopically constructed of little springs and dragging dashpots. When a constitutive equation includes a time derivative of stress, the "response" (in this case $D = [(\nabla\mathbf{v}) + (\nabla\mathbf{v})^T]/2$) to the "stimulus" (T or S) includes factors of the form $e^{-t/c}$ where the c are decay time constants which are functions of the fluid parameters λ_i. In this case, the fluid has memory of its past over a time period less than these time constants, as with the honey example. For flow that is very slow relative to these time constants, the time derivative term may be neglected. In the moderately slow flow case, it can be shown that the distinction between the corotational time derivative $\overset{o}{S}$ and (dS/dt) can be neglected and then the convected Maxwell fluid shown above becomes the traditional linear Maxwell fluid which is modeled on those springs and dashpots with $S + \lambda (\partial S/\partial t) = 2\mu D$. (The time derivatives here are meant to act only on the second argument of $S(\mathbf{x},\tau)$ so may be regarded as partial derivatives.) If $\lambda = 0$, the linear Maxwell fluid becomes an (incompressible) Newtonian fluid like water which has no memory.

<u>Comment:</u> The linear Maxwell fluid equation $S + \lambda (\partial S/\partial t) = 2\mu D$ can be solved for S using the standard Green's Function method and the solution is $S(t) = 2 \int_{-\infty}^{t} dt' \, [(\mu/\lambda)e^{-(t-t')/\lambda}] \, D(t')$ where the bracketed quantity (the Green's Function or kernel) is called the stress relaxation function $\varphi(t-t')$. One can see in this solution the notion of memory (history) with time constant λ: the stress of the present is a function of the rate of deformation D going on in the entire past history. This solution fits into the "simple fluid" form shown earlier, where recall that $D = (1/2)A_1$ and $A_1 = [\, \partial_\tau C_t(\mathbf{x},\tau)]^{\tau=t}$.

Our main point is to demonstrate the construction of constitutive equations which are covariant with respect to rotations, and which therefore can contain only tensors which in fact transform as tensors under rotations. In continuum mechanics, such tensors are said to be objective tensors.

<u>Interpretation of the adjusted stress tensors discussed above.</u>

In Chapter 2 we discuss the notion of the transformation of a contravariant vector $\mathbf{V}' = R\mathbf{V}$ in developmental notation. In x'-space, the vector components are $V'_i = R_{ij}V_i$ where V_i are the components in x-space. If R is a rotation matrix, then the unit basis vectors in the two spaces can be taken as Cartesian, call them $\mathbf{u'_n}$ in x'-space and $\mathbf{u_n}$ in x-space. We have these two expansions of $\mathbf{V}$:

$$\mathbf{V} = \Sigma_n \, V_n \, \mathbf{u_n} = \Sigma_n \, V'_n \, \mathbf{u'_n} \qquad \text{where} \qquad V_n = \mathbf{V} \bullet \mathbf{u_n} \qquad V'_n = \mathbf{V} \bullet \mathbf{u'_n} \ . \tag{K.5.25}$$

In the "active view" of things, we can think of $\mathbf{V}' = R\mathbf{V}$ as creating a new vector $\mathbf{V}'$ in x-space from the old vector $\mathbf{V}$ by rotating the vector $\mathbf{V}$ by R. In the "passive view", we think of the V'_i as the components of the original vector $\mathbf{V}$ projected onto the backwards-rotated basis vectors $\mathbf{u'_n} = R^{-1}\, \mathbf{u_n}$. To verify this relation between the basis vectors, we can write

$$V'_n = \mathbf{V} \bullet \mathbf{u'_n} = \mathbf{V} \bullet R^{-1}\mathbf{u_n} = R\mathbf{V} \bullet RR^{-1}\mathbf{u_n} = R\mathbf{V} \bullet \mathbf{u_n} = \mathbf{V}' \bullet \mathbf{u_n} = V'_n \ . \tag{K.5.26}$$

So one can think *either* of $\mathbf{V}$ being rotated forward in x-space into $\mathbf{V}'$ where $\mathbf{V}'$ has x-space components V'_n , *or* one can think of the V'_n as the components of $\mathbf{V}$ one measures in frame that is backwards rotated by R^{-1} , that is, $\mathbf{u'_n} = R^{-1}\, \mathbf{u_n}$.

Consider then a rank-2 tensor M that transforms as shown in (5.7.1) according to $M' = R\, M\, R^T$. The passive interpretation is that the components M'_{ij} are those one observes in a frame of reference whose basis vectors are rotated by R^{-1} relative to the basis vectors of the unprimed frame, just as in the vector case of the last paragraph. If we now set $R = \mathcal{R}^{-1}$, then $M' = \mathcal{R}^{-1}\, M\, (\mathcal{R}^{-1})^T$ tells us that the components M'_{ij} of tensor M are those measured in a frame whose basis vectors are rotated forward by $\mathcal{R}$ relative to the unprimed frame. If it happens that $\mathcal{R} = \mathcal{R}(t)$, we would say that M'_{ij} are the components of M which are observed in a frame of reference which is rotating by $\mathcal{R}(t)$ relative to the frame of the unprimed components M_{ij}. The basis vectors of the primed frame are then $\mathbf{u'_n} = \mathcal{R}\, \mathbf{u_n}$.

With this long-winded introduction, we now consider the corotational stress tenser $J_t(\tau)$ shown in (K.5.7),

$$J_t(\tau) \equiv R_t^T(\tau)\, T(\tau)\, R_t(\tau) \ . \qquad \textit{(K.5.7)} \qquad \text{(K.5.27)}$$

Since $R_t(\tau)$ is a rotation, $R_t^T(\tau) = R_t^{-1}(\tau)$, so we have, suppressing τ,

$$J_t = R_t^{-1} T\, R_t \ . \qquad \text{(K.5.28)}$$

Therefore, we can regard $(J_t)_{ij}$ as T'_{ij}, the components of stress T measured in a frame which is rotating by R_t relative to the frame in which the T_{ij} are measured. Since this primed frame rotates by R_t relative to the unprimed frame, it is called a corotating frame, and J_t is then called the corotational stress, and its time derivative is called the corotational stress rate.

We next consider the upper Oldroyd stress defined above in (K.5.22),

$$J_U \equiv F_t^{-1}\, T\, (F_t^{-1})^T \ . \qquad \textit{(K.5.22)} \qquad \text{(K.5.29)}$$

In analogy with the above discussion, we can regard $(J_U)_{ij}$ as T'_{ij}, the components of stress T measured in a frame which is *deforming* by F_t relative to the unprimed frame. That is to say, the basis vectors of the primed frame are given by $\mathbf{u'}_n = F_t\, \mathbf{u}_n$. In this case, since F_t is not a rotation, if the $\mathbf{u}_n$ start as unit vectors, then the $\mathbf{u'}_n$ are not unit vectors. One can think of each basis vector $\mathbf{u}_n$ as being aligned with its own dumbbell $d\mathbf{x}^{(n)}$ and then we have $d\mathbf{x}'^{(n)} = F_t\, d\mathbf{x}^{(n)}$. What this says is that the basis vectors are embedded in the fluid which deforms as it flows according to F_t. The basis vectors "convect" with the fluid, so this upper Oldroyd stress is called the upper convective stress tensor.

For the lower Oldroyd stress in (K.5.19) we have

$$J_L \equiv F_t^T\, T\, F_t \qquad \textit{(K.5.19)} \qquad \text{(K.5.30)}$$

and this cannot be written in the form $J_L = \mathcal{R}^{-1} M\, (\mathcal{R}^{-1})^T$ so this does not fit into our interpretive template. But in the next section we show that J_L is the covariant partner to the contravariant tensor J_U so they are both the same animal and we are happy to have the interpretation above for J_U.

The Oldroyd convected stresses in developmental and standard notation

In the previous section two adjusted stress tensors were introduced,

$$J_U \equiv F_t^{-1}\, T\, F_t^{-1T} \qquad \text{// upper}$$
$$J_L \equiv F_t^T\, T\, F_t \ . \qquad \text{// lower} \qquad \text{(K.5.31)}$$

In developmental notation, a covariant tensor gets an overbar while a contravariant one does not. If we assume that frame S is Cartesian, then $T = \overline{T}$ as was discussed for vectors in Section 5.9. We can interpret the above two equations in this manner

$$\begin{aligned} &J \equiv F_t^{-1}\, T\, F_t^{-1T} && \text{// upper} \\ &\overline{J} \equiv F_t^{T}\, \overline{T}\, F_t && \text{// lower} \end{aligned} \tag{K.5.32}$$

which we compare with (5.7.1) where we change generic matrix name M to T,

$$\begin{aligned} &T' = R\, T\, R^{T} && \text{// contravariant rank-2 tensor transforms this way} \\ &\overline{T}' = S^{T}\, \overline{T}\, S && \text{// covariant rank-2 tensor transforms this way} \quad . \end{aligned} \tag{K.5.33}$$

Setting $R = F_t^{-1}$ and $S = R^{-1} = F_t$ gives

$$\begin{aligned} &T' = F_t^{-1}\, T\, F_t^{-1T} && \text{// contravariant rank-2 tensor} \\ &\overline{T}' = F_t^{T}\, \overline{T}\, F_t && \text{// covariant rank-2 tensor} \end{aligned} \tag{K.5.34}$$

Therefore we identify

$$\begin{aligned} &J_U = J = T' = \text{contravariant stress tensor T viewed from a frame convecting at } F_t^{-1} \\ &J_L = \overline{J} = \overline{T}' = \text{covariant stress tensor T viewed from a frame convecting at } F_t^{-1} \end{aligned} \tag{K.5.35}$$

In Standard Notation the two equations

$$\begin{aligned} &J \equiv F_t^{-1}\, T\, F_t^{-1T} && \text{// upper} \\ &\overline{J} \equiv F_t^{T}\, \overline{T}\, F_t && \text{// lower} \end{aligned} \tag{K.5.36}$$

become

$$\begin{aligned} &J^{ij} = (F_t^{-1})^{i}{}_{a}\, (F_t^{-1})^{j}{}_{b}\, T^{ab} && (J_U)_{ij} = J^{ij} = \text{contravariant} \\ &J_{ij} = (F_t^{-1})_{i}{}^{a}\, (F_t^{-1})_{j}{}^{b}\, T_{ab} && (J_L)_{ij} = J_{ij} = \text{covariant} \end{aligned} \tag{K.5.37}$$

and this explains the meaning of the words "upper" and "lower" in respect to the Oldroyd objects. See footnote on Lai page 485.

References

Listed in alphabetical order by last name of first author. A quick web search on article title can usually locate documents with broken web links. Those below were last checked May 18, 2016.

L. A. Ahlfors, *Complex Analysis, 2nd Ed.* (McGraw-Hill, New York, 1966).

G. Backus, *Continuum Mechanics* (Samizdat Press, Golden Colo., 1997).

J.D. Bjorken and S.D. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill, New York, 1964).

E. B. Christoffel, "Ueber die Transformation der homogenen Differentialausdrücke zweiten Grades", *Journal für die reine and angewandte Mathemarik*, **70** (1869), 46–70, 241–245. This paper may be found in Christoffel's Collected Mathematical papers, Gesammelte Mathematische Abhandlungen, 2 vols. (Tuebner, Leipzig-Berlin, 1910), downloadable from Google books.

R. Hermann, *Ricci and Levi-Civita's Tensor Analysis Paper (English translation with comments)* (Math Sci Press, Brookline, MA, 1975, perhaps on-line) First, with much praise, Hermann provides an English translation of the French Ricci & Levi-Civita paper referenced below, updating words, phrases and symbols to the current day. Second, he inserts perhaps 180 pages of inline italicized text which translates the ideas of the paper into mathematical frameworks not known or not connected to by the authors (eg, fiber bundles, direct product spaces, Killing vectors, moving frames, group theory, etc.). Hermann presents material that was more simply understood by later authors (eg, E. Cartan). Earlier in 1966 Hermann wrote a book *Lie Groups for Physicists* which contains the group theory chunk of this added material. One realizes that differential geometry is a very large field touching upon many areas of Mathematics (a house with many mansions).

W.M. Lai, D. Rubin and E. Krempl, *Introduction to Continuum Mechanics, 4th Ed.* (Butterworth-Heinemann/Elsevier, Amsterdam, 2010). This book has had the same three authors since its first edition in 1974, and it is with apologies to the last two authors that we reference the book just as Lai when we should be saying Lai *et.al*. or perhaps LRK.

P. Lucht, *Tensor Products, Wedge Products and Differential Forms* (2016, http://user.xmission.com/~rimrock).

H. Margenau and G.M. Murphy, *The Mathematics of Physics and Chemistry, 2nd Ed.* (D. van Nostrand, London. 1956).

A. Messiah, *Quantum Mechanics* (John Wiley, New York, 1958). Reference is made to page 878 (Vol II) of the North-Holland 1966 fifth printing paperback two-volume set, Chapter XX paragraph 2.

P. Moon and D.E. Spencer, *Field Theory Handbook, Including Coordinate Systems, Differential Equations and their Solutions* (Springer-Verlag, Berlin, 1961). This is the place to find explicit expressions for differential operators in many specific curvilinear coordinate systems. The book writes the Laplace and Helmholtz equations in various systems and discusses separability and the special functions that arise in the separated equations.

P.M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953).

M.M.G. Ricci, T. Levi-Civita, "Méthodes de calcul différentiel absolu et leurs applications", Mathematische Annalen (Springer) 54 (1–2): 125–201 (March 1900). This huge 77 page paper is sometimes referred to as "the bible of tensor analysis". Levi-Civita was a student of Ricci and they worked together on this paper and later elaborations. Their work was instrumental in Einstein's later discovery of general relativity. The title is "Methods of absolute differential calculus and their applications". Absolute differential calculus was the authors' phrase for what is now called Tensor Analysis/Calculus/Algebra. The word absolute referred to the idea of equations being covariant (see Section 7.15 above).

R. Sjamaar, *Manifolds and Differential Forms* (2015, http://www.math.cornell.edu/~sjamaar/manifolds).

I. Stakgold, *Boundary Value Problems of Mathematical Physics, Volumes 1 and 2* (Macmillan, London, 1967).

J.J. Sylvester, "On the General Theory of Associated Algebraical Forms" (Cambridge and Dublin Math. Journal, VI, pp 289-293, 1851). This paper appears in H.F. Baker, Ed., *The Collected Mathematical Papers of James Joseph Sylvester* (Cambridge University Press, 1901).

S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (John Wiley & Sons, New York, 1972).

E. B. Wilson (notes of J.W. Gibbs), *Vector Analysis* (Dover, New York, 1960)