

Tensor Analysis and Curvilinear Coordinates
Phil Lucht
Rimrock Digital Technology, Salt Lake City, Utah 84103
last update: February 16, 2012
rimrock@xmission.com

Overview and Summary	6
1. The Transformation F: invertibility, coordinate lines, and level surfaces.....	9
Example 1: Polar coordinates (N=2).....	9
Example 2: Spherical coordinates (N=3)	10
Cartesian Space and Quasi-Cartesian Space	12
Pictures A,B,C and D.....	12
Coordinate Lines	13
Example 1: Polar coordinates, coordinate lines	14
Example 2: Spherical coordinates, coordinate lines	14
Level Surfaces.....	15
2. Linear Local Transformations associated with F : scalars and two kinds of vectors	17
(a) Scalars.....	18
(b) Contravariant vectors	19
(c) Covariant vectors	19
(d) Bar notation	20
(e) Origin of the names contravariant and covariant.....	20
(f) Other vector types?	21
(g) Linear transformations	21
(h) Vectors that are contravariant by definition	22
(i) Vector Fields	22
(j) Names and symbols.....	23
(k) Definition of the words "scalar" and "vector".....	24
3. Tangent Base Vectors e_n and Inverse Tangent Base Vectors u'_n	25
(a) Definition of the e_n ; the e_n are the columns of S	26
(b) e_n as a contravariant vector	27
(c) a semantic question: unit vectors.....	28
Example 1: Polar coordinates, tangent base vectors	28
Example 2: Spherical Coordinates, tangent base vectors.....	29
(d) The inverse tangent base vectors u'_n and inverse coordinate lines	31
Example 1: Polar coordinates: inverse tangent base vectors and inverse coordinate lines	31
4. Notions of length, distance and scalar product in Cartesian Space.....	33
5. The Metric Tensor	35
(a) Definition of the metric tensor.....	35
(b) Inverse of the metric tensor.....	36
(c) A metric tensor is symmetric.....	37
(d) $\det(g)$ and g_{nn} of a Cartesian-generated metric tensor are non-negative	37
(e) Definition of two kinds of rank-2 tensors.....	38

(f) Proof that the metric tensor and its inverse are both rank-2 tensors	39
(g) Metric tensor converts vector types.....	40
(h) Vectors in Cartesian space	41
(i) Metric tensor: covariant scalar product and norm.....	41
(j) Metric tensor and tangent base vectors	43
(k) The Jacobian J	44
(ℓ) Some relations between g , R and S in Picture C.....	47
Example 1: Polar coordinates: metric tensor and Jacobian.....	48
Example 2: Spherical coordinates: metric tensor and Jacobian	49
(m) Special Relativity and its Metric Tensor: vectors and spinors	49
(n) General Relativity and its Metric Tensor	53
(o) Continuum Mechanics and its Metric Tensor	54
6. Reciprocal Base Vectors E_n and Inverse Reciprocal Base Vectors U'_n.....	55
(a) Definition of the E_n	55
(b) The Dot Products and Reciprocity (Duality).....	56
(c) Covariant partner for E_n	57
(d) Summary of the basic facts:	58
(e) Repeat the above for the inverse transformation: definition of the U'_n	58
(f) Expanding vectors on different sets of basis vectors	59
(g) Another way to write the E_n	62
(h) Comparison of $\underline{e}_n$ and E_n	63
(i) Handedness of the e_n and the sign of $\det(S)$	64
7. Translation to the Standard Notation	67
(a) Outer Products	67
(b) Mixed Tensors and Notation Issues	67
(c) The up/down bell goes off.....	68
(d) Some Preliminary Translations: raising and lowering indices on a vector with g	69
(e) Contraction of a Pair of Indices	70
(f) Dealing with the matrix R	71
(g) Repeat the above section for S	72
(h) About ϵ and δ	72
(i) Further development of the Standard Notation	73
(j) Tensors of Rank n , direct products, Lie groups, symmetry and Ricci-Levi-Civita.....	80
(k) The Contraction Tilt-Reversal Rule	83
(ℓ) The Contraction Neutralization Rule	84
(m) Raising and lowering indices on g	85
(n) Other forms of R	86
(o) Summary of facts about R	86
(p) Repeat all the above for S	86
(q) Theorem: $S^a_b = R^a_b$ and $S^b_a = R^b_a$ (reflect indices in vertical line between them).....	87
(r) Orthogonality Rules	88
(s) The tangent and reciprocal base vectors and expansions on same	89
(t) Comment on Covariant versus Contravariant	92
(u) The Significance of Tensor Analysis	93
(v) The Christoffel Business: covariant derivatives.....	95

(w) Expansions of higher order tensors	97
8. Transformation of Differential Area, Volume and Length	99
(a) The differential N-piped mapping	99
(b) Properties of the finite N-piped spanned by the e_n in x-space.....	101
(c) Objects in x-space.....	101
(d) Tensor density weights of objects	102
(e) Objects in x'-space	103
(f) Evaluation of x-space edge and area magnitudes	104
Examples of area magnitude transformation for $N = 2,3,4$	106
Example 2: Spherical Coordinates: area patches	107
(g) Evaluation of x'-space edge, area and volume magnitudes	108
(h) Summary of length, area and volume transformation results	109
(i) Transformation of Differential Volume applied to Integration.....	110
(j) Interpretations of the Jacobian	112
9. The Divergence in curvilinear coordinates	113
(a) Geometric Derivation of the Curvilinear Divergence Formula	113
(b) Various expressions for $\text{div } B$	116
(c) Translation from Picture B to Picture M&S	118
(d) Comparison of various authors' notations	119
(e) The Christoffel derivation of $\text{div } B$	120
10. The Gradient in curvilinear coordinates.....	122
(a) Expressions for $\text{grad } f$	122
(b) Expressions for $\text{grad } f \bullet B$	124
11. The Laplacian in curvilinear coordinates	126
12. The Curl in curvilinear coordinates	128
(a) Definition of $\text{curl } B$	128
(b) Computation of the line integral.....	129
(c) Solving for the curl.....	131
(d) Various forms of the curl.....	132
(e) The curl in orthogonal coordinate systems.....	133
(f) The curl in $N > 3$ dimensions.....	134
13. The Vector Laplacian in curvilinear coordinates.....	136
(a) Derivation of the Vector Laplacian in general curvilinear coordinates	136
(b) The Vector Laplacian in orthogonal curvilinear coordinates	139
(c) The Vector Laplacian in Cartesian coordinates.....	141
14. Summary of Differential Operators in curvilinear coordinates	142
(a) divergence.....	143
(b) gradient and gradient dot vector	143
(c) Laplacian	144
(d) curl.....	144
(e) vector Laplacian	145
Example 1: Polar coordinates: a practical curvilinear notation.....	146
Appendix A: Reciprocal Base Vectors the Hard Way.....	148
(a) Definition of E	148

(b) Simpler notation	149
(c) Generalized Cross Product of N-1 vectors of dimension N	149
(d) Missing Man Formation	151
(e) Apply this Notation to E	151
(f) Compute $E_m \bullet e_n$	152
(g) Compute $E_n \bullet E_m$	153
(h) Summary of relationship between the tangent and reciprocal base vectors	153
(i) Another Cross Product Notation and another expression for E	154
Appendix B: The Geometry of Parallelepipeds in N dimensions.....	155
(a) Preliminary: Equation of a plane in N dimensions	155
(b) N-pipeds and their Faces in Various Dimensions	156
The 1-piped	156
The 2-piped	156
The 3-piped	158
The N-piped	159
(c) The question of inward versus outward facing normal vectors.	160
(d) The Face Area and Volume of N-pipeds in Various Dimensions	161
The 2-piped	161
The 3-piped	162
The 4-piped	164
The N-piped	166
(e) Summary of Main Results of this Appendix	168
Appendix C: Elliptical Polar Coordinates (N=2, non-orthogonal).....	170
(a) Elliptical polar coordinates	170
(b) Forward coordinate lines	171
(c) Inverse coordinate lines	171
(d) Drawing a contravariant vector V in x-space: the meaning of V'_n	172
(e) Drawing a contravariant vector V' in x'-space: the two "views"	173
(f) Drawing the specific contravariant vector dx in x-space and x'-space	175
(g) Study of how dx transforms in the mapping between x-space and x'-space	176
(h) Derivation of the Jacobian Integration Rule.....	177
Appendix D. Tensor Densities and the ϵ tensor.....	180
(a) Definition of a tensor density	180
(b) A few facts about tensor densities	181
(c) Theorem about Totally Antisymmetric Tensors: there is really only one	183
(d) The contravariant ϵ tensor	184
(e) Some facts about the ϵ tensor	185
(f) The covariant ϵ tensor	187
(g) Generalized cross products.....	188
(h) The tensorial nature of curl B.....	188
(i) Tensor $\mathcal{E}$ as a weight 0 version of ϵ : three conventions.....	189
(j) Representation of ϵ , $\epsilon\epsilon$ and contracted $\epsilon\epsilon$ as determinants	192
(k) Covariant forms of the previous section results	198
Appendix E: Tensor Expansions: direct product, polyadic and operator notation.....	200
(a) Direct Product Notation.....	200

(b) Tensor Expansions and Bases	201
(c) Polyadic Notation	203
(d) Dyadic Products	204
(e) Transpose notation for dyadics	204
(f) Large and small dots used with dyadics	205
(g) Operators and Matrices for Rank-2 tensors	206
Appendix F: Expansion of the gradient of a vector ($\nabla \mathbf{v}$) in curvilinear coordinates	210
(a) Expressing $\partial_c v_d$ in terms of x' -space objects	211
(b) Expressing $u_d u_c^T$ in terms of $e_i e_j^T$	211
(c) Combining the two steps	212
(d) Alternate forms of the ($\nabla \mathbf{v}$) expansion	213
(e) Special case of orthogonal coordinates	214
References	217

Overview and Summary

This paper develops elementary tensor analysis (also known as tensor algebra or tensor calculus) 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. The general mathematical classification of this subject is now called differential geometry.

Two somewhat different applications of tensor analysis are treated concurrently. One is the subject of curvilinear coordinates in N dimensions, while the other involves transformations connecting "frames of reference". These transformations could be spatial rotations, 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 second set of applications. On the other hand, all the basic expressions for the standard differential operators in general curvilinear coordinates are derived from scratch. These results are often stated but not so often derived.

The first six sections 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, the seventh section translates this developmental notation to the Standard Notation in use today. The eighth section treats transformations of length, area and volume and then the curvilinear differential operator expressions are derived, one per section, with a summary in the final section.

The information is presented informally as if it were a set of lectures. Little attention is paid to 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.

Section 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.

Section 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.

Section 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.

Section 4 is brief review of the notions of norm, metric and scalar product in Cartesian Space.

Section 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 shown how $\bar{g}$ applied to a contravariant vector $\mathbf{V}$ produces a vector that

is covariant $\bar{\mathbf{V}} = \bar{\mathbf{g}} \mathbf{V}$, and conversely $\mathbf{g} \bar{\mathbf{V}} = \mathbf{V}$. In Cartesian space $\mathbf{g} = 1$, 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 Section 3. The Jacobian J and associated functions are defined, though the significance of J is deferred to Section 8. The last two subsections briefly discuss the connection between tensor analysis and special and general relativity with a mention of spinor algebra.

Section 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 $\mathbf{e}_n$, the vector components in the expansion are in fact those of the contravariant vector in x' -space, $V'_i = R_{ij} V_j$. This fact proves useful in later sections 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.

Section 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, 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. The third last subsection discusses the significance of tensor analysis with respect to physics in terms of covariant equations, and the second last broaches the topic of the covariant derivative of a vector field with its associated Christoffel symbols. Finally, the last subsection shows how to expand tensors of any rank in various bases and notations, and considers the expression of the non-tensor $\nabla \mathbf{v}$ in curvilinear coordinates.

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

Section 8 shows how differential length, area and volume transform under $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. This section 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 $\sqrt{\text{cof}(g'_{nn})}$. Heavy use is made of the results of Appendices A and B, as outlined below.

Sections 9 through 13 use the information of Section 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.

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

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 Section 6, and the covariant dot products involving $\mathbf{E}_n$ and $\mathbf{e}_n$ are verified.

Appendix B presents the geometry of a parallelepiped 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 $d\mathbf{A}^n = |\det(S)| \mathbf{E}_n(\prod_{i \neq n} dx^i)$, this equation is then used in Section 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 Section 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^m 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 vector densities and the Levi-Civita ϵ tensor, and includes a derivation of all the $\epsilon\epsilon$ contraction formulas and their covariant statements.

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

Appendix F then applies the methods of Appendix E to the expression of the rank-2 tensor-like object $(\nabla \mathbf{v})$ in general curvilinear coordinates, this object being of interest in continuum mechanics.

Notations

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

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")

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

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

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

Maple = a computer algebra system similar to Mathematica

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

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

$$\mathbf{x}' = \mathbf{F}(\mathbf{x}) \quad \mathbf{F}: \mathbb{R}^N \rightarrow \mathbb{R}^N$$

defined by a set of N continuous (C_2) functions F_i , each of N variables,

$$x'_1 = F_1(x_1, x_2, x_3 \dots x_N)$$

$$x'_2 = F_2(x_1, x_2, x_3 \dots x_N)$$

...

$$x'_N = F_N(x_1, x_2, x_3 \dots x_N)$$

If all functions F_i are linear in all of their arguments, then the mapping $\mathbf{F}: \mathbb{R}^N \rightarrow \mathbb{R}^N$ 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,

$$\mathbf{x}' = \mathbf{F}(\mathbf{x}) \quad \mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}') ,$$

or in an equivalent notation

$$\mathbf{x}' = \mathbf{x}'(\mathbf{x}) \quad \mathbf{x} = \mathbf{x}(\mathbf{x}')$$

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

The functions F_i must be C_1 continuous to support the linearization derivatives appearing in Section 2, and they must be C_2 continuous to support some of the differential operators expressed in curvilinear coordinates in Sections 9-14 and the covariant derivative in Section 7 (v).

Example 1: Polar coordinates (N=2)

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

$$\mathbf{x} = (x_1, x_2) = (x, y)$$

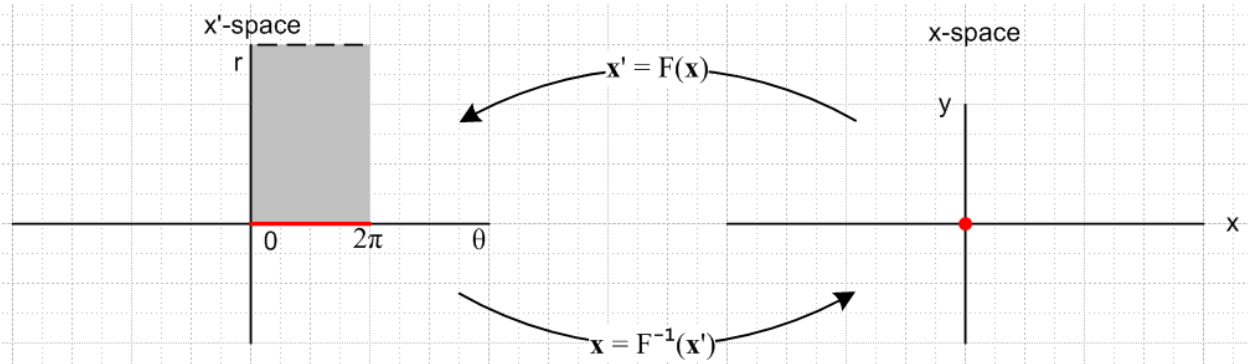
$$\mathbf{x}' = (x'_1, x'_2) = (\theta, r) \quad // \text{ note that } r = x_2'$$

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

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

(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:



The domain of $\mathbf{x}' = F(\mathbf{x})$ in x -space on the right is all of $\mathbb{R}^2$, but the range in x' -space is shown in gray. Imitating the language of complex variables, we can regard this gray range as depicting the principle branch of the multi-variable function $\mathbf{x}' = F(\mathbf{x})$. 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 rectangles. 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.

(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 area 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 fully 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,

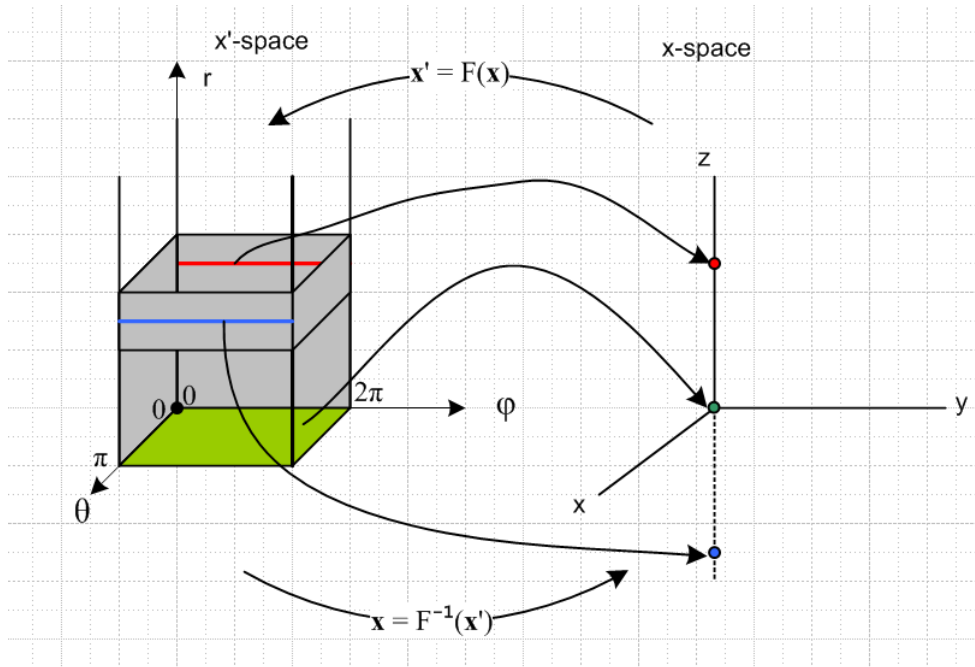
$$\begin{aligned} \mathbf{x} &= (x_1, x_2, x_3) = (x, y, z) \\ \mathbf{x}' &= (x_1', x_2', x_3') = (r, \theta, \phi) \end{aligned}$$

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

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

(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



The domain of $\mathbf{x}' = F(\mathbf{x})$ in x -space on the right is all of E^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 $\mathbf{x}' = F(\mathbf{x})$. 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 E^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 E^N in which the distance between two vectors is given by the formula

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^N (x_i - y_i)^2} \Rightarrow [d(\mathbf{x} + d\mathbf{x}, \mathbf{x})]^2 = \sum_{i=1}^N (dx_i)^2 \quad \text{metric tensor} = \text{diag}(1, 1, 1, \dots, 1)$$

as discussed in Section 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 was $D'^2 = (\theta - \theta')^2 + (r - r')^2$, 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 $\text{diag}(1, 1)$.

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 $g^C = 1 = \text{diag}(1, 1, \dots, 1)$, the above equations can be written

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

where repeated indices are implicitly summed (sometimes called the Einstein 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

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

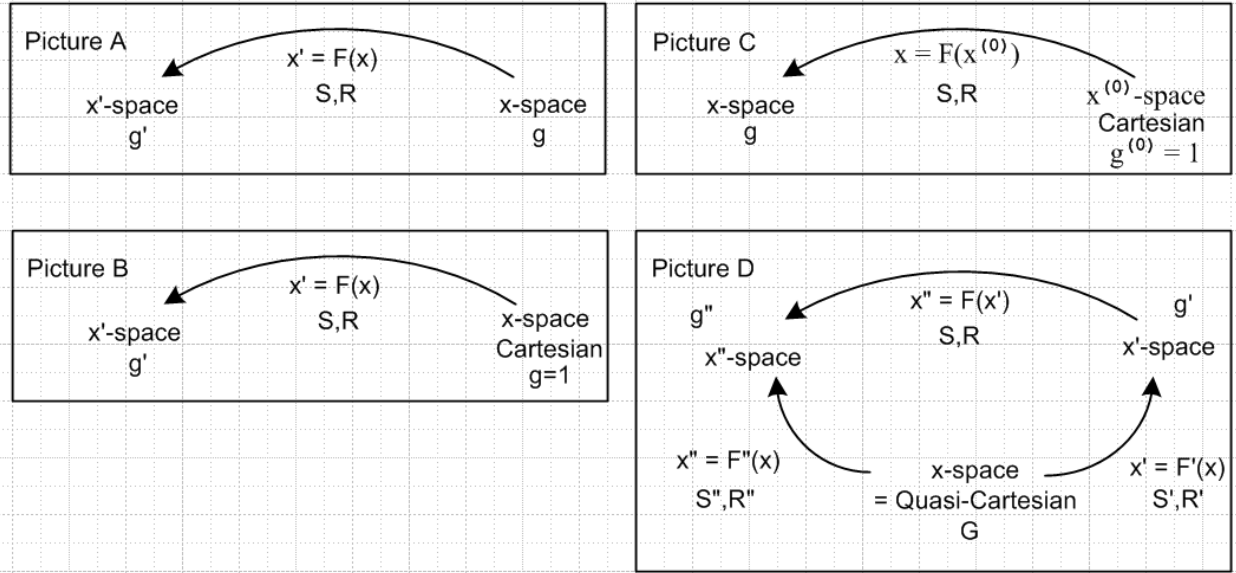
and of course this allows for the possibility of a negative distance squared (see Section 5 (i)).

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 $G = \text{diag}(1, -1, -1, -1)$.

Pictures A,B,C and D

We shall always 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.



The matrices R and S are associated with transformation F as described in Section 2 below, while G and 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 Sections 9-14 Picture M&S (Moon & Spencer) is introduced for the special purpose of displaying the differential operator expressions. This is Picture B with $x' \rightarrow u$ and $g' \rightarrow g$ 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 $x = F^{-1}(x')$ then lies on a unique intersection of coordinate lines.

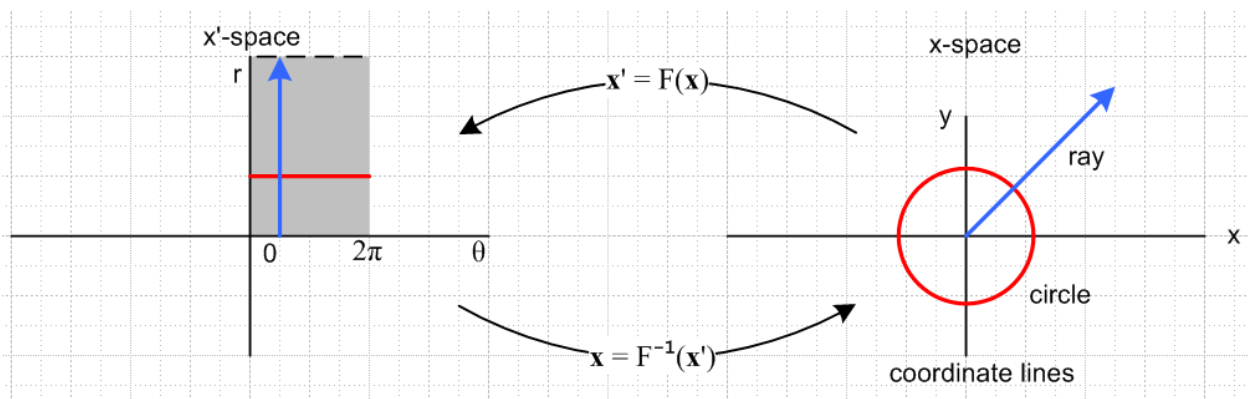
For example, in spherical coordinates we start with some (x, y, z) in x -space and compute the $x'_i = (r, \theta, \phi)$ in x' -space. Our point x in x -space then lies on the r -coordinate line whose label is r , it lies on the θ -coordinate line whose label is θ , and it lies on the ϕ -coordinate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

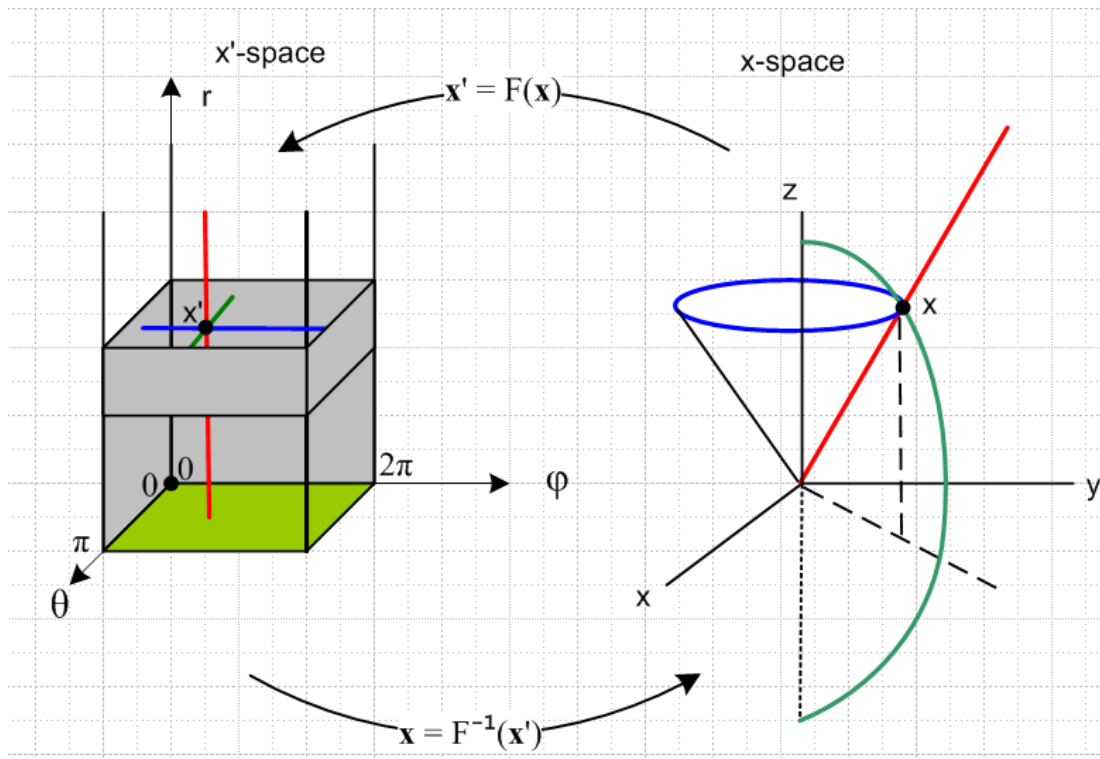
Here are some coordinate lines for our prototype $N=2$ non-linear transformation, Cartesian to polar coordinates:



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.



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

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

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, $\{a'_i{}^{(1)}, a'_i{}^{(2)}, \dots\}$, 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 $f(x_1, x_2, \dots, x_N) = k$, 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:

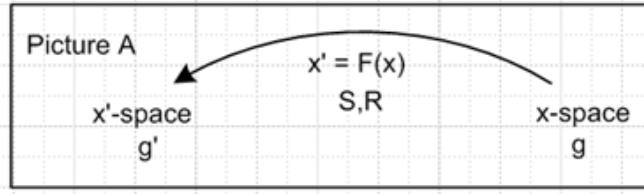
The green coordinate line is the intersection of two level surfaces: half-plane and sphere.

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

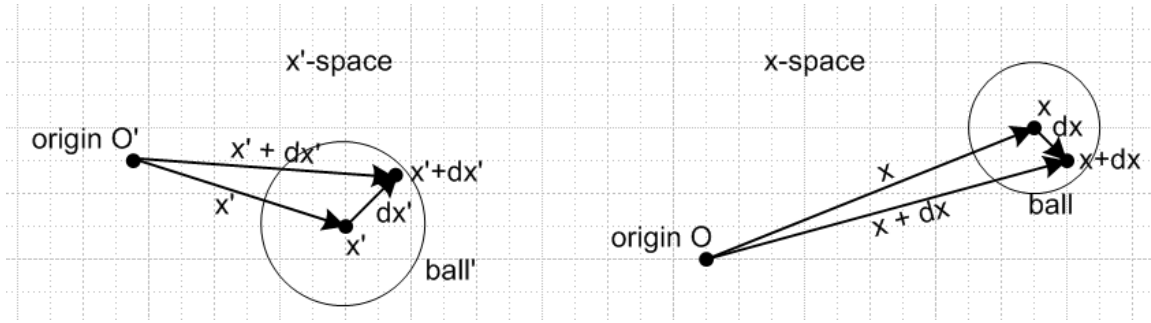
The blue coordinate line is the intersection of two level surfaces: sphere and cone.

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

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



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



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 $d\mathbf{x}$ which maps into a small vector in x' -space called $d\mathbf{x}'$. Since F was assumed invertible, it must be invertible locally in these two balls. That is, given a $d\mathbf{x}$ above, one can determine $d\mathbf{x}'$, 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,

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

$$\Rightarrow dx'_i = \sum_k (\partial F_i(\mathbf{x}) / \partial x_k) dx_k$$

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

$$dx'_i = \sum_k (\partial x'_i / \partial x_k) dx_k = \sum_k R_{ik} dx_k \quad R_{ik} \equiv (\partial x'_i / \partial x_k)$$

Doing the same operation in the other direction gives

$$dx_i = \sum_k (\partial x_i / \partial x'_k) dx'_k = \sum_k S_{ik} dx'_k \quad S_{ik} \equiv (\partial x_i / \partial x'_k)$$

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

$$\begin{array}{llll}
d\mathbf{x}' = \mathbf{R}(\mathbf{x}) d\mathbf{x} & R_{i\mathbf{k}}(\mathbf{x}) \equiv (\partial x'_{i}/\partial x_{\mathbf{k}}) & \mathbf{R} = \mathbf{S}^{-1} & // d\mathbf{x}'_{i} = R_{i\mathbf{j}} d\mathbf{x}_{\mathbf{j}} \\
d\mathbf{x} = \mathbf{S}(\mathbf{x}') d\mathbf{x}' & S_{i\mathbf{k}}(\mathbf{x}') \equiv (\partial x_{i}/\partial x'_{\mathbf{k}}) & \mathbf{S} = \mathbf{R}^{-1} & // d\mathbf{x}_{i} = S_{i\mathbf{j}} d\mathbf{x}'_{\mathbf{j}}
\end{array}$$

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

$$(\mathbf{RS})_{i\mathbf{j}} = \sum_{\mathbf{k}} R_{i\mathbf{k}} S_{\mathbf{k}\mathbf{j}} = \sum_{\mathbf{k}} (\partial x'_{i}/\partial x_{\mathbf{k}}) (\partial x_{\mathbf{k}}/\partial x'_{\mathbf{j}}) = \sum_{\mathbf{k}} \frac{\partial x'_{i}}{\partial x_{\mathbf{k}}} \frac{\partial x_{\mathbf{k}}}{\partial x'_{\mathbf{j}}} = \frac{\partial x'_{i}}{\partial x'_{\mathbf{j}}} = \delta_{i,\mathbf{j}}$$

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

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

Matrices $\mathbf{R}(\mathbf{x})$ and $\mathbf{S}(\mathbf{x}')$ 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 $\mathbf{R}$ and $\mathbf{S}$ are likely to change. So $\mathbf{R}$ and $\mathbf{S}$ 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 $dx_{i\mathbf{j}}$ and the same question for $dx'_{i\mathbf{j}}$ and $dx'_{i\mathbf{j}}$. As will be shown in Section 6 (f),

$$\begin{array}{ll}
d\mathbf{x} = \sum_{\mathbf{n}} dx_{\mathbf{n}} \mathbf{u}_{\mathbf{n}} & \text{where the } \mathbf{u}_{\mathbf{n}} \text{ are } x\text{-space axis-aligned basis vectors of the form } \mathbf{u}_1 = (1,0,0,..0) \\
d\mathbf{x}' = \sum_{\mathbf{n}} dx'_{\mathbf{n}} \mathbf{e}'_{\mathbf{n}} & \text{where the } \mathbf{e}'_{\mathbf{n}} \text{ are } x'\text{-space axis-aligned basis vectors of the form } \mathbf{e}'_1 = (1,0,0,..0)
\end{array}$$

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

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 Section 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 $\mathbf{R}$ or $\mathbf{S}$? The answer is yes, but first we discuss scalars.

(a) Scalars

A quantity is a **scalar** with respect to transformation $\mathbf{F}$ 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 $\mathbf{x}$, while φ' is evaluated at $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. As noted in section (k) below, one could be more precise by referring to the objects described here as a "tensorial scalar" and a "tensorial scalar field".

(b) Contravariant vectors

If transformation F (possibly non-linear) transforms x -space to x' -space without affecting time, then consider the familiar velocity vector,

$$v_i = dx_i/dt \quad \Rightarrow \quad \mathbf{v} = d\mathbf{x}/dt$$

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:

$$\mathbf{v}' = R(\mathbf{x}) \mathbf{v}$$

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

In contrast, unless $R(\mathbf{x})$ is a constant in space (which would be the case only if F were a linear transformation) $\mathbf{x}' \neq R(\mathbf{x}) \mathbf{x}$, so in general $\mathbf{x}$ itself is *not* a contravariant vector although $d\mathbf{x}$ is.

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

(c) Covariant vectors

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

$$\bar{\nabla}_i = \bar{\partial}_i = \partial/\partial x_i$$

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),

$$\bar{\nabla}'_i = \bar{\partial}'_i = \frac{\partial}{\partial x'_i} = \frac{\partial x_k}{\partial x'_i} \frac{\partial}{\partial x_k} = S_{ki} \bar{\partial}_k = S^T_{ik} \bar{\partial}_k = S^T_{ik} \bar{\nabla}_k$$

$$\Rightarrow \bar{\nabla}' = S^T \bar{\nabla}$$

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

$$\bar{\nabla}'_i \phi'(\mathbf{x}') = \frac{\partial}{\partial x'_i} \phi'(\mathbf{x}') = \frac{\partial x_k}{\partial x'_i} \frac{\partial}{\partial x_k} \phi(\mathbf{x}) = S^T_{ik} \bar{\nabla}_k \phi(\mathbf{x})$$

$$\Rightarrow \bar{\nabla}'\phi'(\mathbf{x}') = S^T \bar{\nabla}\phi(\mathbf{x})$$

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

Vectors that transform according to $\mathbf{V}' = \mathbf{S}^T(\mathbf{x}) \mathbf{V}$ such as the gradient operator $\bar{\nabla}$ 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

$$\bar{\mathbf{E}} = -\bar{\nabla} \Phi \quad \bar{E}_i = -\bar{\partial}_i \Phi = -\partial\Phi/\partial x_i$$

(d) 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 $\bar{\mathbf{V}}$ with components $\bar{V}_i$. This is why overbars were placed on $\bar{\nabla}$ and $\bar{\partial}_i$ and $\bar{\mathbf{E}}$ in the previous section. We call this our "developmental notation", as distinct from the Standard Notation introduced in Section 7.

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

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

One could imagine replacing $\mathbf{S}$ with some $\mathbf{Q}^T$ to make the second equation more like the first, but of course then $\mathbf{R}\mathbf{Q}^T = 1$ instead of $\mathbf{R}\mathbf{S} = 1$. In the Standard Notation, where there are four versions of the matrix $\mathbf{R}$, we shall see that $\mathbf{R} \rightarrow R^i_j$ and $\mathbf{S} \rightarrow S^i_j = R^i_j$ and $\mathbf{S}$ can be removed from the picture (see Section 7 (q)).

(e) Origin of the names contravariant and covariant

A justification of the terms covariant and contravariant is presented at the end of Section 7 (t), 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 existed) 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 $\mathbf{A}$ is an 3×3 matrix, then

the pair $F(\mathbf{A}\mathbf{x})$ and $G(\mathbf{A}\mathbf{x})$ are said to be covariant (or concurrent)
the pair $F(\mathbf{A}\mathbf{x})$ and $G(\mathbf{A}^{-1}\mathbf{x})$ 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 (t) this idea is applied to the transformation of two "things", where one thing is the component of a vector like V_n and the other thing is a basis vector onto which a vector is expanded. The connection is a bit distant, but the underlying concept carries through.

Notations like $\mathbf{y} = \mathbf{F}(\mathbf{A}\mathbf{x})$ 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 Gibbs' 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

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.

(f) 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 (j)). Some authors refer to the rank of a tensor as the order of a tensor.)

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 " $(1/2) \oplus (1/2)$ "). 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" $(1/2) \oplus 0$ and $0 \oplus (1/2)$, 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 (m).

(g) Linear transformations

For a *linear* transformation F , the matrix elements of R and S are constants and don't depend on $\mathbf{x}$ or $\mathbf{x}'$. The reason is fairly obvious. For linear $\mathbf{x}' = F(\mathbf{x})$ (an added constant would make F non-linear)

$$F(\alpha \mathbf{x}' + \beta \mathbf{y}') = \alpha F(\mathbf{x}') + \beta F(\mathbf{y}') \quad \Rightarrow \quad x'_i = F_{i1}x_1 + F_{i2}x_2 + \dots + F_{iN}x_N \quad // = F_i(\mathbf{x})$$

where the F_{ij} are constants independent of the coordinates, in which case

$$dx'_i = F_{i1} dx_1 + F_{i2} dx_2 + \dots + F_{iN} dx_N = \sum_k F_{ik} dx_k$$

so

$$R = F \quad \text{and} \quad S = F^{-1}$$

This is the situation with rotations and Lorentz transformations.

(h) Vectors that are contravariant by definition

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

$$\{ dx', dx \} \quad dx' = R(x) dx \quad \text{contravariant vector}$$

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

$$\{ V', V \} \quad V' \equiv R(x)V \quad \text{contravariant vector}$$

Conversely, one could start with some given V' and define $V \equiv S(x)V'$ (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 $x' \equiv R(x)x$, this x' conflicts with the existing $x' = F(x)$, so the method of contravariant by definition is unacceptable.

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 $V \equiv SV'$, 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.

Example 3: This time F is a general transformation and we start with $V' = e'_n$ which are a set of axis-aligned basis vectors in x' -space. We define vectors $V = e_n$ according to $e_n \equiv Se'_n$. Then $\{ e'_n, e_n \}$ form a vector which is "contravariant by definition" and $e'_n = R e_n$ ($R = S^{-1}$). Since the newly defined vector 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 e_n .

(i) Vector Fields

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

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

$$T'(\mathbf{x}') = T(\mathbf{x}) \quad \text{where} \quad \mathbf{x}' = \mathbf{F}(\mathbf{x})$$

If the transformation is a 3D rotation from frame S to frame S' , then T' is the temperature measured in frame S' at point $\mathbf{x}'$ and T is the temperature measured at the corresponding point $\mathbf{x}$ 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

$$T'(r,\theta) = T(x,y) = T(r\cos\theta, r\sin\theta) \neq T(r,\theta)$$

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 $\mathbf{x}' = \mathbf{F}(\mathbf{x})$:

$$\begin{aligned} \mathbf{V}'(\mathbf{x}') &= \mathbf{R} \mathbf{V}(\mathbf{x}) & \text{contravariant} & & \mathbf{R}_{i\mathbf{k}}(\mathbf{x}) &\equiv (\partial x'_i / \partial x_{\mathbf{k}}) & & \mathbf{R} = \mathbf{S}^{-1} \\ \bar{\mathbf{V}}'(\mathbf{x}') &= \mathbf{S}^T \bar{\mathbf{V}}(\mathbf{x}) & \text{covariant} & & \mathbf{S}_{i\mathbf{k}}(\mathbf{x}') &\equiv (\partial x_i / \partial x'_{\mathbf{k}}) &= \mathbf{S}_{\mathbf{k}i}^T(\mathbf{x}') \end{aligned}$$

Similar transformation rules apply to tensors of any rank. For example, the metric tensor g_{ab} (developmental notation) is a rank-2 contravariant tensor field and the transformation rule is this

$$g'_{ab}(\mathbf{x}') = R_{a\mathbf{a}} \cdot R_{b\mathbf{b}} \cdot g_{\mathbf{a}\mathbf{b}}(\mathbf{x}) \quad \text{or} \quad g'_{ab} = R_{a\mathbf{a}} \cdot R_{b\mathbf{b}} \cdot g_{\mathbf{a}\mathbf{b}}$$

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

$$V'^{\mu}(x'^{\alpha}) = \Lambda^{\mu}_{\nu} V^{\nu}(x^{\alpha}) \quad x'^{\mu} = \Lambda^{\mu}_{\nu} x^{\nu}$$

(j) Names and symbols

The matrix $\mathbf{R}_{i\mathbf{k}}(\mathbf{x}) = (\partial x'_i / \partial x_{\mathbf{k}})$ is called the **Jacobian matrix** for the transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$, while the matrix $\mathbf{S}_{i\mathbf{k}}(\mathbf{x}') = (\partial x_i / \partial x'_{\mathbf{k}})$ is then the Jacobian matrix of the inverse transformation $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$. The magnitude of the determinant of Jacobian matrix $\mathbf{S}$ will be shown in Section 8 (e) to have a certain significance, and that determinant is called "**the Jacobian**" $= \det(\mathbf{S}(\mathbf{x}')) \equiv J(\mathbf{x}')$.

The author has anguished over what names to give the matrices $\mathbf{R}$ and $\mathbf{S} = \mathbf{R}^{-1}$. One option was to use $\mathbf{R} = \mathbf{L}$, where $\mathbf{L}$ stands for the fact that this matrix is describing a Local coordinate system at point $\mathbf{x}$, or a

Linearized transformation. But L is always used for differential operators, 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, but J looks too much like "an integer" or angular momentum or "the Jacobian". T for Transformation might have been confused with the transformation F. 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 $RS = 1$.

The fact that vectors are processed by $N \times N$ 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 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**).

(k) Definition of the words "scalar" and "vector".

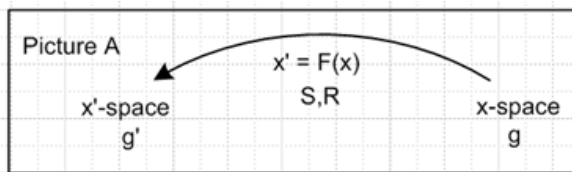
These words have multiple potential definitions. Although we shall lapse very frequently, the following set of definitions would allow for precision statements:

- A "scalar" is a single number (or expression), a 1-tuple.
- A "tensorial scalar" is a scalar that transforms under transformation F as a tensorial scalar, which is also known as a rank-0 tensor. An example would be $m' = m$, mass with respect to 3D rotations.
- A "vector" is an N-tuple of numbers.
- A "tensorial vector" is a vector that transforms under transformation F as either a contravariant vector or a covariant vector, so a tensorial vector is a rank-1 tensor.
- A "scalar field" is a single function of $\mathbf{x}$ (the x-space coordinates).
- A "tensorial scalar field" is a scalar field that transforms under transformation F as a tensorial scalar field, which is also known as a rank-0 tensor field. For example, $f'(\mathbf{x}') = f(\mathbf{x})$ is a scalar field.
- A "vector field" is an N-tuple of functions of $\mathbf{x}$ -- an N-tuple of scalar fields.
- A "tensorial vector field" is a vector field that transforms under transformation F as either a contravariant vector field or a covariant vector field, so a tensorial vector field is a rank-1 tensor field.

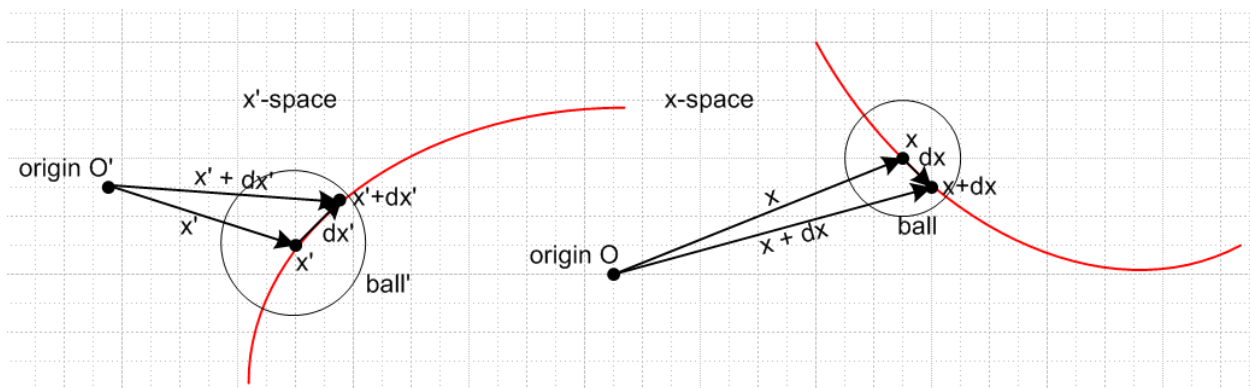
The notion of tensor densities described in Appendix D further complicates the nomenclature. One can have scalar densities and vector densities of various weights.

3. Tangent Base Vectors e_n and Inverse Tangent Base Vectors u'_n

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



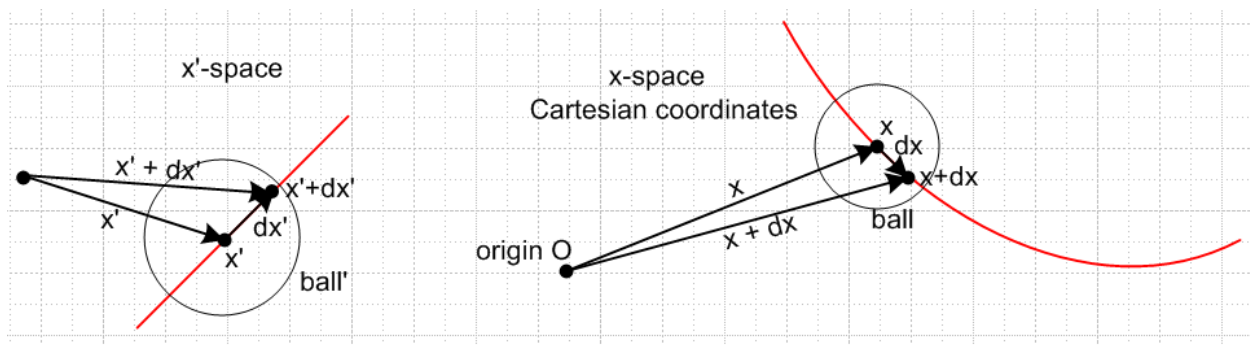
In the previous picture showing dx and dx' , one has much freedom to "try out" different differential vectors. For any dx one picks at point x , one gets some dx' according to $dx' = R(x) dx$. Consider this slightly enhanced version of the previous drawing (red curves added)



The point x in x -space (right side) can be regarded as lying on some arbitrary 1-dimensional curve in $\mathbb{R}^N$ shown on the right in red. Select dx 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 $dx' = R(x) dx$. A similar statement can be made starting instead with an arbitrary curve in x' -space. The tangent dx' there then maps into $dx = S(x') dx'$ 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,



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.

(a) Definition of the $\mathbf{e}_n$; the $\mathbf{e}_n$ are the columns of $\mathbf{S}$

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

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

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

$$d\mathbf{x}' = \mathbf{e}'_n dx'_n \quad // \text{ no implied sum on } n$$

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

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

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

$$\mathbf{e}_n \equiv \mathbf{S} \mathbf{e}'_n$$

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

$$\begin{aligned} d\mathbf{x} &= \mathbf{e}_n dx'_n \\ \Rightarrow dx_i &= (\mathbf{e}_n)_i dx'_n \end{aligned}$$

But of course $dx_i = S_{in} dx'_n$, and therefore

$$\begin{aligned} (\mathbf{e}_n)_i &= S_{in} = \partial x_i / \partial x'_n \\ \Rightarrow (\mathbf{e}_n)_i &= \partial x_i / \partial x'_n = S_{in} \end{aligned}$$

This says that the vectors $\mathbf{e}_n$ are the columns of the matrix $\mathbf{S}$:

$$\mathbf{S} = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \dots \mathbf{e}_N] \quad \text{matrix} = N \text{ columns}$$

These vectors $\mathbf{e}_n$ are called 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.

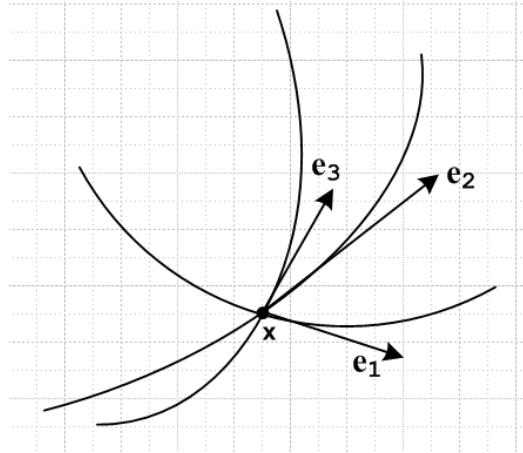
As one moves from $\mathbf{x}$ to some nearby point, the tangent base vectors all change slightly because in general $S = S(\mathbf{x}'(\mathbf{x}))$ and the $\mathbf{e}_n = \mathbf{e}_n(\mathbf{x})$ are the columns of S .

Since $\det(S) \neq 0$ due to our assumption that F was 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 $\hat{\mathbf{e}}_n = \mathbf{e}_n / |\mathbf{e}_n|$.

Some authors refer to the $\mathbf{e}_n$ as $\mathbf{g}_n$, but g will soon be overloaded with other meanings (metric tensors, and determinant of the covariant metric tensor) so we stick with $\mathbf{e}_n$, the notation used by Margenau and Murphy (p 193).

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 $\mathbf{x}$:



(b) $\mathbf{e}_n$ as a contravariant vector

The situation described above was this,

$$\begin{array}{lll} d\mathbf{x}' = \mathbf{e}'_n dx'_n & x'\text{-space} & // \text{ no implied sum on } n \\ d\mathbf{x} = \mathbf{e}_n dx'_n & x\text{-space} & // \text{ no implied sum on } n \end{array}$$

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

$$d\mathbf{x}' = R(\mathbf{x}) d\mathbf{x}$$

$$\mathbf{e}'_n dx'_n = R(\mathbf{x}) \mathbf{e}_n dx'_n$$

$$\mathbf{e}'_n = R(\mathbf{x}) \mathbf{e}_n$$

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

$$(\mathbf{e}'_n)_i = R_{ij} (\mathbf{e}_n)_j$$

$$\delta_{n,i} = R_{ij} S_{jn}$$

recovering the fact that $RS = 1$. This is an example of a vector being "contravariant by definition", as discussed in Section 2 (h).

These two expansions are easy to show just by verifying that components of both sides are the same:

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

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

(c) a semantic question: unit vectors

Above it was noted that $\mathbf{e}'_1 = (1,0,0,\dots)$. Should this be called "a unit vector" ? It will be seen below that in fact $|\mathbf{e}'_1| = \sqrt{g'_{11}} \neq 1$ where 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 Appendix C (e).

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:

$$\mathbf{x} = (x_1, x_2) = (x, y)$$

$$\mathbf{x}' = (x'_1, x'_2) = (\theta, r)$$

$$\mathbf{x} = F^{-1}(\mathbf{x}') \leftrightarrow \begin{aligned} x &= r \cos(\theta) & x_1 &= x'_2 \cos(x'_1) \\ y &= r \sin(\theta) & x_2 &= x'_2 \sin(x'_1) \end{aligned}$$

So

$$S_{11} = (\partial x / \partial \theta) = -r \sin \theta$$

$$S_{12} = (\partial x / \partial r) = \cos \theta$$

$$S_{21} = (\partial y / \partial \theta) = r \cos \theta$$

$$S_{22} = (\partial y / \partial r) = \sin \theta$$

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

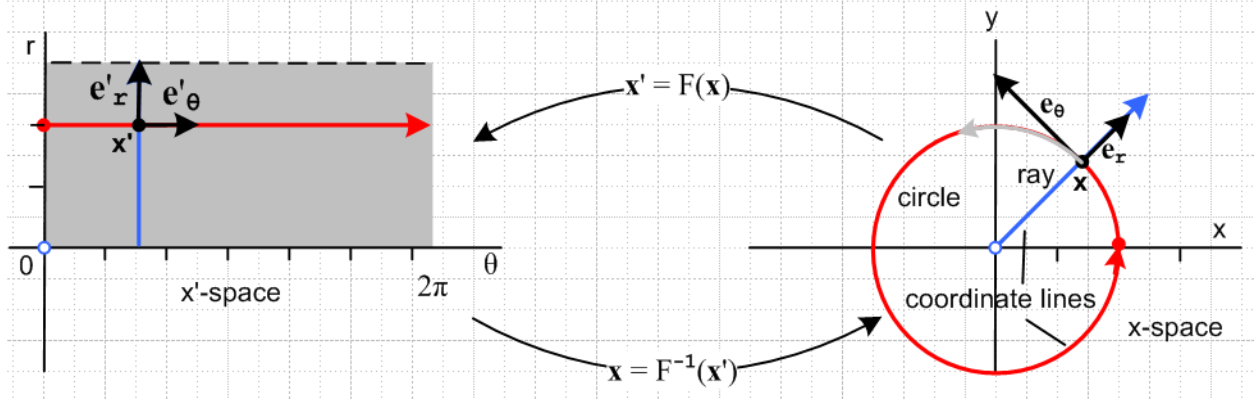
$$S = \begin{pmatrix} -r \sin \theta & \cos \theta \\ r \cos \theta & \sin \theta \end{pmatrix} \quad // \quad \det(S) = -r \quad R = S^{-1} = \begin{pmatrix} -\sin \theta / r & \cos \theta / r \\ \cos \theta & \sin \theta \end{pmatrix}$$

The tangent base vectors $\mathbf{e}_n$ can be read off as the columns of S

$$\mathbf{e}_1 = r(-\sin \theta, \cos \theta) = \mathbf{e}_\theta = r \hat{\mathbf{e}}_\theta \quad // = r \hat{\boldsymbol{\theta}}$$

$$\mathbf{e}_2 = (\cos \theta, \sin \theta) = \mathbf{e}_r = \hat{\mathbf{e}}_r \quad // = \hat{\mathbf{r}}$$

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 $\mathbf{x}'$ -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.



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 $\mathbf{x}$, the $\mathbf{e}_r$ base vector stays the same, but $\mathbf{e}_\theta$ grows longer. If one moves azimuthally from $\mathbf{x}$ to some larger angle $\theta + \Delta\theta$, 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 $\mathbf{v}$ is $v_x^2 + v_y^2$, but that is not the right expression for $|\mathbf{v}|$ in a non-Cartesian space. For example, as will be shown below, $|\mathbf{e}_\theta| = |\mathbf{e}_\theta|$, 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 Appendix C (e) 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 $\mathbf{x}'$ -space on the left map into the two $\mathbf{e}_n$ vectors on the right according to $\mathbf{e}_n \equiv \mathbf{S}'_n$. 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 into the warped gray arrow, while $\mathbf{e}'_r$ maps into $\mathbf{e}_r$.

Example 2: Spherical Coordinates, tangent base vectors

$$\mathbf{x} = (x_1, x_2, x_3) = (x, y, z)$$

$$\mathbf{x}' = (x'_1, x'_2, x'_3) = (r, \theta, \phi)$$

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

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

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

where Maple computes R as S^{-1} and finds as well that

$$\det(S) = r^2 \sin \theta$$

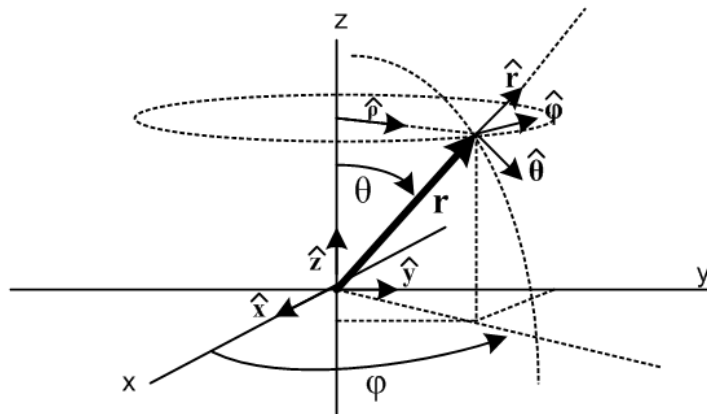
The tangent base vectors are the columns of S , so

$$\begin{aligned}
\mathbf{e}_r &= (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta) & |\mathbf{e}_r| &= 1 & &= h'_r \\
\mathbf{e}_\theta &= r(\cos \theta \cos \varphi, \cos \theta \sin \varphi, -\sin \theta) & |\mathbf{e}_\theta| &= r & &= h'_\theta \\
\mathbf{e}_\varphi &= r \sin \theta (-\sin \varphi, \cos \varphi, 0) & |\mathbf{e}_\varphi| &= r \sin \theta & &= h'_\varphi
\end{aligned}$$

and unit vector versions are then

$$\begin{aligned}
\hat{\mathbf{e}}_r &= (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta) &= \hat{\mathbf{r}} & & \mathbf{e}_r &= \hat{\mathbf{r}} \\
\hat{\mathbf{e}}_\theta &= (\cos \theta \cos \varphi, \cos \theta \sin \varphi, -\sin \theta) &= \hat{\boldsymbol{\theta}} & & \mathbf{e}_\theta &= r \hat{\boldsymbol{\theta}} \\
\hat{\mathbf{e}}_\varphi &= (-\sin \varphi, \cos \varphi, 0) &= \hat{\boldsymbol{\phi}} & & \mathbf{e}_\varphi &= r \sin \theta \hat{\boldsymbol{\phi}}
\end{aligned}$$

The unit vectors can be displayed in this standard picture,



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

(d) 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 $\mathbf{R} \leftrightarrow \mathbf{S}$ in the above discussion. The tangent base vectors for the inverse mapping would then be the columns of matrix $\mathbf{R}$ instead of $\mathbf{S}$. We shall denote these inverse tangent base vectors which exist in $\mathbf{x}'$ -space by the symbol $\mathbf{u}'_n$. Then:

$$\begin{aligned} (\mathbf{e}_n)_i &= S_{in} = \partial x_i / \partial x'_n & // \text{ the tangent base vectors as above} \\ \mathbf{S} &= [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \dots \mathbf{e}_N] & // \text{ are the columns of } \mathbf{S} \\ \\ (\mathbf{u}'_n)_i &= R_{in} = \partial x'_i / \partial x_n & // \text{ inverse tangent base vectors} \\ \mathbf{R} &= [\mathbf{u}'_1, \mathbf{u}'_2, \mathbf{u}'_3 \dots \mathbf{u}'_N] & // \text{ are the columns of } \mathbf{R} \end{aligned}$$

By varying only x_n in $\mathbf{x}$ -space holding all the other $x_i = \text{constant}$, one generates the x_n -coordinate lines in $\mathbf{x}'$ -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 below and another in Appendix C.

In section (b) above the vector $\mathbf{e}_n$ transformed as a contravariant vector into an axis-aligned basis vector $\mathbf{e}'_n$ in $\mathbf{x}'$ -space

$$\mathbf{e}'_n = \mathbf{R} \mathbf{e}_n \quad (\mathbf{e}'_n)_i = R_{ij} (\mathbf{e}_n)_j \quad (\mathbf{e}_n)_i = S_{in} \quad (\mathbf{e}'_n)_i = \delta_{n,i}$$

The same thing happens here, only in reverse :

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

where now the $\mathbf{u}_n$ are axis-aligned basis vectors in $\mathbf{x}$ -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 Section 6 below.

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

It was shown earlier for polar coordinates that,

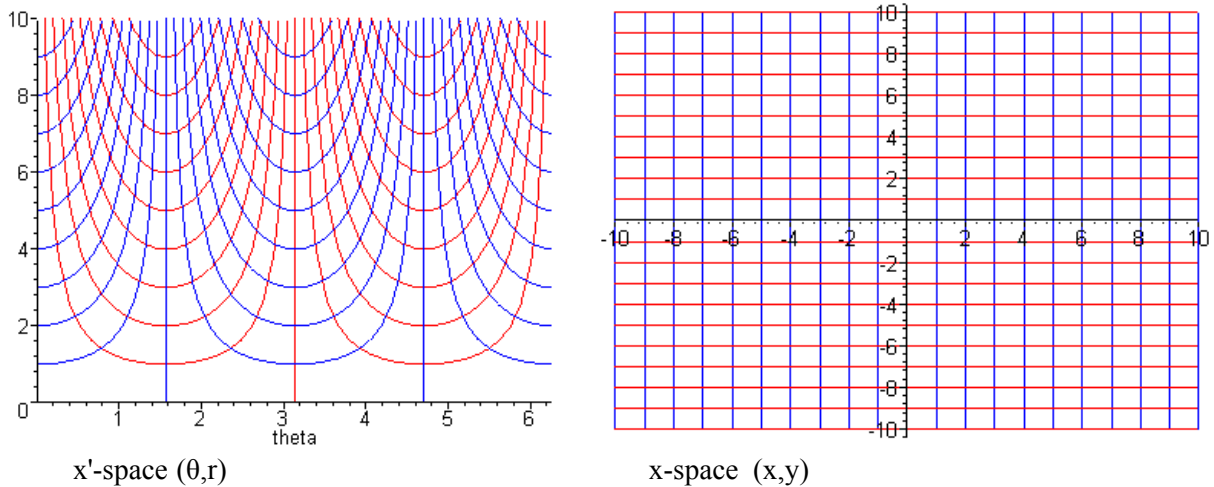
$$\mathbf{R} = \mathbf{S}^{-1} = \begin{pmatrix} -\sin\theta/r & \cos\theta/r \\ \cos(\theta) & \sin\theta \end{pmatrix}$$

so the inverse tangent base vectors are given by the columns of $\mathbf{R}$,

$$\begin{aligned} \mathbf{u}'_x &= (-\sin\theta/r, \cos\theta) & // \text{ note near } \theta = 0 \text{ that } \mathbf{u}'_x \text{ indicates a large negative slope} \\ \mathbf{u}'_y &= (\cos\theta/r, \sin\theta) & // \text{ note near } \theta = 0 \text{ that } \mathbf{u}'_y \text{ indicates a small positive slope} \end{aligned}$$

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). Looking at the small θ region of the plot on the left below, one sees slopes as just described above.

For the polar coordinates mapping discussed above, 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_1) in x' -space has the form $r = y_1/\sin\theta$, where y_1 denotes some selected y value (a red horizontal line), so plotting $r = y_1/\sin\theta$ in x' -space for various values of y_1 displays a set of inverse x -coordinate lines (red). Similarly $r = x_1/\cos\theta$ gives some y -coordinate lines (blue). Here is a Maple plot:



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* $\mathbb{R}^N$ (a vector space is sometimes called a linear space), and have not "endowed" it with a norm, metric or a scalar product. Quantities like $d\mathbf{x}_1$ 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 Section 1.

For $\mathbb{R}^N$ 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

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

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 $\mathbb{R}^N$ the "natural metric" is defined in terms of the above L^2 norm, so that

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

Now our space is both a normed linear space and a metric space.

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

$$(x, y) \equiv \sum_i x_i y_i \equiv \mathbf{x} \bullet \mathbf{y} \quad // = \sum_{i,j} \delta_{i,j} x_i y_j$$

which of course implies this special case,

$$(x, x) = \mathbf{x} \bullet \mathbf{x} = \sum_i x_i^2 = \|\mathbf{x}\|^2 = |\mathbf{x}|^2$$

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 $\mathbb{R}^N$, our "Cartesian Space".

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

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

To avoid dealing with the square root, one usually writes

$$\begin{aligned} (ds)^2 &\equiv \|d\mathbf{x}\|^2 = \sum_i (dx_i)^2 = (dx_1)^2 + (dx_2)^2 + \dots + (dx_N)^2 \\ &= \sum_i dx_i dx_i = \sum_{i,j} \delta_{i,j} dx_i dx_j \end{aligned}$$

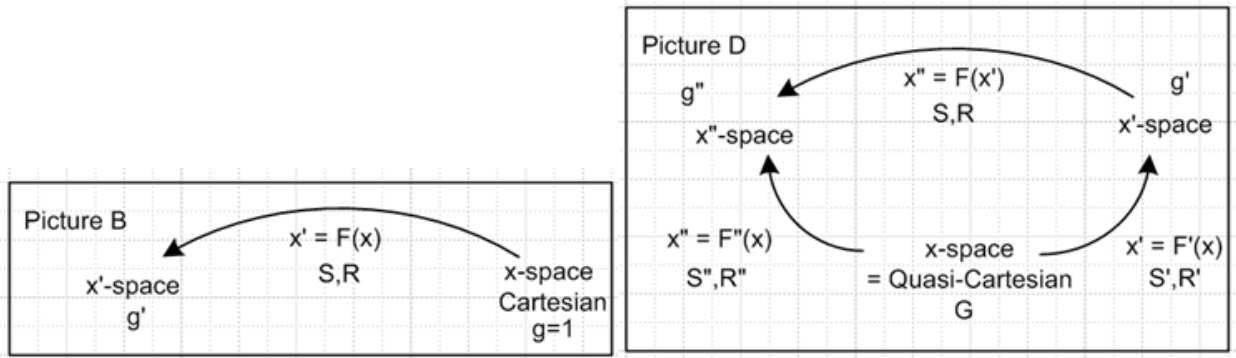
As shown in the next section, one can interpret $\delta_{i,j}$ as the metric tensor in Cartesian Space.

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 $N=\infty$ dimensional Hilbert Spaces used in functional analysis, where basis vectors might be Legendre polynomials $P_n(z)$ on $(-1,1)$, $n = 0,1,2,\dots,\infty$. He has little to say, however, about curvilinear coordinate spaces in this particular book.

5. The Metric Tensor

The metric tensor is the heart of the machine of tensor analysis and we shall have a lot to say about it in this section. Each subsection 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 subsections 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.

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:



The various partial derivatives are determined from their definitions,

$$\begin{aligned} R'_{ik} &\equiv (\partial x'_i / \partial x_k) & R''_{ik} &\equiv (\partial x''_i / \partial x_k) & R_{ik} &\equiv (\partial x''_i / \partial x'_k) \\ S'_{ik} &\equiv (\partial x_i / \partial x'_k) & S''_{ik} &\equiv (\partial x_i / \partial x''_k) & S_{ik} &\equiv (\partial x'_i / \partial x''_k) \end{aligned}$$

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

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

(a) Definition of the metric tensor

The metric or distance between vectors $\mathbf{x}$ and $\mathbf{x}+d\mathbf{x}$ can be specified as done in Section 4 in terms of the norm of differential vector $d\mathbf{x}$,

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

with the caveat that this is not an official norm, see section (i) 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)

$$(ds)^2 = \sum_{i=1}^N \sum_{j=1}^N [\text{metric tensor}]_{ij} dx_i dx_j$$

This is a bit of chicken and egg because one is really defining "distance" and "metric tensor" at the same time. Each selection of a metric tensor defines the meaning of distance ds in the space of interest.

Suppose the length of a small vector $d\mathbf{x}$ in a Quasi-Cartesian x -space is known to be ds . Recall from Section 1 that such a space has a diagonal metric tensor G whose diagonal elements are independently either $+1$ or -1 . How might one express this same ds in terms of the other spaces' coordinates $\mathbf{x}'$ and $\mathbf{x}''$? (see Picture D) Going to x' -space one finds, since $d\mathbf{x} = S'(\mathbf{x}') d\mathbf{x}'$,

$$\begin{aligned} (ds)^2 &= \sum_i G_{ii} dx_i dx_i = \sum_i G_{ii} (\sum_k S'_{ik} dx'_k) (\sum_m S'_{im} dx'_m) \\ &= \sum_k \sum_m \{ \sum_i G_{ii} S'_{ik} S'_{im} \} dx'_k dx'_m \end{aligned}$$

Defining the metric tensor in x' -space to be (comment on the bar below)

$$\bar{g}'_{km} \equiv \sum_i G_{ii} S'_{ik} S'_{im} = \sum_{ij} S'^T_{ki} G_{ij} S'_{jm} \quad \Rightarrow \quad \bar{\mathbf{g}}' = S'^T \mathbf{G} S'$$

one then has, with implied summation on the right,

$$(ds)^2 = \sum_k \sum_m \bar{g}'_{km} dx'_k dx'_m = \bar{\mathbf{g}}'_{km} dx'_k dx'_m$$

For the transformation from x -space to x'' -space in Picture D, a similar result is obtained,

$$\begin{aligned} \bar{g}''_{km} &\equiv \sum_i G_{ii} S''_{ik} S''_{im} \quad \Rightarrow \quad \bar{\mathbf{g}}'' = S''^T \mathbf{G} S'' \\ (ds)^2 &= \bar{\mathbf{g}}''_{km} dx''_k dx''_m \end{aligned}$$

Since $(ds)^2$ is a number which is the same in all three systems (that number is the distance between two points in x -space), the quantity $\bar{\mathbf{g}}'_{km} dx'_k dx'_m$ is a tensorial scalar.

The metric tensor is specific to a space; it is a property of the space; it is part of the space's definition.

We have placed bars over the g 's anticipating what will soon be shown, that these matrices are "covariant" matrices. Then we won't have to go back and fix things up.

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

$$\bar{\mathbf{g}} = \mathbf{G} \quad \bar{\mathbf{g}}' = S'^T \mathbf{G} S' \quad \bar{\mathbf{g}}'' = S''^T \mathbf{G} S''$$

(b) Inverse of the metric tensor

The inverses of the three metric tensors shall be indicated *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:

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

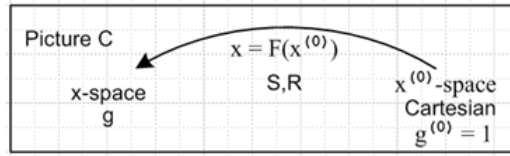
$$g' \equiv \bar{g}'^{-1} = (S'^T G S')^{-1} = S'^{-1} G (S'^T)^{-1} = R' G R'^T$$

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

Here are the collected facts from above:

$g = G$	$g' = R' G R'^T$	$g'' = R'' G R''^T$	$S = R' S''$
$\bar{g} = G$	$\bar{g}' = S'^T G S'$	$\bar{g}'' = S''^T G S''$	$R = R'' S'$
$\bar{g}g = 1$	$\bar{g}'g' = 1$	$\bar{g}''g = 1$	

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



$$\begin{aligned} g &= RGR^T \\ \bar{g} &= S^TGS \\ \bar{g}g &= 1 \end{aligned} \quad (ds)^2 = \bar{g}_{km} dx_k dx_m$$

where now the clutter of primes is gone. If x -space is Cartesian so $G = 1$, then $g = RR^T$ and $\bar{g} = S^TS$. But we continue with **Picture D**.

(c) A metric tensor is symmetric

Any matrix of the form $M = A^TDA$ where D is a diagonal matrix (so $D=D^T$) is symmetric:

$$M^T = (A^TDA)^T = A^TDA = M \quad // \text{ and similarly with } A \rightarrow A^T$$

Since all metric tensors shown above match this form, they are all symmetric: $g_{ab} = g_{ba}$ for any g (with or without an overbar).

Another view is to again consider $(ds)^2 = \bar{g}_{km} dx_k dx_m$. An arbitrary matrix $\bar{g}_{km}$ can always be decomposed into the sum of a symmetric matrix and an antisymmetric matrix. The antisymmetric part makes no contribution to $(ds)^2$ since $A_{ij}a_i a_j = 0$. Therefore, any such antisymmetric part might as well be thrown out, leaving a symmetric $\bar{g}_{km}$.

(d) $\det(g)$ 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 $\bar{g} = S^TS$. Any matrix of either of these forms has positive diagonal elements and positive determinant:

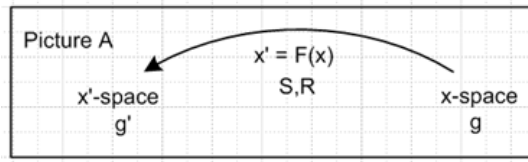
$$(A^T A)_{aa} = \sum_b (A^T)_{ab} A_{ba} = \sum_b (A)_{ba} A_{ba} = \sum_b (A_{ba})^2 \geq 0 \quad // \text{ diagonal elements } \geq 0$$

$$\det(A^T A) = \det(A^T) \det(A) = \det(A) \det(A) = [\det(A)]^2 \geq 0 \quad // \det \geq 0$$

To show these results for the AA^T form, just replace $A \rightarrow A^T$ everywhere. Recall that transformation F maps $\mathbb{R}^N \rightarrow \mathbb{R}^N$ 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 $\bar{g} = S^T GS$ and $G \neq 1$.

(c) Definition of two kinds of rank-2 tensors

We now switch to **Picture A**,



Recall the vector transformation rules from Section 2 (d),

$$\begin{aligned} \mathbf{V}' &= \mathbf{R} \mathbf{V} & \text{contravariant} & & \mathbf{R}_{ik}(\mathbf{x}) &\equiv (\partial x'_i / \partial x_k) & & \mathbf{R} = \mathbf{S}^{-1} \\ \bar{\mathbf{V}}' &= \mathbf{S}^T \bar{\mathbf{V}} & \text{covariant} & & \mathbf{S}_{ik}(\mathbf{x}') &\equiv (\partial x_i / \partial x'_k) &= \mathbf{S}_{ki}^T(\mathbf{x}') \end{aligned}$$

which can be written out in components

$$\begin{aligned} V'_a &= R_{aa'} V_{a'} & \text{contravariant} & & \mathbf{R}_{ik}(\mathbf{x}) &\equiv (\partial x'_i / \partial x_k) & & \mathbf{R} = \mathbf{S}^{-1} \\ \bar{V}'_a &= S^T_{aa'} \bar{V}_{a'} & \text{covariant} & & \mathbf{S}_{ik}(\mathbf{x}') &\equiv (\partial x_i / \partial x'_k) &= \mathbf{S}_{ki}^T(\mathbf{x}') \end{aligned}$$

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:

$$\begin{aligned} M'_{ab} &= R_{aa'} R_{bb'} M_{a'b'} & // \text{ contravariant rank-2 tensor} \\ \bar{M}'_{ab} &= S^T_{aa'} S^T_{bb'} \bar{M}_{a'b'} & // \text{ covariant rank-2 tensor} \end{aligned}$$

and again we put a bar over the covariant objects.

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

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

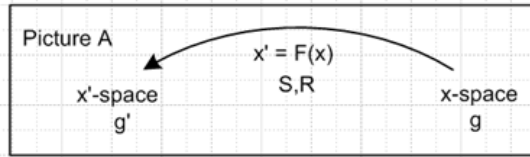
This fact is used many times in the manipulations below.

(f) Proof that the metric tensor and its inverse are both rank-2 tensors

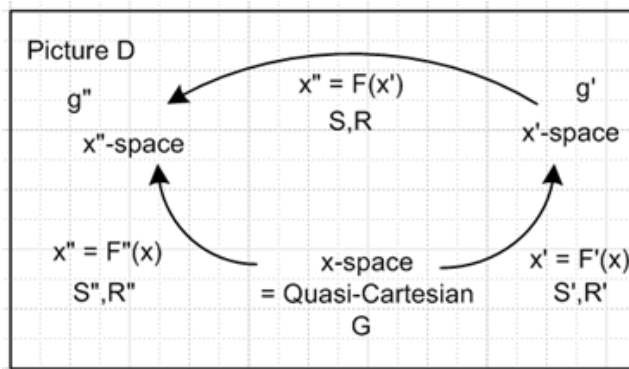
The above rank-2 tensor transformation rules can be written in the following matrix form (something not possible with higher-rank tensors),

$$\begin{aligned} M' &= R M R^T & // \text{contravariant rank-2 tensor} \\ \bar{M}' &= S^T \bar{M} S & // \text{covariant rank-2 tensor} \end{aligned}$$

where recall



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



$$\begin{aligned} M'' &= R M' R^T & // \text{contravariant rank-2 tensor} \\ \bar{M}'' &= S^T \bar{M}' S & // \text{covariant rank-2 tensor} \end{aligned}$$

Consider then this sequence of steps:

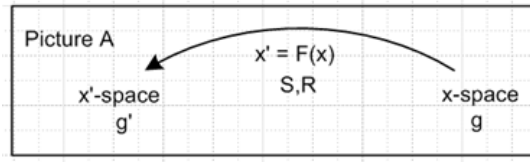
$$\begin{aligned} 1 * G * 1 &= 1 * G * 1 \\ (S''R'') G (S''R'')^T &= (S'R') G (S'R')^T & // S''R'' = 1 \\ S'' (R''G R''^T) S''^T &= S'(R'G R'^T) S'^T & // \text{regroup} \\ S'' g'' S''^T &= S' g' S'^T & // \text{since } g'' = R''G R''^T \text{ and } g' = R'G R'^T \\ g'' S''^T &= R'' S' g' S'^T & // \text{left multiply by } S''^{-1} = R'' \\ g'' &= R'' S' g' S'^T R''^T & // \text{right multiply by } S''^T, -1 = R''^T \\ g'' &= (R'' S') g' (S'^T R''^T) & // \text{regroup} \\ g'' &= (R'' S') g' (R'' S')^T & // (AB)^T = B^T A^T \\ g'' &= R g' R^T & // \text{expressions in section (b) above for R and S} \end{aligned}$$

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,

$$\begin{aligned} g'' &= R g' R^T \\ g''^{-1} &= (R g' R^T)^{-1} \\ g''^{-1} &= S^T g'^{-1} S & // R^{T,-1} = S^T \text{ etc} \\ \overline{g}'' &= S^T \overline{g}' S & // \overline{g}' = g'^{-1} \end{aligned}$$

and this last result shows that $\overline{g}'$ is a *covariant* rank-2 tensor with respect to the transformation F taking x' -space to x'' -space. 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,



$$\begin{aligned} g' &= R g R^T & g'_{ab} &= R_{aa'} R_{bb'} g_{a'b'} & // g \text{ is a contravariant rank-2 tensor} \\ \overline{g}' &= S^T \overline{g} S & \overline{g}'_{ab} &= S^T_{aa'} S^T_{bb'} \overline{g}_{a'b'} & // \overline{g} \text{ is a covariant rank-2 tensor} \end{aligned}$$

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

$$\begin{aligned} g &= S g' S^T & g_{ab} &= S_{aa'} S_{bb'} g'_{a'b'} \\ \overline{g} &= R^T \overline{g}' R & \overline{g}_{ab} &= R^T_{aa'} R^T_{bb'} \overline{g}'_{a'b'} \end{aligned}$$

Two more useful variations of the above are

$$\begin{aligned} Rg &= g' S^T & R_{ab} g_{bc} &= g'_{ab} S_{cb} \\ \overline{g} S &= R^T \overline{g}' & \overline{g}_{ab} S_{bc} &= R_{ba} \overline{g}'_{bc} \end{aligned}$$

(g) Metric tensor converts vector types

We continue in **Picture A**. Suppose V is a contravariant vector so $V' = RV$. Construct a new vector W with the following properties (see Section 7 (u) concerning "covariant equations")

$$\begin{aligned} W &= \overline{g} V & \text{x-space} \\ W' &= \overline{g}' V' & \text{x'-space} \end{aligned}$$

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

$$W' = \overline{g}' V' = (S^T \overline{g} S) (RV) = S^T \overline{g} (SR) V = S^T \overline{g} V = S^T W$$

Therefore this new vector W is a *covariant* vector under F , so it should have an overbar,

$$\bar{\mathbf{W}} \equiv \bar{g} \mathbf{V}$$

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

$$\begin{aligned} \bar{\mathbf{V}} &= \bar{g} \mathbf{V} & \mathbf{V} &= g \bar{\mathbf{V}} \\ \bar{V}_i &= \bar{g}_{ij} V_j & V_i &= g_{ij} \bar{V}_j \end{aligned}$$

(h) Vectors in Cartesian space

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

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 also covariant and vice versa. In other words, if $g = 1$, every vector is the same as its covariant partner vector. The transformation rules in this case are

$$\begin{aligned} \mathbf{V}' &= \mathbf{R} \mathbf{V} \\ \bar{\mathbf{V}}' &= \mathbf{S}^T \bar{\mathbf{V}} = \mathbf{S}^T \mathbf{V} \end{aligned}$$

Although the vectors $\mathbf{V}$ and $\bar{\mathbf{V}}$ are the same, eliminating $\mathbf{V}$ shows that $\bar{\mathbf{V}}'$ and $\mathbf{V}'$ are *not* the same. One finds that $\bar{\mathbf{V}}' = (\mathbf{S}^T \mathbf{S}) \mathbf{V}' = \bar{g}' \mathbf{V}'$, so $\bar{\mathbf{V}}' = \bar{g}' \mathbf{V}' \neq \mathbf{V}'$.

(i) Metric tensor: covariant scalar product and norm

For a Cartesian space, Section 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 $\mathbf{A} \bullet \mathbf{B}$ as defined below, as well as the notation $|\mathbf{A}|^2 \equiv \mathbf{A} \bullet \mathbf{A}$ where $|\mathbf{A}|$ will be called the length, magnitude or norm of $\mathbf{A}$, 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 $\mathbf{A} \bullet \mathbf{A}$ as "the norm" of $\mathbf{A}$ (e.g., Messiah bottom p 878 discussing special relativity), which is our $|\mathbf{A}|^2$. For a general 4-vector $\mathbf{A}$ in special or

general relativity, most authors just write $\mathbf{A} \cdot \mathbf{A}$ ($A_\mu A^\mu$ in standard notation), note that the quantity is invariant under transformations, but don't give it a name.

Whereas we use the bold $\bullet$ 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 $A \cdot B$ or $A \cdot B$. Typical usage then in standard notation would be $\mathbf{p} \cdot \mathbf{p} = p^\mu p_\mu = p_0^2 - \mathbf{p} \cdot \mathbf{p}$ (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 (u)) as:

$$\mathbf{A} \bullet \mathbf{B} \equiv \bar{g}_{ab} A_a B_b = \bar{g}_{ab} B_b A_a = \bar{g}_{ba} B_a A_b = \bar{g}_{ab} B_a A_b = \mathbf{B} \bullet \mathbf{A}$$

This covariant scalar (or dot) product is more interesting and useful than the object $A_a B_a$ because the covariant scalar product of two contravariant vectors is a tensorial scalar, as we now show (Picture A)

$$\begin{aligned} \mathbf{A}' \bullet \mathbf{B}' &= \bar{g}'_{ab} A'_a B'_b = \bar{g}'_{ab} (R_{aa'} A_a) (R_{bb'} B_b) = \bar{g}'_{ab} R_{aa'} R_{bb'} A_a B_b \\ &= [(R^T)_{a'}{}_a \bar{g}'_{ab} R_{bb'}] A_a B_b = [R^T \bar{g}' R]_{a'b'} A_a B_b = \bar{g}_{a'b'} A_a B_b \\ &= \bar{g}_{ab} A_a B_b = \mathbf{A} \bullet \mathbf{B} \end{aligned}$$

Recall that for any contravariant vector $\mathbf{B}$, there is a partner covariant vector $\bar{\mathbf{B}}_a = \bar{g}_{ab} B_b$. Using this partner $\bar{\mathbf{B}}$ one can restate the above covariant scalar product as

$$\mathbf{A} \bullet \mathbf{B} = \bar{g}_{ab} A_a B_b = A_a \bar{B}_a$$

or, taking instead $\bar{\mathbf{A}}_b = \bar{g}_{ba} A_a$,

$$\mathbf{A} \bullet \mathbf{B} = \bar{A}_b B_b = \bar{A}_a B_a$$

And finally, if in $\mathbf{A} \bullet \mathbf{B} = A_a \bar{B}_a$ we write $A_a = g_{ab} \bar{A}_b$, we get

$$\mathbf{A} \bullet \mathbf{B} = A_a \bar{B}_a = g_{ab} \bar{A}_b \bar{B}_a = g_{ab} \bar{A}_a \bar{B}_b$$

where the scalar product is now expressed in terms of the covariant partner vectors $\bar{\mathbf{A}}$ and $\bar{\mathbf{B}}$. To summarize, there are four different ways to write this covariant scalar product :

$$\mathbf{A} \bullet \mathbf{B} = \bar{g}_{ab} A_a B_b = A_a \bar{B}_a = \bar{A}_a B_a = g_{ab} \bar{A}_a \bar{B}_b = \mathbf{B} \bullet \mathbf{A}$$

Using the appropriate expressions on the above line, one may conclude that the covariant dot product of any two tensorial vectors is a tensorial scalar.

In the special case that $\mathbf{A} = \mathbf{B}$, we use the shorthand notation (with caveat as noted above)

$$|\mathbf{A}|^2 \equiv \mathbf{A} \bullet \mathbf{A}$$

Going back to the a result of section (a), one sees the $(ds)^2$ distance squared in a new light,

$$(ds)^2 = \bar{g}'_{km} dx'_k dx'_m = dx' \bullet dx' = dx \bullet dx = \text{a scalar with respect to } F$$

so ds is sometimes called "the invariant distance". In special relativity, using the Bjorken and Drell notation noted above where $g'_{\mu\nu} = \text{diag}(1, -1, -1, -1)$ and $c=1$, one writes (standard notation)

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

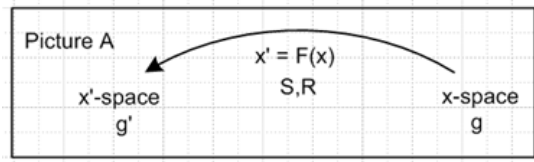
and $d\tau$ is called "the proper time", a particular case of the invariant distance ds . Notice that $(d\tau)^2 < 0$ for a spacelike vector dx^μ , meaning one that lies outside the future and past lightcones ($|d\mathbf{x}| > |dt|$). We now restore $\bullet$ to our covariant definition.

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

$$|\mathbf{e}'_n|^2 = \mathbf{e}'_n \bullet \mathbf{e}'_n = \mathbf{e}_n \bullet \mathbf{e}_n = |\mathbf{e}_n|^2 \quad \Rightarrow \quad |\mathbf{e}'_n| = |\mathbf{e}_n|$$

(j) Metric tensor and tangent base vectors

The context of Picture A continues,



Recall this fact from Section 3,

$$S = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \dots \mathbf{e}_N]$$

where the columns of S are the tangent base vectors. It follows that (see end of section (h))

$$\bar{g}' = S^T \bar{g} S = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \dots \mathbf{e}_N]^T \bar{g} [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \dots \mathbf{e}_N]$$

so

$$\bar{g}' = \begin{matrix} \mathbf{e}_1 \bullet \mathbf{e}_1 & \mathbf{e}_1 \bullet \mathbf{e}_2 & \mathbf{e}_1 \bullet \mathbf{e}_3 & \dots & \mathbf{e}_1 \bullet \mathbf{e}_N \\ \mathbf{e}_2 \bullet \mathbf{e}_1 & \mathbf{e}_2 \bullet \mathbf{e}_2 & \mathbf{e}_2 \bullet \mathbf{e}_3 & \dots & \mathbf{e}_2 \bullet \mathbf{e}_N \\ \mathbf{e}_3 \bullet \mathbf{e}_1 & \mathbf{e}_3 \bullet \mathbf{e}_2 & \mathbf{e}_3 \bullet \mathbf{e}_3 & \dots & \mathbf{e}_3 \bullet \mathbf{e}_N \\ \dots & \dots & \dots & \dots & \dots \\ \mathbf{e}_N \bullet \mathbf{e}_1 & \mathbf{e}_N \bullet \mathbf{e}_2 & \mathbf{e}_N \bullet \mathbf{e}_3 & \dots & \mathbf{e}_N \bullet \mathbf{e}_N \end{matrix}$$

since,

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

using the covariant scalar product defined in the previous section. Taking the n,m component of the above matrix equation, one gets

$$\bar{g}'_{mn} = \mathbf{e}_m \bullet \mathbf{e}_n$$

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 is

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

Therefore, the tangent base vectors will only be mutually orthogonal when the metric tensor $\bar{\mathbf{g}}'$ of x'-space is a *diagonal matrix*. We refer to the coordinates of an x'-space having a diagonal metric tensor as comprising an **orthogonal coordinate system**. At any point $\mathbf{x}$ 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 particular, $\bar{g}'_{nn} = \mathbf{e}_n \bullet \mathbf{e}_n$ lets us write the length of a tangent base vector in terms of the corresponding diagonal element of $\bar{\mathbf{g}}'$,

$$|\mathbf{e}_n|^2 = \mathbf{e}_n \bullet \mathbf{e}_n = \bar{g}'_{nn} \quad \Rightarrow \quad |\mathbf{e}_n| = \sqrt{\bar{g}'_{nn}} \quad \Rightarrow \quad \hat{\mathbf{e}}_n = \mathbf{e}_n / \sqrt{\bar{g}'_{nn}}$$

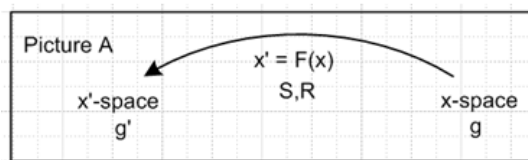
The quantities $|\mathbf{e}_n| = \sqrt{\bar{g}'_{nn}}$ are called **scale factors** and are sometimes written h'_n or Q'_n

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

Section (d) 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.

(k) The Jacobian J

The context of Picture A continues,



First of all, note that since $RS = 1$,

$$\det(S) = 1/\det(R)$$

The Jacobian $J(\mathbf{x}')$ is defined as follows,

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

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 $S_{ij}(\mathbf{x}') = \partial x_i / \partial x'_j$, 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

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

as if $\mathbf{x}$ and $\mathbf{x}'$ were independent variables. In our presentation, $\mathbf{x}' = F(\mathbf{x})$ is not an independent variable but is determined by $F(\mathbf{x})$. 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 $J(\mathbf{x}') = \mathcal{J}(\mathbf{x} = F^{-1}(\mathbf{x}'), \mathbf{x}') = \mathcal{J}(\mathbf{x}(\mathbf{x}'), \mathbf{x}')$.

Note 3: 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 $N \times N$ matrix S may be written ($\epsilon_{abc\dots x}$ is the permutation tensor, Section 7 (h)),

$$\det(S) = \epsilon_{abc\dots x} S_{a1} S_{b2} \dots S_{xN}$$

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

$$\det(S) = \epsilon_{abc\dots x} (\mathbf{e}_1)_a (\mathbf{e}_2)_b \dots (\mathbf{e}_N)_x$$

so J is related to the tangent base vectors by

$$J = \epsilon_{abc\dots x} (\mathbf{e}_1)_a (\mathbf{e}_2)_b \dots (\mathbf{e}_N)_x$$

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

$$\begin{aligned} \det(\bar{\mathbf{g}}') &= \det(S^T \bar{\mathbf{g}} S) = \det(S^T) \det(\bar{\mathbf{g}}) \det(S) = \det(S) \det(S) \det(\bar{\mathbf{g}}) = J^2 \det(\bar{\mathbf{g}}) \\ \det(\mathbf{g}') &= \det(R \mathbf{g} R^T) = \det(R) \det(\mathbf{g}) \det(R^T) = \det(R) \det(R) \det(\mathbf{g}) = J^{-2} \det(\mathbf{g}) \end{aligned}$$

or

$$\begin{aligned} \det(\bar{\mathbf{g}}') &= J^2 \det(\bar{\mathbf{g}}) & \Rightarrow & & J^2 &= \det(\bar{\mathbf{g}}') / \det(\bar{\mathbf{g}}) = [\det(S)]^2 \\ \det(\mathbf{g}') &= J^{-2} \det(\mathbf{g}) \end{aligned}$$

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

$$\begin{aligned} g(\mathbf{x}) &\equiv \det(\overline{g}(\mathbf{x})) = 1/\det(g(\mathbf{x})) && // \text{ in } \mathbf{x}\text{-space} \\ g'(\mathbf{x}') &\equiv \det(\overline{g}'(\mathbf{x}')) = 1/\det(g'(\mathbf{x}')) && // \text{ in } \mathbf{x}'\text{-space} \end{aligned}$$

So that

$$J^2(\mathbf{x}') = \det(\overline{g}'(\mathbf{x}')) / \det(\overline{g}(\mathbf{x})) = g'(\mathbf{x}') / g(\mathbf{x})$$

Normally the argument dependence is suppressed and one then writes

$$J^2 = \det(\overline{g}') / \det(\overline{g}) = g'/g$$

As explained in Appendix D (a), the equation $g' = J^2 g$ says that g , instead of being a tensorial scalar, is a scalar density of weight -2. One must be a little careful to distinguish the scalars g and g' from the tensors g_{ij} and g'_{ij} expressed in matrix notation as g and g' .

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

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

Since $g \overline{g} = 1$, one has $\det(g)\det(\overline{g}) = 1$ so that $\text{sign}[\det(\overline{g})] = \text{sign}[\det(g)]$.

Since $\det(\overline{g}') / \det(\overline{g}) = [\det(S)]^2$, one has $\text{sign}[\det(\overline{g}')] = [\det(\overline{g})]$. Therefore:

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

Since transformation F is assumed invertible in its domain and range, one cannot have $\det(S) = 0$ anywhere except perhaps on a boundary. Since $\det(\overline{g}') = [\det(S)]^2 \det(\overline{g})$, if we assume $\det(\overline{g})$ vanishes nowhere in the $\mathbf{x}$ -space domain of F , then $\det(\overline{g}') \neq 0$ 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

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

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

$$|J| = \sqrt{g'} = |\det(S)| \quad // \text{ curvilinear}$$

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

$$|J| = \sqrt{-g'} = |\det(S)| \quad // \text{ relativity}$$

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

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

$$g \equiv \det(\bar{g}) \quad g' \equiv \det(\bar{g}')$$

$$g' = J^2 g \quad \Rightarrow \quad g \text{ is a tensor density of weight } -2$$

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

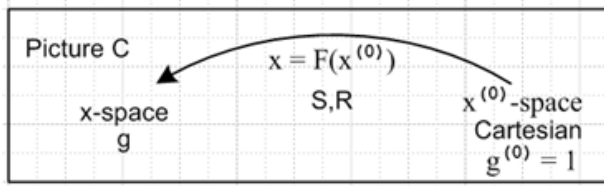
$$|J| = \sqrt{sg'} / \sqrt{sg} = |\det(S)| = \sqrt{g'/g}$$

Note: 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.

Carl Gustav Jacob Jacobi (1804 –1851). 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 $\text{sn}(x;k)$, 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.

(l) Some relations between g , R and S in Picture C

In Picture C,



the statement of the rank-2 tensor transformation of g and $\bar{g}$ becomes

$$g = R R^T$$

$$\bar{g} = S^T S$$

which can be written in a variety of ways,

$$\begin{aligned} R^T &= (SR)R^T = S(RR^T) = Sg & \Rightarrow & R = g S^T & \Rightarrow & 1 = Sg S^T \\ S^T &= S^T(R^T S^T) = (S^T S)R = \bar{g} R & \Rightarrow & S = R^T \bar{g} & \Rightarrow & 1 = R^T \bar{g} R \end{aligned}$$

In summary:

$$\begin{aligned} g &= R R^T & R^T &= S^T g & R &= g^{-1} S^T & 1 &= S^T g S \\ \bar{g} &= S^T S & S^T &= \bar{g}^{-1} R & S &= R^T \bar{g}^{-1} & 1 &= R^T \bar{g} R \end{aligned}$$

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

$$\begin{aligned} \bar{g}_{nn} &= \sum_i S_{ni}^T S_{in} = \sum_i (S_{in})^2 = \sum_i (\partial x_i / \partial x'_n)^2 \\ g_{nn} &= \sum_i R_{ni} R_{in}^T = \sum_i (R_{ni})^2 = \sum_i (\partial x'_n / \partial x_i)^2 \end{aligned}$$

If the x'_i are orthogonal coordinates, then $\bar{g}_{nm} = h_n'^{-2} \delta_{nm}$ and $g_{nm} = h_n'^2 \delta_{nm}$ where the h'_n are the scale factors mentioned above in section (j). These scale factors may then be expressed as (M&F p 23 1.3.4)

$$h_n'^2 = \bar{g}_{nn} = \sum_i (\partial x_i / \partial x'_n)^2 \quad h_n'^{-2} = g_{nn} = \sum_i (\partial x'_n / \partial x_i)^2$$

Example 1: Polar coordinates: metric tensor and Jacobian

Picture C continues (so now $\theta = x_1$ and $r = x_2$) and the metric tensor for polar coordinates will be computed in two ways. On the last visit to this example (end of Section 3), it was shown that

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

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

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

Another way is this:

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

Notice that this metric tensor is in fact symmetric, and that one of its elements is a function of the coordinates. The length² of a small vector dx can be written

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

The Jacobian is given by

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

Example 2: Spherical coordinates: metric tensor and Jacobian

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

In our last visit to this example (end of Section 3) it was found that

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

The metric tensor is then given by Maple as

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

so that

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

The Jacobian is found by Maple to be,

$$J(r, \theta, \varphi) = \det(S) = r^2\sin\theta$$

Differential distance is then

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

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

(m) Special Relativity and its Metric Tensor: vectors and spinors

In this section the Standard Notation introduced below in Section 7 is used. In that notation $R_{i,j}$ is written R^i_j , contravariant vectors V_i are written V^i , and covariant vectors $\bar{V}_j$ 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 $G = \text{diag}(1, -1, -1, -1)$ and the ordering of 4-vectors as $x^\mu = (t, x, y, z)$ where $c=1$ (speed of light) and $\mu = 0, 1, 2, 3$ (Bjorken and Drell p 281). General relativity people often use $G = \text{diag}(-1, 1, 1, 1) \equiv \eta$ instead (Weinberg p 26). Still other authors use $G = 1$ and $x^\mu = (it, x, y, z)$ 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^μ_ν

$$x'^{\mu} = F^{\mu}_{\nu} x^{\nu} = R^{\mu}_{\nu} x^{\nu} \quad \Rightarrow \quad x^{\nu} \text{ is a contravariant vector}$$

and, as is always the case with any linear transformation, invariant length is preserved $x'.x' = x.x = \text{scalar}$. Special relativity 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 $R G R^T = G$. This condition restricts the (proper) Lorentz transformations to be rotations, boosts (velocity transformations), or any combination of the two. In particular,

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

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

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

and second, a rotation transformation about the x axis,

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

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,

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

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 ($g = G = \text{diag}(1, -1, -1, -1)$)

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

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 $F^\mu_\nu(\mathbf{r}, \mathbf{b}) = [\exp \{-i(\mathbf{r} \cdot \mathbf{J} + \mathbf{b} \cdot \mathbf{K})\}]^\mu_\nu$ where the 6 numbers $\mathbf{r}$ and $\mathbf{b}$ 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 2×2 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 $1/2 \otimes 1/2$ there are $2 \times 2 = 4$ 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.

One can also consider two other representations of the Lorentz group which are pretty obvious: $1/2 \otimes 0$ and $0 \otimes 1/2$. These are 2×2 matrix representations and they are *different* 2×2 representations. For each representation one can construct a whole tensor analysis edifice 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 $0 \otimes 1/2$ 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

$$\begin{array}{lll} V'^a = R^a_b V^b & V'^{\dot{a}} = R^{\dot{a}}_{\dot{b}} V^{\dot{b}} & \text{contravariant 2-vectors} \\ V'_a = R_a^b V_b & V'_{\dot{a}} = R_{\dot{a}}^{\dot{b}} V_{\dot{b}} & \text{covariant 2-vectors} \end{array}$$

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

$$X'^{a\dot{b}} = R^a_a R^{\dot{b}}_{\dot{b}} X^{a\dot{b}}, \quad X'^{a'b'} = R^a_a R^{b'}_{b'} X^{a'b'}$$

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

$$\text{evalm}(Bx \&* G \&* \text{transpose}(Bx)) \quad \text{means} \quad B_x G B_x^T$$

and Maple is just verifying that $B_{\mathbf{x}} G B_{\mathbf{x}}^T = G$ and similarly $R_{\mathbf{x}} G R_{\mathbf{x}}^T = G$:

```
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));
```

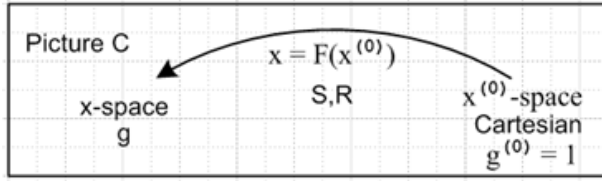
$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -\cos(r)^2 - \sin(r)^2 & 0 \\ 0 & 0 & 0 & -\cos(r)^2 - \sin(r)^2 \end{bmatrix}$$

```
simplify(%);
```

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

(n) General Relativity and its Metric Tensor

In general relativity a Picture of interest is Picture C



but the $x^{(0)}$ space is replaced by a Quasi-Cartesian space with coordinates ξ^i with the metric tensor of special relativity.

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 $G = \text{diag}(-1, 1, 1, 1) \equiv \eta$.

The x^i are the coordinates of *some other* coordinate system. There is some transformation $\mathbf{x} = F(\xi)$ which defines the relationship between these two systems. The covariant metric tensor in x -space is written $\bar{g} = S^T G S = S^T \eta S$. Using the Standard Notation introduced in Section 7 below, this is usually written as

$$\bar{g} = S^T \eta S \quad // \text{ result from Section 5 (b)}$$

$$g_{dn} = S^T \eta_{dn} S \quad // \text{ where } (\eta_{dn})_{ab} = \eta_{ab}$$

$$g_{\mu\nu} = (S^T)^\mu_\alpha \eta_{\alpha\beta} S^\beta_\nu = S^\alpha_\mu \eta_{\alpha\beta} S^\beta_\nu \quad // \text{ Standard Notation as in Section 7}$$

$$g_{\mu\nu} = (\partial \xi^\alpha / \partial x^\mu) \eta_{\alpha\beta} (\partial \xi^\beta / \partial x^\nu)$$

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

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 $m\mathbf{a} = \mathbf{F}$ appears this way in general relativity

$$m (\partial^2 x^\mu / \partial \tau^2) = F^\mu - m \Gamma^\mu_{\nu\lambda} (\partial x^\nu / \partial \tau) (\partial x^\lambda / \partial \tau) \quad // \text{ p 123 (5.1.11 following)}$$

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.

$$\Gamma^\mu_{\nu\lambda} = \frac{1}{2} g^{\mu\sigma} (\partial_\nu g_{\lambda\sigma} + \partial_\lambda g_{\nu\sigma} - \partial_\sigma g_{\nu\lambda}) \quad // \text{ p 75 (3.3.7)}$$

These 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 discussed in Section 7 (v).

(o) Continuum Mechanics and its Metric Tensor

One can describe (Lai) the forward "flow" of a continuous blob of matter by $\mathbf{x} = \mathbf{x}(\mathbf{X}, t)$ where $\mathbf{X} = \mathbf{x}(\mathbf{X}, t_0)$. The "particle" of matter that starts at location $\mathbf{X}$ at time t_0 ends up at $\mathbf{x}$ at time t . Two points separated by $d\mathbf{X}$ at t_0 end up separated by some $d\mathbf{x}$ at t . The relation between them is given by $d\mathbf{x} = \mathbf{F} d\mathbf{X}$ where $\mathbf{F}$ is called the deformation gradient (a rank-2 tensor). $\mathbf{F}$ describes how a particle starting say with a cubic shape at t_0 gets deformed into some parallelepiped shape at t .

The finite flow $\mathbf{x} = \mathbf{x}(\mathbf{X}, t)$ from time t_0 to time t can be thought of as a generic (generally non-linear) transformation of the form $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ as in Section 1 above. The two times are regarded as fixed parameters and the starting coordinates $\mathbf{X}$ are assumed to be Cartesian. In order to put this flow into the notation of this document, let $\mathbf{X} \rightarrow \mathbf{x}$ and $\mathbf{x} \rightarrow \mathbf{x}'$ so that

<u>cont mechanics</u>		<u>this document</u>	
$\mathbf{x}, \mathbf{X}$	$\leftrightarrow$	$\mathbf{x}', \mathbf{x}$	
$\mathbf{x} = \mathbf{x}(\mathbf{X}, t)$	$\leftrightarrow$	$\mathbf{x}' = \mathbf{F}(\mathbf{x})$	// Lai (3.1.4)
$d\mathbf{x} = \mathbf{F} d\mathbf{X}$	$\leftrightarrow$	$d\mathbf{x}' = \mathbf{R} d\mathbf{x}$	// as in Section 2 // Lai (3.7.6), (3.18.3)
$\mathbf{X} = \text{Cartesian}$	$\leftrightarrow$	$g = 1$	
$\mathbf{B} = \mathbf{F}\mathbf{F}^T$	$\leftrightarrow$	$g' = \mathbf{R}\mathbf{R}^T$	// as in Section 5 (ℓ) // Lai (3.25.2)

Thus, the deformation gradient $\mathbf{F}$ is just the $\mathbf{R}$ matrix of the forward transformation $\mathbf{x} = \mathbf{x}(\mathbf{X}, t)$. The matrix $\mathbf{B}$ is called the left Cauchy-Green deformation tensor and is seen to be the metric tensor g' in $\mathbf{x}$ -space. A distance element in this space would then be $(ds)^2 = B_{ij}dx^i dx^j$.

One can then consider the time-reversed flow $\mathbf{X} = \mathbf{X}(\mathbf{x}, t)$ as the inverse of the above transformation by letting $\mathbf{x} \rightarrow \mathbf{x}$ and $\mathbf{X} \rightarrow \mathbf{x}'$, now regarding the $\mathbf{x}$ coordinates as Cartesian :

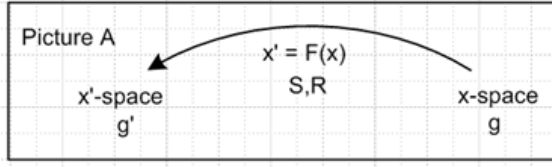
<u>cont mechanics</u>		<u>this document</u>	
$\mathbf{X}, \mathbf{x}$	$\leftrightarrow$	$\mathbf{x}', \mathbf{x}$	
$\mathbf{X} = \mathbf{X}(\mathbf{x}, t)$	$\leftrightarrow$	$\mathbf{x}' = \mathbf{F}(\mathbf{x})$	
$d\mathbf{x} = \mathbf{F} d\mathbf{X}$	$\leftrightarrow$	$d\mathbf{x} = \mathbf{S} d\mathbf{x}'$	// as in Section 2
$\mathbf{x} = \text{Cartesian}$	$\leftrightarrow$	$g = 1$	
$\mathbf{C} = \mathbf{F}^T\mathbf{F}$	$\leftrightarrow$	$g' = \mathbf{S}^T\mathbf{S}$	// as in Section 5 (ℓ) // Lai (3.23.2)

For the transformation in this reversed direction, the same deformation tensor $\mathbf{F}$ is now the $\mathbf{S}$ matrix, and the matrix $\mathbf{C} = \mathbf{F}^T\mathbf{F}$ is then the metric tensor g' in $\mathbf{X}$ space. $\mathbf{C}$ is called the right Cauchy-Green deformation tensor and $(dS)^2 = C_{ij}dX^i dX^j$.

In (reverse) dyadic notation the deformation gradient is written $\mathbf{F} = (\nabla \mathbf{x})$ where ∇ means $\nabla^{(\mathbf{x})}$ so that $d\mathbf{x} = (\nabla \mathbf{x}) d\mathbf{X}$. The $(\nabla \mathbf{x})$ notation is explained in Appendix E, and in Appendix F the object $(\nabla \mathbf{v})$ for an arbitrary vector field $\mathbf{v}(\mathbf{x})$ is expressed in general curvilinear coordinates.

6. Reciprocal Base Vectors E_n and Inverse Reciprocal Base Vectors U'_n

This entire Section uses the Picture A context,



(a) Definition of the E_n

Although various definitions are possible, we shall define the reciprocal tangent vectors E_n in the following manner (implied sum on i)

$$E_n \equiv g'_{ni} e_i \quad \Rightarrow \quad e_n = \bar{g}'_{ni} E_i \quad // \text{ since } \bar{g}' = g'^{-1}$$

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

$$\bar{V}_n = \bar{g}'_{ni} V_i$$

In the last equation, the right side is a linear combination of vector *components* V_i , while in the previous equation the right side is a linear combination of *vectors* e_i . In this case, i is a label on e_i , whereas in the other case i is an index on V_i . Labels and indices are different.

Since the tangent base vectors e_i are contravariant vectors in x -space (Section 3 (b)), and since E_n is a linear combination of the e_i , the E_n are also contravariant vectors in x -space. Notice that in the definition $E_n \equiv g'_{ni} e_i$, these two x -space vectors are related by the metric tensor of the *other* space.

One can express the components of E_n in two ways,

$$\begin{aligned} (E_n)_k &\equiv g'_{ni} (e_i)_k = S_{ki} g'_{ni} & // \text{ since } (e_i)_k &\equiv S_{ki}, \text{ Section 3} \\ &= g'_{ni} S_{ki} = R_{ni} g_{ik} & // \text{ since } R_{ab} g_{bc} &= g'_{ab} S_{cb}, \text{ end of Section 5 (f)} \end{aligned}$$

so that

$$(E_n)_i = g'_{na} S_{ia} = g_{ia} R_{na} \quad // \text{ sum on second indices}$$

Applying R to both sides of $E_n \equiv g'_{ni} e_i$ gives E_n transformed into x' -space,

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

so

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

(b) The Dot Products and Reciprocity (Duality)

Three covariant dot products are of great interest. The first is this (Section 5 (f) for last step)

$$\mathbf{e}_n \bullet \mathbf{e}_m = \bar{g}_{ij} (\mathbf{e}_n)_i (\mathbf{e}_m)_j = \bar{g}_{ij} S_{in} S_{jm} = S_{ni}^T \bar{g}_{ij} S_{jm} = (S^T \bar{g} S)_{nm} = \bar{g}'_{nm}$$

The second is

$$\mathbf{E}_n \bullet \mathbf{e}_m = \bar{g}_{ij} (\mathbf{E}_n)_i (\mathbf{e}_m)_j = \bar{g}_{ij} g_{ia} R_{na} S_{jm} = \delta_{j,a} R_{na} S_{jm} = R_{nj} S_{jm} = (RS)_{nm} = \delta_{n,m}$$

and the third is

$$\begin{aligned} \mathbf{E}_n \bullet \mathbf{E}_m &= \bar{g}_{ij} (\mathbf{E}_n)_i (\mathbf{E}_m)_j = \bar{g}_{ij} g_{ia} R_{na} g_{jb} R_{mb} = \delta_{j,a} R_{na} g_{jb} R_{mb} = R_{nj} g_{jb} R_{mb} \\ &= R_{nj} g_{jb} R_{bm}^T = (R g R^T)_{nm} = g'_{nm} \end{aligned}$$

To summarize,

$$\begin{aligned} \mathbf{e}_n \bullet \mathbf{e}_m &= \bar{g}'_{nm} \\ \mathbf{E}_n \bullet \mathbf{e}_m &= \delta_{n,m} \\ \mathbf{E}_n \bullet \mathbf{E}_m &= g'_{nm} \end{aligned}$$

Using the transformed $\mathbf{E}'_n$ defined above, one finds that

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

which is consistent with the fact that this is a covariant dot product of tensorial vectors:

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

The other two dot products above work this same way, so

$$\begin{aligned} \mathbf{e}'_n \bullet \mathbf{e}'_m &= \bar{g}'_{nm} \\ \mathbf{E}'_n \bullet \mathbf{e}'_m &= \delta_{n,m} \\ \mathbf{E}'_n \bullet \mathbf{E}'_m &= g'_{nm} \end{aligned}$$

Notes on Reciprocity (Duality)

1. The reciprocal base vectors $\mathbf{E}_n$ are more usually defined as being those vectors which satisfy the equations $\mathbf{E}_n \bullet \mathbf{e}_m = \delta_{n,m}$ where the $\mathbf{e}_m$ are known. Each $\mathbf{E}_n$ vector has N components so the full set of $\mathbf{E}_n$ vectors has N^2 components. As n and m take all values, $\mathbf{E}_n \bullet \mathbf{e}_m = \delta_{n,m}$ represents N^2 linear equations. A solution exists since S is invertible (the $\mathbf{e}_m$ form a complete set). The solution is unique and in fact gives

our assumed definition above $\mathbf{E}_n \equiv g'_{ni} \mathbf{e}_i$. Here is a fast solution of this Cramer's Rule problem using matrix notation:

$$\begin{aligned} \mathbf{E}_n \bullet \mathbf{e}_m &= \delta_{n,m} \Rightarrow \bar{g}_{ij}(\mathbf{E}_n)_i (\mathbf{e}_m)_j = \delta_{n,m} \Rightarrow (\mathbf{E}_n)_i \bar{g}_{ij} S_{jm} = \delta_{n,m} \quad \text{Let } A_{ni} \equiv (\mathbf{E}_n)_i. \\ \text{Then have } A \bar{g} S &= 1, \Rightarrow A = (\bar{g} S)^{-1} = S^{-1} \bar{g}^{-1} = R \quad g = g' S^T \quad (\text{end Sec 5f}). \text{ Therefore} \\ A &= g' S^T \Rightarrow (\mathbf{E}_n)_i = A_{ni} = g'_{nj} (S^T)_{ji} = g'_{nj} (\mathbf{e}_j)_i \Rightarrow \mathbf{E}_n = g'_{nj} \mathbf{e}_j. \quad \text{QED} \end{aligned}$$

2. In general, if one has $\mathbf{A}_n \bullet \mathbf{a}_m = \delta_{n,m}$, the vectors $\mathbf{A}_n$ are said to be "reciprocal" to the $\mathbf{a}_m$ and vice versa, so the vectors $\mathbf{E}_n$ are reciprocal to the tangent base vectors $\mathbf{e}_n$.

3. Some authors refer to $\mathbf{A}_n \bullet \mathbf{a}_m = \delta_{n,m}$ as a "duality relation" and either set of vectors is "dual to" the other set. The $\mathbf{E}_n$ are referred to as the dual vectors to $\mathbf{e}_n$.

4. If the $\mathbf{a}_m$ are contravariant vectors, then $\mathbf{A}_n$ will also be contravariant and then $\mathbf{A}_n \bullet \mathbf{a}_m$ is a tensorial scalar. Therefore if $\mathbf{A}_n \bullet \mathbf{a}_m = \delta_{n,m}$, then so also $\mathbf{A}'_n \bullet \mathbf{a}'_m = \delta_{n,m}$ in x' -space, where $\mathbf{a}'_m = R \mathbf{a}_m$ and $\mathbf{A}'_n = R \mathbf{A}_n$. For example, $\mathbf{E}'_n \bullet \mathbf{e}'_m = \delta_{n,m}$ in x' -space where $\mathbf{e}'_m = R \mathbf{e}_m$ and $\mathbf{E}'_n = R \mathbf{E}_n$.

5. In section (e) 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 $\mathbf{x} = F^{-1}(\mathbf{x}')$.

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

$$\begin{aligned} \mathbf{V} &= \sum_n k_n \mathbf{A}_n & \text{where } k_m &= \mathbf{V} \bullet \mathbf{a}_m \\ \mathbf{V} &= \sum_n c_n \mathbf{a}_n & \text{where } c_m &= \mathbf{V} \bullet \mathbf{A}_m \end{aligned}$$

so that for example $\mathbf{a}_m \bullet \mathbf{V} = \mathbf{a}_m \bullet [\sum_n k_n \mathbf{A}_n] = \sum_n k_n \mathbf{a}_m \bullet \mathbf{A}_n = \sum_n k_n \delta_{m,n} = k_m$. These expansions are explored in section (f) below for the two dual sets $\mathbf{E}_n, \mathbf{e}_n$ and $\mathbf{U}_n, \mathbf{u}_n$.

(c) Covariant partner for $\mathbf{E}_n$

The covariant partner for $\mathbf{E}_n$ is given by

$$(\bar{\mathbf{E}}_n)_i = \bar{g}_{ij} (\mathbf{E}_n)_j$$

so that

$$(\bar{\mathbf{E}}_n)_i = \bar{g}_{ij} (\mathbf{E}_n)_j = \bar{g}_{ij} g_{ja} R_{na} = \delta_{i,a} R_{na} = R_{ni}$$

Thus, one can regard the covariant vectors $\bar{\mathbf{E}}_n$ as being the rows of matrix R

$$R = \begin{bmatrix} \bar{\mathbf{E}}_1 \\ \bar{\mathbf{E}}_2 \\ \bar{\mathbf{E}}_3 \\ \dots \\ \bar{\mathbf{E}}_N \end{bmatrix} = [\bar{\mathbf{E}}_1, \bar{\mathbf{E}}_2, \bar{\mathbf{E}}_3 \dots \bar{\mathbf{E}}_N]^T$$

which compare to $S = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \dots \mathbf{e}_N]$

(d) Summary of the basic facts:

$$\begin{aligned} (\mathbf{e}_n)_k &= S_{kn} & \mathbf{e}_n \bullet \mathbf{e}_m &= \bar{g}'_{nm} & |\mathbf{e}_n| &= \sqrt{g'_{nn}} = h'_n & S &= [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \dots \mathbf{e}_N] \\ (\bar{\mathbf{E}}_n)_i &= g_{ia} R_{na} & \mathbf{E}_n \bullet \mathbf{E}_m &= g'_{nm} & |\mathbf{E}_n| &= \sqrt{g'_{nn}} & R &= [\bar{\mathbf{E}}_1, \bar{\mathbf{E}}_2, \bar{\mathbf{E}}_3 \dots \bar{\mathbf{E}}_N]^T \\ &= g'_{na} S_{ia} & \mathbf{e}_n \bullet \mathbf{E}_m &= \delta_{n,m} & \mathbf{E}_n &\equiv g'_{ni} \mathbf{e}_i & \mathbf{e}_n &= \bar{g}'_{ni} \mathbf{E}_i \quad (\bar{\mathbf{E}}_n)_i = R_{ni} \end{aligned}$$

$$\mathbf{e}'_n = R \mathbf{e}_n \quad \text{where} \quad (\mathbf{e}'_n)_i = \delta_{n,i} \quad // \text{ from Section 3 (b)}$$

In general, neither set of base vectors -- the tangent $\{\mathbf{e}_n\}$ or the reciprocal $\{\mathbf{E}_n\}$ -- is orthogonal, since the metric tensor g' in general is not diagonal. And in general none of these vectors is a unit vector.

(e) Repeat the above for the inverse transformation: definition of the $\mathbf{U}'_n$

Section 3 (d) introduced the *inverse* tangent base vectors called $\mathbf{u}'_n$. The prime indicates that these vectors exist in x' -space. In analogy with what was done above, one can define the inverse reciprocal base vectors $\mathbf{U}'_n$ according to

$$\mathbf{U}'_n \equiv g_{ni} \mathbf{u}'_i \quad \Rightarrow \quad \mathbf{u}'_n = \bar{g}_{ni} \mathbf{U}'_i \quad // \text{ since } \bar{g} = g^{-1}$$

Everything goes along as in the previous sections, but with these changes:

$$g' \leftrightarrow g \quad R \leftrightarrow S \quad \mathbf{e}_n \rightarrow \mathbf{u}'_n \quad \mathbf{e}'_n \rightarrow \mathbf{u}_n \quad \mathbf{E}_n \rightarrow \mathbf{U}'_n \quad \mathbf{E}'_n \rightarrow \mathbf{U}_n$$

Here are the key results, translated from above,

$$\mathbf{U}'_n \equiv g_{ni} \mathbf{u}'_i$$

$$(S\mathbf{U}'_n) \equiv g_{ni} (S\mathbf{u}'_i) \quad \Rightarrow \quad \mathbf{U}_n = g_{ni} \mathbf{u}_i \quad S = R^{-1}$$

$$(\mathbf{U}'_n)_i = g_{na} R_{ia} = g'_{ia} S_{na} \quad // \text{ sum on second indices}$$

$$(\mathbf{u}_n)_k = S_{ki} (\mathbf{u}'_n)_i = \delta_{n,k}$$

$$(\mathbf{U}_n)_k = S_{ki} (\mathbf{U}'_n)_i = S_{ki} g_{na} R_{ia} = S_{ki} R_{ia} g_{na} = \delta_{k,a} g_{na} = g_{nk}$$

$$(\bar{\mathbf{U}}'_n)_i = \bar{g}'_{ij} (\mathbf{U}'_n)_j$$

$$\begin{aligned}
(\mathbf{u}'_n)_k &= R_{kn} & \mathbf{u}'_n \bullet \mathbf{u}'_m &= \bar{g}_{nm} |\mathbf{u}'_n| = \sqrt{\bar{g}_{nn}} = h_n & R &= [\mathbf{u}'_1, \mathbf{u}'_2, \mathbf{u}'_3 \dots \mathbf{u}'_N] \\
(\mathbf{U}'_n)_i &= g'_{ia} S_{na} & \mathbf{U}'_n \bullet \mathbf{U}'_m &= g_{nm} & |\mathbf{U}'_n| &= \sqrt{g_{nn}} & S &= [\bar{\mathbf{U}}'_1, \bar{\mathbf{U}}'_2, \bar{\mathbf{U}}'_3 \dots \bar{\mathbf{U}}'_N]^T \\
&= g_{na} R_{ia} & \mathbf{u}'_n \bullet \mathbf{U}'_m &= \delta_{n,m} & \mathbf{U}'_n &\equiv g_{ni} \mathbf{u}'_i & \mathbf{u}'_n &= \bar{g}_{ni} \mathbf{U}'_n \quad (\bar{\mathbf{U}}'_n)_i = S_{ni}
\end{aligned}$$

$$\begin{aligned}
\mathbf{u}_n \bullet \mathbf{u}_m &= \bar{g}_{nm} \\
\mathbf{U}_n \bullet \mathbf{u}_m &= \delta_{n,m} \\
\mathbf{U}_n \bullet \mathbf{U}_m &= g_{nm}
\end{aligned}$$

$$\mathbf{u}_n = S \mathbf{u}'_n \quad \text{where} \quad (\mathbf{u}_n)_i = \delta_{n,i} \quad // \text{ from Section 3 (translated)}$$

It is helpful to keep all these eight vector symbol names in mind (and each has a covariant partner)

	<u>x'-space</u>	<u>x-space</u>		
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	$\mathbf{U}'_n$	$\mathbf{E}_n$	$(\mathbf{U}'_n)_i = g'_{ia} S_{na}$ $= g_{na} R_{ia}$	$(\mathbf{E}_n)_i = g_{ia} R_{na}$ $= g'_{na} S_{ia}$

and recall that $\mathbf{A}_n \bullet \mathbf{a}_m = \delta_{n,m}$ for each of the four dual pairs (two primed, two unprimed).

(f) 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 $\mathbf{V} = (V_1, V_2, \dots, V_N)$. 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,

$$\mathbf{V} = V_1 \mathbf{u}_1 + V_2 \mathbf{u}_2 + \dots = \sum_n V_n \mathbf{u}_n \quad \text{where} \quad \mathbf{U}_n \bullet \mathbf{V} = V_n \quad \mathbf{U}_n = g_{ni} \mathbf{u}_i$$

The components V_n are $\mathbf{U}_n \bullet \mathbf{V}$ because $\mathbf{U}_n \bullet \mathbf{u}_m = \delta_{n,m}$. From Section 5 (g), one can write $V_n = g_{nm} \bar{V}_m$ (regarded here as a definition of the $\bar{V}_m$) so one finds that

$$\mathbf{V} = \sum_n V_n \mathbf{u}_n = \sum_n g_{nm} \bar{V}_m \mathbf{u}_n = \sum_n \bar{V}_m g_{mn} \mathbf{u}_n = \sum_n \bar{V}_m \mathbf{U}_m$$

and thus another expansion for $\mathbf{V}$ is this

$$\mathbf{V} = \bar{V}_1 \mathbf{U}_1 + \bar{V}_2 \mathbf{U}_2 + \dots = \sum_n \bar{V}_n \mathbf{U}_n \quad \text{where} \quad \mathbf{u}_n \bullet \mathbf{V} = \bar{V}_n$$

Comments:

1. If $\mathbf{V}$ 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 $\mathbf{V}$ is a contravariant vector, the expansion above $\mathbf{V} = \sum_n V_n \mathbf{u}_n$ displays the contravariant components of $\mathbf{V}$. The second expansion $\mathbf{V} = \sum_n \bar{V}_m \mathbf{U}_m$ is still an expansion for contravariant vector $\mathbf{V}$, but it displays the components of the covariant vector $\bar{\mathbf{V}}$ which is the "partner" to $\mathbf{V}$ by $\bar{V}_n = \bar{g}_{nm} V_m$. It would be incorrect to write this second expansion as $\bar{\mathbf{V}} = \sum_n \bar{V}_m \mathbf{U}_m$ since that would say $\bar{g} \mathbf{V} = \sum_n \bar{V}_m \mathbf{U}_m$ 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 $\mathbf{V}$ on the tangent basis vectors $\mathbf{e}_n$, and we denote the components with a script lower-case v ($\mathbf{u}$).

$$\mathbf{V} = \mathbf{u}_1 \mathbf{e}_1 + \mathbf{u}_2 \mathbf{e}_2 + \dots = \sum_n \mathbf{u}_n \mathbf{e}_n$$

Using $\mathbf{e}_n \bullet \mathbf{E}_m = \delta_{n,m}$ one finds that

$$\mathbf{u}_n = \mathbf{E}_n \bullet \mathbf{V} = (\mathbf{E}_n)_k V_k = R_{nk} V_k = V'_n \quad // \bar{g} = 1 \text{ so } \mathbf{A} \bullet \mathbf{B} = \bar{g}_{ab} A_a B_b = A_k B_k$$

Therefore, the expansion is

$$\mathbf{V} = V'_1 \mathbf{e}_1 + V'_2 \mathbf{e}_2 + \dots = \sum_n V'_n \mathbf{e}_n \quad \text{where } \mathbf{E}_n \bullet \mathbf{V} = V'_n$$

If it happens that the N-tuple $\mathbf{V} = (V_1, V_2, \dots, V_N)$ transforms as a contravariant vector, then V_n are the contravariant components of that vector, and V'_n are the contravariant components of $\mathbf{V}'$ in x' -space. On the other hand, if $\mathbf{V}$ 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 $\mathbf{V}'$.

Writing $V'_n = g'_{nm} \bar{V}'_m$ the above expansion can be expressed as

$$\mathbf{V} = \sum_n V'_n \mathbf{e}_n = \sum_{n,m} g'_{nm} \bar{V}'_m \mathbf{e}_n = \sum_m \bar{V}'_m \sum_n g'_{mn} \mathbf{e}_n = \sum_m \bar{V}'_m \mathbf{E}_m$$

so that

$$\mathbf{V} = \bar{V}'_1 \mathbf{E}_1 + \bar{V}'_2 \mathbf{E}_2 + \dots = \sum_n \bar{V}'_n \mathbf{E}_n \quad \text{where } \mathbf{e}_n \bullet \mathbf{V} = \bar{V}'_n$$

Summary of x-space expansions:

$$\begin{array}{llll}
 \mathbf{V} = V_1 \mathbf{u}_1 + V_2 \mathbf{u}_2 + \dots & = \sum_n V_n \mathbf{u}_n & \text{where } \mathbf{U}_n \bullet \mathbf{V} = V_n & \mathbf{U}_n = g_{ni} \mathbf{u}_i \\
 \mathbf{V} = \bar{V}_1 \mathbf{U}_1 + \bar{V}_2 \mathbf{U}_2 + \dots & = \sum_n \bar{V}_n \mathbf{U}_n & \text{where } \mathbf{u}_n \bullet \mathbf{V} = \bar{V}_n & \\
 \mathbf{V} = V'_1 \mathbf{e}_1 + V'_2 \mathbf{e}_2 + \dots & = \sum_n V'_n \mathbf{e}_n & \text{where } \mathbf{E}_n \bullet \mathbf{V} = V'_n & \mathbf{E}_n = g'_{ni} \mathbf{e}_i \\
 \mathbf{V} = \bar{V}'_1 \mathbf{E}_1 + \bar{V}'_2 \mathbf{E}_2 + \dots & = \sum_n \bar{V}'_n \mathbf{E}_n & \text{where } \mathbf{e}_n \bullet \mathbf{V} = \bar{V}'_n &
 \end{array}$$

Expanding on unit vectors. The covariant lengths of the different basis vectors are given by

$$\begin{array}{ll}
 |\mathbf{e}_n| = |\mathbf{e}'_n| = \sqrt{g'_{nn}} & |\mathbf{u}_n| = |\mathbf{u}'_n| = \sqrt{g_{nn}} \\
 |\mathbf{E}_n| = |\mathbf{E}'_n| = \sqrt{g'_{nn}} & |\mathbf{U}_n| = |\mathbf{U}'_n| = \sqrt{g_{nn}}
 \end{array}$$

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

$$\hat{\mathbf{e}}_n \equiv \mathbf{e}_n / \sqrt{g'_{nn}}$$

the third expansion above becomes (AaSansOutline font for unit-vector components)

$$\mathbf{V} = \mathcal{V}'_1 \hat{\mathbf{e}}_1 + \mathcal{V}'_2 \hat{\mathbf{e}}_2 + \dots = \sum_n \mathcal{V}'_n \hat{\mathbf{e}}_n \quad \text{where} \quad \sqrt{g'_{nn}} \mathbf{E}_n \bullet \mathbf{V} = \mathcal{V}'_n = \sqrt{g'_{nn}} V'_n$$

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 $(\mathcal{V}'_1, \mathcal{V}'_2 \dots \mathcal{V}'_N)$, although related to contravariant vector $\mathbf{V}$ ($V'_n = R_{nk} V_k$), is not itself a contravariant vector since it does not obey the rule $\mathcal{V}'_n = R_{nk} \mathcal{V}'_k$. In fact

$$V'_n = R_{nk} V_k \Rightarrow (1/\sqrt{g'_{nn}}) \mathcal{V}'_n = R_{nk} (1/\sqrt{g_{kk}}) \mathcal{V}'_k \Rightarrow \mathcal{V}'_n = R_{nk} (\sqrt{g'_{nn}}/\sqrt{g_{kk}}) \mathcal{V}'_k$$

x'-space expansions

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,

$$g' \leftrightarrow g \quad R \leftrightarrow S \quad \mathbf{u}'_n \rightarrow \mathbf{e}_n \quad \mathbf{u}_n \rightarrow \mathbf{e}'_n \quad \mathbf{U}'_n \rightarrow \mathbf{E}_n \quad \mathbf{U}_n \rightarrow \mathbf{E}'_n \quad V'_n \leftrightarrow V_n \quad \bar{V}'_n \leftrightarrow \bar{V}_n$$

and here then are the x'-space expansions:

$$\begin{array}{llll}
 \mathbf{V}' = V'_1 \mathbf{e}'_1 + V'_2 \mathbf{e}'_2 + \dots & = \sum_n V'_n \mathbf{e}'_n & \text{where } \mathbf{E}'_n \bullet \mathbf{V}' = V'_n & \mathbf{E}'_n = g'_{ni} \mathbf{e}'_i \\
 \mathbf{V}' = \bar{V}'_1 \mathbf{E}'_1 + \bar{V}'_2 \mathbf{E}'_2 + \dots & = \sum_n \bar{V}'_n \mathbf{E}'_n & \text{where } \mathbf{e}'_n \bullet \mathbf{V}' = \bar{V}'_n & \\
 \mathbf{V}' = V_1 \mathbf{u}'_1 + V_2 \mathbf{u}'_2 + \dots & = \sum_n V_n \mathbf{u}'_n & \text{where } \mathbf{U}'_n \bullet \mathbf{V}' = V_n & \mathbf{U}'_n = g_{ni} \mathbf{u}'_i \\
 \mathbf{V}' = \bar{V}_1 \mathbf{U}'_1 + \bar{V}_2 \mathbf{U}'_2 + \dots & = \sum_n \bar{V}_n \mathbf{U}'_n & \text{where } \mathbf{u}'_n \bullet \mathbf{V}' = \bar{V}_n &
 \end{array}$$

(g) Another way to write the E_n

The reciprocal base vectors were defined above as linear combinations of the tangent base vectors, all in the general Picture A context,

$$\mathbf{E}_k \equiv g'_{ki} \mathbf{e}_i$$

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 $g=1$)

$$\mathbf{E}_k = \det(R) (-1)^{k-1} \mathbf{e}_1 \times \mathbf{e}_2 \times \dots \times \mathbf{e}_N \quad // \mathbf{e}_k \text{ missing} \quad k = 1, 2, 3 \dots N$$

This is a generalized cross product (Appendix A) 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

$$(\mathbf{E}_k)_\alpha \equiv \det(R) (-1)^{k-1} \varepsilon_{\alpha abc \dots x} (\mathbf{e}_1)_a (\mathbf{e}_2)_b \dots (\mathbf{e}_N)_x \quad // \kappa \text{ and } (\mathbf{e}_\kappa)_\kappa \text{ are missing}$$

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 $(\mathbf{e}_\kappa)_\kappa$ 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,

$$\begin{aligned} \mathbf{E}_1 &= \det(R) \mathbf{e}_2 \times \mathbf{e}_3 \\ \mathbf{E}_2 &= \det(R) \mathbf{e}_3 \times \mathbf{e}_1 \\ \mathbf{E}_3 &= \det(R) \mathbf{e}_1 \times \mathbf{e}_2 \end{aligned}$$

These equations can be verified in a simple manner. To show an equation is true, it 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

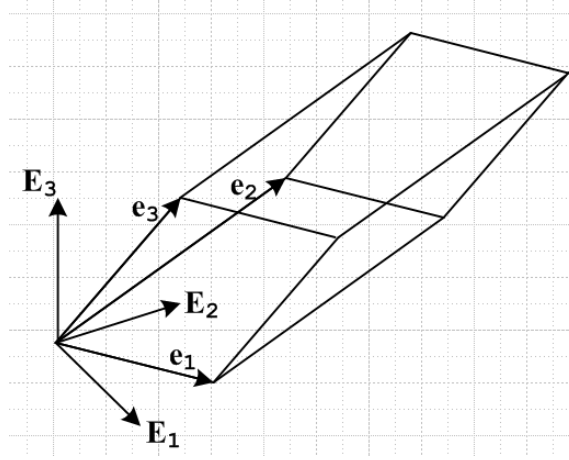
$$\mathbf{E}_1 \cdot \mathbf{e}_n = \det(R) \mathbf{e}_2 \times \mathbf{e}_3 \cdot \mathbf{e}_n \quad \text{for } n = 1, 2, 3$$

If $n=2$ or $n=3$, both sides vanish, according to $\mathbf{e}_n \cdot \mathbf{E}_m = \delta_{n,m}$ 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 $\mathbf{e}_2 \times \mathbf{e}_3 \cdot \mathbf{e}_1 = 1/\det(R) = \det(S)$. But

$$\begin{aligned} \mathbf{e}_1 \cdot \mathbf{e}_2 \times \mathbf{e}_3 &= (\mathbf{e}_1)_i (\mathbf{e}_2 \times \mathbf{e}_3)_i = (\mathbf{e}_1)_i \varepsilon_{ijk} (\mathbf{e}_2)_j (\mathbf{e}_3)_k = \varepsilon_{ijk} (\mathbf{e}_1)_i (\mathbf{e}_2)_j (\mathbf{e}_3)_k \\ &= \varepsilon_{ijk} S_{i1} S_{j2} S_{k3} = \det(S) \quad \text{QED.} \end{aligned}$$

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 $\det(R) > 0$.



The reader is invited to exercise his or her right hand to confirm the directions of the arrows,

$$\mathbf{E}_1 = \det(R) \mathbf{e}_2 \times \mathbf{e}_3$$

$$\mathbf{E}_2 = \det(R) \mathbf{e}_3 \times \mathbf{e}_1$$

$$\mathbf{E}_3 = \det(R) \mathbf{e}_1 \times \mathbf{e}_2$$

(h) Comparison of $\bar{\mathbf{e}}_n$ and $\mathbf{E}_n$

One could create a covariant partner to $\mathbf{e}_n$ which would be $\bar{\mathbf{e}}_n = \bar{g} \mathbf{e}_n$ as described in Section 5 (g). This $\bar{\mathbf{e}}_n$ is not the same as the reciprocal base vector $\mathbf{E}_n \equiv g'_{nk} \mathbf{e}_k$. The comparison is interesting:

$$(\bar{\mathbf{e}}_n)_i \equiv \bar{g}_{ik} (\mathbf{e}_n)_k \quad // \text{ matrix acts on vector index}$$

$$(\mathbf{E}_n)_i \equiv g'_{nk} (\mathbf{e}_k)_i \quad // \text{ matrix acts on } \mathbf{e}_k \text{ label}$$

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 $g' \neq 1$. We mention this to head off a possible confusion when the Standard Notation is introduced in the next Section and the above two equations become

$$(\mathbf{e}_n)_i \equiv g_{ik} (\mathbf{e}_n)^k \quad // \text{ Standard Notation, } g \text{ lowers an index}$$

$$(\mathbf{e}^n)^i \equiv g^{ik} (\mathbf{e}_k)^i \quad // \text{ Standard Notation, } k \text{ is a label on } \mathbf{e}_k, \text{ not an index}$$

The mapping to standard notation does *not* include $\bar{\mathbf{e}}_n \rightarrow \mathbf{e}^n$, for example. One fact about the standard notation is that, unlike the developmental notation, one cannot look at a vector in bold like $\mathbf{e}_n$ and determine whether it is contravariant or covariant. Only when the index is displayed can one tell. One can think of $\mathbf{e}_n$ as representing both its contravariant self and its covariant partner ($\mathbf{e}_n$ is a tensorial vector).

(i) Handedness of the $\mathbf{e}_n$ and the sign of $\det(S)$.

Handedness

As noted in Section 3 (a), since transformation F is invertible, the tangent basis vectors $\mathbf{e}_n$ are linearly independent and form a basis in x -space. In general the $\mathbf{e}_n$ are non-orthogonal, and one might ask what it means for the set of basis vectors $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_N\}$ to comprise a "right handed coordinate system".

For $N=3$, we shall define right and left-handed coordinate systems in this way,

$$\begin{array}{ll} \mathbf{e}_2 \times \mathbf{e}_3 \cdot \mathbf{e}_1 > 0 & \text{right handed system} \\ \mathbf{e}_2 \times \mathbf{e}_3 \cdot \mathbf{e}_1 < 0 & \text{left handed system} \end{array}$$

Section 5 (j) showed that

$$|\mathbf{e}_n| = \sqrt{g'_{nn}} \equiv h'_n > 0 \quad \text{"scale factor } n"$$

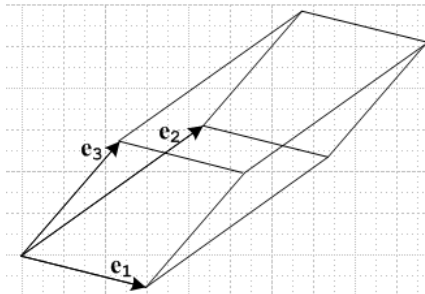
One can then define unit-vector versions of the $\mathbf{e}_n$

$$\hat{\mathbf{e}}_n = \mathbf{e}_n / h'_n$$

and then the above definition of right-handedness can equivalently be written,

$$\hat{\mathbf{e}}_2 \times \hat{\mathbf{e}}_3 \cdot \hat{\mathbf{e}}_1 > 0$$

If $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\}$ were basis vectors for an orthogonal coordinate system with $\hat{\mathbf{e}}_2 \times \hat{\mathbf{e}}_3 = \hat{\mathbf{e}}_1$ (and cyclic), one would find $\hat{\mathbf{e}}_2 \times \hat{\mathbf{e}}_3 \cdot \hat{\mathbf{e}}_1 = 1 > 0$, so the above definition agrees with the usual meaning of a right handed system in the orthogonal case. For a skewed system as shown in this picture,



it will still be true that $\hat{\mathbf{e}}_2 \times \hat{\mathbf{e}}_3 \cdot \hat{\mathbf{e}}_1 > 0$ but one will find $\hat{\mathbf{e}}_2 \times \hat{\mathbf{e}}_3 \cdot \hat{\mathbf{e}}_1 < 1$. This picture illustrates the notion that an orthogonal system can be deformed into a non-orthogonal one in such a way as to remain right handed.

Now returning to $\mathbf{e}_2 \times \mathbf{e}_3 \cdot \mathbf{e}_1 > 0$, one can write this as

$$\mathbf{e}_2 \times \mathbf{e}_3 \bullet \mathbf{e}_1 = [\mathbf{e}_2 \times \mathbf{e}_3]_k \bullet (\mathbf{e}_1)_k = \varepsilon_{kij} (\mathbf{e}_2)_i (\mathbf{e}_3)_j (\mathbf{e}_1)_k = \varepsilon_{kij} S_{i1} S_{j2} S_{k3} = \det(S)$$

Therefore, for N=3 the vectors $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\}$ form a right-handed coordinate system if $\det(S) > 0$. An example is provided by spherical coordinates where $\{\hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}}\} = \{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3\}$ form a right-handed orthogonal system and $\det(S) = r^2 \sin\theta > 0$ (Section 3 Example 2).

The above discussion can be reframed in terms of the reciprocal base vectors $\mathbf{E}_n$ along with the $\mathbf{e}_n$. Appendix A shows that

$$\mathbf{E}_k \equiv \det(R) (-1)^{k-1} \mathbf{e}_1 \times \mathbf{e}_2 \times \dots \times \mathbf{e}_N \quad // \mathbf{e}_k \text{ missing; } N > 2$$

so that for N = 3 one has

$$\mathbf{E}_1 \equiv \det(R) \mathbf{e}_2 \times \mathbf{e}_3$$

The rule for right-handedness given above

$$\mathbf{e}_2 \times \mathbf{e}_3 \bullet \mathbf{e}_1 > 0$$

can then be restated as

$$\mathbf{E}_1 \bullet \mathbf{e}_1 / \det(R) > 0$$

But $\mathbf{E}_1 \bullet \mathbf{e}_1 = 1$ from section (b) above, so this just says

$$\det(S) > 0$$

This is the same conclusion reached earlier for a right-handed system, but this second method shows the way to generalizing the idea to $N > 3$. Working backwards,

$$\det(S) > 0$$

$$\mathbf{E}_k \bullet \mathbf{e}_k / \det(R) > 0$$

$$\{ \det(R) (-1)^{k-1} \mathbf{e}_1 \times \mathbf{e}_2 \times \dots \times \mathbf{e}_N \} \bullet \mathbf{e}_k / \det(R) > 0$$

$$(-1)^{k-1} [\mathbf{e}_1 \times \mathbf{e}_2 \times \dots \times \mathbf{e}_N] \bullet \mathbf{e}_k > 0 \quad // \mathbf{e}_k \text{ missing in cross product}$$

If this last inequality is true for any k in the range (1,2..N), it is true for all k, so one could just test the case $k = 1$,

$$[\mathbf{e}_2 \times \mathbf{e}_3 \dots \times \mathbf{e}_N] \bullet \mathbf{e}_1 > 0$$

If this is true, then we call $\{ \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_N \}$ a right-handed coordinate system and $\det(S) > 0$. Otherwise the coordinate system is left-handed and $\det(S) < 0$.

In various locations we have used the symbol $\sigma = \text{sign}[\det(S)]$.

The sign of $\det(S)$

For a given ordering of the x'_i 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 $S_{i\mathbf{k}}(\mathbf{x}') \equiv (\partial x_i / \partial x'_k)$ 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 $\mathbf{x}$ (or range $\mathbf{x}'$) of $\mathbf{x}' = F(\mathbf{x})$, 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 $\det S < 0$ for both these systems.

7. Translation to the Standard Notation

In this Section 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.

(a) Outer Products

It is possible to form larger tensors from smaller ones using the "outer product" method. For example, consider,

$$T_{ab} \equiv U_a V_b$$

where U and V are assumed to be contravariant vectors. One then has

$$T'_{ab} = U'_a V'_b = (R_{aa'} U_{a'}) (R_{bb'} V_{b'}) = R_{aa'} R_{bb'} U_{a'} V_{b'} = R_{aa'} R_{bb'} T_{a'b'}$$

so in this way a contravariant rank-2 tensor (Section 5 (e)) has been successfully constructed from two contravariant vectors. Similarly,

$$\bar{T}_{ab} \equiv \bar{U}_a \bar{V}_b \quad \Rightarrow \quad \bar{T}'_{ab} = S^T_{aa'} S^T_{bb'} \bar{T}_{a'b'}$$

so the outer product of two covariant vectors transforms as a covariant rank-2 tensor.

(b) Mixed Tensors and Notation Issues

Suppose we take the "outer product" of a contravariant vector with a covariant vector,

$$[\dots]_{ab} \equiv U_a \bar{V}_b$$

where we are not sure what to call this thing, so we just call it $[\dots]$. Here is how this new object transforms (always: with respect to the underlying transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$)

$$[\dots]'_{ab} = U'_a \bar{V}'_b = (R_{aa'} U_{a'}) (S^T_{bb'} \bar{V}_{b'}) = R_{aa'} S^T_{bb'} U_{a'} \bar{V}_{b'} = R_{aa'} S^T_{bb'} [\dots]_{a'b'}$$

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

$$[.....]_{abcd} = U_a \bar{V}_b X_c \bar{Y}_d$$

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 hand side of the equation, 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 $\bar{V} \equiv \bar{g} V$ 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

$$\bar{W}_{ab} \equiv U_a \bar{V}_b$$

$$\bar{\bar{W}}_{abcd} = U_a \bar{V}_b X_c \bar{Y}_d$$

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,

$$W_{a\bar{b}c\bar{d}} = U_a V_{\bar{b}} X_c Y_{\bar{d}}$$

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 $\bar{V}_a = \bar{g}_{ab} V_b \rightarrow V_{\bar{a}} = g_{\bar{a}b} V_b$.

There are several problems with this scheme. One is that in the spinor application of tensor analysis used in special relativity (see Section 5 (m)), dots are placed on certain indices 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.

(c) 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 section (j) 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,

$$W_a{}^b{}_c{}^d = U_a V^b X_c Y^d \quad // \text{ a path not taken}$$

This is *almost* what happened, but the up/down decision went the other way and we now have:

superscripts = contravariant = up
subscript = covariant = down

and then we get this translation

$$W_{a\bar{b}c\bar{d}} = U_a V_{\bar{b}} X_c Y_{\bar{d}} \quad \rightarrow \quad W^a{}_b{}^c{}_d = U^a V_b X^c Y_d \quad // \text{ the path taken}$$

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 dx , 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 Sections into this new notation. There are some subtle details that will be discussed in the following sections.

(d) Some Preliminary Translations: raising and lowering indices on a vector with g

In the entire rest of this entire Section, anything to the left of a $\rightarrow$ arrow is in "developmental notation", while anything to the right of $\rightarrow$ is in "Standard Notation".

So we start translating some of the results above:

$$\begin{aligned} s &\rightarrow s && // \text{ a scalar} \\ V_a &\rightarrow V^a && // \text{ a contravariant rank-1 tensor (vector)} \\ \bar{V}_a &\rightarrow V_a && // \text{ a covariant rank-1 tensor (vector)} \\ M_{ab} &\rightarrow M^{ab} && // \text{ a contravariant rank-2 tensor} \\ \bar{M}_{ab} &\rightarrow M_{ab} && // \text{ a covariant rank-2 tensor} \\ g_{ab} &\rightarrow g^{ab} && // \text{ the contravariant rank-2 metric tensor} \\ \bar{g}_{ab} &\rightarrow g_{ab} && // \text{ the covariant rank-2 metric tensor} \\ g \text{ is inverse of } \bar{g} &\rightarrow g^{ab} \text{ is inverse of } g_{ab} \end{aligned}$$

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. Still, one cannot just write $\det(g)$ because $\det(g)$ is representation dependent, so one must say something like $\det(g_{ab})$ or $\det(g^{**})$ to denote a particular determinant.

As for converting a vector from one type to the other,

$$\begin{aligned}\bar{V}_a &= \bar{g}_{ab} V_b & \rightarrow & V_a = g_{ab} V^b & // g_{ab} \text{ "lowers" a contravariant index} \\ V_a &= g_{ab} \bar{V}_b & \rightarrow & V^a = g^{ab} V_b & // g^{ab} \text{ "raises" a covariant index ,}\end{aligned}$$

and so in this new 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, as will be seen below.

As for bolded vectors, the translation rule is,

$$\begin{aligned}\mathbf{V} &\rightarrow \mathbf{V} \\ \bar{\mathbf{V}} &\rightarrow \mathbf{V}\end{aligned}$$

The reason is that the overbar is no longer used to denote covariance. The above lines show a subtle change in the interpretation of the bolded symbol $\mathbf{V}$ in the standard notation: the single symbol $\mathbf{V}$ stands for both the developmental vector $\mathbf{V}$ and for its developmental covariant partner vector $\bar{\mathbf{V}}$. The new symbol $\mathbf{V}$ is both contravariant with components V^a and it is covariant with components V_a .

The invariant distance and covariant dot products:

$$\begin{aligned}dx_i &\rightarrow dx^i \\ (ds)^2 &= \bar{g}_{ab} dx_a dx_b \quad \rightarrow g_{ab} dx^a dx^b \\ \mathbf{A} \bullet \mathbf{B} &= \bar{g}_{ab} A_a B_b \quad \rightarrow \mathbf{A} \bullet \mathbf{B} = g_{ab} A^a B^b = A_b B^b = A^a B_a = g^{ab} A_a B_b\end{aligned}$$

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 for a short while.

(e) Contraction of a Pair of Indices

When two indices are summed together in a tensor expression and one is up and the other down, one says that the two indices are **contracted**. Here is an example, where the index b is contracted,

$$V_a = g_{ab}V^b$$

It is shown below that contracted indices neutralize each other in terms of how an object transforms. Thus, for example, the RHS above $g_{ab}V^b$ transforms as a covariant vector, which conveniently matches the LHS. Similarly, A^aB_a transforms as a scalar.

(f) Dealing with the matrix R

Consider the translation of this partial derivative into the new up/down notation. Since the differential dx element is contravariant and is now written dx^i ,

$$(\partial x'_i / \partial x_k) \rightarrow (\partial x'^i / \partial x^k)$$

In terms of "existence", this object has one leg in each space of Picture A. The gradient operator $\partial / \partial x_k$ is an x -space thing, while x'^i is an x' -space thing. Since this object 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 that the gradient is a covariant vector, so we regard k as a covariant index. And of course dx'^i is a contravariant vector, so i is a contravariant index. Here then is the proper translation:

$$R_{ik} \equiv (\partial x'_i / \partial x_k) \rightarrow R^i_k \equiv (\partial x'^i / \partial x^k)$$

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 very typical for an object to have up and down indices but the object is not a tensor. 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 a contravariant vector.

Consider now the translation of the transformation rule for a contravariant vector

$$V'_a = R_{ab}V_b \rightarrow V'^a = R^a_b V^b \quad // \text{ contravariant}$$

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. So 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) 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 one space, and V' and the other half of R are in a different space. There is no object called R' , as if R were in x -space and R' were in x' -space.

(g) Repeat the above section for S

We omit the words and just show the translations

$$\begin{aligned}
 (\partial x_i / \partial x'^k) & \rightarrow (\partial x^i / \partial x'^k) \\
 S_{ik} \equiv (\partial x_i / \partial x'^k) & \rightarrow S^i_k \equiv (\partial x^i / \partial x'^k) \\
 \bar{V}'_a = S^T_{ab} \bar{V}_b = S_{ba} \bar{V}_b & \rightarrow V'_a = S^b_a V_b \quad // \text{ covariant}
 \end{aligned}$$

(h) About ϵ and δ

The Kronecker δ is sometimes written in different ways to make things "look nice",

$$\delta_a^b = \delta^a_b = \delta_b^a = \delta^b_a = \delta_{a,b}$$

Section (m) will show that one can regard the above sequence of equalities as saying

$$g_a^b = g^a_b = g_b^a = g^b_a = \delta_{a,b}$$

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 tensor, but tradition is to write the diagonal objects using the δ symbol.

The object $\epsilon_{abc} \dots$ 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

$$\det(M) = \epsilon_{abc} \dots x M_{1a} M_{2b} \dots M_{Nx} = \epsilon_{abc} \dots x M_{a1} M_{b2} \dots M_{xN}$$

or in an ordinary cross product

$$A_a = \epsilon_{abc} B_b C_c$$

This permutation tensor has the usual properties that $\epsilon_{123 \dots N} = +1$, 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, $\epsilon_{abc} \dots x$ 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 (d) just as with any tensor. Therefore, in Cartesian space with $g = 1$, indices on ϵ are raised and lowered with no consequence, and then one can identify any form of ϵ as being the permutation tensor. For example, $\epsilon^{abc} = \epsilon_{abc} = \epsilon_a^b{}_c$ and so on. In a non-Cartesian x -space, however, one would say that $\epsilon_a^b{}_c = g^{bb'} \epsilon_{ab'c} \neq \epsilon_{abc}$. In the Weinberg convention, one sets $\epsilon^{123 \dots N} = \epsilon'_{123 \dots N} = 1$ and $\epsilon^{abc \dots x} = \epsilon'_{abc \dots x}$ 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 $g \neq 1$, $\epsilon_{abc\dots x}$ 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 is always just the permutation tensor, but in the cross product that is not the case, and one would properly write

$$A_a = \epsilon_{abc} B^b C^c$$

and conclude that the cross product of two ordinary contravariant vectors is a covariant vector density (Appendix D (g)). Again, in Cartesian space where one often works, this would be the same as $A_a = \epsilon^{abc} B^b C^c = \epsilon_a^b{}^c B^b C^c$, but the "properly tilted form" $A_a = \epsilon_{abc} B^b C^c$ reveals the tensor nature of the object A_a . As mentioned below in section (u), this "covariant" equation would appear as $A'_a = \epsilon'_{abc} B'^b C'^c$ in x' -space, but since A'_a is a covariant vector density, $A'_a \neq R_a^b A_b$, and in fact $A'_a = |J| R_a^b A_b$.

The permutation tensor $\epsilon_{abc\dots}$ and the contravariant Levi-Civita tensor $\epsilon^{abc\dots x}$ are both "totally antisymmetric" which just means ϵ changes sign if any pair of indices is swapped. In fact, as discussed in Appendix D (c), there IS only one antisymmetric tensor of rank N apart from a multiplicative scalar factor, and $\epsilon^{abc\dots x}$ is it. This fact simplifies various calculations. Technically, $\epsilon^{abc\dots x}$ is a totally antisymmetric tensor *density*, but normally it is just called "the totally antisymmetric tensor". As shown in Appendix D, the covariant Levi-Civita tensor $\epsilon_{abc\dots x}$ is also totally antisymmetric and is therefore a multiple of $\epsilon^{abc\dots x}$.

The reader is invited to peruse Appendix D at some appropriate time for more about tensor densities and the ϵ tensor.

(i) Further development of the Standard Notation

This long section contains a veritable grab-bag of important Standard Notation facts. Each such fact is proven and not just quoted. Many of the results presented here anticipate more formal presentations of the same results in later sections.

Translation of determinants. Section (g) showed that $R_{ab} \rightarrow R^a{}_b$, so one translates from old to new notation,

$$\begin{aligned} \det(R) = \epsilon_{abc\dots} R_{1a} R_{2b} \dots R_{Nx} &\rightarrow \det(R^i{}_j) = \epsilon_{abc\dots} R^1{}_a R^2{}_b \dots R^N{}_x \\ \det(S) = \epsilon_{abc\dots} S_{1a} S_{2b} \dots S_{Nx} &\rightarrow \det(S^i{}_j) = \epsilon_{abc\dots} S^1{}_a S^2{}_b \dots S^N{}_x \\ \\ \det(R) = \epsilon_{abc\dots} R_{a1} R_{b2} \dots R_{xN} &\rightarrow \det(R^i{}_j) = \epsilon_{abc\dots} R^a{}_1 R^b{}_2 \dots R^x{}_N \\ \det(S) = \epsilon_{abc\dots} S_{a1} S_{b2} \dots S_{xN} &\rightarrow \det(S^i{}_j) = \epsilon_{abc\dots} S^a{}_1 S^b{}_2 \dots S^x{}_N \end{aligned}$$

where ϵ is the bookkeeping permutation tensor discussed in section (h).

Inverse of R and S. Again, section (g) showed that $R_{ab} \rightarrow R^a{}_b$. 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

$$(\partial x^c / \partial x'^a) (\partial x'^a / \partial x^b) = \delta^c{}_b \quad \text{or} \quad S^c{}_a R^a{}_b = \delta^c{}_b$$

and therefore it must be that $(R^{-1})^c_a = S^c_a$. A similar argument shows that $(S^{-1})^c_a = R^c_a$. Using the results of the next section which allow us to raise and lower indices on both sides of an equation, this relationships $R^{-1} = S$ is valid for all four matrix position possibilities,

$$\begin{aligned}(R^{-1})^{ik} &= S^{ik} \\ (R^{-1})^i_k &= S^i_k \\ (R^{-1})_i^k &= S_i^k \\ (R^{-1})_{ik} &= S_{ik}\end{aligned}$$

and of course the same is true for $S^{-1} = R$. Thus arise these translations from old to new notation:

$$\begin{array}{lll} R^{-1} = S & \rightarrow & (R^{-1})^i_k = S^i_k \text{ and all other index combinations} \\ S^{-1} = R & \rightarrow & (S^{-1})^i_k = R^i_k \text{ and all other index combinations} \\ RR^{-1} = RS = 1 \text{ etc} & \rightarrow & R^i_k (R^{-1})^k_a = R^i_k S^k_a = \delta^i_a \text{ etc} \end{array} \quad (1)$$

Tensor g raises and lowers any index. So far the following translation rules have been established:

$$g_{ab} \rightarrow g^{ab} \quad \bar{g}_{ab} \rightarrow g_{ab} \quad R_{ik} \rightarrow R^i_k \quad S_{ik} \rightarrow S^i_k$$

It was shown in developmental notation Section 5 (e) how a rank-2 contravariant tensor transforms. Here then is how that statement translates to the new notation

$$\begin{array}{lll} M'_{ab} = R_{aa'} R_{bb'} M_{a'b'} & \rightarrow & M'^{ab} = R^a_{a'} R^b_{b'} M^{a'b'} \quad \text{contravariant rank-2 tensor} \\ \bar{M}'_{ab} = S_{a'a} S_{b'b} \bar{M}_{a'b} & \rightarrow & M'_{ab} = S^{a'}_a S^{b'}_b M_{a'b} \quad \text{covariant rank-2 tensor} \end{array} \quad (2)$$

Since g is such a rank-2 tensor, replace M by g to get

$$\begin{array}{lll} g'_{ab} = R_{aa'} R_{bb'} g_{a'b'} & \rightarrow & g'^{ab} = R^a_{a'} R^b_{b'} g^{a'b'} \\ \bar{g}'_{ab} = R_{aa'} R_{bb'} \bar{g}_{a'b} & \rightarrow & g'_{ab} = S^{a'}_a S^{b'}_b g_{a'b} \end{array} \quad (3)$$

It was shown in section (d) that $V_a = g_{aa'} V^{a'}$ and $V^a = g^{aa'} V_{a'}$, so that $g_{aa'}$ lowers a vector index and $g^{aa'}$ raises a vector index. That is to say, $g_{aa'}$ converts a contravariant vector index into a covariant one, and $g^{aa'}$ does the reverse.

What does $g_{aa'}$ do to the index of a rank-2 tensor? Consider the following definition:

$$M^a_b \equiv g^{aa'} M_{a'b} \quad (4)$$

Since g^{ab} and g_{ab} are inverses, it follows that

$$M_{ab} = g_{aa'} M^{a'}_b \quad (5)$$

How does this new object M^a_b transform? The *claim* is that it transforms as a mixed rank-2 tensor, which would mean that

$$M'^a_b = R^a_{a'} S^{b'}_b M^{a'}_{b'}$$

The upper index gets a factor $R^a_{a'}$ and the lower index gets a factor $S^{b'}_b$, consistent with (*) above. It is not hard to prove this claim:

$$\begin{aligned} M'^a_b &= g'^{aa'} M'_{a'b} = \overset{(4)}{(R^a_c R^{a'}_d g^{cd})} \overset{(3)}{(S^e_{a'} S^f_b M_{ef})} = \overset{(2)}{(R^a_c R^{a'}_d g^{cd})} \overset{(5)}{(S^e_{a'} S^f_b g_{ei} M^i_f)} \\ &= [R^a_c R^{a'}_d g^{cd} S^e_{a'} S^f_b g_{ei}] M^i_f \\ &= [R^a_c (S^e_{a'} R^{a'}_d) g^{cd} S^f_b g_{ei}] M^i_f = [R^a_c (SR)^e_d g^{cd} S^f_b g_{ei}] M^i_f \\ &= [R^a_c \delta^e_d g^{cd} S^f_b g_{ei}] M^i_f = [R^a_c g^{cd} S^f_b g_{di}] M^i_f \quad (1) \\ &= [R^a_c (g^{cd} g_{di}) S^f_b] M^i_f = [R^a_c \delta^c_i S^f_b] M^i_f = [R^a_i S^f_b] M^i_f \quad \text{inverses} \\ &= R^a_{a'} S^{b'}_b M^{a'}_{b'}, \quad \text{QED} \end{aligned}$$

Similarly one could define $M_a^b \equiv g_{aa'} M^{a'b}$ and one would find that

$$M'^a_b = S^{a'}_a R^b_b M_{a'}^{b'}$$

so M_a^b is then another member of the family of rank-2 tensors. Finally were one to define $M_{ab} \equiv g_{bb'} M_a^{b'}$ one would find that M_{ab} transforms as in (2). To summarize the four transformation results

$$\begin{aligned} M'^{ab} &= R^a_{a'} R^b_{b'} M^{a'b'} \\ M'^a_b &= R^a_{a'} S^{b'}_b M^{a'}_{b'} \\ M'^a_b &= S^{a'}_a R^b_{b'} M_{a'}^{b'} \\ M'_{ab} &= S^{a'}_a S^{b'}_b M_{a'}^{b'} \end{aligned}$$

One sees then a family of four tensors associated with M . One is contravariant, one is covariant, and the other two are mixed.

Let $[---^i---$] represent a tensor with a certain contravariant index i and dashes indicate other indices which could be up or down. Similarly define $[---_i---$] as another tensor in the same family where the index i that was up is now down.

Raising Lowering Rule:

$$\begin{aligned} g_{aa'} [---^{a'}---] &= [---_a---] \\ \text{and } g^{aa'} [---_a---] &= [---^{a'}---] \end{aligned}$$

The notion of higher rank tensors is coming soon, but we just want to establish the general idea that ANY index on ANY tensor can be raised or lowered by an appropriate g tensor. For the rank-2 tensors this was demonstrated explicitly above, and section (d) showed it was valid for rank-1 tensors (vectors),

$$\begin{aligned} g_{aa'} V^{a'} &= V_a \\ g^{aa'} V_a &= V^{a'} \end{aligned}$$

Comment: Notice that in every equation shown above, the summed indices always occur in the contracted form discussed in section (e) above, which is to say, one index is up and the other is down.

Contraction Tilt Reversal Rule: $[\text{-----}^a\text{-----}_a\text{-----}] = [\text{-----}_a\text{-----}^a\text{-----}]$

This is proved in section (k) below, but since we are going to need it right now, here is a preview of that proof: (note that $g^{ab} g_{ac} = g^{ba} g_{ac} = \delta^b_c$)

$$[\text{-----}^a\text{-----}_a\text{-----}] = g^{ab} g_{ac} [\text{-----}_b\text{-----}^c\text{-----}] = \delta^b_c [\text{-----}_b\text{-----}^c\text{-----}] = [\text{-----}_b\text{-----}^b\text{-----}]$$

The upshot is that one can "reverse the tilt" on any pair of contracted indices.

The Diagonal g Rule: $g^a_b = \delta^a_b$ and $g_a^b = \delta_a^b$ // and same for g'

This is proved in section (m) below, but since we are going to need it right now, here is a preview of that proof:

$$\begin{aligned} g^a_b &= g^{aa'} g_{a'b} && // g^{aa'} \text{ raises the first index of tensor } g_{a'b} \\ &= \delta^a_b && // \text{ because } g_{ij} \text{ and } g^{ij} \text{ are inverses of each other} \end{aligned}$$

Matrix Multiplication in the Standard Notation. Although there are various forms of matrix multiplication, the most standard form is that obtained when all matrices have a "down-tilt" form. Consider this example,

$$C^a_b = A^a_c B^c_b$$

where it is assumed that all three objects are down-tilt rank-2 tensors (or objects like R^a_c and S^a_c whose indices behave as if they rank-2 tensors). Although these are "split level" matrices, once can see that the index c has the correct "adjacency" property to justify matrix multiplication. A second requirement is that any matrix summed index must be a genuine contraction with one index up and the other down. The above tensor transformation rule can then be written in this more compact matrix notation,

$$C = AB \quad // \text{ all down-tilt}$$

An example appears in (1) above:

$$\delta^i_a = R^i_k S^k_a = \quad \leftrightarrow \quad 1 = RS \quad S^{-1} = R$$

By application of suitable g tensors to the first equation above (or by lowering index a and raising index b), one gets $C_a^b = A_{ac}B^{cb}$. Application of the Tilt Reversal Rule on index c then gives

$$C^a_b = A^a_c B^c_b$$

Again the adjacency and contraction rules are met, so this equation can be represented also by

$$C = AB \quad // \text{ all up-tilt}$$

and so

$$\delta_i^a = R_i^k S_k^a = \leftrightarrow \quad 1 = RS \quad S^{-1} = R$$

Comment: Notice that the equation $1 = RS$ is valid both in the standard notation (providing R,S and 1 all have the same tilt) and in the developmental notation.

The upshot is that in Standard Notation, matrix notation can be used if all matrices in the equation being represented are either all down-tilt or all up-tilt. As will be shown later, such matrix equations are all "covariant" in that both sides of the equation have the same tensor transformation property, and this is due to the fact that the matrix summation index is a contraction. One could *talk about* matrix multiplication in other cases, such as

$$C^{ab} = A^{ac} B^{cb}$$

but since index c is not a contraction, if A and B are tensors, then C cannot be a tensor and we don't even want to think about such equations.

Transpose of a rank-2 tensor. If A is a rank-2 tensor, the translation mapping

$$(A^T)_{ab} = A_{ba} \quad \rightarrow \quad (A^T)^{ab} = A^{ba}$$

seems obvious, and the object A^T therefore also transforms as a rank-2 tensor. Once $(A^T)^{ab} = A^{ba}$ is established in standard notation, one can apply the metric tensor g to lower either or both of the indices of this equation, to get $(A^T)_b^a = A_b^a$, $(A^T)_a^b = A_a^b$, and $(A^T)_{ab} = A_{ba}$. Notice in all four equations that the indices on the two sides of the equation are reflected in a vertical axis passing between the indices. This causes a left index to become a right index and vice versa, as one would expect for transposing a matrix. Moreover, on each side of all four equations, each index has the same contravariant/covariant sense.

This same argument also applies to R and S even though they are not tensors. The only difference is that the first index of R^{ba} is lowered by g' while the second by g. For example, $(R^T)_b^a = R_b^a$ where b is a g' type index and a is a g type index, as will be elaborated in section (o) below. Similarly $(S^T)_b^a = S_b^a$.

To summarize the situation with transposes in Standard Notation:

$(A^T)^{ab} = A^{ba}$	$(R^T)^{ab} = R^{ba}$	$(S^T)^{ab} = S^{ba}$
$(A^T)_b^a = A_b^a$	$(R^T)_b^a = R_b^a$	$(S^T)_b^a = S_b^a$
$(A^T)_a^b = A_a^b$	$(R^T)_a^b = R_a^b$	$(S^T)_a^b = S_a^b$
$(A^T)_{ab} = A_{ba}$	$(R^T)_{ab} = R_{ba}$	$(S^T)_{ab} = S_{ba}$

Notice that the rule is *not* $(A^T)^a_b = A^b_a$ which would be a straight swap of indices (and would result in A^T not being a tensor). The straight swap idea works for the pure contravariant and pure covariant forms of A , but not for the tilted forms! As shown in Theorem 4 below, this tilted transpose form has some interesting implications.

Theorem 1: For either the down-tilt or up-tilt version of S , S is a real-orthogonal matrix.

This means all of the following: $S^{-1} = S^T \quad SS^T = 1 \quad S^TS = 1$

Proof of theorem: start with (3) above which says g_{ij} is a covariant rank-2 tensor,

$$\begin{aligned}
 g'_{ab} &= S^{a'}_a S^{b'}_b g_{a'b'} && // \text{ next, apply } g^{b'b} \text{ to both sides (or just raise index b on both sides)} \\
 g'^a_b &= S^{a'}_a S^{b'b} g_{a'b'} && // \text{ next, reverse the tilt of the b' index} \\
 g'^a_b &= S^{a'}_a S_{b'}^{b'} g_{a'b'} && // \text{ next, use the Diagonal g rule in two places} \\
 \delta_a^b &= S^{a'}_a S_{b'}^{b'} \delta_{a'}^{b'} = S^{a'}_a S_{a'}^b && // \text{ next, use } (S^T)_a^{a'} = S^{a'}_a \\
 \delta_a^b &= (S^T)_a^{a'} S_{a'}^b && // \text{ next use matrix notation as per above} \\
 1 &= S^TS && \Rightarrow \quad S^T = S^{-1} \quad \Rightarrow \quad 1 = SS^T
 \end{aligned}$$

Theorem 2: For either the down-tilt or up-tilt version of R , R is a real-orthogonal matrix.

This means all of the following: $R^{-1} = R^T \quad RR^T = 1 \quad R^TR = 1$

In other words, the previous theorem also applies to R . Again, this fact is not true of the developmental notation matrix R_{ab} . The proof is very similiar but just different enough to warrant showing it :

Proof of theorem: start with (3) above which says g^{ij} is a covariant rank-2 tensor,

$$\begin{aligned}
 g'^{ab} &= R^a_{a'} R^b_{b'} g^{a'b'} && // \text{ next, apply } g_{b'b} \text{ to both sides (or just lower index b on both sides)} \\
 g'^a_b &= R^a_{a'} R_{b'b} g^{a'b'} && // \text{ next, reverse the tilt of the b' index} \\
 g'^a_b &= R^a_{a'} R_b^{b'} g^{a'b'} && // \text{ next, use the Diagonal g rule in two places} \\
 \delta^a_b &= R^a_{a'} R_b^{b'} \delta^{a'}_{b'} = R^a_{a'} R_b^{a'} && // \text{ next, use } (R^T)^{a'}_b = R_b^{a'} \\
 \delta^a_b &= R^a_{a'} (R^T)^{a'}_b && // \text{ next use matrix notation as per above} \\
 1 &= RR^T && \Rightarrow \quad R^T = R^{-1} \quad \Rightarrow \quad 1 = R^TR
 \end{aligned}$$

Comment: In the developmental notation, neither R nor S is a real-orthogonal matrix, unless by accident, but in the tilted standard notation they both are!

Theorem 3: For either the down-tilt or up-tilt version of S ,

$$\begin{aligned}
 (a) \quad & S = R^T \quad \text{and} \quad R = S^T \\
 (b) \quad & S^a_b = R_b^a \quad \text{and} \quad S_a^b = R^b_a \quad \quad \quad (\text{ reflect indices in vertical line between them})
 \end{aligned}$$

Proof of theorem: From the inverse discussion at the start of this section, $S = R^{-1}$ for any index positions in the standard notation, including the up-tilt and down-tilt positions. From Theorem 2, $R^T = R^{-1}$ in either up-tilt or down-tilt forms. Therefore $S = R^T$ (hence $S^T = R^{TT} = R$) in either up-tilt or down-tilt form. Thus

$$\begin{aligned} S = R^T &\Rightarrow S_a^b = (R^T)_a^b = R_b^a \\ S = R^T &\Rightarrow S_b^a = (R^T)_b^a = R_a^b \end{aligned} \quad \text{QED}$$

An implication of this theorem is that one can completely eliminate references matrix S in tensor analysis and that is what is usually done. This is like replacing S with R^{-1} in the developmental notation.

Theorem 4: For a standard notation tilted matrix, $\det(A) \neq \det(A^T)$.

This surprising result points out a potential hazard of using matrices in the standard notation, and perhaps is an indication of why people avoid matrix concepts and just write out all the components.

For a traditional matrix A_{ab} the determinant is given by either of these forms (figure on rows or columns)

$$\begin{aligned} \det(A) &= \varepsilon_{ab\dots} A_{1a} A_{2b} \dots = \varepsilon_{ab\dots} A_{a1} A_{b2} \dots \\ \det(A^T) &= \varepsilon_{ab\dots} A_{1a}^T A_{2b}^T \dots = \varepsilon_{ab\dots} A_{a1} A_{b2} \dots = \det(A) \end{aligned}$$

In the tilted standard notation, however, one has

$$\begin{aligned} \det[A^i_j] &= \varepsilon_{ab\dots} A^1_a A^1_b \dots = \varepsilon_{ab\dots} A^a_1 A^b_2 \dots \\ \det[(A^T)^i_j] &= \varepsilon_{ab\dots} (A^T)^1_a (A^T)^1_b \dots = \varepsilon_{ab\dots} A^1_a A^2_b \end{aligned}$$

but this is not the same as either of the $\det[A^i_j]$ forms! Here is a simple example:

$$\begin{aligned} \det(A) &= \varepsilon_{ab\dots} A^1_a A^1_b \dots = \begin{vmatrix} A^1_1 & A^1_2 \\ A^2_1 & A^2_2 \end{vmatrix} = A^1_1 A^2_2 - A^2_1 A^1_2 \\ \det(A^T) &= \varepsilon_{ab\dots} A^1_a A^2_b = \begin{vmatrix} A^1_1 & A^2_1 \\ A^1_2 & A^2_2 \end{vmatrix} = A^1_1 A^2_2 - A^2_1 A^1_2 \neq \det(A) \end{aligned}$$

The point is that the standard notation transpose rule $(A^T)^i_j = A_i^j$ doesn't just swap the indices, it also changes the tilt. That means that the columns of one determinant matrix are not the same as the rows of the other and that is why $\det(A) \neq \det(A^T)$.

Corollary: $R^T R = 1$ does not imply that $\det(R) = \pm 1$.

The usual proof goes that $\det(R^T R) = \det(R^T) \det(R) = \det(R) \det(R) = [\det(R)]^2 = 1$, but of course the part saying $\det(R^T) = \det(R)$ is no longer valid. Therefore, the fact that $R^T R = 1$ does not lead us to the

false conclusion that the Jacobian of Section 5 (k) is somehow forced to be ± 1 ! So although R and S are real orthogonal matrices, they are not "rotation matrices".

Orthogonality rules. The first two rules are easy to show,

$$\begin{aligned} R^T R = 1 & \Rightarrow (R^T)^a_b R^b_c = \delta^a_c & \Rightarrow R_b^a R^b_c = \delta^a_c \\ R R^T = 1 & \Rightarrow R^a_b (R^T)^b_c = \delta^a_c & \Rightarrow R^a_b R^b_c = \delta^a_c \end{aligned}$$

Lowering a and raising c on both sides and reversing the b tilt then gives the other two rules

$$\begin{aligned} R^b_a R_b^c &= \delta_a^c \\ R_a^b R^c_b &= \delta_a^c \end{aligned}$$

These orthogonality rules are derived in a slightly different manner and order in section (r) below.

Variations on the relation between g and g'. Item (3) above gave the basic statement of the transformation properties of tensor g. These can be inverted as follows

$$\begin{aligned} g'^{ab} &= R^a_a R^b_b g^{a'b'} & \Rightarrow g^{ab} &= (R^{-1})^a_a (R^{-1})^b_b g'^{a'b'} = S^a_a S^b_b g'^{a'b'} \\ g'_{ab} &= S^{a'}_a S^{b'}_b g_{a'b'} & \Rightarrow g_{ab} &= (S^{-1})^{a'}_a (S^{-1})^{b'}_b g'_{a'b'} = R^{a'}_a R^{b'}_b g'_{a'b'} \end{aligned}$$

and here then is a summary,

$$\begin{aligned} g'^{ab} &= R^a_a R^b_b g^{a'b'} & g^{ab} &= S^a_a S^b_b g'^{a'b'} \\ g'_{ab} &= S^{a'}_a S^{b'}_b g_{a'b'} & g_{ab} &= R^{a'}_a R^{b'}_b g'_{a'b'} \end{aligned}$$

If x-space is Cartesian with $g = 1$, the first column above simplifies to

$$\begin{aligned} g'^{ab} &= R^a_c R^b_c & // g &= 1 \\ g'_{ab} &= S^c_a S^c_b = R^c_a R^c_b & // g &= 1 \end{aligned}$$

Notice that the summation index c is not a contraction here. Also, although R and S are not tensors, the sums shown produce the true tensors g'^{ab} and g'_{ab} .

(j) 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 $S^{a'}_a (= R^{a'}_a)$ as shown below in section (q)) for each covariant index, as illustrated by this example :

$$T'^{abc}_{de} = R^a_a R^b_b R^c_c S^{d'}_d S^{e'}_e T^{a'b'c'}_{d'e'}$$

$$T'^{abc}_{de} = R^a_a R^b_b R^c_c R^{d'}_d R^{e'}_e T^{a'b'c'}_{d'e'}$$

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

$$A^a B^b C^c D_d E_e$$

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 $T^{abc}_{de}(x)$, then of course all family members are tensor fields.

Direct Products. Consider again the outer product of vectors $A^a B^b C^c D_d E_e$. 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 $B'^a = R^a_c B^c$ occurs in an identical copy of the space $\mathcal{R}$, but transformation $D'_d = S^{d'}_d D_d = R_d^{d'} D_d$ occurs in a covariant version of $\mathcal{R}$ we call $\bar{\mathcal{R}}$. 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 (i).

The transformation of the outer product object, as already noted, is given by

$$A'^a B'^b C'^c D'_d E'_e = R^a_a R^b_b R^c_c R_d^{d'} R_e^{e'} A^a B^b C^c D_d E_e$$

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

$$\mathcal{R}_{dp} = \mathcal{R} \otimes \mathcal{R} \otimes \mathcal{R} \otimes \bar{\mathcal{R}} \otimes \bar{\mathcal{R}}$$

One could then define

$$(\mathcal{R}_{dp})^{abc}_{de} ; a'b'c', d'e' \equiv R^a_a R^b_b R^c_c R_d^{d'} R_e^{e'}$$

so that

$$A'^a B'^b C'^c D'_d E'_e = (\mathcal{R}_{dp})^{abc}_{de} ; a'b'c', d'e' A^a B^b C^c D_d E_e$$

and of course this would apply to any tensor of the same index configuration, such as

$$T^{abc}_{de} = \mathcal{R}_{dp}^{abc}_{de} ; a'b'c', d'e' T^{a'b'c'}_{d'e'}$$

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 $\mathbf{1}$ 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 $\mathcal{R}_{\mathbf{a}\mathbf{p}}$ 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 $\mathcal{R}_{\mathbf{a}\mathbf{p}} = \mathcal{R} \otimes \mathcal{R}$ can be written as a sum of "irreducible" representations of that group. What this means is that the transformation elements of $\mathcal{R}_{\mathbf{a}\mathbf{p}}$ and the objects $T^{\mathbf{a}\mathbf{b}}$ can be shuffled around with linear combinations so that $(\mathcal{R}_{\mathbf{a}\mathbf{p}})^{\mathbf{a}\mathbf{b}}_{\mathbf{a}'\mathbf{b}'}$, when thought of as a matrix with columns labeled by N^2 $\mathbf{a}\mathbf{b}$ possibilities and rows labeled by the N^2 $\mathbf{a}'\mathbf{b}'$ possibilities, appears in "block diagonal form" with all zeros outside the blocks. In this case, the shuffled components of tensor $T^{\mathbf{a}\mathbf{b}}$ can be regarded as a non-interacting assembly of pieces each of which transforms according to one of those blocks of the shuffled $(\mathcal{R}_{\mathbf{a}\mathbf{p}})^{\mathbf{a}\mathbf{b}}_{\mathbf{a}'\mathbf{b}'}$.

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 $\oplus$ symbols indicate this block diagonal form. In this case the blocks are 5×5 , 3×3 and 1×1 , fitting onto the diagonal of the 9×9 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 5×5 block are called $D^{(2)}_{\mathbf{M}, \mathbf{M}'}(\phi, \theta, \psi)$ where $\mathbf{M}, \mathbf{M}' = 2, 1, 0, -1, -2$, 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 $\mathbf{A}$ and $\mathbf{B}$, one can identify $\mathbf{A} \cdot \mathbf{B}$ as the combination transforming according to $\mathcal{R}^{(0)}$ ("scalar") and $\mathbf{A} \times \mathbf{B}$ (linearly combined) as that transforming as $\mathcal{R}^{(1)}$ ("vector"). The traceless matrix $A_i B_j - \delta_{ij} \mathbf{A} \cdot \mathbf{B}$ 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 $T^{\mathbf{a}\mathbf{b}\mathbf{c}}$.

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 $\mathbf{M}' = 0$, and then $D^{(L)}_{\mathbf{M}, 0}(\phi, \theta, \psi) \sim Y_{\mathbf{LM}}(\theta, \phi)$, the famous spherical harmonics that describe the "orbitals" which have mystified first-year chemistry students for the last 100 years.

Historical Note: 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 $X_{r_1 r_2 \dots r_m}$, ($r_1, r_2, \dots, r_m$ pouvant prendre chacun toutes les valeurs $1, 2, \dots, n$), si les éléments $Y_{r_1 r_2 \dots r_m}$ du système transformé sont donnés par les formules

$$(6) \quad Y_{r_1 r_2 \dots r_m} = \sum_{s_1 s_2 \dots s_m}^n X_{s_1 s_2 \dots s_m} \frac{\partial x_{s_1}}{\partial y_{r_1}} \frac{\partial x_{s_2}}{\partial y_{r_2}} \dots \frac{\partial x_{s_m}}{\partial y_{r_m}}.$$

Nous désignerons au contraire par des symboles tels que $X^{(r_1 r_2 \dots r_m)}$ les éléments d'un système *contravariant*, c'est à dire d'un système, dont la transformation est représentée par les formules

$$(7) \quad Y^{(r_1 r_2 \dots r_m)} = \sum_{s_1 s_2 \dots s_m}^n X^{(s_1 s_2 \dots s_m)} \frac{\partial y_{r_1}}{\partial x_{s_1}} \frac{\partial y_{r_2}}{\partial x_{s_2}} \dots \frac{\partial y_{r_m}}{\partial x_{s_m}},$$

les éléments X et Y se rapportant respectivement aux variables x et y . —

We will say that a system of order m is covariant (and in this case we will designate its elements by the symbol $X_{r_1, r_2, \dots}$) ($r_1, r_2, \dots$ can each take all the values $1 \dots n$), if the elements $Y_{r_1, r_2, \dots}$ of the transformed system are given by the formulas (6).

We will designate on the contrary by the symbols $X^{(r_1, r_2, \dots)}$ 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.

(k) The Contraction Tilt-Reversal Rule

In some complicated combination of multiple tensors and perhaps some R and S objects, imagine there is somewhere a pair of summed indices where one is up and the other is down. As noted above, 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,

$$[\text{-----}^a \text{-----}_a \text{-----}]$$

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,

$$[\text{-----}^a \text{-----}_a \text{-----}] = g^{ab} g_{ac} [\text{-----}_b \text{-----}^c \text{-----}]$$

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, $g^{ab} g_{ac} = g^{ba} g_{ac} = \delta_{a,c}$, which at once gives the desired result

$$[\text{-----}^a\text{-----}_a\text{-----}] = [\text{-----}_a\text{-----}^a\text{-----}] \quad // \text{ the Contraction Tilt-Reversal Rule}$$

A notable example of course is this:

$$A^a Y_a = A_a Y^a = "A \bullet Y" \quad // \text{ or perhaps } "A.Y" \text{ as noted in Section 5 (i)}$$

(l) The Contraction Neutralization Rule

A contracted index pair plays no role in how an object transforms, the two indices neutralize each other, as we now show.

First, recall that the indices on a general rank-n tensor (perhaps formed from several tensors) transform the same way an outer product of n vectors transforms, where the vector index types match those of the tensor. The vectors transform this way:

$$V'^a = R^a_b V^b \quad V'_a = S^b_a V_b$$

So, we take our same "big object" above and now ask how it transforms.

In the following, the X's represent either R or S factors for the dash indices (each of which might be up or down):

$$\begin{aligned} [\text{-----}^a\text{-----}_a\text{-----}]' &= \text{XXXXX } R^a_b \text{XXXXXXXXXXXX } S^c_a \text{XXXX } [\text{-----}^b\text{-----}_c\text{-----}] \\ &= S^c_a R^a_b \text{XXXXX } \text{XXXXXXXXXXXX } \text{XXXX } [\text{-----}^b\text{-----}_c\text{-----}] \\ &= \delta^c_b \text{XXXXX } \text{XXXXXXXXXXXX } \text{XXXX } [\text{-----}^b\text{-----}_c\text{-----}] \\ &= \text{XXXXX } \text{XXXXXXXXXXXX } \text{XXXX } [\text{-----}^a\text{-----}_a\text{-----}] \end{aligned}$$

where now the only X's left are for the *other* indices. Again we look at our canonical example,

$$A^a Y_a = A_a Y^a = A.Y$$

The contracted vector indices cancel each other out and the resulting object transforms as a scalar.

Here are some examples of tensor transformations with 0,1 and 2 index pairs contracted:

$$\begin{aligned} T'^{abc}_{de} &= R^a_a, R^b_b, R^c_c, S^d_d S^e_e T^{a'b'c'}_{d'e'} & // \text{ no pairs contracted} \\ T'^{abc}_{ae} &= R^b_b, R^c_c, S^e_e T^{a'b'c'}_{a'e'} & // \text{ index a contracted} \\ T'^{abc}_{ab} &= R^c_c, T^{a'b'c'}_{a'e'} & // \text{ index a and index b contracted} \\ Q' &= Q \quad \text{where } Q = A^a Y_a \text{ and } Q' = A'^a Y'_a & // \text{ index a contracted} \end{aligned}$$

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

$$D^{bc}_e \equiv T^{abc}_{ae} = \text{a mixed rank-3 tensor}$$

$$E^c = T^{abc}_{ab} = \text{a contravariant vector (rank-1 tensor)}$$

$$Q = A^a Y_a = \text{a scalar (rank-0 tensor)}$$

It is similarly possible to build larger tensors from smaller ones, for example

$$Z^{abc}_{de} = V^a W_e g^{ab} L^c$$

which goes under the same rubric "outer product" mentioned earlier.

(m) Raising and lowering indices on g

On the one hand, since g_{ab} and g^{ab} are inverses of each other (formerly $\bar{g}$ and g), one has

$$g_{ab} g^{bc} = \delta_{a,c} = \delta_a^c$$

where the above-mentioned "look-nice" form of $\delta_{a,c}$ makes indices match. On the other hand,

$$g_{ab} g^{bc} = g_b^c \quad // \text{ left } g \text{ lowers the left index of the right } g, \text{ or the converse}$$

Comparison shows that

$$g_b^c = \delta_b^c$$

As a sanity check, consider

$$V_a = g_{ab} V^b$$

Applying our Contraction Tilt-reversal Rule, this can be written

$$V_a = g_a^b V_b \quad \text{but this is } = \delta_a^b V_b = V_a$$

There are many ways to write things, here is a collection ($g_{ab} = g_{ba}$!)

$$g_{ab} g^{bc} = \delta_a^c = g_a^c \quad // \text{ start out}$$

$$g_a^b g_b^c = \delta_a^c = g_a^c \quad // \text{ tilt reversal of the line above; just says } \delta_a^b \delta_b^c = \delta_a^c$$

$$g^{ab} g_{bc} = \delta^a_c = g^a_c$$

$$g^a_b g^b_c = \delta^a_c = g^a_c$$

(n) Other forms of R

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

$$\begin{aligned} R^a_b &= (\partial x'^a / \partial x^b) && // \text{original object (formerly } R_{ab}) \\ R^{ab} = R^a_b \cdot g^{b'b} &= (\partial x'^a / \partial x_b) && // g \text{ pulls up the second index of } R^a_b \\ R_{ab} = g'_{aa'} R^{a'}_b &= (\partial x'_a / \partial x^b) && // g' \text{ pulls down the first index of } R^a_b \\ R_a^b = g'_{aa'} R^{a'}_b \cdot g^{b'b} &= (\partial x'_a / \partial x_b) && // \text{both actions at once} \end{aligned}$$

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.

(o) Summary of facts about R

$$\begin{aligned} R_{ik} &\equiv (\partial x'_i / \partial x_k) && \rightarrow && R^i_k \equiv (\partial x'^i / \partial x^k) \\ V'_a = R_{ab} V_b &&& \rightarrow && V'^a = R^a_b V^b && V^a = S^a_b V'^b [= R_b^a V'^b] \end{aligned}$$

$$\begin{aligned} R^a_b(x) &= (\partial x'^a / \partial x^b) && // \text{original object (formerly } R_{ab}) \\ R^{ab} = R^a_b \cdot g^{b'b} &= (\partial x'^a / \partial x_b) && // g \text{ pulls up the second index} \\ R_{ab} = g'_{aa'} R^{a'}_b &= (\partial x'_a / \partial x^b) && // g' \text{ pulls down the first index} \\ R_a^b = g'_{aa'} R^{a'}_b \cdot g^{b'b} &= (\partial x'_a / \partial x_b) && // \text{both actions at once} \end{aligned}$$

$$R^a_b = g'^{aa'} R_{a'}^{b'} g_{b'b} \quad // \text{the inverse of the previous line (using } g^{ab} g_{bc} = \delta^a_c \text{ twice)}$$

(p) Repeat all the above for S

$$\begin{aligned} S_{ik} &\equiv (\partial x_i / \partial x'^k) && \rightarrow && S^i_k \equiv (\partial x^i / \partial x'^k) \\ \bar{V}'_a = S^T_{ab} \bar{V}_b = S_{ba} \bar{V}_b &&& \rightarrow && V'_a = S^b_a V_b && V_a = R^b_a V'_b [= S_a^b V'_b] \end{aligned}$$

$$\begin{aligned} S^a_b(x) &= (\partial x^a / \partial x'^b) && // \text{original object (formerly } S_{ab}) \\ S^{ab} \equiv S^a_b \cdot g'^{b'b} &= (\partial x^a / \partial x'_b) && // g' \text{ pulls the second index up} \\ S_{ab} \equiv g_{aa'} S^{a'}_b &= (\partial x_a / \partial x'^b) && // g \text{ pulls the first index down} \\ S_a^b \equiv g_{aa'} S^{a'}_b \cdot g'^{b'b} &= (\partial x_a / \partial x'_b) && // \text{both actions at once} \end{aligned}$$

$$S^a_b \equiv g^{aa'} S_{a'}^{b'} g'_{b'b} \quad // \text{the inverse of the previous line (using } g^{ab} g_{bc} = \delta^a_c \text{ twice)}$$

(q) Theorem: $S^a_b = R_b^a$ and $S_a^b = R_a^b$ (reflect indices in vertical line between them)

This theorem has already been proven in section (i) Theorem 3 as part of the discussion there of the fact that, when tilted matrix forms are consider for R and S, one has all of the following matrix results:

$$\begin{array}{llll} S^{-1} = S^T & SS^T = 1 & S^T S = 1 & S = R^T \\ R^{-1} = R^T & RR^T = 1 & R^T R = 1 & R = S^T \end{array} \quad \begin{array}{ll} S = R^{-1} \\ R = S^{-1} \end{array}$$

Here two slightly different lower-level proofs of this theorem that $S^a_b = R_b^a$. Once this is established, one can raise and lower indices on either side to get all of the following

$$S^a_b = R_b^a \quad S_a^b = R_a^b \quad S^{ab} = R^{ba} \quad S_{ab} = R_{ba}$$

As noted earlier, an implication is that one can completely eliminate references to matrix S in tensor analysis and that is what is usually done!

Proof of Theorem: This proof is a bit long-winded, but brings in many earlier results:

$$\begin{array}{ll} \delta_b^{a''} S^a_{a''} = S^a_b & // \text{ introduce a } \delta . \text{ Remember all g's are symmetric.} \\ (g'_{bb'}, g'^{a''b'}) S^a_{a''} = S^a_b & // \text{ since } g'_{ab} \text{ and } g'^{ab} \text{ are inverses of each other.} \\ g'_{bb'} \delta^{b'}_{b''} S^a_{a''} g'^{a''b''} = S^a_b & // \text{ reorder and introduce another } \delta \\ g'_{bb'} (R^{b'}_{a'} S^{a'}_{b''}) S^a_{a''} g'^{a''b''} = S^a_b & // 1 = RS \text{ so } \delta^{b'}_{b''} = (R^{b'}_{a'} S^{a'}_{b''}) \\ g'_{bb'} R^{b'}_{a'} (S^a_{a''} S^{a'}_{b''} g'^{a''b''}) = S^a_b & // \text{ regroup} \\ g'_{bb'} R^{b'}_{a'} (g^{aa'}) = S^a_b & // \text{ use } g^{aa'} = S^a_{a''} S^{a'}_{b''} g'^{a''b''}, \text{ see end of (i) above} \\ (g'_{bb'} R^{b'}_{a'} g^{aa'}) = S^a_b & // \text{ regroup} \\ R_b^a = S^a_b & // g \text{ and } g' \text{ raise and lower R's indices, see (n) above} \end{array}$$

Notice that the above theorem says

$$S^a_b = (\partial x^a / \partial x'^b) = (\partial x'_b / \partial x_a) = R_b^a$$

A faster way to derive this result is to differentiate $dx^c dx_c = dx'^c dx'_c$ and use the chain rule:

$$\begin{aligned} dx'_b &= (\partial (dx'^c dx'_c) / \partial x'^b) = (\partial (dx^c dx_c) / \partial x^a) (\partial x^a / \partial x'^b) = dx_a (\partial x^a / \partial x'^b) \\ \Rightarrow (\partial x'_b / \partial x_a) &= (\partial x^a / \partial x'^b) \quad \Rightarrow \quad R_b^a = S^a_b \end{aligned}$$

Similar results can be derived for other index positions (or we can just raise and lower indices!) to get

$$\begin{aligned}
 S^a_b &= R^a_b &= (\partial x^a / \partial x'^b) &= (\partial x'_b / \partial x_a) \\
 S_{ab} &= R_{ba} &= (\partial x_a / \partial x'^b) &= (\partial x'_b / \partial x^a) \\
 S^{ab} &= R^{ba} &= (\partial x^a / \partial x'^b) &= (\partial x'^b / \partial x_a) \\
 S_a^b &= R^b_a &= (\partial x_a / \partial x'^b) &= (\partial x'^b / \partial x^a)
 \end{aligned}$$

Here index a is always in x-space, while index b is in x'-space.

The two vector transformation rules

$$V'^a = R^a_b V^b \quad V'_a = S^b_a V_b$$

can now be written

$$V'^a = R^a_b V^b \quad V'_a = R_a^b V_b$$

which has the advantage that the indices are properly arranged for matrix multiplication in both cases. Here then is a restatement of the transformation of the example given in section (ℓ) ,

$$T'^{abc}_{de} = R^a_a \cdot R^b_b \cdot R^c_c \cdot S^{d'}_d S^{e'}_e T^{a'b'c'}_{d'e'} \quad // \text{ no pairs contracted}$$

becomes

$$T'^{abc}_{de} = R^a_a \cdot R^b_b \cdot R^c_c \cdot R^{d'}_d R^{e'}_e T^{a'b'c'}_{d'e'} \quad // \text{ no pairs contracted}$$

It is easy to remember since the second index is always the summed index and the other index has to match (up or down) the left side of the equation.

(r) Orthogonality Rules

The above theorem $S^a_b = R^a_b$ can be used to eliminate S in various forms of $RS = 1$:

$SR = 1$	$S^a_b R^b_c = \delta^a_c$	$R^a_b R^b_c = \delta^a_c$	$R^a_b R^b_c = \delta^a_c$	$\Sigma \text{ 1st}$
$R^T S^T = 1$	$R^b_a S^c_b = \delta^c_a$	$R^b_a R^c_b = \delta^c_a$	$R^b_a R^c_b = \delta^c_a$	$\Sigma \text{ 1st}$
$RS = 1$	$R^a_b S^b_c = \delta^a_c$	$R^a_b R^c_b = \delta^a_c$	$R^c_b R^a_b = \delta^c_a$	$\Sigma \text{ 2nd}$
$S^T R^T = 1$	$S^b_a R^c_b = \delta^c_a$	$R^a_b R^c_b = \delta^c_a$	$R^c_b R^a_b = \delta^c_a$	$\Sigma \text{ 2nd}$

The four results in the right column are called **orthogonality rules** for R. The first pair is summed on the first index, the second on the second. In section (i) it was shown that these rules are just statements of the fact that in up or down tilted standard notation R is a real-orthogonal matrix so $RR^T = R^T R = 1$.

(s) The tangent and reciprocal base vectors and expansions on same

Tangent and reciprocal base vectors

Here are some basic translations:

$$\begin{aligned}
 (\mathbf{e}_n)_i &\rightarrow (\mathbf{e}_n)^i && // \text{contravariant index } i \\
 (\bar{\mathbf{e}}_n)_i &\rightarrow (\mathbf{e}_n)_i && // \text{covariant index } i \\
 (\mathbf{E}_n)_i &\rightarrow (\mathbf{e}^n)^i && // \text{contravariant index } i \\
 (\bar{\mathbf{E}}_n)_i &\rightarrow (\mathbf{e}^n)_i && // \text{covariant index } i \\
 (\mathbf{e}_n)_i = S_{in} &\rightarrow (\mathbf{e}_n)^i = S^i_n = R_n^i && // \text{contravariant index } i \\
 (\bar{\mathbf{E}}_n)_i = R_{ni} &\rightarrow (\mathbf{e}^n)_i = R^n_i && // \text{covariant index } i \\
 (\mathbf{E}_n)_i = R_{nk}g_{ki} &\rightarrow (\mathbf{e}^n)^i = R^n_k g^{ki} = R^{ni} && // \text{contravariant index } i
 \end{aligned}$$

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. If one form or the other is intended, one must show an index up or down, even if it is just a dummy placeholder index. As examples,

$$\begin{aligned}
 S = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \dots \mathbf{e}_N] &\rightarrow S^i_j = [(\mathbf{e}_1)^k, (\mathbf{e}_2)^k, (\mathbf{e}_3)^k \dots (\mathbf{e}_N)^k] \\
 R = [\bar{\mathbf{E}}_1, \bar{\mathbf{E}}_2, \bar{\mathbf{E}}_3 \dots \bar{\mathbf{E}}_N]^T &\rightarrow R^i_j = [(\mathbf{e}^1)_k, (\mathbf{e}^2)_k, (\mathbf{e}^3)_k \dots (\mathbf{e}^N)_k]^T
 \end{aligned}$$

The relationship between $\mathbf{e}^n$ and $\mathbf{e}_n$ is very simple,

$$\mathbf{E}_n \equiv g'_{ni} \mathbf{e}_i \rightarrow \mathbf{e}^n = g^{ni} \mathbf{e}_i \quad \text{and} \quad \mathbf{e}_n = g'_{ni} \mathbf{e}^i$$

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.

The dot products are

$$\begin{aligned}
 \mathbf{e}_n \bullet \mathbf{e}_m &= \bar{g}'_{nm} \rightarrow \mathbf{e}_n \bullet \mathbf{e}_m = g'_{nm} \\
 \mathbf{E}_n \bullet \mathbf{e}_m &= \delta_{n,m} \rightarrow \mathbf{e}^n \bullet \mathbf{e}_m = \delta^n_m \\
 \mathbf{E}_n \bullet \mathbf{E}_m &= g'_{nm} \rightarrow \mathbf{e}^n \bullet \mathbf{e}^m = g'^{nm}
 \end{aligned}$$

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, $\mathbf{e}^n \bullet \mathbf{e}_m = g'^n_m = \delta^n_m$.

Inverse tangent and reciprocal base vectors

Using the rules given above,

$$g' \leftrightarrow g \quad R \leftrightarrow S \quad \mathbf{e}_n \rightarrow \mathbf{u}'_n \quad \mathbf{e}'_n \rightarrow \mathbf{u}_n \quad \mathbf{E}_n \rightarrow \mathbf{U}'_n \quad \mathbf{E}'_n \rightarrow \mathbf{U}_n$$

we can obtain the corresponding results for the inverse tangent and reciprocal base vectors:

$$(\mathbf{u}'_n)_i \rightarrow (\mathbf{u}_n)^i \quad // \text{ contravariant index } i$$

$$(\bar{\mathbf{u}}'_n)_i \rightarrow (\mathbf{u}_n)_i \quad // \text{ covariant index } i$$

$$(\mathbf{U}'_n)_i \rightarrow (\mathbf{u}^n)^i \quad // \text{ contravariant index } i$$

$$(\bar{\mathbf{U}}'_n)_i \rightarrow (\mathbf{u}^n)_i \quad // \text{ covariant index } i$$

$$(\mathbf{u}'_n)_i = R_{in} \rightarrow (\mathbf{u}'_n)^i = R^i_n = S_n^i \quad // \text{ contravariant index } i$$

$$(\bar{\mathbf{U}}'_n)_i = S_{ni} \rightarrow (\mathbf{u}^n)_i = S^n_i \quad // \text{ covariant index } i$$

$$(\mathbf{U}'_n)_i = S_{nk} g'_{ki} \rightarrow (\mathbf{u}^n)^i = S^n_k g'^{ki} = S^{ni} \quad // \text{ contravariant index } i$$

$$R = [\mathbf{u}'_1, \mathbf{u}'_2, \mathbf{u}'_3 \dots \mathbf{u}'_N] \rightarrow R^i_j = [(\mathbf{u}'_1)^k, (\mathbf{u}'_2)^k, (\mathbf{u}'_3)^k \dots (\mathbf{u}'_N)^k]$$

$$S = [\bar{\mathbf{U}}'_1, \bar{\mathbf{U}}'_2, \bar{\mathbf{U}}'_3 \dots \bar{\mathbf{U}}'_N]^T \rightarrow S^i_j = [(\bar{\mathbf{U}}'^1)_k, (\bar{\mathbf{U}}'^2)_k, (\bar{\mathbf{U}}'^3)_k \dots (\bar{\mathbf{U}}'^N)_k]^T$$

$$\mathbf{U}'_n \equiv g_{ni} \mathbf{u}'_n \rightarrow \mathbf{u}^n = g^{ni} \mathbf{u}'_i \quad \text{and} \quad \mathbf{u}'_n = g_{ni} \mathbf{u}^i$$

$$\mathbf{u}'_n \bullet \mathbf{u}'_m = \bar{g}_{nm} \rightarrow \mathbf{u}'_n \bullet \mathbf{u}'_m = g_{nm}$$

$$\mathbf{U}'_n \bullet \mathbf{u}'_m = \delta_{n,m} \rightarrow \mathbf{u}^n \bullet \mathbf{u}'_m = \delta^n_m$$

$$\mathbf{U}'_n \bullet \mathbf{U}'_m = g_{nm} \rightarrow \mathbf{u}^n \bullet \mathbf{u}^m = g^{nm}$$

Summary table.

The summary table given at the end of Section 6 (e) was this

	<u>x'-space</u>	<u>x-space</u>		
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	$\mathbf{U}'_n$	$\mathbf{E}_n$	$(\mathbf{U}'_n)_i = g'_{ia} S_{na}$	$(\mathbf{E}_n)_i = g_{ia} R_{na}$

$$= g_{na} R_{ia} \quad = g'_{na} S_{ia}$$

which translates into this $\rightarrow$:

	<u>x'-space</u>	<u>x-space</u>		
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	$\mathbf{u}'_n$	$\mathbf{e}_n$	$(\mathbf{u}'_n)^i = R^i_n$	$(\mathbf{e}_n)^i = S^i_n$
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 Section 6 (f) were

$$\begin{aligned}
 \mathbf{V} &= V_1 \mathbf{u}_1 + V_2 \mathbf{u}_2 + \dots &= \sum_n V_n \mathbf{u}_n & \text{where } \mathbf{U}_n \bullet \mathbf{V} = V_n & \mathbf{U}_n = g_{ni} \mathbf{u}_i \\
 \mathbf{V} &= \bar{V}_1 \mathbf{U}_1 + \bar{V}_2 \mathbf{U}_2 + \dots &= \sum_n \bar{V}_n \mathbf{U}_n & \text{where } \mathbf{u}_n \bullet \mathbf{V} = \bar{V}_n & \\
 \mathbf{V} &= V'_1 \mathbf{e}_1 + V'_2 \mathbf{e}_2 + \dots &= \sum_n V'_n \mathbf{e}_n & \text{where } \mathbf{E}_n \bullet \mathbf{V} = V'_n & \mathbf{E}_n = g'_{ni} \mathbf{e}_i \\
 \mathbf{V} &= \bar{V}'_1 \mathbf{E}_1 + \bar{V}'_2 \mathbf{E}_2 + \dots &= \sum_n \bar{V}'_n \mathbf{E}_n & \text{where } \mathbf{e}_n \bullet \mathbf{V} = \bar{V}'_n &
 \end{aligned}$$

and they now become $\rightarrow$:

$$\begin{aligned}
 \mathbf{V} &= V^1 \mathbf{u}_1 + V^2 \mathbf{u}_2 + \dots &= \sum_n V^n \mathbf{u}_n & \text{where } \mathbf{u}^n \bullet \mathbf{V} = V^n & \mathbf{u}^n = g^{ni} \mathbf{u}_i \\
 \mathbf{V} &= V_1 \mathbf{u}^1 + V_2 \mathbf{u}^2 + \dots &= \sum_n V_n \mathbf{u}^n & \text{where } \mathbf{u}_n \bullet \mathbf{V} = V_n & \\
 \mathbf{V} &= V'^1 \mathbf{e}_1 + V'^2 \mathbf{e}_2 + \dots &= \sum_n V'^n \mathbf{e}_n & \text{where } \mathbf{e}^n \bullet \mathbf{V} = V'^n & \mathbf{e}^n = g'^{ni} \mathbf{e}_i \\
 \mathbf{V} &= V'_1 \mathbf{e}^1 + V'_2 \mathbf{e}^2 + \dots &= \sum_n V'_n \mathbf{e}^n & \text{where } \mathbf{e}_n \bullet \mathbf{V} = V'_n &
 \end{aligned}$$

x'-space expansions

Similarly, the x'-space expansions of Section 6 (f) were

$$\begin{aligned}
 \mathbf{V}' &= V'_1 \mathbf{e}'_1 + V'_2 \mathbf{e}'_2 + \dots &= \sum_n V'_n \mathbf{e}'_n & \text{where } \mathbf{E}'_n \bullet \mathbf{V}' = V'_n & \mathbf{E}'_n = g'_{ni} \mathbf{e}'_i \\
 \mathbf{V}' &= \bar{V}'_1 \mathbf{E}'_1 + \bar{V}'_2 \mathbf{E}'_2 + \dots &= \sum_n \bar{V}'_n \mathbf{E}'_n & \text{where } \mathbf{e}'_n \bullet \mathbf{V}' = \bar{V}'_n & \\
 \mathbf{V}' &= V_1 \mathbf{u}'_1 + V_2 \mathbf{u}'_2 + \dots &= \sum_n V_n \mathbf{u}'_n & \text{where } \mathbf{U}'_n \bullet \mathbf{V}' = V_n & \mathbf{U}'_n = g_{ni} \mathbf{u}'_i \\
 \mathbf{V}' &= \bar{V}_1 \mathbf{U}'_1 + \bar{V}_2 \mathbf{U}'_2 + \dots &= \sum_n \bar{V}_n \mathbf{U}'_n & \text{where } \mathbf{u}'_n \bullet \mathbf{V}' = \bar{V}_n &
 \end{aligned}$$

and they now become $\rightarrow$:

$$\begin{aligned}
 \mathbf{V}' &= V'^1 \mathbf{e}'_1 + V'^2 \mathbf{e}'_2 + \dots &= \sum_n V'^n \mathbf{e}'_n & \text{where } \mathbf{e}'^n \bullet \mathbf{V}' = V'^n & \mathbf{e}'^n = g'^{ni} \mathbf{e}'_i \\
 \mathbf{V}' &= V'_1 \mathbf{e}'^1 + V'_2 \mathbf{e}'^2 + \dots &= \sum_n V'_n \mathbf{e}'^n & \text{where } \mathbf{e}'_n \bullet \mathbf{V}' = V'_n & \\
 \mathbf{V}' &= V^1 \mathbf{u}'_1 + V^2 \mathbf{u}'_2 + \dots &= \sum_n V^n \mathbf{u}'_n & \text{where } \mathbf{u}'^n \bullet \mathbf{V}' = V^n & \mathbf{u}'^n = g^{ni} \mathbf{u}'_i \\
 \mathbf{V}' &= V_1 \mathbf{u}'^1 + V_2 \mathbf{u}'^2 + \dots &= \sum_n V_n \mathbf{u}'^n & \text{where } \mathbf{u}'_n \bullet \mathbf{V}' = V_n &
 \end{aligned}$$

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 :

$$\begin{aligned} \mathbf{V} &= V^n \mathbf{u}_n = V_n \mathbf{u}^n = V'^n \mathbf{e}_n = V'_n \mathbf{e}^n = \mathcal{V}^n \hat{\mathbf{e}}_n & // \text{ x-space expansions, } \mathcal{V}^n &= h_n V^n \\ \mathbf{V}' &= V'^n \mathbf{e}'_n = V'_n \mathbf{e}^{'n} = V^n \mathbf{u}'_n = V_n \mathbf{u}^{'n} & // \text{ x'-space expansions} \end{aligned}$$

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. Corresponding expansions of higher rank tensors are presented in section (w) below.

(t) Comment on Covariant versus Contravariant

Consider this expansion for a vector $\mathbf{V}$ in x-space,

$$\mathbf{V} = V^n \mathbf{c}_n \qquad V^n = \mathbf{V} \bullet \mathbf{c}^n$$

where $\mathbf{c}_n$ is some basis having dual basis $\mathbf{c}^n$ where as usual $\mathbf{c}_n \bullet \mathbf{c}^m = \delta_n^m$. Imagine taking $V^n \rightarrow V'^n = R^n_m V^m$ and $\mathbf{c}_i \rightarrow \mathbf{c}'_i = Q_i^j \mathbf{c}_j$. What Q would cause the following to be true?

$$\mathbf{V} = V^n \mathbf{c}_n = V'^n \mathbf{c}'_n$$

In other words, how does one transform that basis $\mathbf{c}_n$ such that the vector $\mathbf{V}$ 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 as in section (r))

$$V'^n \mathbf{c}'_n = [R^n_m V^m] [R_n^j \mathbf{c}_j] = (R^n_m R_n^j) V^m \mathbf{c}_j = \delta_m^j V^m \mathbf{c}_j = V^j \mathbf{c}_j = V^n \mathbf{c}_n$$

Compare then the transformation of V^n with that of the basis $\mathbf{c}_n$:

$$\begin{aligned} V'^n &= R^n_m V^m \\ \mathbf{c}'_n &= R_n^m \mathbf{c}_m \end{aligned}$$

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 $\mathbf{V}$. 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 $\mathbf{c}_m$.

If one starts over with V_n components and the $\mathbf{c}^n$ "dual" (reciprocal) expansion vectors and asks for a solution to this corresponding problem,

$$\mathbf{V} = V_n \mathbf{c}^n = V'_n \mathbf{c}'^n$$

one finds not surprisingly that the dual basis must vary as $\mathbf{c}^{n'} = R^n_m \mathbf{c}^m$ and then one has

$$\begin{aligned} V'_{n'} &= R^n_m V_m \\ \mathbf{c}^{n'} &= R^n_m \mathbf{c}^m \end{aligned}$$

which is the previous result with all indices up $\leftrightarrow$ down. Comparing the tilts, one would say that the V_m again "contra vary" with the way the $\mathbf{c}^m$ vary to maintain invariance of $\mathbf{V}$. But one does not care about the dual basis, one cares about the *basis*, so relative to the basis $\mathbf{c}_n$ one has

$$\begin{aligned} V'_n &= R^n_m V_m \\ \mathbf{c}_n' &= R^n_m \mathbf{c}_m \end{aligned}$$

If the basis $\mathbf{c}_m$ is varied as shown here, then the dual basis $\mathbf{c}^m$ varies as shown above and $\mathbf{V}$ remains invariant. Comparing now the way the V_n transform with the way the basis vectors $\mathbf{c}_m$ transform, 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 $\mathbf{c}_m$.

(u) 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:

$$Q_a{}^d(x) = H_{ab}(x) T^b{}_c(x) B^d(x)$$

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

$$Q'_a{}^d(\mathbf{x}') = H'_{ab}(\mathbf{x}') T'^b{}_c(\mathbf{x}') B'^d(\mathbf{x}')$$

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".)

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", 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. By then applying the same inverse transformation to both sides of

the resulting equation to "clear out" the transformation factors, one is left with the initial equation but with everything now primed.

If this explanation is not convincing, a brute force demonstration can substitute. The following is also a good exercise is using the two tilt forms of the R matrix. Recall from section (q) that $S^b_a = R^b_a$ and that $SR = 1$ is replaced by the various orthogonality rules of section (r).

We shall process the primed equation into the unprimed one, being careful to give new summation indices unique names:

$$Q'^d_{a'c'}(x') = H'_{ab}(x') T'^b_c(x') B'^d(x') \quad // \text{ x'-space equation}$$

$$\begin{aligned} [R^{a'}_{a'} R^d_{d'} R^{c'}_{c'} Q_{a',d',c'}(x)] &= [R^{a'}_{a'} R^{b'}_{b'} H_{a',b'}(x)] [R^{b'}_{b'} R^{c'}_{c'} T^{b''}_{c'}(x)] [R^d_{d'} B^{d'}(x)] \\ &= R^{a'}_{a'} R^d_{d'} R^{c'}_{c'} (R^{b'}_{b'} R^{b''}_{b''}) H_{a',b'}(x) T^{b''}_{c'}(x) B^{d'}(x) \end{aligned}$$

where we have reordered the R factors to (1) match the LHS as much as possible, and (2) to group any pairs that came from a contracted index. Using one of the orthogonality rules of section (r),

$$\begin{aligned} &= R^{a'}_{a'} R^d_{d'} R^{c'}_{c'} (\delta^{b''}_{b'}) H_{a',b'}(x) T^{b''}_{c'}(x) B^{d'}(x) \\ &= R^{a'}_{a'} R^d_{d'} R^{c'}_{c'} H_{a',b'}(x) T^{b'}_{c'}(x) B^{d'}(x) \end{aligned}$$

so that

$$(R^{a'}_{a'} R^d_{d'} R^{c'}_{c'}) Q_{a',d',c'}(x) = (R^{a'}_{a'} R^d_{d'} R^{c'}_{c'}) H_{a',b'}(x) T^{b'}_{c'}(x) B^{d'}(x)$$

Now apply to both sides the factor $R^a_A R^D_D R^C_C$ and sum on a,c,d.

$$\begin{aligned} &R^a_A R^D_D R^C_C (R^{a'}_{a'} R^d_{d'} R^{c'}_{c'}) Q_{a',d',c'}(x) \\ &= R^a_A R^D_D R^C_C (R^{a'}_{a'} R^d_{d'} R^{c'}_{c'}) H_{a',b'}(x) T^{b'}_{c'}(x) B^{d'}(x) \end{aligned}$$

or

$$\begin{aligned} &(R^a_A R^{a'}_{a'}) (R^D_D R^d_{d'}) (R^C_C R^{c'}_{c'}) Q_{a',d',c'}(x) \\ &= (R^a_A R^{a'}_{a'}) (R^D_D R^d_{d'}) (R^C_C R^{c'}_{c'}) H_{a',b'}(x) T^{b'}_{c'}(x) B^{d'}(x) \end{aligned}$$

Then use a section (r) orthogonality rule in each (R R) to get

$$(\delta^a_{a'}) (\delta^D_{d'}) (\delta^C_{c'}) Q_{a',d',c'}(x) = (\delta^a_{a'}) (\delta^D_{d'}) (\delta^C_{c'}) H_{a',b'}(x) T^{b'}_{c'}(x) B^{d'}(x)$$

or

$$Q^D_{ac}(x) = H_{Ab'}(x) T^{b'}_c(x) B^D(x)$$

Finally, rename the indices A,D,C,b' to be a,d,c,b

$$Q^d_{ac}(x) = H_{ab}(x) T^b_c(x) B^d(x) \quad // \text{ x-space equation}$$

Thus it has been shown that, if all the objects transform as tensors, the equation is covariant.

Examples: 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 (m)).

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 :

$$F^{\mu\nu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu$$

where ∂^μ means $g^{\mu\alpha}\partial_\alpha$, the contravariant form of the gradient operator. The components are then

$$F^{\mu\nu} = \begin{bmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & -B_z & B_y \\ E_y/c & B_z & 0 & -B_x \\ E_z/c & -B_y & B_x & 0 \end{bmatrix}$$

where c is the speed of light and of course $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic fields. Two of Maxwell's equations are (in SI units where $\epsilon_0\mu_0 = 1/c^2$)

$$\partial_\nu F^{\mu\nu} = \mu_0 J^\mu \quad J^\mu = (c\rho, \mathbf{J})$$

while the other two are

$$\partial_\alpha F_{\mu\nu} + \partial_\mu F_{\nu\alpha} + \partial_\nu F_{\alpha\mu} = 0 \quad \text{or} \quad \partial_\alpha F_{\mu\nu} + \text{cyclic} = 0$$

One can see that each of these equations involves only tensors and we expect that in x' -space these equations will take the form

$$\begin{aligned} \partial'_\nu F'^{\mu\nu} &= \mu_0 J'^\mu & J'^\mu &= (c\rho', \mathbf{J}') \\ \partial'_\alpha F'_{\mu\nu} + \partial'_\mu F'_{\nu\alpha} + \partial'_\nu F'_{\alpha\mu} &= 0 & \text{or} \quad \partial'_\alpha F'_{\mu\nu} + \text{cyclic} &= 0 \end{aligned}$$

Objects like $\partial_\nu F^{\mu\nu}$ and $\partial_\alpha F_{\mu\nu}$ are true rank-3 tensors because the transformation F is linear. The next section shows the pain encountered when F is not linear.

(v) The Christoffel Business: covariant derivatives

When a transformation F is non-linear, the matrix R^a_b is a function of $\mathbf{x}$. 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,

$$(\partial'_a V'_b) = (R^d_a \partial_d) (R^c_b V_c) = R^d_a R^c_b (\partial_d V_c) + R^d_a (\partial_d R^c_b) V_c$$

This second term did not arise earlier when we looked at ∂_a on a *scalar* field $\phi(\mathbf{x}') = \phi(\mathbf{x})$,

$$(\partial'_a \phi) = (R_a^d \partial_d) \phi = R_a^d (\partial_d \phi)$$

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,

$$(\partial'_a V'_b) = R_a^d R_b^c (\partial_d V_c) \quad , \quad // F \text{ is a linear transformation}$$

but in the general case the second term is present, so $\partial_a V_c$ fails to transform as a rank-2 covariant tensor. In this case, one defines a certain "covariant derivative" which itself has an extra piece

$$V_{b;a} \equiv \partial_a V_b - \Gamma_{ab}^c V_c$$

where Γ_{ab}^c is a certain function of the metric tensor g . One then finds that

$$V'_{b;a} = R_a^d R_b^c V_{d;c}$$

or

$$[\partial'_a V'_b - \Gamma'^{c'}_{a'b'} V'_{c'}] = R_a^d R_b^c [\partial_d V_c - \Gamma_{ab}^c V_c] \quad (*)$$

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 Γ_{ab}^c (the "Christoffel connection") is given by

$$\Gamma_{ab}^c \equiv \{ab, c\} \equiv \left\{ \begin{matrix} c \\ ab \end{matrix} \right\} \equiv g^{cd} [ab, d] = \frac{1}{2} g^{cd} (\partial_a g_{bd} + \partial_b g_{ad} - \partial_d g_{ab}) \quad // \text{Christoffel 2nd kind}$$

$$\Gamma_{dab} \equiv [ab, d] \equiv \frac{1}{2} (\partial_a g_{bd} + \partial_b g_{ad} - \partial_d g_{ab}) \quad // \text{Christoffel 1st kind}$$

and this is where the various "Christoffel symbols" come into play. In general relativity, Γ_{ab}^c is also 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(n).

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,

$$\partial'_a g'_{bd}(\mathbf{x}') = (R_a^d \partial_d)(R_b^{b'} R_d^{d'} g_{b'd'}) = R_b^{b'} R_d^{d'} (\partial_d g_{b'd'}) + \text{other terms}$$

and therefore neither of the Christoffel symbols Γ_{dab} or Γ_{ab}^c transforms as a tensor in this case.

Verifying the claim (*) takes a bit of algebra which is left to the reader (see Weinberg p 103). There are various other notations for the covariant derivative, such as $\nabla_a B_b$ and $B_{b;a}$.

(w) 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 (s) above. There it was shown that a vector $\mathbf{A}$ can be expanded in two interesting ways :

$$\begin{aligned}\mathbf{A} &= \sum_n \mathbf{A}^n \mathbf{u}_n & \mathbf{A}^n \text{ are the contravariant components of } \mathbf{A} \text{ in } x\text{-space} \\ \mathbf{A} &= \sum_n \mathbf{A}'^n \mathbf{e}_n & \mathbf{A}'^n \text{ are the contravariant components of } \mathbf{A} \text{ in } x'\text{-space}\end{aligned}$$

In the first, $\mathbf{u}_n$ are axis aligned basis vectors, and in the second $\mathbf{e}_n$ are the tangent base vectors. If $\mathbf{A}$ were instead a tensor of rank n , these expansions would be replaced by

$$\begin{aligned}A &= \sum_{ijk\dots} A^{ijk\dots} (\mathbf{u}_i \otimes \mathbf{u}_j \otimes \mathbf{u}_k \dots) & A^{ijk\dots} \text{ are the contravariant components of } A \text{ in } x\text{-space} \\ A &= \sum_{ijk\dots} A'^{ijk\dots} (\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \dots) & A'^{ijk\dots} \text{ are the contravariant components of } A \text{ in } x'\text{-space}\end{aligned}$$

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

$$\begin{aligned}A &= \sum_{ijk\dots} A^{ijk\dots} \mathbf{u}_i \mathbf{u}_j \mathbf{u}_k \dots & A^{ijk\dots} \text{ are the contravariant components of } A \text{ in } x\text{-space} \\ A &= \sum_{ijk\dots} A'^{ijk\dots} \mathbf{e}_i \mathbf{e}_j \mathbf{e}_k \dots & A'^{ijk\dots} \text{ are the contravariant components of } A \text{ in } x'\text{-space} .\end{aligned}$$

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 App. E). In this case (only) the product can be visualized as $\mathbf{u}_i \mathbf{u}_j^T$ which is a matrix constructed from a column vector to the left of a row vector. Thus one can write

$$\begin{aligned}A &= \sum_{ij} A^{ij} \mathbf{u}_i \mathbf{u}_j^T & A^{ij} \text{ are the contravariant components of } A \text{ in } x\text{-space} \\ A &= \sum_{ij} A'^{ij} \mathbf{e}_i \mathbf{e}_j^T & A'^{ij} \text{ are the contravariant components of } A \text{ in } x'\text{-space} .\end{aligned}$$

Appendix E encourages 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,

$$\begin{aligned}A^{nm} &= \langle \mathbf{u}^n | A | \mathbf{u}^m \rangle = \text{the } x\text{-space components of tensor } A \\ A'^{nm} &= \langle \mathbf{e}^n | A | \mathbf{e}^m \rangle = \text{the } x'\text{-space components of tensor } A\end{aligned}$$

As Appendix E shows, these two matrices are related to each other by a similarity transformation $A' = R A R^{-1}$.

Appendix F applies the notation shown above to the problem of writing the dyadic product $(\nabla \mathbf{v})$ in arbitrary curvilinear coordinates. One can think of $(\nabla \mathbf{v})$ as a matrix which acts on some vector to its right. In Cartesian coordinates one has $(\nabla \mathbf{v})_{ij} \equiv (\partial_j v_i)$. Since $(\nabla \mathbf{v})$ is not a differential operator acting on whatever vector lies to its right, we did not include an analysis of $(\nabla \mathbf{v})$ in Section 10 on the gradient.

The main result of Appendix F is this: (x-space is Cartesian, x'-space is that of the curvilinear coordinates):

$$(\nabla \mathbf{v}) = \Sigma_{ij} [(\nabla \mathbf{v})^{(u)}]^{ij} \mathbf{u}_i \mathbf{u}_j^T \quad [(\nabla \mathbf{v})^{(u)}]^{ij} = (\partial^j v^i)$$

$$(\nabla \mathbf{v}) = \Sigma_{ij} [(\nabla \mathbf{v})^{(e)}]^{ij} \mathbf{e}_i \mathbf{e}_j^T \quad [(\nabla \mathbf{v})^{(e)}]^{ij} = \Sigma_{ab} g'^{ja} [g'^{ib} (\partial'_a v'_b) + \Sigma_d R^i_d (\partial'_a R^b_d) v'_b]$$

Notice that the expression for the second matrix $[(\nabla \mathbf{v})^{(e)}]^{ij}$ contains only x'-space references ($R(\mathbf{x}')$, $v'_b(\mathbf{x}')$, $g'(\mathbf{x}')$, ∂'). Again, one can think of $(\nabla \mathbf{v})$ as an operator in a Hilbert Space and the two matrices shown above are then associated with this operator in the $\mathbf{u}_i$ and $\mathbf{e}_i$ bases,

$$[(\nabla \mathbf{v})^{(u)}]^{ij} = \langle \mathbf{u}^n | (\nabla \mathbf{v}) | \mathbf{u}^m \rangle$$

$$[(\nabla \mathbf{v})^{(e)}]^{ij} = \langle \mathbf{e}^n | (\nabla \mathbf{v}) | \mathbf{e}^m \rangle$$

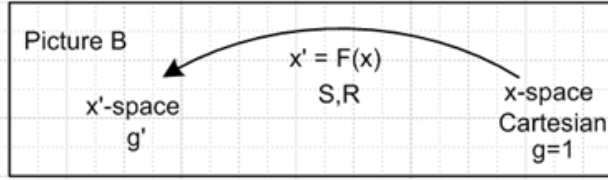
The object $(\nabla \mathbf{v})$ looks like a rank-2 tensor but of course it is not a rank-2 tensor for the reason discussed in section (v). The complicated second term shown above in $[(\nabla \mathbf{v})^{(e)}]^{ij}$ is similar to that "second term" which appears in the expansion of $(\partial'_a V'_b)$ at the start of section (v). Since $(\nabla \mathbf{v})$ is not a tensor, the coefficients shown in the above expansions cannot be interpreted as the x-space and x'-space components of a tensor. Nevertheless, being able to express $(\nabla \mathbf{v})$ in terms of curvilinear coordinates finds great usefulness in continuum mechanics.

8. Transformation of Differential Area, Volume and Length

This Section and all remaining Sections use the Standard Notation introduced in Section 7.

The term N-piped is short for N dimensional parallelepiped.

The context is Picture B:



Since this Section is lengthy, a brief overview is in order:

Overview

The transformation of differential length, area, and volume is framed in terms of the mapping of a differential orthogonal N-piped in x'-space to a likely rotated and possibly 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 $e_n dx'^n$ where e_n are the tangent base vectors introduced in Section 3. We want to know what happens to the volume, edges and face areas as one N-piped is mapped into the other by an inverse curvilinear coordinates transformation $x = F^{-1}(x')$.

After defining the two N-pipeds of interest in section (a), the focus is on the x-space N-piped. In section (b) results are quoted from Appendix B which describe the geometry of finite parallelepipeds in N dimensions. In Section (c) these results are adapted to the differential x-space N-piped and expressions are obtained for the three x-space objects $dx^{(n)}$, dA^n and dV (edge, face area, volume).

Section (d) addresses the tensor density weights of the area vector dA^n and the volume scalar dV . These weights are found to be -1 causing these two objects to transform in a non-tensorial manner.

Section (e) finds expressions for the corresponding x'-space objects $dx'^{(n)}$, dA'^n and dV' .

Section (f) evaluates the covariant magnitudes of the x-space edge and area vectors, and then provides Examples of area magnitude transformations.

Section (g) evaluates the corresponding x'-space edge and area vector magnitudes, and then Section (h) summarizes all the results in two simple tables followed by a set of interpretative comments.

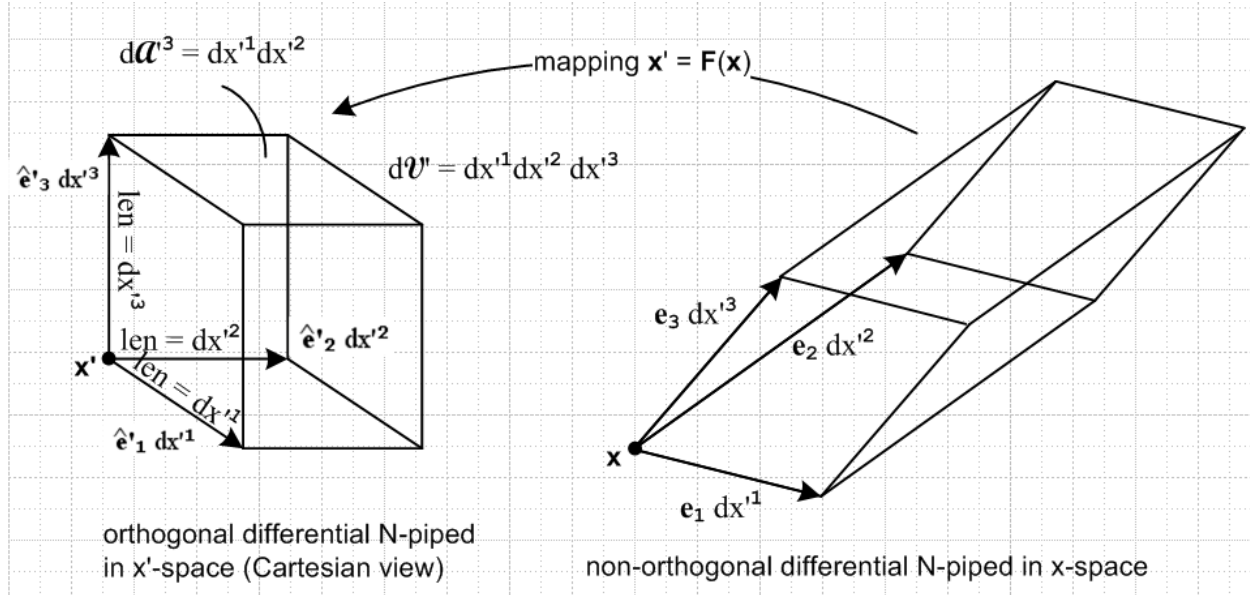
Section (i) derives the Jacobian Integration Rule as a distribution, and finally Section (j) gives a list of different interpretations of the Jacobian.

(a) The differential N-piped mapping

Section 3 (a) above considered the following situation ($dx'^n > 0$) :

$dx'^{(n)} = e'_n dx'^n$	x'-space n-axis-aligned differential vector, $(e'_n)^i = \delta_n^i$
$dx^{(n)} = e_n dx'^n$	x-space mapping of the above vector under F^{-1} or R^{-1}
$dx'^{(n)} = R(x) dx^{(n)}$	relation of the two differential vectors (contravariant rule)

A superscript (n) on the differentials makes clear there is no implied sum on n. The vectors $dx^{(n)}$ span a differential N-piped in x-space, while the $dx'^{(n)}$ span a corresponding differential N-piped in x'-space, and the two N-pipeds are related by the mapping $x' = F(x)$. In Appendix C (e) we discuss the somewhat confusing issue of the "two views" of x'-space, one Cartesian ($g'=1$) and one Curvilinear ($g'=g'$). The N-piped in x'-space just mentioned is considered in this Cartesian view of x'-space in which the basis vectors e'_n [$e'_1 = (1,0,0..)$, etc.] are orthonormal Cartesian vectors. Here is a depiction of the mapping for N=3:



The true Curvilinear-view x'-space N-piped has axes e'_n and not $\hat{e}'_n$ and $g' \neq 1$. The false Cartesian view N-piped merely allows visualization of these curvilinear coordinate variations, all $dx'^k > 0$,

$$\begin{aligned} d\mathcal{L}^n &\equiv dx'^n \\ d\mathcal{A}^n &\equiv \prod_{i \neq n} dx'^i \\ d\mathcal{V}'' &\equiv \prod_i dx'^i = d\mathcal{A}^n d\mathcal{L}^n \end{aligned}$$

For example, for N=3 one would have

$$\begin{aligned} d\mathcal{L}^{i1} &\equiv dx'^1 \\ d\mathcal{A}^{i3} &= dx'^1 dx'^2 & d\mathcal{A}^{i1} &= dx'^2 dx'^3 & d\mathcal{A}^{i2} &= dx'^3 dx'^1 \\ d\mathcal{V}'' &= dx'^1 dx'^2 dx'^3 \end{aligned}$$

The Cartesian-view x'-space N-piped is always orthogonal since its spanning vectors $\hat{e}'_n$ are orthogonal. In contrast, the x-space N-piped is typically rotated and possibly skewed as well (if the coordinates x'^1 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 $e_n(x)$ are functions of x .

(b) 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).
- There are 2^N vertices.
- 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).
- 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.
- 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.
- 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.
- The vector area $\mathbf{A}^n$ is given by several equivalent expressions:

$$\mathbf{A}^n = |\det(\mathbf{S}^a_b)| \mathbf{e}^n$$

$$\mathbf{A}^n = \sigma (-1)^{n-1} \prod_{i \neq n}^x \mathbf{e}_i$$

$$\mathbf{A}^n = \sigma (-1)^{n-1} \mathbf{e}_1 \times \mathbf{e}_2 \dots \times \mathbf{e}_N \quad // \mathbf{e}_n \text{ missing} \quad \sigma \equiv \text{sign}[\det(\mathbf{S}^a_b)] = \text{sign}[\det(\mathbf{R}^a_b)]$$

$$(\mathbf{A}^n)_i = \sigma (-1)^{n-1} \epsilon_{iabc\dots x} (e_1)^a (e_2)^b \dots (e_N)^x \quad // \mathbf{e}_n \text{ missing}$$

- The volume of the N-piped is given by (see Section 5 (k) concerning J)

$$V = |\det [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \dots \mathbf{e}_N]| = |\det(\mathbf{S}^a_b)| = g^{1/2} = |J|$$

(c) Objects in x-space

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 the properties listed above with the following adjustments:

$$d\mathbf{x}^{(n)} = \mathbf{e}_n dx'_n \quad // \text{an edge vector, as in section (a) above}$$

$$d\mathbf{A}^n = |\det(\mathbf{S}^a_b)| \mathbf{e}^n \prod_{i \neq n} dx'^i = g^{1/2} \mathbf{e}^n \prod_{i \neq n} dx'^i \quad // \text{label on } d\mathbf{A}^n \text{ is placed up to match that on } \mathbf{e}^n$$

$$d\mathbf{A}^n = \sigma (-1)^{n-1} \Pi_{i \neq n}^{\mathbf{x}} (\mathbf{e}_i dx'^i) = \sigma (-1)^{n-1} \Pi_{i \neq n}^{\mathbf{x}} \mathbf{e}_i \Pi_{i \neq n} dx'^i$$

$$d\mathbf{A}^n = \sigma (-1)^{n-1} \mathbf{e}_1 \times \mathbf{e}_2 \dots \times \mathbf{e}_N \Pi_{i \neq n} dx'^i \quad // \mathbf{e}_n \text{ missing} \quad \sigma \equiv \text{sign}[\det(\mathbf{S}^{\mathbf{a}}_{\mathbf{b}})] = \text{sign}[\det(\mathbf{R}^{\mathbf{a}}_{\mathbf{b}})]$$

$$dV = |\det [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \dots \mathbf{e}_N]| \Pi_i dx'^i = |\det(\mathbf{S}^{\mathbf{a}}_{\mathbf{b}})| \Pi_i dx'^i = g'^{1/2} \Pi_i dx'^i$$

Notice, for example, that scaling the edges of an N-piped by a set of N numbers does not alter the direction of the faces' area normal vectors. A suitable normal for face pair n can always be taken as $\mathbf{e}^n$. Using the curvilinear coordinate variation definitions from section (a), these results may be restated as

$$\begin{aligned} d\mathbf{x}^{(n)} &= \mathbf{e}_n d\mathcal{L}^n & d\mathcal{L}'^1 &\equiv dx'^1 \\ d\mathbf{A}^n &= \sigma (-1)^{n-1} \Pi_{i \neq n}^{\mathbf{x}} \mathbf{e}_i d\mathcal{A}^n & d\mathcal{A}^n &\equiv \Pi_{i \neq n} dx'^i \\ dV &= g'^{1/2} d\mathcal{V} & d\mathcal{V} &\equiv \Pi_i dx'^i \end{aligned}$$

These three items are the x-space objects of interest.

(d) Tensor density weights of objects

Appendix D discusses the notion of tensor densities and their weights and lists various properties of and theorems about tensor densities. Weights are additive when smaller tensors are assembled to form larger ones, and ordinary tensors, such as tensorial vectors, have weight 0. In Appendix D (f) it is noted that the covariant Levi-Civita tensor written as $\varepsilon_{\mathbf{a}\mathbf{b}\mathbf{c}} \dots$ has weight -1. Since

$$(\mathbf{A}^n)_i = \sigma (-1)^{n-1} \varepsilon_{i\mathbf{a}\mathbf{b}\mathbf{c}} \dots_{\mathbf{x}} (\mathbf{e}_1)^{\mathbf{a}} (\mathbf{e}_2)^{\mathbf{b}} \dots (\mathbf{e}_N)^{\mathbf{c}} \quad // \mathbf{e}_n \text{ missing}$$

and since the $\mathbf{e}_n$ are ordinary tensorial vectors, it follows that the vector $\mathbf{A}^n$ is in fact a vector density of weight -1. A vector density $\mathbf{V}$ of weight W transforms from x-space to x'-space in this manner

$$\mathbf{V}' = |J|^{-W} \mathbf{R} \mathbf{V}$$

and therefore

$$\mathbf{A}'^n = |J|^{-(-1)} \mathbf{R} \mathbf{A}^n = |J| \mathbf{R} \mathbf{A}^n = g'^{1/2} \mathbf{R} \mathbf{A}^n$$

and this applies also to $d\mathbf{A}^n$.

The object $d\mathbf{x}^{(n)}$ transforms as an ordinary tensorial vector (weight 0).

The object dV may be written as

$$dV = |d\mathbf{A}^n \bullet d\mathbf{x}^{(n)}|$$

which can be verified as follows:

$$\begin{aligned}
dV &= | d\mathbf{A}^n \bullet d\mathbf{x}^{(n)} | = | (d\mathbf{A}^n)_i (d\mathbf{x}^{(n)})^i | \\
&= | \sigma (-1)^{n-1} \varepsilon_{iabc\dots x} (e_1)^a (e_2)^b \dots (e_N)^x \Pi_{i \neq n} dx'^i dx'^n (e_n)^i | \quad // e_n \text{ missing in } x \text{ prod} \\
&= | (-1)^{n-1} \varepsilon_{iabc\dots x} (e_1)^a (e_2)^b \dots (e_n)^i \dots (e_N)^x | \Pi_i dx'^i \\
&= | \varepsilon_{iabc\dots i\dots x} (e_1)^a (e_2)^b \dots (e_n)^i \dots (e_N)^x | \Pi_i dx'^i \\
&= | \det(S^a_b) | \Pi_i dx'^i \quad // \text{Section 3 (a)}
\end{aligned}$$

which agrees with the form given above.

Appendix D (b) 6 shows that the covariant dot product of vector densities $\mathbf{A}$ and $\mathbf{B}$ of weights W and w is a scalar density of weight $W + w$ and therefore $\mathbf{A}' \bullet \mathbf{B}' = |J|^{W+w} \mathbf{A} \bullet \mathbf{B}$. Since $dV = | d\mathbf{A}^n \bullet d\mathbf{x}^{(n)} |$, one concludes that dV is a scalar density of weight $-1+0 = -1$ and therefore

$$dV' = J dV = g'^{1/2} dV$$

To summarize the weight situation for our length, area and volume objects,

$d\mathbf{x}^{(n)}$	weight = 0	ordinary vector
$d\mathbf{A}^n$	weight = -1	vector density
dV	weight = -1	scalar density

(e) Objects in x' -space

Assuming all three objects are tensors of weights shown, one finds that

$$\begin{aligned}
d\mathbf{x}'^{(n)} &= R d\mathbf{x}^{(n)} \\
d\mathbf{A}'^n &= J R d\mathbf{A}^n = g'^{1/2} R d\mathbf{A}^n \\
dV' &= J dV = g'^{1/2} dV
\end{aligned}$$

The first object may be written

$$d\mathbf{x}'^{(n)} = R [e_n dx'^n] = [R e_n] d\mathcal{L}^n = e'_n d\mathcal{L}^n$$

so that

$$(d\mathbf{x}'^{(n)})^i = (e'_n)^i d\mathcal{L}^n = \delta_n^i d\mathcal{L}^n$$

The differential vectors $d\mathbf{x}'^{(n)}$ are thus axis-aligned in x' -space.

For the third object, using the expression $dV = (g')^{1/2} d\mathcal{V}$ from section (c),

$$dV' = g'^{1/2} dV = g' d\mathcal{V}'$$

The remaining second object $d\mathbf{A}^n$ object may be calculated as follows:

$$\begin{aligned}
(d\mathbf{A}^n)_i &= g'^{1/2} R_i^k (d\mathbf{A}^n)_k = & // g'^{1/2} = |J| = |J|^{-(-1)}, \text{ weight} = -1 \\
&= g'^{1/2} R_i^k \sigma (-1)^{n-1} \{ \epsilon_{kabc\dots x} (\mathbf{e}_1)^a (\mathbf{e}_2)^b \dots (\mathbf{e}_N)_x \} d\mathcal{A}^n & // n \text{ missing} \\
&= g'^{1/2} R_i^k \sigma (-1)^{n-1} \{ \epsilon_{kabc\dots x} S^a_1 S^b_2 \dots S^x_N \} d\mathcal{A}^n & // n \text{ missing} \\
&= g'^{1/2} S^k_i \sigma (-1)^{n-1} \{ \epsilon_{kabc\dots x} S^a_1 S^b_2 \dots S^x_N \} d\mathcal{A}^n & // n \text{ missing} \\
&= g'^{1/2} \sigma \{ (-1)^{n-1} \epsilon_{kabc\dots x} S^a_1 S^b_2 \dots S^k_i \dots S^x_N \} d\mathcal{A}^n \\
&= g'^{1/2} \sigma \{ \epsilon_{abc\dots k\dots x} S^a_1 S^b_2 \dots S^k_i \dots S^x_N \} d\mathcal{A}^n
\end{aligned}$$

If $i \neq k$, then two S second indices must be the same and result is 0. If $i = k$, then $\{\} = \det(S^a_b) = \sigma g'^{1/2}$.
Therefore

$$(d\mathbf{A}^n)_i = \delta^n_i g' d\mathcal{A}^n$$

So the area vectors $d\mathbf{A}^n$ are also axis-aligned in x' -space.

To summarize, the x' -space objects can be written in this manner

$(dx'^{(n)})^i = \delta_n^i d\mathcal{L}^n$	length	edge is axis-aligned
$(d\mathbf{A}^n)_i = \delta^n_i g' d\mathcal{A}^n$	area	area vector is axis-aligned
$dV' = g' d\mathcal{V}$	volume	

(f) Evaluation of x -space edge and area magnitudes

Edge

Consider from section (c),

$$d\mathbf{x}^{(n)} = \mathbf{e}_n d\mathcal{L}^n .$$

The magnitude squared is given by (vertical bars always indicate covariant length, Section 5 (i))

$$|d\mathbf{x}^{(n)}|^2 = d\mathbf{x}^{(n)} \bullet d\mathbf{x}^{(n)} = (d\mathcal{L}^n)^2 \mathbf{e}_n \bullet \mathbf{e}_n = g'_{nn} (d\mathcal{L}^n)^2 = h_n'^2 (d\mathcal{L}^n)^2$$

so that

$$dx^{(n)} \equiv |d\mathbf{x}^{(n)}| = h_n' d\mathcal{L}^n$$

Area

Turning now to the area, section (c) showed that

$$d\mathbf{A}^n = g^{1/2} \mathbf{e}^n d\mathcal{A}^n,$$

so the magnitude squared is given by

$$|d\mathbf{A}^n|^2 = d\mathbf{A}^n \bullet d\mathbf{A}^n = (g^{1/2} d\mathcal{A}^n)^2 \mathbf{e}^n \bullet \mathbf{e}^n = (d\mathcal{A}^n)^2 g' g'^{nn} = (d\mathcal{A}^n)^2 g' (1/h'_n)^2.$$

Therefore one way to write the magnitude is this

$$d\mathbf{A}^n \equiv |d\mathbf{A}^n| = (1/h'_n) g^{1/2} d\mathcal{A}^n$$

Another way to write this same $d\mathbf{A}^n$ arises from the following fact:

$$g' g'^{nn} = \text{cof}(g'_{nn})$$

which we now prove:

$$(g'_{up})_{ab} \equiv g'^{ab} \quad (g'_{dn})_{ab} \equiv g'_{ab}$$

$$g'_{up} = (g'_{dn})^{-1} = \text{cof}(g'_{dn}^T) / \det(g'_{dn}) = \text{cof}(g'_{dn}) / \det(g'_{dn})$$

$$\Rightarrow (g'_{up})_{nn} = \text{cof}[(g'_{dn})_{nn}] / \det(g'_{dn})$$

or

$$g'^{nn} = \text{cof}[g'_{nn}] / g' \quad \text{QED}$$

Therefore, one can write

$$|d\mathbf{A}^n|^2 = (d\mathcal{A}^n)^2 g' g'^{nn} = (d\mathcal{A}^n)^2 \text{cof}(g'_{nn})$$

and then one ends up with two forms for the area magnitude $d\mathbf{A}^n \equiv |d\mathbf{A}^n|$ and one form for $|d\mathbf{x}^{(n)}|$

$$d\mathbf{A}^n = (1/h'_n) g^{1/2} d\mathcal{A}^n$$

$$d\mathbf{A}^n = \sqrt{\text{cof}(g'_{nn})} d\mathcal{A}^n$$

$$d\mathbf{x}^{(n)} = h'_n d\mathcal{L}^n$$

When $d\mathbf{A}^n$ is written in its alternative form

$$d\mathbf{A}^n = \sigma (-1)^{n-1} \Pi_{i \neq n}^x (\mathbf{e}_i dx^{i1}) = \sigma (-1)^{n-1} \Pi_{i \neq n}^x \mathbf{e}_i d\mathcal{A}^n$$

one finds that

$$dA^n = |\Pi_{i \neq n}^x \mathbf{e}_i| d\mathcal{A}^n$$

Therefore it must be true that

$$|\Pi_{i \neq n}^x \mathbf{e}_i| = \sqrt{\text{cof}(g'_{nn})}$$

A proof of this fact is not too difficult:

$$\begin{aligned} |\Pi_{i \neq n}^x (\mathbf{e}_i)|^2 &= \Pi_{i \neq n}^x (\mathbf{e}_i) \bullet \Pi_{j \neq n}^x (\mathbf{e}_j) = [\Pi_{i \neq n}^x (\mathbf{e}_i)]_k [\Pi_{j \neq n}^x (\mathbf{e}_j)]_k \\ &= [\varepsilon_{kab c \dots x} (\mathbf{e}_1)^a (\mathbf{e}_2)^b \dots (\mathbf{e}_N)_x] [\varepsilon_{ka' b' c' \dots x} (\mathbf{e}_1)^{a'} (\mathbf{e}_2)^{b'} \dots (\mathbf{e}_N)^{x'}] \quad // \mathbf{e}_n \text{ missing in both} \\ &= \varepsilon_{kab c \dots x} \varepsilon_{ka' b' c' \dots x} \{(\mathbf{e}_1)^a (\mathbf{e}_2)^b \dots (\mathbf{e}_N)_x\} (\mathbf{e}_1)^{a'} (\mathbf{e}_2)^{b'} \dots (\mathbf{e}_N)^{x'} \quad // \mathbf{e}_n \text{ missing} \\ &= \mathbf{e}_1 \bullet \mathbf{e}_1 \mathbf{e}_2 \bullet \mathbf{e}_2 \dots \mathbf{e}_N \bullet \mathbf{e}_N + \text{all signed permutations of 2nd labels} \quad // \mathbf{e}_n \text{ missing} \\ &= g'_{11} g'_{12} \dots g'_{NN} + \text{all signed permutations of 2nd indices} \quad // \mathbf{e}_n \text{ missing} \end{aligned}$$

But this is last object is the determinant of the g'_{ij} matrix with g'_{nn} crossed out, which is to say, it is the minor of g'_{nn} . Since g'_{nn} is a diagonal element, the minor and cofactor are the same. Thus, this last object is in fact just $\text{cof}(g'_{nn})$. QED.

The fact that $|\Pi_{i \neq n}^x \mathbf{e}_i| = \sqrt{\text{cof}(g'_{nn})}$ can be shown for $N=3$ using normal vector algebra. Setting $n = 1$, for example, one needs to show that

$$|\Pi_{i \neq 1}^x \mathbf{e}_i|^2 = |\mathbf{e}_2 \times \mathbf{e}_3|^2 = \text{cof}(g'_{11})$$

To this end, use the vector identity

$$(\mathbf{A} \times \mathbf{B}) \bullet (\mathbf{A} \times \mathbf{B}) = A^2 B^2 - (\mathbf{A} \bullet \mathbf{B})^2$$

to show that

$$|\mathbf{e}_2 \times \mathbf{e}_3|^2 = (\mathbf{e}_2 \times \mathbf{e}_3) \bullet (\mathbf{e}_2 \times \mathbf{e}_3) = |\mathbf{e}_2|^2 |\mathbf{e}_3|^2 - (\mathbf{e}_2 \bullet \mathbf{e}_3)^2 = g'_{22} g'_{33} - (g'_{23})^2 = \text{cof}(g'_{11}).$$

and the cases $n = 2$ and 3 are similar.

Examples of area magnitude transformation for $N = 2, 3, 4$

In the previous section it was shown that $dA^n = \sqrt{\text{cof}(g'_{nn})} d\mathcal{A}^n$. 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 $d\mathcal{A}^n$ are the curvilinear coordinate variations

For $N=2$ the area magnitude transformation results are (for a general non-orthogonal x' -space system)

$$\begin{aligned} dA^1 &= \sqrt{g'_{22}} d\mathbf{a}^1 = h'_2 d\mathbf{a}^1 & d\mathbf{a}^1 &= dx'^1 = d\mathcal{L}^1 \\ dA^2 &= \sqrt{g'_{11}} d\mathbf{a}^2 = h'_1 d\mathbf{a}^2 & d\mathbf{a}^2 &= dx'^2 = d\mathcal{L}^2 \end{aligned}$$

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

$$\begin{aligned} dA^1 &= \sqrt{g'_{22} g'_{33} - (g'_{23})^2} d\mathbf{a}^1 & d\mathbf{a}^1 &= dx'^2 dx'^3 \\ dA^2 &= \sqrt{g'_{33} g'_{11} - (g'_{31})^2} d\mathbf{a}^2 & d\mathbf{a}^2 &= dx'^3 dx'^1 \\ dA^3 &= \sqrt{g'_{11} g'_{22} - (g'_{12})^2} d\mathbf{a}^3 & d\mathbf{a}^3 &= dx'^1 dx'^2 \end{aligned}$$

For an *orthogonal* N=3 system the metric tensor g'_{ab} is diagonal, and then the above simplifies to

$$\begin{aligned} dA^1 &= \sqrt{g'_{22} g'_{33}} d\mathbf{a}^1 = h'_2 h'_3 d\mathbf{a}^1 & d\mathbf{a}^1 &= dx'^2 dx'^3 \\ dA^2 &= \sqrt{g'_{33} g'_{11}} d\mathbf{a}^2 = h'_3 h'_1 d\mathbf{a}^2 & d\mathbf{a}^2 &= dx'^3 dx'^1 \\ dA^3 &= \sqrt{g'_{11} g'_{22}} d\mathbf{a}^3 = h'_1 h'_2 d\mathbf{a}^3 & d\mathbf{a}^3 &= dx'^1 dx'^2 \end{aligned}$$

For an N=4 orthogonal system,

$$\begin{aligned} dA^1 &= \sqrt{\text{cof}(g'_{11})} d\mathbf{a}^1 = h'_2 h'_3 h'_4 d\mathbf{a}^1 & d\mathbf{a}^1 &= dx'^2 dx'^3 dx'^4 \\ dA^2 &= \sqrt{\text{cof}(g'_{22})} d\mathbf{a}^2 = h'_1 h'_3 h'_4 d\mathbf{a}^2 & d\mathbf{a}^2 &= dx'^3 dx'^4 dx'^1 \\ dA^3 &= \sqrt{\text{cof}(g'_{33})} d\mathbf{a}^3 = h'_1 h'_2 h'_4 d\mathbf{a}^3 & d\mathbf{a}^3 &= dx'^4 dx'^1 dx'^2 \\ dA^4 &= \sqrt{\text{cof}(g'_{44})} d\mathbf{a}^4 = h'_1 h'_2 h'_3 d\mathbf{a}^4 & d\mathbf{a}^4 &= dx'^1 dx'^2 dx'^3 \end{aligned}$$

Example 2: Spherical Coordinates: area patches

Consider again $dA^n = \sqrt{\text{cof}(g'_{nn})} d\mathbf{a}^n$. Since spherical coordinates are orthogonal, the orthogonal N=3 example above may be used. Example 2 of Section 5 showed that $[1,2,3 = r,\theta,\varphi]$

$$\begin{aligned} h'_1 &= h'_r = 1 & d\mathbf{a}^1 &= dx'^2 dx'^3 = d\theta d\varphi \\ h'_2 &= h'_\theta = r & d\mathbf{a}^2 &= dx'^3 dx'^1 = r dr d\varphi \\ h'_3 &= h'_\varphi = r \sin\theta & d\mathbf{a}^3 &= dx'^1 dx'^2 = r dr d\theta \end{aligned}$$

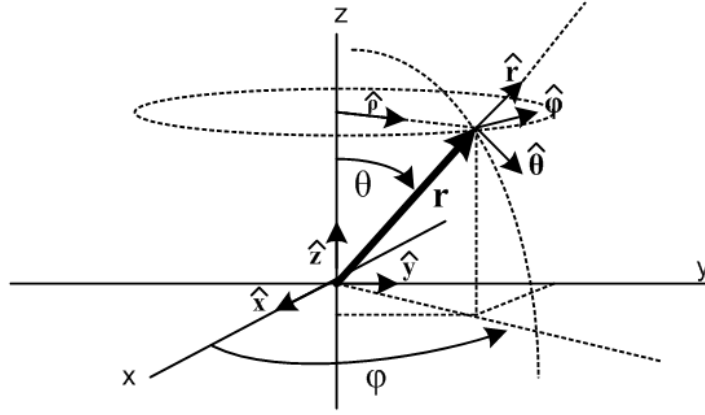
Therefore

$$\begin{aligned} dA^1 &= dA^1 \hat{\mathbf{e}}_1 \Rightarrow d\mathbf{A}^r = dA^r \hat{\mathbf{e}}_r = dA^r \hat{\mathbf{r}} \text{ with } dA^r = h'_2 h'_3 d\mathbf{a}^1 = r^2 \sin\theta d\theta d\varphi \\ dA^2 &= dA^2 \hat{\mathbf{e}}_2 \Rightarrow d\mathbf{A}^\theta = dA^\theta \hat{\mathbf{e}}_\theta = dA^\theta \hat{\boldsymbol{\theta}} \text{ with } dA^\theta = h'_3 h'_1 d\mathbf{a}^2 = r \sin\theta r dr d\varphi \\ dA^3 &= dA^3 \hat{\mathbf{e}}_3 \Rightarrow d\mathbf{A}^\varphi = dA^\varphi \hat{\mathbf{e}}_\varphi = dA^\varphi \hat{\boldsymbol{\phi}} \text{ with } dA^\varphi = h'_1 h'_2 d\mathbf{a}^3 = r dr d\theta \end{aligned}$$

so that

$$\begin{array}{lll}
 d\mathbf{A}^r = r^2 \sin\theta \, d\theta d\varphi \hat{\mathbf{r}} & \rho d\varphi \, r d\theta & \rho = r \sin\theta \\
 d\mathbf{A}^\theta = r \sin\theta \, dr d\varphi \hat{\boldsymbol{\theta}} & \rho d\varphi \, dr & \\
 d\mathbf{A}^\varphi = r dr d\theta \hat{\boldsymbol{\varphi}} & r d\theta \, dr &
 \end{array}$$

where all three vectors are seen to have the correct dimensions L^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,



(g) Evaluation of x' -space edge, area and volume magnitudes

In the previous section, the x -space edge and area magnitudes were given as

$$\begin{aligned}
 dx^{(n)} &= h'_n d\mathcal{L}^n \\
 dA^n &= (1/h'_n) g'^{1/2} d\mathcal{A}^n = \sqrt{\text{cof}(g'_{nn})} d\mathcal{A}^n
 \end{aligned}$$

Appendix D (b) 6 shows that $|\mathbf{A}'| = |\mathbf{J}|^{-W} |\mathbf{A}|$ for a vector density of weight W . This rule can then be used to trivially compute the edge and area magnitudes in x' -space ($W = -1 \Rightarrow |\mathbf{J}|^{-W} = |\mathbf{J}| = g'^{1/2}$):

$$\begin{array}{ll}
 dx'^{(n)} = dx^{(n)} & \text{since weight} = 0 \\
 dA'^{(n)} = g'^{1/2} dA^{(n)} & \text{since weight} = -1 \\
 dV' = g'^{1/2} dV & \text{since weight} = -1
 \end{array}$$

and therefore

$$\begin{aligned}
 dx'^{(n)} &= h'_n d\mathcal{L}^n \\
 dA'^{(n)} &= (1/h'_n) g' d\mathcal{A}^n = g'^{1/2} \sqrt{\text{cof}(g'_{nn})} d\mathcal{A}^n \\
 dV' &= g' d\mathcal{V}
 \end{aligned}$$

(h) Summary of length, area and volume transformation results

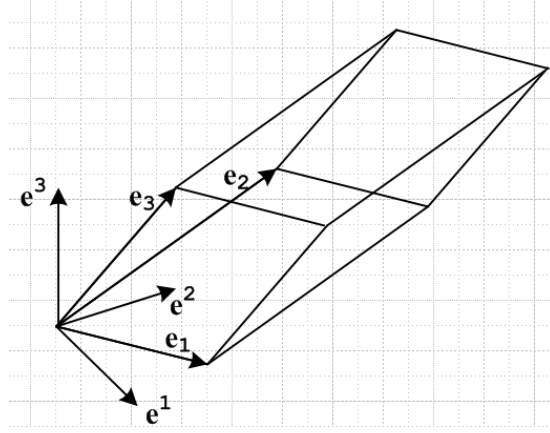
	<u>x-space expression</u>	<u>x-space magnitude</u>
length	$d\mathbf{x}^{(n)} = \mathbf{e}_n d\mathcal{L}^n$	$d\mathbf{x}^{(n)} = h_n' d\mathcal{L}^n$
area	$d\mathbf{A}^n = g'^{1/2} \mathbf{e}^n d\mathcal{A}^n$	$d\mathbf{A}^n = \sqrt{\text{cof}(g'_{nn})} d\mathcal{A}^n = (1/h_n') g'^{1/2} d\mathcal{A}^n$
volume	$dV = (g')^{1/2} d\mathcal{V}$	$dV = (g')^{1/2} d\mathcal{V}$
	<u>x'-space expression</u>	<u>x'-space magnitude</u>
length	$(d\mathbf{x}'^{(n)})^i = \delta_n^i d\mathcal{L}^n$	$d\mathbf{x}'^{(n)} = h_n' d\mathcal{L}^n$
area	$(d\mathbf{A}'^n)_i = \delta_i^n g' d\mathcal{A}^n$	$d\mathbf{A}'^{(n)} = g'^{1/2} \sqrt{\text{cof}(g'_{nn})} d\mathcal{A}^n = (1/h_n') g' d\mathcal{A}^n$
volume	$dV' = g' d\mathcal{V}$	$dV' = g' d\mathcal{V}$
where	$d\mathcal{L}^n \equiv d\mathbf{x}^n$ $d\mathcal{A}^n \equiv \Pi_{i \neq n} d\mathbf{x}^i$ $d\mathcal{V} \equiv \Pi_i d\mathbf{x}^i = d\mathcal{A}^n d\mathcal{L}^n$	$ J = \sigma J = g'^{1/2} \quad // \text{ Section 5 (k)}$ $// \text{ no implied sum}$

Comments:

1. In the N-piped transformation described above, the edge $d\mathbf{x}'^{(n)} = \mathbf{e}'_n d\mathbf{x}^n$ which is axis-aligned in x'-space is mapped (rotated and stretched) into the x-space edge $d\mathbf{x}^{(n)} = \mathbf{e}_n d\mathbf{x}^n$ which is of course aligned with the tangent base vector $\mathbf{e}_n$ in x-space. The length of the edge $d\mathbf{x}^{(n)}$ in x-space is given by $h_n' d\mathbf{x}^n$. In the visualization mapping drawing above, one would say that the Cartesian-view edge length $d\mathbf{x}^n$ on the left got scaled by factor h_n' by the transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$, and this is why h_n' is called a "scale factor".

2. The transformation rule $dV = (g')^{1/2} d\mathcal{V}$ relates the visualized orthogonal volume $d\mathcal{V} = \Pi_i d\mathbf{x}^i$ on the left of the mapping drawing to the volume dV of the rotated and possibly skewed N-piped on the right. The equation $dV = (g')^{1/2} d\mathcal{V}$ is therefore *not* the famous Jacobian Integration Rule since that rule involves an orthogonal differential volume in x-space. (The Rule is given in the next section).

3. The x'-space axis-aligned differential area vector $(d\mathbf{A}'^n)_i = \delta_i^n g' d\mathcal{A}^n$ gets rotated and scaled on its way to x-space and ends up there as $d\mathbf{A}^n = g'^{1/2} \mathbf{e}^n d\mathcal{A}^n$ where it points in the direction of the reciprocal base vector $\mathbf{e}^n$. The $\mathbf{e}_n$ and $\mathbf{e}^n$ are illustrated in this picture for $N=3$:



In this picture the $\mathbf{e}_n$ form a right-handed coordinate system as described in Section 6 (i). As shown there, this means that $\sigma = \text{sign}(\det S) = \text{sign}(\det R) = +1$. Appendix A shows that, with $\mathbf{E}_n \rightarrow \mathbf{e}^n$,

$$\mathbf{e}^n = \det(R) (-1)^{n-1} \mathbf{e}_1 \times \mathbf{e}_2 \times \dots \times \mathbf{e}_N \quad // \mathbf{e}_n \text{ missing}$$

and for $N=2$

$$\begin{aligned} \mathbf{e}^1 &= \det(R) \mathbf{e}_2 \times \mathbf{e}_3 \\ \mathbf{e}^2 &= \det(R) \mathbf{e}_3 \times \mathbf{e}_1 \\ \mathbf{e}^3 &= \det(R) \mathbf{e}_1 \times \mathbf{e}_2 \end{aligned}$$

so the $\mathbf{e}^n$ vectors really are perpendicular to the vectors which span their areas, as the figure attempts to show. The $\mathbf{e}^n$ imagined on the three "near" face centers all point "inward" toward the N-piped center, in agreement with the claim above concerning the directions of the $d\mathbf{A}^n$ vectors. The general perpendicularity rule of course is that $\mathbf{e}^n \cdot \mathbf{e}_m = 0$ for all $m \neq n$, see Appendix A (f).

(i) Transformation of Differential Volume applied to Integration

As discussed in Appendix C (h), 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 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. Moreover, the shape of the differential volume N -piped varies over the region of integration. Nevertheless, this "rag-tag band" of differential volumes, as noted in Appendix C for the 2D case, fills the integration region perfectly.

Alternatively one could consider dV (call it dV_b) to be the usual $dx_1 dx_2 \dots dx_N$ differential volume elements, 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 integral as the limit of a sum,

that $\lim \sum_i dV_a(\mathbf{x}_i) f(\mathbf{x}_i) = \lim \sum_i dV_b(\mathbf{x}_i) f(\mathbf{x}_i)$

$$\int_D dV_a h(\mathbf{x}) = \int_D dV_b h(\mathbf{x})$$

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 $h(\mathbf{x})$. [Formally one might have to require $h(\mathbf{x})$ to be a "test function" $\phi(\mathbf{x})$. Certainly one would require that both integrals converge.]

What has been shown in the *previous* section, regarding the Jacobian, is that

$$dV_a = |J(\mathbf{x}')| d\mathbf{V}' = |J(\mathbf{x}')| \left(\prod_{i=1}^N dx'^i \right) \quad |J(\mathbf{x}')| = \sqrt{g'(\mathbf{x})} \quad // g = +1$$

Combining this with the distributional symbolic equation $dV_a = dV_b$ gives

$$dV_a = dV_b$$

$$|J(\mathbf{x}')| \left(\prod_{i=1}^N dx'^i \right) = \left(\prod_{i=1}^N dx^i \right)$$

or

$$|J(\mathbf{x}')| d\mathbf{V}' = dV_b$$

Now overriding our previous notation, we can make these new commonly used definitions

$$dV \equiv \left(\prod_{i=1}^N dx^i \right)$$

$$dV' \equiv \left(\prod_{i=1}^N dx'^i \right)$$

and express the distributional result as

$$|J(\mathbf{x}')| dV' = dV$$

We refer to this distributional equality in Appendix C as the Jacobian Integration Rule. The symbolic equation is a shorthand for this equation

$$\int_D dV h(\mathbf{x}) = \int_{D'} dV' |J(\mathbf{x}')| h(\mathbf{x})$$

where $h(\mathbf{x}) = h(\mathbf{x}(\mathbf{x}'))$, and where region D' is the same region as D expressed in terms of the $\mathbf{x}'$ coordinates. Writing out the volume elements this says

$$\int_D \left(\prod_{i=1}^N dx^i \right) h(\mathbf{x}) = \int_{D'} \left(\prod_{i=1}^N dx'^i \right) |J(\mathbf{x}')| h(\mathbf{x}(\mathbf{x}'))$$

and finally using Section 5 (k),

$$\int_D (\prod_{i=1}^N dx^i) h(\mathbf{x}) = \int_{D'} (\prod_{i=1}^N dx'^i) [\sqrt{\det(g'_{ab})}] h(\mathbf{x}(\mathbf{x}'))$$

For example, when applied to polar and spherical coordinates, one gets

$$\int_D dx dy h(\mathbf{x}) = \int_{D'} dr d\theta [r] h(\mathbf{x}(r,\theta)) \quad \sqrt{\det(g'_{ab})} = r$$

$$\int_D dx dy dz h(\mathbf{x}) = \int_{D'} dr d\theta d\phi [r^2 \sin\theta] h(\mathbf{x}(r,\theta,\phi)) \quad \sqrt{\det(g'_{ab})} = r^2 \sin\theta$$

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 $h(\mathbf{x}(r,\theta,\phi)) = h(r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta)$.

(j) Interpretations of the Jacobian

Using Section 5 (k) facts (in standard notation) and the above sections, one can produce various expressions and interpretations for the Jacobian J and its absolute value $|J|$:

$$J(\mathbf{x}') \equiv \det(S^i_j(\mathbf{x}')) = \det(\partial x^i / \partial x'^j) = 1 / \det(R^i_j(\mathbf{x}')) = 1 / \det(\partial x'^i / \partial x^j) \quad // \text{ Section 5 (k)}$$

$$|J(\mathbf{x}')| = \sqrt{\det(g'_{ab}(\mathbf{x}'))} = \sqrt{g'(\mathbf{x}')} \quad // \text{ Section 5 (k)}$$

$$|J(\mathbf{x}')| = \text{the volume of the N-piped in x-space spanned by the } \mathbf{e}_n(\mathbf{x}), \text{ where } \mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$$

$$|J(\mathbf{x}')| = dV_{N\text{-piped}} / dV' = \text{ratio of differential x-space N-piped volume} / (\prod_{i=1}^N dx'^i)$$

$$|J(\mathbf{x}')| = dV / dV' = (\prod_{i=1}^N dx^i) / (\prod_{i=1}^N dx'^i) \quad // \text{ distributional Jacobian Integration Rule}$$

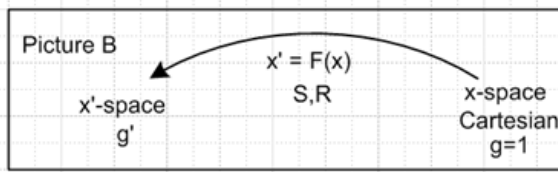
As discussed in Section 6 (i), if the curvilinear coordinates are ordered so that the $\mathbf{e}_n$ form a right handed coordinate system, then $\det(S) > 0$, $\sigma = \text{sign}(\det(S)) = +1$, and $|J| = J$.

9. The Divergence in curvilinear coordinates

(a) Geometric Derivation of the Curvilinear Divergence Formula

In Cartesian coordinates $\text{div } \mathbf{B} = \nabla \cdot \mathbf{B} = \partial_n B^n$, but expressed in curvilinear coordinates the right side has a more complicated form.

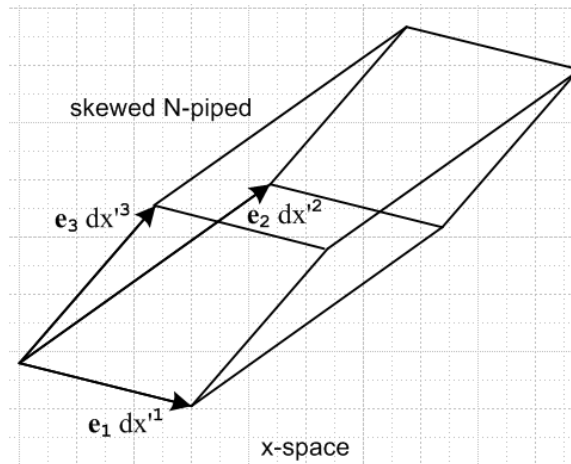
We provide here a geometric derivation of the formula for the divergence of a contravariant vector field expressed in curvilinear coordinates, which means x' -space coordinates with Picture B.



This derivation is an exercise in using the transformation results obtained in Section 8 above, and in understanding the meaning of the components of a vector, as discussed in Appendix C (d).

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 $\mathbf{x}$. Being a scalar field, the divergence is a property of the vector field at some point $\mathbf{x}$ 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.

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 (c) above:



In order to avoid confusion with volume V or area A , we name the vector field $\mathbf{B}$. As just noted, the divergence of a vector field $\mathbf{B}$ 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,

$$[\text{div } \mathbf{B}](\mathbf{x}) = (1/dV) \int d\mathbf{A} \bullet \mathbf{B} = (1/dV) \int d\mathbf{A}(\mathbf{x}) \bullet \mathbf{B}(\mathbf{x})$$

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 $d\mathbf{A}$ of a face changes, hence $d\mathbf{A}(\mathbf{x})$.

Comment on the div $\mathbf{B}$ as a scalar. If $\mathbf{B}$ is a tensorial vector field, then $\text{div } \mathbf{B}$ is a tensorial scalar field, and one can write

$$[\text{div } \mathbf{B}]'(\mathbf{x}') = [\text{div } \mathbf{B}](\mathbf{x})$$

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 $\mathbf{B}$ is a tensorial scalar. In Section 8 it was shown that $d\mathbf{A}$ is a vector density of weight -1 and so is dV . This means that $dV' = J dV$ and $d\mathbf{A}' = J R d\mathbf{A}$ so the ratio $d\mathbf{A}/dV$ is a tensorial vector. The fact that $\text{div } \mathbf{B}$ is a tensorial scalar is more obvious from the alternative divergence derivation given in section (e) below.

The task is now to compute the integral $\int d\mathbf{A}(\mathbf{x}) \bullet \mathbf{B}(\mathbf{x})$.

Appendix B shows that the N-piped faces come in parallel pairs, so we start by considering pair n. As shown in Section 8 (c), the vector area of the far face of pair n is given by

$$d\mathbf{A}^n(\mathbf{x}) = |\det(S^i_j(\mathbf{x}'))| \mathbf{e}^n(\mathbf{x}) (\prod_{i \neq n} dx'^i) = \sqrt{g'(\mathbf{x}')} \mathbf{e}^n(\mathbf{x}) (\prod_{i \neq n} dx'^i)$$

Here $|\det(S)| = |J| = g'^{1/2}$ (Section 5 (k)), and $\mathbf{e}^n$ are the reciprocal base vectors (Section 6). Quantities $d\mathbf{A}^n$, S , $\mathbf{e}^n$ and g' are explicitly shown as functions of space. Appendix B (c) shows that 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 $-d\mathbf{A}^n$.

The contribution to the above divergence integral from "far face n" is, approximately,

$$\begin{aligned} [d\mathbf{A}^n] \bullet \mathbf{B}(\mathbf{x}) &\approx [\sqrt{g'(\mathbf{x}'_{\text{far}})} \mathbf{e}^n(\mathbf{x}_{\text{far}}) (\prod_{i \neq n} dx'^i)] \bullet \mathbf{B}(\mathbf{x}_{\text{far}}) \\ &= \sqrt{g'(\mathbf{x}'_{\text{far}})} (\prod_{i \neq n} dx'^i) \mathbf{e}^n(\mathbf{x}_{\text{far}}) \bullet \mathbf{B}(\mathbf{x}_{\text{far}}) \end{aligned}$$

where $\mathbf{x}_{\text{far}}$ is taken to be a point at the center of far face n. Recall from Section 7 (s) that

$$\mathbf{B}(\mathbf{x}) = B^m(\mathbf{x}') \mathbf{e}_m \quad \text{where} \quad B^m(\mathbf{x}') = \mathbf{e}^n(\mathbf{x}) \bullet \mathbf{B}(\mathbf{x})$$

which says that, when $\mathbf{B}$ is expanded on the $\mathbf{e}_n$, the coefficients B^m of the expansion are the contravariant components of vector $\mathbf{B}$ transformed into $\mathbf{B}'$ in x' -space (the curvilinear coordinate space) by $\mathbf{B}' = R\mathbf{B}$.

Inserting the last equation above applied at $\mathbf{x}' = \mathbf{x}'_{\text{far}}$,

$$\mathbf{B}^{\mathbf{n}}(\mathbf{x}'_{\text{far}}) = \mathbf{e}^{\mathbf{n}}(\mathbf{x}_{\text{far}}) \bullet \mathbf{B}(\mathbf{x}_{\text{far}})$$

into the previous $\approx$ equation then gives

$$d\mathbf{A}^{\mathbf{n}} \bullet \mathbf{B}(\mathbf{x}_{\text{far}}) \approx \sqrt{g'(\mathbf{x}'_{\text{far}})} (\Pi_{i \neq n} dx'^i) \mathbf{B}^{\mathbf{n}}(\mathbf{x}'_{\text{far}})$$

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 $d\mathbf{A}^{\mathbf{n}}$ is reversed. Adding the contributions of these two faces of pair n gives

$$\int_{\text{two faces } n} d\mathbf{A} \bullet \mathbf{B}(\mathbf{x}) = \{ \sqrt{g'(\mathbf{x}'_{\text{far}})} \mathbf{B}^{\mathbf{n}}(\mathbf{x}'_{\text{far}}) - \sqrt{g'(\mathbf{x}'_{\text{near}})} \mathbf{B}^{\mathbf{n}}(\mathbf{x}'_{\text{near}}) \} (\Pi_{i \neq n} dx'^i)$$

In x' -space, if $\mathbf{x}'_{\text{cen}}$ is a point at the center of the near face of face pair n , then

$$\begin{aligned} \mathbf{x}'_{\text{near}} &= \mathbf{x}'_{\text{cen}} \\ \mathbf{x}'_{\text{far}} &= \mathbf{x}'_{\text{cen}} + \mathbf{e}'_{\mathbf{n}} dx'^{\mathbf{n}} \end{aligned} \quad \text{where } \mathbf{e}'_{\mathbf{n}} = \text{axis-aligned basis vector in } x'\text{-space,}$$

since these two points map into the near and far face n centers in x -space. For any function f ,

$$f(\mathbf{x}'_{\text{far}}) - f(\mathbf{x}'_{\text{near}}) = (\partial'_{\mathbf{n}} f(\mathbf{x}')) dx'^{\mathbf{n}} \quad // \text{ no implied sum on } n$$

where a change is made only in coordinate $x'^{\mathbf{n}}$ by amount $dx'^{\mathbf{n}}$. Applying to $f = J \mathbf{B}^{\mathbf{n}}$ yields

$$\{ \sqrt{g'(\mathbf{x}'_{\text{far}})} \mathbf{B}^{\mathbf{n}}(\mathbf{x}'_{\text{far}}) - \sqrt{g'(\mathbf{x}'_{\text{near}})} \mathbf{B}^{\mathbf{n}}(\mathbf{x}'_{\text{near}}) \} \approx \partial'_{\mathbf{n}} [\sqrt{g'(\mathbf{x}'_{\text{cen}})} \mathbf{B}^{\mathbf{n}}(\mathbf{x}'_{\text{cen}})] dx'^{\mathbf{n}}$$

In the limit that differentials are very close to 0, replace $\mathbf{x}_{\text{cen}}$ by $\mathbf{x}$. Then

$$\begin{aligned} \int_{\text{two faces } n} d\mathbf{A} \bullet \mathbf{B}(\mathbf{x}) &= \partial'_{\mathbf{n}} [\sqrt{g'(\mathbf{x}')} \mathbf{B}^{\mathbf{n}}(\mathbf{x}')] dx'^{\mathbf{n}} (\Pi_{i \neq n} dx'^i) \\ &= \partial'_{\mathbf{n}} [\sqrt{g'(\mathbf{x}')} \mathbf{B}^{\mathbf{n}}(\mathbf{x}')] (\Pi_{\mathbf{i}} dx'^i) \end{aligned}$$

where now all N differentials are present in $(\Pi_{\mathbf{i}} dx'^i)$. 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

$$\text{total flux} = \int d\mathbf{A} \bullet \mathbf{B}(\mathbf{x}) = \partial'_{\mathbf{n}} [\sqrt{g'(\mathbf{x}')} \mathbf{B}^{\mathbf{n}}(\mathbf{x}')] (\Pi_{\mathbf{i}} dx'^i) = \partial'_{\mathbf{n}} [\sqrt{g'(\mathbf{x}')} \mathbf{B}^{\mathbf{n}}(\mathbf{x}')] d\mathbf{v}'$$

where $d\mathbf{v}' = \Pi_{\mathbf{i}} dx'^i$ is the volume of the differential N -piped in Cartesian-view x' -space (Section 8 (a)). The divergence of $\mathbf{B}$ from the defining symbolic expression is then

$$[\text{div } \mathbf{B}](\mathbf{x}) = \int d\mathbf{A} \bullet \mathbf{B}(\mathbf{x}) / dV = \partial'_n [\sqrt{g'(\mathbf{x}')}] B'^n(\mathbf{x}') (d\mathcal{V}'/dV)$$

where dV is the volume of the N-piped shown above. In Section 8 (h) it is shown that

$$dV = (g')^{1/2} d\mathcal{V}' \quad \Rightarrow \quad (d\mathcal{V}'/dV) = 1/\sqrt{g'(\mathbf{x}')}$$

so that

$$[\text{div } \mathbf{B}](\mathbf{x}) = [1/\sqrt{g'(\mathbf{x}')}] \partial'_n [\sqrt{g'(\mathbf{x}')}] B'^n(\mathbf{x}') \quad // \text{ all } \mathbf{x}'\text{-space coordinates and objects}$$

$$[\text{div } \mathbf{B}](\mathbf{x}) = \partial_n B^n(\mathbf{x}) \quad // \text{ all } \mathbf{x}\text{-space coordinates and objects}$$

The added second line just shows $[\text{div } \mathbf{B}](\mathbf{x})$ expressed in terms of the Cartesian $\mathbf{x}$ -space coordinates and objects, while the first line resulting from our derivation shows *the same* $[\text{div } \mathbf{B}](\mathbf{x})$ expressed in terms of only $\mathbf{x}'$ -space objects and coordinates. The claim advertised above has been fulfilled.

If $\mathbf{B}$ is a tensorial vector, then as noted above $\text{div } \mathbf{B}$ is a tensorial scalar,

$$[\text{div } \mathbf{B}](\mathbf{x}) = [\text{div } \mathbf{B}'](\mathbf{x}')$$

and thus the left sides of both equations above could be replaced by $[\text{div } \mathbf{B}'](\mathbf{x}')$.

(b) Various expressions for $\text{div } \mathbf{B}$

It is shown above that

$$[\text{div } \mathbf{B}](\mathbf{x}) = [1/\sqrt{g'(\mathbf{x}')}] \partial'_n [\sqrt{g'(\mathbf{x}')}] B'^n(\mathbf{x}')$$

To obtain $\text{div } \mathbf{B}$ written in terms of *covariant* components B'_n , one sets $B'^n = g'^{nm} B'_m$ to get

$$[\text{div } \mathbf{B}](\mathbf{x}) = [1/\sqrt{g'(\mathbf{x}')}] \partial'_n [\sqrt{g'(\mathbf{x}')}] g'^{nm}(\mathbf{x}') B'_m(\mathbf{x}')$$

and recall that the B'_m are the coefficients of $\mathbf{B}$ when expanded on the $\mathbf{e}^n$, $\mathbf{B} = B'_n \mathbf{e}^n$.

In practical work $\mathbf{B}$ is expanded on the *unit* vectors $\hat{\mathbf{e}}_n \equiv \mathbf{e}_n / |\mathbf{e}_n| = \mathbf{e}_n / h'_n$ so that

$$\mathbf{B} = B'^n \mathbf{e}_n = B'^n h'_n \hat{\mathbf{e}}_n = \mathcal{B}^n \hat{\mathbf{e}}_n \quad \text{where } \mathcal{B}^n \equiv B'^n h'_n$$

and then

$$[\text{div } \mathbf{B}](\mathbf{x}) = [1/\sqrt{g'(\mathbf{x}')}] \partial'_n [\sqrt{g'(\mathbf{x}')}] \mathcal{B}^n(\mathbf{x}') / h'_n(\mathbf{x}')]$$

For example, spherical coordinate work might use $\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3 = \hat{\mathbf{r}}, \hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}}$. As noted earlier, the components $\mathcal{B}^n(\mathbf{x})$ are not contravariant vector components since they don't quite transform properly:

$$\mathbf{B}^n = \mathbf{R}^n_m \mathbf{B}^m \quad (\mathcal{B}^n / h'_n) = \mathbf{R}^n_m (\mathcal{B}^m / 1) \quad \mathcal{B}^n = h'_n (\mathbf{R}^n_m \mathcal{B}^m)$$

Our Picture B results so far are these, assuming $\mathbf{B}$ is a tensorial vector,

General: $\mathbf{B}^n(\mathbf{x}') = \mathbf{R}^n_m \mathbf{B}^m(\mathbf{x}) \quad \mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}') \equiv \mathbf{x}(\mathbf{x}') \quad \mathbf{x}' = \mathbf{F}(\mathbf{x}) \equiv \mathbf{x}'(\mathbf{x})$

$$\begin{aligned} [\text{div } \mathbf{B}](\mathbf{x}) &= [1/\sqrt{g'(\mathbf{x}')}] \partial'_n [\sqrt{g'(\mathbf{x}')} \mathbf{B}^n(\mathbf{x}')] & \mathbf{B} &= \mathbf{B}^n \mathbf{e}_n \\ [\text{div } \mathbf{B}](\mathbf{x}) &= [1/\sqrt{g'(\mathbf{x}')}] \partial'_n [\sqrt{g'(\mathbf{x}')} \mathcal{B}^n(\mathbf{x}')/ h'_n(\mathbf{x}')] & \mathbf{B} &= \mathcal{B}^n \hat{\mathbf{e}}_n \\ [\text{div } \mathbf{B}](\mathbf{x}) &= [1/\sqrt{g'(\mathbf{x}')}] \partial'_n [\sqrt{g'(\mathbf{x}')} g^{mm}(\mathbf{x}') \mathbf{B}_m(\mathbf{x}')] & \mathbf{B} &= \mathbf{B}'_n \mathbf{e}^n \\ [\text{div } \mathbf{B}](\mathbf{x}) &= \partial_n \mathbf{B}^n(\mathbf{x}) \quad // \mathbf{B}^n = \text{Cartesian components of } \mathbf{B} & \mathbf{B} &= \mathbf{B}^n \hat{\mathbf{n}} \\ [\text{div } \mathbf{B}](\mathbf{x}) &= [\text{div } \mathbf{B}](\mathbf{x}') \end{aligned}$$

For orthogonal curvilinear coordinates, one has

$$g'_{ij} = h'_i{}^2 \delta_{ij} \quad g'^{ij} = h'_i{}^{-2} \delta_{ij} \quad \det(g'_{ij}) = \Pi_i h'_i{}^2 \quad \sqrt{g'} = (\Pi_i h'_i) = h'_1 h'_2 \dots h'_N$$

so the above expressions can be written (the arguments $\mathbf{x}'$ of the h'_n are now suppressed)

Orthogonal:

$$\begin{aligned} [\text{div } \mathbf{B}](\mathbf{x}) &= [1/(\Pi_i h'_i)] \partial'_n [(\Pi_i h'_i) \mathbf{B}^n(\mathbf{x}')] & \mathbf{B} &= \mathbf{B}^n \mathbf{e}_n \\ [\text{div } \mathbf{B}](\mathbf{x}) &= [1/(\Pi_i h'_i)] \partial'_n [(\Pi_i h'_i) \mathcal{B}^n(\mathbf{x}')/ h'_n] & \mathbf{B} &= \mathcal{B}^n \hat{\mathbf{e}}_n \\ [\text{div } \mathbf{B}](\mathbf{x}) &= [1/(\Pi_i h'_i)] \partial'_n [(\Pi_i h'_i) \mathbf{B}'_n(\mathbf{x}')/ h'^n_n] & \mathbf{B} &= \mathbf{B}'_n \mathbf{e}^n \\ [\text{div } \mathbf{B}](\mathbf{x}) &= \partial_n \mathbf{B}^n(\mathbf{x}) \quad // \mathbf{B}^n = \text{Cartesian components of } \mathbf{B} & \mathbf{B} &= \mathbf{B}^n \hat{\mathbf{n}} \\ [\text{div } \mathbf{B}](\mathbf{x}) &= [\text{div } \mathbf{B}](\mathbf{x}') \end{aligned}$$

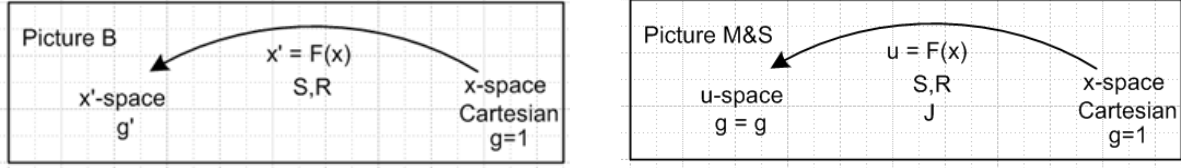
Comment: Are the equations of the above "General" block valid if $\mathbf{B}$ is *not* a tensorial vector? For such a $\mathbf{B}$ one might try to make it be tensorial "by definition" as discussed in Section 2 (h). One would then go ahead and define $\mathbf{B}^n(\mathbf{x}') \equiv \mathbf{R}^n_m \mathbf{B}^m(\mathbf{x})$ and claim success. If such a definition does not result in an inconsistency, then such a $\mathbf{B}$ has been moved into the class of tensorial vectors. Example 1 of Section 2 (h) shows how such an inconsistency might arise, and it is interesting to see how that plays out here. Suppose $\mathbf{F}$ is non-linear so that $\mathbf{R}$ and $g' = \mathbf{R}\mathbf{R}^T$ are functions of $\mathbf{x}'$ and are not constants. Take $\mathbf{B}(\mathbf{x}) = \mathbf{x}$ (the identity field) and try to make it be contravariant by definition, $\mathbf{x}^n \equiv \mathbf{R}^n_m \mathbf{x}^m$. The Cartesian divergence is then $\text{div } \mathbf{B} = \partial_n \mathbf{B}^n(\mathbf{x}) = \partial_n x^n = 3$. But the first equation of the General block says

$$\text{div } \mathbf{B} = [1/\sqrt{g'(\mathbf{x}')}] \partial'_n [\sqrt{g'(\mathbf{x}')} x^n] = \partial'_n x^n + [1/\sqrt{g'(\mathbf{x}')}] x^n \partial'_n \sqrt{g'(\mathbf{x}')} = 3 + \text{other stuff}$$

and thus the two calculations for $\text{div } \mathbf{B}$ disagree. As noted earlier, $\mathbf{x}' \equiv \mathbf{R}\mathbf{x}$ conflicts with $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ in the case of non-linear $\mathbf{F}$.

This comment can be applied to the tensorial character of the differential operators treated in later Sections.

(c) Translation from Picture B to Picture M&S



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 $\mathbf{x}'$ by $\mathbf{u}$ everywhere
- replace ∂'_n by ∂_n meaning $\partial/\partial u^n$ (exception: on a "Cartesian" line ∂_n means $\partial/\partial x^n$)
- 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, $B'^m \rightarrow B^n$, $f' \rightarrow f$)

After this translation, all unprimed tensor components are functions of $\mathbf{x}$, while all italicized tensor components are functions of $\mathbf{u}$.

Here then are the translations of the two blocks above: (implied summation everywhere)

General: now $B^n(\mathbf{u}) = R^n_m B^m(\mathbf{x})$ $\mathbf{x} = F^{-1}(\mathbf{u}) \equiv \mathbf{x}(\mathbf{u})$ $\mathbf{u} = F(\mathbf{x}) \equiv \mathbf{u}(\mathbf{x})$

$$\begin{aligned}
 [\text{div } \mathbf{B}](\mathbf{x}) &= [1/\sqrt{g}] \partial_n [\sqrt{g} B^n] & \mathbf{B} &= B^n \mathbf{e}_n \\
 [\text{div } \mathbf{B}](\mathbf{x}) &= [1/\sqrt{g}] \partial_n [\sqrt{g} \mathcal{B}^n / h_n] & \mathbf{B} &= \mathcal{B}^n \hat{\mathbf{e}}_n \quad // \text{ M\&S 1.06} \\
 [\text{div } \mathbf{B}](\mathbf{x}) &= [1/\sqrt{g}] \partial_n [\sqrt{g} g^{nm} B_m] & \mathbf{B} &= B_n \mathbf{e}^n \\
 [\text{div } \mathbf{B}](\mathbf{x}) &= \partial_n B^n & \mathbf{B} &= B^n \hat{\mathbf{n}} = B_n \hat{\mathbf{n}} \\
 [\text{div } \mathbf{B}](\mathbf{x}) &= [\text{div } \mathbf{B}](\mathbf{u}) & & // \text{ transformation (scalar)}
 \end{aligned}$$

Orthogonal:

$$\begin{aligned}
 [\text{div } \mathbf{B}](\mathbf{x}) &= [1/(\Pi_i h_i)] \partial_n [(\Pi_i h_i) B^n] & \mathbf{B} &= B^n \mathbf{e}_n \\
 [\text{div } \mathbf{B}](\mathbf{x}) &= [1/(\Pi_i h_i)] \partial_n [(\Pi_i h_i) \mathcal{B}^n / h_n] & \mathbf{B} &= \mathcal{B}^n \hat{\mathbf{e}}_n \\
 [\text{div } \mathbf{B}](\mathbf{x}) &= [1/(\Pi_i h_i)] \partial_n [(\Pi_i h_i) B_n / h_n^2] & \mathbf{B} &= B_n \mathbf{e}^n
 \end{aligned}$$

Notice that the scalar function $[\text{div } \mathbf{B}](\mathbf{x}) \rightarrow [\text{div } \mathbf{B}](\mathbf{u})$ according to the fourth rule above, and that the arguments of all $h_k(\mathbf{u})$ are suppressed.

As an example, for $N=3$ the second line above becomes

$$[\text{div } \mathbf{B}](\mathbf{x}) = [1/(h_1 h_2 h_3)] \{ \partial_1 [h_2 h_3 \mathcal{B}^1(\mathbf{u})] + \text{cyclic} \} \quad \mathbf{B} = \mathcal{B}^n \hat{\mathbf{e}}_n$$

where + cyclic means two other terms with 1,2,3 cyclically permuted.

With the replacements

$$\mathbf{B} \rightarrow \mathbf{E}, \quad \mathcal{B}^n \rightarrow E_n \quad h_n \rightarrow \sqrt{g_{nn}}$$

the equation marked above agrees with Moon & Spencer p 2 (1.06).

Comment: The AaSansOutline font $\mathcal{B}^n$ used above for components of vectors expanded onto $\hat{\mathbf{e}}_n$ is a bit clumsy and does not reproduce well in PDF files. It had to be something in upper case distinct from B^n and B^n . In practice one can replace $\mathcal{B}^n$ with a different symbol and then $\mathcal{B}^n$ is just a formal notation appearing in formulas. For example, in spherical coordinates $(1,2,3) = (r,\theta,\phi)$ one can make the replacements $\mathcal{B}^1, \mathcal{B}^2, \mathcal{B}^3 \rightarrow B_r, B_\theta, B_\phi$ and these then do not conflict with B_x, B_y, B_z or B_r, B_θ, B_ϕ . See Section 14 Example 1 for another example.

(d) 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	p 115	ξ_n	x_n	ξ -space	h_n
Margenau & Murphy	p 192	q^n	x^n	q-space	Q_n

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

$$\text{div } \mathbf{A} = \nabla \cdot \mathbf{A} = \frac{1}{h_1 h_2 h_3} \sum_n \frac{\partial}{\partial \xi_n} \left(h_1 h_2 h_3 \frac{A_n}{h_n} \right)$$

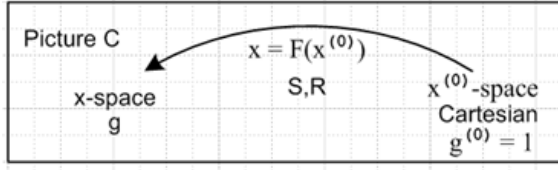
which compare to the above

$$[\text{div } \mathbf{A}](\mathbf{x}) = [\text{div } \mathbf{A}](\mathbf{u}) = [1/(h_1 h_2 h_3)] \partial_n [h_1 h_2 h_3 \mathcal{A}^n(\mathbf{u}) / h_n] \quad \mathbf{A} = \mathcal{A}^n \hat{\mathbf{e}}_n$$

so presumably one should identify the M&F A_n with $\mathcal{A}^n$, the coefficient of $\hat{\mathbf{e}}_n$.

(e) The Christoffel derivation of div **B**

This derivation is done in Picture C where the curvilinear coordinates are x so that ∂_a means $\partial/\partial x^a$ and $g \neq 1$ is the curvilinear metric tensor,



Mention of $x^{(0)}$ -space is completely avoided. In fact, even the x coordinate symbol rarely appears and could also have been removed. Function $\text{div } \mathbf{B}$ is in x -space, and once seen as a tensorial scalar, can be regarded as also being in $x^{(0)}$ space,

$$[\text{div } \mathbf{B}]^{(0)}(x^{(0)}) = [\text{div } \mathbf{B}](x) \quad // = B^a{}_{,a}$$

We start with the covariant derivative of a contravariant vector field component, which object is known to be a mixed rank-2 tensor (see Section 7 (v)),

$$B^b{}_{,a} \equiv \partial_a B^b + \Gamma^b_{an} B^n$$

$$\text{where } \Gamma^c_{ab} = \frac{1}{2} g^{cd} (\partial_a g_{bd} + \partial_b g_{ad} - \partial_d g_{ab}).$$

Then the *tensorial scalar* object $\text{div } \mathbf{B}$ is defined by index contraction to be

$$\text{div } \mathbf{B} \equiv B^a{}_{,a} \equiv \partial_a B^a + \Gamma^a_{an} B^n$$

Evaluation of Γ^a_{an} gives,

$$\Gamma^a_{an} = \frac{1}{2} g^{ad} (\partial_a g_{nd} + \partial_n g_{ad} - \partial_d g_{an}) = \frac{1}{2} g^{ad} \partial_n g_{ad} = \frac{1}{2} (1/g) \partial_n g = (1/\sqrt{g}) \partial_n (\sqrt{g})$$

where the first and third terms cancel due to symmetry and the fact that $(1/g) \partial_n g = g^{ab} (\partial_n g_{ab})$ is proved as follows:

$$(1) g^{ab} = (g^{-1})_{ab} = \text{cof}(g_{ab})^T / \det(g_{ab}) = \text{cof}(g_{ab}) / g \Rightarrow \text{cof}(g_{ab}) = g g^{ab}$$

$$(2) g = \det(g_{ab}) = \Sigma_a g_{ab} \text{cof}(g_{ab}) \Rightarrow \partial g / \partial g_{ab} = \text{cof}(g_{ab}) = g g^{ab}$$

$$(3) \partial_n g = \partial g / \partial x^n = (\partial g / \partial g_{ab}) (\partial g_{ab} / \partial x^n) = g g^{ab} (\partial_n g_{ab}) \Rightarrow (1/g) \partial_n g = g^{ab} (\partial_n g_{ab})$$

Therefore

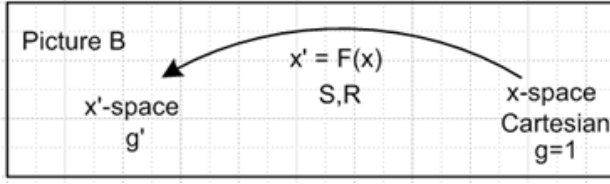
$$\text{div } \mathbf{B} = B^a{}_{,a} \equiv \partial_a B^a + \Gamma^a_{an} B^n = \partial_n B^n + (1/\sqrt{g}) \partial_n (\sqrt{g}) B^n$$

$$= [1/\sqrt{g}] \partial_n [\sqrt{g} B^n]$$

in agreement with the geometric derivation. Thus the entire distinction from $\partial_n B^n$ is the second term above which arises from the affine connection Γ^b_{an} part of the covariant derivative.

10. The Gradient in curvilinear coordinates

This Section is considered in the Picture B context,



(a) Expressions for grad f

The gradient of f is defined in Cartesian space by

$$[\text{grad } f]_{\mathbf{n}} \equiv G_{\mathbf{n}} \equiv \partial_{\mathbf{n}} f(\mathbf{x}) \quad \mathbf{G} = \text{grad } f = \partial_{\mathbf{n}} f(\mathbf{x}) \hat{\mathbf{n}}$$

Assuming f is a tensorial scalar field under F, then $G_{\mathbf{n}} = \partial_{\mathbf{n}} f(\mathbf{x})$ are covariant vector field components under F. Since the equation above is then a tensor equation, we know it is covariant in the sense of Section 7 (u). Therefore in x'-space it becomes

$$[\text{grad } f]_{\mathbf{n}} = G'_{\mathbf{n}} \equiv \partial'_{\mathbf{n}} f(\mathbf{x}')$$

According to Section 7 (s) (based on Section 6 (f)) vector $\mathbf{G}$ can be expanded as

$$\mathbf{G} = G'_1 \mathbf{e}^1 + G'_2 \mathbf{e}^2 + \dots = \sum_{\mathbf{n}} G'_{\mathbf{n}} \mathbf{e}^{\mathbf{n}} \quad \text{where} \quad \mathbf{e}_{\mathbf{n}} \cdot \mathbf{G} = G'_{\mathbf{n}}$$

where the coefficients $G'_{\mathbf{n}}$ are the covariant components of $\mathbf{G}'$ in x'-space, which is the transformed $\mathbf{G}$. One can thus write

$$\mathbf{G}(\mathbf{x}) = [\text{grad } f](\mathbf{x}) = G'_{\mathbf{n}} \mathbf{e}^{\mathbf{n}} = \partial'_{\mathbf{n}} f(\mathbf{x}') \mathbf{e}^{\mathbf{n}} = \mathbf{e}^{\mathbf{n}} \partial'_{\mathbf{n}} f(\mathbf{x}') \equiv \nabla'_{\text{CL}} f(\mathbf{x}') \quad \nabla'_{\text{CL}} \equiv \mathbf{e}^{\mathbf{n}} \partial'_{\mathbf{n}}$$

$$\mathbf{G}(\mathbf{x}) = [\text{grad } f](\mathbf{x}) = \partial_{\mathbf{n}} f(\mathbf{x}) \hat{\mathbf{n}} = \hat{\mathbf{n}} \partial_{\mathbf{n}} f(\mathbf{x}) = \nabla f(\mathbf{x}) \quad \nabla \equiv \hat{\mathbf{n}} \partial_{\mathbf{n}}$$

where the second line shows the usual Cartesian form of the gradient. The first line shows how one *could* define a "curvilinear gradient" operator $\nabla'_{\text{CL}} \equiv \mathbf{e}^{\mathbf{n}} \partial'_{\mathbf{n}}$, but this does not seem particular useful. To restate,

$$\begin{aligned} [\text{grad } f](\mathbf{x}) &= \partial'_{\mathbf{n}} f(\mathbf{x}') \mathbf{e}^{\mathbf{n}} & // \text{ Curvilinear} \\ [\text{grad } f](\mathbf{x}) &= \partial_{\mathbf{n}} f(\mathbf{x}) \hat{\mathbf{n}} & // \text{ Cartesian} \end{aligned}$$

In the first line, the reciprocal vectors $\mathbf{e}^{\mathbf{n}}$ 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

$$\partial'_{\mathbf{n}} f(\mathbf{x}') = \partial'_{\mathbf{n}} f(\mathbf{x}(\mathbf{x}')) \quad \text{where} \quad \mathbf{x}(\mathbf{x}') = F^{-1}(\mathbf{x}')$$

The *contravariant* components of $\text{grad } f$ are then easily obtained as

$$G'^i(\mathbf{x}') = [\text{grad } f]^i(\mathbf{x}') = \partial'^i f'(\mathbf{x}') = g'^{ij}(\mathbf{x}') \partial'_j f'(\mathbf{x}') \quad \mathbf{G} = \text{grad } f = G'^i(\mathbf{x}') \mathbf{e}_i$$

If an expansion on unit vectors is desired, the right equation on the last line can be written,

$$\mathbf{G} = [\text{grad } f](\mathbf{x}) = G'^i \mathbf{e}_i = (G'^i h'_{ij}) \hat{\mathbf{e}}_i = \mathcal{G}'^i \hat{\mathbf{e}}_i \quad \text{where } \mathcal{G}'^i \equiv h'_{ij} G'^j$$

so then

$$\mathcal{G}'^i(\mathbf{x}')/h'_{ii} = [\text{grad } f]^i(\mathbf{x}') = \partial'^i f'(\mathbf{x}') = g'^{ij}(\mathbf{x}') \partial'_j f'(\mathbf{x}') \quad \mathbf{G} = \text{grad } f = \mathcal{G}'^i \hat{\mathbf{e}}_i$$

Gathering up these results one gets

$$\begin{aligned} G'_i(\mathbf{x}') &= [\text{grad } f]_i(\mathbf{x}') = \partial'_i f'(\mathbf{x}') & \mathbf{G} = \text{grad } f &= G'_i(\mathbf{x}') \mathbf{e}^i \\ G'^i(\mathbf{x}') &= [\text{grad } f]^i(\mathbf{x}') = \partial'^i f'(\mathbf{x}') = g'^{ij}(\mathbf{x}') \partial'_j f'(\mathbf{x}') & \mathbf{G} = \text{grad } f &= G'^i(\mathbf{x}') \mathbf{e}_i \\ \mathcal{G}'^i(\mathbf{x}') &= h'_{ij} [\text{grad } f]^j(\mathbf{x}') = h'_{ij} \partial'^j f'(\mathbf{x}') = h'_{ij} \mathcal{G}'^{ij}(\mathbf{x}') \partial'_j f'(\mathbf{x}') & \mathbf{G} = \text{grad } f &= \mathcal{G}'^i(\mathbf{x}') \hat{\mathbf{e}}_i \end{aligned}$$

which can be rewritten

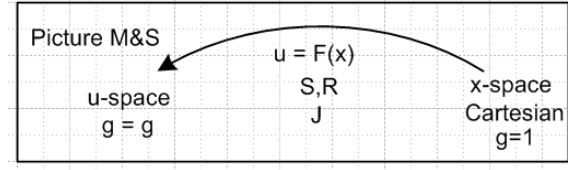
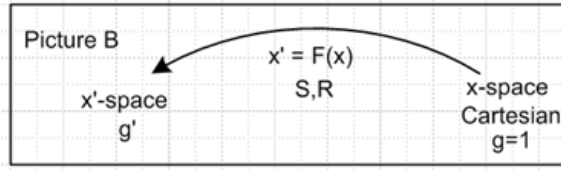
$$\begin{aligned} [\text{grad } f](\mathbf{x}) &= (\partial'_i f'(\mathbf{x}')) \mathbf{e}^i \\ [\text{grad } f](\mathbf{x}) &= (\partial'^i f'(\mathbf{x}')) \mathbf{e}_i = g'^{ij}(\mathbf{x}') (\partial'_j f'(\mathbf{x}')) \mathbf{e}_i \\ [\text{grad } f](\mathbf{x}) &= h'_{ij} (\partial'^i f'(\mathbf{x}')) \hat{\mathbf{e}}_i = h'_{ij} g'^{ij}(\mathbf{x}') (\partial'_j f'(\mathbf{x}')) \hat{\mathbf{e}}_i \\ [\text{grad } f](\mathbf{x}) &= (\partial_i f(\mathbf{x})) \hat{\mathbf{n}} = \nabla f(\mathbf{x}) \quad // \text{ Cartesian} \\ [\text{grad } f]_i(\mathbf{x}') &= R_i^j [\text{grad } f]_j(\mathbf{x}) \quad // \text{ transformation} \end{aligned} \quad f'(\mathbf{x}') = f(\mathbf{x})$$

Again, in each of the first three forms above, $\mathbf{G} = [\text{grad } f](\mathbf{x})$ is being expressed as a linear combination of $\mathbf{e}$ vectors which are in x -space, but the coefficients are given entirely in terms of x' -space coordinates and objects. Since $\text{div } \mathbf{B}$ was a scalar quantity, this mixture of vectors in x -space with components in x' -space did not arise.

For the orthogonal case $g'^{ij} = (1/h'^i_i)^2 \delta_{i,j}$ so the block above becomes

$$\begin{aligned} [\text{grad } f](\mathbf{x}) &= (\partial'_i f'(\mathbf{x}')) \mathbf{e}^i \\ [\text{grad } f](\mathbf{x}) &= (\partial'^i f'(\mathbf{x}')) \mathbf{e}_i = (1/h'^i_i)^2 (\partial'_i f'(\mathbf{x}')) \mathbf{e}_i \\ [\text{grad } f](\mathbf{x}) &= h'_{ii} (\partial'^i f'(\mathbf{x}')) \hat{\mathbf{e}}_i = (1/h'_{ii}) (\partial'_i f'(\mathbf{x}')) \hat{\mathbf{e}}_i \\ [\text{grad } f](\mathbf{x}) &= (\partial_i f(\mathbf{x})) \hat{\mathbf{n}} \quad // \text{ Cartesian} \\ [\text{grad } f]_i(\mathbf{x}') &= R_i^j [\text{grad } f]_j(\mathbf{x}) \quad // \text{ transformation} \end{aligned} \quad f'(\mathbf{x}') = f(\mathbf{x})$$

This above equations can be converted from Picture B to Picture M&S using the same rules given in Section 9 (c), which we repeat below



- replace $\mathbf{x}'$ by $\mathbf{u}$ everywhere
- replace ∂'_n by ∂_n meaning $\partial/\partial u^n$ (exception: on a "Cartesian" line ∂_n means $\partial/\partial x^n$)
- 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 $\mathbf{x}$, while all italicized tensor components are functions of $\mathbf{u}$.

The translated results are then (all implied sums)

$$\begin{aligned}
 [\text{grad } f](\mathbf{x}) &= (\partial_i f) \mathbf{e}^i \\
 [\text{grad } f](\mathbf{x}) &= (\partial^i f) \mathbf{e}_i = g^{ij} (\partial_j f) \mathbf{e}_i \\
 [\text{grad } f](\mathbf{x}) &= h_i (\partial^i f) \hat{\mathbf{e}}_i = h_i g^{ij} (\partial_j f) \hat{\mathbf{e}}_i \\
 [\text{grad } f](\mathbf{x}) &= (\partial_i f) \hat{\mathbf{n}} \quad // \text{ Cartesian} \\
 [grad f]_i(\mathbf{u}) &= R_i^j [\text{grad } f]_j(\mathbf{x}) \quad // \text{ transformation (vector)} \quad f(\mathbf{u}) = f(\mathbf{x}) = f(\mathbf{x}(\mathbf{u}))
 \end{aligned}$$

Notice that $[\text{grad } f]_i(\mathbf{x}') \rightarrow [grad f]_i(\mathbf{u})$ according to the fourth rule, meaning $G'_i(\mathbf{x}') \rightarrow G_i(\mathbf{u})$.

For *orthogonal* curvilinear coordinates,

$$\begin{aligned}
 [\text{grad } f](\mathbf{x}) &= (\partial_i f) \mathbf{e}^i \\
 [\text{grad } f](\mathbf{x}) &= (\partial^i f) \mathbf{e}_i = (1/h_i^2) (\partial_i f) \mathbf{e}_i \\
 [\text{grad } f](\mathbf{x}) &= h_i (\partial^i f) \hat{\mathbf{e}}_i = (1/h_i) (\partial_i f) \hat{\mathbf{e}} \quad // \text{ M\&S 1.05}
 \end{aligned}$$

One can always make the replacement $f(\mathbf{u}) = f(\mathbf{x}(\mathbf{u}))$ in any of the above equations (f scalar). And one more time: the various $\mathbf{e}$ vectors are in $\mathbf{x}$ -space, but all the coefficients are expressed in curvilinear $\mathbf{u}$ -space coordinates and components. With the replacements

$$f \rightarrow \varphi \quad \hat{\mathbf{e}}_i \rightarrow \mathbf{a}_i \quad h_i \rightarrow \sqrt{g_{ii}}$$

the equation marked agrees with Moon & Spencer p 2 (1.05).

(b) Expressions for $\text{grad } f \bullet \mathbf{B}$

Sometimes one is interested in the following quantity (back to Picture B)

$$\text{grad } f \bullet \mathbf{B}$$

where f is a tensorial scalar field and $\mathbf{B}$ is a tensorial vector field. In this case, since $\text{grad } f$ is a tensorial vector field, the quantity $\text{grad } f \bullet \mathbf{B}$ is a tensorial scalar, and so $(\text{grad } f)' \bullet \mathbf{B}' = (\text{grad } f) \bullet \mathbf{B}$.

This quantity $\text{grad } f \bullet \mathbf{B}$ can be written several ways depending on how $\mathbf{B}$ is expanded:

$$\begin{aligned} \mathbf{B} = \sum_i B'_i \mathbf{e}^i &\Rightarrow \text{grad } f \bullet \mathbf{B} = (\partial'^n f) \mathbf{e}_n \bullet \sum_i B'_i \mathbf{e}^i = (\partial'^n f) B'_n \\ \mathbf{B} = \sum_i B'^i \mathbf{e}_i &\Rightarrow \text{grad } f \bullet \mathbf{B} = (\partial'_n f) \mathbf{e}^n \bullet \sum_i B'^i \mathbf{e}_i = (\partial'_n f) B'^n \end{aligned}$$

The last line can be written, using $\mathcal{B}^i \equiv B'^i h'_i$,

$$\mathbf{B} = \sum_i [B'^i h'_i] \hat{\mathbf{e}}_i \equiv \sum_i \mathcal{B}^i \hat{\mathbf{e}}_i \Rightarrow \text{grad } f \bullet \mathbf{B} = (\partial'_n f) B'^n = (\partial'_n f) (\mathcal{B}^n / h'_n)$$

To summarize:

$$\begin{aligned} [\text{grad } f](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) &= (\partial'^n f(\mathbf{x}')) B'_n(\mathbf{x}') && \text{for } \mathbf{B} = \sum_i B'_i \mathbf{e}^i \\ [\text{grad } f](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) &= (\partial'_n f(\mathbf{x}')) B'^n(\mathbf{x}') && \text{for } \mathbf{B} = \sum_i B'^i \mathbf{e}_i \\ [\text{grad } f](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) &= (\partial'_n f(\mathbf{x}')) \mathcal{B}^n(\mathbf{x}') / h'_n(\mathbf{x}') && \text{for } \mathbf{B} = \sum_i \mathcal{B}^i \hat{\mathbf{e}}_i \\ [\text{grad } f](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) &= (\partial_n f(\mathbf{x})) B^n(\mathbf{x}) && \text{for } \mathbf{B} = B^n \hat{\mathbf{n}} \quad // \text{ Cartesian} \\ [\text{grad } f](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) &= [\text{grad } f](\mathbf{x}') \bullet \mathbf{B}'(\mathbf{x}') \end{aligned}$$

Since $\text{grad } f \bullet \mathbf{B}$ 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 $\mathbf{x}'$ -space coordinates and components.

The conversion from Picture B to Picture M&F is straightforward (see Sections 8 and 9)

$$\begin{aligned} [\text{grad } f](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) &= (\partial^n f) B_n && \mathbf{B} = \sum_i B_i \mathbf{e}^i \\ [\text{grad } f](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) &= (\partial_n f) B^n && \mathbf{B} = \sum_i B^i \mathbf{e}_i \\ [\text{grad } f](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) &= (\partial_n f) \mathcal{B}^n / h_n && \mathbf{B} = \sum_i \mathcal{B}^i \hat{\mathbf{e}}_i \\ [\text{grad } f](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) &= (\partial_n f) B^n && // \text{ Cartesian} \quad \mathbf{B} = B^n \hat{\mathbf{n}} \\ [\text{grad } f](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) &= [\text{grad } f](\mathbf{u}) \bullet \mathbf{B}(\mathbf{u}) && // \text{ transformation (scalar)} \end{aligned}$$

where once again $f(\mathbf{u}) = f(\mathbf{x}(\mathbf{u}))$.

Comment: According to the second line above, one can write

$$\text{grad } f \bullet d\mathbf{x} = \partial_n f(\mathbf{x}) dx^n = df = f(\mathbf{x}+d\mathbf{x}) - f(\mathbf{x})$$

This equation $df = [\text{grad } f](\mathbf{x}) \bullet d\mathbf{x}$ is sometimes used as an alternate definition of the gradient. If $d\mathbf{x}$ is selected to be in the direction of $\text{grad } f$, 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 $f(x,y)$, the gradient then points "uphill", and the negative of the gradient then points "downhill".

11. The Laplacian in curvilinear coordinates

The Laplacian (also known as the Laplace-Beltrami operator) is defined by

$$\text{lap } f = \text{div} (\text{grad } f) = \text{div } \mathbf{G} \quad \text{where} \quad \mathbf{G} \equiv \text{grad } f$$

Since f is (by assumption) a tensorial scalar field, $\text{grad } f$ is a tensorial vector. Then, as found in Section 9, $\text{div}(\text{grad } f)$ is a tensorial scalar, meaning $[\text{lap } f](\mathbf{x}) = [\text{lap } f](\mathbf{x}')$.

In Cartesian coordinates one writes

$$\text{lap } f = \nabla^2 f = \nabla \bullet \nabla f = \sum_n \partial_n^2 f$$

but this form gets modified when $\text{lap } f$ is expressed in curvilinear coordinates. Section 9 showed that

$$\text{div } \mathbf{G} = [1/\sqrt{g'}] \partial'_m [\sqrt{g'} G'^m] \quad \text{where} \quad \mathbf{G} = G'^m \mathbf{e}_m \quad g' = \det(g')$$

Section 10 showed that

$$\text{grad } f = \mathbf{G} = [g'^{nm} (\partial'_n f')] \mathbf{e}_m = G'^m \mathbf{e}_m \quad G'^m = g'^{nm} (\partial'_n f') \quad \partial'_n f' = \partial'_n f(\mathbf{x}') = \partial'_n f'(\mathbf{x}')$$

Therefore

$$\text{lap } f = \text{div} (\text{grad } f) = \text{div } \mathbf{G} = [1/\sqrt{g'}] \partial'_m [\sqrt{g'} G'^m] = [1/\sqrt{g'}] \partial'_m [\sqrt{g'} g'^{nm} (\partial'_n f')]$$

so the general results can be concisely stated:

$$\begin{aligned} [\text{lap } f](\mathbf{x}) &= [1/\sqrt{g'(\mathbf{x}')}] \partial'_m [\sqrt{g'(\mathbf{x}')} g'^{nm}(\mathbf{x}') (\partial'_n f'(\mathbf{x}'))] & // \text{ implied sum on } n \text{ and } m \\ [\text{lap } f](\mathbf{x}) &= \sum_n \partial_n^2 f(\mathbf{x}) & // \text{ Cartesian} \\ [\text{lap } f](\mathbf{x}) &= [\text{lap } f](\mathbf{x}') & f'(\mathbf{x}') = f(\mathbf{x}) = f(\mathbf{x}(\mathbf{x}')) \end{aligned}$$

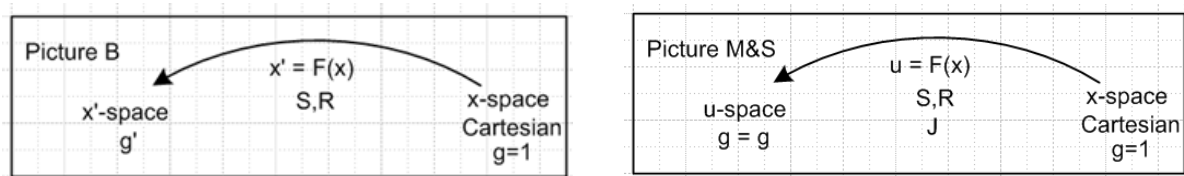
For an *orthogonal* coordinate system,

$$g'_{nm} = h'^2_n \delta_{n,m} \quad g'^{nm} = (1/h'^2_n) \delta_{n,m} \quad \sqrt{g'} = \prod_i h'_i$$

and the first line above simplifies to

$$[\text{lap } f](\mathbf{x}) = [1/(\prod_i h'_i)] \partial'_m [(\prod_i h'_i) (1/h'^2_m) (\partial'_m f')]]$$

Converting from Picture B to Picture MS gives (see Section 9 (c))



$$[\text{lap } f](\mathbf{x}) = [1/\sqrt{g}] \partial_m [\sqrt{g} g^{nm} (\partial_n f)]$$

$$[\text{lap } f](\mathbf{x}) = \sum_n \partial_n^2 f(\mathbf{x}) \quad // \text{ Cartesian}$$

$$[\text{lap } f](\mathbf{x}) = [\text{lap } f](\mathbf{u}) \quad // \text{ transformation (scalar)}$$

$$f(\mathbf{u}) = f(\mathbf{x}) = f(\mathbf{x}(\mathbf{u}))$$

The first line simplifies in the orthogonal case to

$$[\text{lap } f](\mathbf{x}) = [1/(\Pi_i h_i)] \partial_m [(\Pi_i h_i) (1/h_m^2) (\partial_m f)] \quad // \text{ orthogonal} \quad // \text{ M\&S 1.09}$$

For N=3 this says,

$$[\text{lap } f](\mathbf{x}) = 1/(h_1 h_2 h_3) \{ \partial_1 [(h_2 h_3 / h_1) \partial_1 f] + \text{cyclic} \}$$

With the replacements

$$f \rightarrow \phi \quad h_i^2 \rightarrow g_{ii} \quad (\Pi_i h_i) \rightarrow g$$

the equation marked above agrees with Moon & Spencer p 3 (1.09).

12. The Curl in curvilinear coordinates

The vector curl is defined only in N=3 dimensions (but see section (f) below). Picture B is used.

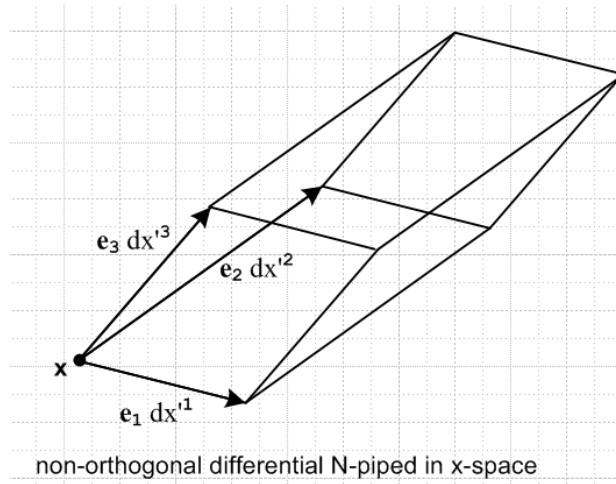
In Cartesian coordinates one writes

$$[\text{curl } \mathbf{B}]_i(\mathbf{x}) = [\nabla \times \mathbf{B}(\mathbf{x})]_i = \varepsilon_{ijk} \partial_j B_k(\mathbf{x})$$

but when expressed in terms of curvilinear coordinates and components, the form is different.

(a) Definition of curl $\mathbf{B}$

Consider the x-space differential 3-piped shown on the right side of the figure in Section 8 (a),



This 3-piped has three pairs of parallel faces. Within each pair, the "near" face touches the point $\mathbf{x}$ at which the tails of the spanning vectors meet, while the "far" face does not. As shown in Section 8 (c), the vector area associated with the face pair n is ($|J| = g'^{1/2}$ when $g = 1$)

$$d\mathbf{A}^n = |\det(S^i_j)| \mathbf{e}^n (\Pi_{i \neq n} dx^i) = |J| \mathbf{e}^n (\Pi_{i \neq n} dx^i) = g'^{1/2} \mathbf{e}^n (\Pi_{i \neq n} dx^i) \quad n = 1, 2, 3$$

where $J = \det(S^i_j)$ is the Jacobian, as in Section 5 (k), and $\mathbf{e}^n$ is a reciprocal base vector, as in Section 6 or Appendix A (called $\mathbf{E}_n$). Area $d\mathbf{A}^n$ is the out-facing vector area for the far face of pair n , while $-d\mathbf{A}^n$ is the out-facing vector for the near face.

Consider now the line integral of a vector field $\mathbf{B}(\mathbf{x})$ 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 $d\mathbf{A}^n$ which is the same as the direction of $\mathbf{e}^n$. For example, for the bottom face (near face 3) of the 3-piped shown above, this vector points "up", or toward the center of the 3-piped. Denote this line integral by

$$(\oint \mathbf{B} \cdot d\mathbf{x})^n$$

Sometimes this line integral is referred to as "the circulation" or "the rotation" of $\mathbf{B}$ around near face n (and $\text{rot } \mathbf{B}$ is another notation used for $\text{curl } \mathbf{B}$).

In x -space the quantity $\mathbf{C}(\mathbf{x}) \equiv \text{curl } \mathbf{B}(\mathbf{x})$ is a vector field defined in the following manner in the limit that all the differentials $dx'^n \rightarrow 0$:

$$\mathbf{C} \cdot d\mathbf{A}^n = (\oint \mathbf{B} \cdot d\mathbf{x})^n \quad \mathbf{C} \equiv \text{curl } \mathbf{B}$$

Since $d\mathbf{A}^n$ is given above in terms of $\mathbf{e}^n$, $\mathbf{C}$ should be expanded on the $\mathbf{e}_k$ and one writes

$$\mathbf{C} = \sum_{k=1}^3 C'^k \mathbf{e}_k$$

so that

$$\mathbf{C} \cdot d\mathbf{A}^n = (\sum_k C'^k \mathbf{e}_k) \cdot (g'^{1/2} \mathbf{e}^n (\prod_{i \neq n} dx'^i)) = C'^n g'^{1/2} (\prod_{i \neq n} dx'^i)$$

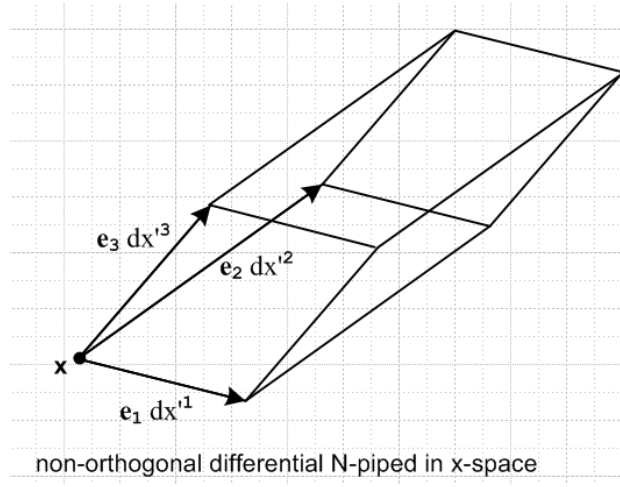
and then

$$C'^n(\mathbf{x}') \sqrt{g'(\mathbf{x}')} (\prod_{i \neq n} dx'^i) = (\oint \mathbf{B} \cdot d\mathbf{x})^n$$

Our task is to compute this line integral and thereby come up with an expression for $C'^n(\mathbf{x}')$, the components of $\text{curl } \mathbf{B}$ when $\mathbf{B}$ is expanded onto the $\mathbf{e}_k$ in x -space.

(b) Computation of the line integral

This shall be done for the bottom face ($n=3$) of the x -space differential 3-piped. Since $\mathbf{e}^3$ is "up", the circulation is a counterclockwise line integral around the boundary of the bottom face. It is useful to have the above picture near at hand to allow visualization of the four contributions to the line integral:



$$(\oint \mathbf{B} \cdot d\mathbf{s})^3 \approx [\mathbf{B}(\mathbf{x}_{\text{front}}) - \mathbf{B}(\mathbf{x}_{\text{back}})] \cdot (\mathbf{e}_1 dx^1) + [\mathbf{B}(\mathbf{x}_{\text{right}}) - \mathbf{B}(\mathbf{x}_{\text{left}})] \cdot (\mathbf{e}_2 dx^2)$$

where $\mathbf{B}(\mathbf{x}_{\text{front}})$ refers to the value of $\mathbf{B}$ 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 $dx^i \rightarrow 0$, this simple approximation of the line integral is "good enough" to produce the desired results.

Motivated by $\mathbf{e}_i \cdot \mathbf{e}^j = \delta_i^j$, expand $\mathbf{B}$ as follows

$$\mathbf{B} = B'_j \mathbf{e}^j \quad \text{where} \quad B'_j(\mathbf{x}') = \mathbf{B}(\mathbf{x}) \cdot \mathbf{e}_j$$

where the B'_j are the covariant components of $\mathbf{B}$ in x' -space. This gives

$$(\oint \mathbf{B} \cdot d\mathbf{x})^3 = [B'_1(\mathbf{x}'_{\text{front}}) - B'_1(\mathbf{x}'_{\text{back}})] dx^1 + [B'_2(\mathbf{x}'_{\text{right}}) - B'_2(\mathbf{x}'_{\text{left}})] dx^2$$

where $\mathbf{x}'_{\text{front}} = F(\mathbf{x}_{\text{front}})$ and similarly for the other three points. In x -space one has

$$d\mathbf{x}_{\text{BF}} \equiv \mathbf{x}_{\text{back}} - \mathbf{x}_{\text{front}} = \mathbf{e}_2 dx^2$$

$$d\mathbf{x}_{\text{RL}} \equiv \mathbf{x}_{\text{right}} - \mathbf{x}_{\text{left}} = \mathbf{e}_1 dx^1$$

Applying matrix R gives the corresponding x' -space equations (recall $d\mathbf{x}' = R d\mathbf{x}$ and $\mathbf{e}'_n = R\mathbf{e}_n$)

$$d\mathbf{x}'_{\text{BF}} \equiv \mathbf{x}'_{\text{back}} - \mathbf{x}'_{\text{front}} = \mathbf{e}'_2 dx'^2 \quad \mathbf{e}'_2 = (0, 1, 0, \dots)$$

$$d\mathbf{x}'_{\text{RL}} \equiv \mathbf{x}'_{\text{right}} - \mathbf{x}'_{\text{left}} = \mathbf{e}'_1 dx'^1 \quad \mathbf{e}'_1 = (1, 0, 0, \dots)$$

Using the fact that $f(\mathbf{x} + d\mathbf{x}') \approx f(\mathbf{x}') + \Sigma_n \partial_n f(\mathbf{x}') dx'^n$ one finds

$$B'_1(\mathbf{x}'_{\text{back}}) \approx B'_1(\mathbf{x}'_{\text{front}}) + (\partial B'_1 / \partial x'^2) dx'^2$$

$$B'_2(\mathbf{x}'_{\text{right}}) \approx B'_2(\mathbf{x}'_{\text{left}}) + (\partial B'_2 / \partial x'^1) dx'^1$$

$$d\mathbf{x}' = \mathbf{e}'_2 dx'^2$$

$$d\mathbf{x}' = \mathbf{e}'_1 dx'^1$$

so the circulation integral is then

$$\begin{aligned}
(\oint \mathbf{B} \cdot d\mathbf{x})^3 &\approx -(\partial B'_1 / \partial x'^2) dx'^2 dx'^1 + (\partial B'_2 / \partial x'^1) dx'^1 dx'^2 \\
&= [-(\partial B'_1 / \partial x'^2) + (\partial B'_2 / \partial x'^1)] dx'^1 dx'^2 = [-\partial'_2 B'_1 + \partial'_1 B'_2] dx'^1 dx'^2 \\
&= [\partial'_1 B'_2 - \partial'_2 B'_1] dx'^1 dx'^2 \\
&= \varepsilon^{3ab} \partial'_a B'_b (\Pi_{i \neq 3} dx'^i)
\end{aligned}$$

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 $dx'^i \rightarrow 0$)

$$(\oint \mathbf{B} \cdot d\mathbf{x})^n = \varepsilon^{nab} \partial'_a B'_b (\Pi_{i \neq n} dx'^i)$$

Appendix D discusses the tensor ε known as the Levi-Cevita 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 that $\varepsilon'^{abc\dots} = \varepsilon^{abc\dots}$ where ε' is the tensor in x' -space, as shown in Appendix D (d). In Cartesian space one can regard $\varepsilon^{abc\dots} = \varepsilon_{abc\dots}$ as a bookkeeping permutation tensor with the properties given in Section 7 (h). Installing the prime on ε ,

$$(\oint \mathbf{B} \cdot d\mathbf{x})^n = \varepsilon'^{nab} \partial'_a B'_b (\Pi_{i \neq n} dx'^i)$$

and this integral is then given entirely in terms of x' -space coordinates and objects.

(c) Solving for the curl

The equation for the curl obtained at the end of section (a) was

$$C'^n(\mathbf{x}') \sqrt{g'(\mathbf{x}')} (\Pi_{i \neq n} dx'^i) = (\oint \mathbf{B} \cdot d\mathbf{x})^n$$

Insert the section (b) result for $(\oint \mathbf{B} \cdot d\mathbf{x})^n$ to get

$$C'^n(\mathbf{x}') \sqrt{g'(\mathbf{x}')} (\Pi_{i \neq n} dx'^i) = \varepsilon'^{nab} \partial'_a B'_b (\Pi_{i \neq n} dx'^i)$$

The differentials cancel out, so *then* take $dx'^i \rightarrow 0$ and thus shrink the 3-piped around the point of interest $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$ so that

$$\begin{aligned}
C'^n &= [(1/\sqrt{g'}) \varepsilon'^{nab} \partial'_a B'_b] & \mathbf{C} &= C'^n \mathbf{e}_n & \mathbf{B} &= B'_n \mathbf{e}^n \\
\text{curl } \mathbf{B} &= \mathbf{C} = [(1/\sqrt{g'}) \varepsilon'^{nab} \partial'_a B'_b] \mathbf{e}_n
\end{aligned}$$

The comparison between the curvilinear and Cartesian expressed curls is this:

$$[\text{curl } \mathbf{B}](\mathbf{x}) = [(1/\sqrt{g'}) \varepsilon^{nab} \partial'_a B'_b(\mathbf{x}')] \mathbf{e}_n = (1/\sqrt{g'}) \{ [\partial'_1 B'_2 - \partial'_2 B'_1] \mathbf{e}_3 + \text{cyclic} \}$$

$$[\text{curl } \mathbf{B}](\mathbf{x}) = \varepsilon^{nab} \partial_a B_b(\mathbf{x}) \hat{\mathbf{n}} = \{ [\partial_1 B_2 - \partial_2 B_1] \hat{\mathbf{3}} + \text{cyclic} \}$$

Appendix D (h) shows that the vector $\mathbf{C} = \text{curl } \mathbf{B}$ really is a tensorial vector if $\mathbf{B}$ is a tensorial vector. Therefore the notation $\mathbf{C} = C^n \mathbf{e}_n$ is justified.

Comment: In the first line above one can replace $[\text{curl } \mathbf{B}](\mathbf{x})$ by $[\nabla \times \mathbf{B}](\mathbf{x})$ with the understanding that the LHS is the curl in Cartesian x-space and the RHS is expressing this LHS in terms of x'-space coordinates and objects. The RHS is certainly not equal to $\nabla' \times \mathbf{B}' = \varepsilon^{nab} \partial'_a B'_b(\mathbf{x}') \hat{\mathbf{n}}'$. It is to avoid this possible confusion that the curl is written out as the word curl, and the same comment applies to the other differential operators.

(d) Various forms of the curl

The first form is that just presented above,

$$\begin{aligned} C^n &= [(1/\sqrt{g'}) \varepsilon^{nab} \partial'_a B'_b] & \mathbf{C} &= C^n \mathbf{e}_n & \mathbf{B} &= B'^n \mathbf{e}_n \\ \text{curl } \mathbf{B} &= [(1/\sqrt{g'}) \varepsilon^{nab} \partial'_a B'_b] \mathbf{e}_n \end{aligned}$$

If it is desired to have contravariant components of $\mathbf{B}$, one gets

$$\begin{aligned} C^n &= [(1/\sqrt{g'}) \varepsilon^{nab} \partial'_a (g'_{bc} B'^c)] & \mathbf{C} &= C^n \mathbf{e}_n & \mathbf{B} &= B'^n \mathbf{e}_n \\ \text{curl } \mathbf{B} &= [(1/\sqrt{g'}) \varepsilon^{nab} \partial'_a (g'_{bc} B'^c)] \mathbf{e}_n \end{aligned}$$

For practical applications, one usually wants both vectors expanded on the $\hat{\mathbf{e}}_n$ unit vectors in this way

$$\begin{aligned} \mathbf{C} &= C^n \mathbf{e}_n = (C^n h'_n) \hat{\mathbf{e}}_n \equiv \mathcal{C}^n \hat{\mathbf{e}}_n & \mathcal{C}^n &= h'_n C^n \\ \mathbf{B} &= B'^n \mathbf{e}_n = (B'^n h'_n) \hat{\mathbf{e}}_n \equiv \mathcal{B}^n \hat{\mathbf{e}}_n & \mathcal{B}^n &= h'_n B'^n \end{aligned}$$

so that

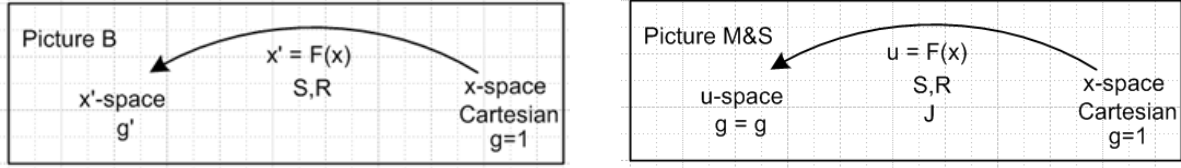
$$\begin{aligned} \mathcal{C}^n &= [(1/\sqrt{g'}) h'_n \varepsilon^{nab} \partial'_a (g'_{bc} b'^c/h'_c)] & \mathbf{C} &= \mathcal{C}^n \hat{\mathbf{e}}_n & \mathbf{B} &= \mathcal{B}^n \hat{\mathbf{e}}_n \\ \text{curl } \mathbf{B} &= [(1/\sqrt{g'}) h'_n \varepsilon^{nab} \partial'_a (g'_{bc} b'^c/h'_c)] \hat{\mathbf{e}}_n & \text{curl } \mathbf{B} &= \mathbf{C} \end{aligned}$$

To summarize: $B'^c = R^c_d B^d$ $g' = g'(\mathbf{x}')$ etc.

$$\begin{aligned} [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} [(1/\sqrt{g'}) \partial'_a B'_b] \mathbf{e}_n & \mathbf{B} &= B'^n \mathbf{e}_n \\ [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} [(1/\sqrt{g'}) \partial'_a (g'_{bc} B'^c)] \mathbf{e}_n & \mathbf{B} &= B'^n \mathbf{e}_n \end{aligned}$$

$$\begin{aligned}
[\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} [(1/\sqrt{g'}) h'_n \partial'_a (g'_{bc} \mathcal{B}^c / h'_c)] \hat{\mathbf{e}}_n \\
[\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} \partial_a B_b(\mathbf{x}) \hat{\mathbf{n}} \quad // \text{ Cartesian} & \mathbf{B} = \mathcal{B}^n \hat{\mathbf{e}}_n \\
& & \mathbf{B} = B^n \hat{\mathbf{n}}
\end{aligned}$$

Converting from Picture B to Picture MS (see Section 9 (c))



one gets :

$$\begin{aligned}
[\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} [(1/\sqrt{g}) \partial_a B_b] \mathbf{e}_n & \mathbf{B} &= B_n \mathbf{e}^n \\
[\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} [(1/\sqrt{g}) \partial_a (g_{bc} B^c)] \mathbf{e}_n & \mathbf{B} &= B^n \mathbf{e}_n \\
[\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} [(1/\sqrt{g}) h_n \partial_a (g_{bc} \mathcal{B}^c / h_c)] \hat{\mathbf{e}}_n & \mathbf{B} &= \mathcal{B}^n \hat{\mathbf{e}}_n \\
[\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} \partial_a B_b(\mathbf{x}) \hat{\mathbf{n}} \quad // \text{ Cartesian} & \mathbf{B} &= B^n \hat{\mathbf{n}}
\end{aligned}$$

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.

Each of the above results can be written as a determinant using the idea $\det(\mathbf{Q}) \equiv \Sigma_i \mathbf{Q}_{1i} \text{cof}(\mathbf{Q}_{1i})$:

$$\begin{aligned}
[\text{curl } \mathbf{B}](\mathbf{x}) &= (1/\sqrt{g}) \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ B_1 & B_2 & B_3 \end{vmatrix} & \mathbf{B} &= B_n \mathbf{e}^n \\
[\text{curl } \mathbf{B}](\mathbf{x}) &= (1/\sqrt{g}) \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ g_{1c} B^c & g_{2c} B^c & g_{3c} B^c \end{vmatrix} & \mathbf{B} &= B^n \mathbf{e}_n \\
[\text{curl } \mathbf{B}](\mathbf{x}) &= (1/\sqrt{g}) \begin{vmatrix} h_1 \hat{\mathbf{e}}_1 & h_2 \hat{\mathbf{e}}_2 & h_3 \hat{\mathbf{e}}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ (g_{1c}/h_c) \mathcal{B}^c & (g_{2c}/h_c) \mathcal{B}^c & (g_{3c}/h_c) \mathcal{B}^c \end{vmatrix} & \mathbf{B} &= \mathcal{B}^n \hat{\mathbf{e}}_n \quad // \text{ M\&S 1.07} \\
[\text{curl } \mathbf{B}](\mathbf{x}) &= \begin{vmatrix} \hat{\mathbf{1}} & \hat{\mathbf{2}} & \hat{\mathbf{3}} \\ \partial_1 & \partial_2 & \partial_3 \\ B_1 & B_2 & B_3 \end{vmatrix} \quad // \text{ here } \partial_n = \partial/\partial x^n \text{ and } B_i = B_i(\mathbf{x}) & \mathbf{B} &= B^n \hat{\mathbf{n}}
\end{aligned}$$

(e) 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: $\mathbf{B}'^c(\mathbf{x}') = \mathbf{R}^c{}_d \mathbf{B}^d(\mathbf{x})$ $h_i' = h_i'(\mathbf{x}') \text{ etc.}$

$$\begin{aligned} [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon'^{nab} [(1/(h_1'h_2'h_3')) \partial'_a B'_b] \mathbf{e}_n & \mathbf{B} &= B'_n \mathbf{e}_n \\ [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon'^{nab} [(1/(h_1'h_2'h_3')) \partial'_a (h_b'^2 B'^b)] \mathbf{e}_n & \mathbf{B} &= B'^n \mathbf{e}_n \\ [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon'^{nab} [(1/(h_1'h_2'h_3')) h'_n \partial'_a (h'_b \mathcal{B}^b)] \hat{\mathbf{e}}_n & \mathbf{B} &= \mathcal{B}^n \hat{\mathbf{e}}_n \\ [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon'^{nab} \partial_a B_b(\mathbf{x}) \hat{\mathbf{n}} & // & \text{Cartesian} & \mathbf{B} &= B^n \hat{\mathbf{n}} \end{aligned}$$

Picture M&S: $\mathcal{B}^c(\mathbf{u}) = \mathbf{R}^c{}_d \mathbf{B}^d(\mathbf{x})$ $h_i = h_i(\mathbf{u}) \text{ etc.}$

$$\begin{aligned} [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} [(h_1 h_2 h_3)^{-1} \partial_a B_b] \mathbf{e}_n & \mathbf{B} &= B_n \mathbf{e}_n \\ [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} [(h_1 h_2 h_3)^{-1} \partial_a (h_b^2 B^b)] \mathbf{e}_n & \mathbf{B} &= B^n \mathbf{e}_n \\ [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} [(h_1 h_2 h_3)^{-1} h_n \partial_a (h_b \mathcal{B}^b)] \hat{\mathbf{e}}_n & \mathbf{B} &= \mathcal{B}^n \hat{\mathbf{e}}_n \end{aligned}$$

$$[\text{curl } \mathbf{B}](\mathbf{x}) = (h_1 h_2 h_3)^{-1} \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ B_1 & B_2 & B_3 \end{vmatrix} \quad \mathbf{B} = B_n \mathbf{e}^n$$

$$[\text{curl } \mathbf{B}](\mathbf{x}) = (h_1 h_2 h_3)^{-1} \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ h_1^2 B^1 & h_2^2 B^2 & h_3^2 B^3 \end{vmatrix} \quad \mathbf{B} = B^n \mathbf{e}_n$$

$$[\text{curl } \mathbf{B}](\mathbf{x}) = (h_1 h_2 h_3)^{-1} \begin{vmatrix} h_1 \hat{\mathbf{e}}_1 & h_2 \hat{\mathbf{e}}_2 & h_3 \hat{\mathbf{e}}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ h_1 \mathcal{B}^1 & h_2 \mathcal{B}^2 & h_3 \mathcal{B}^3 \end{vmatrix} \quad \mathbf{B} = \mathcal{B}^n \hat{\mathbf{e}}_n \quad // \text{ M\&S 1.07a}$$

With the replacements

$$\mathbf{B} \rightarrow \mathbf{E} \quad \mathcal{B}^n \rightarrow E_n \quad \hat{\mathbf{e}}_n \rightarrow \mathbf{a}_n \quad h_i \rightarrow \sqrt{g_{ii}} \quad (h_1 h_2 h_3)^{-1} \rightarrow (1/\sqrt{g})$$

the equations marked above agree with Moon & Spencer p 2 (1.07) and p 3 (1.07a).

(f) The curl in $N > 3$ dimensions

Looking at the basic form of the curl above

$$[\text{curl } \mathbf{B}]^n(\mathbf{x}) = \varepsilon^{nab} \partial_a B_b(\mathbf{x}) \quad // \text{ Cartesian}$$

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

$$[\text{curl } \mathbf{B}]^n(\mathbf{x}) = (1/\sqrt{g}) \varepsilon^{nabc} \partial_a (\partial_b B_c) = (1/\sqrt{g}) \varepsilon^{nabc} \partial_a \partial_b B_c \quad ?$$

$$[\text{curl } \mathbf{B}]^n(\mathbf{x}) = (1/\sqrt{g}) \varepsilon^{nabc} \partial_a (B_b B_c) \quad ?$$

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.

The rank-2 tensor $B_{\mathbf{b};\mathbf{a}} - B_{\mathbf{a};\mathbf{b}} = \partial_{\mathbf{a}}B_{\mathbf{b}} - \partial_{\mathbf{b}}B_{\mathbf{a}}$ discussed in Appendix D (h) 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,

$$[\text{curl } \mathbf{B}]^{\mathbf{n}} = \epsilon^{\mathbf{nab}} [B_{\mathbf{b};\mathbf{a}} - B_{\mathbf{a};\mathbf{b}}]/2 = \epsilon^{\mathbf{nab}} B_{\mathbf{b};\mathbf{a}} = \epsilon^{\mathbf{nab}} [\partial_{\mathbf{a}}B_{\mathbf{b}} - \partial_{\mathbf{b}}B_{\mathbf{a}}]/2 = \epsilon^{\mathbf{nab}} \partial_{\mathbf{a}}B_{\mathbf{b}}.$$

In relativity work, since $N=4$, there *is* no vector curl, and one sees $B_{\mathbf{b};\mathbf{a}} - B_{\mathbf{a};\mathbf{b}}$ referred to as the covariant curl, and $\partial_{\mathbf{a}}B_{\mathbf{b}} - \partial_{\mathbf{b}}B_{\mathbf{a}}$ as the ordinary curl (Weinberg p 106).

Writing the N -dimensional covariant curl in this manner

$$[\text{curl } B]_{\mathbf{ij}}(\mathbf{x}) = [B_{\mathbf{j};\mathbf{i}}(\mathbf{x}) - B_{\mathbf{i};\mathbf{j}}(\mathbf{x})]$$

one can ask how this same curl would be expressed in terms of $\mathbf{x}'$ -space coordinates and objects. In analogy with $N=3$ curl above, one might conjecture that

$$[\text{curl } B]_{\mathbf{ij}}(\mathbf{x}) = [B'_{\mathbf{b};\mathbf{a}}(\mathbf{x}') - B'_{\mathbf{a};\mathbf{b}}(\mathbf{x}')] (\mathbf{e}^{\mathbf{a}})_{\mathbf{i}} (\mathbf{e}^{\mathbf{b}})_{\mathbf{j}}$$

where one could regard $\mathbf{e}^{\mathbf{ab}}_{\mathbf{ij}} \equiv (\mathbf{e}^{\mathbf{a}})_{\mathbf{i}} (\mathbf{e}^{\mathbf{b}})_{\mathbf{j}}$ as a basis vector of the direct product space $\bar{\mathcal{R}} \times \bar{\mathcal{R}}$ mentioned in Section 7 (j). The conjecture can be quickly verified. Start with the transformation rule for a rank-2 covariant tensor, which is what $[B_{\mathbf{j};\mathbf{i}}(\mathbf{x}) - B_{\mathbf{i};\mathbf{j}}(\mathbf{x})]$ is,

$$[B'_{\mathbf{b};\mathbf{a}}(\mathbf{x}') - B'_{\mathbf{a};\mathbf{b}}(\mathbf{x}')] = R_{\mathbf{a}}^{\mathbf{i}} R_{\mathbf{b}}^{\mathbf{j}} [B_{\mathbf{j};\mathbf{i}}(\mathbf{x}) - B_{\mathbf{i};\mathbf{j}}(\mathbf{x})] = R_{\mathbf{a}}^{\mathbf{i}} R_{\mathbf{b}}^{\mathbf{j}} [\text{curl } B]_{\mathbf{ij}}$$

Apply $R^{\mathbf{a}}_{\mathbf{a}'} R^{\mathbf{b}}_{\mathbf{b}'}$ to the above line,

$$\begin{aligned} R^{\mathbf{a}}_{\mathbf{a}'} R^{\mathbf{b}}_{\mathbf{b}'} [B'_{\mathbf{b};\mathbf{a}}(\mathbf{x}') - B'_{\mathbf{a};\mathbf{b}}(\mathbf{x}')] &= R^{\mathbf{a}}_{\mathbf{a}'} R^{\mathbf{b}}_{\mathbf{b}'} R_{\mathbf{a}}^{\mathbf{i}} R_{\mathbf{b}}^{\mathbf{j}} [\text{curl } B]_{\mathbf{ij}} \\ &= (R^{\mathbf{a}}_{\mathbf{a}'} R_{\mathbf{a}}^{\mathbf{i}}) (R^{\mathbf{b}}_{\mathbf{b}'} R_{\mathbf{b}}^{\mathbf{j}}) [\text{curl } B]_{\mathbf{ij}} = \delta_{\mathbf{a}'}^{\mathbf{i}} \delta_{\mathbf{b}'}^{\mathbf{j}} [\text{curl } B]_{\mathbf{ij}} = [\text{curl } B]_{\mathbf{a}'\mathbf{b}'} \end{aligned}$$

so that (See Section 7 (s))

$$[\text{curl } B]_{\mathbf{a}'\mathbf{b}'} = R^{\mathbf{a}}_{\mathbf{a}'} R^{\mathbf{b}}_{\mathbf{b}'} [B'_{\mathbf{b};\mathbf{a}}(\mathbf{x}') - B'_{\mathbf{a};\mathbf{b}}(\mathbf{x}')] = (\mathbf{e}^{\mathbf{a}})_{\mathbf{a}'} (\mathbf{e}^{\mathbf{b}})_{\mathbf{b}'} [B'_{\mathbf{b};\mathbf{a}}(\mathbf{x}') - B'_{\mathbf{a};\mathbf{b}}(\mathbf{x}')]_{\mathbf{a}'\mathbf{b}'}$$

and changing index names,

$$[\text{curl } B]_{\mathbf{ij}} = (\mathbf{e}^{\mathbf{a}})_{\mathbf{i}} (\mathbf{e}^{\mathbf{b}})_{\mathbf{j}} [B'_{\mathbf{b};\mathbf{a}}(\mathbf{x}') - B'_{\mathbf{a};\mathbf{b}}(\mathbf{x}')] \quad \text{QED.}$$

13. The Vector Laplacian in curvilinear coordinates

This operator is defined in terms of the vector curl which is only defined for $N=3$.
The context is Picture B.

(a) Derivation of the Vector Laplacian in general curvilinear coordinates

The definition of the vector Laplacian of a vector field $\mathbf{B}(\mathbf{x})$ is

$$\nabla^2 \mathbf{B} \equiv \text{grad}(\text{div } \mathbf{B}) - \text{curl}(\text{curl } \mathbf{B}),$$

so, as expected, the vector Laplacian is a vector field. In Cartesian coordinates, one finds that

$$[\nabla^2 \mathbf{B}]^i = \nabla^2 (B^i) \equiv \Sigma_n \partial_n^2 B^i$$

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, M&S use $\star$ in place of ∇^2 and we will honor these authors by using that symbol here, so

$$\star \mathbf{B} \equiv \text{grad}(\text{div } \mathbf{B}) - \text{curl}(\text{curl } \mathbf{B})$$

In order to make use of the results of earlier sections, define

$$\mathbf{G} \equiv \text{grad}(f) \quad \text{where } f = \text{div } \mathbf{B}$$

$$\mathbf{V} \equiv \text{curl } \mathbf{C} \quad \text{where } \mathbf{C} \equiv \text{curl } \mathbf{B}$$

so that

$$\star \mathbf{B} = \mathbf{G} - \mathbf{V}$$

Section 10 (c) gives this expression for $\mathbf{G}$,

$$\mathbf{G} = \text{grad}(f) = (\partial'_k f') \mathbf{e}^k$$

in which expression Section 9 (b) allows replacement of f' as follows,

$$f' = f'(\mathbf{x}') = f(\mathbf{x}) = \text{div } \mathbf{B} = [1/\sqrt{g'}] \partial'_i [\sqrt{g'} B'^i]$$

so that

$$\Rightarrow \mathbf{G} = \partial'_k \{ (1/\sqrt{g'}) \partial'_i (\sqrt{g'} B'^i) \} \mathbf{e}^k \quad // \text{ 2 implied sums, } i \text{ and } k$$

The second term $\mathbf{V}$ is little more complicated. First, from Section 12 (d),

$$\mathbf{C} = \text{curl } \mathbf{B} = \varepsilon^{nab} [(1/\sqrt{g'}) \partial'_a \{B'_b\}] \mathbf{e}_n = C'^n \mathbf{e}_n$$

$$\mathbf{V} = \text{curl } \mathbf{C} = \varepsilon^{mcd} [(1/\sqrt{g'}) \partial'_c \{C'_d\}] \mathbf{e}_n = V'^m \mathbf{e}_n$$

where recall from Appendix D (d) that $\varepsilon^{abc\dots} = \varepsilon^{abc\dots} = \varepsilon_{abc\dots}$ = the usual permutation tensor, but written up and primed so as to be in covariant form. Appendix D (h) shows that $\mathbf{C}$ and $\mathbf{V}$ are both true contravariant vectors, with components C'^n and V'^n in x' -space.

The first line above says $C'^e = \varepsilon^{eab} [(1/\sqrt{g'}) \partial'_a B'_b]$ so that

$$C'_d = g'_{de} C'^e = g'_{de} \varepsilon^{eab} [(1/\sqrt{g'}) \partial'_a B'_b],$$

which can then be inserted then into the second line to get

$$\begin{aligned} \mathbf{V} = \text{curl } \mathbf{C} &= \varepsilon^{mcd} [(1/\sqrt{g'}) \partial'_c \{ g'_{de} \varepsilon^{eab} [(1/\sqrt{g'}) \partial'_a B'_b] \}] \mathbf{e}_n = V'^m \mathbf{e}_n \\ &= (1/\sqrt{g'}) \varepsilon^{mcd} \varepsilon^{eab} \partial'_c \{ (1/\sqrt{g'}) g'_{de} (\partial'_a B'_b) \} \mathbf{e}_n \quad // \text{ note } g'_{dc} = g'_{dc}(\mathbf{x}') , \text{ etc} \end{aligned}$$

which has 6 implied sums, a,b,c,d,e, and n. For each value of n, there are not really 3^5 terms because most terms vanish due to the ε factors. Looking at $\varepsilon^{mcd} \varepsilon^{eab} = \varepsilon^{mcd} \varepsilon^{eab}$, one sees that for each n, the c and d sums generate only 2 terms, and for each of these ε^{eab} generates $3*2*1 = 6$ terms, so there are 12 terms total for each n. Later when g'_{de} is assumed diagonal, the effective factor is $\varepsilon^{mcd} \varepsilon^{eab}$ implying $2 * (2*1) = 4$ terms, which shall be written out in that case.

Combining these terms, the vector Laplacian is now

$$\star \mathbf{B} = \mathbf{G} - \mathbf{V} =$$

$$\partial'_k \{ (1/\sqrt{g'}) \partial'_i (\sqrt{g'} B'^i) \} \mathbf{e}^k - (1/\sqrt{g'}) \varepsilon^{mcd} \varepsilon^{eab} \partial'_c \{ (1/\sqrt{g'}) g'_{de} (\partial'_a B'_b) \} \mathbf{e}_n$$

Setting $\mathbf{e}^k = g'^{kn} \mathbf{e}_n$ in the first term gives

$$\begin{aligned} \star \mathbf{B} &= \partial'_k \{ (1/\sqrt{g'}) \partial'_i (\sqrt{g'} B'^i) \} g'^{kn} \mathbf{e}_n - (1/\sqrt{g'}) \varepsilon^{mcd} \varepsilon^{eab} \partial'_c \{ (1/\sqrt{g'}) g'_{de} (\partial'_a B'_b) \} \mathbf{e}_n \\ &= \mathbf{e}_n [g'^{kn} \partial'_k \{ (1/\sqrt{g'}) \partial'_i (\sqrt{g'} B'^i) \} - (1/\sqrt{g'}) \varepsilon^{mcd} \varepsilon^{eab} \partial'_c \{ (1/\sqrt{g'}) g'_{de} (\partial'_a B'_b) \}] \end{aligned}$$

so at least now both terms use the same expansion base vector $\mathbf{e}_n$. As a next step, $\mathbf{e}_n = h'_n \hat{\mathbf{e}}_n$ so

$$\star \mathbf{B} = h'_n \hat{\mathbf{e}}_n [g'^{kn} \partial'_k \{ (1/\sqrt{g'}) \partial'_i (\sqrt{g'} B'^i) \} - (1/\sqrt{g'}) \varepsilon^{mcd} \varepsilon^{eab} \partial'_c \{ (1/\sqrt{g'}) g'_{de} (\partial'_a B'_b) \}]$$

The component B'_b in the second term can be made contravariant using $B'_b = g'_{bf} B'^f$ to get

$$\star \mathbf{B} = h'_n \hat{\mathbf{e}}_n [g'^{kn} \partial'_k \{ (1/\sqrt{g'}) \partial'_i (\sqrt{g'} B^i) \} - (1/\sqrt{g'}) \varepsilon^{ncd} \varepsilon^{eab} \partial'_c \{ (1/\sqrt{g'}) g'_{de} (\partial'_a [g'_{bf} B^f]) \}]$$

and then, as was done in earlier sections, replace

$$B^m = b^m/h'_n \quad \mathbf{B} = \mathcal{B}^n \hat{\mathbf{e}}_n$$

to get this final form in "practical units",

$$\star \mathbf{B} = h'_n \hat{\mathbf{e}}_n [g'^{kn} \partial'_k \{ (1/\sqrt{g'}) \partial'_i (\sqrt{g'} \mathcal{B}^i/h'_i) \} - (1/\sqrt{g'}) \varepsilon^{ncd} \varepsilon^{eab} \partial'_c \{ (1/\sqrt{g'}) g'_{de} (\partial'_a [g'_{bf} \mathcal{B}^f/h'_f]) \}]$$

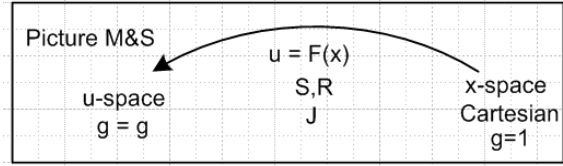
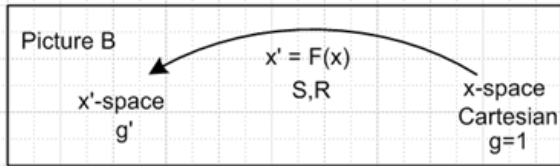
There are so many options here it is difficult to summarize, but here are two forms from above:

$$[\star \mathbf{B}](\mathbf{x}) = \mathbf{e}_n [g'^{kn} \partial'_k \{ (1/\sqrt{g'}) \partial'_i (\sqrt{g'} B^i(\mathbf{x})) \} - (1/\sqrt{g'}) \varepsilon^{ncd} \varepsilon^{eab} \partial'_c \{ (1/\sqrt{g'}) g'_{de} (\partial'_a [g'_{bf} B^f(\mathbf{x})]) \}] \quad \mathbf{B} = B^n \mathbf{e}_n$$

$$[\star \mathbf{B}](\mathbf{x}) = \hat{\mathbf{e}}_n h'_n [g'^{kn} \partial'_k \{ (1/\sqrt{g'}) \partial'_i (\sqrt{g'} \mathcal{B}^i(\mathbf{x})/h'_i) \} - (1/\sqrt{g'}) \varepsilon^{ncd} \varepsilon^{eab} \partial'_c \{ (1/\sqrt{g'}) g'_{de} (\partial'_a [g'_{bf} \mathcal{B}^f(\mathbf{x})/h'_f]) \}] \quad \mathbf{B} = \mathcal{B}^n \hat{\mathbf{e}}_n$$

$$[\star \mathbf{B}](\mathbf{x}) = \nabla^2(\mathbf{B}(\mathbf{x})) \quad // \text{ Cartesian, meaning } [\star \mathbf{B}]^i(\mathbf{x}) = \nabla^2 B^i(\mathbf{x}) \quad \mathbf{B} = B^n \hat{\mathbf{n}}$$

Converting from Picture B to Picture MS gives (see Section 9 (c))



$$[\star \mathbf{B}](\mathbf{x}) = \mathbf{e}_n [g^{kn} \partial_k \{ (1/\sqrt{g}) \partial_i (\sqrt{g} B^i) \} - (1/\sqrt{g}) \varepsilon^{ncd} \varepsilon^{eab} \partial_c \{ (1/\sqrt{g}) g_{de} (\partial_a [g_{bf} B^f]) \}] \quad \mathbf{B} = B^n \mathbf{e}_n$$

$$[\star \mathbf{B}](\mathbf{x}) = \hat{\mathbf{e}}_n h_n [g^{kn} \partial_k \{ (1/\sqrt{g}) \partial_i (\sqrt{g} \mathcal{B}^i/h_i) \} - (1/\sqrt{g}) \varepsilon^{ncd} \varepsilon^{eab} \partial_c \{ (1/\sqrt{g}) g_{de} (\partial_a [g_{bf} \mathcal{B}^f/h_f]) \}] \quad \mathbf{B} = \mathcal{B}^n \hat{\mathbf{e}}_n$$

$$[\star \mathbf{B}](\mathbf{x}) = \nabla^2(\mathbf{B}(\mathbf{x})) \quad // \text{ Cartesian, meaning } [\star \mathbf{B}]^i(\mathbf{x}) = \nabla^2[B^i(\mathbf{x})] \quad \mathbf{B} = B^n \hat{\mathbf{n}}$$

In the first two equations above, all the ∂_i mean $\partial/\partial u^i$ and the argument $\mathbf{u}$ of all functions is suppressed. The Cartesian form will be verified below.

(b) The Vector Laplacian in orthogonal curvilinear coordinates

We continue in Picture M&A and process only the second equation of the above block, since it is the one with the practical components $\mathbf{B} = \mathcal{B}^n \hat{\mathbf{e}}_n$. Setting $g_{ij} = h_i^2 \delta_{i,j}$ and $g^{ij} = (1/h_i)^2 \delta_{i,j}$ things simplify somewhat

$$\begin{aligned}
 \star \mathbf{B} &= h_n \hat{\mathbf{e}}_n [g^{kn} \partial_k \{ (1/\sqrt{g}) \partial_i (\sqrt{g} \mathcal{B}^i(\mathbf{u})/h_i) \} \\
 &\quad - (1/\sqrt{g}) \varepsilon^{ncd} \varepsilon^{eab} \partial_c \{ (1/\sqrt{g}) g_{de} (\partial_a [g_{bf} \mathcal{B}^f(\mathbf{u})/h_f]) \}] \\
 &= h_n \hat{\mathbf{e}}_n [\delta_{k,n} \partial_k \{ (1/\sqrt{g}) \partial_i (\sqrt{g} \mathcal{B}^i(\mathbf{u})/h_i) \} (1/h_n)^2 \\
 &\quad - (1/\sqrt{g}) \varepsilon^{ncd} \varepsilon^{eab} \partial_c \{ (1/\sqrt{g}) h_d^2 \delta_{d,e} (\partial_a [h_b^2 \delta_{b,f} \mathcal{B}^f(\mathbf{u})/h_f]) \}] \\
 &= h_n \hat{\mathbf{e}}_n [\partial_n \{ (1/\sqrt{g}) \partial_i (\sqrt{g} \mathcal{B}^i(\mathbf{u})/h_i) \} (1/h_n)^2 \\
 &\quad - (1/\sqrt{g}) \varepsilon^{ncd} \varepsilon^{eab} \partial_c \{ (1/\sqrt{g}) h_d^2 (\partial_a [h_b \mathcal{B}^b(\mathbf{u})]) \}] \\
 &= \hat{\mathbf{e}}_n [(1/h_n) \partial_n \{ (1/\sqrt{g}) \partial_i (\sqrt{g} \mathcal{B}^i/h_i) \} - (h_n/\sqrt{g}) \varepsilon^{ncd} \varepsilon^{eab} \partial_c \{ (1/\sqrt{g}) h_d^2 (\partial_a [h_b \mathcal{B}^b]) \}]
 \end{aligned}$$

The first term can be written as

$$\hat{\mathbf{e}}_n (1/h_n) \partial_n T \quad \text{where } T = (1/\sqrt{g}) \partial_i (\sqrt{g} \mathcal{B}^i/h_i)$$

To expand the second term, set $n = 1$ and then write things out explicitly. For the moment, we suppress the leading factor $-(h_1/\sqrt{g})$ and write

$$\begin{aligned}
 &\varepsilon^{1cd} \varepsilon^{dab} \partial_c \{ (1/\sqrt{g}) h_d^2 (\partial_a [h_b \mathcal{B}^b]) \} \\
 &= \varepsilon^{3ab} \partial_2 \{ (1/\sqrt{g}) h_3^2 (\partial_a [h_b \mathcal{B}^b]) \} - \varepsilon^{2ab} \partial_3 \{ (1/\sqrt{g}) h_2^2 (\partial_a [h_b \mathcal{B}^b]) \} \\
 &\quad \quad \quad c=2 \quad d=3 \quad \quad \quad c=3 \quad d=2 \\
 &= [\partial_2 \{ (1/\sqrt{g}) h_3^2 (\partial_1 [h_2 \mathcal{B}^2]) \} - \partial_2 \{ (1/\sqrt{g}) h_3^2 (\partial_2 [h_1 \mathcal{B}^1]) \}] \\
 &\quad \quad \quad a=1 \quad b=2 \quad \quad \quad a=2 \quad b=1 \\
 &- [\partial_3 \{ (1/\sqrt{g}) h_2^2 (\partial_3 [h_1 \mathcal{B}^1]) \} - \partial_3 \{ (1/\sqrt{g}) h_2^2 (\partial_1 [h_3 \mathcal{B}^3]) \}] \\
 &\quad \quad \quad a=3 \quad b=1 \quad \quad \quad a=1 \quad b=3 \\
 &= \partial_2 \{ (1/\sqrt{g}) h_3^2 (\partial_1 [h_2 \mathcal{B}^2] - \partial_2 [h_1 \mathcal{B}^1]) \} \\
 &- \partial_3 \{ (1/\sqrt{g}) h_2^2 (\partial_3 [h_1 \mathcal{B}^1] - \partial_1 [h_3 \mathcal{B}^3]) \}
 \end{aligned}$$

Define now Γ_n in cyclic fashion.

$$\Gamma_1 \equiv (1/\sqrt{g}) h_1^2 (\partial_2 [h_3 \mathcal{B}^3] - \partial_3 [h_2 \mathcal{B}^2])$$

$$\begin{aligned}\Gamma_2 &\equiv (1/\sqrt{g}) h_2^2 (\partial_3 [h_1 \mathcal{B}^1] - \partial_1 [h_3 \mathcal{B}^3]) \\ \Gamma_3 &\equiv (1/\sqrt{g}) h_3^2 (\partial_1 [h_2 \mathcal{B}^2] - \partial_2 [h_1 \mathcal{B}^1])\end{aligned}$$

and then we have shown that

$$\begin{aligned}\text{2nd term (n=1)} &= -(h_1/\sqrt{g}) \varepsilon^{1cd} \varepsilon^{dab} \partial_c \{ (1/\sqrt{g}) h_d^2 (\partial_a [h_b \mathcal{B}^b]) \} \hat{e}_1 \\ &= -(h_1/\sqrt{g}) (\partial_2 \Gamma_3 - \partial_3 \Gamma_2) \hat{e}_1 = + (h_1/\sqrt{g}) (\partial_3 \Gamma_2 - \partial_2 \Gamma_3) \hat{e}_1\end{aligned}$$

Therefore the entire first term (n=1) of $\star \mathbf{B}$ is given by

$$\star \mathbf{B} \text{ (first term)} = [(1/h_1) \partial_1 T + (h_1/\sqrt{g}) (\partial_3 \Gamma_2 - \partial_2 \Gamma_3)] \hat{e}_1$$

The other two terms are obtained by cyclic permutation so the final result is then

$$\begin{aligned}[\star \mathbf{B}](\mathbf{x}) &= [(1/h_1) \partial_1 T + (h_1/\sqrt{g}) (\partial_3 \Gamma_2 - \partial_2 \Gamma_3)] \hat{e}_1 + \text{cyclic} \\ &= [(1/h_1) \partial_1 T + (h_1/\sqrt{g}) (\partial_3 \Gamma_2 - \partial_2 \Gamma_3)] \hat{e}_1 \\ &\quad + [(1/h_2) \partial_2 T + (h_2/\sqrt{g}) (\partial_1 \Gamma_3 - \partial_3 \Gamma_1)] \hat{e}_2 \\ &\quad + [(1/h_3) \partial_3 T + (h_3/\sqrt{g}) (\partial_2 \Gamma_1 - \partial_1 \Gamma_2)] \hat{e}_3\end{aligned} \quad // \text{ M\&S 1.11}$$

where

$$\begin{aligned}T &= (1/\sqrt{g}) \partial_i (\sqrt{g} \mathcal{B}^i/h_i) \\ \Gamma_1 &= (1/\sqrt{g}) h_1^2 (\partial_2 [h_3 \mathcal{B}^3] - \partial_3 [h_2 \mathcal{B}^2]) \\ \Gamma_2 &= (1/\sqrt{g}) h_2^2 (\partial_3 [h_1 \mathcal{B}^1] - \partial_1 [h_3 \mathcal{B}^3]) \\ \Gamma_3 &= (1/\sqrt{g}) h_3^2 (\partial_1 [h_2 \mathcal{B}^2] - \partial_2 [h_1 \mathcal{B}^1])\end{aligned}$$

With the replacements

$$\mathbf{B} \rightarrow \mathbf{E} \quad \mathcal{B}^n \rightarrow E_n \quad \hat{e}_n \rightarrow \mathbf{a}_n \quad h_i \rightarrow \sqrt{g_{ii}} \quad T \rightarrow Y$$

the result agrees with M&S p 3 (1.11). A more compact summary is this:

$$\begin{aligned}\star \mathbf{B} &= [(1/h_n) \partial_n T - (h_n/\sqrt{g}) \varepsilon^{nab} \partial_a \Gamma_b] \hat{e}_n \\ T &= (1/\sqrt{g}) \partial_i (\sqrt{g} \mathcal{B}^i(\mathbf{u})/h_i) \\ \Gamma_b &= (1/\sqrt{g}) h_b^2 \varepsilon^{bcd} (\partial_c [h_d \mathcal{B}^d(\mathbf{u})])\end{aligned}$$

(c) The Vector Laplacian in Cartesian coordinates

First, one can verify that the last result of section (b) gives the starting point formula for $\star \mathbf{B}$ if $g = 1$ (in u-space). One then has,

$$h_i = 1 \quad g = 1 \quad \mathbf{u} = \mathbf{F}(\mathbf{x}) = \mathbf{x}$$

$$(\hat{\mathbf{e}}_n)^i = S_n^i = \delta_n^i \quad \Rightarrow \quad \hat{\mathbf{e}}_n = \hat{\mathbf{n}},$$

$$\mathbf{B} = \mathcal{B}^n \hat{\mathbf{e}}_n = B^n \hat{\mathbf{n}} \quad \Rightarrow \quad \mathcal{B}^n = B^n$$

so the above 3-line equation block becomes

$$\star \mathbf{B} = [\partial_n T - \varepsilon^{nab} \partial_a \Gamma_b] \hat{\mathbf{n}}$$

$$T = \partial_i (B^i(\mathbf{u}))$$

$$\Gamma_b = \varepsilon^{bcd} (\partial_c B^d(\mathbf{u}))$$

or

$$\begin{aligned} \star \mathbf{B} &= [\partial_n \{ \partial_i B^i \} - \varepsilon^{nab} \partial_a \{ \varepsilon^{bcd} (\partial_c B^d) \}] \hat{\mathbf{n}} & (*) \\ &= [\partial_n \{ \text{div } \mathbf{B} \} - \varepsilon^{nab} \partial_a \{ (\text{curl } \mathbf{B})_b \}] \hat{\mathbf{n}} \\ &= [\partial_n \{ \text{div } \mathbf{B} \} - [\text{curl } (\text{curl } \mathbf{B})]_n] \hat{\mathbf{n}} \\ &= \nabla \{ \text{div } \mathbf{B} \} - [\text{curl } (\text{curl } \mathbf{B})] \quad \text{QED} \end{aligned}$$

Second, one can verify the claim made earlier that in Cartesian coordinates

$$[\star \mathbf{B}]_n = \nabla^2 B_n .$$

To show this, it is necessary to show that (left side from (*) above)

$$\begin{aligned} \partial_n \partial_i B_i - \varepsilon^{nab} \partial_a \varepsilon^{bcd} (\partial_c B_d) &= \partial_i^2 B_n & ? \\ \varepsilon^{nab} \varepsilon^{bcd} \partial_a (\partial_c B_d) &= \partial_n \partial_i B_i - \partial_i^2 B_n & ? \\ \varepsilon^{bna} \varepsilon^{bcd} \partial_a (\partial_c B_d) &= \partial_n \partial_i B_i - \partial_i^2 B_n & ? \end{aligned}$$

But since index b now appears only in the ε 's, use (up and down indices same in Cartesian x-space)

$$\varepsilon_{bna} \varepsilon_{bcd} = \delta_{nc} \delta_{ad} - \delta_{nd} \delta_{ac} \quad // \text{ Appendix D (j) item 4}$$

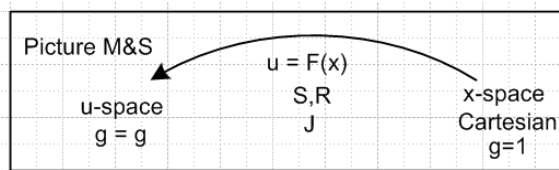
so

$$\begin{aligned} (\delta_{nc} \delta_{ad} - \delta_{nd} \delta_{ac}) \partial_a (\partial_c B_d) &= \partial_n \partial_i B_i - \partial_i^2 B_n & ? \\ \delta_{nc} \delta_{ad} \partial_a (\partial_c B_d) - \delta_{nd} \delta_{ac} \partial_a (\partial_c B_d) &= \partial_n \partial_i B_i - \partial_i^2 B_n & ? \\ \partial_a (\partial_n B_a) - \partial_a (\partial_a B_n) &= \partial_n \partial_i B_i - \partial_i^2 B_n & ? \\ \partial_n (\partial_a B_a) - \partial_a^2 B_n &= \partial_n (\partial_i B_i) - \partial_i^2 B_n & ? \end{aligned}$$

Since this last equation is true on inspection, QED.

14. Summary of Differential Operators in curvilinear coordinates

The results are given in the Picture M&S context, and are copied from Sections 9-13.



The Standard Notation of Section 7 is used throughout.

In all the differential operator equations below, an operator acts either on a tensorial vector field $\mathbf{B}$ or on a tensorial scalar field f . On the right side of the drawing above, objects are said to be in x-space, and $f(\mathbf{x})$ and $B_n(\mathbf{x})$ (components of $\mathbf{B}$) 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 $\mathbf{B}$, 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 $\mathbf{B} = R\mathbf{B}$, contravariant vector transformation).

See comment at the end of Section 9 (c) concerning the font used for $\mathcal{B}^n(\mathbf{u})$.

On lines marked "Cartesian", $\partial_n = \partial/\partial x^n$ and $f(\mathbf{x})$ 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 $\mathbf{u}$. The other functions like h_n , g_{ab} and g are also functions of $\mathbf{u}$.

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 Section 3, and the $\mathbf{e}^n$ are the reciprocal base vectors of Section 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 $\mathbf{A} \bullet \mathbf{B}$ is the covariant one of Section 5 (i).

For each differential operator, the object on the LHS of the equations is always the same: it is a differential operator acting on $f(\mathbf{x})$ or $\mathbf{B}(\mathbf{x})$ 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 .

The expressions marked below appear on pages 2 or 3 of Moon & Spencer (M&S).

$$\begin{array}{llll} \text{general:} & g \equiv \det(g_{ab}) & h_n^2 \equiv g_{nn} & \partial^i = g^{ij} \partial_j & \mathcal{B}^n = h_n B^n & \mathbf{e}_n = h_n \hat{\mathbf{e}}_n \\ \text{orthogonal:} & \sqrt{g} = (\Pi_i h_i) = h_1 h_2 \dots h_N & g_{nm} = h_n^2 \delta_{n,m} & & g^{nm} = h_n^{-2} \delta_{n,m} \end{array}$$

(a) divergence

divergence general:

$$\begin{array}{ll} [\text{div } \mathbf{B}](\mathbf{x}) = [1/\sqrt{g}] \partial_n [\sqrt{g} B^n] & \mathbf{B} = B^n \mathbf{e}_n \\ [\text{div } \mathbf{B}](\mathbf{x}) = [1/\sqrt{g}] \partial_n [\sqrt{g} \mathcal{B}^n / h_n] & \mathbf{B} = \mathcal{B}^n \hat{\mathbf{e}}_n \quad // \text{ M\&S 1.06} \\ [\text{div } \mathbf{B}](\mathbf{x}) = [1/\sqrt{g}] \partial_n [\sqrt{g} g^{nm} B_m] & \mathbf{B} = B_n \mathbf{e}^n \\ [\text{div } \mathbf{B}](\mathbf{x}) = \partial_n B^n & // B^n = \text{Cartesian components of } \mathbf{B} & \mathbf{B} = B^n \hat{\mathbf{n}} = B_n \hat{\mathbf{n}} \\ [\text{div } \mathbf{B}](\mathbf{x}) = [\text{div } \mathbf{B}](\mathbf{u}) & // \text{transformation (scalar)} \end{array}$$

divergence orthogonal:

$$\begin{array}{ll} [\text{div } \mathbf{B}](\mathbf{x}) = [1/(\Pi_i h_i)] \partial_n [(\Pi_i h_i) B^n] & \mathbf{B} = B^n \mathbf{e}_n \\ [\text{div } \mathbf{B}](\mathbf{x}) = [1/(\Pi_i h_i)] \partial_n [(\Pi_i h_i) \mathcal{B}^n / h_n] & \mathbf{B} = \mathcal{B}^n \hat{\mathbf{e}}_n \\ [\text{div } \mathbf{B}](\mathbf{x}) = [1/(\Pi_i h_i)] \partial_n [(\Pi_i h_i) B_n / h_n^2] & \mathbf{B} = B_n \mathbf{e}^n \end{array}$$

divergence orthogonal N=3:

$$[\text{div } \mathbf{B}](\mathbf{x}) = [1/(h_1 h_2 h_3)] \{ \partial_1 [h_2 h_3 \mathcal{B}^1(\mathbf{u})] + \text{cyclic} \} \quad \mathbf{B} = \mathcal{B}^n \hat{\mathbf{e}}_n$$

(b) gradient and gradient dot vector

gradient general:

$$\begin{array}{llll} [\text{grad } f](\mathbf{x}) = (\partial_i f) \mathbf{e}^i & & & \\ [\text{grad } f](\mathbf{x}) = (\partial^i f) \mathbf{e}_i & = g^{ij} (\partial_j f) \mathbf{e}_i & & \\ [\text{grad } f](\mathbf{x}) = h_i (\partial^i f) \hat{\mathbf{e}}_i & = h_i g^{ij} (\partial_j f) \hat{\mathbf{e}}_i & & \\ [\text{grad } f](\mathbf{x}) = (\partial_i f) \hat{\mathbf{n}} & // \text{Cartesian} & & \\ [grad f]_i(\mathbf{u}) = R_i^j [\text{grad } f]_j(\mathbf{x}) & // \text{transformation (vector)} & f(\mathbf{u}) = f(\mathbf{x}) = f(\mathbf{x}(\mathbf{u})) \end{array}$$

gradient orthogonal:

$$\begin{array}{llll} [\text{grad } f](\mathbf{x}) = (\partial_i f) \mathbf{e}^i & & & \\ [\text{grad } f](\mathbf{x}) = (\partial^i f) \mathbf{e}_i & = (1/h_i^2) (\partial_i f) \mathbf{e}_i & & \\ [\text{grad } f](\mathbf{x}) = h_i (\partial^i f) \hat{\mathbf{e}}_i & = (1/h_i) (\partial_i f) \hat{\mathbf{e}} & // \text{M\&S 1.05} \end{array}$$

gradient dotted with a vector:

$$\begin{aligned}
 [\text{grad } f](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) &= (\partial_{\mathbf{n}} f) B_{\mathbf{n}} & \mathbf{B} &= \Sigma_{\mathbf{i}} B_{\mathbf{i}} \mathbf{e}^{\mathbf{i}} \\
 [\text{grad } f](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) &= (\partial_{\mathbf{n}} f) B^{\mathbf{n}} & \mathbf{B} &= \Sigma_{\mathbf{i}} B^{\mathbf{i}} \mathbf{e}_{\mathbf{i}} \\
 [\text{grad } f](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) &= (\partial_{\mathbf{n}} f) \mathcal{B}^{\mathbf{n}}/h_{\mathbf{n}} & \mathbf{B} &= \Sigma_{\mathbf{i}} \mathcal{B}^{\mathbf{i}} \hat{\mathbf{e}}_{\mathbf{i}} \\
 [\text{grad } f](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) &= (\partial_{\mathbf{n}} f) B^{\mathbf{n}} & // \text{ Cartesian} & \mathbf{B} = B^{\mathbf{n}} \hat{\mathbf{n}} \\
 [\text{grad } f](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) &= [\text{grad } f](\mathbf{u}) \bullet \mathbf{B}(\mathbf{u}) & // \text{ transformation (scalar)} &
 \end{aligned}$$

(c) Laplacian

Laplacian general:

$$\begin{aligned}
 [\text{lap } f](\mathbf{x}) &= [1/\sqrt{g}] \partial_{\mathbf{m}} [\sqrt{g} g^{\mathbf{nm}} (\partial_{\mathbf{n}} f)] \\
 [\text{lap } f](\mathbf{x}) &= \partial_{\mathbf{n}}^2 f(\mathbf{x}) & // \text{ Cartesian} & f(\mathbf{u}) = f(\mathbf{x}) = f(\mathbf{x}(\mathbf{u})) \\
 [\text{lap } f](\mathbf{x}) &= [\text{lap } f](\mathbf{u}) & // \text{ transformation (scalar)} &
 \end{aligned}$$

Laplacian orthogonal:

$$[\text{lap } f](\mathbf{x}) = [1/(\Pi_{\mathbf{i}} h_{\mathbf{i}})] \partial_{\mathbf{m}} [(\Pi_{\mathbf{i}} h_{\mathbf{i}}) (1/h_{\mathbf{m}}^2) (\partial_{\mathbf{m}} f)] \quad // \text{ orthogonal} \quad // \text{ M\&S 1.09}$$

Laplacian orthogonal N=3:

$$[\text{lap } f](\mathbf{x}) = 1/(h_1 h_2 h_3) \{ \partial_1 [(h_2 h_3/h_1) \partial_1 f] + \text{cyclic} \}$$

(d) curl

curl general:

// N=3 only

$$\begin{aligned}
 [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{\mathbf{nab}} [(1/\sqrt{g}) \partial_{\mathbf{a}} B_{\mathbf{b}}] \mathbf{e}_{\mathbf{n}} & \mathbf{B} &= B_{\mathbf{n}} \mathbf{e}^{\mathbf{n}} \\
 [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{\mathbf{nab}} [(1/\sqrt{g}) \partial_{\mathbf{a}} (g_{\mathbf{bc}} B^{\mathbf{c}})] \mathbf{e}_{\mathbf{n}} & \mathbf{B} &= B^{\mathbf{n}} \mathbf{e}_{\mathbf{n}} \\
 [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{\mathbf{nab}} [(1/\sqrt{g}) h_{\mathbf{n}} \partial_{\mathbf{a}} (g_{\mathbf{bc}} \mathcal{B}^{\mathbf{c}}/h_{\mathbf{c}})] \hat{\mathbf{e}}_{\mathbf{n}} & \mathbf{B} &= \mathcal{B}^{\mathbf{n}} \hat{\mathbf{e}}_{\mathbf{n}} \\
 [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{\mathbf{nab}} \partial_{\mathbf{a}} B_{\mathbf{b}}(\mathbf{x}) \hat{\mathbf{n}} & // \text{ Cartesian} & \mathbf{B} = B^{\mathbf{n}} \hat{\mathbf{n}}
 \end{aligned}$$

$$[\text{curl } \mathbf{B}](\mathbf{x}) = (1/\sqrt{g}) \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ B_1 & B_2 & B_3 \end{vmatrix} \quad \mathbf{B} = B_{\mathbf{n}} \mathbf{e}^{\mathbf{n}}$$

$$[\text{curl } \mathbf{B}](\mathbf{x}) = (1/\sqrt{g}) \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ g_{1c} B^c & g_{2c} B^c & g_{3c} B^c \end{vmatrix} \quad \mathbf{B} = B^{\mathbf{n}} \mathbf{e}_{\mathbf{n}}$$

$$[\text{curl } \mathbf{B}](\mathbf{x}) = (1/\sqrt{g}) \begin{vmatrix} h_1 \hat{\mathbf{e}}_1 & h_2 \hat{\mathbf{e}}_2 & h_3 \hat{\mathbf{e}}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ (g_{1c}/h_c) \mathcal{B}^c & (g_{2c}/h_c) \mathcal{B}^c & (g_{3c}/h_c) \mathcal{B}^c \end{vmatrix} \quad \mathbf{B} = \mathcal{B}^n \hat{\mathbf{e}}_n \quad // \text{M\&S 1.07}$$

$$[\text{curl } \mathbf{B}](\mathbf{x}) = \begin{vmatrix} \hat{\mathbf{1}} & \hat{\mathbf{2}} & \hat{\mathbf{3}} \\ \partial_1 & \partial_2 & \partial_3 \\ B_1 & B_2 & B_3 \end{vmatrix} \quad // \text{here } \partial_n = \partial/\partial x^n \text{ and } B_i = B_i(\mathbf{x}) \quad \mathbf{B} = B^n \hat{\mathbf{n}}$$

curl orthogonal:

$$\begin{aligned} [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} [(h_1 h_2 h_3)^{-1} \partial_a B_b] \mathbf{e}_n & \mathbf{B} &= B_n \mathbf{e}^n \\ [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} [(h_1 h_2 h_3)^{-1} \partial_a (h_b^2 B^b)] \mathbf{e}_n & \mathbf{B} &= B^n \mathbf{e}_n \\ [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} [(h_1 h_2 h_3)^{-1} h_n \partial_a (h_b \mathcal{B}^b)] \hat{\mathbf{e}}_n & \mathbf{B} &= \mathcal{B}^n \hat{\mathbf{e}}_n \end{aligned}$$

$$[\text{curl } \mathbf{B}](\mathbf{x}) = (h_1 h_2 h_3)^{-1} \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ B_1 & B_2 & B_3 \end{vmatrix} \quad \mathbf{B} = B_n \mathbf{e}^n$$

$$[\text{curl } \mathbf{B}](\mathbf{x}) = (h_1 h_2 h_3)^{-1} \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ h_1^2 B^1 & h_2^2 B^2 & h_3^2 B^3 \end{vmatrix} \quad \mathbf{B} = B^n \mathbf{e}_n$$

$$[\text{curl } \mathbf{B}](\mathbf{x}) = (h_1 h_2 h_3)^{-1} \begin{vmatrix} h_1 \hat{\mathbf{e}}_1 & h_2 \hat{\mathbf{e}}_2 & h_3 \hat{\mathbf{e}}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ h_1 \mathcal{B}^1 & h_2 \mathcal{B}^2 & h_3 \mathcal{B}^3 \end{vmatrix} \quad \mathbf{B} = \mathcal{B}^n \hat{\mathbf{e}}_n \quad // \text{M\&S 1.07a}$$

(e) vector Laplacian

vector Laplacian general:

// N=3 only

$$[\star \mathbf{B}](\mathbf{x}) = \mathbf{e}_n [g^{kn} \partial_k \{ (1/\sqrt{g}) \partial_i (\sqrt{g} B^i) \} - (1/\sqrt{g}) \varepsilon^{ncd} \varepsilon^{eab} \partial_c \{ (1/\sqrt{g}) g_{de} (\partial_a [g_{bf} B^f]) \}] \quad \mathbf{B} = B^n \mathbf{e}_n$$

$$[\star \mathbf{B}](\mathbf{x}) = \hat{\mathbf{e}}_n h_n [g^{kn} \partial_k \{ (1/\sqrt{g}) \partial_i (\sqrt{g} \mathcal{B}^i/h_i) \} - (1/\sqrt{g}) \varepsilon^{ncd} \varepsilon^{eab} \partial_c \{ (1/\sqrt{g}) g_{de} (\partial_a [g_{bf} \mathcal{B}^f/h_f]) \}] \quad \mathbf{B} = \mathcal{B}^n \hat{\mathbf{e}}_n$$

$$[\star \mathbf{B}](\mathbf{x}) = \nabla^2(\mathbf{B}(\mathbf{x})) \quad // \text{Cartesian, meaning } [\star \mathbf{B}]^i(\mathbf{x}) = \nabla^2[B^i(\mathbf{x})] \quad \mathbf{B} = B^n \hat{\mathbf{n}}$$

vector Laplacian orthogonal:

$$\begin{aligned} [\star \mathbf{B}](\mathbf{x}) &= [(1/h_1) \partial_1 T + (h_1/\sqrt{g}) (\partial_3 \Gamma_2 - \partial_2 \Gamma_3)] \hat{\mathbf{e}}_1 \\ &\quad + [(1/h_2) \partial_2 T + (h_2/\sqrt{g}) (\partial_1 \Gamma_3 - \partial_3 \Gamma_1)] \hat{\mathbf{e}}_2 \\ &\quad + [(1/h_3) \partial_3 T + (h_3/\sqrt{g}) (\partial_2 \Gamma_1 - \partial_1 \Gamma_2)] \hat{\mathbf{e}}_3 \end{aligned} \quad // \text{ M\&S 1.11}$$

$$\begin{aligned} T &= (1/\sqrt{g}) \partial_i (\sqrt{g} \mathcal{B}^i / h_i) \\ \Gamma_1 &= (1/\sqrt{g}) h_1^2 (\partial_2 [h_3 \mathcal{B}^3] - \partial_3 [h_2 \mathcal{B}^2]) \\ \Gamma_2 &= (1/\sqrt{g}) h_2^2 (\partial_3 [h_1 \mathcal{B}^1] - \partial_1 [h_3 \mathcal{B}^3]) \\ \Gamma_3 &= (1/\sqrt{g}) h_3^2 (\partial_1 [h_2 \mathcal{B}^2] - \partial_2 [h_1 \mathcal{B}^1]) \end{aligned}$$

or

$$\begin{aligned} [\star \mathbf{B}](\mathbf{x}) &= [(1/h_n) \partial_n T - (h_n/\sqrt{g}) \varepsilon^{nab} \partial_a \Gamma_b] \hat{\mathbf{e}}_n \\ T &= (1/\sqrt{g}) \partial_i (\sqrt{g} \mathcal{B}^i / h_i) \\ \Gamma_b &= (1/\sqrt{g}) h_b^2 \varepsilon^{bcd} (\partial_c [h_d \mathcal{B}^d]) \end{aligned}$$

Example 1: Polar coordinates: a practical curvilinear notation

From earlier versions of this example we know that

$$\begin{aligned} \text{general:} \quad g &\equiv \det(g_{ab}) \quad h_n^2 \equiv g_{nn} \quad \partial^i = g^{ij} \partial_j \quad \mathcal{B}^n = h_n B^n \quad \mathbf{e}_n = h_n \hat{\mathbf{e}}_n \\ \text{orthogonal:} \quad \sqrt{g} &= (\Pi_i h_i) = h_1 h_2 \dots h_N \quad g_{nm} = h_n^2 \delta_{n,m} \quad g^{nm} = h_n^{-2} \delta_{n,m} \end{aligned}$$

$$\begin{aligned} \mathbf{e}_1 = r(-\sin\theta, \cos\theta) &= \mathbf{e}_\theta = r \hat{\mathbf{e}}_\theta & // = r \hat{\boldsymbol{\theta}} \\ \mathbf{e}_2 = (\cos\theta, \sin\theta) &= \mathbf{e}_r = \hat{\mathbf{e}}_r & // = \hat{\mathbf{r}} \end{aligned}$$

$$g_{ij} = \begin{pmatrix} \theta & r \\ r^2 & 0 \\ 0 & 1 \end{pmatrix} \begin{matrix} \theta \\ r \end{matrix} \quad R = \begin{pmatrix} x & y \\ -\sin\theta/r & \cos\theta/r \\ \cos(\theta) & \sin\theta \end{pmatrix} \begin{matrix} \theta \\ r \end{matrix} \quad u_1 = \theta \quad u_2 = r$$

$$h_1 = h_\theta = \sqrt{g_{\theta\theta}} = r \quad h_2 = h_r = \sqrt{g_{rr}} = 1$$

Assume one is working with a 2D vector velocity field $\mathbf{v}(\mathbf{x})$, 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^θ v^r	v_θ v_r	$\mathbf{v}_\theta$ $\mathbf{v}_r$	v_x v_y
v^1 v^2	v_1 v_2	$\mathbf{u}^1$ $\mathbf{u}^2$	$v_1=v^1$ $v_2=v^2$
contravariant (italic)	covariant (italic)	unit vector (non-italic)	Cartesian (non-italic)

One ends up with the comfortable $\mathbf{v} = v_\theta \hat{\boldsymbol{\theta}} + v_r \hat{\mathbf{r}}$ notation as shown below.

$$\begin{aligned} v_\theta &\equiv \mathbf{u}^1 = h_1 v^1 = h_\theta v^\theta = r v^\theta & // \text{ unit vector projection components} \\ v_r &\equiv \mathbf{u}^2 = h_2 v^2 = h_r v^r = v^r \end{aligned}$$

$$\begin{aligned} v^\theta &= R^\theta_x v_x + R^\theta_y v_y = -\sin\theta/r v_x + \cos\theta/r v_y \\ v^r &= R^r_x v_x + R^r_y v_y = \cos\theta v_x + \sin\theta v_y \end{aligned}$$

so

$$\begin{aligned} v_\theta &= r v^\theta = -\sin\theta v_x + \cos\theta v_y \\ v_r &= v^r = \cos\theta v_x + \sin\theta v_y \end{aligned}$$

$$\begin{aligned} \mathbf{v} &= v^n \mathbf{e}_n = v^\theta \mathbf{e}_\theta + v^r \mathbf{e}_r \\ \mathbf{v} &= \mathbf{u}^n \hat{\mathbf{e}}_n = v_\theta \hat{\mathbf{e}}_\theta + v_r \hat{\mathbf{e}}_r = v_\theta \hat{\boldsymbol{\theta}} + v_r \hat{\mathbf{r}} \\ \mathbf{v} &= v_n \hat{\mathbf{n}} = v_x \hat{\mathbf{x}} + v_y \hat{\mathbf{y}} \end{aligned}$$

As an example of a differential operator, consider the divergence for orthogonal coordinates from the above table,

$$[\text{div } \mathbf{B}](\mathbf{x}) = [1/(\Pi_i h_i)] \partial_n [(\Pi_i h_i) \mathcal{B}^n / h_n] \quad \mathbf{B} = \mathcal{B}^n \hat{\mathbf{e}}_n$$

which applied to the present situation reads ($h_1 = h_\theta = r$ and $h_2 = h_r = 1$),

$$\begin{aligned} [\text{div } \mathbf{v}](\mathbf{x}) &= [1/(h_1 h_2)] \{ \partial_1 [h_2 \mathbf{u}^1] + \partial_2 [h_1 \mathbf{u}^2] \} & \mathbf{V} = \mathbf{u}^n \hat{\mathbf{e}}_n \\ &= [1/(h_\theta h_r)] \{ \partial_\theta [h_r v_\theta] + \partial_r [h_\theta v_r] \} \\ &= (1/r) \{ \partial_\theta v_\theta + \partial_r (r v_r) \} \end{aligned}$$

Suppose v_x and v_y are constants. Then the Cartesian expression says

$$[\text{div } \mathbf{v}](\mathbf{x}) = \partial_n v^n = \partial_x v_x + \partial_y v_y = 0 + 0 = 0$$

The above curvilinear expression gives

$$\begin{aligned} [\text{div } \mathbf{v}](\mathbf{x}) &= (1/r) \{ \partial_\theta v_\theta + \partial_r (r v_r) \} \\ &= (1/r) \{ \partial_\theta [-\sin\theta v_x + \cos\theta v_y] + \partial_r (r [\cos\theta v_x + \sin\theta v_y]) \} \\ &= (1/r) \{ [-\cos\theta v_x - \sin\theta v_y] + [\cos\theta v_x + \sin\theta v_y] \} \\ &= 0 \end{aligned}$$

The notation illustrated here works for any curvilinear coordinates.

Appendix A: Reciprocal Base Vectors the Hard Way

Note: This Appendix is written in the development notation, not the Standard Notation, though a few equations are translated to the latter form. The rules for translation to Standard Notation are

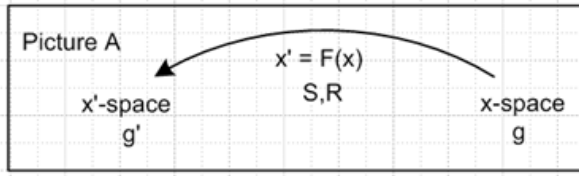
$$\mathbf{E}_n \rightarrow \mathbf{e}^n \quad R_{i,j} \rightarrow R^i_j \quad S_{i,j} \rightarrow S^i_j \quad \bar{g}'_{nm} \rightarrow g'_{nm} \quad g'_{nm} \rightarrow g'^{nm} .$$

Introduction

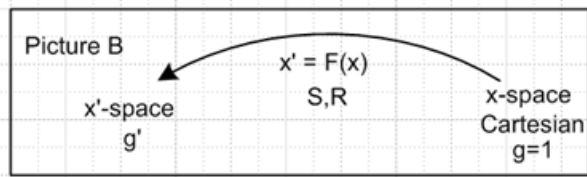
In Section 6 of the main text the reciprocal base vectors are defined as $\mathbf{E}_n \equiv g'_{ni} \mathbf{e}_i$, and the results given in that section,

$$\begin{aligned} (\mathbf{e}_n)_k &= S_{kn} & \mathbf{e}_n \bullet \mathbf{e}_m &= \bar{g}'_{nm} & |\mathbf{e}_n| &= \sqrt{\bar{g}'_{nn}} = h'_n & S &= [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \dots \mathbf{e}_N] \\ (\mathbf{E}_n)_i &\equiv g'_{ia} R_{na} & \mathbf{E}_n \bullet \mathbf{E}_m &= g'_{nm} & |\mathbf{E}_n| &= \sqrt{g'_{nn}} & R &= [\bar{\mathbf{E}}_1, \bar{\mathbf{E}}_2, \bar{\mathbf{E}}_3 \dots \bar{\mathbf{E}}_N]^T \\ &= g'_{na} S_{ia} & \mathbf{e}_n \bullet \mathbf{E}_m &= \delta_{n,m} & \mathbf{E}_n &\equiv g'_{ni} \mathbf{e}_i & \mathbf{e}_n &= \bar{g}'_{ni} \mathbf{E}_i , \end{aligned}$$

are all applicable in the Picture A context with arbitrary metric tensors g' and g ,



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 Section 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. The main 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) Definition of $\mathbf{E}_n$

The reciprocal base vectors are defined in the following very strange looking and clumsy manner,

$$(\mathbf{E}_k)_\alpha \equiv \det(R) (-1)^{k-1} \varepsilon_{\alpha i_1 i_2 i_3 \dots \cancel{i_k} \dots i_N} (\mathbf{e}_1)_{i_1} (\mathbf{e}_2)_{i_2} \dots (\mathbf{e}_\cancel{k})_{\cancel{i_k}} \dots (\mathbf{e}_N)_{i_N}$$

where N is the number of dimensions of the Cartesian x-space $\mathbb{R}^N$ 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: $\varepsilon_{123\dots N} \equiv +1$, 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 $\varepsilon = 0$.

(b) Simpler notation

To avoid dealing with subscripts on subscripts, one can rewrite the above definition in a less precise but simpler notation

$$(\mathbf{E}_k)_\alpha \equiv \det(R) (-1)^{k-1} \varepsilon_{\alpha abc\dots x} (\mathbf{e}_1)_a (\mathbf{e}_2)_b \dots (\mathbf{e}_N)_x \quad // \kappa(k) \text{ and } (\mathbf{e}_\kappa)_\kappa \text{ are missing}$$

In this notation, subscript x stands for the N^{th} letter of the alphabet (imagine $N \leq 26$). If κ is the k^{th} letter of the alphabet, then κ is missing from the indices on ε , and the factor $(\mathbf{e}_\kappa)_\kappa$ is missing from the product of factors. For example, if $k = 2$, then summation index $\kappa = b$ 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,

$$(\mathbf{E}_k)_\alpha \equiv \det(R) \varepsilon_{abc\dots\alpha\dots x} (\mathbf{e}_1)_a (\mathbf{e}_2)_b \dots (\mathbf{e}_N)_x \quad // (\mathbf{e}_\kappa)_\kappa \text{ is missing, } \alpha \text{ in } \kappa \text{ position (the } k^{\text{th}})$$

Example: For $N = 3$ the above becomes,

$$\begin{aligned} (\mathbf{E}_1)_\alpha &\equiv \det(R) \varepsilon_{\alpha bc} (\mathbf{e}_2)_b (\mathbf{e}_3)_c \Rightarrow \mathbf{E}_1 = \det(R) \mathbf{e}_2 \times \mathbf{e}_3 & a \text{ is missing} \\ (\mathbf{E}_2)_\alpha &\equiv \det(R) \varepsilon_{a\alpha c} (\mathbf{e}_1)_a (\mathbf{e}_3)_c \Rightarrow \mathbf{E}_2 = \det(R) \mathbf{e}_3 \times \mathbf{e}_1 & b \text{ is missing} \\ (\mathbf{E}_3)_\alpha &\equiv \det(R) \varepsilon_{ab\alpha} (\mathbf{e}_1)_a (\mathbf{e}_2)_b \Rightarrow \mathbf{E}_3 = \det(R) \mathbf{e}_1 \times \mathbf{e}_2 & c \text{ is missing} \end{aligned}$$

and the results are cyclic. Here is a detail from the middle line

$$\varepsilon_{a\alpha c} (\mathbf{e}_1)_a (\mathbf{e}_3)_c = -\varepsilon_{\alpha ac} (\mathbf{e}_1)_a (\mathbf{e}_3)_c = +\varepsilon_{\alpha ca} (\mathbf{e}_1)_a (\mathbf{e}_3)_c = \varepsilon_{\alpha ca} (\mathbf{e}_3)_c (\mathbf{e}_1)_a = [\mathbf{e}_3 \times \mathbf{e}_1]_\alpha$$

(c) 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:

$$\mathbf{Q}_a \equiv \varepsilon_{abc\dots x} B_b C_c D_d \dots X_x$$

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

$$\mathbf{Q} = \mathbf{B} \times \mathbf{C} \times \mathbf{D} \times \dots \times \mathbf{X} \quad / \text{N-1 factors, N-2 crosses}$$

This vector notation is *defined* by the previous line.

The vector $\mathbf{Q}$ is orthogonal to all the vectors from which it is constructed! For example (here is the point where $\mathbf{Q} \cdot \mathbf{C} \equiv g_{ab} Q_a C_b$ needs to be $Q_a C_a$, so $g = 1$ is required in x-space)

$$\mathbf{Q} \cdot \mathbf{C} = C_a Q_a = C_a \varepsilon_{abc\dots x} B_b C_c D_d \dots X_x = B_b D_d \dots X_x \{ \varepsilon_{abc\dots x} C_a C_c \}$$

But $\{..\}$ is the contraction of something symmetric under $a \leftrightarrow c$ ($C_a C_c$) with something antisymmetric under $a \leftrightarrow c$ ($\varepsilon_{\alpha abc\dots x}$) and therefore $\{..\} = 0$. In general

$$\begin{aligned} S_{ac} A_{ac} &= S_{ca} A_{ca} && // \text{relabel both dummy summation indices} \\ &= S_{ac} (-A_{ac}) && // S \text{ is Symmetric, } A \text{ is antisymmetric} \\ &= -S_{ac} A_{ac} && // = \text{the negative of the starting expression} \\ &= 0 \end{aligned}$$

Similarly, $\mathbf{Q} \cdot \mathbf{A} = 0$, $\mathbf{Q} \cdot \mathbf{B} = 0$ and so on.

Swapping the position of any two vectors in the generalized cross product causes $\mathbf{Q}$ to change sign. For example, swapping $\mathbf{B}$ and $\mathbf{C}$,

$$\begin{aligned} Q_a &\equiv \varepsilon_{abc\dots x} C_b B_c D_d \dots X_x = \varepsilon_{acb\dots x} C_c B_b D_d \dots X_x && // b \leftrightarrow c \\ &= -\varepsilon_{abc\dots x} B_b C_c D_d \dots X_x = -Q_a && // \text{swap indices on } \varepsilon \end{aligned}$$

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

$$Q_a \equiv \varepsilon_{ab} B_b \quad \Rightarrow \quad Q_1 = B_2 \quad \text{and} \quad Q_2 = -B_1$$

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 $\mathbf{B}$ 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:

$$\mathbf{Q} = \mathbf{B} \times \hat{\mathbf{3}} \quad \Rightarrow \quad Q_a \equiv \varepsilon_{abc} B_b (\hat{\mathbf{3}})_c = \varepsilon_{abc} B_b \delta_{3,c} = \varepsilon_{ab3} B_b = \varepsilon_{ab} B_b$$

Thus, this trick reproduces the correct component equation, and it makes more obvious the fact that $\mathbf{Q}$ is orthogonal to $\mathbf{B}$.

Summary: The generalized cross product $\mathbf{Q}$ of $N-1$ vectors each of dimension N can be expressed in both component and vector notation:

$$\begin{aligned} Q_a &\equiv \varepsilon_{abc\dots x} B_b C_c D_d \dots X_x \\ \mathbf{Q} &= \mathbf{B} \times \mathbf{C} \times \mathbf{D} \times \dots \times \mathbf{X} \quad / \text{N-1 factors, N-2 crosses} \end{aligned}$$

$\mathbf{Q}$ is orthogonal to all the vectors from which it is composed. Swapping any two vectors negates $\mathbf{Q}$. When $N=2$, one can rescue the otherwise failing vector notation by thinking of it as saying $\mathbf{Q} = \mathbf{B} \times \hat{\mathbf{z}}$.

Comment: Notice that $\mathbf{Q} = \mathbf{B} \times \mathbf{C} \times \mathbf{D}$ is defined for 4-vectors only. This is a completely different animal from the object $\mathbf{Q} = \mathbf{B} \times (\mathbf{C} \times \mathbf{D})$ which is defined for 3-vectors only. This latter object contains two ε factors, while the former only one.

(d) Missing Man Formation

We now make a small variation in the notation. Start with the above equation,

$$Q_a \equiv \varepsilon_{abc\dots x} B_b C_c D_d \dots X_x ,$$

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,

$$Q_\alpha \equiv \varepsilon_{\alpha ac\dots x} A_a B_b C_c \dots X_x \quad // \kappa \text{ and } K_\kappa \text{ are missing}$$

There are still $N-1$ factors, and one can still write this in vector notation

$$\mathbf{Q} = \mathbf{A} \times \mathbf{B} \times \mathbf{C} \times \dots \times \mathbf{X} \quad // \mathbf{K} \text{ is missing}$$

and of course it is still true that $\mathbf{Q} \bullet \mathbf{C} = 0$, etc. For $N=2$ the vector notation is rescued as in (c) above.

(e) Apply this Notation to $\mathbf{E}$

Compare the above Q_α to the section (a) definition of $(\mathbf{E}_\kappa)_\alpha$,

$$(\mathbf{E}_\kappa)_\alpha \equiv \det(R) (-1)^{k-1} \{ \varepsilon_{\alpha abc\dots x} (\mathbf{e}_1)_a (\mathbf{e}_2)_b \dots (\mathbf{e}_N)_x \} \quad // \kappa(k) \text{ and } (\mathbf{e}_\kappa)_\kappa \text{ are missing; } N \geq 2$$

Therefore, the definition of $\mathbf{E}_\kappa$ for $N > 2$ can be written in this vector notation,

$$\mathbf{E}_\kappa \equiv \det(R) (-1)^{k-1} \mathbf{e}_1 \times \mathbf{e}_2 \times \dots \times \mathbf{e}_N \quad // \mathbf{e}_\kappa \text{ missing; } N > 2$$

The reciprocal base vector $\mathbf{E}_\kappa$ is thus orthogonal to all the tangent base vectors from which it is constructed (remember $\mathbf{e}_\kappa$ is missing)! For example, for $N=3$ the three $\mathbf{E}$ vectors are given by

$$\begin{aligned} \mathbf{E}_1 &= \det(R) (-1)^{1-1} \mathbf{e}_2 \times \mathbf{e}_3 = \det(R) \mathbf{e}_2 \times \mathbf{e}_3 \\ \mathbf{E}_2 &= \det(R) (-1)^{2-1} \mathbf{e}_1 \times \mathbf{e}_3 = \det(R) \mathbf{e}_3 \times \mathbf{e}_1 \\ \mathbf{E}_3 &= \det(R) (-1)^{3-1} \mathbf{e}_1 \times \mathbf{e}_2 = \det(R) \mathbf{e}_1 \times \mathbf{e}_2 \end{aligned}$$

which agrees with the results quoted above. For $N=2$ ($\mathbf{E}$'s label corresponds to the missing $\mathbf{e}$'s label),

$$\begin{aligned} \mathbf{E}_1 &= \det(R) (-1)^{1-1} \mathbf{e}_2 \times \hat{\mathbf{z}} = \det(R) \mathbf{e}_2 \times \hat{\mathbf{z}} & \text{or} & & (\mathbf{E}_1)_k &= \det(R) \varepsilon_{ka} (\mathbf{e}_2)_a \\ \mathbf{E}_2 &= \det(R) (-1)^{2-1} \mathbf{e}_1 \times \hat{\mathbf{z}} = -\det(R) \mathbf{e}_1 \times \hat{\mathbf{z}} & \text{or} & & (\mathbf{E}_2)_k &= -\det(R) \varepsilon_{ka} (\mathbf{e}_1)_a \end{aligned}$$

One can combine these two lines into one as follows (eg, $k = 1$, then $3-1 = 2$, etc)

$$\mathbf{E}_k = \det(R) (-1)^{k-1} \mathbf{e}_{3-k} \times \hat{\mathbf{z}} = \det(R) \mathbf{e}_{3-k} \times \hat{\mathbf{z}} \quad \text{or} \quad (\mathbf{E}_1)_k = \det(R) (-1)^{k-1} \varepsilon_{ka} (\mathbf{e}_{3-k})_a$$

The vector "trick" notation shows that $\mathbf{E}_1 \bullet \mathbf{e}_2 = 0$ and $\mathbf{E}_2 \bullet \mathbf{e}_1 = 0$,

$$\begin{aligned} \mathbf{E}_1 \bullet \mathbf{e}_2 &= \det(R) \mathbf{e}_2 \times \hat{\mathbf{z}} \bullet \mathbf{e}_2 = 0 \\ \mathbf{E}_2 \bullet \mathbf{e}_1 &= -\det(R) \mathbf{e}_1 \times \hat{\mathbf{z}} \bullet \mathbf{e}_1 = 0 \end{aligned}$$

and also

$$\begin{aligned} \mathbf{E}_1 \bullet \mathbf{e}_1 &= \det(R) \varepsilon_{ka} (\mathbf{e}_2)_a (\mathbf{e}_1)_k = \det(R) \det[\mathbf{e}_1, \mathbf{e}_2] = \det(R) \det(S) = 1 \\ \mathbf{E}_2 \bullet \mathbf{e}_2 &= -\det(R) \varepsilon_{ka} (\mathbf{e}_1)_a (\mathbf{e}_2)_k = -\det(R) \det[\mathbf{e}_2, \mathbf{e}_1] = \det(R) \det(S) = 1 \end{aligned}$$

It is shown next that these $N=2$ results are special cases of a general fact: $\mathbf{E}_m \bullet \mathbf{e}_n = \delta_{m,n}$.

Section 5 (j) showed that $\mathbf{e}_m \bullet \mathbf{e}_n = \bar{g}'_{mn}$. The other two dot products are now considered.

(f) Compute $\mathbf{E}_m \bullet \mathbf{e}_n$

One can now compute, for general N ,

$$\mathbf{E}_k \bullet \mathbf{e}_k = (\mathbf{E}_k)_\alpha (\mathbf{e}_k)_\alpha = \{ \det(R) (-1)^{k-1} \varepsilon_{\alpha abc \dots x} (\mathbf{e}_1)_a (\mathbf{e}_2)_b \dots (\mathbf{e}_N)_x \} \quad (\mathbf{e}_k)_\alpha$$

k is missing

Slide α to the right in the ε subscript field and put it into the hole of the missing subscript k , picking up $(-1)^{k-1}$. At the same time, move the $(\mathbf{e}_k)_\alpha$ to the left and position it in its proper place in the product of factors,

$$\begin{aligned} \mathbf{E}_k \bullet \mathbf{e}_k &= (\mathbf{E}_k)_\alpha (\mathbf{e}_k)_\alpha = \{ \det(R) \varepsilon_{abc \dots \alpha \dots x} (\mathbf{e}_1)_a (\mathbf{e}_2)_b \dots (\mathbf{e}_k)_\alpha \dots (\mathbf{e}_N)_x \} \\ &= \det(R) \det[\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \dots \mathbf{e}_N] = \det(R) \det(S) = 1 \quad // \text{ since } RS = 1 \end{aligned}$$

We already know that $\mathbf{E}_k$ is orthogonal to all the $\mathbf{e}_n$ which form the generalized cross product, therefore

$$\mathbf{E}_m \bullet \mathbf{e}_n = \delta_{m,n}$$

which is the "duality relation" discussed more generally in Section 6 (b)

(g) Compute $\mathbf{E}_n \bullet \mathbf{E}_m$

Since the vectors $\{ \mathbf{e}_n \}$ are linearly independent and thus form a basis in $\mathbb{R}^N$, $\mathbf{E}_m$ can be expanded onto the $\mathbf{e}_n$,

$$\mathbf{E}_m = \sum_n A_n^{(m)} \mathbf{e}_n$$

$$\delta_{m,k} = \mathbf{E}_m \bullet \mathbf{e}_k = \sum_n A_n^{(m)} \mathbf{e}_n \bullet \mathbf{e}_k = \sum_n A_n^{(m)} \bar{g}'_{nk}$$

Multiplying both sides by g'_{ki} and summing on k gives

$$\text{LHS} = \sum_k g'_{ki} \delta_{m,k} = g'_{mi}$$

$$\text{RHS} = \sum_n A_n^{(m)} (\sum_k \bar{g}'_{nk} g'_{ki}) = \sum_n A_n^{(m)} (\bar{g}' g')_{ni} = \sum_n A_n^{(m)} \delta_{n,i} = A_i^{(m)}$$

Therefore $A_i^{(m)} = g'_{mi}$ so,

$$\mathbf{E}_m = \sum_n A_n^{(m)} \mathbf{e}_n = \sum_n g'_{mn} \mathbf{e}_n$$

which is to say $\mathbf{E}_k$ is this linear combination of the $\mathbf{e}_i$ (this is the definition used in Section 6 (a))

$$\mathbf{E}_k = \sum_i g'_{ki} \mathbf{e}_i = g'_{ki} \mathbf{e}_i \quad // \text{implied sum on } i \quad // \text{Std Notation: } \mathbf{e}^k = \sum_i g'^{ki} \mathbf{e}_i$$

which may be compared with the previous result

$$\mathbf{E}_k \equiv \det(R) (-1)^{k-1} \mathbf{e}_1 \times \mathbf{e}_2 \times \dots \times \mathbf{e}_N \quad // \mathbf{e}_k \text{ missing;}$$

It seems rather impressive that these two dissimilar ways of writing $\mathbf{E}$ are equal. Finally,

$$\mathbf{E}_n \bullet \mathbf{E}_m = \mathbf{E}_n \bullet (g'_{mi} \mathbf{e}_i) = g'_{mi} (\mathbf{E}_n \bullet \mathbf{e}_i) = g'_{mi} \delta_{n,i} = g'_{mn} = g'_{nm} \quad // \text{recall } g' \text{ symmetric}$$

(h) Summary of relationship between the tangent and reciprocal base vectors

$$\begin{array}{lll} \mathbf{e}_n \bullet \mathbf{e}_m = \bar{g}'_{nm} & \mathbf{E}_n \bullet \mathbf{E}_m = g'_{nm} & \mathbf{e}_n \bullet \mathbf{E}_m = \delta_{n,m} \\ \mathbf{E}_n = \sum_i g'_{ni} \mathbf{e}_i & \mathbf{e}_n = \sum_i \bar{g}'_{ni} \mathbf{E}_i & \bar{g}' = g'^{-1} \end{array}$$

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 Section 6 in which the equation $\mathbf{E}_n = \sum_i g'_{ni} \mathbf{e}_i$ 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.

In Standard Notation, the summary above can be restated as

$$\begin{array}{lll} \mathbf{e}_n \bullet \mathbf{e}_m = g'_{nm} & \mathbf{e}^n \bullet \mathbf{e}^m = g'^{nm} & \mathbf{e}_n \bullet \mathbf{e}^m = \delta_n^m \\ \mathbf{e}^n = \sum_i g'^{ni} \mathbf{e}_i & \mathbf{e}_n = \sum_i g'_{ni} \mathbf{e}^i & g'_{ab} = (g'^{ab})^{-1} \end{array}$$

(i) Another Cross Product Notation and another expression for E

Go back to the general cross product of N-1 vectors each of dimension N,

$$\mathbf{Q} = \mathbf{B} \times \mathbf{C} \times \mathbf{D} \times \dots \times \mathbf{X} \quad // \text{ N-1 factors, N-2 crosses}$$

Replace $\mathbf{B}, \mathbf{C}, \mathbf{D} \dots$ by vectors $\mathbf{A}^{(n)}$,

$$\mathbf{Q} = \mathbf{A}^{(1)} \times \mathbf{A}^{(2)} \times \mathbf{A}^{(3)} \times \dots \times \mathbf{A}^{(N-1)} \quad // \text{ N-1 factors, N-2 crosses}$$

It is convenient to write this as

$$\mathbf{Q} = \Pi_{i=1}^{N-1} \mathbf{A}^{(i)} = \Pi_i^x \mathbf{A}^{(i)}$$

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

$$\mathbf{Q} = \Pi_{i=1}^1 \mathbf{A}^{(i)} = \mathbf{A}^{(1)} \times \hat{\mathbf{3}} \quad // \neq \mathbf{A}^{(1)}$$

as discussed in section (c) above.

This same Π^x notation can be applied to the "missing man formation" of (d) above. Suppose

$$\mathbf{Q} = \mathbf{A}^{(1)} \times \mathbf{A}^{(2)} \times \mathbf{A}^{(3)} \times \dots \times \mathbf{A}^{(N)} \quad // \mathbf{A}^{(n)} \text{ is missing}$$

One can write this as

$$\mathbf{Q} = \Pi_{i=1 \dots N, i \neq n}^x \mathbf{A}^{(i)} \equiv \Pi_{i \neq n}^x \mathbf{A}^{(i)}$$

And of course this idea can be applied to the expression for $\mathbf{E}_k$

$$\mathbf{E}_k = \det(\mathbf{R}) (-1)^{k-1} \mathbf{e}_1 \times \mathbf{e}_2 \times \dots \times \mathbf{e}_N \quad // \mathbf{e}_k \text{ missing;}$$

$$\mathbf{E}_k = \det(\mathbf{R}) (-1)^{k-1} \Pi_{i \neq k}^x \mathbf{e}_i$$

Once again, for $N=2$ the Π^x symbol implies that ($\mathbf{e}_k$ = "missing", $\mathbf{e}_{3-k}$ = the one not missing)

$$\Pi_{i \neq k}^x \mathbf{e}_i = \Pi_{i=1 \dots 2, i \neq k}^x \mathbf{e}_i = \mathbf{e}_{3-k} \times \hat{\mathbf{3}}$$

$$\mathbf{E}_k = \det(\mathbf{R}) (-1)^{k-1} \mathbf{e}_{3-k} \times \hat{\mathbf{3}}$$

which is the "trick" notation of subsection (c) above for the $N=2$ case.

Appendix B: The Geometry of Parallelepipeds in N dimensions

Introduction

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^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^{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^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 Section 3, the normal vectors for the pairs of parallel faces are just the reciprocal base vectors $\mathbf{E}_n$ of Section 6.

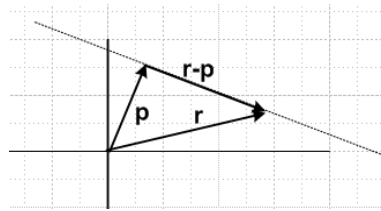
Section (d) focuses on the area and volume of N-pipeds in various dimensions, and simple expressions for the volume and vector areas of the faces of an N-piped are obtained.

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 $N = 2, 3, \dots$ are treated with nearly identical boilerplate templates to build up the inductive case.

All major results of this Appendix are concisely stated in Summary section (e). Since this is a very *long* section (~ 15 p), a reader not interested in details would do well to simply read that summary and skip the rest of this Appendix.

(a) 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 $\mathbf{p}$ be a vector from the origin to that closest point, and let $\mathbf{r}$ represent a point lying on the plane,



Since $\mathbf{p}$ is normal to the plane, and since $\mathbf{r}-\mathbf{p}$ is a vector lying *in* the plane, it follows that

$$\mathbf{p} \bullet (\mathbf{r}-\mathbf{p}) = 0 \quad \Rightarrow \quad \mathbf{r} \bullet \mathbf{p} = p^2 \quad \Rightarrow \quad \mathbf{r} \bullet \hat{\mathbf{p}} = p$$

Therefore, one way to write the equation of a plane in N dimensions is

$$\mathbf{r} \cdot \hat{\mathbf{p}} = p \quad \mathbf{r} = (x_1, x_2, \dots, x_N)$$

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 $\mathbf{r} \cdot \hat{\mathbf{p}} = 0$ where $\hat{\mathbf{p}}$ is either normal to the plane.

(b) N-pipeds and their Faces in Various Dimensions

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.

$$N=1 \quad \mathbf{r}_{\text{volume}1} = \alpha_1 \mathbf{e}_1 \quad 0 \leq \alpha_1 \leq 1$$

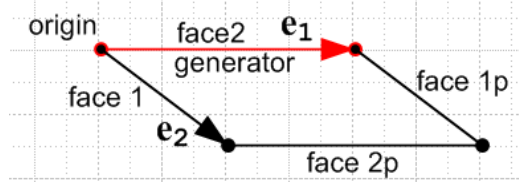
This piped has two vertices located at $\mathbf{v}_1 = 0$ and $\mathbf{v}_2 = \mathbf{e}_1$. These two vertices are also the "faces" of this 1-piped, so there are two faces (one pair of faces). These faces are 0 dimensional and therefore don't point in any direction (they are the endpoints of the segment). 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 volume of this 1-piped is e_1 .

The 2-piped

Now add another dimension, going to $N=2$. Introduce a unit vector $\mathbf{e}_2$ in some arbitrary direction in $\mathbb{R}^2$ 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 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.



The loci of the 2-piped's volume and of its four 1-piped faces are given by

$$\begin{aligned}
 \mathbf{r}_{\text{volume2}} &= \alpha_1 \mathbf{e}_1 + \alpha_2 \mathbf{e}_2 & 0 \leq \alpha_1, \alpha_2 \leq 1 \\
 \mathbf{r}_{\text{face2}} &= \alpha_1 \mathbf{e}_1 & 0 \leq \alpha_1 \leq 1 & \text{// the generator face} \\
 \mathbf{r}_{\text{face2p}} &= \alpha_1 \mathbf{e}_1 + \mathbf{e}_2 & 0 \leq \alpha_1 \leq 1 & \text{// partner of the generator face} \\
 \mathbf{r}_{\text{face1}} &= \alpha_2 \mathbf{e}_2 & 0 \leq \alpha_2 \leq 1 & \text{// side face touching the origin} \\
 \mathbf{r}_{\text{face1p}} &= \alpha_2 \mathbf{e}_2 + \mathbf{e}_1 & 0 \leq \alpha_2 \leq 1 & \text{// partner of the above side face}
 \end{aligned}$$

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.

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 (see Section 6 (b))

$$\mathbf{E}_i \cdot \mathbf{e}_j = \delta_{i,j} \quad \text{// } \mathbf{E}_k = \sum_{i=1}^2 g'_{ki} \mathbf{e}_i, \text{ see Section 6 (a)}$$

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 Section 6 and Appendix A.

Consider now these dot products:

$$\begin{aligned}
 \mathbf{E}_2 \cdot \mathbf{r}_{\text{face2}} &= \mathbf{E}_2 \cdot \alpha_1 \mathbf{e}_1 = 0 & \Rightarrow & \hat{\mathbf{E}}_2 \cdot \mathbf{r}_{\text{face2}} = 0 \\
 \mathbf{E}_2 \cdot \mathbf{r}_{\text{face2p}} &= \mathbf{E}_2 \cdot [\alpha_1 \mathbf{e}_1 + \mathbf{e}_2] = 1 & \Rightarrow & \hat{\mathbf{E}}_2 \cdot \mathbf{r}_{\text{face2p}} = 1/E_2
 \end{aligned}$$

The first line says (section (a) above) that face 2 lies on a plane which passes through the origin and which has **normal vector** $\hat{\mathbf{E}}_2$. 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 $1/|E_1|$. Similarly,

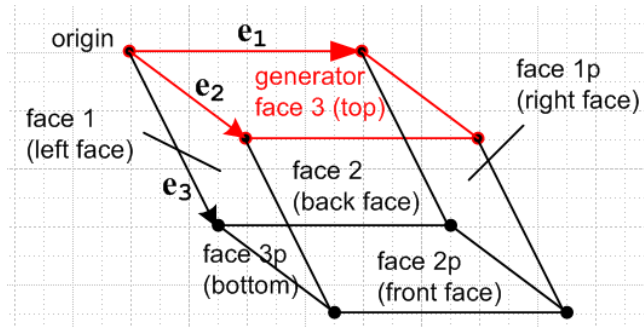
$$\begin{aligned}
 \mathbf{E}_1 \cdot \mathbf{r}_{\text{face1}} &= \mathbf{E}_1 \cdot \alpha_2 \mathbf{e}_2 = 0 & \Rightarrow & \hat{\mathbf{E}}_1 \cdot \mathbf{r}_{\text{face1}} = 0 \\
 \mathbf{E}_1 \cdot \mathbf{r}_{\text{face1p}} &= \mathbf{E}_1 \cdot [\alpha_2 \mathbf{e}_2 + \mathbf{e}_1] = 1 & \Rightarrow & \hat{\mathbf{E}}_1 \cdot \mathbf{r}_{\text{face1p}} = 1/E_1
 \end{aligned}$$

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.

The 3-piped

Now add another dimension, going to $N=3$. Introduce a unit vector $\hat{\mathbf{e}}_3$ in some arbitrary direction in $\mathbf{R}^3$ 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.



The loci of the 3-piped's volume and of its six 2-piped faces are given by

$$\mathbf{r}_{\text{volume3}} = \alpha_1 \mathbf{e}_1 + \alpha_2 \mathbf{e}_2 + \alpha_3 \mathbf{e}_3 \quad 0 \leq \alpha_1, \alpha_2, \alpha_3 \leq 1$$

$\mathbf{r}_{\text{face3}} = \alpha_1 \mathbf{e}_1 + \alpha_2 \mathbf{e}_2$	$0 \leq \alpha_1, \alpha_2 \leq 1$	// the generator face
$\mathbf{r}_{\text{face3p}} = \alpha_1 \mathbf{e}_1 + \alpha_2 \mathbf{e}_2 + \mathbf{e}_3$	$0 \leq \alpha_1, \alpha_2 \leq 1$	// partner face to the above
$\mathbf{r}_{\text{face2}} = \alpha_1 \mathbf{e}_1 + \alpha_3 \mathbf{e}_3$	$0 \leq \alpha_1, \alpha_3 \leq 1$	// the generator face
$\mathbf{r}_{\text{face2p}} = \alpha_1 \mathbf{e}_1 + \alpha_3 \mathbf{e}_3 + \mathbf{e}_2$	$0 \leq \alpha_1, \alpha_3 \leq 1$	// partner face to the above
$\mathbf{r}_{\text{face1}} = \alpha_2 \mathbf{e}_2 + \alpha_3 \mathbf{e}_3$	$0 \leq \alpha_2, \alpha_3 \leq 1$	// the generator face
$\mathbf{r}_{\text{face1p}} = \alpha_2 \mathbf{e}_2 + \alpha_3 \mathbf{e}_3 + \mathbf{e}_1$	$0 \leq \alpha_2, \alpha_3 \leq 1$	// partner face to the above

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. 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:

$$\mathbf{r}_{\text{volume}3} = \sum_n \alpha_n \mathbf{e}_n \quad 0 \leq \alpha_n \leq 1$$

$$\mathbf{r}_{\text{face}(i)} = \sum_{n \neq i} \alpha_n \mathbf{e}_n \quad 0 \leq \alpha_n \leq 1 \quad i = 1, 2, 3$$

$$\mathbf{r}_{\text{face}(ip)} = \sum_{n \neq i} \alpha_n \mathbf{e}_n + \mathbf{e}_i \quad 0 \leq \alpha_n \leq 1 \quad i = 1, 2, 3$$

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 (see Section 6 (b))

$$\mathbf{E}_i \bullet \mathbf{e}_j = \delta_{ij} \quad // \quad \mathbf{E}_k = \sum_i g'_{ki} \mathbf{e}_i, \text{ see Section 6 (a)}$$

If one interprets the $\mathbf{e}_n$ vectors as tangent base vectors, then the three vectors $\mathbf{E}_n$ are the corresponding reciprocal base vectors which are discussed in Section 6 and Appendix A. Consider now these dot products:

$$\mathbf{E}_1 \bullet \mathbf{r}_{\text{face}1} = \mathbf{E}_1 \bullet [\alpha_2 \mathbf{e}_2 + \alpha_3 \mathbf{e}_3] = 0 \quad \Rightarrow \quad \hat{\mathbf{E}}_1 \bullet \mathbf{r}_{\text{face}1} = 0$$

$$\mathbf{E}_1 \bullet \mathbf{r}_{\text{face}1p} = \mathbf{E}_1 \bullet [\alpha_2 \mathbf{e}_2 + \alpha_3 \mathbf{e}_3 + \mathbf{e}_1] = 1 \quad \Rightarrow \quad \hat{\mathbf{E}}_1 \bullet \mathbf{r}_{\text{face}1p} = 1/|\mathbf{E}_1|$$

The first line says that face 1 lies on a plane which passes through the origin and which has normal vector $\hat{\mathbf{E}}_1$. 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 $1/|\mathbf{E}_1|$

A similar pair of equations obtains for each of the other face pairs.

The N-piped

Now add another dimension, going from N-1 to N. Introduce a unit vector $\hat{\mathbf{e}}_N$ in some arbitrary direction in $\mathbf{R}^N$ so that $(\hat{\mathbf{e}}_1 \dots \hat{\mathbf{e}}_N)$ are linearly independent. Take the (N-1)-piped described above and translate it by distance e_N in the $\hat{\mathbf{e}}_N$ 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^{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 $\mathbf{e}_i$ 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".

The loci of the N-piped's volume and of its $2N$ (N-1)-piped faces are given by:

$$\mathbf{r}_{\text{volume}N} = \sum_n \alpha_n \mathbf{e}_n \quad 0 \leq \alpha_n \leq 1$$

$$\mathbf{r}_{\text{face}(i)} = \sum_{n \neq i} \alpha_n \mathbf{e}_n \quad 0 \leq \alpha_n \leq 1 \quad i = 1, 2, \dots, N$$

$$\mathbf{r}_{\text{face}(ip)} = \sum_{n \neq i} \alpha_n \mathbf{e}_n + \mathbf{e}_i \quad 0 \leq \alpha_n \leq 1 \quad i = 1, 2, \dots, N$$

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.

It is possible to construct vectors $\mathbf{E}_1 \dots \mathbf{E}_N$ as linear combinations of $\mathbf{e}_1 \dots \mathbf{e}_N$ such that the following is true (see Section 6 (b))

$$\mathbf{E}_i \bullet \mathbf{e}_j = \delta_{ij} \quad // \quad \mathbf{E}_k = \sum_i g'_{ki} \mathbf{e}_i$$

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 Section 6 and Appendix A. Consider now these dot products:

$$\begin{aligned} \mathbf{E}_i \bullet \mathbf{r}_{\text{face}(i)} &= \mathbf{E}_i \bullet [\sum_{n \neq i} a_n \mathbf{e}_n] = 0 & \Rightarrow & \hat{\mathbf{E}}_i \bullet \mathbf{r}_{\text{face}(i)} = 0 \\ \mathbf{E}_i \bullet \mathbf{r}_{\text{face}(ip)} &= \mathbf{E}_i \bullet [\sum_{n \neq i} a_n \mathbf{e}_n + \mathbf{e}_i] = 1 & \Rightarrow & \hat{\mathbf{E}}_i \bullet \mathbf{r}_{\text{face}(ip)} = 1/E_i \end{aligned}$$

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 $1/|E_i|$

(c) 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

$$\begin{aligned} (\text{piped})_{\text{center}} &= \sum_n (1/2) \mathbf{e}_n & // & \text{vector from origin to piped center} \\ (\text{face } i)_{\text{center}} &= \sum_{n \neq i} (1/2) \mathbf{e}_n & // & \text{vector from origin to center of face } i \\ (\text{face } ip)_{\text{center}} &= \sum_{n \neq i} (1/2) \mathbf{e}_n + \mathbf{e}_i & // & \text{vector from origin to center of face } ip \end{aligned}$$

Construct vectors from piped center to face centers (results here are fairly obvious)

$$\begin{aligned} (\text{face } i)_{\text{center}} - (\text{piped})_{\text{center}} &= [\sum_{n \neq i} (1/2) \mathbf{e}_n] - \sum_n (1/2) \mathbf{e}_n = - (1/2) \mathbf{e}_i \\ (\text{face } ip)_{\text{center}} - (\text{piped})_{\text{center}} &= [\sum_{n \neq i} (1/2) \mathbf{e}_n + \mathbf{e}_i] - \sum_n (1/2) \mathbf{e}_n = + (1/2) \mathbf{e}_i \end{aligned}$$

Then compute

$$\begin{aligned} \mathbf{E}_i \bullet \{(\text{face } i)_{\text{center}} - (\text{piped})_{\text{center}}\} &= \mathbf{E}_i \bullet [- (1/2) \mathbf{e}_i] = - (1/2) < 0 \\ \mathbf{E}_i \bullet \{(\text{face } ip)_{\text{center}} - (\text{piped})_{\text{center}}\} &= \mathbf{E}_i \bullet [+ (1/2) \mathbf{e}_i] = + (1/2) > 0 \end{aligned}$$

One may conclude that $\mathbf{E}_i$ is an outward pointing normal for face i_p (far face). Therefore, $-\mathbf{E}_i$ is an outward pointing normal for face i , which recall is the face which touches the origin (near face).

(d) 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.

The 2-piped

Face Area. The area of a 2-piped face (a line segment) is just the length of the edge which is the face,

$$\begin{aligned} A_1 &= |\mathbf{e}_2| \\ A_2 &= |\mathbf{e}_1| \end{aligned}$$

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,

$$\begin{aligned} \mathbf{A}_1 &= |\mathbf{e}_2| \hat{\mathbf{E}}_1 & // \mathbf{E}_k &= \sum_i g'_{ki} \mathbf{e}_i, \text{ see Appendix A (g)} \\ \mathbf{A}_2 &= |\mathbf{e}_1| \hat{\mathbf{E}}_2 \end{aligned}$$

and these vectors are out-facing for faces 1_p and 2_p . We claim that both these results can be expressed in a single formula

$$\mathbf{A}_n = |\det(S)| \mathbf{E}_n .$$

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

$$\begin{aligned} |\det(S)| |\mathbf{E}_1| &= |\mathbf{e}_2| & \text{and} & & |\det(S)| |\mathbf{E}_2| &= |\mathbf{e}_1| \\ \text{or} & & & & & \\ |\mathbf{E}_k| &= |\det(R)| |\mathbf{e}_{3-k}| & k=1,2 & & // RS = 1 \end{aligned}$$

Using the $N=2$ trick notation from Appendix A (c),

$$\begin{aligned} \mathbf{E}_k &= \det(R) (-1)^{k-1} \mathbf{e}_{3-k} \times \hat{\mathbf{3}} \\ \text{so that} & & & & \\ |\mathbf{E}_k| &= |\det(R)| |\mathbf{e}_{3-k} \times \hat{\mathbf{3}}| = |\det(R)| |\mathbf{e}_{3-k}| & k=1,2 \end{aligned}$$

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 $N \geq 2$.

One can restate $\mathbf{A}_n = |\det(S)| \mathbf{E}_n$ using the cross product notation presented in Appendix A (i):

$$\begin{aligned}\mathbf{A}_n &= |\det(S)| \mathbf{E}_n = |\det(S)| \det(R) (-1)^{n-1} \prod_{i \neq n} \mathbf{e}_i \\ &= \sigma (-1)^{n-1} \prod_{i \neq n} \mathbf{e}_i \quad \sigma \equiv \text{sign}(\det(S)) = \text{sign}(\det(R))\end{aligned}$$

Volume. The volume of a 2-piped is the base times the height of a parallelogram, familiarly given as by the cross product of the edges,

$$\text{volume}(2) = | \mathbf{e}_1 \times \mathbf{e}_2 | = | \varepsilon_{ab} (\mathbf{e}_1)_a (\mathbf{e}_2)_b | = | \det [\mathbf{e}_1, \mathbf{e}_2] | = | \det(S) |$$

where S is the linearized transformation matrix for $N=2$, see Section 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 $\hat{3}$ dimension to give it meaning. The second form *does* have a meaning for $N=2$, and that meaning is $|(\mathbf{e}_1)_1 (\mathbf{e}_2)_2 - (\mathbf{e}_1)_2 (\mathbf{e}_2)_1|$.

The 3-piped

Face Area. The faces of a 3-piped are 2-pipeds. For $N=2$, the 2-piped volume was

$$\text{volume}(2) = | \varepsilon_{ab} (\mathbf{e}_1)_a (\mathbf{e}_2)_b |,$$

where $\mathbf{e}_1$ and $\mathbf{e}_2$ were 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 $\mathbf{e}_n$ coordinate system -- vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ are 3D vectors. The first 2 components of each of these 3D vectors are the same as the components of the 2D $\mathbf{e}_i$ 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 $| \varepsilon_{ab3} (\mathbf{e}_1)_a (\mathbf{e}_2)_b |$, where ε is now a 3D ε tensor. The conclusion is that

$$A_3 = | \varepsilon_{ab3} (\mathbf{e}_1)_a (\mathbf{e}_2)_b |$$

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

$$\begin{aligned}A_1 &= | \varepsilon_{1ab} (\mathbf{e}_2)_a (\mathbf{e}_3)_b | \\ A_2 &= | \varepsilon_{a2b} (\mathbf{e}_3)_a (\mathbf{e}_1)_b |\end{aligned}$$

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

$$\begin{aligned}A_1 &= | \mathbf{e}_2 \times \mathbf{e}_3 | && \text{face 1 and face 1p} \\ A_2 &= | \mathbf{e}_3 \times \mathbf{e}_1 | && \text{face 2 and face 2p} \\ A_3 &= | \mathbf{e}_1 \times \mathbf{e}_2 | && \text{face 3 and face 3p}\end{aligned}$$

In Appendix A (e) it was shown that $\mathbf{E}_1 = \det(\mathbf{R}) \mathbf{e}_2 \times \mathbf{e}_3$ so that $\mathbf{e}_2 \times \mathbf{e}_3$ lines up with $\mathbf{E}_1$. Regardless of the sign of $\det(\mathbf{R})$, we define the *vector* areas to point in the $+\hat{\mathbf{E}}_n$ directions. Thus,

$$\begin{aligned} \mathbf{A}_1 &\equiv |\mathbf{e}_2 \times \mathbf{e}_3| \hat{\mathbf{E}}_1 && \text{face 1p out-facing} \\ \mathbf{A}_2 &\equiv |\mathbf{e}_3 \times \mathbf{e}_1| \hat{\mathbf{E}}_2 && \text{face 2p out-facing} \\ \mathbf{A}_3 &\equiv |\mathbf{e}_1 \times \mathbf{e}_2| \hat{\mathbf{E}}_3 && \text{face 3p out-facing} \end{aligned}$$

These equations can be combined into the following single formula

$$\mathbf{A}_n = |\mathbf{e}_1 \times \dots \times \mathbf{e}_3| \hat{\mathbf{E}}_n \quad // \mathbf{e}_n \text{ missing}$$

where the $\mathbf{e}_i$ are reordered for free due to the absolute value signs. But Appendix A says

$$\mathbf{E}_n = \det(\mathbf{R}) (-1)^{n-1} \mathbf{e}_1 \times \dots \times \mathbf{e}_3 \quad // \mathbf{e}_n \text{ missing}$$

so

$$|\mathbf{E}_n| = |\det(\mathbf{R})| |\mathbf{e}_1 \times \dots \times \mathbf{e}_3| \quad // \mathbf{e}_n \text{ missing}$$

Thus,

$$\begin{aligned} \mathbf{A}_n &= \hat{\mathbf{E}}_n |\mathbf{E}_n| / |\det(\mathbf{R})| = |\det(\mathbf{S})| \mathbf{E}_n \\ &= |\det(\mathbf{S})| \det(\mathbf{R}) (-1)^{n-1} \mathbf{e}_1 \times \dots \times \mathbf{e}_3 && // \mathbf{e}_n \text{ missing} \\ &= \sigma (-1)^{n-1} \mathbf{e}_1 \times \dots \times \mathbf{e}_3 && // \mathbf{e}_n \text{ missing} \\ \text{where} \quad &\sigma \equiv \text{sign}[\det(\mathbf{S})] = \text{sign}[\det(\mathbf{R})] \end{aligned}$$

To summarize, for $N=3$ one has

$$\begin{aligned} \mathbf{A}_n &= |\det(\mathbf{S})| \mathbf{E}_n = \sigma (-1)^{n-1} \mathbf{e}_1 \times \dots \times \mathbf{e}_3 && // \mathbf{e}_n \text{ missing} \\ &= \sigma (-1)^{n-1} \Pi_{i \neq n}^{\mathbf{x}} \mathbf{e}_i \end{aligned}$$

where the last line uses the shorthand notation of Appendix A (i). These expressions have the same form as those of the 2-piped.

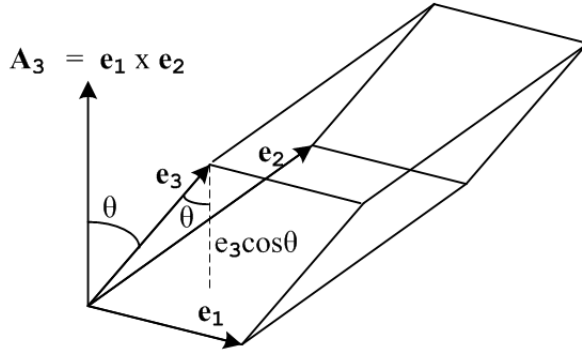
Volume. The volume of a 3-piped (with each of the above $\mathbf{A}_n$ in turn treated as the base) is base times height, so

$$\text{volume}(3) = |\mathbf{A}_1 \bullet \mathbf{e}_1| = |\mathbf{A}_2 \bullet \mathbf{e}_2| = |\mathbf{A}_3 \bullet \mathbf{e}_3|$$

or

$$\text{volume}(3) = |\mathbf{e}_2 \times \mathbf{e}_3 \bullet \mathbf{e}_1| = |\mathbf{e}_3 \times \mathbf{e}_1 \bullet \mathbf{e}_2| = |\mathbf{e}_1 \times \mathbf{e}_2 \bullet \mathbf{e}_3|$$

Here is a drawing showing the last case ($\sigma = +1$), where "base" is $\mathbf{A}_3 = |\mathbf{e}_1 \times \mathbf{e}_2|$ and "height" is $\mathbf{e}_3 \cos \theta$,



Using ϵ notation one can write

$$\begin{aligned} \mathbf{e}_3 \bullet \mathbf{e}_1 \times \mathbf{e}_2 &= (\mathbf{e}_3)_i \epsilon_{ijk} (\mathbf{e}_1)_j (\mathbf{e}_2)_k = \epsilon_{ijk} (\mathbf{e}_1)_j (\mathbf{e}_2)_k (\mathbf{e}_3)_i = \epsilon_{jki} (\mathbf{e}_1)_j (\mathbf{e}_2)_k (\mathbf{e}_3)_i \\ &= \det [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3] = \det(S) \end{aligned}$$

so that

$$\text{volume}(3) = |\mathbf{e}_3 \bullet \mathbf{e}_1 \times \mathbf{e}_2| = |\epsilon_{abc} (\mathbf{e}_1)_a (\mathbf{e}_2)_b (\mathbf{e}_3)_c| = |\det [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]| = |\det(S)|$$

These expressions have the same form as those of the 2-piped.

The 4-piped

Face Area: The faces of a 4-piped are 3-pipeds. For $N=3$, the 3-piped volume was

$$\text{volume}(3) = |\epsilon_{abc} (\mathbf{e}_1)_a (\mathbf{e}_2)_b (\mathbf{e}_3)_c|$$

where $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ were 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 $|\epsilon_{abc4} (\mathbf{e}_1)_a (\mathbf{e}_2)_b (\mathbf{e}_3)_c|$, where ϵ is now a 4D ϵ tensor. The conclusion is that

$$A_4 = |\epsilon_{abc4} (\mathbf{e}_1)_a (\mathbf{e}_2)_b (\mathbf{e}_3)_c|$$

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

$$A_1 = |\epsilon_{1abc} (\mathbf{e}_2)_a (\mathbf{e}_3)_b (\mathbf{e}_4)_c|$$

$$A_2 = |\epsilon_{a2bc} (\mathbf{e}_3)_a (\mathbf{e}_4)_b (\mathbf{e}_1)_c|$$

$$A_3 = |\epsilon_{ab3c} (\mathbf{e}_4)_a (\mathbf{e}_1)_b (\mathbf{e}_2)_c|$$

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

$$\begin{aligned} A_1 &= |\mathbf{e}_2 \times \mathbf{e}_3 \times \mathbf{e}_4| && \text{face 1 and face 1p} \\ A_2 &= |\mathbf{e}_3 \times \mathbf{e}_4 \times \mathbf{e}_1| && \text{face 2 and face 2p} \\ A_3 &= |\mathbf{e}_4 \times \mathbf{e}_1 \times \mathbf{e}_2| && \text{face 3 and face 3p} \\ A_4 &= |\mathbf{e}_1 \times \mathbf{e}_2 \times \mathbf{e}_3| && \text{face 4 and face 4p} \end{aligned}$$

where, as discussed in Appendix A (c),

$$\mathbf{Q} = \mathbf{A} \times \mathbf{B} \times \mathbf{C} \quad \text{is defined by} \quad Q_k = \varepsilon_{kabc} A_a B_b C_c .$$

In Appendix A (e) it was shown that $\mathbf{E}_1 = \det(R) \mathbf{e}_2 \times \mathbf{e}_3 \times \mathbf{e}_4$ so that $\mathbf{e}_2 \times \mathbf{e}_3 \times \mathbf{e}_4$ 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

$$\mathbf{A}_n = |\mathbf{e}_1 \times \dots \times \mathbf{e}_3| \hat{\mathbf{E}}_n \quad // \mathbf{e}_n \text{ missing} \quad n = 1, 2, 3, 4$$

where the $\mathbf{e}_i$ are reordered for free due to the absolute value signs. But Appendix A says

$$\mathbf{E}_n = \det(R) (-1)^{n-1} \mathbf{e}_1 \times \dots \times \mathbf{e}_4 \quad // \mathbf{e}_n \text{ missing}$$

so

$$|\mathbf{E}_n| = |\det(R)| |\mathbf{e}_1 \times \dots \times \mathbf{e}_4| \quad // \mathbf{e}_n \text{ missing}$$

Thus,

$$\begin{aligned} \mathbf{A}_n &= |\mathbf{E}_n| \hat{\mathbf{E}}_n / |\det(R)| = |\det(S)| \mathbf{E}_n \\ &= |\det(S)| \det(R) (-1)^{n-1} \mathbf{e}_1 \times \dots \times \mathbf{e}_4 \quad // \mathbf{e}_n \text{ missing} \\ &= \sigma (-1)^{n-1} \mathbf{e}_1 \times \dots \times \mathbf{e}_4 \quad // \mathbf{e}_n \text{ missing} \quad \sigma \equiv \text{sign}[\det(S)] = \text{sign}[\det(R)] \end{aligned}$$

To summarize, for $N=4$ one has

$$\begin{aligned} \mathbf{A}_n &= |\det(S)| \mathbf{E}_n = \sigma (-1)^{n-1} \mathbf{e}_1 \times \dots \times \mathbf{e}_4 \quad // \mathbf{e}_n \text{ missing} \\ &= \sigma (-1)^{n-1} \prod_{i \neq n} \mathbf{e}_i \quad \sigma \equiv \text{sign}[\det(S)] = \text{sign}[\det(R)] \end{aligned}$$

These expressions have the same form as those of the 2-piped and the 3-piped.

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

$$\text{volume}(4) = |\mathbf{A}_1 \bullet \mathbf{e}_1| = |\mathbf{A}_2 \bullet \mathbf{e}_2| = |\mathbf{A}_3 \bullet \mathbf{e}_3| = |\mathbf{A}_4 \bullet \mathbf{e}_4|$$

or

$$\text{volume}(4) = |\mathbf{e}_2 \times \mathbf{e}_3 \times \mathbf{e}_4 \bullet \mathbf{e}_1| = |\mathbf{e}_3 \times \mathbf{e}_4 \times \mathbf{e}_1 \bullet \mathbf{e}_2| = |\mathbf{e}_4 \times \mathbf{e}_1 \times \mathbf{e}_2 \bullet \mathbf{e}_3| = |\mathbf{e}_1 \times \mathbf{e}_2 \times \mathbf{e}_3 \bullet \mathbf{e}_4|$$

Using ε notation one can write the first case as

$$\begin{aligned} \mathbf{e}_2 \times \mathbf{e}_3 \times \mathbf{e}_4 \cdot \mathbf{e}_1 &= (\mathbf{e}_1)_a \varepsilon_{abcd} (\mathbf{e}_2)_b (\mathbf{e}_3)_c (\mathbf{e}_4)_d = \varepsilon_{abcd} (\mathbf{e}_1)_a (\mathbf{e}_2)_b (\mathbf{e}_3)_c (\mathbf{e}_4)_d \\ &= \det [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4] = \det(S) \end{aligned}$$

so that

$$\text{volume}(4) = |\det(S)| = |\det [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4]| = |\varepsilon_{abcd} (\mathbf{e}_1)_a (\mathbf{e}_2)_b (\mathbf{e}_3)_c (\mathbf{e}_4)_d|$$

These expressions have the same form as those of the 2-piped and the 3-piped.

The N-piped

Face Area: The faces of a 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

$$\text{volume}(N-1) = |\varepsilon_{abc\dots x} (\mathbf{e}_1)_a (\mathbf{e}_2)_b \dots (\mathbf{e}_{N-1})_x|$$

where $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_{N-1}$ were (N-1)D vectors. 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 $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_{N-1}$ 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 $\mathbf{e}_1$ 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 $|\varepsilon_{abc\dots xN} (\mathbf{e}_1)_a (\mathbf{e}_2)_b \dots (\mathbf{e}_{N-1})_x|$, where ε is now an ND ε tensor. The conclusion is that

$$A_N = |\varepsilon_{abc\dots xN} (\mathbf{e}_1)_a (\mathbf{e}_2)_b \dots (\mathbf{e}_{N-1})_x|$$

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,

$$\begin{aligned} A_1 &= |\varepsilon_{1abc\dots x} (\mathbf{e}_2)_a (\mathbf{e}_3)_b \dots (\mathbf{e}_{N-1})_w (\mathbf{e}_N)_x| \\ A_2 &= |\varepsilon_{a2bc\dots x} (\mathbf{e}_3)_a (\mathbf{e}_4)_b \dots (\mathbf{e}_N)_w (\mathbf{e}_1)_x| \end{aligned}$$

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

$$\begin{aligned} A_1 &= |\mathbf{e}_2 \times \mathbf{e}_3 \times \mathbf{e}_4 \times \mathbf{e}_5 \dots \times \mathbf{e}_N| && \text{face 1 and face 1p} \\ A_2 &= |\mathbf{e}_3 \times \mathbf{e}_4 \times \mathbf{e}_5 \times \mathbf{e}_6 \dots \times \mathbf{e}_1| && \text{face 2 and face 2p} \\ A_3 &= |\mathbf{e}_4 \times \mathbf{e}_5 \times \mathbf{e}_6 \times \mathbf{e}_7 \dots \times \mathbf{e}_2| && \text{face 3 and face 3p} \\ &\dots && \\ A_N &= |\mathbf{e}_5 \times \mathbf{e}_6 \times \mathbf{e}_7 \times \mathbf{e}_8 \dots \times \mathbf{e}_3| && \text{face N and face Np} \end{aligned}$$

where, as discussed in Appendix A (c),

$$\mathbf{Q} = \mathbf{A} \times \mathbf{B} \times \mathbf{C} \dots \times \mathbf{X} \quad \text{is defined by} \quad Q_k = \varepsilon_{kabc\dots x} A_a B_b C_c \dots X_x$$

In Appendix A (e) it was shown that $\mathbf{E}_1 = \det(\mathbf{R}) \mathbf{e}_2 \times \mathbf{e}_3 \dots \mathbf{e}_N$ so that $\mathbf{e}_2 \times \mathbf{e}_3 \dots \mathbf{e}_N$ lines up with $\mathbf{E}_1$. Regardless of the sign of $\det(\mathbf{R})$, we define the *vector* areas to point in the $+\hat{\mathbf{E}}_n$ directions. Thus

$$\mathbf{A}_n = |\mathbf{e}_1 \times \dots \times \mathbf{e}_N| \hat{\mathbf{E}}_n \quad // \mathbf{e}_n \text{ missing} \quad n = 1, 2, 3, \dots, N$$

where the $\mathbf{e}_i$ are reordered for free due to the absolute value signs. But Appendix A says

$$\mathbf{E}_n = \det(\mathbf{R}) (-1)^{n-1} \mathbf{e}_1 \times \dots \times \mathbf{e}_N \quad // \mathbf{e}_n \text{ missing}$$

so

$$|\mathbf{E}_n| = |\det(\mathbf{R})| |\mathbf{e}_1 \times \dots \times \mathbf{e}_N| \quad // \mathbf{e}_n \text{ missing}$$

Thus,

$$\begin{aligned} \mathbf{A}_n &= |\mathbf{E}_n| \hat{\mathbf{E}}_n / |\det(\mathbf{R})| = |\det(\mathbf{S})| \mathbf{E}_n \\ &= |\det(\mathbf{S})| \det(\mathbf{R}) (-1)^{n-1} \mathbf{e}_1 \times \dots \times \mathbf{e}_N \quad // \mathbf{e}_n \text{ missing} \\ &= \sigma (-1)^{n-1} \mathbf{e}_1 \times \dots \times \mathbf{e}_N \quad // \mathbf{e}_n \text{ missing} \quad \sigma \equiv \text{sign}[\det(\mathbf{S})] = \text{sign}[\det(\mathbf{R})] \end{aligned}$$

To summarize, for $N=N$ one has

$$\begin{aligned} \mathbf{A}_n &= |\det(\mathbf{S})| \mathbf{E}_n = \sigma (-1)^{n-1} \mathbf{e}_1 \times \dots \times \mathbf{e}_N \quad // \mathbf{e}_n \text{ missing} \\ &= \sigma (-1)^{n-1} \prod_{i \neq n} \mathbf{e}_i \quad \sigma \equiv \text{sign}[\det(\mathbf{S})] = \text{sign}[\det(\mathbf{R})] \end{aligned}$$

These expressions have 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

$$\text{volume}(N) = |\mathbf{A}_1 \bullet \mathbf{e}_1| = |\mathbf{A}_2 \bullet \mathbf{e}_2| = \dots = |\mathbf{A}_N \bullet \mathbf{e}_N|$$

or

$$\text{volume}(N) = |\mathbf{e}_2 \times \mathbf{e}_3 \times \mathbf{e}_4 \dots \mathbf{e}_N \bullet \mathbf{e}_1| = \dots$$

Using ε notation one can write the first case as

$$\begin{aligned} \mathbf{e}_2 \times \mathbf{e}_3 \times \mathbf{e}_4 \dots \mathbf{e}_N \bullet \mathbf{e}_1 &= (\mathbf{e}_1)_a \varepsilon_{abc \dots x} (\mathbf{e}_2)_b (\mathbf{e}_3)_c \dots (\mathbf{e}_N)_x = \varepsilon_{abc \dots x} (\mathbf{e}_1)_a (\mathbf{e}_2)_b (\mathbf{e}_3)_c \dots (\mathbf{e}_N)_x \\ &= \det [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \dots, \mathbf{e}_N] = \det(\mathbf{S}) \end{aligned}$$

where x is the N^{th} letter of the alphabet, so that

$$\text{volume}(N) = |\det(\mathbf{S})| = |\det [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \dots, \mathbf{e}_N]| = |\varepsilon_{abc \dots x} (\mathbf{e}_1)_a (\mathbf{e}_2)_b (\mathbf{e}_3)_c \dots (\mathbf{e}_N)_x|$$

These expressions have 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 above.

(e) Summary of Main Results of this Appendix

1. One way to write the equation of a plane in N dimensions is

$$\mathbf{r} \cdot \hat{\mathbf{p}} = p \quad \mathbf{r} = (x_1, x_2, \dots, x_N)$$

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 $\mathbf{r} \cdot \hat{\mathbf{p}} = 0$ where $\hat{\mathbf{p}}$ is either normal to the plane.

2. An N-piped has 2^N vertices as demonstrated by the inductive construction method presented above.

3. The locus of points making up the (closed) interior of an N-piped spanned by $\mathbf{e}_1 \dots \mathbf{e}_N$ is given by

$$\mathbf{r}_{\text{volumeN}} = \sum_{n=1}^N \alpha_n \mathbf{e}_n \quad 0 \leq \alpha_n \leq 1$$

The tails of all the vectors $\mathbf{e}_1 \dots \mathbf{e}_N$ 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

$$\begin{aligned} \mathbf{r}_{\text{face}(i)} &= \sum_{n \neq i} \alpha_n \mathbf{e}_n & 0 \leq \alpha_n \leq 1 & \quad i = 1, 2, \dots, N \\ \mathbf{r}_{\text{face}(ip)} &= \sum_{n \neq i} \alpha_n \mathbf{e}_n + \mathbf{e}_i & 0 \leq \alpha_n \leq 1 & \quad i = 1, 2, \dots, N \end{aligned}$$

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 $\mathbf{E}_i \cdot \mathbf{e}_j = \delta_{i,j}$ where $\mathbf{E}_i$ are the reciprocal base vectors. In this case, the equations of the faces of the N-piped can be written in the form shown in item 1 above,

$$\hat{\mathbf{E}}_i \cdot \mathbf{r}_{(\text{face } i)} = 0 \quad \hat{\mathbf{E}}_i \cdot \mathbf{r}_{(\text{face } ip)} = 1/E_i \quad i = 1, 2, \dots, N$$

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).

7. The out-facing vector area of face ip of an N-piped can be expressed as

$$\mathbf{A}_i = |\det(\mathbf{S})| \mathbf{E}_i$$

$$\mathbf{A}_i = \sigma (-1)^{i-1} \prod_{j \neq i} \mathbf{e}_j$$

$$\sigma \equiv \text{sign}[\det(\mathbf{S})] = \text{sign}[\det(\mathbf{R})]$$

$$\mathbf{A}_i = \sigma (-1)^{i-1} \mathbf{e}_1 \times \mathbf{e}_2 \dots \times \mathbf{e}_N \quad // \mathbf{e}_i \text{ missing}$$

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 Appendix A (i),

$$(\mathbf{A}_i)_\alpha = \sigma (-1)^{i-1} \varepsilon_{\alpha a b c \dots x} (e_1)_a (e_2)_b \dots (e_N)_x \quad // \text{where } (e_i)_i \text{ and } i \text{ are missing}$$

For $N=2$ the last two expressions for $\mathbf{A}_i$ are interpreted as shown in Appendix A (c)

$$\mathbf{A}_i = \sigma (-1)^{i-1} \mathbf{e}_{3-i} \times \hat{\mathbf{3}} \quad i = 1, 2$$

8. The volume of an N -piped spanned by $\mathbf{e}_1 \dots \mathbf{e}_N$ is given by

$$\text{volume}(N) = |\det(\mathbf{S})| = |\det [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \dots, \mathbf{e}_N]| = |\varepsilon_{a b c \dots x} (e_1)_a (e_2)_b (e_3)_c \dots (e_N)_x|$$

where one can regard the tangent base vectors as the columns of the linearized transformation matrix $\mathbf{S}$.

Appendix C: Elliptical Polar Coordinates (N=2, non-orthogonal)

This Appendix is written in the developmental notation of Sections 1-6.

(a) 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 $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ is given by

		<u>x'-space</u>	<u>x-space (Cartesian)</u>
$\rho^2 = x^2/a^2 + y^2/b^2$	$x^2 + y^2 = r^2$ still	$x'_1 = \theta$	$x_1 = x$
$\tan\theta = y/x$		$x'_2 = \rho$	$x_2 = y$

Writing the first equation above as

$$1 = x^2/(\rho a)^2 + y^2/(\rho b)^2$$

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 $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$ is given by

$x = \rho a \cos\theta$	$x/a = \rho \cos\theta$	$\Rightarrow$	$x^2/a^2 + y^2/b^2 = \rho^2$
$y = \rho b \sin\theta$	$y/b = \rho \sin\theta$	$\Rightarrow$	$\tan\theta = y/x$

The matrix S is given by

$S_{11} = (\partial x / \partial \theta) = -\rho a \sin\theta$	$S_{i\mathbf{k}} \equiv (\partial x_i / \partial x'_{\mathbf{k}})$
$S_{12} = (\partial x / \partial \rho) = a \cos\theta$	
$S_{21} = (\partial y / \partial \theta) = \rho b \cos\theta$	
$S_{22} = (\partial y / \partial \rho) = b \sin\theta$	

$$S = \begin{pmatrix} -\rho a \sin\theta & a \cos\theta \\ \rho b \cos\theta & b \sin\theta \end{pmatrix} \quad \Rightarrow \quad \det(S) = -\rho ab \quad \text{and} \quad R = S^{-1} = \begin{pmatrix} -\sin\theta/(\rho a) & \cos\theta/(\rho b) \\ \cos\theta/a & \sin\theta/b \end{pmatrix}$$

The tangent base vectors $\mathbf{e}_n$ can be read off as the columns of S

$\mathbf{e}_1 = \rho(-a \sin\theta, b \cos\theta) = \mathbf{e}_\theta$	$ \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$	$\mathbf{e}_\rho = \mathbf{e}_\rho \hat{\mathbf{e}}_\rho$

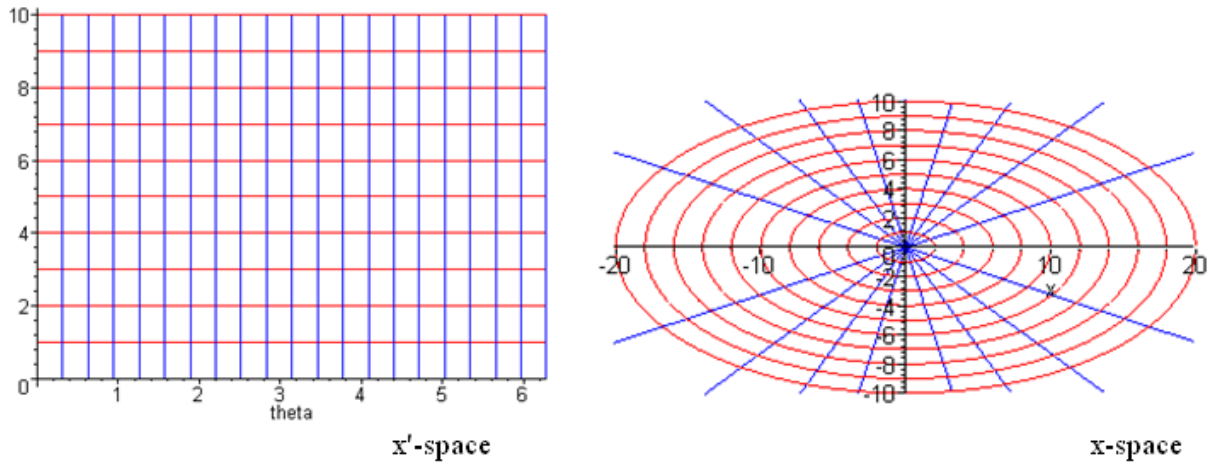
The covariant metric tensor is,

$$\bar{g}' = S^T S = \begin{pmatrix} \rho^2 \{a^2 \sin^2(\theta) + b^2 \cos^2(\theta)\} & [b^2 - a^2] \rho \sin(\theta) \cos(\theta) \\ [b^2 - a^2] \rho \sin(\theta) \cos(\theta) & a^2 \cos^2(\theta) + b^2 \sin^2(\theta) \end{pmatrix} = \begin{pmatrix} e_1 \bullet e_1 & e_1 \bullet e_2 \\ e_2 \bullet e_1 & e_2 \bullet e_2 \end{pmatrix}$$

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 $\bar{g}'$ is non-diagonal.

(b) Forward coordinate lines

Here is a Maple plot of some x-space forward coordinate lines (parameters $a = 2$ and $b = 1$)



The coordinate lines in x-space are plotted using these equations,

$y = b \sqrt{\rho_i^2 - (x/a)^2}$	// ellipses	$\rho_i = 1, 2, \dots, 10$	10 ellipses
$y = x \tan \theta_i$	// rays	$\theta_i = 2\pi (i/20), i = 1, 2, \dots, 20$	20 rays

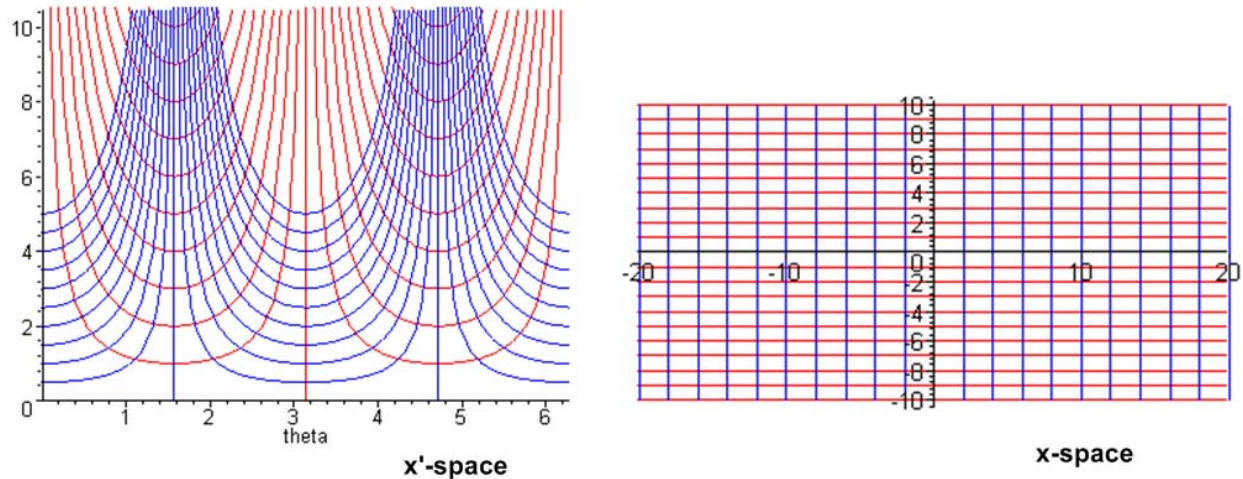
which are obtained from the forward transformation equations

$$\rho^2 = x^2/a^2 + y^2/b^2$$

$$\tan \theta = y/x$$

(c) 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

$$\begin{array}{lll} \rho = x_i / (a \cos \theta) & x_i = -10 \text{ to } +10 & 21 \text{ blue curves (one is a boxy U)} \\ \rho = y_i / (b \sin \theta) & y_i = -10 \text{ to } +10 & 21 \text{ red curves (one is a boxy U)} \end{array}$$

which are obtained from the inverse transformation equations

$$\begin{aligned} x &= a\rho \cos \theta \\ y &= b\rho \sin \theta \end{aligned}$$

The $\sec \theta$ and $\csc \theta$ 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.

(d) Drawing a contravariant vector $\mathbf{V}$ in x -space: the meaning of $\mathbf{V}'_n$.

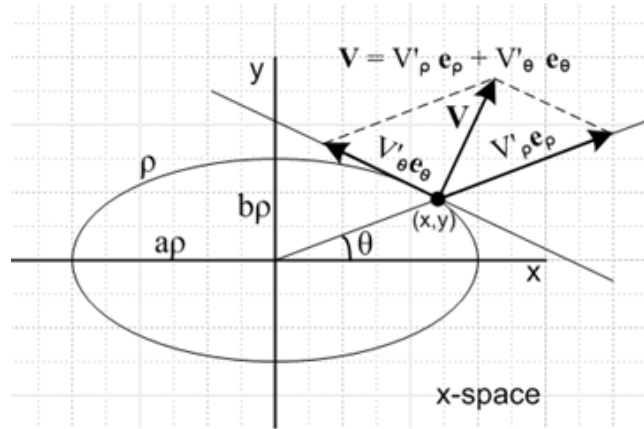
A contravariant vector field $\mathbf{V}(\mathbf{x})$ can be expanded in these two ways (Section 6 (f))

$$\mathbf{V} = V_1(\mathbf{x}) \hat{\mathbf{1}} + V_2(\mathbf{x}) \hat{\mathbf{2}} = V_x(\mathbf{x}) \hat{\mathbf{x}} + V_y(\mathbf{x}) \hat{\mathbf{y}} \quad // \quad \mathbf{u}_n = \hat{\mathbf{n}} \text{ for Cartesian}$$

$$\mathbf{V} = V'_1(\mathbf{x}') \mathbf{e}_1 + V'_2(\mathbf{x}') \mathbf{e}_2 = V'_\theta(\mathbf{x}') \mathbf{e}_\theta + V'_\rho(\mathbf{x}') \mathbf{e}_\rho \quad \mathbf{V}'(\mathbf{x}') = \mathbf{R}(\mathbf{x}) \mathbf{V}(\mathbf{x})$$

where the V'_n are the components of $\mathbf{V}$ transformed into x' -space where $\mathbf{V}$ becomes $\mathbf{V}'$. The prime is not necessary on V'_ρ but we maintain it as a reminder that it is an x' -space component. $\mathbf{R}(\mathbf{x})$ is the matrix of Section 2 and $\mathbf{e}_n$ are the tangent base vectors of Section 3. The fields $V'_n(\mathbf{x}')$ are "the components of vector field $\mathbf{V}'$ in x' -space", since $\mathbf{V}' = \mathbf{R}\mathbf{V}$, or $V'_i = R_{ij}V_j$. Moreover, these $V'_n(\mathbf{x}')$ are "expressed in terms of curvilinear coordinates" (Section 2(i)). If one is asked to "express a vector $\mathbf{V}$ in curvilinear coordinates", one is usually being asked to write $\mathbf{V}$ as the second expansion above. The vectors $\mathbf{e}_n$ and $\mathbf{V}$ exist *in x -space*, and in the second expansion it just happens that the coefficients $V'_n(\mathbf{x}')$ are the components of $\mathbf{V}'$, which is a vector in x' -space, and are functions of the curvilinear coordinates.

Here is a graphical representation of this vector $\mathbf{V}$ in x-space:

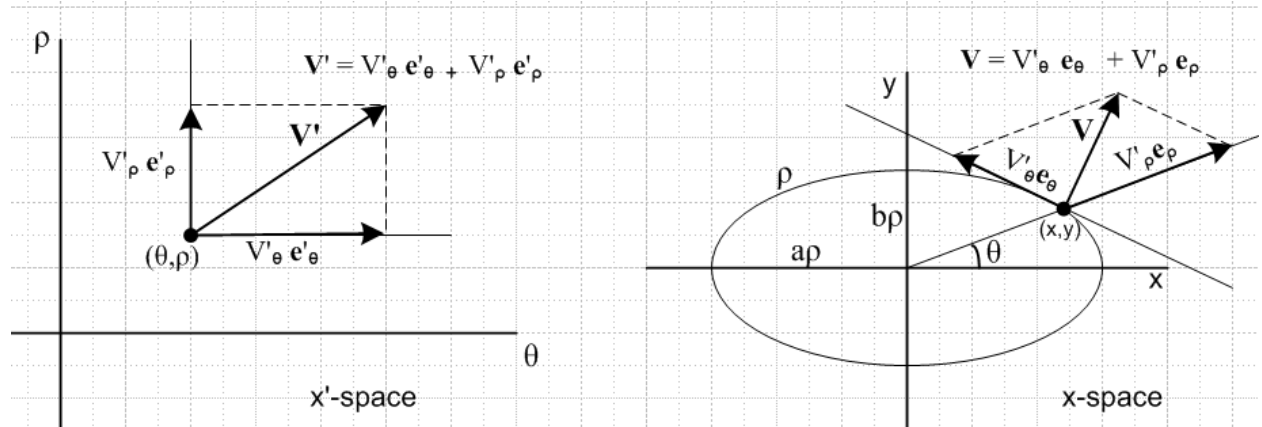


As advertised, the tangent base vectors are not at right angles. The $\mathbf{V}$ parallelogram accurately illustrates the equation $\mathbf{V} = V'_\theta \mathbf{e}_\theta + V'_\rho \mathbf{e}_\rho$. Since x-space is Cartesian, there is no distinction between Cartesian length (graphical length) and 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 $\mathbf{V}$ 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 $\mathbf{R}$ and use $V'_i = R_{ij} V_j$.

(e) Drawing a contravariant vector $\mathbf{V}'$ in x'-space: the two "views"

As with previous examples, the above picture is drawn to the right of an x'-space picture as follows:



In Section 3 the basis vectors $\mathbf{e}'_n$ were introduced as

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

and $\mathbf{e}_n$ was shown to be a contravariant vector,

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

Applying matrix $R(\mathbf{x})$ to the above equation $\mathbf{V} = V'_\theta \mathbf{e}_\theta + V'_\rho \mathbf{e}_\rho$ one gets

$$\mathbf{V}' = V'_\theta \mathbf{e}'_\theta + V'_\rho \mathbf{e}'_\rho$$

which appears first in the list of expansions of $\mathbf{V}'$ in Section 6 (f). 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 $g' \neq 1$. These views are really two *different* x' -spaces since the metric tensors are different.

In the Cartesian view of x' -space the length (norm) of a vector is given by $|\mathbf{A}|^2 = \delta_{ij} A_i A_j = \Sigma A_i^2$, so one has, since $(\mathbf{e}'_n)_i = \delta_{n,i}$,

$$\bar{g}' = 1 \quad |\mathbf{e}'_n| = 1 \quad \mathbf{e}'_n = \hat{\mathbf{e}}'_n = \hat{\mathbf{n}}' = \text{the usual axis-aligned unit vectors in } x'\text{-space}$$

$$\mathbf{V}' = V'_\theta \hat{\boldsymbol{\theta}} + V'_\rho \hat{\boldsymbol{\rho}} \quad \hat{\boldsymbol{\theta}} = \hat{\mathbf{e}}'_1 \quad \hat{\boldsymbol{\rho}} = \hat{\mathbf{e}}'_2$$

The left-side graph shown above 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

$$\hat{\boldsymbol{\theta}} \cdot \hat{\boldsymbol{\theta}} = 1$$

$$\hat{\boldsymbol{\rho}} \cdot \hat{\boldsymbol{\rho}} = 1$$

$$\hat{\boldsymbol{\theta}} \cdot \hat{\boldsymbol{\rho}} = 0$$

Although easy to understand, the graph in this Cartesian view has little utility except in a special case to be mentioned below involving integration. For example, the quantity $|\mathbf{V}'|$ shown above is not very useful since it is the Cartesian norm in a non-Cartesian space.

In the Curvilinear view of x' -space the length $|\mathbf{A}|$ of a contravariant vector $\mathbf{A}$ is the "covariant length" given by $|\mathbf{A}|^2 = \bar{g}'_{ij} A_i A_j$, and one has

$$\bar{g}' \neq 1 \quad |\mathbf{e}'_n| = |\mathbf{e}_n| = h'_n \equiv \sqrt{\bar{g}'_{nn}} \quad n = 1, 2 \text{ for } \theta, \rho$$

$$\mathbf{V}' = V'_\theta \mathbf{e}'_\theta + V'_\rho \mathbf{e}'_\rho = V'_\theta h'_\theta \hat{\mathbf{e}}'_\theta + V'_\rho h'_\rho \hat{\mathbf{e}}'_\rho \quad \hat{\mathbf{e}}'_n \equiv \mathbf{e}'_n / |\mathbf{e}'_n| = \mathbf{e}'_n / h'_n$$

$$|\mathbf{V}'|^2 = |\mathbf{V}|^2 = V_x^2 + V_y^2 \neq (V'_\theta h'_\theta)^2 + (V'_\rho h'_\rho)^2 \quad // \text{ unless } x'_i \text{ are orthogonal coordinates}$$

This last inequality says that in the curvilinear view the Pythagorean Theorem is invalid. In fact

$$\begin{aligned}
|\mathbf{V}'|^2 &= \overline{g}'_{ij} V'_i V'_j = \overline{g}'_{\theta\theta} V'^2_{\theta} + \overline{g}'_{\rho\rho} V'^2_{\rho} + 2 \overline{g}'_{\theta\rho} V'_{\theta} V'_{\rho} \\
&= (V'_{\theta} h'_{\theta})^2 + (V'_{\rho} h'_{\rho})^2 + 2 \overline{g}'_{\theta\rho} V'_{\theta} V'_{\rho}
\end{aligned}$$

In writing $|\mathbf{e}'_n| = |\mathbf{e}_n|$ and $|\mathbf{V}'|^2 = |\mathbf{V}|^2$ above, we use the rule shown in Section 5 (i) which says $|\mathbf{A}'|^2 = |\mathbf{A}|^2$ for any contravariant vector $\mathbf{A}$ ($|\mathbf{A}|^2$ is a scalar). Moreover,

$$\begin{aligned}
\hat{\mathbf{e}}'_{\theta} \cdot \hat{\mathbf{e}}'_{\theta} &= \mathbf{e}'_{\theta} \cdot \mathbf{e}'_{\theta} / (h'^2_{\theta}) = \mathbf{e}_{\theta} \cdot \mathbf{e}_{\theta} / (h'^2_{\theta}) = \overline{g}'_{\theta\theta} / (h'^2_{\theta}) = 1 \\
\hat{\mathbf{e}}'_{\rho} \cdot \hat{\mathbf{e}}'_{\rho} &= \mathbf{e}'_{\rho} \cdot \mathbf{e}'_{\rho} / (h'^2_{\rho}) = \mathbf{e}_{\rho} \cdot \mathbf{e}_{\rho} / (h'^2_{\rho}) = \overline{g}'_{\rho\rho} / (h'^2_{\rho}) = 1 \\
\hat{\mathbf{e}}'_{\theta} \cdot \hat{\mathbf{e}}'_{\rho} &= \mathbf{e}'_{\theta} \cdot \mathbf{e}'_{\rho} / (h'_{\theta} h'_{\rho}) = \mathbf{e}_{\theta} \cdot \mathbf{e}_{\rho} / (h'_{\theta} h'_{\rho}) = \overline{g}'_{\theta\rho} / (h'_{\theta} h'_{\rho}) \neq 0 \quad \Leftarrow !!
\end{aligned}$$

so that the $\hat{\mathbf{e}}'_n$ are unit vectors having unit covariant length, but $\hat{\mathbf{e}}'_{\theta} \cdot \hat{\mathbf{e}}'_{\rho} \neq 0$ despite the fact that these vectors are drawn at right angles in the x' -space graph above, $\mathbf{e}_n = h'_n \hat{\mathbf{e}}'_n$. At least in an orthogonal curvilinear system $\hat{\mathbf{e}}'_{\theta} \cdot \hat{\mathbf{e}}'_{\rho} = 0$ since then $\overline{g}'$ is diagonal. In the Curvilinear view, one regards the x' -space graph above as representing the fact that $\mathbf{V}' = V'_{\rho} \mathbf{e}'_{\rho} + V'_{\theta} \mathbf{e}'_{\theta}$, but one must realize that the drawing is inaccurate graphically speaking. One might imagine 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.

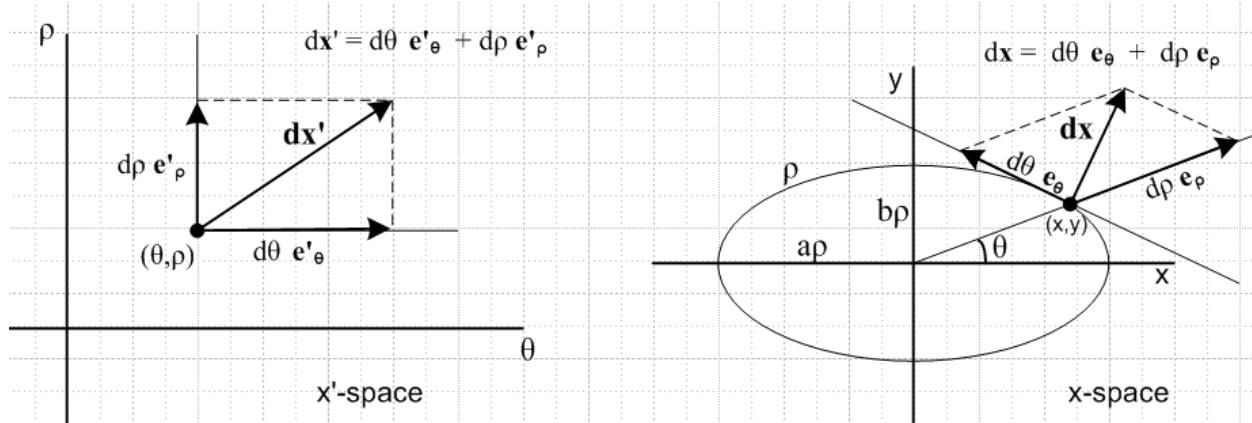
We have discussed the two views of the x' -space graph to avoid confusion between them. It is normally the Curvilinear view that we take of x' -space since it really does have $\overline{g}' \neq 1$. However, in the integration discussion of section (h) below, the Cartesian-view will be applied for differential vectors.

(f) Drawing the specific contravariant vector $d\mathbf{x}$ in x -space and x' -space

Since $d\mathbf{x}$ is the primordial contravariant vector, everything stated in the last two sections applies with $\mathbf{V} \rightarrow d\mathbf{x}$ and $V'_{\theta} \rightarrow dx'_{\theta} = d\theta$, $V'_{\rho} \rightarrow dx'_{\rho} = d\rho$, where we finally drop the primes on $d\theta$ and $d\rho$. The expansions of $d\mathbf{x}$ and $d\mathbf{x}'$ are,

$$\begin{aligned}
d\mathbf{x} &= d\theta \mathbf{e}_{\theta} + d\rho \mathbf{e}_{\rho} & // \text{ in } x\text{-space} \\
d\mathbf{x}' &= d\theta \mathbf{e}'_{\theta} + d\rho \mathbf{e}'_{\rho} & // \text{ in } x'\text{-space}
\end{aligned}$$

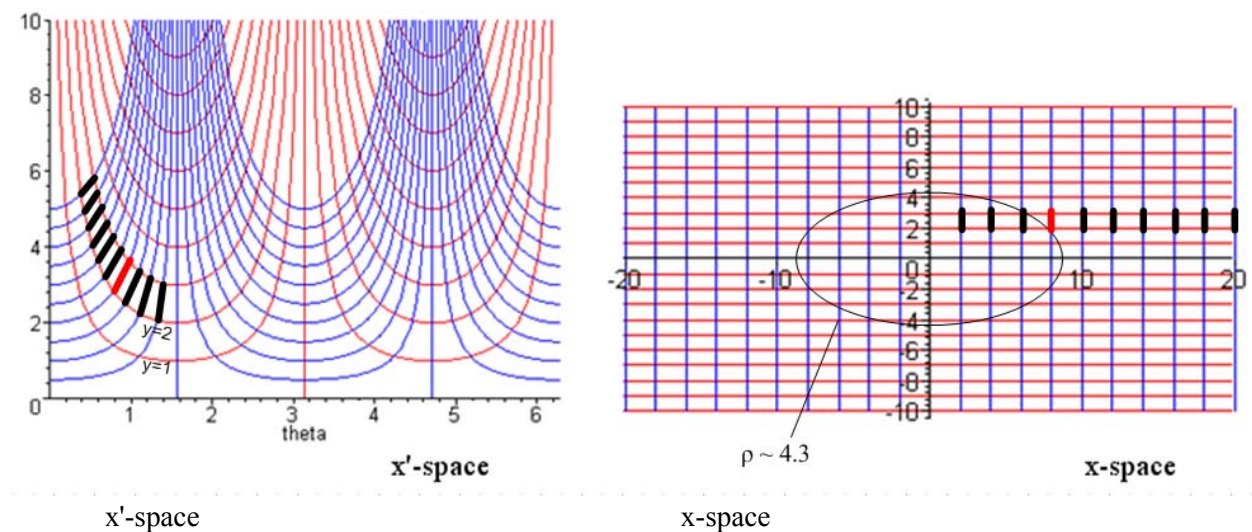
For $\mathbf{V} = d\mathbf{x}$ the picture above becomes



It must be understood that now the vector arrows like dx are highly magnified and in reality are very small compared to, say, the curvature of the ellipse.

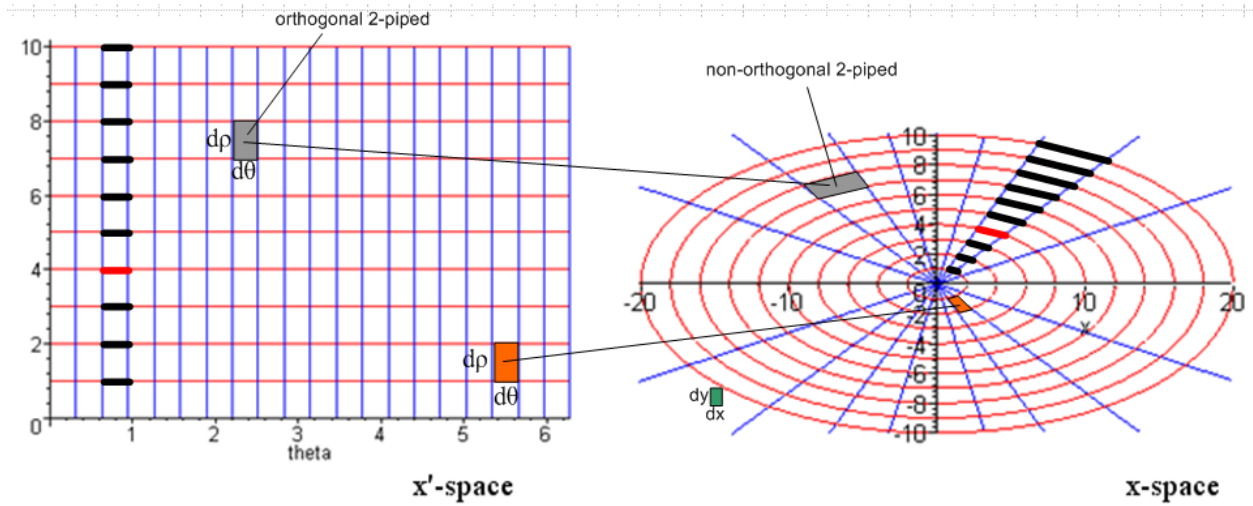
(g) Study of how dx transforms in the mapping between x -space and x' -space

Consider this drawing which shows a representative set of vectors dx in x -space (the bars), along with the forward mappings ($dx' = F(dx)$ or $dx' = Rdx$) of the corresponding vectors dx' in x' -space. The vectors on the right all point up, those on the left point generally to the northeast.



Now select the red dx bar on the right and operationally apply *the previous* picture. First determine the tangent base vectors e_θ and e_ρ at the location of the red bar. Then setting $dx = d\theta e_\theta + d\rho e_\rho$, consider the value of the two numbers $d\theta$ and $d\rho$ for this red bar. Graphically, knowing which way e_θ and e_ρ point at the bottom of the red dx , one expects $d\theta > 0$ and $d\rho > 0$. The red dx' bar on the left has these Cartesian values $d\theta$ and $d\rho$, and has a Cartesian-view length of $|dx|^2 = (d\theta)^2 + (d\rho)^2$. 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 *covariant* lengths of the x' -space bars are all the same, and are equal to the Cartesian length of those bars in x -space, since $dx' \cdot dx' = dx \cdot dx$.

Consider now some bar mapping in the other direction:



Now the dx 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 covariant length matches that of its corresponding dx 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.

(h) Derivation of the Jacobian Integration Rule.

Consider now an integral $\int d\theta dp f(\theta, \rho)$. The tiny rectangles of area $d\theta dp$, 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 Section 8]

$$dA' \equiv dp d\theta = \text{the area of a differential patch in the } \textit{Cartesian} \text{ view of } x'\text{-space}$$

This is the graphical area one sees in the picture. There is no need to define or consider any "covariant area" in x' -space because the Cartesian area is being used.

In the limiting process which defines the integration, each $d\theta dp$ patch on the left has the same area $dp d\theta$. 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 Section 8 (e), the ratio of the two patch areas is the absolute value of the Jacobian $|J(\mathbf{x})|$,

$$(\text{area of skewed patch on the right at location } \mathbf{x}) = |J(\mathbf{x})| dA' = |J(\mathbf{x})| dp d\theta$$

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 $dx dy$.

The mapping shown above between patches is an $N=2$ example of the general N -dimensional discussion in Section 8 (a) 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:

$$\text{integral of } f(\mathbf{x}) = \lim \sum_i dA_1(\mathbf{x}_i) f(\mathbf{x}_i) \quad dA_1(\mathbf{x}_i) = \text{patches shown on the right above}$$

$$\text{integral of } f(\mathbf{x}) = \lim \sum_i dA_2(\mathbf{x}_i) f(\mathbf{x}_i) \quad dA_2(\mathbf{x}_i) = dx dy$$

In the first integral, 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 $|J(\mathbf{x}')|d\theta d\rho$ and the areas are different because the Jacobian is a function of $\mathbf{x} = \mathbf{x}(\mathbf{x}')$.

In the second integral, every patch $dA_1(\mathbf{x}_i)$ has the same area $dx dy$, so really $dA_1(\mathbf{x}_i)$ does not depend on $\mathbf{x}_i$ in this form of the integration. One such $dx dy$ patch is shown in green above. The coverage of the dA_1 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 $dx dy$ patch that approximates it in area, like the green patch. The statement is about an integration. Thus one has

$$\lim \sum_i dA_1(\mathbf{x}_i) f(\mathbf{x}_i) = \lim \sum_i dA_2(\mathbf{x}_i) f(\mathbf{x}_i)$$

or

$$\int [|J(\mathbf{x}')| d\theta d\rho] f(\mathbf{x}(\mathbf{x}')) = \int [dx dy] f(\mathbf{x})$$

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

$$|J(\theta, \rho)| d\rho d\theta = dx dy$$

where the meaning of this symbolic equality is the integral statement above,

$$\int_D dx dy f(\mathbf{x}) = \int_{D'} d\theta d\rho |J(\mathbf{x}')| f(\mathbf{x}(\mathbf{x}')) ,$$

valid for any integrable $f(\mathbf{x})$ and any integration region D (region D' corresponds to D in $\mathbf{x}'$ -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

$$dA \equiv dx dy$$

so the above equality of integrals can be written

$$\int dA f(\mathbf{x}) = \int dA' |J(\mathbf{x}')| f(\mathbf{x}(\mathbf{x}')) \quad |J(\mathbf{x}')| dA' = dA$$

In N dimensions, dA and dA' are differential "volumes", and the general Jacobian Integration rule takes the form

$$\int dV f(\mathbf{x}) = \int dV' |J(\mathbf{x}')| f(\mathbf{x}(\mathbf{x}')) \quad |J(\mathbf{x}')| dV' = dV$$

$dV' \equiv dx'_1 dx'_2 \dots dx'_N$ = the volume of an orthogonal differential N-piped
in the Cartesian-view x' -space

$dV \equiv dx_1 dx_2 \dots dx_N$ = the volume of an orthogonal differential N-piped in x -space.

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 with that mapping which involved a non-orthogonal N-piped in x -space.

To finish off our sample N=2 case, recall from earlier that

$$|J'(x')| = |\det(S)| = ab\rho$$

and therefore

$$\int dx dy f(x,y) = \int d\theta d\rho |J(\mathbf{x}')| f(a\rho \cos\theta, b\rho \sin\theta) = ab \int d\theta d\rho \rho f(a\rho \cos\theta, b\rho \sin\theta)$$

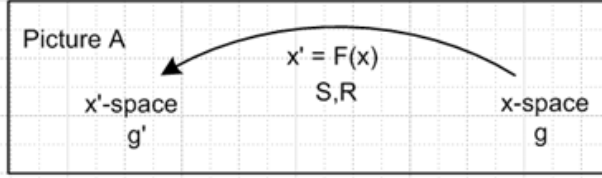
In the limit of regular polar coordinates, one then has $a = b = 1$ and $\rho = r$ so

$$\int dx dy f(x,y) = \int r dr d\theta f(r \cos\theta, r \sin\theta)$$

which is the familiar result.

Appendix D. Tensor Densities and the ϵ tensor

Picture A is used in this Appendix:



(a) Definition of a tensor density

First, recall from the Section 5 (k) discussion of the Jacobian J ,

$$J \equiv \det(S^i_j) = \sigma \sqrt{sg'} / \sqrt{sg} = \sigma (sg'/sg)^{1/2} = \sigma (g'/g)^{1/2} \quad \Rightarrow \quad (g'/g)^{1/2} = \sigma J = |J| > 0$$

$$\begin{aligned} s &= \text{sign}[\det(g_{ij})] = \text{sign}(g) = \text{sign}(g') & g &= \det(g_{ij}) & sg &= |g| > 0 & S^i_j &\equiv (\partial x^i / \partial x'^j) \\ \sigma &= \text{sign}[\det(S^i_j)] = \text{sign}(J) & g' &= \det(g'_{ij}) & sg' &= |g'| > 0 \end{aligned}$$

For proper Lorentz transformations of special relativity, $\det(S) = 1$ so $\sigma = +1$. For curvilinear coordinates, one normally selects an ordering of the x_i so that $\sigma = +1$, such as r, θ, ϕ in spherical coordinates. Nevertheless, to allow for the possibility of $J < 0$, $|J|$ is used below in place of the usual J seen in other sources. (There are other ways to handle this issue.)

Second, recall our generic sample tensor transformation from Section 7 (j)

$$T'^{abc}_{de} = R^a_a R^b_b R^c_c S^{d'}_d S^{e'}_e T^{a'b'c'}_{d'e'}$$

which can be rewritten in a more standard way using the theorem of Section 7 (q) that $S^\mu_\nu = R_\nu^\mu$,

$$T'^{abc}_{de} = R^a_a R^b_b R^c_c R^{d'}_d R^{e'}_e T^{a'b'c'}_{d'e'}$$

T is a mixed rank-5 tensor, meaning it transforms as shown above with respect to the underlying transformation F . T is an "ordinary" standard-issue tensorial tensor.

Now suppose instead that the object T were to transform like this,

$$T'^{abc}_{de} = |J|^{-W} R^a_a R^b_b R^c_c R^{d'}_d R^{e'}_e T^{a'b'c'}_{d'e'}$$

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

$$|\partial x' / \partial x|^W \equiv [\det(\partial x' / \partial x)]^{+W} = [\det(\partial x'_i / \partial x_k)]^{+W} = J^{-W}$$

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

$$g' = |J|^2 g = |J|^{-(-2)} g \Rightarrow \text{weight}(g) = -2$$

This is the scalar density mentioned in Section 5 (k) of weight -2 . Therefore,

$$g'^{-1} = |J|^{-(2)} g^{-1} \Rightarrow \text{weight}(g^{-1}) = +2$$

(b) 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 .
2. **Contracting** indices within a tensor does not alter its weight W . If indices a and d are contracted in the example above, one gets

$$\begin{aligned} T^{abc}{}_{ae} &= |J|^{-W} R^a{}_a R^b{}_b R^c{}_c R^{d'}{}_{e'} T^{a'b'c'}{}_{d'e'} \\ &= |J|^{-W} (R^a{}_a R^{d'}) R^b{}_b R^c{}_c R^{e'}{}_{e'} T^{a'b'c'}{}_{d'e'} \\ &= |J|^{-W} \delta_a{}^{d'} R^b{}_b R^c{}_c R^{e'}{}_{e'} T^{a'b'c'}{}_{d'e'} \\ &= |J|^{-W} R^b{}_b R^c{}_c R^{e'}{}_{e'} T^{a'b'c'}{}_{a'e'} \end{aligned}$$

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. For example,

$$\begin{aligned} A^{ia} &= |J|^{-W1} R^a{}_a A^{a'} \\ B^{ic}{}_d &= |J|^{-W2} R^c{}_c R^{d'}{}_d B^{c'}{}_{d'} \\ \Rightarrow (A^{ia} B^{ic}{}_d) &= |J|^{-(W1+W2)} R^a{}_a R^c{}_c R^{d'}{}_d (A^{a'} B^{c'}{}_{d'}) \end{aligned}$$

4. Although sometimes authors take a differing stance for certain tensors, 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. This means the g^{ab} **raises** an index and g_{ab} **lowers** an index.

5. Raising or lowering an index does not change the weight of a tensor density. Again, using our generic example above,

$$T^{abc}_{de} = |J|^{-W} R^a_a, R^b_b, R^c_c, R^d_{d'} R^{e'}_e T^{a'b'c'}_{d'e'}$$

$$T^{abc}_{d^e} = g^{ex} T^{abc}_{dx} \quad // \text{raise last index in x'-space}$$

$$T^{a'b'c'}_{d'e'} = g_{e'e''} T^{a'b'c'}_{d^e} \quad // \text{lower last index in x-space}$$

Therefore

$$\begin{aligned} T^{abc}_{d^e} &= g^{ex} [|J|^{-W} R^a_a, R^b_b, R^c_c, R^d_{d'} R^{e'}_x T^{a'b'c'}_{d'e'}] \\ &= g^{ex} [|J|^{-W} R^a_a, R^b_b, R^c_c, R^d_{d'} R^{e'}_x (g_{e'e''} T^{a'b'c'}_{d^e})] \\ &= |J|^{-W} R^a_a, R^b_b, R^c_c, R^d_{d'} (g^{ex} R^{e'}_x g_{e'e''}) T^{a'b'c'}_{d^e} \\ &= |J|^{-W} R^a_a, R^b_b, R^c_c, R^d_{d'} (R^{e'}_{e''}) T^{a'b'c'}_{d^e} \quad // \text{Section 7 (o)} \end{aligned}$$

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 .

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 $\mathbf{A} \bullet \mathbf{B}' = |J|^{W+w} \mathbf{A} \bullet \mathbf{B}$.

Proof: First form the rank-2 tensor density $A^i B_j$ which by item 3 has weight $W+w$. Lower the second index and the mixed rank-2 tensor $A^i B_j$ by item 5 still has weight $W+w$. Then contract to get $\mathbf{A} \bullet \mathbf{B} = A^i B_i$ and by item 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 $|\mathbf{A}'| = |J|^{-W} |\mathbf{A}|$

Proof: $|\mathbf{A}|^2 = \mathbf{A} \bullet \mathbf{A}$ which has weight $2W$ meaning $|\mathbf{A}'|^2 = |J|^{-2W} |\mathbf{A}|^2$. Therefore $|\mathbf{A}'| = |J|^{-W} |\mathbf{A}|$.

7. Weight Changing Theorem: Suppose T is a tensor density of weight W , and suppose some tensor-like object U is related to T in this manner, where k and w are any constants,

$$\begin{aligned} U^{abc}_{de} &= k (sg)^{w/2} T^{abc}_{de} = k |g|^{w/2} T^{abc}_{de} \\ U^{abc}_{de} &= k (sg')^{w/2} T^{abc}_{de} = k |g'|^{w/2} T^{abc}_{de} \end{aligned}$$

The theorem states that U transforms as a tensor density of weight $W-w$.

One would say according to Section 7 (u) that this tensor density equation is "covariant".

Proof: Start as in item 5 with a generic sample tensor density

$$T^{abc}_{de} = |J|^{-W} R^a_a R^b_b R^c_c R^{d'}_d R^{e'}_e T^{a'b'c'}_{d'e'}$$

Then starting with the second line of the pair of lines above,

$$\begin{aligned} U^{abc}_{de} &= k(sg')^{w/2} T^{abc}_{de} = k(sg'/sg)^{w/2} (sg)^{w/2} T^{abc}_{de} = k(g'/g)^{w/2} (sg)^{w/2} T^{abc}_{de} \\ &= |J|^W k(sg)^{w/2} \{ |J|^{-W} R^a_a R^b_b R^c_c R^{d'}_d R^{e'}_e T^{a'b'c'}_{d'e'} \} \\ &= |J|^W \{ |J|^{-W} R^a_a R^b_b R^c_c R^{d'}_d R^{e'}_e k(sg)^{w/2} T^{a'b'c'}_{d'e'} \} \\ &= |J|^W \{ |J|^{-W} R^a_a R^b_b R^c_c R^{d'}_d R^{e'}_e U^{a'b'c'}_{d'e'} \} \\ &= |J|^{-(W-w)} R^a_a R^b_b R^c_c R^{d'}_d R^{e'}_e U^{a'b'c'}_{d'e'} \end{aligned} \quad \text{QED}$$

8. As $|J| \rightarrow 1$, all tensor tensities become true tensors. One could imagine some limiting process on the underlying transformation F such that the linearized transformation matrix R approaches an orthogonal matrix $RR^T = 1$ which then must have $\det(R) = \pm 1$. and thus $|\det(S)| = |J| = 1$. In this case $|J|^{-W} \rightarrow 1^{-W} = 1$ and therefore any tensor density, regardless of its weight W , becomes an ordinary tensor. An example is that the cross product considered in section (g) below of $N-1$ contravariant vectors becomes then an ordinary covariant vector. If $g=1$ in x -space, then $g'=RR^T = 1$ in x' -space and then that resulting vector can be considered either contravariant or covariant since both spaces are then Cartesian. This is the case with the ordinary 3D $\mathbf{A} = \mathbf{B} \times \mathbf{C}$ under rotations. One can think of the ϵ_{abc} as moving in this limit from a tensor density of weight -1 to an ordinary tensor of effective weight 0 .

(c) Theorem about Totally Antisymmetric Tensors: there is really only one

Theorem: Apart from a scalar factor, there exists only one totally antisymmetric (TA) tensor.

Proof: Suppose there were two TA tensors called $\epsilon^{abc\dots}$ and $r^{abc\dots}$. If two or more of the indices are equal, both tensors are 0, so for such index sets, one can say $r^{abc\dots} = f \epsilon^{abc\dots}$ where f is any finite function whatsoever. Consider now the case where all the indices are distinct, and therefore exhaust the set $123\dots N$, and consider $abc\dots$ to be a permutation of $123\dots N$ obtained by doing S pairwise swaps,

$$abc\dots = P(123\dots) \quad p = (-1)^S$$

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

$$\begin{aligned} r^{abc\dots} &= p r^{123\dots} & // \text{ for example, } r^{2134\dots} &= (-1)^1 r^{1234\dots} \\ \epsilon^{abc\dots} &= p \epsilon^{123\dots} \end{aligned}$$

Define scalar function $f \equiv r^{123\dots} / \epsilon^{123\dots}$, whatever it might be. Then

$$\begin{aligned} r^{abc\dots} &= p(f e^{123\dots}) \\ \varepsilon^{abc\dots} &= p e^{123\dots} \end{aligned}$$

and dividing these two equations one finds,

$$r^{abc\dots} = f \varepsilon^{abc\dots}$$

which has now been shown valid for all index sets $abc\dots$. Therefore, any "other" totally antisymmetric tensor is just a scalar function times the ε tensor.

(d) The contravariant ε tensor

Knowing nothing to start, assume that the famous $\varepsilon^{abc\dots}$ totally antisymmetric tensor transforms under F as a tensor density of some weight W which we hope to determine. Then

$$\varepsilon'^{abc\dots} = |J|^{-W} R^a_a{}^b_b{}^c_c{}^{\dots} \varepsilon^{a'b'c'\dots} \quad (*)$$

Assume that $\varepsilon^{abc\dots}$ is the usual permutation tensor normalized to $\varepsilon^{123\dots N} = +1$. 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, $\varepsilon = 0$. This is an important starting assumption, and from it most everything follows.

Given this assumption, the RHS of (*) is totally antisymmetric (TA). The argument is given once here and then used later several times. Consider an $a \leftrightarrow b$ swap. Then

$$\begin{aligned} \varepsilon'^{bac\dots} &= |J|^{-W} R^b_a{}^a_b{}^c_c{}^{\dots} \varepsilon^{a'b'c'\dots} = |J|^{-W} R^b_b{}^a_a{}^c_c{}^{\dots} \varepsilon^{b'a'c'\dots} \\ &= |J|^{-W} R^a_a{}^b_b{}^c_c{}^{\dots} (-\varepsilon^{a'b'c'\dots}) = -\varepsilon'^{abc\dots} \end{aligned}$$

The same result is true for any swap, thus RHS (*) = TA. Since according to section (b) there is only one TA tensor available apart from a scalar function factor, it follows that

$$|J|^{-W} R^a_a{}^b_b{}^c_c{}^{\dots} (-\varepsilon^{a'b'c'\dots}) = K \varepsilon^{abc\dots} = \text{RHS of equation } (*)$$

where K is some scalar function, perhaps just a constant. Equation (*) above then reads

$$K \varepsilon^{abc\dots} = |J|^{-W} R^a_a{}^b_b{}^c_c{}^{\dots} \varepsilon^{a'b'c'\dots} \quad (**)$$

Setting in the standard order, one finds that

$$K \varepsilon^{123\dots} = |J|^{-W} R^1_a{}^a R^2_b{}^b R^3_c{}^c{}^{\dots} \varepsilon^{a'b'c'\dots}$$

or

$$K = |J|^{-W} \det(R^i_j) = |J|^{-W} |J|^{-1} = |J|^{-(W+1)}$$

A priori, the general solution of this last equation is,

$$K = |J|^{-M} \quad W+1 = M \quad \text{or} \quad W = M-1$$

Since (*) and (**) have the same RHS, equating the LHS's says

$$\epsilon'^{abc\dots} = K \epsilon^{abc\dots}$$

A second assumption is now made: that $\epsilon^{abc\dots}$ 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. This forces $K = 1$, so that

$$\epsilon'^{abc\dots} = \epsilon^{abc\dots}$$

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 $\epsilon_{123\dots N} = \epsilon'_{123\dots N} = +1$, but we shall follow Weinberg.

Since $K = 1$, it follows that $M = 0$ so $W = M - 1 = -1$.

To summarize, assuming that $\epsilon^{abc\dots}$ is the usual permutation tensor normalized in the usual way, and assuming that $\epsilon'^{abc\dots} = \epsilon^{abc\dots}$ so this tensor is the same in all frames or spaces, THEN one concludes that $\epsilon^{abc\dots}$ must transform as a rank-N tensor density of weight $W = -1$. That is to say,

$$\epsilon'^{abc\dots} = |J| R^a{}_a{}' R^b{}_b{}' \dots \epsilon^{a'b'c'\dots}$$

This then is our second example of a tensor density. Viewed in this light, the tensor $\epsilon^{abc\dots}$ 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 is called the Levi-Civita connection.

(e) Some facts about the ϵ tensor

1. Consider (based on section (b) 4 above),

$$\epsilon_{abc\dots} = g_{aa'} g_{bb'} \dots \epsilon^{a'b'c'\dots}$$

This is again in the convention of Weinberg p 99 (4.4.10).

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)

$$\epsilon_{abc\dots} = s g_{aa'} g_{bb'} \dots \epsilon^{a'b'c'\dots} \quad s = \text{sign}[\det(g_{ij})]$$

Inserting such a sign renders any ϵ mixed tensor like $\epsilon^a{}_{bc\dots}$ ambiguous when $s = -1$, but is acceptable if one promises never to make use of a mixed ϵ tensor. We shall refer to these two methods as "the Weinberg convention" and the "added sign s convention", the former being assumed unless otherwise stated.

Install the reference sequence to obtain

$$\epsilon_{123\dots} = g_{1a} g_{2b} \dots \epsilon^{a'b'c'\dots} = \det(g_{ij}) = g \quad // \text{ Section 5 (k)}$$

Similarly, $\epsilon'_{123\dots} = \det(g'_{ij})$. To summarize,

$$\begin{aligned} \epsilon_{123\dots} &= \det(g_{ij}) = g & // \text{ } \epsilon \text{ all-down index reference values} \\ \epsilon'_{123\dots} &= \det(g'_{ij}) = g' \end{aligned}$$

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

$$\epsilon_{abc\dots} = g_{aa} g_{bb} \dots \epsilon^{a'b'c'\dots}$$

The RHS is a totally antisymmetric in indices abc... (see above) and can therefore be written

$$\text{RHS} = C \epsilon^{abc\dots}$$

since we showed earlier that there is only one TA tensor apart from scalar C. Therefore

$$\epsilon_{abc\dots} = C \epsilon^{abc\dots} \quad (*)$$

Insert the reference sequence

$$\epsilon_{123\dots} = C \epsilon^{123\dots} = C$$

But in 1 it was just showed that $\epsilon_{123\dots} = \det(g_{ij})$. Therefore

$$C = \det(g_{ij})$$

and then (*) says for the "Weinberg convention",

$$\begin{aligned} \epsilon_{abc\dots} &= \det(g_{ij}) \epsilon^{abc\dots} = g \epsilon^{abc\dots} & // \text{ relating all down to all up} \\ \epsilon'_{abc\dots} &= \det(g'_{ij}) \epsilon'^{abc\dots} = g' \epsilon'^{abc\dots} & // = g' \epsilon^{abc\dots} \end{aligned}$$

where the second line follows by the same argument. These equations relate all down to all up in the same space. Notice that both $\epsilon_{abc\dots}$ and $\epsilon'_{abc\dots}$ are totally antisymmetric.

In the "added sign s convention" the above equations are instead

$$\begin{aligned} \epsilon_{abc\dots} &= |\det(g_{ij})| \epsilon^{abc\dots} = |g| \epsilon^{abc\dots} & // \text{ relating all down to all up} \\ \epsilon'_{abc\dots} &= |\det(g'_{ij})| \epsilon'^{abc\dots} = |g'| \epsilon'^{abc\dots} & // = |g'| \epsilon^{abc\dots} \end{aligned}$$

3. Divide the last two Weinberg convention equations to find that

$$\varepsilon'_{abc\dots} = [\det(g'_{ij}) / \det(g_{ij})] \varepsilon_{abc\dots} = (g'/g) \varepsilon_{abc\dots}$$

From Section 5 (k) this says

$$\varepsilon'_{abc\dots} = |J|^2 \varepsilon_{abc\dots}$$

and this same conclusion is valid for the "added sign s convention" as well since $\det(g)$ and $\det(g')$ always have the same sign as shown in Section 5 (k).

Two comments:

- Although we set $\varepsilon'^{abc\dots} = \varepsilon^{abc\dots}$ by fiat, we cannot similarly set $\varepsilon'_{abc\dots} = \varepsilon_{abc\dots}$ by fiat. This latter result comes out being $\varepsilon'_{abc\dots} = |J|^2 \varepsilon_{abc\dots}$ as just shown.
- The fact that $\varepsilon'_{abc\dots} = |J|^2 \varepsilon_{abc\dots}$ does *not* say that ε_{abc} is a tensor density of weight -2 because there are no R factors showing. (See the section (a) definition of a tensor density transformation.)

(f) The covariant ε tensor

According to section (b) 5, lowering indices does not change the weight of a tensor density. Section (d) showed that $\varepsilon^{abc\dots}$ is a tensor density of weight -1, so we know right away that $\varepsilon_{abc\dots}$ is also a tensor density of weight -1. Nevertheless, it is interesting to see what happens when the same method used in section (c) for $\varepsilon^{abc\dots}$ is applied to $\varepsilon_{abc\dots}$.

We start by assuming $\varepsilon_{abc\dots}$ is a tensor density of some unknown weight W,

$$\varepsilon'_{abc\dots} = |J|^{-W} [R_a^{a'} R_b^{b'} \dots \varepsilon_{a'b'c'\dots}] \quad (*)$$

Section (e) 2 noted that $\varepsilon_{a'b'c'\dots}$ is a totally antisymmetric tensor, and therefore as in section (d) one concludes that the RHS of (*) is also totally antisymmetric and can be written as $\text{RHS}(*) = K \varepsilon_{abc\dots}$ so (*) then says

$$K \varepsilon_{abc\dots} = |J|^{-W} [R_a^{a'} R_b^{b'} \dots \varepsilon_{a'b'c'\dots}] \quad (**)$$

Use section (e) 2 to set $\varepsilon_{a'b'c'\dots} = \det(g_{ij}) \varepsilon^{a'b'c'\dots}$ inside the bracket,

$$K \varepsilon_{abc\dots} = |J|^{-W} [R_a^{a'} R_b^{b'} \dots \det(g_{ij}) \varepsilon^{a'b'c'\dots}] ,$$

and then install the reference sequence on both sides

$$K \varepsilon_{123\dots} = |J|^{-W} [R_1^{a'} R_2^{b'} \dots \det(g_{ij}) \varepsilon^{a'b'c'\dots}]$$

But section (e) 1 says $\varepsilon_{123\dots} = \det(g_{ij})$, so cancel $\det(g_{ij})$ on both sides to get

$$K = |J|^{-W} [R_1^{a'} R_2^{b'} \dots \varepsilon^{a'b'c' \dots}] = |J|^{-W} \det(R_i^j) = |J|^{-W} \det(S_i^j) = |J|^{-W} |J| = |J|^{-(W-1)}$$

So here the result is $K = |J|^{-(W-1)}$ whereas in section (d) the result was $K = |J|^{-(W+1)}$. In the current case, since (*) and (**) have the same RHS, setting the LHS's equal says

$$\varepsilon'_{abc\dots} = K \varepsilon_{abc\dots}$$

But from section (e) 3 one must select $K = |J|^2$ and not $K=1$. Then

$$K = |J|^{-(W-1)} \Rightarrow |J|^{-(-2)} = |J|^{-(W-1)} \Rightarrow W = -1$$

and the conclusion is that $\varepsilon_{abc\dots}$ transforms with weight -1, the same as $\varepsilon^{abc\dots}$, so (*) becomes

$$\varepsilon'_{abc\dots} = |J| [R_a^{a'} R_b^{b'} \dots \varepsilon_{a'b'c' \dots}]$$

(g) Generalized cross products

In Appendix A (c) the following cross product of N-1 vectors is considered (converted now to standard notation)

$$Q_a \equiv \varepsilon_{abc\dots x} B^b C^c D^d \dots X^x \quad \text{or} \quad \mathbf{Q} = \mathbf{B} \times \mathbf{C} \times \mathbf{D} \dots \times \mathbf{X}$$

If the vectors B,C,D...X are all contravariant vectors, then applying the rule of section (b) 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 tensor density of weight = -1, and thus has this transformation rule

$$Q'_a = |J| R_a^b Q_b$$

Similarly, one may consider

$$Q^a \equiv \varepsilon^{abc\dots x} B_b C_c D_d \dots X_x$$

If vectors B,C,D...X are covariant vectors, then Q^a is a vector density of weight -1 and

$$Q'^a = |J| R^a_b Q^b$$

(h) The tensorial nature of curl B

In Section 12 involving $\mathbf{C} \equiv \text{curl } \mathbf{B}$, it was found that

$$\begin{aligned} C^n &= \varepsilon^{nab} \partial_a B_b \\ C'^n &= (1/\sqrt{g'}) \varepsilon'^{nab} \partial'_a B'_b \end{aligned}$$

where $\mathbf{B}$ is assumed to be an ordinary tensorial vector. It is helpful to write these in the following manner

$$\begin{aligned} C^n &= (1/\sqrt{g}) \varepsilon^{nab} [\partial_a B_b - \partial_b B_a]/2 & // g = 1 \\ C'^n &= (1/\sqrt{g'}) \varepsilon'^{nab} [\partial'_a B'_b - \partial'_b B'_a]/2 \end{aligned}$$

Recall now from Section 7 (v) that the covariant derivative of a vector is given by

$$B_{b;a} = \partial_a B_b - \Gamma^c_{ab} B_c$$

where the affine connection Γ^c_{ab} is symmetric under $a \leftrightarrow b$. Therefore

$$B_{b;a} - B_{a;b} = [\partial_a B_b - \Gamma^c_{ab} B_c] - [\partial_b B_a - \Gamma^c_{ba} B_c] = \partial_a B_b - \partial_b B_a \equiv E_{ab}$$

The object E_{ab} is a rank-2 covariant tensor because both $B_{b;a}$ and $B_{a;b}$ are rank-2 covariant tensors,

$$\begin{aligned} E'_{ab} &= B'_{b;a} - B'_{a;b} = R_b^{b'} R_a^{a'} B_{b';a'} - R_a^{a'} R_b^{b'} B_{a';b'} \\ &= R_a^{a'} R_b^{b'} [B_{b';a'} - B_{a';b'}] = R_a^{a'} R_b^{b'} E_{ab} \end{aligned}$$

Therefore, our pair of equations can be written

$$\begin{aligned} C^n &= (1/\sqrt{g}) \varepsilon^{nab} [B_{b;a} - B_{a;b}]/2 \\ C'^n &= (1/\sqrt{g'}) \varepsilon'^{nab} [B'_{b;a} - B'_{a;b}]/2 \end{aligned}$$

and one more time using the ε antisymmetry ($s \equiv \text{sign}(\det(g)) = \text{sign}(\det(1)) = \text{sign } 1 = 1$.)

$$\begin{aligned} C^n &= (sg)^{-1/2} \varepsilon^{nab} B_{b;a} \\ C'^n &= (sg')^{-1/2} \varepsilon'^{nab} B'_{b;a} \end{aligned}$$

The object $B_{b;a}$ is an ordinary rank-2 tensor (weight 0), while ε^{nab} is a tensor density of weight -1. Therefore, from section (b) 3, the object $\varepsilon^{nab} B_{b;a}$ is a contravariant vector density of $W = -1$. The Weight Changing Theorem from section (b) 7 can now be applied with $w = -1$ and $k = 1$, so C^n has weight $W - w = (-1) - (-1) = 0$. The conclusion is that $\mathbf{C} = \text{curl } \mathbf{B}$ transforms as a regular tensorial vector under any transformation F . Similarly, $\mathbf{V} = \text{curl } \mathbf{C} = \text{curl}[\text{curl } \mathbf{B}]$, as encountered in Section 13, is also an ordinary tensorial vector.

(i) Tensor $\mathcal{E}$ as a weight 0 version of ε : three conventions

Equations in the "Weinberg Convention"

In this section it is assumed that g_{ij} and g^{ij} raise and lower indices of the ε tensor just as they do for any other tensor (Weinberg convention). In sections (d) and (e) it was established that

$$\begin{aligned} \varepsilon_{abc\dots} &= g \varepsilon^{abc\dots} & \varepsilon^{123\dots} &= +1 & \varepsilon_{123\dots} &= g \\ \varepsilon'_{abc\dots} &= g' \varepsilon'^{abc\dots} & \varepsilon'^{123\dots} &= +1 & \varepsilon'_{123\dots} &= g' \end{aligned}$$

$$\varepsilon'_{abc\dots} = |J|^2 \varepsilon_{abc\dots} = (g'/g) \varepsilon_{abc\dots} \quad \varepsilon \text{ is a rank-}N \text{ tensor of weight } W = -1$$

Again, just in passing, notice that $\varepsilon'_{abc\dots} = |J|^2 \varepsilon_{abc\dots}$ 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 $(sg = |g|, s = \text{sign}(g) = \text{sign}(g'))$ as in Section 5 (k))

$$\begin{aligned} \mathcal{E}_{abc\dots} &\equiv |g|^{-1/2} \varepsilon_{abc\dots} &\Rightarrow \quad \mathcal{E}_{123\dots} &= |g|^{-1/2} g = |g|^{-1/2} s |g| = s |g|^{1/2} \\ \mathcal{E}'_{abc\dots} &\equiv |g'|^{-1/2} \varepsilon'_{abc\dots} &\Rightarrow \quad \mathcal{E}'_{123\dots} &= |g'|^{-1/2} g' = |g'|^{-1/2} s |g'| = s |g'|^{1/2} \end{aligned}$$

According to the Weight Changing Theorem of section (b) 7 with $w = -1$ and $k = 1$ the object $\mathcal{E}_{abc\dots}$ is a tensor density with weight $W - w = (-1) - (-1) = 0$. Thus, $\mathcal{E}_{abc\dots}$ is an ordinary tensor.

Raising indices on both sides gives

$$\begin{aligned} \mathcal{E}^{abc\dots} &\equiv |g|^{-1/2} \varepsilon^{abc\dots} \Rightarrow \quad \mathcal{E}^{123\dots} = |g|^{-1/2} \\ \mathcal{E}'^{abc\dots} &\equiv |g'|^{-1/2} \varepsilon'^{abc\dots} \Rightarrow \quad \mathcal{E}'^{123\dots} = |g'|^{-1/2} \end{aligned}$$

To compare $\mathcal{E}_{abc\dots}$ and $\mathcal{E}^{abc\dots}$,

$$\begin{aligned} \mathcal{E}_{abc\dots} &= |g|^{-1/2} \varepsilon_{abc\dots} = |g|^{-1/2} g \varepsilon^{abc\dots} = s |g|^{-1/2} |g| \varepsilon^{abc\dots} = s |g|^{1/2} \varepsilon^{abc\dots} \\ \mathcal{E}^{abc\dots} &= |g|^{-1/2} \varepsilon^{abc\dots} \end{aligned}$$

so that,

$$\mathcal{E}_{abc\dots} = s |g| \mathcal{E}^{abc\dots} = g \mathcal{E}^{abc\dots}$$

Summarizing,

$$\begin{aligned} \mathcal{E}^{123\dots} &= |g|^{-1/2} & \mathcal{E}_{123\dots} &= s |g|^{+1/2} & \mathcal{E}_{abc\dots} &= g \mathcal{E}^{abc\dots} = s |g| \mathcal{E}^{abc\dots} \\ \mathcal{E}'^{123\dots} &= |g'|^{-1/2} & \mathcal{E}'_{123\dots} &= s |g'|^{+1/2} \end{aligned}$$

$$\begin{aligned} \mathcal{E}^{abc\dots} \mathcal{E}_{ABC\dots} &= |g|^{-1} \varepsilon^{abc\dots} \varepsilon_{ABC\dots} \\ \mathcal{E}'^{abc\dots} \mathcal{E}'_{ABC\dots} &= |g'|^{-1} \varepsilon'^{abc\dots} \varepsilon'_{ABC\dots} \end{aligned}$$

As a check, apply the Weight Changing Theorem with $w = -2$ and $k = 1$. Object $\mathcal{E}^{abc\dots} \mathcal{E}_{ABC\dots}$ has weight -2 since it is formed from two weight -1 ε densities, and then $\text{weight(LHS)} = W - w = (-2) - (-2) = 0$, so the object $\mathcal{E}^{abc\dots} \mathcal{E}_{ABC\dots}$ is an ordinary mixed rank-2N tensor.

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 (e). Here is the previous section redone in the "added sign s convention":

In section (e) it was established that $(g = \det(g_{ij}))$

$$\begin{aligned}
\varepsilon_{abc\dots} &= sg \varepsilon^{abc\dots} & \varepsilon^{123\dots} &= +1 & \varepsilon_{123\dots} &= sg = |g| & // g \rightarrow sg \\
\varepsilon'_{abc\dots} &= sg' \varepsilon'^{abc\dots} & \varepsilon'^{123\dots} &= +1 & \varepsilon'_{123\dots} &= sg' = |g'| & // g' \rightarrow sg' \\
\varepsilon'_{abc\dots} &= |J|^2 \varepsilon_{abc\dots} = (g'/g) \varepsilon_{abc\dots} & & & \varepsilon & \text{ is rank-N tensor of weight } W = -1 & // \text{ same}
\end{aligned}$$

Consider the following new objects defined by ($sg = |g|$, $s = \text{sign}(g) = \text{sign}(g')$ as in Section 5 (k))

$$\begin{aligned}
\mathcal{E}_{abc\dots} &\equiv |g|^{-1/2} \varepsilon_{abc\dots} \Rightarrow & \mathcal{E}_{123\dots} &= |g|^{-1/2} |g| = |g|^{1/2} \\
\mathcal{E}'_{abc\dots} &\equiv |g'|^{-1/2} \varepsilon'_{abc\dots} \Rightarrow & \mathcal{E}'_{123\dots} &= |g'|^{-1/2} |g'| = |g'|^{1/2}
\end{aligned}$$

According to the Weight Changing Theorem of section (b) 7 with $w = -1$ and $k = 1$ the object $\mathcal{E}_{abc\dots}$ is a covariant tensor density with weight $W - w = (-1) - (-1) = 0$. Thus, $\mathcal{E}_{abc\dots}$ is an ordinary covariant tensor except the indices cannot be raised by g . In this convention then one must make independent definitions of the contravariant components as follows,

$$\begin{aligned}
\mathcal{E}^{abc\dots} &\equiv |g|^{-1/2} \varepsilon^{abc\dots} \Rightarrow & \mathcal{E}^{123\dots} &= |g|^{-1/2} \\
\mathcal{E}'^{abc\dots} &\equiv |g'|^{-1/2} \varepsilon'^{abc\dots} \Rightarrow & \mathcal{E}'^{123\dots} &= |g'|^{-1/2}
\end{aligned}$$

To compare $\mathcal{E}_{abc\dots}$ and $\mathcal{E}^{abc\dots}$,

$$\begin{aligned}
\mathcal{E}_{abc\dots} &= |g|^{-1/2} \varepsilon_{abc\dots} = |g|^{-1/2} |g| \varepsilon^{abc\dots} = |g|^{-1/2} |g| \varepsilon^{abc\dots} \\
\mathcal{E}^{abc\dots} &= |g|^{-1/2} \varepsilon^{abc\dots}
\end{aligned}$$

so that

$$\mathcal{E}_{abc\dots} = |g| \mathcal{E}^{abc\dots}$$

Summarizing,

$$\begin{aligned}
\mathcal{E}^{123\dots} &= |g|^{-1/2} & \mathcal{E}_{123\dots} &= |g|^{+1/2} & \mathcal{E}_{abc\dots} &= |g| \mathcal{E}^{abc\dots} \\
\mathcal{E}'^{123\dots} &= |g'|^{-1/2} & \mathcal{E}'_{123\dots} &= |g'|^{+1/2} & \mathcal{E}'_{abc\dots} &= |g'| \mathcal{E}'^{abc\dots} \\
\mathcal{E}^{abc\dots} \mathcal{E}_{ABC\dots} &= |g|^{-1} \varepsilon^{abc\dots} \varepsilon_{ABC\dots} \\
\mathcal{E}'^{abc\dots} \mathcal{E}'_{ABC\dots} &= |g'|^{-1} \varepsilon'^{abc\dots} \varepsilon'_{ABC\dots}
\end{aligned}$$

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 (k) below.

Equations in the "Ricci-Levi-Civita convention"

Ricci and Levi-Civita use the "added s convention" but add a factor $\sigma = \text{sign}(\det(S))$ into their definition of $\mathcal{E}$ (see their paper p 135 or Hermann pp 31-21) so that

$$\begin{aligned}
\mathcal{E}_{abc\dots} &\equiv \sigma|g|^{-1/2} \varepsilon_{abc\dots} &\Rightarrow & \mathcal{E}_{123\dots} = \sigma|g|^{-1/2} |g| = \sigma|g|^{1/2} \\
\mathcal{E}'_{abc\dots} &\equiv \sigma|g'|^{-1/2} \varepsilon'_{abc\dots} &\Rightarrow & \mathcal{E}'_{123\dots} = \sigma|g'|^{-1/2} |g'| = \sigma|g'|^{1/2} \\
\mathcal{E}^{abc\dots} &\equiv \sigma|g|^{-1/2} \varepsilon^{abc\dots} &\Rightarrow & \mathcal{E}^{123\dots} = \sigma|g|^{-1/2} \\
\mathcal{E}'^{abc\dots} &\equiv \sigma|g'|^{-1/2} \varepsilon'^{abc\dots} &\Rightarrow & \mathcal{E}'^{123\dots} = \sigma|g'|^{-1/2}
\end{aligned}$$

Summarizing,

$$\begin{aligned}
\mathcal{E}^{123\dots} &= \sigma|g|^{-1/2} & \mathcal{E}_{123\dots} &= \sigma|g|^{+1/2} & \mathcal{E}_{abc\dots} &= |g| \mathcal{E}^{abc\dots} \\
\mathcal{E}'_{123\dots} &= \sigma|g'|^{-1/2} & \mathcal{E}'_{123\dots} &= \sigma|g'|^{+1/2} & \mathcal{E}'_{abc\dots} &= |g'| \mathcal{E}'^{abc\dots} \\
\mathcal{E}^{abc\dots} \mathcal{E}_{ABC\dots} &= |g|^{-1} \varepsilon^{abc\dots} \varepsilon_{ABC\dots} \\
\mathcal{E}'^{abc\dots} \mathcal{E}'_{ABC\dots} &= |g'|^{-1} \varepsilon'^{abc\dots} \varepsilon'_{ABC\dots}
\end{aligned}$$

Notice that in all three conventions, 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.

(j) Representation of ε , $\varepsilon\varepsilon$ and contracted $\varepsilon\varepsilon$ as determinants

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} ,

$$Q_{abc\dots x} \equiv \sum_P p P_2(M_{a1}M_{b2} M_{c3}\dots M_{xN})$$

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 $2 \leftrightarrow 3$ on the right above, each permutation would have $S \rightarrow S+1$ since an extra swap is needed to undo $2 \leftrightarrow 3$. Thus, all parities $p \rightarrow -p$ and in fact the whole object negates. But the swap $2 \leftrightarrow 3$ is the same as $b \leftrightarrow c$ since $M_{b3} M_{c2} = M_{c2} M_{b3}$. Applying this argument to any pair of indices, one concludes that $Q_{abc\dots x}$ is totally antisymmetric and therefore can be written as $K \varepsilon_{abc\dots x}$:

$$\sum_P p P_2(M_{a1}M_{b2} M_{c3}\dots M_{xN}) = K \varepsilon_{abc\dots x}$$

Setting $abc\dots x$ to 123...N, one gets.

$$\sum_P p P_2(M_{11}M_{22} M_{33}\dots M_{NN}) = K$$

The left side of this last equation can be written as

$$\sum_P p P_2(M_{11}M_{22} M_{33}.....M_{NN}) = \sum_{abc...x} \epsilon_{abc...x} M_{1a}M_{2b} M_{3c}.....M_{Nx}$$

because $p = \epsilon_{abc...x}$ 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

$$\sum_P p P_2(M_{a1}M_{b2} M_{c3}.....M_{xN}) = \det(M) \epsilon_{abc...x}$$

Consider now the following matrix where $abc...x$ is some permutation of $123...x$,

$$\text{where } M^{(abc...)} = \begin{matrix} M_{a1} & M_{a2} & M_{a3} & \dots & M_{aN} \\ M_{b1} & M_{b2} & M_{b3} & \dots & M_{bN} \\ M_{c1} & M_{c2} & M_{c3} & \dots & M_{cN} \\ \dots & \dots & \dots & \dots & \dots \\ M_{x1} & M_{x2} & M_{x3} & \dots & M_{xN} \end{matrix}$$

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 $\epsilon_{abc...x}$. Therefore

$$\det(M^{(abc...)}) = \epsilon_{abc...x} \det(M)$$

and therefore

$$\sum_P p P_2(M_{a1}M_{b2} M_{c3}.....M_{xN}) = \det(M^{(abc...)}) = \det(M) \epsilon_{abc...x}$$

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 that matrix.

2. Application of the theorem to $M = \delta$: a representation of ϵ

Apply the above theorem to matrix $M = 1 \equiv \delta$, the identity matrix, having $M_{ij} = \delta_{i,j}$. Clearly $\det(\delta) = 1$ and one then has

$$\sum_P p P_2(\delta_{a,1}\delta_{b,2} \delta_{c,3}.....\delta_{x,N}) = \det[\delta^{(abc...)}] = \epsilon_{abc...x}$$

Thus is obtained the famous representation of $\epsilon_{abc...x}$ as a certain determinant of Kronecker deltas,

$$\epsilon_{abc...x} = \det[\delta^{(abc...)}]$$

$$\text{where } \delta^{(abc...)} = \begin{matrix} \delta_{a,1} & \delta_{a,2} & \delta_{a,3} & \dots & \delta_{a,N} & = \mathbf{R}_a \\ \delta_{b,1} & \delta_{b,2} & \delta_{b,3} & \dots & \delta_{b,N} & = \mathbf{R}_b \\ \delta_{c,1} & \delta_{c,2} & \delta_{c,3} & \dots & \delta_{c,N} & = \mathbf{R}_c \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \delta_{x,1} & \delta_{x,2} & \delta_{x,3} & \dots & \delta_{x,N} & = \mathbf{R}_x \end{matrix}$$

where, 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

$$\varepsilon_{abc\dots x} = \begin{vmatrix} \delta_{a,1} & \delta_{a,2} & \delta_{a,3} & \dots & \delta_{a,N} \\ \delta_{b,1} & \delta_{b,2} & \delta_{b,3} & \dots & \delta_{b,N} \\ \delta_{c,1} & \delta_{c,2} & \delta_{c,3} & \dots & \delta_{c,N} \\ \dots & \dots & \dots & \dots & \dots \\ \delta_{x,1} & \delta_{x,2} & \delta_{x,3} & \dots & \delta_{x,N} \end{vmatrix} = |\delta^{(abc\dots)}|$$

which is the same as

$$\varepsilon_{abc\dots x} = \sum_P P_2(\delta_{a,1}\delta_{b,2}\delta_{c,3}\dots\delta_{x,N}) .$$

3. Outer product of two ε tensors.

Consider now

$$\varepsilon_{abc\dots x} = \det \begin{pmatrix} \mathbf{R}_a \\ \mathbf{R}_b \\ \dots \\ \mathbf{R}_x \end{pmatrix} \quad \text{and} \quad \varepsilon_{a'b'c'\dots x'} = \det \begin{pmatrix} \mathbf{R}_{a'} \\ \mathbf{R}_{b'} \\ \dots \\ \mathbf{R}_{x'} \end{pmatrix}$$

Then

$$\begin{aligned} \varepsilon_{abc\dots x} \varepsilon_{a'b'c'\dots x'} &= \det \begin{pmatrix} \mathbf{R}_a \\ \mathbf{R}_b \\ \dots \\ \mathbf{R}_x \end{pmatrix} \det \begin{pmatrix} \mathbf{R}_{a'} \\ \mathbf{R}_{b'} \\ \dots \\ \mathbf{R}_{x'} \end{pmatrix} = \det \begin{pmatrix} \mathbf{R}_a \\ \mathbf{R}_b \\ \dots \\ \mathbf{R}_x \end{pmatrix} \det (\mathbf{R}_a, \mathbf{R}_b, \dots \mathbf{R}_{x'}) \\ &= \det \left\{ \begin{pmatrix} \mathbf{R}_a \\ \mathbf{R}_b \\ \dots \\ \mathbf{R}_x \end{pmatrix} (\mathbf{R}_a, \mathbf{R}_b, \dots \mathbf{R}_{x'}) \right\} \end{aligned}$$

which is the determinant of this matrix

$$\begin{array}{cccc} \mathbf{R}_a \bullet \mathbf{R}_{a'} & \mathbf{R}_a \bullet \mathbf{R}_{b'} & \mathbf{R}_a \bullet \mathbf{R}_{c'} & \dots \mathbf{R}_a \bullet \mathbf{R}_{x'} \\ \mathbf{R}_b \bullet \mathbf{R}_{a'} & \mathbf{R}_b \bullet \mathbf{R}_{b'} & \mathbf{R}_b \bullet \mathbf{R}_{c'} & \dots \mathbf{R}_b \bullet \mathbf{R}_{x'} \\ \mathbf{R}_c \bullet \mathbf{R}_{a'} & \mathbf{R}_c \bullet \mathbf{R}_{b'} & \mathbf{R}_c \bullet \mathbf{R}_{c'} & \dots \mathbf{R}_c \bullet \mathbf{R}_{x'} \\ \dots & \dots & \dots & \dots \\ \mathbf{R}_x \bullet \mathbf{R}_{a'} & \mathbf{R}_x \bullet \mathbf{R}_{b'} & \mathbf{R}_x \bullet \mathbf{R}_{c'} & \dots \mathbf{R}_x \bullet \mathbf{R}_{x'} \end{array}$$

A typical element of this matrix is given by

$$\mathbf{R}_c \bullet \mathbf{R}_{b'} = (\mathbf{R}_c)_i (\mathbf{R}_{b'})_i = \delta_{c,i} \delta_{b',i} = \delta_{c,b'}$$

so that matrix can be written as

$$\begin{array}{cccc} \delta_{a,a'} & \delta_{a,b'} & \delta_{a,c'} & \dots \delta_{a,x'} \\ \delta_{b,a'} & \delta_{b,b'} & \delta_{b,c'} & \dots \delta_{b,x'} \\ \delta_{c,a'} & \delta_{c,b'} & \delta_{c,c'} & \dots \delta_{c,x'} \\ \dots & & & \\ \delta_{x,a'} & \delta_{x,b'} & \delta_{x,c'} & \dots \delta_{x,x'} \end{array} \equiv \delta^{(abc\dots x; a'b'c'\dots x')}$$

where we have made up a name for this matrix as shown.

The conclusion then is that

$$\varepsilon_{abc\dots x} \varepsilon_{a'b'c'\dots x'} = \begin{vmatrix} \delta_{a,a'} & \delta_{a,b'} & \delta_{a,c'} & \dots \delta_{a,x'} \\ \delta_{b,a'} & \delta_{b,b'} & \delta_{b,c'} & \dots \delta_{b,x'} \\ \delta_{c,a'} & \delta_{c,b'} & \delta_{c,c'} & \dots \delta_{c,x'} \\ \dots & & & \\ \delta_{x,a'} & \delta_{x,b'} & \delta_{x,c'} & \dots \delta_{x,x'} \end{vmatrix} \equiv \left| \delta^{(abc\dots x; a'b'c'\dots x')} \right|$$

which is the same as

$$\varepsilon_{abc\dots x} \varepsilon_{a'b'c'\dots x'} = \sum_P p P_2(\delta_{a,a'} \delta_{b,b'} \delta_{c,c'} \dots \delta_{x,x'})$$

As usual, the argument of P_2 is the product of the diagonal elements of the matrix.

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 $\varepsilon\varepsilon$ product:

$$\sum_a \varepsilon_{abc\dots x} \varepsilon_{ab'c'\dots x'}$$

For fixed given values of $bc\dots x$ and $b'c'\dots x'$, there is only one way this sum can be non-zero. In that one way, $bc\dots x$ and $b'c'\dots x'$ must each be permutations of the set $\{1, 2, \dots, N \text{ exclude } A\}$ where A is the "hit value" of a in the sum on a . Then

$$\sum_a \varepsilon_{abc\dots x} \varepsilon_{ab'c'\dots x'} = \varepsilon_{Abc\dots x} \varepsilon_{Ab'c'\dots x'}$$

$$= \sum_P p P_2(\delta_{A,A} \delta_{b,b'} \delta_{c,c'} \dots \delta_{x,x'}) = \sum_P p P_2(\delta_{b,b'} \delta_{c,c'} \dots \delta_{x,x'}) \quad (*)$$

where in this last expression the sum can be regarded as being over permutations where $b'c'\dots x'$ is a permutation of $b,c\dots x$. Each of these lists of integers is in turn a permutation of $\{1, 2, \dots, N \text{ exclude } A\}$. Now, parity $p = (-1)^S$ where S is a number of swaps it takes to connect $b'c'\dots x'$ with $b,c\dots x$, 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 $\delta_{b,b'}\delta_{c,c'}\dots\delta_{x,x'}$ 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,

$$\sum_a \epsilon_{abc} \epsilon_{ab'c'} = \sum_P p P_2(\delta_{b,b'}\delta_{c,c'}) = \delta_{b,b'}\delta_{c,c'} - \delta_{b,c'}\delta_{c,b'}$$

The permutation sum shown on the right side of (*) is the determinant of $\delta^{(abc\dots x; a'b'c'\dots x')}$ 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,

$$\sum_a \epsilon_{abc\dots x} \epsilon_{ab'c'\dots x'} = [\text{cof } \delta^{(abc\dots x; a'b'c'\dots x')}]_{aa}$$

where the notation cofM refers to a matrix of cofactors with elements $[\text{cofM}]_{ij}$. 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)

$$\epsilon_{abc\dots x} \epsilon_{ab'c'\dots x'} = \begin{vmatrix} \delta_{b,b'} & \delta_{b,c'} & \dots & \delta_{b,x'} \\ \delta_{c,b'} & \delta_{c,c'} & \dots & \delta_{c,x'} \\ \delta_{x,b'} & \delta_{x,c'} & \dots & \delta_{x,x'} \end{vmatrix} = [\text{cof } \delta^{(abc\dots x; a'b'c'\dots x')}]_{aa}$$

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 $\epsilon\epsilon$ product:

$$\sum_{a,b} \epsilon_{abcd\dots x} \epsilon_{abc'd'\dots x'}$$

For fixed given values of $c,d\dots x$ and $c'd'\dots x'$, in order for this double sum to be non-zero, the index sets $cd\dots x$ and $c'd'\dots x'$ must each be permutations of the set $\{12\dots N \text{ exclude } A,B\}$ where A,B are a pair of hit values for the a and b sums. If $a=A$ and $b=B$ is 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

$$\begin{aligned} \sum_{a,b} \epsilon_{abcd\dots x} \epsilon_{abc'd'\dots x'} &= 2! \epsilon_{ABc\dots x} \epsilon_{ABc'\dots x'} \\ &= 2! \sum_P p P_2(\delta_{A,A}\delta_{B,B'}\delta_{c,c'}\delta_{d,d'}\dots\delta_{x,x'}) = 2! \sum_P p P_2(\delta_{c,c'}\delta_{d,d'}\dots\delta_{x,x'}) \end{aligned} \quad (*)$$

where in this last expression the sum is over permutations where $c'd'\dots x'$ is a permutation of $c,d\dots x$. Each of these lists of integers is in turn a permutation of $\{12\dots N \text{ exclude } A,B\}$. Now parity $p = (-1)^S$ where S is a number of swaps it takes to connect $c'd'\dots x'$ with $c,d\dots x$. 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,

$$\sum_{a,b} \epsilon_{abc} \epsilon_{abc'} = 2! \sum_P p P_2(\delta_{c,c'}) = 2 \delta_{c,c'}$$

If $c = c' = 2$, then this says

$$\sum_{a,b} \varepsilon_{ab2} \varepsilon_{ab2} = \varepsilon_{132} \varepsilon_{132} + \varepsilon_{312} \varepsilon_{312} = 1 + 1 = 2$$

The permutation sum shown on the right side of (*) is the determinant of $\delta^{(abc \dots x; a'b'c' \dots x')}$ but with the first 2 rows and columns crossed out. Therefore,

$$\sum_{a,b} \varepsilon_{abcd \dots x} \varepsilon_{abc'd' \dots x'} = 2! \{ [\text{cof } \delta^{(abc \dots x; a'b'c' \dots x')}]_{aa} \}_{bb}$$

The conclusion then is that (implied summation on a,b on the LHS)

$$\varepsilon_{abcd \dots x} \varepsilon_{abc'd' \dots x'} = 2! \begin{vmatrix} \delta_{c,c'} & \dots & \delta_{c,x'} \\ \dots & & \dots \\ \delta_{x,c'} & \dots & \delta_{x,x'} \end{vmatrix} = 2! \{ [\text{cof } \delta^{(abc \dots x; a'b'c' \dots x')}]_{aa} \}_{bb}$$

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

$$\sum_{a,b,c} \varepsilon_{abcd \dots x} \varepsilon_{abcd' \dots x'} = 3! \{ \{ [\text{cof } \delta^{(abc \dots x; a'b'c' \dots x')}]_{aa} \}_{bb} \}_{cc}$$

The conclusion then is that (implied summation on a,b,c on the LHS)

$$\varepsilon_{abcd \dots x} \varepsilon_{abcd' \dots x'} = 3! \begin{vmatrix} \delta_{d,d'} & \dots & \delta_{d,x'} \\ \dots & & \dots \\ \delta_{x,d'} & \dots & \delta_{x,x'} \end{vmatrix} = 3! \{ \{ [\text{cof } \delta^{(abc \dots x; a'b'c' \dots x')}]_{aa} \}_{bb} \}_{cc}$$

Eventually one arrives at a point where all but one of the indices are summed, so that

$$\varepsilon_{abcd \dots x} \varepsilon_{abcd \dots x'} = (N-1)! |\delta_{x,x'}| = (N-1)! \delta_{x,x'}$$

an example being

$$\varepsilon_{abc2} \varepsilon_{abc2} = 3! \delta_{22} = 3! = \varepsilon_{1342} \varepsilon_{1342} + \varepsilon_{1432} \varepsilon_{1432} + 4 \text{ more terms} = 1+1+4 = 6$$

The final point is that at which *all* indices are summed, with result

$$\varepsilon_{abcd \dots x} \varepsilon_{abcd \dots x} = N!$$

and example of which is

$$\varepsilon_{abc} \varepsilon_{abc} = \varepsilon_{123}^2 + \varepsilon_{213}^2 + 4 \text{ more terms} = 1 + 1 + 4 = 6$$

6. Summary of Results

$$\begin{aligned} \varepsilon_{abc\dots x} \varepsilon_{a'b'c'\dots x'} &= \begin{vmatrix} \delta_{a,a'} & \delta_{a,b'} & \delta_{a,c'} & \dots & \delta_{a,x'} \\ \delta_{b,a'} & \delta_{b,b'} & \delta_{b,c'} & \dots & \delta_{b,x'} \\ \delta_{c,a'} & \delta_{c,b'} & \delta_{c,c'} & \dots & \delta_{c,x'} \\ \dots & \dots & \dots & \dots & \dots \\ \delta_{x,a'} & \delta_{x,b'} & \delta_{x,c'} & \dots & \delta_{x,x'} \end{vmatrix} \equiv \left| \delta^{(abc\dots x; a'b'c'\dots x')} \right| \\ \varepsilon_{abc\dots x} \varepsilon_{ab'c'\dots x'} &= \begin{vmatrix} \delta_{b,b'} & \delta_{b,c'} & \dots & \delta_{b,x'} \\ \delta_{c,b'} & \delta_{c,c'} & \dots & \delta_{c,x'} \\ \dots & \dots & \dots & \dots \\ \delta_{x,b'} & \delta_{x,c'} & \dots & \delta_{x,x'} \end{vmatrix} = [\text{cof } \delta^{(abc\dots x; a'b'c'\dots x')}]_{aa} \\ \varepsilon_{abcd\dots x} \varepsilon_{abc'd'\dots x'} &= 2! \begin{vmatrix} \delta_{c,c'} & \dots & \delta_{c,x'} \\ \dots & \dots & \dots \\ \delta_{x,c'} & \dots & \delta_{x,x'} \end{vmatrix} = 2! \{ [\text{cof } \delta^{(abc\dots x; a'b'c'\dots x')}]_{aa} \}_{bb} \\ \varepsilon_{abcd\dots x} \varepsilon_{abcd'\dots x'} &= 3! \begin{vmatrix} \delta_{d,d'} & \dots & \delta_{d,x'} \\ \dots & \dots & \dots \\ \delta_{x,d'} & \dots & \delta_{x,x'} \end{vmatrix} = 3! \{ \{ [\text{cof } \delta^{(abc\dots x; a'b'c'\dots x')}]_{aa} \}_{bb} \}_{cc} \\ &\dots \\ \varepsilon_{abcd\dots x} \varepsilon_{abcd\dots x'} &= (N-1)! \delta_{x,x'} \\ \varepsilon_{abcd\dots x} \varepsilon_{abcd\dots x} &= N! \end{aligned}$$

(k) Covariant forms of the previous section results

The results above were all developed in Cartesian x-space where up and down indices on the ε 's did not matter. The rules for converting any result above to covariant form are as follows:

Weinberg convention:

- write the left side as either $|g|^{-1} \varepsilon^{*****} \varepsilon_{*****}$ or as $\mathcal{E}^{*****} \mathcal{E}_{*****}$. The objects with indices as shown by asterisks are true tensors (weight 0).
- write the right side replacing every $\delta_{a,b}$ by $\delta^a_b \rightarrow g^a_b$ as shown in Section 7 (m). Then the right side will also be a true tensor.

Example: The $\varepsilon\varepsilon$ product for $N=2$ with no indices summed was written above as ($g=1$)

$$\varepsilon_{ab} \varepsilon_{a'b'} = \begin{vmatrix} \delta_{aa'} & \delta_{ab'} \\ \delta_{ba'} & \delta_{bb'} \end{vmatrix} = \delta_{a,a'} \delta_{b,b'} - \delta_{a,b'} \delta_{b,a'}$$

The covariant form is as follows, where now g is some arbitrary metric tensor for x-space,

$$\mathcal{E}^{ab}\mathcal{E}_{a'b'} = |g|^{-1} \varepsilon^{ab}\varepsilon_{a'b'} = \begin{vmatrix} g^a_{a'} & g^a_{b'} \\ g^b_{a'} & g^b_{b'} \end{vmatrix} = g^a_{a'} g^b_{b'} - g^a_{b'} g^b_{a'}$$

The equation in x'-space would then be

$$\mathcal{E}'^{ab}\mathcal{E}'_{a'b'} = |g'|^{-1} \varepsilon'^{ab}\varepsilon'_{a'b'} = \begin{vmatrix} g'^a_b & g'^a_{b'} \\ g'^a_{b'} & g'^a_{b'} \end{vmatrix} = g'^a_b g'^a_{b'} - g'^a_{b'} g'^a_b$$

because true tensor equations are "covariant"(Section 7 (u)). 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:

$$\mathcal{E}^{ab}\mathcal{E}_{a'b'} = |g|^{-1} \varepsilon^{ab}\varepsilon_{a'b'} = \begin{vmatrix} g^a_{a'} & g^a_{b'} \\ g^b_{a'} & g^b_{b'} \end{vmatrix} = g^a_{a'} g^b_{b'} - g^a_{b'} g^b_{a'}$$

$$\mathcal{E}^a_b \mathcal{E}_{a'b'} = |g|^{-1} \varepsilon^a_b \varepsilon_{a'b'} = \begin{vmatrix} g^a_{a'} & g^a_{b'} \\ g_{ba'} & g_{bb'} \end{vmatrix} = g^a_{a'} g_{bb'} - g^a_{b'} g_{ba'}$$

$$\mathcal{E}_{ab} \mathcal{E}_{a'b'} = |g|^{-1} \varepsilon_{ab} \varepsilon_{a'b'} = \begin{vmatrix} g_{aa'} & g_{ab'} \\ g_{ba'} & g_{bb'} \end{vmatrix} = g_{aa'} g_{bb'} - g_{ab'} g_{ba'}$$

and of course in x'-space the equations are the same but everything is primed.

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 $\varepsilon_{abc\dots} = s g \varepsilon^{abc\dots}$ in these conventions instead of $\varepsilon_{abc\dots} = g \varepsilon^{abc\dots}$ so that $\varepsilon_{abc\dots}(\text{Weinberg}) = s\varepsilon_{abc\dots}(\text{added-sign})$.

Example: The first equation above becomes ($\varepsilon_{a'b'} \rightarrow s \varepsilon_{a'b'}$)

$$\mathcal{E}^{ab}\mathcal{E}_{a'b'} = |g|^{-1} \varepsilon^{ab}\varepsilon_{a'b'} = s \begin{vmatrix} g^a_{a'} & g^a_{b'} \\ g^b_{a'} & g^b_{b'} \end{vmatrix} = s (g^a_{a'} g^b_{b'} - g^a_{b'} g^b_{a'})$$

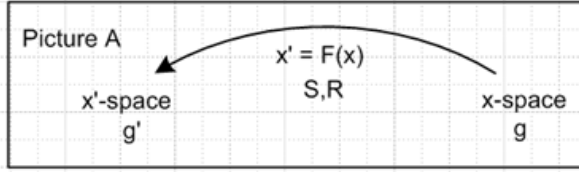
The second equation is *undefined* (when $s=-1$), and the third equation is

$$\mathcal{E}_{ab} \mathcal{E}_{a'b'} = |g|^{-1} \varepsilon_{ab} \varepsilon_{a'b'} = \begin{vmatrix} g_{aa'} & g_{ab'} \\ g_{ba'} & g_{bb'} \end{vmatrix} = g_{aa'} g_{bb'} - g_{ab'} g_{ba'}$$

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.

Appendix E: Tensor Expansions: direct product, polyadic and operator notation

This entire section uses the general Picture A context where x-space need not be Cartesian,



The Standard Notation is used throughout.

(a) Direct Product Notation

The key tool required for the expression of tensor expansions is the notion of a direct product of n tensorial vectors defined in this simple way,

$$\begin{aligned} (A \otimes B \otimes C \dots)^{abc\dots} &\equiv A^a B^b C^c \dots \\ (A \otimes B \otimes C \dots)_b^{a \dots c\dots} &\equiv A^a B_b C^c \dots \quad \text{etc} \end{aligned}$$

The tensor $A \otimes B \otimes C \dots$ is nothing more than the outer product of vectors A,B,C as discussed in Section 7 (a) for contravariant vectors, but later extended to any mixture of vector types. As noted in Section 7 (j), one can define a direct product of two rank-2 tensors in this way,

$$\begin{aligned} (M \otimes N)_{ab, AB} &\equiv M_{aA} N_{bB} & // \text{rank} = n = 2; \text{ number of tensors} = I = 2 \\ (M \otimes N)^{ab, AB} &\equiv M^a_A N^b_B \quad \text{etc} \end{aligned}$$

and then the same idea can be applied to form a direct product of tensors of any rank, for example

$$(M \otimes N)_{ab, AB, \alpha\beta} = M_{aA\alpha} N_{bB\beta} \quad \text{etc} \quad // \text{rank} = n = 3; \text{ number of tensors} = I = 2$$

On the left side the number of groups of indices equals the tensor rank n, and the number of indices within each group matches the number I of tensors being direct-product-multiplied.

In what follows, only the direct product of vectors shall be considered. One can define the dot product of two direct-product-space vectors in this obvious manner,

$$\begin{aligned} (A \otimes B \otimes C \dots) \bullet (A' \otimes B' \otimes C' \dots) &\equiv (A \otimes B \otimes C \dots)^{abc\dots} (A' \otimes B' \otimes C' \dots)_{abc} \\ &= A^a B^b C^c \dots A'_a B'_b C'_c \dots = A \bullet A' B \bullet B' C \bullet C' \dots \end{aligned}$$

where of course the indices abc can be tilted in any way desired according to Section 7 (k).

(b) Tensor Expansions and Bases

Let $\mathbf{b}_i$ be an arbitrary complete set of basis vectors in x-space. As shown in Section 6 (b) there exists a unique set of *dual* basis vectors $\mathbf{b}^i$ (also in x-space) such that $\mathbf{b}^i \bullet \mathbf{b}_j = \delta^i_j$. Consider then the following expansion of a rank-3 tensor A

$$A = \sum_{ijk} \alpha^{ijk} (\mathbf{b}_i \otimes \mathbf{b}_j \otimes \mathbf{b}_k) \quad \text{where} \quad (\mathbf{b}_i \otimes \mathbf{b}_j \otimes \mathbf{b}_k \dots)^{abc} = (b_i)^a (b_j)^b (b_k)^c$$

The coefficients α^{ijk} can be obtained by dotting both sides with $(\mathbf{b}^{i'} \otimes \mathbf{b}^{j'} \otimes \mathbf{b}^{k'})$ and using

$$(\mathbf{b}^{i'} \otimes \mathbf{b}^{j'} \otimes \mathbf{b}^{k'}) \bullet (\mathbf{b}_i \otimes \mathbf{b}_j \otimes \mathbf{b}_k) = \mathbf{b}^{i'} \bullet \mathbf{b}_i \mathbf{b}^{j'} \bullet \mathbf{b}_j \mathbf{b}^{k'} \bullet \mathbf{b}_k = \delta^{i'}_i \delta^{j'}_j \delta^{k'}_k .$$

The result is then

$$\alpha^{ijk} = A \bullet (\mathbf{b}^i \otimes \mathbf{b}^j \otimes \mathbf{b}^k) = A^{abc} (\mathbf{b}^i \otimes \mathbf{b}^j \otimes \mathbf{b}^k)_{abc} = A^{abc} (b^i)_a (b^j)_b (b^k)_c . \quad (*)$$

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.

In this manner, a tensor A of any rank can be expanded on an arbitrary complete set of basis vectors, and the coefficients of that expansion can be obtained by the inversion shown above for rank 3.

Two special bases are of interest.

The $\mathbf{u}_i$ are the axis-aligned basis vectors in x-space as discussed in see Section 7 (s). For these basis vectors, one has $(\mathbf{u}_i)^a = \delta_i^a$ and $(\mathbf{u}^i)_a = \delta^i_a$. If one considers this expansion,

$$A = \sum_{ijk} \alpha^{ijk} (\mathbf{u}_i \otimes \mathbf{u}_j \otimes \mathbf{u}_k)$$

$$\text{where} \quad (\mathbf{u}_i \otimes \mathbf{u}_j \otimes \mathbf{u}_k \dots)^{abc} = (u_i)^a (u_j)^b (u_k)^c = \delta_i^a \delta_j^b \delta_k^c$$

then the coefficients are found to be

$$\alpha^{ijk} = A \bullet (\mathbf{u}^i \otimes \mathbf{u}^j \otimes \mathbf{u}^k) = A^{abc} \delta^i_a \delta^j_b \delta^k_c = A^{ijk}$$

so the coefficients are exactly the x-space contravariant components of the tensor A. Thus

$$A = \sum_{ijk} A^{ijk} (\mathbf{u}_i \otimes \mathbf{u}_j \otimes \mathbf{u}_k)$$

On the other hand, if $\mathbf{e}_i$ are the tangent base vectors in x-space (see Sections 3), the dual vectors are the $\mathbf{e}^i$ and from Section 7 (s) one has $(\mathbf{e}_i)^a = S^a_i = R_i^a$ and $(\mathbf{e}^i)_a = S_a^i = R^i_a$. If one considers the expansion

$$A = \sum_{ijk} \alpha^{ijk} (\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k)$$

$$\text{where} \quad (\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \dots)^{abc} = (e_i)^a (e_j)^b (e_k)^c = R_i^a R_j^b R_k^c$$

then the coefficients are found to be

$$\begin{aligned}\alpha^{ijk} &= A \bullet (e^i \otimes e^j \otimes e^k) = A^{abc} (e^i)_a (e^j)_b (e^k)_c = A^{abc} R^i_a R^j_b R^k_c \\ &= R^i_a R^j_b R^k_c A^{abc} = A'^{ijk}\end{aligned}$$

and thus the coefficients in this case are exactly the x'-space contravariant components of tensor A, as shown in Section 7 (j). Thus,

$$A = \sum_{ijk} A'^{ijk} (e_i \otimes e_j \otimes e_k)$$

Expansions like the above are the generalizations to tensors of any rank of these vector expansions stated in Section 7 (s),

$$A = \sum_i \alpha^i \mathbf{b}_i \quad \alpha^i = \mathbf{b}^i \bullet A \quad // \text{arbitrary basis}$$

$$A = \sum_i A^i \mathbf{u}_i \quad // \text{axis aligned unit vectors}$$

$$A = \sum_i A'^i \mathbf{e}_i \quad // \text{tangent base vectors}$$

where we continue to write rank-1 tensors 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:

$$A = \sum_{ijk\dots} \alpha^{ijk\dots} (\mathbf{b}_i \otimes \mathbf{b}_j \otimes \mathbf{b}_k \dots) \quad \alpha^{ijk\dots} = A^{abc\dots} (b^i)_a (b^j)_b (b^k)_c \dots$$

$$A = \sum_{ijk\dots} A^{ijk\dots} (\mathbf{u}_i \otimes \mathbf{u}_j \otimes \mathbf{u}_k \dots) \quad A^{ijk\dots} = \text{contravariant components of } A \text{ in } x\text{-space}$$

$$A = \sum_{ijk\dots} A'^{ijk\dots} (\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \dots) \quad A'^{ijk\dots} = \text{contravariant components of } A \text{ in } x'\text{-space}$$

Orthonormal basis: If the basis vectors $\mathbf{b}_i$ happen to be orthonormal as defined by $\mathbf{b}_i \bullet \mathbf{b}_j = \delta_{ij}$ then $\mathbf{b}^i = \mathbf{b}_i$ because the dual basis is unique. As indicated in (*) above, this implies that coefficient α^{ijk} is unchanged if any or all indices are lowered, *as if* these α^{ijk} were components of a tensor in some Cartesian space. That Cartesian space is in fact the x'-space that *would* arise if transformation F were custom-selected such that the $\mathbf{b}_i$ were the tangent base vectors $\mathbf{e}_i$ for that F, for then $g'_{ij} = \mathbf{e}_i \bullet \mathbf{e}_j = \delta_{i,j}$ so *that* x'-space would in fact be Cartesian. But for a pre-determined F, the α^{ijk} are just some coefficients and are not components of a tensor relative to F, and it just happens that $\alpha^{ijk} = \alpha_i{}^{jk}$ etc.

An example of orthonormal basis vectors arises if $\mathbf{b}_i = \hat{\mathbf{e}}_i \equiv \mathbf{e}_i / h'_i$ and x'-space has a diagonal metric tensor $g'_{ab} = h'^2_{a,b} \delta_{a,b}$. One then has $\hat{\mathbf{e}}_i \bullet \hat{\mathbf{e}}_j = \delta_{i,j}$ since

$$\hat{\mathbf{e}}_i \bullet \hat{\mathbf{e}}_j = \mathbf{e}_i \bullet \mathbf{e}_j / (h'_i h'_j) = g'_{ij} / (h'_i h'_j) = h'^2_{ij} \delta_{ij} / (h'_i h'_j) = \delta_{i,j} .$$

Then since the dual basis is unique, one has $\hat{\mathbf{e}}^i = \hat{\mathbf{e}}_i$ and then

$$\alpha^{ijk}(\text{any up/down}) = A^{abc} (\hat{\mathbf{e}}^i)_a (\hat{\mathbf{e}}^j)_b (\hat{\mathbf{e}}^k)_c = A^{ijk} (h'_i h'_j h'_k)$$

where the last expression comes from the third expansion shown above.

Expansion of tensor-like objects. If A^{ijk} is some "tensor like" object having three indices (such as $\partial^i T^{jk}$) one can still do the three expansions shown above but the results would have to be restated this way:

$$A = \sum_{ijk} \dots \alpha^{ijk} \dots (\mathbf{b}_i \otimes \mathbf{b}_j \otimes \mathbf{b}_k \dots) \quad \alpha^{ijk} \dots = A^{abc} \dots (b^i)_a (b^j)_b (b^k)_c \dots$$

$$A = \sum_{ijk} \dots A^{ijk} \dots (\mathbf{u}_i \otimes \mathbf{u}_j \otimes \mathbf{u}_k \dots) \quad A^{ijk} \dots = \text{components of } A \text{ in } x\text{-space}$$

$$A = \sum_{ijk} \dots \mathcal{A}^{ijk} \dots (\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \dots) \quad \mathcal{A}^{ijk} \dots = R^i_a R^j_b R^k_c A^{abc}$$

Since A is not a tensor, in this case $\mathcal{A}^{ijk} \dots$ are not the contravariant components of tensor A in x' -space.

(c) Polyadic Notation

Some fields of study historically use "polyadic notation" as follows,

$$(\mathbf{ABC} \dots) \equiv \mathbf{A} \otimes \mathbf{B} \otimes \mathbf{C} \dots$$

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,

$$(\mathbf{b}_i \mathbf{b}_j \mathbf{b}_k \dots) \equiv \mathbf{b}_i \otimes \mathbf{b}_j \otimes \mathbf{b}_k \dots$$

Sometimes this basis vector notation is compressed even more, to wit,

$$\mathbf{i} \mathbf{j} \mathbf{k} \dots \equiv (\mathbf{b}_i \mathbf{b}_j \mathbf{b}_k \dots) \equiv \mathbf{b}_i \otimes \mathbf{b}_j \otimes \mathbf{b}_k \dots$$

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 do not "commute". For example

$$\begin{aligned} \mathbf{i} \mathbf{j} &= (\mathbf{b}_i \mathbf{b}_j) = \mathbf{b}_i \otimes \mathbf{b}_j \quad \Rightarrow \quad (\mathbf{i} \mathbf{j})^{nm} = (\mathbf{b}_i \mathbf{b}_j)^{nm} = (\mathbf{b}_i \otimes \mathbf{b}_j)^{nm} = (b_i)^n (b_j)^m \\ (\mathbf{j} \mathbf{i})^{nm} &= (\mathbf{b}_j \mathbf{b}_i)^{nm} = (\mathbf{b}_j \otimes \mathbf{b}_i)^{nm} = (b_j)^n (b_i)^m \neq (\mathbf{i} \mathbf{j})^{nm} \end{aligned}$$

and therefore one cannot write $\mathbf{i} \mathbf{j} = \mathbf{j} \mathbf{i}$.

The general expansion stated above now appears as

$$\begin{aligned} A &= \sum_{ijk} \dots \alpha^{ijk} \dots (\mathbf{b}_i \otimes \mathbf{b}_j \otimes \mathbf{b}_k \dots) \\ &= \sum_{ijk} \dots \alpha^{ijk} \dots (\mathbf{b}_i \mathbf{b}_j \mathbf{b}_k \dots) \end{aligned}$$

$$= \Sigma_{ijk...} \alpha^{ijk...} (i j k ...)$$

where

$$\begin{aligned} \alpha^{ijk...} &= A \bullet (b^i \otimes b^j \otimes b^k ...) \\ &= A \bullet (b^i b^j b^k ...) \\ &= A \bullet (i^d j^d k^d ...) \\ &= A^{abc...} (b^i)_a (b^j)_b (b^k)_c ... \end{aligned}$$

where we have just made up a notation i^d to stand for the dual vector b^i .

One can find further discussion of polyadic notation for example in Backus.

(d) 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]

$$(\mathbf{AB})^{ij} \equiv A^i B^j \quad // = (\mathbf{A} \otimes \mathbf{B})^{ij}$$

In this notation, the expansion given above for a rank-2 tensor becomes

$$A = \Sigma_{ij} \alpha^{ij} (b_i b_j) \quad \alpha^{ij} = A^{ab} (b^i)_a (b^j)_b = A^{ab} (b^i b^j)_{ab}$$

Notice 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 (g) below 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.

(e) Transpose notation for dyadics

Superscript T as usual indicates the transpose of a vector or matrix.

For a vector **V**, certainly $V^i = (V^T)^i$, meaning the object in the i^{th} column of **V** is the same as the object in the i^{th} row of $\mathbf{V}^T$. Therefore one can express the dyadic product in this more down-to-earth manner,

$$(\mathbf{ab})^{ij} \equiv a^i b^j = a^i (b^T)^j = (\mathbf{ab}^T)^{ij}$$

or

$$\mathbf{ab} = \mathbf{ab}^T$$

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,

$$\mathbf{ab}^T = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} (b_1 \ b_2) = \begin{pmatrix} a_1 b_1 & a_1 b_2 \\ a_2 b_1 & a_2 b_2 \end{pmatrix} = \text{a matrix}$$

and one sees that in fact

$$(\mathbf{ab})_{ij} = (\mathbf{ab}^T)_{ij} = a_i b_j^T = a_i b_j \ .$$

Meanwhile, the object $\mathbf{a}^T \mathbf{b}$ is just a number,

$$\mathbf{a}^T \mathbf{b} = \mathbf{a} \bullet \mathbf{b} = (a_1 \ a_2) \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = a_1 b_1 + a_2 b_2 = \text{a scalar (if } \mathbf{a} \text{ and } \mathbf{b} \text{ are vectors)}$$

This transpose notation can then be applied to the dyadic expansion of a 2x2 matrix A,

$$A = \sum_{ij} \alpha^{ij} \mathbf{b}_i \mathbf{b}_j = \sum_{ij} \alpha^{ij} \mathbf{b}_i \mathbf{b}_j^T = \alpha^{11} \mathbf{b}_1 \mathbf{b}_1^T + \alpha^{12} \mathbf{b}_1 \mathbf{b}_2^T \dots$$

In the special case that the $\mathbf{b}_i$ are the unit vectors $\mathbf{u}_i$, and assuming N = 2 dimensions, one has

$$A = \sum_{nm} A^{nm} \mathbf{u}_n \mathbf{u}_m = \sum_{nm} A^{nm} \mathbf{u}_n \mathbf{u}_m^T = A^{11} \mathbf{u}_1 \mathbf{u}_1^T + A^{12} \mathbf{u}_1 \mathbf{u}_2^T + A^{21} \mathbf{u}_2 \mathbf{u}_1^T + A^{22} \mathbf{u}_2 \mathbf{u}_2^T \\ = \text{a matrix with } A^{12} \text{ in the upper right corner}$$

where $\mathbf{u}_n$ is a column unit vector and $\mathbf{u}_n^T$ is the corresponding row unit vector (see comments in Section 3 (c) about "unit" vectors). For example,

$$\mathbf{u}_1 \mathbf{u}_2^T = \begin{pmatrix} 1 \\ 0 \end{pmatrix} (0 \ 1) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

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 $\mathbf{u}_a \mathbf{u}_b \mathbf{u}_c^T$ for example. The direct product or polyadic notation seems clearest for rank $n > 2$.

(f) Large and small dots used with dyadics

Sometimes a small-size dot $\bullet$ is used to indicate the action of a dyadic (matrix) on a vector. If A is a dyadic then one defines:

$$\begin{aligned} \mathbf{A} \bullet \mathbf{c} &\equiv \mathbf{Ac} = \text{a column vector} &\Rightarrow (\mathbf{A} \bullet \mathbf{c})^i &= (\mathbf{Ac})^i = A^{ij} c_j &\Rightarrow \mathbf{A} \bullet \mathbf{c} &= \sum_{ij} A^{ij} c_j \mathbf{u}_i \\ \mathbf{c} \bullet \mathbf{A} &\equiv \mathbf{c}^T \mathbf{A} = \text{a row vector} &\Rightarrow (\mathbf{c} \bullet \mathbf{A})^i &= (\mathbf{c}^T \mathbf{A})^i = c_j A^{ji} &\Rightarrow \mathbf{c} \bullet \mathbf{A} &= \sum_{ij} c_j A^{ji} \mathbf{u}_i \\ \mathbf{d} \bullet \mathbf{A} \bullet \mathbf{c} &= \mathbf{d}^T \mathbf{Ac} = \text{a number} &&= d_i A^{ij} c_j \end{aligned}$$

It then follows that, for the particular dyadic $A = \mathbf{ab}$,

$$(\mathbf{ab}) \bullet \mathbf{c} \equiv (\mathbf{ab}) \mathbf{c} = (\mathbf{ab}^T) \mathbf{c} = \mathbf{a}(\mathbf{b}^T \mathbf{c}) = \mathbf{a}(\mathbf{b} \bullet \mathbf{c}) = (\mathbf{b} \bullet \mathbf{c}) \mathbf{a} = \text{a column vector}$$

$$\begin{aligned}\mathbf{c} \bullet (\mathbf{ab}) &\equiv \mathbf{c}^T (\mathbf{ab}) = \mathbf{c}^T (\mathbf{ab}^T) = (\mathbf{c}^T \mathbf{a}) \mathbf{b}^T = (\mathbf{c} \bullet \mathbf{a}) \mathbf{b}^T = \text{a row vector} \\ \mathbf{d} \bullet (\mathbf{ab}) \bullet \mathbf{c} &= \mathbf{d}^T (\mathbf{ab}) \mathbf{c} = \mathbf{d}^T \mathbf{ab}^T \mathbf{c} = (\mathbf{d}^T \mathbf{a})(\mathbf{b}^T \mathbf{c}) = (\mathbf{d} \bullet \mathbf{a})(\mathbf{b} \bullet \mathbf{c}) = \text{a number}\end{aligned}$$

Here is more detail on the first line of the above group showing a skeletal matrix structure,

$$\begin{aligned}(\mathbf{ab}) \mathbf{c} &= (\mathbf{a} \mathbf{b}^T) \mathbf{c} = \mathbf{ab}^T \mathbf{c} = \mathbf{a}(\mathbf{b}^T \mathbf{c}) = \mathbf{a}(\mathbf{b} \bullet \mathbf{c}) \\ \left\{ \begin{pmatrix} x & x \\ x & x \end{pmatrix} \right\} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} &= \left\{ \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} (b_1 \ b_2) \right\} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} (b_1 \ b_2) \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \{ (b_1 \ b_2) \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \mathbf{b} \bullet \mathbf{c}\end{aligned}$$

The same small dot is used to indicate the product of two dyadics, which is to say, matrix multiplication

$$\mathbf{A} \bullet \mathbf{B} \equiv \mathbf{AB}$$

Regarding this small size dot $\bullet$: (1) from a matrix algebra point of view, it is completely superfluous except in the case $\mathbf{c} \bullet \mathbf{A} \equiv \mathbf{c}^T \mathbf{A}$; (2) it is completely different from the dot $\bullet$ used in $\mathbf{b}^T \mathbf{c} = \mathbf{b} \bullet \mathbf{c}$. It is this larger dot $\bullet$ which was the subject of Section 6 (b); (3) The next section provides an explanation of the small dot as part of an operator interpretation for dyadics.

(g) Operators and Matrices for Rank-2 tensors

Operator concept. As discussed in Section 5 (i), x-space and x'-space of Picture A are both N-dimensional real Hilbert Spaces with scalar product indicated by the large dot $\bullet$, and one can regard $\mathbf{V}$ as a vector in either space. Expressed as a "vector" in x-space one can write, as done above with generic basis $\mathbf{b}_i$,

$$\mathbf{V} = \sum_i \alpha^i \mathbf{b}_i$$

Moreover, one can regard a rank-2 tensor A as an "operator" in this Hilbert space,

$$\mathbf{A} = \sum_{ij} \alpha^{ij} \mathbf{b}_i \mathbf{b}_j^T$$

Application of $(\mathbf{b}^T)^n$ on the left and $\mathbf{b}^m$ on the right, and then a double use of $(\mathbf{b}^T)^n \mathbf{b}_i = \mathbf{b}^n \bullet \mathbf{b}_i = \delta^n_i$ gives

$$\alpha^{nm} = (\mathbf{b}^T)^n \mathbf{A} \mathbf{b}^m$$

Here, one regards A as an operator in the x Hilbert space, whereas α^{nm} is a "matrix" which is *associated* with the operator A in the particular $\mathbf{b}^n$ basis. The idea of A as operator has an abstract meaning distinct from the matrix A_{ij} . In the above equation the symbol A is this abstract operator and $(\mathbf{b}^T)^n \mathbf{A} \mathbf{b}^m$ has a meaning distinct from our interpretation of it in terms of the matrix combination of three objects. In the matrix interpretation, one writes $(\mathbf{b}^T)^n \mathbf{A} \mathbf{b}^m = [(\mathbf{b}^T)^n]_i A^{ij} [\mathbf{b}^m]_j = [\mathbf{b}^T]_i A^{ij} [\mathbf{b}^m]_j$ and only then does A become a "matrix". This matrix happens to be the contravariant A^{ij} matrix because we happened to select the $\mathbf{u}_n$ basis to write the components like $[\mathbf{b}^m]_j = \mathbf{u}_j \bullet \mathbf{b}^m$.

Bra-ket Notation. For the author of this document, the bra-ket notation commonly used in quantum mechanics (Paul Dirac 1939) provides a clean 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. In bra-ket notation one writes $\mathbf{b}_i \rightarrow |\mathbf{b}_i\rangle$, $\mathbf{b}_i^T \rightarrow \langle \mathbf{b}_i|$, so that the above equations become

$$\begin{aligned} |V\rangle &= \sum_i \alpha^i |\mathbf{b}_i\rangle & \langle \mathbf{b}_j | \mathbf{b}^i \rangle &= \delta_j^i & \text{orthogonality of the basis} \\ A &= \sum_{i,j} \alpha^{ij} |\mathbf{b}_i\rangle \langle \mathbf{b}_j| & 1 &= \sum_i |\mathbf{b}_i\rangle \langle \mathbf{b}^i| & \text{completeness of the basis} \\ \alpha^{nm} &= \langle \mathbf{b}^n | A | \mathbf{b}^m \rangle & \langle U | V \rangle &= \mathbf{U} \bullet \mathbf{V} = \text{scalar product} \end{aligned}$$

In this notation, the $N |\mathbf{b}_i\rangle$ are a set of basis vectors which span an N -dimensional real Hilbert Space, while $\langle \mathbf{b}_i|$ span the so-called adjoint (or transpose in our case) Hilbert Space. One then refers to α^{nm} as the "matrix element of the operator A in the $\mathbf{b}^i$ basis ". In this notation, based on what was presented earlier, one can write,

$$A^{nm} = \langle \mathbf{u}^n | A | \mathbf{u}^m \rangle = \text{the } x\text{-space components of tensor } A$$

$$A'^{nm} = \langle \mathbf{e}^n | A | \mathbf{e}^m \rangle = \text{the } x'\text{-space components of tensor } A$$

In all these equations the covariant indices can be moved up and down in the usual manner. Notice in the last two lines that the operator A between the vertical bars is *the exact same operator* on both lines. The matrices are different not because the operator has changed, but because the basis vectors are different.

The point here is that one really can regard a rank-2 tensor A as an operator and not a specific matrix. In a given basis, that operator "has" a certain matrix. One might write the associated matrix as

$$[A^{(b)}]^{nm} = \langle \mathbf{b}^n | A | \mathbf{b}^m \rangle.$$

That is to say, the matrix needs some label like (b) to indicate the basis used to define the matrix. In our notation, A^{nm} with no label refers to the matrix associated with the $\mathbf{u}^m$ basis, while A'^{nm} goes with the $\mathbf{e}^m$ basis.

Bases are related by a transformation. Consider again (note that $|i\rangle \equiv |\mathbf{u}_i\rangle$ on the next line)

$$[A^{(b)}]^{nm} = \langle \mathbf{b}^n | A | \mathbf{b}^m \rangle = (\mathbf{b}^n)^T A \mathbf{b}^m = [\mathbf{b}^n]_i A^{ij} [\mathbf{b}^m]_j \quad // = \langle \mathbf{b}^n | i \rangle \langle i | A | j \rangle \langle j | \mathbf{b}^m \rangle$$

where the subscripts i and j are those associated with the $\mathbf{u}_n$ basis. Defining $B_j^m \equiv [\mathbf{b}^m]_j$ one has

$$\begin{aligned} [A^{(b)}]^{nm} &= (B^T)^n_i A^{ij} B_j^m \\ \text{or} \quad [A^{(b)}]_m^n &= (B^T)^n_i A^{ij} B_j^m & // \text{ lower m and reverse the tilt of index j} \\ \text{or} \quad A^{(b)} &= B^T A B & // \text{ tilted matrix multiplication as per Section 7 (i)} \end{aligned}$$

which shows that the matrix elements $[A^{(b)}]_m^n$ are related to the A^i_j by a "congruence transformation" with a matrix B whose columns are the basis vectors $\mathbf{b}^m$. When $\mathbf{b}_m = \mathbf{u}_m$, matrix B is the identity matrix, and when $\mathbf{b}_m = \mathbf{e}_m$, one has $B_j^m \equiv [\mathbf{b}^m]_j = (\mathbf{e}^m)_j = R_j^m = S_j^m$. Thus $B = S$ in this case. It was shown in Section 7 (i) that in standard notation S is real orthogonal, so in fact one has for the $\mathbf{b}_m = \mathbf{e}_m$ basis,

$$A^{(e)} = B^T A B = S^T A S = S^{-1} A S = R A R^{-1}$$

and our two matrices are related by similarity by our old friends S and R as shown.

More on bra-ket notation and its relation to the small dyadic dot.

Consider the following facts,

$$\langle \mathbf{d} | A | \mathbf{c} \rangle = \mathbf{d}^T A \mathbf{c} = \mathbf{d}^T [A \mathbf{c}] = \langle \mathbf{d} | A \mathbf{c} \rangle$$

$$\langle \mathbf{d} | A | \mathbf{c} \rangle = \mathbf{d}^T A \mathbf{c} = [\mathbf{d}^T A] \mathbf{c} = [A^T \mathbf{d}]^T \mathbf{c} = \langle A^T \mathbf{d} | \mathbf{c} \rangle$$

where

$|A\mathbf{c}\rangle$ = a new Hilbert space vector which results when A is applied to $|\mathbf{c}\rangle$

$\langle A^T \mathbf{d} |$ = a new transpose Hilbert space vector which results when A is applied to $\langle \mathbf{d} |$

So one has this general idea that

$$\langle \mathbf{d} | A | \mathbf{c} \rangle = \langle \mathbf{d} | A \mathbf{c} \rangle = \langle A^T \mathbf{d} | \mathbf{c} \rangle$$

$$A | \mathbf{c} \rangle = |A\mathbf{c}\rangle \quad \langle \mathbf{d} | A = \langle A^T \mathbf{d} |$$

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 $|A\mathbf{c}\rangle \equiv |\mathbf{e}\rangle$ is some different vector in the Hilbert space (different from $|\mathbf{c}\rangle$), call it $|\mathbf{e}\rangle$, and the grouping (Ac) labels this vector. Similarly, $\langle A^T \mathbf{d} |$ is vector $\langle \mathbf{e} |$ in the transpose Hilbert space. The distinction between A as an abstract operator in the Hilbert space, and the A in (Ac) and $(A^T \mathbf{d}) = (\mathbf{d}^T A)^T$ 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:

$$A \bullet \mathbf{c} = A \mathbf{c} \quad \mathbf{d} \bullet A = (A^T \mathbf{d})^T = \mathbf{d}^T A \quad \mathbf{d} \bullet A \bullet \mathbf{c} = \mathbf{d}^T A \mathbf{c} \quad A \bullet B \bullet \mathbf{c}$$

$$A | \mathbf{c} \rangle = |A\mathbf{c}\rangle \quad \langle \mathbf{d} | A = \langle A^T \mathbf{d} | \quad \langle \mathbf{d} | A | \mathbf{c} \rangle = \langle \mathbf{d} | A \mathbf{c} \rangle \quad A B | \mathbf{c} \rangle$$

In the rightmost column operator B is applied first to $|\mathbf{c}\rangle$ to get vector $|B\mathbf{c}\rangle$, and then operator A is applied to $|B\mathbf{c}\rangle$ to give yet another vector $|A(B\mathbf{c})\rangle$. In bra-ket notation the product of two abstract operators is given just as AB, but in dyadic notation it is written $A \bullet B$.

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 (again, one might have $g \neq 1$)

$$\begin{aligned} T^{nm} &= (\mathbf{AB})^{nm} = \langle \mathbf{u}^n | (\mathbf{AB}) | \mathbf{u}^m \rangle = (\mathbf{u}^n)^T \mathbf{A} \mathbf{B}^T \mathbf{u}^m \\ &= [(\mathbf{u}^n)^T]_a A^a (\mathbf{B}^T)^b [\mathbf{u}^m]_b = \delta^n_a A^a B^b \delta^m_b = A^n B^m. \end{aligned}$$

In the generic $\mathbf{b}_n$ basis one might then write something like

$$[(\mathbf{AB})^{(b)}]^{nm} = \langle \mathbf{b}^n | (\mathbf{AB}) | \mathbf{b}^m \rangle$$

It is to emphasize this operator view of a dyadic that Morse and Feshbach use fancy letters like $\mathfrak{U}$ to represent dyadics. Then the small-dot notation $\mathfrak{U} \cdot \mathbf{B}$ emphasizes the idea of an operator acting on a vector, equivalent to $\mathfrak{U} | \mathbf{B} \rangle$. Here then are a few quotes from Morse and Feshbach ($\mathbf{a}_n = \mathbf{u}_n$) to illustrate some of the notation described above. These authors are working in Cartesian space ($g=1$) where up and down indices don't matter. (The first item here is our $\mathbf{A} \cdot \mathbf{c} = \sum_{ij} A^{ij} c_j \mathbf{u}_i$ from above.)

$$\mathfrak{U} \cdot \mathbf{B} = \sum_{mn} a_m A_{mn} B_n; \quad \mathbf{B} \cdot \mathfrak{U} = \sum_{mn} B_m A_{mn} a_n \quad (1.6.5)$$

$$\begin{aligned} \mathfrak{U} &= a_1 A_{11} a_1 + a_1 A_{12} a_2 + a_1 A_{13} a_3 + a_2 A_{21} a_1 + a_2 A_{22} a_2 + a_2 A_{23} a_3 \\ &\quad + a_3 A_{31} a_1 + a_3 A_{32} a_2 + a_3 A_{33} a_3 \quad (1.6.6) \end{aligned}$$

$$(\mathfrak{U}^{-1}) \cdot \mathfrak{U} = \mathfrak{U} \cdot (\mathfrak{U}^{-1}) = \mathfrak{I}$$

There is, of course, a zero dyadic $\mathfrak{O}$ and a unity dyadic $\mathfrak{I}$ called the *idemfactor*:

$$\mathfrak{O} \cdot \mathbf{F} = 0; \quad \mathfrak{I} \cdot \mathbf{F} = \mathbf{F}; \quad \mathfrak{I} = a_1 a_1 + a_2 a_2 + a_3 a_3$$

where $\mathbf{F}$ is any vector.

$$\mathfrak{U}_s = A_x \mathbf{ji} + B_z \mathbf{ij} + B_y \mathbf{ik} + B_x \mathbf{ji} + A_y \mathbf{jj} + B_x \mathbf{jk} + B_y \mathbf{ki} + B_x \mathbf{kj} + A_z \mathbf{kk}$$

Notice the impressive name "idemfactor" for the identity operator $1 = \sum_i | a_i \rangle \langle a_i | = \sum_i a_i a_i^T = \sum_i a_i a_i$.

Appendix F: Expansion of the gradient of a vector ($\nabla \mathbf{v}$) in curvilinear coordinates

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 rank-2 tensors.

In this literature one sometimes sees $(\nabla \mathbf{A})_{ij} \equiv \partial_j A_i$ where the indices are the *reverse* of the normal dyadic definition of Appendix E because this makes certain equations look simpler.

For example, in continuum mechanics one encounters the so-called convective or material derivative of an arbitrary vector field $\mathbf{A}(\mathbf{x}, t)$ in the Eulerian or spatial "view" of the motion of a blob of continuous matter (eg, Lai (3.4.3) and (3.4.8)),

$$\begin{aligned} DA_i/Dt &= \partial_t A_i + \nabla A_i \bullet \mathbf{v} = \partial_t A_i + (\partial_j A_i) v_j = \partial_t A_i + (\nabla \mathbf{A})_{ij} v_j = \partial_t A_i + [(\nabla \mathbf{A}) \mathbf{v}]_i \\ \Rightarrow \quad D\mathbf{A}/Dt &= \partial_t \mathbf{A} + (\nabla \mathbf{A}) \mathbf{v} \end{aligned}$$

Here $\mathbf{v}(\mathbf{x}, t)$ is the velocity field of the moving matter blob. The notation DA_i/Dt is just a historical notation for the total derivative $d_t A_i = dA_i(\mathbf{x}, t)/dt$.

The matrix $\nabla \mathbf{A}$ is not a differential operator since the derivative does not act on the vector standing to the right of $\nabla \mathbf{A}$, but one is still often interested in expressing $\nabla \mathbf{A}$ in curvilinear coordinates.

We now carry out this task as an illustration of the various notations presented in Appendix E. We replace the generic vector $\mathbf{A}$ by generic vector $\mathbf{v}$ and this $\mathbf{v}$ has nothing to do with the $\mathbf{v}$ shown in the example above. As usual for the curvilinear coordinates application, $\mathbf{x}$ -space is taken to be Cartesian, and $\mathbf{x}'$ -space is that of the curvilinear coordinates $\mathbf{x}'$.

Consider then the dyadic $(\nabla \mathbf{v})$ (a dyadic with indices reversed from the usual dyadic sense)

$$(\nabla \mathbf{v})_{ij} \equiv \partial_j v_i$$

As shown in Appendix E, this dyadic can be expanded on the Cartesian $\mathbf{x}$ -space basis as

$$(\nabla \mathbf{v}) = \Sigma_{cd} (\nabla \mathbf{v})_{dc} \mathbf{u}_d \mathbf{u}_c^T = \Sigma_{cd} (\partial_c v_d) \mathbf{u}_d \mathbf{u}_c^T .$$

The general plan for expressing $(\nabla \mathbf{v})$ in curvilinear coordinates is this:

(1) to cause $\mathbf{u}_d \mathbf{u}_c^T$ to be replaced by $\mathbf{e}_d \mathbf{e}_c^T$, where $\mathbf{e}_n$ are the tangent basis vectors of the transformation from Cartesian $\mathbf{x}$ -space coordinates to curvilinear $\mathbf{x}'$ -space coordinates under the non-linear transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. As shown in the various Sections of this document on differential operators, the $\mathbf{e}_n$ vectors in $\mathbf{x}$ -space (or their unit vector versions $\hat{\mathbf{e}}_n$) are the appropriate basis vectors for expressing any tensor (or tensor-like object) where $\mathbf{x}'$ -space coefficients are desired. That is to say, as discussed in Appendix E,

$$\mathbf{A} = \Sigma_{ij} A'^{ij} \mathbf{e}_i \mathbf{e}_j^T$$

(2) to write out $(\partial_c v_d)$ in terms of $\mathbf{x}'$ -space coordinates and objects (like v'_i)

The second part of this plan is quite simple so we do it will be done first.

(a) Expressing $\partial_c v_d$ in terms of x' -space objects

One can express $\partial_c v_d$ in terms of x' -space coordinates and objects in exactly the Christoffel manner of Section 7 (v). Using the covariant vector rule $V_a = R^b_a V'_b$ shown in Section 7 (p) one easily gets

$$\partial_c v_d = \Sigma_{ij} (R^i_c \partial'_i)(R^j_d v'_j) = \Sigma_{ij} R^i_c [(\partial'_i R^j_d) v'_j + R^j_d (\partial'_i v'_j)] \quad (*)$$

Comment: As noted in Section 7 (v), for a non-linear transformation F the object $\partial_c v_d$ is *not* a rank-2 tensor. If it *were* a rank 2 tensor, as in the case of linear F , it would be very easy to express $\partial_c v_d$ in terms of x' -space objects and coordinates, according to the expansion of Appendix E,

$$A = \Sigma_{ij} A'_{ij} \mathbf{e}_i \mathbf{e}_j^T \quad A'_{ij} = R_i^a R_j^b A_{ab} \quad // A \text{ is a rank-2 tensor}$$

$$(\nabla v) = \Sigma_{ij} (\nabla v)'_{ij} \mathbf{e}_i \mathbf{e}_j^T \quad (\nabla v)'_{ij} = R_i^a R_j^b (\nabla v)_{ab} = R_i^a R_j^b \partial_b v_a = R_j^b \partial_b R_i^a v_a = (\partial'_j v'_i)$$

so that $(\nabla v) = \Sigma_{ij} (\partial'_j v'_i) \mathbf{e}_i \mathbf{e}_j^T$ and we are all done. When F is non-linear, as is the case for curvilinear coordinate transformations, the extra work shown below or its equivalent is required.

(b) Expressing $u_d u_c^T$ in terms of $e_i e_j^T$

In Section 3 (b) it was shown that $\mathbf{e}'_n \equiv R \mathbf{e}_n = \Sigma_i R_{in} \mathbf{e}_i$ which in standard notation reads

$$\mathbf{e}'_n = \Sigma_i R^n_i \mathbf{e}_i \quad \text{where} \quad (\mathbf{e}'_n)^i = \delta_n^i \quad (*)$$

This can quickly be verified by dotting both sides with $\mathbf{e}^m$ (the dual basis is a complete basis)

$$\mathbf{e}'_n \bullet \mathbf{e}^m = \Sigma_i R^n_i \mathbf{e}_i \bullet \mathbf{e}^m$$

$$\text{LHS} = \mathbf{e}'_n \bullet \mathbf{e}^m = g_{ab} (\mathbf{e}'_n)^a (\mathbf{e}^m)^b = g_{ab} \delta_n^a (\mathbf{e}^m)^b = g_{nb} (\mathbf{e}^m)^b = g_n^b (\mathbf{e}^m)_b = (\mathbf{e}^m)_n = R^m_n$$

$$\text{RHS} = \Sigma_i R^n_i \mathbf{e}_i \bullet \mathbf{e}^m = \Sigma_i R^n_i \delta_i^m = R^m_n$$

For present purposes, (*) can be re-expressed as,

$$\mathbf{u}_d = \Sigma_e R^e_d \mathbf{e}_e \quad \text{where} \quad (\mathbf{u}_d)^e = \delta_d^e$$

Then

$$\mathbf{u}_c^T = \Sigma_f R^f_c \mathbf{e}_f^T$$

and so

$$\mathbf{u}_d \mathbf{u}_c^T = \sum_{ef} R_d^e R_c^f \mathbf{e}_e \mathbf{e}_f^T$$

and our section (b) task is completed.

(c) Combining the two steps

It has now been shown in section (a) that

$$\partial_c v_d = \sum_{ij} R_c^i [(\partial'_i R_d^j) v'_j + R_d^j (\partial'_i v'_j)]$$

and in section (b) that

$$\mathbf{u}_d \mathbf{u}_c^T = \sum_{ef} R_d^e R_c^f \mathbf{e}_e \mathbf{e}_f^T$$

Therefore, carefully doing one step at a time,

$$\begin{aligned} (\nabla \mathbf{v}) &= \sum_{cd} (\nabla \mathbf{v})_{dc} \mathbf{u}_d \mathbf{u}_c^T = \sum_{cd} (\partial_c v_d) \mathbf{u}_d \mathbf{u}_c^T \\ &= \sum_{cd} \{ \sum_{ij} R_c^i [(\partial'_i R_d^j) v'_j + R_d^j (\partial'_i v'_j)] \} \{ \sum_{ef} R_d^e R_c^f \mathbf{e}_e \mathbf{e}_f^T \} \\ &= \sum_{cd} \sum_{ef} \{ \sum_{ij} R_c^i [(\partial'_i R_d^j) v'_j + R_d^j (\partial'_i v'_j)] \} \{ R_d^e R_c^f \mathbf{e}_e \mathbf{e}_f^T \} \\ &= \sum_{cd} \sum_{ef} \{ \sum_{ij} R_c^f R_c^i [R_d^e (\partial'_i R_d^j) v'_j + R_d^e R_d^j (\partial'_i v'_j)] \} \mathbf{e}_e \mathbf{e}_f^T \\ &= \sum_{ef} \sum_{ij} (\sum_c R_c^f R_c^i) [\sum_d R_d^e (\partial'_i R_d^j) v'_j + (\sum_d R_d^e R_d^j) (\partial'_i v'_j)] \mathbf{e}_e \mathbf{e}_f^T \end{aligned}$$

Since the metric tensor is in fact a tensor, one knows that $g^{fi} = R_c^f R_c^i g^{ad}$. But $g = 1$ so this becomes $g^{fi} = R_c^f R_c^i$ and so one has

$$\begin{aligned} (\sum_c R_c^f R_c^i) &= g^{fi} \\ (\sum_d R_d^e R_d^j) &= g^{ej} \end{aligned}$$

which can be inserted into the above expression to give,

$$= \sum_{ef} \sum_{ij} g^{fi} [\sum_d R_d^e (\partial'_i R_d^j) v'_j + g^{ej} (\partial'_i v'_j)] \mathbf{e}_e \mathbf{e}_f^T$$

Change ij to ab

$$= \sum_{ef} \sum_{ab} g^{fa} [\sum_d R_d^e (\partial'_a R_d^b) v'_b + g^{eb} (\partial'_a v'_b)] \mathbf{e}_e \mathbf{e}_f^T$$

and then change ef to ij

$$= \sum_{ij} \{ \sum_{ab} g^{ja} [\sum_d R_d^i (\partial'_a R_d^b) v'_b + g^{ib} (\partial'_a v'_b)] \} \mathbf{e}_i \mathbf{e}_j^T .$$

Therefore, it has been shown that

$$(\nabla \mathbf{v}) = \Sigma_{ij} Q^{ij} \mathbf{e}_i \mathbf{e}_j^T = \Sigma_{ij} [(\nabla \mathbf{v})^{(e)}]^{ij} \mathbf{e}_i \mathbf{e}_j^T$$

where

$$Q^{ij} \equiv \Sigma_{ab} g'^{ja} [\Sigma_d R^i_d (\partial'_a R^b_d) v'_b + g'^{ib} (\partial'_a v'_b)] = [(\nabla \mathbf{v})^{(e)}]^{ij}$$

Here the notation $(\nabla \mathbf{v})^{(e)}$ indicates, as suggested in Appendix E, that this matrix is associated with the $\mathbf{e}_i$ basis which in turn is associated with the curvilinear coordinates $\mathbf{x}'$. Since g' and R are functions of the curvilinear coordinates $\mathbf{x}'$, our program of expressing $(\nabla \mathbf{v})$ entirely in terms of $\mathbf{x}'$ -space coordinates and objects is now complete. Notice that $\mathbf{x}'$ is an arbitrary not necessarily orthogonal coordinate system.

(d) Alternate forms of the $(\nabla \mathbf{v})$ expansion

In terms of *covariant* components of vector $\mathbf{v}'$, it was shown above that

$$Q^{ij} = \Sigma_{ab} g'^{ja} [\Sigma_d R^i_d (\partial'_a R^b_d) v'_b + g'^{ib} (\partial'_a v'_b)] \quad (\nabla \mathbf{v}) = \Sigma_{ij} Q^{ij} \mathbf{e}_i \mathbf{e}_j^T$$

If one wants to see *contravariant* components of vector $\mathbf{v}'$, replace $v'_b = \Sigma_c g'_{bc} v'^c$ to get

$$Q^{ij} = \Sigma_{ab} g'^{ja} [\Sigma_d R^i_d (\partial'_a R^b_d) \Sigma_c g'_{bc} v'^c + g'^{ib} \partial'_a (\Sigma_c g'_{bc} v'^c)] .$$

Or, if one wants to see the *unit-vector* expansion components of vector $\mathbf{v}'$, replace $v'^c = h'^c_{c} u'^c$ to get

$$Q^{ij} = \Sigma_{ab} g'^{ja} [\Sigma_d R^i_d (\partial'_a R^b_d) \Sigma_c g'_{bc} h'^c_{c} u'^c + g'^{ib} \partial'_a (\Sigma_c g'_{bc} h'^c_{c} u'^c)] .$$

As a reminder, these u'^n arise as follows:

$$\mathbf{v} = \Sigma_n v'^n \mathbf{e}_n = \Sigma_n v'^n (h'_n \hat{\mathbf{e}}_n) = \Sigma_n (h'_n v'^n) \hat{\mathbf{e}}_n = \Sigma_n u'^n \hat{\mathbf{e}}_n \quad \Rightarrow \quad u'^n \equiv h'_n v'^n .$$

As discussed in Section 14 Example 1, the components u'^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 u'^n and just write $u'^n = v_{\mathbf{x}', (n)}$. For example, in spherical coordinates r, θ, ϕ :

$$u'^1 = v_r \quad u'^2 = v_\theta \quad u'^3 = v_\phi \quad \mathbf{v} = \Sigma_n u'^n \hat{\mathbf{e}}_n = v_r \hat{\mathbf{r}} + v_\theta \hat{\boldsymbol{\theta}} + v_\phi \hat{\boldsymbol{\phi}} .$$

Once one is talking u'^n and $\hat{\mathbf{e}}_n$, one usually wants to see the $(\nabla \mathbf{v})$ expansion in terms of $\hat{\mathbf{e}}_i \hat{\mathbf{e}}_j^T$. Since

$$\mathbf{e}_i \mathbf{e}_j^T = h'_i h'_j \hat{\mathbf{e}}_i \hat{\mathbf{e}}_j^T$$

our expansion can be rephrased in this manner:

$$(\nabla \mathbf{v}) = \Sigma_{ij} Q^{ij} \mathbf{e}_i \mathbf{e}_j^T = \Sigma_{ij} (Q^{ij} h'_i h'_j) \hat{\mathbf{e}}_i \hat{\mathbf{e}}_j^T = \Sigma_{ij} P^{ij} \hat{\mathbf{e}}_i \hat{\mathbf{e}}_j^T$$

with

$$P^{ij} = h'_i h'_j Q^{ij} = [(\nabla \mathbf{v})^{(\mathbf{e}^\wedge)}]^{ij}$$

where Q^{ij} can be any of the three forms shown above.

(e) Special case of orthogonal coordinates

In this case $g'_{nm} = h_n'^2 \delta_{n,m}$ and $g'^{nm} = h_n'^{-2} \delta_{n,m}$. We consider only the last form for Q^{ij} above. Since the algebra is tedious, here are all the steps, where the next item to be processed is shown in red.

$$\begin{aligned} Q^{ij} &= \Sigma_{ab} \textcolor{red}{g'^{ja}} [\Sigma_d R^i_d (\partial'_a R^b_d) \Sigma_c g'_{bc} h'_c{}^{-1} \mathbf{u}^c + g'^{ib} \partial'_a (\Sigma_c g'_{bc} h'_c{}^{-1} \mathbf{u}^c)] \\ &= \Sigma_{ab} \textcolor{red}{h'_j{}^{-2} \delta_{j,a}} [\Sigma_d R^i_d (\partial'_a R^b_d) \Sigma_c g'_{bc} h'_c{}^{-1} \mathbf{u}^c + g'^{ib} \partial'_a (\Sigma_c g'_{bc} h'_c{}^{-1} \mathbf{u}^c)] \\ &= \Sigma_b \textcolor{red}{h'_j{}^{-2}} [\Sigma_d R^i_d (\partial'_j R^b_d) \Sigma_c g'_{bc} h'_c{}^{-1} \mathbf{u}^c + g'^{ib} \partial'_j (\Sigma_c g'_{bc} h'_c{}^{-1} \mathbf{u}^c)] \\ &= h'_j{}^{-2} [\Sigma_d R^i_d \Sigma_b (\partial'_j R^b_d) \Sigma_c g'_{bc} h'_c{}^{-1} \mathbf{u}^c + \Sigma_b \textcolor{red}{g'^{ib}} \partial'_j (\Sigma_c g'_{bc} h'_c{}^{-1} \mathbf{u}^c)] \\ &= h'_j{}^{-2} [\Sigma_d R^i_d \Sigma_b (\partial'_j R^b_d) \Sigma_c g'_{bc} h'_c{}^{-1} \mathbf{u}^c + \Sigma_b \textcolor{red}{h'_i{}^{-2} \delta_{i,b}} \partial'_j (\Sigma_c g'_{bc} h'_c{}^{-1} \mathbf{u}^c)] \\ &= h'_j{}^{-2} [\Sigma_d R^i_d \Sigma_b (\partial'_j R^b_d) \Sigma_c \textcolor{red}{g'_{bc}} h'_c{}^{-1} \mathbf{u}^c + h'_i{}^{-2} \partial'_j (\Sigma_c \textcolor{red}{g'_{ic}} h'_c{}^{-1} \mathbf{u}^c)] \\ &= h'_j{}^{-2} [\Sigma_d R^i_d \Sigma_b (\partial'_j R^b_d) \Sigma_c \textcolor{red}{h'_c{}^2 \delta_{b,c}} h'_c{}^{-1} \mathbf{u}^c + h'_i{}^{-2} \partial'_j (\Sigma_c \textcolor{red}{h'_i{}^2 \delta_{i,c}} h'_c{}^{-1} \mathbf{u}^c)] \\ &= h'_j{}^{-2} [\Sigma_d R^i_d \Sigma_b (\partial'_j R^b_d) h'_b \mathbf{u}^b + h'_i{}^{-2} \partial'_j (h'_i \mathbf{u}^i)] \\ &= h'_j{}^{-2} h'_i{}^{-2} [h'_i{}^2 \Sigma_d R^i_d \Sigma_b (\partial'_j R^b_d) h'_b \mathbf{u}^b + \partial'_j (h'_i \mathbf{u}^i)] \\ &= h'_j{}^{-2} h'_i{}^{-2} [h'_i{}^2 \Sigma_d R^i_d \Sigma_b (\partial'_j R^b_d) h'_b \mathbf{u}^b + (\partial'_j h'_i) \mathbf{u}^i + h'_i (\partial'_j \mathbf{u}^i)] \end{aligned}$$

so that then

$$P^{ij} = h'_j{}^{-1} h'_i{}^{-1} [h'_i{}^2 \Sigma_d R^i_d \Sigma_b (\partial'_j R^b_d) h'_b \mathbf{u}^b + (\partial'_j h'_i) \mathbf{u}^i + h'_i (\partial'_j \mathbf{u}^i)]$$

This object P^{ij} can be computed in Maple by the following code,

Computation of (grad v) in spherical coordinates 1.11.12.

This file is also the template for any other cases.

```
> restart;
> with(linalg);
> N := 3;
> xp[1] := r;
> xp[2] := theta;
> xp[3] := phi;
> assume(r>0,theta>0,theta<Pi);
> x[1] := r*sin(theta)*cos(phi);
> x[2] := r*sin(theta)*sin(phi);
> x[3] := r*cos(theta);
> Vp := vector( [v[xp[1]],v[xp[2]],v[xp[3]] ] );
> S_ := (i,j) -> diff(x[i],xp[j]);
> S := matrix(N,N,S_);
R is not used anywhere, but we compute it anyway.
> R := simplify(inverse(S));
> gcov := simplify(evalm( transpose(S) &* S ));
> for n from 1 to N do hp[n] := simplify(sqrt(gcov[n,n])) od;
> 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 );
> T2_ := (i,j) -> Vp[i] * diff(hp[i],xp[j]) ;
> T3_ := (i,j) -> hp[i] * Diff(Vp[i],xp[j]) ;
> P_ := (i,j) -> (1/hp[i])*(1/hp[j])*(value(T1_(i,j)) + (T2_(i,j)) + T3_(i,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);
```

which is easily modified for other orthogonal curvilinear systems. Here are some sample results:

$$(\nabla v) = \sum_{i,j} P^{ij} \hat{e}_i \hat{e}_j^T \quad P^{ij} = [(\nabla v)^{(\hat{e})}]^{ij}$$

- P^{ij} in **polar** coordinates (where $1,2 = r,\theta$) :

$$\begin{bmatrix} \frac{\partial}{\partial r} v_r & -\frac{v_\theta}{r} + \frac{\frac{\partial}{\partial \theta} v_r}{r} \\ \frac{\partial}{\partial r} v_\theta & \frac{v_r}{r} + \frac{\frac{\partial}{\partial \theta} v_\theta}{r} \end{bmatrix}$$

// agrees with Lai (2.23.23)

- P^{ij} in **cylindrical** coordinates (where $1,2,3 = r,\theta,z$) :

$$\begin{bmatrix} \frac{\partial}{\partial r} v_r & -\frac{v_\theta}{r} + \frac{\frac{\partial}{\partial \theta} v_r}{r} & \frac{\partial}{\partial z} v_r \\ \frac{\partial}{\partial r} v_\theta & \frac{v_r}{r} + \frac{\frac{\partial}{\partial \theta} v_\theta}{r} & \frac{\partial}{\partial z} v_\theta \\ \frac{\partial}{\partial r} v_z & \frac{\frac{\partial}{\partial \theta} v_z}{r} & \frac{\partial}{\partial z} v_z \end{bmatrix}$$

// agrees with Lai (2.34.5)

The polar coordinates results are seen to be upper left 2x2 piece of the cylindrical results.

• P^{ij} in **spherical** coordinates (where 1,2,3 = r,θ,φ) :

$$\begin{bmatrix} \frac{\partial}{\partial r} v_r & -\frac{v_\theta}{r} + \frac{\frac{\partial}{\partial \theta} v_r}{r} & -\frac{v_\phi}{r} + \frac{\frac{\partial}{\partial \phi} v_r}{r \sin(\theta)} \\ \frac{\partial}{\partial r} v_\theta & \frac{v_r}{r} + \frac{\frac{\partial}{\partial \theta} v_\theta}{r} & -\frac{\cos(\theta) v_\phi}{r \sin(\theta)} + \frac{\frac{\partial}{\partial \phi} v_\theta}{r \sin(\theta)} \\ \frac{\partial}{\partial r} v_\phi & \frac{\frac{\partial}{\partial \theta} v_\phi}{r} & \frac{v_r}{r} + \frac{\cos(\theta) v_\theta}{r \sin(\theta)} + \frac{\frac{\partial}{\partial \phi} v_\phi}{r \sin(\theta)} \end{bmatrix}$$

// agrees with Lai (2.35.25)

References

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 Mathematik*, **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).

M. Lai, E. Krempf and D. Ruben, *Introduction to Continuum Mechanics, 4th Ed.* (Elsevier, Amsterdam, 2009).

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 go to find explicit expressions for differential operators in specific curvilinear coordinate systems (and very much more).

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 (u) above).

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)