

Tensor Analysis and Curvilinear Coordinates

Phil Lucht

Rimrock Digital Technology, Salt Lake City, Utah 84103

last update: Feb 10, 2014

rimrock@xmission.com

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

The material in this document is copyrighted by the author.

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

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

Overview and Summary	7
1. The Transformation F: invertibility, coordinate lines, and level surfaces	11
Example 1: Polar coordinates (N=2)	11
Example 2: Spherical coordinates (N=3)	12
Cartesian Space and Quasi-Cartesian Space	14
Pictures A,B,C and D	14
Coordinate Lines	15
Example 1: Polar coordinates, coordinate lines	16
Example 2: Spherical coordinates, coordinate lines	16
Level Surfaces	17
2. Linear Local Transformations associated with F : scalars and two kinds of vectors	19
(a) Scalars	20
(b) Contravariant vectors	21
(c) Covariant vectors	21
(d) Bar notation	22
(e) Origin of the names contravariant and covariant	22
(f) Other vector types?	23
(g) Linear transformations	23
(h) Vectors that are contravariant by definition	24
(i) Vector Fields	25
(j) Names and symbols	25
(k) Definition of the words "scalar", "vector" and "tensor"	26
3. Tangent Base Vectors e_n and Inverse Tangent Base Vectors u'_n	28
(a) Definition of the e_n ; the e_n are the columns of S	29
(b) e_n as a contravariant vector	30
(c) a semantic question: unit vectors	31
Example 1: Polar coordinates, tangent base vectors	31
Example 2: Spherical Coordinates, tangent base vectors	33
(d) The inverse tangent base vectors u'_n and inverse coordinate lines	34
Example 1: Polar coordinates: inverse tangent base vectors and inverse coordinate lines	35
4. Notions of length, distance and scalar product in Cartesian Space	36

5. The Metric Tensor	38
(a) Definition of the metric tensor.....	38
(b) Inverse of the metric tensor	40
(c) A metric tensor is symmetric.....	41
(d) $\det(g)$ and g_{nn} of a Cartesian-generated metric tensor are non-negative	41
(e) Definition of two kinds of rank-2 tensors.....	41
(f) Proof that the metric tensor and its inverse are both rank-2 tensors	42
(g) Metric tensor converts vector types.....	44
(h) Vectors in Cartesian space	44
(i) Metric tensor: covariant scalar product and norm.....	45
(j) Metric tensor and tangent base vectors	47
(k) The Jacobian J	49
(l) Some relations between g , R and S in Pictures B and C (Cartesian x -space).....	52
Example 1: Polar coordinates: metric tensor and Jacobian.....	53
Example 2: Spherical coordinates: metric tensor and Jacobian	54
(m) Special Relativity and its Metric Tensor: vectors and spinors	54
(n) General Relativity and its Metric Tensor	58
(o) Continuum Mechanics and its Metric Tensors	59
6. Reciprocal Base Vectors E_n and Inverse Reciprocal Base Vectors U'_n.....	63
(a) Definition of the E_n	63
(b) The Dot Products and Reciprocity (Duality).....	64
(c) Covariant partner for E_n	66
(d) Summary of the basic facts:	66
(e) Repeat the above for the inverse transformation: definition of the U'_n	67
(f) Expanding vectors on different sets of basis vectors	68
(g) Another way to write the E_n	70
(h) Comparison of $\bar{e}_n$ and E_n	71
(i) Handedness of coordinate systems: the e_n , the sign of $\det(S)$, and Parity.....	72
7. Translation to the Standard Notation	76
(a) Outer Products	76
(b) Mixed Tensors and Notation Issues	76
(c) The up/down bell goes off.....	77
(d) Some Preliminary Translations: raising and lowering indices on a vector with g	78
(e) Contraction of a Pair of Indices.....	79
(f) Dealing with the matrix R	80
(g) Repeat the above section for S	81
(h) About ε and δ	81
(i) Matrix Multiplication, the meaning of R^T , and Rotations in Standard Notation.....	82
1. Translation of determinants.	82
2. Inverse of R and S	82
3. Tensor g raises and lowers any index.	83
4. Raising Lowering Rule:	84
5. Contraction Tilt Reversal Rule:	85
6. The Diagonal g Rule:	85
7. Matrix Multiplication in the Standard Notation.....	85
8. Transpose of a rank-2 tensor.....	89

9. Orthogonality rules.	91
10. Rotation matrices.	91
11. Variations on the relation between g and g'	94
(j) Tensors of Rank n , direct products, Lie groups, symmetry and Ricci-Levi-Civita.....	94
(k) The Contraction Tilt-Reversal Rule	97
(ℓ) The Contraction Neutralization Rule	99
(m) Raising and lowering indices on g	101
(n) Other forms of R	101
(o) Summary of facts about R	102
(p) Repeat all the above for S	102
(q) Theorem: $S^a_b = R^a_c R^c_b$ and $S^b_a = R^b_c R^c_a$ (reflect indices in vertical line between them)	102
(r) Orthogonality Rules, the Inversion Rule, and the Cancellation Rule	104
(s) The tangent and reciprocal base vectors and expansions on same	105
(t) Comment on Covariant versus Contravariant	109
(u) The Significance of Tensor Analysis	110
(v) The Christoffel Business: covariant derivatives.....	113
(w) Expansions of higher order tensors	114
8. Transformation of Differential Length, Area and Volume	116
Overview	116
(a) The differential N-piped mapping	117
(b) Properties of the finite N-piped spanned by the e_n in x -space.....	119
(c) Back to the differential N-piped mapping: how edges, areas and volume transform.....	121
1. The Setup.	121
2. Edge Transformation.....	122
3. Area Transformation.....	122
4. Volume Transformation.....	124
5. Covariant Magnitudes.....	125
6. Two Theorems : $g' / h'^2_n = g'^{nn} g' = \text{cof}(g'_{nn})$ and $ (II^x_{i \neq n} e_i) = \sqrt{\text{cof}(g'_{nn})}$	126
7. Cartesian-View Magnitude Ratios.....	127
8. Nested Cofactor Formulas.....	128
9. Transformation of <i>arbitrary</i> differential vectors, areas and volume.....	128
10. Concatenation of Transformations.....	130
Examples of area magnitude transformation for $N = 2,3,4$	131
Example 2: Spherical Coordinates: area patches	132
(d) Transformation of Differential Volume applied to Integration.....	133
(e) Interpretations of the Jacobian.....	135
(f) Volume integration of a tensor field under linear transformations	135
9. The Divergence in curvilinear coordinates	137
(a) Geometric Derivation of the Curvilinear Divergence Formula	137
(b) Various expressions for $\text{div } B$	140
(c) Translation from Picture B to Picture M&S	142
(d) Comparison of various authors' notations	143
10. The Gradient in curvilinear coordinates.....	145
(a) Expressions for $\text{grad } f$	145
(b) Expressions for $\text{grad } f \bullet B$	147
11. The Laplacian in curvilinear coordinates	149

12. The Curl in curvilinear coordinates.....	151
(a) Definition of curl $\mathbf{B}$	151
(b) Computation of the line integral.....	152
(c) Solving for the curl.....	154
(d) Various forms of the curl.....	155
(e) The curl in orthogonal coordinate systems.....	157
(f) The curl in $N > 3$ dimensions.....	157
13. The Vector Laplacian in curvilinear coordinates.....	159
(a) Derivation of the Vector Laplacian in general curvilinear coordinates.....	159
(b) The Vector Laplacian in orthogonal curvilinear coordinates	161
(c) The Vector Laplacian in Cartesian coordinates.....	163
14. Summary of Differential Operators in curvilinear coordinates	165
(a) divergence.....	166
(b) gradient and gradient dot vector	166
(c) Laplacian	167
(d) curl.....	167
(e) vector Laplacian	168
Example 1: Polar coordinates: a practical curvilinear notation.....	169
15. Covariant derivation of all curvilinear differential operator expressions	171
(a) Review of Sections 9 through 13.....	171
(b) The Covariant Method.....	172
(c) divergence (Section 9)	174
(d) gradient and gradient dot vector (Section 10)	174
(e) Laplacian (Section 11).....	174
(f) curl (Section 12).....	175
(g) vector Laplacian (Section 13).....	175
Appendix A: Reciprocal Base Vectors the Hard Way.....	179
(a) Definition of $\mathbf{E}_n$	179
(b) Simpler notation	180
(c) Generalized Cross Product of $N-1$ vectors of dimension N	180
(d) Missing Man Formation	182
(e) Apply this Notation to $\mathbf{E}$	182
(f) Compute $\mathbf{E}_m \bullet \mathbf{e}_n$	183
(g) Compute $\mathbf{E}_n \bullet \mathbf{E}_m$	184
(h) Summary of relationship between the tangent and reciprocal base vectors	184
(i) Another Cross Product Notation and another expression for $\mathbf{E}$	185
Appendix B: The Geometry of Parallelepipeds in N dimensions.....	186
(a) Preliminary: Equation of a plane in N dimensions	186
(b) N -pipeds and their Faces in Various Dimensions	187
The 1-piped	187
The 2-piped	187
The 3-piped	189
The N -piped	190
(c) The question of inward versus outward facing normal vectors.	191
(d) The Face Area and Volume of N -pipeds in Various Dimensions	192
The 2-piped	192

The 3-piped.....	193
The 4-piped.....	195
The N-piped.....	197
(e) Summary of Main Results of this Appendix.....	199
Appendix C: Elliptical Polar Coordinates (N=2, non-orthogonal).....	201
(a) Elliptical polar coordinates.....	201
(b) Forward coordinate lines.....	202
(c) Inverse coordinate lines.....	202
(d) Drawing a contravariant vector V in x-space: the meaning of V'_n	203
(e) Drawing a contravariant vector V' in x'-space: two "Views".....	204
(f) Drawing the specific contravariant vector dx in x-space and x'-space.....	207
(g) Study of how dx transforms in the mapping between x-space and x'-space.....	207
(h) Derivation of the Jacobian Integration Rule.....	209
Appendix D: Tensor Densities and the ϵ tensor	212
(a) Definition of a tensor density.....	212
(b) A few facts about tensor densities.....	213
(c) Theorem about Totally Antisymmetric Tensors: there is really only one: $\epsilon^{abc}\dots$	215
(d) The contravariant ϵ tensor.....	216
(e) Some facts about the ϵ tensor.....	217
(f) The covariant ϵ tensor : repeat section (d) as if its weight were not known.....	219
(g) Generalized cross products.....	220
(h) The tensorial nature of curl B.....	220
(i) Tensor $\mathcal{E}$ as a weight 0 version of ϵ : three conventions.....	221
(j) Representation of ϵ , $\epsilon\epsilon$ and contracted $\epsilon\epsilon$ as determinants.....	224
(k) Covariant forms of the previous section results.....	230
Appendix E: Tensor Expansions: direct product, polyadic and operator notation.....	232
(a) Direct Product Notation.....	232
(b) Tensor Expansions and Bases.....	233
(c) Polyadic Notation.....	235
(d) Dyadic Products.....	236
(e) Transpose notation for dyadics.....	237
(f) Large and small dots used with dyadics.....	238
(g) Operators and Matrices for Rank-2 tensors.....	239
(h) Expansions of tensors on unit tangent base vectors.....	243
(i) Tensor expansions in a mixed basis.....	251
(j) What is a tensor?.....	253
Appendix F: The Affine Connection Γ^c_{ab} and Covariant Derivatives	255
(a) Definition and Interpretation of Γ : $\Gamma^c_{ab} = e^c \bullet (\partial_a e_b) = R^c_i (\partial_a R_b^i)$	255
(b) Identities of the form $(\partial_a R^d_n) = -R^e_n R^d_m (\partial_a R_e^m)$	256
(c) Identities of the form $(\partial_c g^{ab}) = -[g^{an} \Gamma^b_{cn} + g^{bn} \Gamma^a_{cn}]$	257
(d) Identity: $\Gamma^d_{ab} = (1/2) g^{dc} [\partial_a g_{bc} + \partial_b g_{ca} - \partial_c g_{ab}]$	258
(e) Picture D1 Context.....	260
(f) Relations between Γ and Γ'	261
(g) Statement and Proof of the Covariant Derivative Theorem.....	262
(h) Rule for raising any index on a covariant derivative of a covariant tensor density.....	267
(i) Examples of covariant derivative expressions.....	268

(j) The Leibniz rule for the covariant derivative of the product of two tensor densities.....	271
Appendix G: Expansion of $(\nabla \mathbf{v})$ in curvilinear coordinates ($\mathbf{v}$ = vector)	275
(a) Continuum Mechanics motivation.....	275
(b) Expansion of $\nabla \mathbf{v}$ on $\mathbf{e}^i \otimes \mathbf{e}^j$ by Method 1: Use fact that $v_{b;a}$ is a tensor.....	275
(c) Expansion of $\nabla \mathbf{v}$ on $\mathbf{e}^i \otimes \mathbf{e}^j$ by Method 2: Use brute force.	277
(d) Expansion on $\mathbf{e}_i \otimes \mathbf{e}_j$ and $\hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j$	279
(e) Orthogonal coordinate systems	280
(f) Maple evaluation of $(\nabla \mathbf{v})$ in several coordinate systems	280
Appendix H: Expansion of $\text{div}(\mathbf{T})$ in curvilinear coordinates ($\mathbf{T}$ = rank-2 tensor)	284
(a) Continuum Mechanics motivation.....	284
(b) Expansion of $\text{div} \mathbf{T}$ on $\mathbf{e}_n$ by Method 1: Use fact that $T^{ab}{}_{;\alpha}$ is a tensor.....	284
(c) Expansion of $\text{div} \mathbf{T}$ on $\mathbf{e}_n$ by Method 2: Use brute force	285
(d) Adjustment for $\mathbf{T}$ expanded on $(\hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j)$ and $\text{div} \mathbf{T}$ expanded on $\hat{\mathbf{e}}_a$	287
(e) Maple: $\text{div} \mathbf{T}$ in cylindrical and spherical coordinates.....	288
Appendix I : The Vector Laplacian in Spherical and Cylindrical Coordinates.....	290
(a) The first method : a review.....	290
(b) The first method in spherical coordinates: Maple speaks	292
(c) The first method in spherical coordinates: putting results in traditional form.....	295
(d) The second method : Part I.....	297
(e) The second method : Part II.....	298
(f) The second method in spherical coordinates: Maple speaks again.....	299
(g) Results for Cylindrical Coordinates from both methods.....	302
Appendix J: Expansion of $(\nabla \mathbf{T})$ in curvilinear coordinates ($\mathbf{T}$ = rank-2 tensor).....	307
(a) Total time derivative as prototype equation	307
(b) Computation of components $(\nabla \mathbf{T})'_{ijk}$	308
(c) Tensor expansions of $\nabla \mathbf{T}$ on the $\mathbf{u}_n$ and $\mathbf{e}_n$ base vectors.....	309
(d) Tensor expansions of $\nabla \mathbf{T}$ on the $\hat{\mathbf{e}}_n$ base vectors.....	309
(e) Total time derivative equation written in unit-base-vector curvilinear components	311
(f) Shorthand notations and a continuum mechanics application	312
(g) Maple computation of the $(\nabla \mathcal{T})'_{ijk}$ components in spherical coordinates	314
(h) Maple computation of the $(\nabla \mathcal{T})'_{ijk}$ components in cylindrical coordinates	318
Appendix K: Deformation Tensors in Continuum Mechanics	320
(a) A Preliminary Deformation Flow Picture.....	320
(b) A More Complicated Deformation Flow Picture	325
(c) Form of a solid constitutive equation involving the deformation tensor	330
(d) Some fluid constitutive equations	331
(e) Corotational and other objective time derivatives of the Cauchy stress tensor	332
References	340

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.

Three somewhat different applications of tensor analysis are treated concurrently.

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

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

The third application deals with material flows in continuum mechanics.

The first six 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 penultimate section. The final section rederives all the same results using the notion of covariance and associated covariant derivatives.

The information is presented informally as if it were a set of lectures. Little attention is paid to 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 demonstrated how $\bar{g}$ applied to a contravariant vector V produces a vector that is covariant $\bar{V} = \bar{g} V$, and conversely $g \bar{V} = V$. In Cartesian space $g = 1$, so the two types of vectors coincide. The role of the metric tensor in the covariant vector dot product is stated, and the metric tensor is related to the tangent base vectors of Section 3. The Jacobian J and associated functions are defined, though the significance of J is deferred to Section 8. The last three subsections briefly discuss the connection between tensor algebra and special relativity (with a mention of spinor algebra), general relativity, and continuum mechanics.

Section 6 introduces the reciprocal (dual) base vectors E_n which are later called e^n in the Standard Notation. Of special interest are the covariant dot products among the e_n and E_n . It is shown how an arbitrary vector can be expanded onto different basis sets. It is found that when a contravariant vector in x-space is expanded on the tangent base vectors 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 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 describes how to expand tensors of any rank in various bases and notations.

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 $x' = F(x)$. 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.

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

Much of our content is contained in a set of Appendices.

Appendix A develops an alternative expression for the reciprocal base vector $\mathbf{E}_n$ as a generalized cross product of the tangent base vectors $\mathbf{e}_n$, applicable when x-space is Cartesian. This alternate $\mathbf{E}_n$ is shown to match the $\mathbf{E}_n$ defined in 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 tensor densities and their rules of the road. Special attention is given to the Levi-Civita ϵ tensor, including a derivation of all the $\epsilon\epsilon$ contraction formulas and their covariant statements. It is noted that the curl of a vector is a vector density.

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

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

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

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

Appendix H does the same for the vector object $\mathbf{div} \mathbf{T}$ where $\mathbf{T}$ is a rank-2 tensor.

Appendix I deals with the vector Laplacian $\star \mathbf{B}$ where $\mathbf{B}$ is a vector.

Appendix J treats the object $(\nabla \mathbf{T})$, where $\mathbf{T}$ is a rank-2 tensor.

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

Notations

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

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

$a \equiv \text{stuff}$ means a is defined by stuff.

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

$\text{diag}(a,b,c..)$ means a diagonal matrix with diagonal elements $a,b,c..$

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

Maple = a computer algebra system similar to Mathematica and MATLAB

vectors (like $\mathbf{v}$) are indicated in bold; all other tensors (like T) are unbolded.

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

$\partial_{\mathbf{i}} = \partial/\partial x^{\mathbf{i}}$ and $\partial_{\mathbf{t}} = \partial/\partial t$ are partial derivatives

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

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

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

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) real functions F_i , each of N real 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 F which is all of $\mathbb{R}^N$, and the range of mapping F which is this subset.

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

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,

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

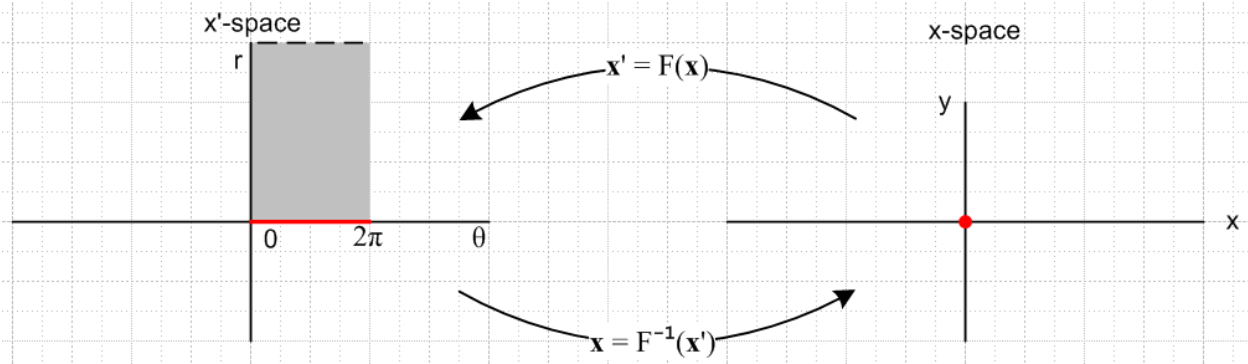
$$\mathbf{x}' = (x'_1, x'_2) = (\theta, r) \quad // \text{ note that } r = x_2'$$

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

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

(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)

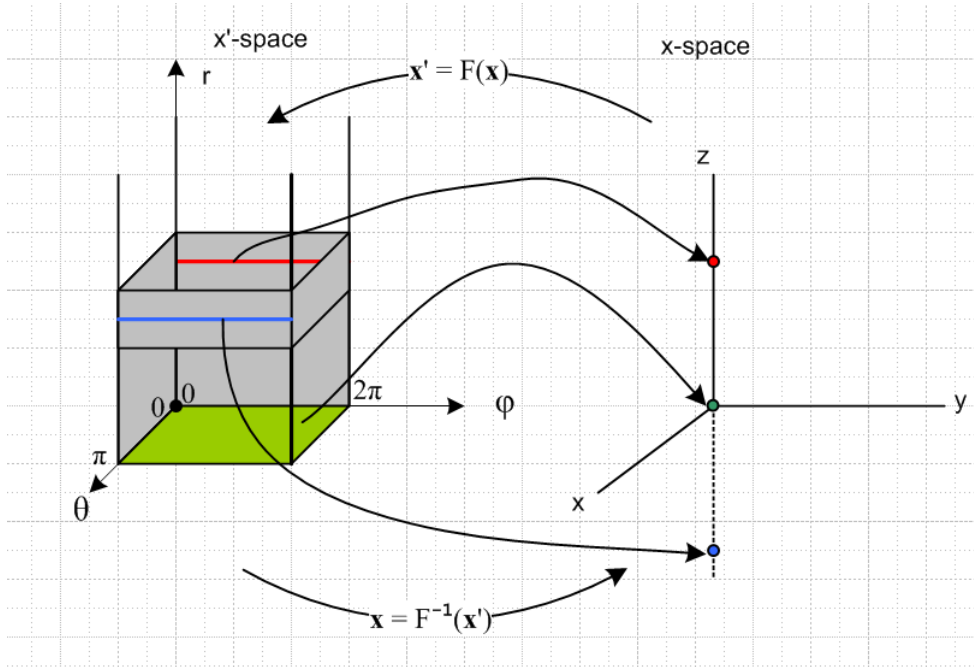
(a) The transformation from Cartesian to spherical coordinates is given by,

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

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

(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

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

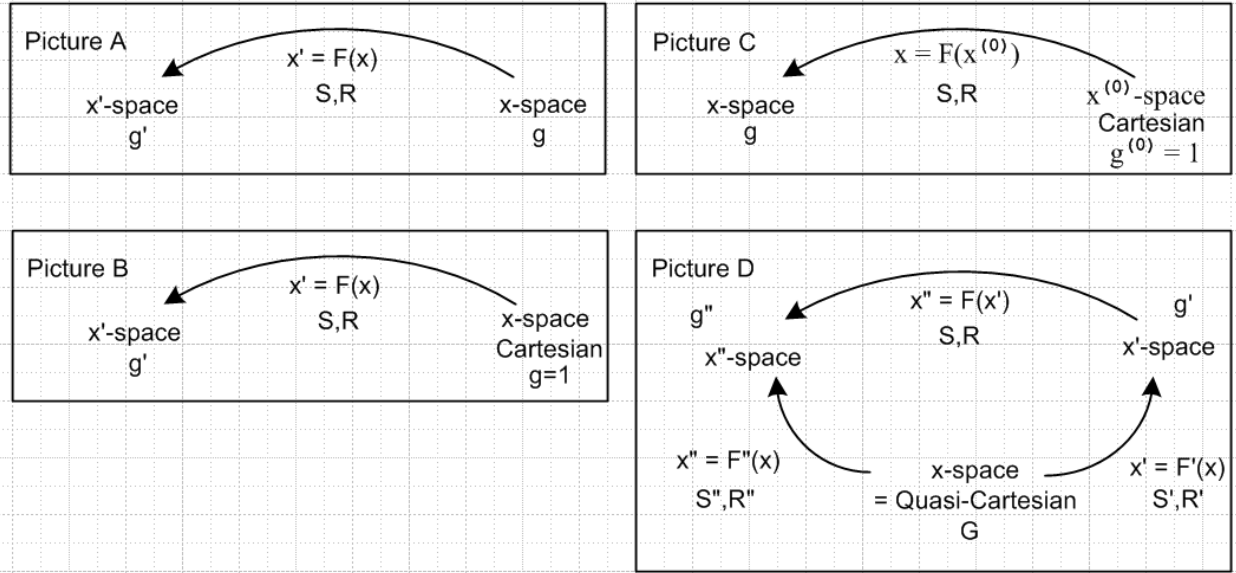
and of course this allows for the possibility of a negative distance squared (see Section 5 (i)).

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.

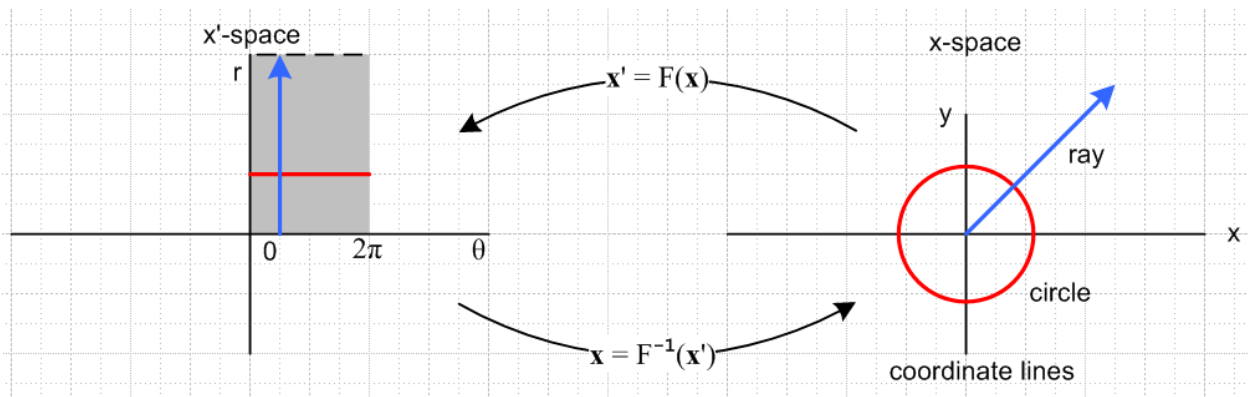
For example, in spherical coordinates we start with some (x, y, z) in x -space and compute the $x'_i = (r, \theta, \phi)$ in x' -space. Our point x in x -space then lies on the r -coordinate line whose label is r , it lies on the θ -coordinate line whose label is θ , and it lies on the ϕ -coordinate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

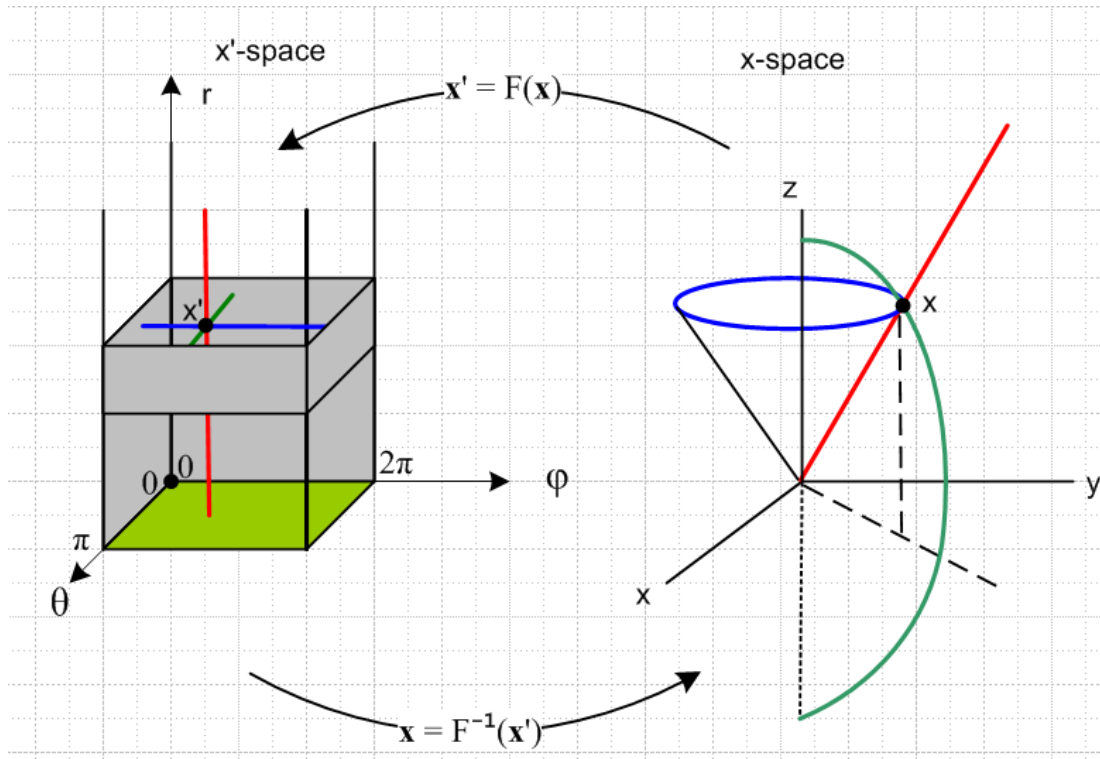
Here are some coordinate lines for our prototype $N=2$ non-linear transformation, Cartesian to polar coordinates:



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:

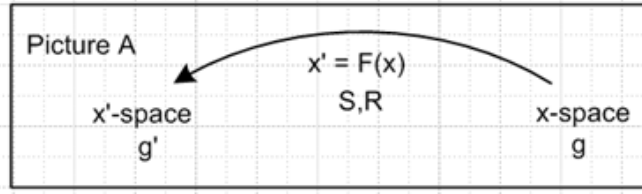
The green coordinate line is the intersection of two level surfaces: half-plane and sphere.

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

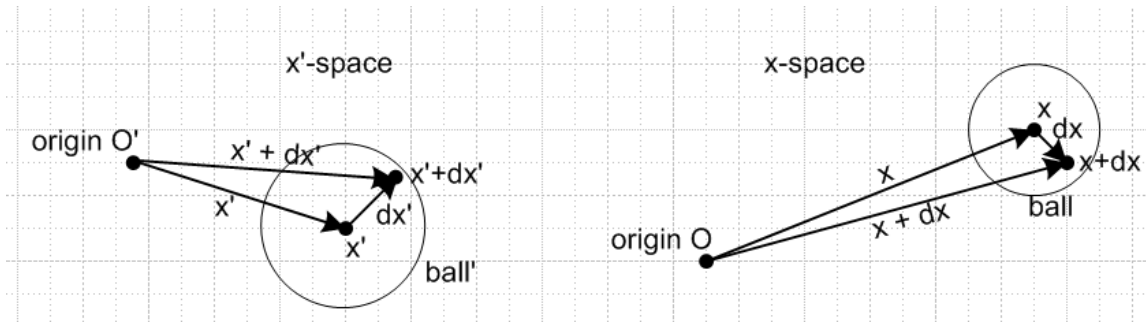
The blue coordinate line is the intersection of two level surfaces: sphere and cone.

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 $\mathbf{x}$ to map into 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,

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

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

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

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 $dx_{\mathbf{i}}$ and the same question for $dx'_{\mathbf{i}}$ and $dx'_{\mathbf{i}}$. As will be shown in Section 6 (f),

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

If x -space and 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 $\phi(\mathbf{x})$ is a "field" and it would be a "scalar field" if $\phi'(\mathbf{x}') = \phi(\mathbf{x})$. 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, and if F is time independent so $F = F(\mathbf{x})$, 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, so

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

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

$$\bar{\mathbf{E}} = -\bar{\nabla} \Phi \quad \bar{\mathbf{E}}_{\mathbf{i}} = -\bar{\partial}_{\mathbf{i}} \Phi = -\partial \Phi / \partial x_{\mathbf{i}}$$

(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 $V_{\mathbf{i}}$ and covariant vectors shall be written $\bar{\mathbf{V}}$ with components $\bar{V}_{\mathbf{i}}$. This is why overbars were placed on $\bar{\nabla}$ and $\bar{\partial}_{\mathbf{i}}$ 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:

$$\begin{array}{lll} \mathbf{V}' = \mathbf{R} \mathbf{V} & \text{contravariant} & R_{\mathbf{i}\mathbf{k}}(\mathbf{x}) \equiv (\partial x'_{\mathbf{i}} / \partial x_{\mathbf{k}}) \\ \bar{\mathbf{V}}' = \mathbf{S}^T \bar{\mathbf{V}} & \text{covariant} & S_{\mathbf{i}\mathbf{k}}(\mathbf{x}') \equiv (\partial x_{\mathbf{i}} / \partial x'_{\mathbf{k}}) = \mathbf{S}^T_{\mathbf{k}\mathbf{i}}(\mathbf{x}') \end{array} \quad \mathbf{R} = \mathbf{S}^{-1}$$

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 $\mathbf{R} \rightarrow R^{\mathbf{i}}_{\mathbf{j}}$ and $\mathbf{S} \rightarrow S^{\mathbf{i}}_{\mathbf{j}} = R_{\mathbf{j}}^{\mathbf{i}}$ 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) 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 $V_{\mathbf{n}}$ 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 Gibb's lectures in a Vector Analysis book that the idea was propagated to a wider circle. Wilson converted those Greek letters to bolded ones,

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

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)

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

where the $F_{i\mathbf{j}}$ 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 $\mathbf{x}' = \mathbf{F}(\mathbf{x})$:

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

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

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

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 $\mathbf{R}_{ik}(\mathbf{x}) = (\partial x'_i / \partial x_k)$ is called the **Jacobian matrix** for the transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$, while the matrix $\mathbf{S}_{ik}(\mathbf{x}') = (\partial x_i / \partial x'_k)$ is then the Jacobian matrix of the inverse transformation $\mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$. The determinant of the Jacobian matrix $\mathbf{S}$ 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 R and $S = R^{-1}$. One option was to use $R = L$, where L stands for the fact that this matrix is describing a Local coordinate system at point $\mathbf{x}$, 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 matrix, but J looks too much like "an integer" or angular momentum or "the Jacobian". T for Transformation might have been confused with the full transformation F , or Cauchy stress T . Our chosen notation R makes one think perhaps R is a Rotation, but that won't in general be the case. For the moment we will continue to use R and S , where recall $RS = 1$. We shall refer to R simply as "the R matrix for transformation $\mathbf{F}$ ".

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 objects of direct product algebras). Of course the differential calculus aspect of the subject is already highly visible, there are ∂ symbols everywhere (hence the name **tensor calculus**).

(k) Definition of the words "scalar", "vector" and "tensor"

In section (a) a "scalar" was defined as something that is invariant under some transformation F , and this was identified with a "rank-0 tensor". Similarly, a "vector" is either a contravariant vector or a covariant vector and both of these are "rank-1 tensors". In Section 5 (e) certain "rank-2" tensors will appear -- they are matrices that transform in a certain way under a transformation F . In Section 7 (j) tensors of rank- n will appear, and these are objects with n indices which transform in a certain manner under F .

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

"tensorial scalar" $\equiv$ rank-0 tensor with respect to some transformation F

"tensorial vector" $\equiv$ rank-1 tensor with respect to some transformation F

"tensorial tensor" $\equiv$ rank- n tensor with respect to some transformation F

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

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

- A "scalar" is a single expression, a 1-tuple. No invariance under any transformation is implied.
- A "vector" is an N -tuple of expressions. No transformation rule is implied.
- A "second order tensor" is a matrix of expressions. No transformation rule is implied.
- A "tensor" is an object with n indices, $n = 2, 3, 4, \dots$ which includes the previous item. A tensor is therefore a collection of expressions which are labeled by n indices each of which goes 1 to N . No transformation rule is implied.

To these definitions we can add another list:

- A "scalar field" is a single function of $\mathbf{x}$ (the $\mathbf{x}$ -space coordinates). No implication of invariance.
- A "vector field" is an N-tuple of functions of $\mathbf{x}$ -- an N-tuple of scalar fields. No transform implied.
- A "tensor field" of order n is a set of N^n scalar functions, for example, $T_{abc} \dots (\mathbf{x})$. Same.

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

In Section 7 (u) we shall discuss the notion of an equation being "covariant", which means it has the exact same form in different frames of reference which are related by a transformation. For example, one might have $\mathbf{F} = m\mathbf{a}$ in frame S, and $\mathbf{F}' = m'\mathbf{a}'$ in frame S', where these frames are related by a static rotation. $\mathbf{F}$ and $\mathbf{a}$ are tensorial vectors with respect to this rotation, and m and m' are tensorial scalars, and $m = m'$ for that reason. Both sides of $\mathbf{F} = m\mathbf{a}$ transform as tensorial vectors. Since rotations are an invariance of Newtonian physics, any valid equation of motion *must* be "covariant", and this applies of course to particle, rigid body and continuum mechanics.

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

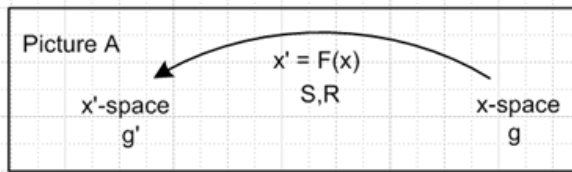
In this document we shall follow the time-honored tradition of being inconsistent in our use of the words scalar, vector and tensor, but the reader is now at least warned.

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

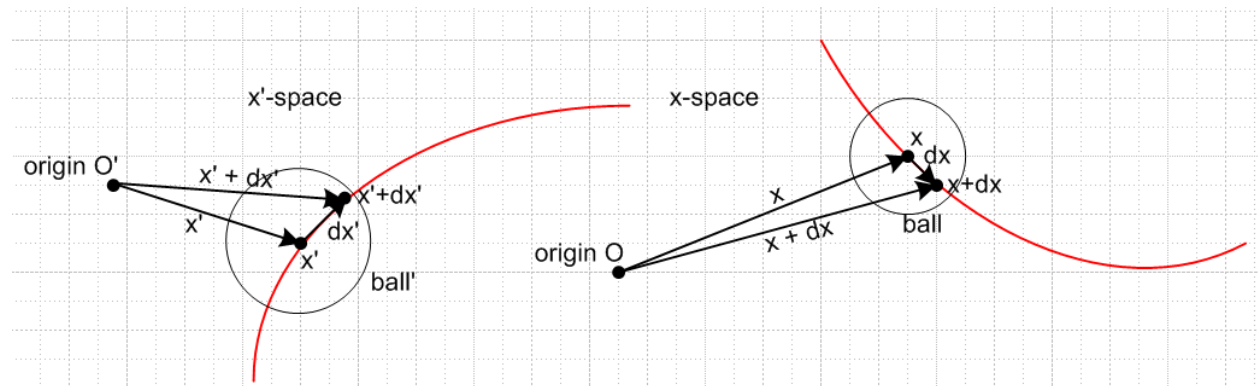
Appendix K explores a few commonly used tensors in continuum mechanics and determines which of these tensors actually transform as tensors (are objective), and which tensors do not transform as tensors (are non-objective).

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

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



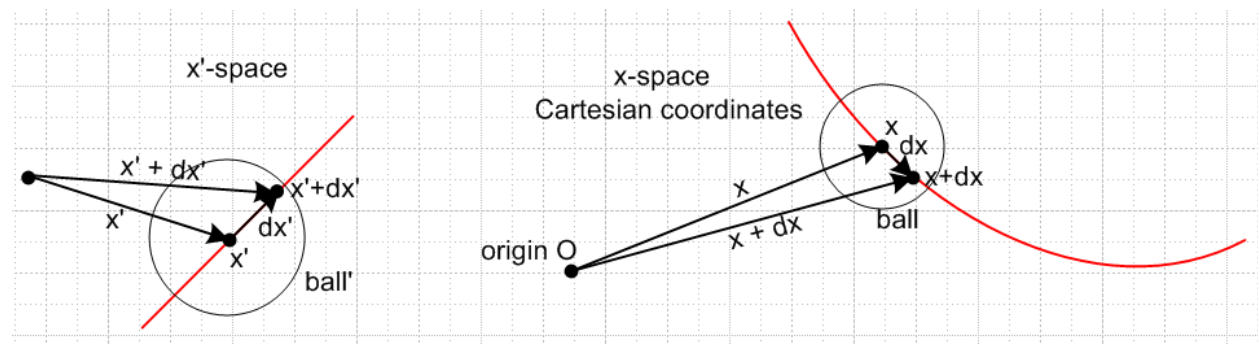
In the previous picture showing dx and dx' , one has much freedom to "try out" different differential vectors. For any dx one picks at point x , one gets some dx' according to $dx' = R(x) dx$. Consider this slightly enhanced version of the previous drawing (red curves added)



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

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

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

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

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

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

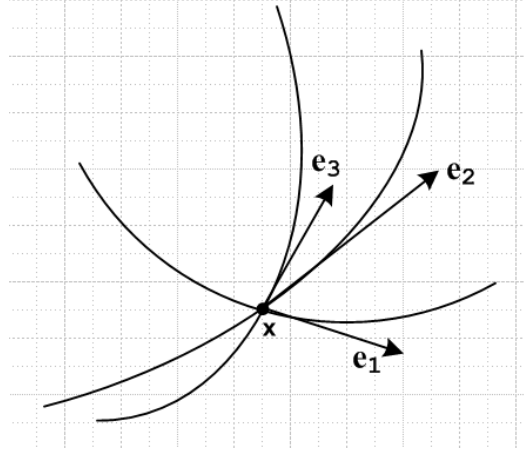
If the points on the x'_n -coordinate line were labeled with the values of x'_n from which they came, one would find that $\mathbf{e}_n$ points in the direction in which those labels increase.

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

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

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 $\mathbf{x}$:



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

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

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

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

The situation described above was this,

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

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

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

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

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

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:

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

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

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

So

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

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

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

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

$$S_{i\mathbf{k}} \equiv (\partial x_{\mathbf{i}} / \partial x'_{\mathbf{k}})$$

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

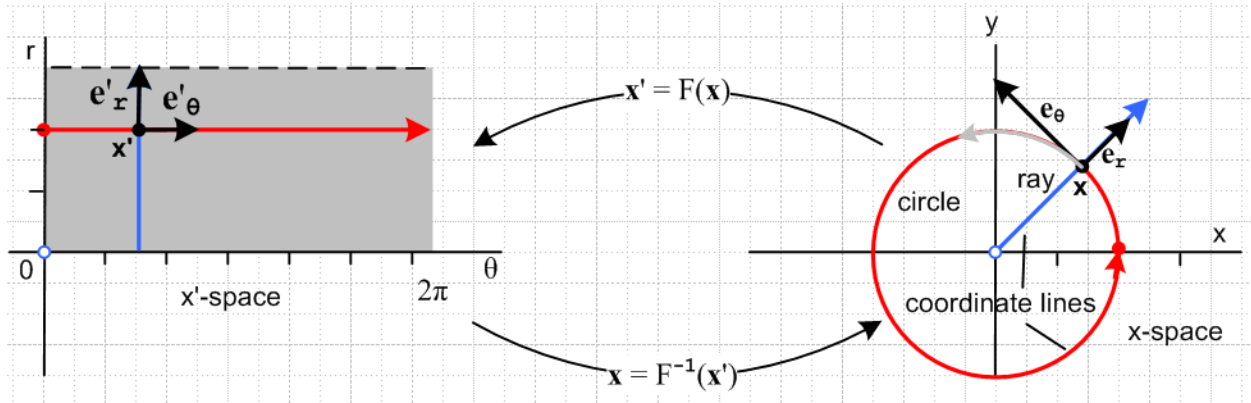
[Note: The above S and R are stated for the ordering 1,2 = θ, r . For the more usual ordering 1,2 = r, θ the columns of S should be swapped, and the rows of R should be swapped.]

The tangent base vectors $\mathbf{e}_{\mathbf{n}}$ 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 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 $(v_x^2 + v_y^2)^{1/2}$, 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 x' -space on the left map into the two $\mathbf{e}_n$ vectors on the right according to $\mathbf{e}_n \equiv S\mathbf{e}'_n$. If one were to apply the *full* mapping $\mathbf{x} = F^{-1}(\mathbf{x}')$ to each point along the arrows $\mathbf{e}'_n$, for some general non-linear F one would find that these arrows map into warped arrows on the right whose bases are tangent to those of the $\mathbf{e}_n$. Those warped arrows lie on the coordinate lines. For this particular mapping, $\mathbf{e}'_\theta$ maps under F^{-1} into the warped gray arrow, while $\mathbf{e}'_r$ maps into $\mathbf{e}_r$.

Example 2: Spherical Coordinates, tangent base vectors

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

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

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

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

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

$$S_{13} = (\partial x / \partial \varphi) = -r \sin \theta \sin \varphi$$

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

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

$$S_{23} = (\partial y / \partial \varphi) = r \sin \theta \cos \varphi$$

$$S_{31} = (\partial z / \partial r) = \cos \theta$$

$$S_{32} = (\partial z / \partial \theta) = -r \sin \theta$$

$$S_{33} = (\partial z / \partial \varphi) = 0$$

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

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

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

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

$$\mathbf{e}_r = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta)$$

$$\mathbf{e}_\theta = r(\cos \theta \cos \varphi, \cos \theta \sin \varphi, -\sin \theta)$$

$$\mathbf{e}_\varphi = r \sin \theta (-\sin \varphi, \cos \varphi, 0)$$

$$|\mathbf{e}_r| = 1 = h'_r$$

$$|\mathbf{e}_\theta| = r = h'_\theta$$

$$|\mathbf{e}_\varphi| = r \sin \theta = h'_\varphi$$

and unit vector versions are then

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

$$\hat{\mathbf{e}}_\theta = (\cos \theta \cos \varphi, \cos \theta \sin \varphi, -\sin \theta) = \hat{\boldsymbol{\theta}}$$

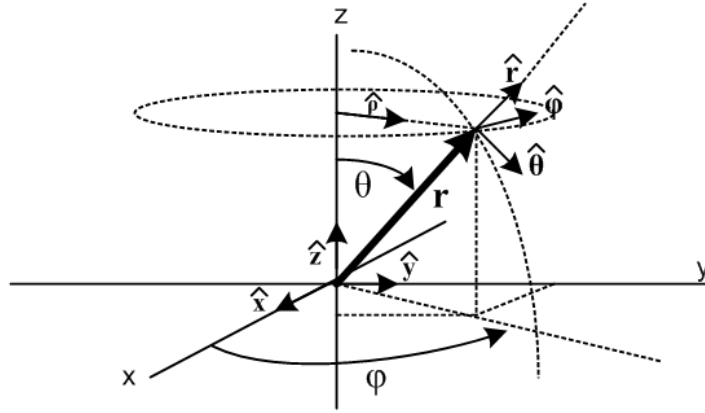
$$\hat{\mathbf{e}}_\varphi = (-\sin \varphi, \cos \varphi, 0) = \hat{\boldsymbol{\phi}}$$

$$\mathbf{e}_r = \hat{\mathbf{r}}$$

$$\mathbf{e}_\theta = r \hat{\boldsymbol{\theta}}$$

$$\mathbf{e}_\varphi = r \sin \theta \hat{\boldsymbol{\phi}}$$

The unit vectors can be displayed in this standard picture,



Notice that $(\hat{r}, \hat{\theta}, \hat{\phi}) = (\hat{e}_1, \hat{e}_2, \hat{e}_3)$ 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 $R \leftrightarrow S$ in the above discussion. The tangent base vectors for the inverse mapping would then be the columns of matrix R instead of S . We shall denote these inverse tangent base vectors which exist in $\mathbf{x}'$ -space by the symbol $\mathbf{u}'_n$. Then:

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

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

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

The same thing happens here, only in reverse :

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

where now the $\mathbf{u}_n$ are axis-aligned basis vectors in 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,

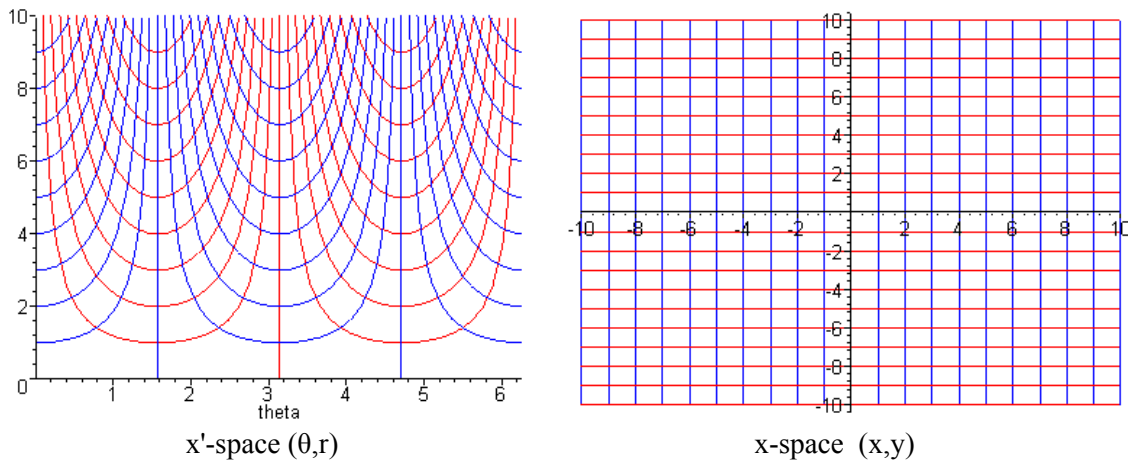
$$R = S^{-1} = \begin{pmatrix} -\sin\theta/r & \cos\theta/r \\ \cos(\theta) & \sin\theta \end{pmatrix}$$

so the inverse tangent base vectors are given by the columns of 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_i) in x'-space has the form $r = y_i/\sin\theta$, where y_i denotes some selected y value (a red horizontal line), so plotting $r = y_i/\sin\theta$ in x'-space for various values of y_i displays a set of inverse x-coordinate lines (red). Similarly $r = x_i/\cos\theta$ gives some y-coordinate lines (blue). Here is a Maple plot:



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 $d\mathbf{x}_i$ above were just little vectors and $\mathbf{x} + d\mathbf{x}$ was vector addition.

Now, for the first time (officially), we discuss length and distance, such as they are in a Cartesian Space, as defined in 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, a combo known as a Banach 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

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

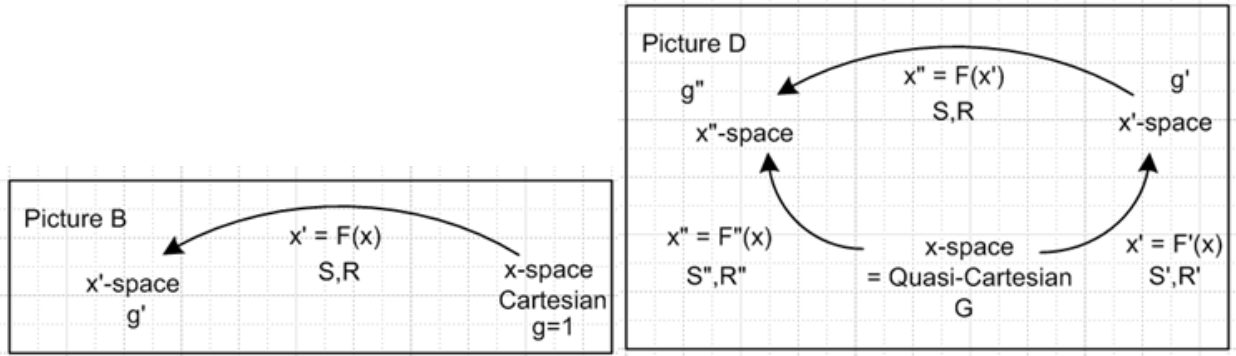
As shown in the next section, one can interpret $\delta_{i,j}$ as the metric tensor in Cartesian Space.

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 $d\mathbf{x} = S'(\mathbf{x}') d\mathbf{x}'$,

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

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

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

one then has, with implied summation on the right,

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

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

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

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 $\bar{g}'_{km} dx'_k dx'_m$ 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 :

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

Concerning the invariance of (ds) . In the above discussion, it was assumed that distance $(ds)^2$ is the same in x' -space as it is in x -space. As will be seen soon, this is equivalent to saying that the covariant dot product of any two vectors gives the same number regardless of which space is used to compute the dot product: $\mathbf{A} \bullet \mathbf{B} = \mathbf{A}' \bullet \mathbf{B}'$. This in turn implies that $|\mathbf{A}| = |\mathbf{A}'|$. In other words, it was assumed above that the dot product of two tensorial vectors is a scalar with respect to the underlying transformation F .

In our major application, where x -space is Cartesian and x' -space is that of some curvilinear coordinates, it is a *requirement* that $|\mathbf{A}| = |\mathbf{A}'|$. The length of a physical vector is the same no matter how one chooses to describe that vector. Imagine that $\mathbf{A}$ is a velocity vector $\mathbf{v}$. The speed $|\mathbf{v}|$ of an object is the same number whether one represents $\mathbf{v}$ in Cartesian or spherical coordinates.

In special relativity one again wants dot products to be scalars and the notion that $(ds)^2$ is a scalar under Lorentz transformations (that is, $\mathbf{dx} \bullet \mathbf{dx} = \mathbf{dx}' \bullet \mathbf{dx}'$) is a key assumption/requirement of the theory.

There are, however, applications of transformations where the scalarity of $(ds)^2$ is not required and in fact it is crucial that $(ds)^2$ can change under a transformation. For example, in continuum mechanics one can consider x -space to be a space describing a flow of continuous matter at some initial time t_0 and x' -space to be the same flow at a later time t . A general flow has $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ where $\mathbf{x}$ is the position of a continuum "particle" at time t_0 and $\mathbf{x}'$ is the position of that same particle at time t . In general $\mathbf{F}$ is non-linear. The distance between two differentially spaced particles at the two times is $\mathbf{dx}$ and $\mathbf{dx}'$, and one has $\mathbf{dx}' = \mathbf{R} \mathbf{dx}$ where $\mathbf{R}$ is not simply a rotation. The whole point here is that during the flow, the distance vector between two close particles rotates and stretches in some manner, and in general (due to this stretch) $|\mathbf{dx}| \neq |\mathbf{dx}'|$, so $(ds)^2$ is definitely not invariant under the flow (ie, under the transformation $\mathbf{F}$).

In Lai (p 105), the idea $\mathbf{dx}' = \mathbf{R} \mathbf{dx}$ translates to $\mathbf{dx} = \mathbf{F} d\mathbf{X}$, and $\mathbf{F}$ is called the deformation gradient and is written $\mathbf{F} = (\nabla \mathbf{x})$ which is a dyadic like notation discussed in Appendices E and G. In Section (i) below we show that in this flow situation there are in fact two distinct metric tensors involved, each having its own role to play. The continuum mechanics application is discussed further in Section (o) below, and the deformation tensor is considered in some detail in Appendix K.

Except when dealing with continuous mechanics flows, we shall treat $(ds)^2 = (ds')^2$ as a scalar.

(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:

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

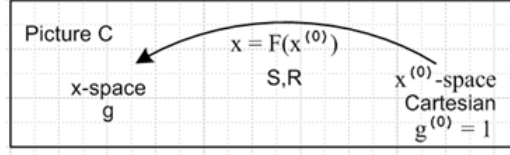
$$g' \equiv \overline{g'}^{-1} = (\mathbf{S}^T G \mathbf{S})^{-1} = \mathbf{S}'^{-1} G (\mathbf{S}'^T)^{-1} = \mathbf{R}' G \mathbf{R}'^T$$

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

Here are the collected facts from above:

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

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,



$$g = RGR^T$$

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

$$\bar{g}g = 1$$

$$(ds)^2 = \bar{g}_{km} dx_k dx_m$$

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).

(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:

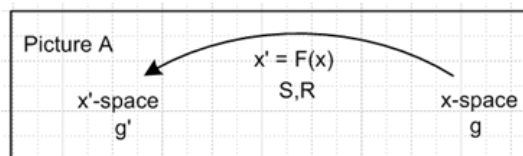
$$(A^TA)_{aa} = \sum_b (A^T)_{ab}A_{ba} = \sum_b (A)_{ba}A_{ba} = \sum_b (A_{ba})^2 \geq 0 \quad // \text{ diagonal elements } \geq 0$$

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

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

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

We now switch to **Picture A**,



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

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

which can be written out in components

$$\begin{aligned} V'_{\mathbf{a}} &= \mathbf{R}_{\mathbf{a}\mathbf{a}'} V_{\mathbf{a}'} & \text{contravariant} & & \mathbf{R}_{\mathbf{i}\mathbf{k}}(\mathbf{x}) &\equiv (\partial x'_{\mathbf{i}} / \partial x_{\mathbf{k}}) & & \mathbf{R} = \mathbf{S}^{-1} \\ \bar{V}'_{\mathbf{a}} &= \mathbf{S}^T_{\mathbf{a}\mathbf{a}'} \bar{V}_{\mathbf{a}'} & \text{covariant} & & \mathbf{S}_{\mathbf{i}\mathbf{k}}(\mathbf{x}') &\equiv (\partial x_{\mathbf{i}} / \partial x'_{\mathbf{k}}) &= \mathbf{S}^T_{\mathbf{k}\mathbf{i}}(\mathbf{x}') \end{aligned}$$

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:

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

and again we put a bar over the covariant objects.

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

- $\det(\mathbf{A}) = \det(\mathbf{A}^T)$
- $\text{cof}(\mathbf{A}^T) = [\text{cof}(\mathbf{A})]^T$ since $[\text{cof}(\mathbf{A}^T)]_{\mathbf{a}\mathbf{b}} = \text{cof}(\mathbf{A}^T_{\mathbf{a}\mathbf{b}}) = \text{cof}(\mathbf{A}_{\mathbf{b}\mathbf{a}}) = [\text{cof}(\mathbf{A})]_{\mathbf{b}\mathbf{a}} = [\text{cof}(\mathbf{A})]^T_{\mathbf{a}\mathbf{b}}$
- $(\mathbf{A}^{-1})^T = \{ [\text{cof}(\mathbf{A})]^T / \det(\mathbf{A}) \}^T = [\text{cof}(\mathbf{A}^T)]^T / \det(\mathbf{A}^T) = (\mathbf{A}^T)^{-1}$

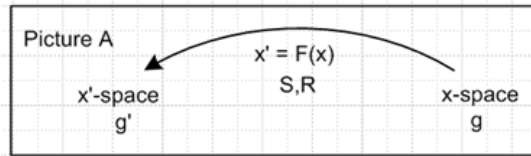
This fact is used many times in the manipulations below.

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

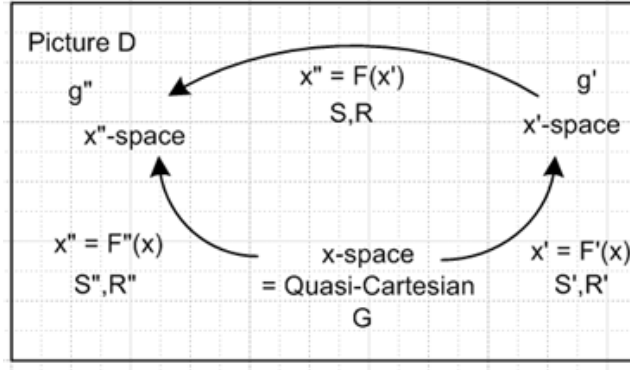
The above rank-2 tensor transformation rules can be written in the following matrix form (something not possible with higher-rank tensors),

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

where recall



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



$$M'' = R M' R^T \quad // \text{ contravariant rank-2 tensor}$$

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

Consider then this sequence of steps:

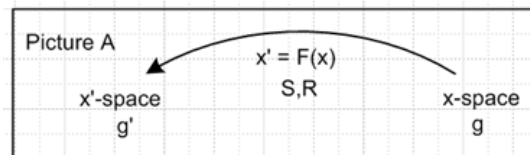
$$\begin{aligned}
 1 * G * 1 &= 1 * G * 1 \\
 (S''R'') G (S''R'')^T &= (S'R') G (S'R')^T && // S''R'' = 1 \\
 S'' (R'G R'^T) S''^T &= S'(R'G R'^T) S'^T && // \text{ regroup} \\
 S'' g'' S''^T &= S' g' S'^T && // \text{ since } g'' = R''G R''^T \text{ and } g' = R'G R'^T \\
 g'' S''^T &= R'' S' g' S'^T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

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,

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

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



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

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

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

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

$$\begin{aligned} Rg &= g' S^T & \bar{g}' R g &= S^T & g' S^T \bar{g} &= R \\ \bar{g} S &= R^T \bar{g}' & g R^T \bar{g}' &= S & \bar{g} S g' &= R^T \end{aligned}$$

(g) Metric tensor converts vector types

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

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

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

$$\mathbf{W}' = \bar{g}' \mathbf{V}' = (S^T \bar{g} S) (R\mathbf{V}) = S^T \bar{g} (SR)\mathbf{V} = S^T \bar{g} \mathbf{V} = S^T \mathbf{W}$$

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

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

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 $\mathbf{A} \bullet \mathbf{A}$ ($A_\mu A^\mu$ 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 $A.B$ or $A \cdot B$. Typical usage then in standard notation would be $p \cdot p = p^\mu p_\mu = p_0^2 - \mathbf{p} \cdot \mathbf{p}$ (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 $\bar{\mathbf{B}}_{\mathbf{a}} = \bar{g}_{\mathbf{a}\mathbf{b}}B_{\mathbf{b}}$. Using this partner $\bar{\mathbf{B}}$ one can restate the above covariant scalar product as

$$\mathbf{A} \bullet \mathbf{B} = \bar{g}_{\mathbf{a}\mathbf{b}}A_{\mathbf{a}}B_{\mathbf{b}} = A_{\mathbf{a}} \bar{\mathbf{B}}_{\mathbf{a}}$$

or, taking instead $\bar{\mathbf{A}}_{\mathbf{b}} = \bar{g}_{\mathbf{b}\mathbf{a}}A_{\mathbf{a}}$,

$$\mathbf{A} \bullet \mathbf{B} = \bar{\mathbf{A}}_{\mathbf{b}} B_{\mathbf{b}} = \bar{\mathbf{A}}_{\mathbf{a}} B_{\mathbf{a}}$$

And finally, if in $\mathbf{A} \bullet \mathbf{B} = A_{\mathbf{a}} \bar{\mathbf{B}}_{\mathbf{a}}$ we write $A_{\mathbf{a}} = g_{\mathbf{a}\mathbf{b}}\bar{\mathbf{A}}_{\mathbf{b}}$, we get

$$\mathbf{A} \bullet \mathbf{B} = A_{\mathbf{a}} \bar{\mathbf{B}}_{\mathbf{a}} = g_{\mathbf{a}\mathbf{b}}\bar{\mathbf{A}}_{\mathbf{b}}\bar{\mathbf{B}}_{\mathbf{a}} = g_{\mathbf{a}\mathbf{b}}\bar{\mathbf{A}}_{\mathbf{a}}\bar{\mathbf{B}}_{\mathbf{b}}$$

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 :

$$\mathbf{A} \bullet \mathbf{B} = \bar{g}_{\mathbf{a}\mathbf{b}}A_{\mathbf{a}}B_{\mathbf{b}} = A_{\mathbf{a}} \bar{\mathbf{B}}_{\mathbf{a}} = \bar{\mathbf{A}}_{\mathbf{a}} B_{\mathbf{a}} = g_{\mathbf{a}\mathbf{b}}\bar{\mathbf{A}}_{\mathbf{a}}\bar{\mathbf{B}}_{\mathbf{b}} = \mathbf{B} \bullet \mathbf{A}$$

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)

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

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

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

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

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

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

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 4-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 after temporarily using it above for a 3-space Cartesian dot product.]

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

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

In applications in which $(ds)^2$ is NOT regarded as a scalar with respect to transformation F, which includes continuum mechanics flows, things are a little different. We have here

$$(ds)^2 \equiv d\mathbf{x} \bullet d\mathbf{x} = d\mathbf{x}' \bullet d\mathbf{x}' \neq (ds')^2 \quad \text{if } (ds') \equiv \text{physical distance in } x'\text{-space}$$

If x' -space has Cartesian coordinates, we have $(ds')^2 = dx'_i dx'_i = \delta_{i,j} dx'_i dx'_j$ and this is different from the quantity $d\mathbf{x}' \bullet d\mathbf{x}' = \bar{g}'_{ij} dx'_i dx'_j$. One might go so far as to define two different $\bullet$ symbols:

$$d\mathbf{x}' \bullet_g d\mathbf{x}' \equiv \bar{g}'_{ij} dx'_i dx'_j \neq (ds')^2$$

$$d\mathbf{x}' \bullet_1 d\mathbf{x}' \equiv \delta_{i,j} dx'_i dx'_j = dx'_i dx'_i = (ds')^2 = \text{physical distance squared}$$

By default, we have $\bullet = \bullet_g$, so then $d\mathbf{x}' \bullet d\mathbf{x}' \neq (ds')^2$.

In this situation, x' -space in effect has *two different* "metric tensors". The first is the $\bar{g}'_{km}$ that appears in the entire discussion of this Section 5 and which is involved in raising and lowering indices on a tensor and satisfies $\bar{g}' = S^T \bar{g} S$ (the fact that g transforms as a rank-2 tensor), and so on. The second we might call $\bar{g}'_{km}$ and this is what determines physical distance in x' -space, $(ds')^2 = \bar{g}'_{km} dx'_k dx'_m$. In continuum mechanics with Cartesian x' -space, we then have $\bar{g}'_{km} = \delta_{k,m}$.

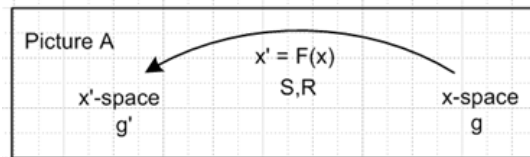
There are two different "metric tensors" because there are two different "metrics" which are of interest in x' -space :

$$\begin{aligned} d_g^2(\mathbf{x}, \mathbf{y}) &= (\mathbf{x} - \mathbf{y}) \bullet_g (\mathbf{x} - \mathbf{y}) = \bar{g}_{km} (x - y)_k (x - y)_m & // \text{raising/lowering metric tensor} \\ d_{\bar{g}}^2(\mathbf{x}, \mathbf{y}) &= (\mathbf{x} - \mathbf{y}) \bullet_{\bar{g}} (\mathbf{x} - \mathbf{y}) = \bar{g}_{km} (x - y)_k (x - y)_m & // \text{physical distance metric tensor} \end{aligned}$$

For Cartesian coordinates in x' -space, $\bar{g}'_{km} = \delta_{k,m}$, and our two metric tensors may be identified with those of "Cartesian View x' -space" and "Curvilinear View x' -space" as discussed in Section 8 and Appendix C (e).

(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 (f))

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

so

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

since,

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

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

$$\bar{g}'_{mn} = \mathbf{e}_m \bullet \mathbf{e}_n \quad \text{or} \quad \bar{g}'_{mn} = \partial'_m \mathbf{x} \bullet \partial'_n \mathbf{x}$$

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

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

Therefore, the tangent base vectors will only be mutually orthogonal when the metric tensor $\bar{g}'$ 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 $\bar{g}'$,

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

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

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

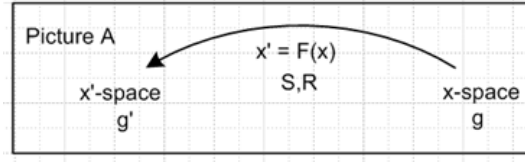
As a reminder, had we called x' -space something like ξ -space, there would be no primes on these symbols, but then if $\bar{g} \neq 1$ there would be confusion as to which space the symbols applied.

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.

Note: Some authors refer to the scale factors h'_n as the **Lamé coefficients**, while other authors refer to R_{ij} as the Lamé coefficients which they call h_{ji} . (Lame)

(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,

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

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 $S_{ij}(\mathbf{x}') = \partial x_i / \partial x'_k$, although a matrix, is not a rank-2 tensor. Similarly, $J(\mathbf{x}')$, while a "scalar" function, is not a rank-0 tensorial scalar. One does not ask how S and J themselves "transform" in going from x -space to x' -space.

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

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

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

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 ($\varepsilon_{abc\dots}$ is the permutation tensor, Section 7 (h)),

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

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

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

so J is related to the tangent base vectors by

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

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

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

or

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

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

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

So that

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

Normally the argument dependence is suppressed and one then writes

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

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,

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

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

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

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

Since transformation F is assumed invertible in its domain and range, one cannot have $\det(S) = 0$ anywhere except perhaps on a boundary. Since $\det(\bar{g}') = [\det(S)]^2 \det(\bar{g})$, if we assume $\det(\bar{g})$ vanishes

nowhere in the x -space domain of F , then $\det(\bar{g}') \neq 0$ everywhere in the range of F . The conclusion with this assumption is that the signature s is always well-defined.

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

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

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

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

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

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

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

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

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

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

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

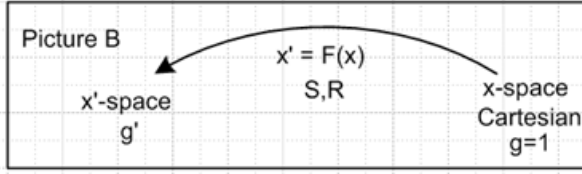
$$|J| = \sqrt{sg'} / \sqrt{sg} = |\det(S)| = \sqrt{g'/g}$$

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 Königsberg, 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 Pictures B and C (Cartesian x -space).

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



the statement of the rank-2 tensor transformation of g' and $\bar{g}'$ becomes (from section (f) above)

$$g' = RR^T$$

$$\bar{g}' = S^T S$$

which can be written in a variety of ways,

$$\begin{aligned} R^T &= (SR)R^T = S(RR^T) = S g' \Rightarrow R = g' S^T \Rightarrow 1 = S g' S^T \\ S^T &= S^T(R^T S^T) = (S^T S)R = \bar{g}' R \Rightarrow S = R^T \bar{g}' \Rightarrow 1 = R^T \bar{g}' R. \end{aligned}$$

In summary:

$$\begin{aligned} g' &= RR^T & R^T &= S g' & R &= g' S^T & 1 &= S g' S^T \\ \bar{g}' &= S^T S & S^T &= \bar{g}' R & S &= R^T \bar{g}' & 1 &= R^T \bar{g}' R. \end{aligned}$$

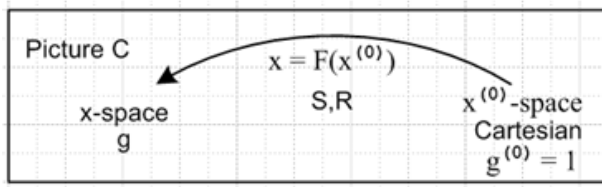
The diagonal elements of $\bar{g}'$ and g' are given by

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

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)

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

In Picture C (sometimes $x^{(0)}$ -space is called ξ -space and $g^{(0)}$ is called η)



the above equations are identical except all the g 's are now unprimed, so

$$\mathbf{g} = \mathbf{R}\mathbf{R}^T$$

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

$$\begin{array}{llll} \mathbf{g} = \mathbf{R}\mathbf{R}^T & \mathbf{R}^T = \mathbf{S} \mathbf{g} & \mathbf{R} = \mathbf{g} \mathbf{S}^T & 1 = \mathbf{S} \mathbf{g} \mathbf{S}^T \\ \bar{\mathbf{g}} = \mathbf{S}^T \mathbf{S} & \mathbf{S}^T = \bar{\mathbf{g}} \mathbf{R} & \mathbf{S} = \mathbf{R}^T \bar{\mathbf{g}} & 1 = \mathbf{R}^T \bar{\mathbf{g}} \mathbf{R} \end{array}$$

$$\bar{g}_{nn} = \sum_n \mathbf{S}_{ni}^T \mathbf{S}_{in} = \sum_n (\mathbf{S}_{in})^2 = \sum_i (\partial \xi_i / \partial x_n)^2$$

$$g_{nn} = \sum_n \mathbf{R}_{ni} \mathbf{R}_{in}^T = \sum_n (\mathbf{R}_{ni})^2 = \sum_i (\partial x_n / \partial \xi_i)^2$$

$$h_n^2 = \bar{g}_{nn} = \sum_i (\partial \xi_i / \partial x_n)^2 \quad h_n^{-2} = g_{nn} = \sum_i (\partial x_n / \partial \xi_i)^2$$

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

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

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

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

Another way is this:

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

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 $d\mathbf{x}$ 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

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

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 $R_{\hat{i}\hat{j}}$ is written $R^{\hat{i}}_{\hat{j}}$, contravariant vectors $V_{\hat{i}}$ are written $V^{\hat{i}}$, and covariant vectors $\bar{V}_{\hat{j}}$ are written $V_{\hat{j}}$. 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 a theory requirement is that invariant length be preserved $x' \cdot x' = x \cdot x = \text{scalar}$. Special relativity also requires that the metric tensor G be the same in all frames, since no frame is special, so $G' = G$. But this says, in our old notation, that $R G R^T = G$. 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,

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

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,

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

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

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

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


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

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

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

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

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

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

```
simplify(%);
```

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

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

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

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

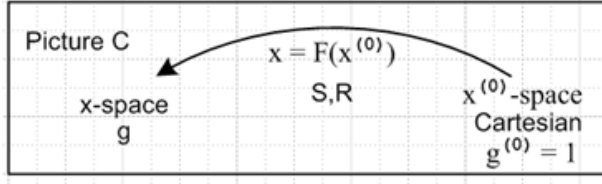
$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -\cos(r)^2 - \sin(r)^2 & 0 \\ 0 & 0 & 0 & -\cos(r)^2 - \sin(r)^2 \end{bmatrix}$$

```
simplify(%);
```

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

(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

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

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

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

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

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

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.

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

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 Appendix F (f).

(o) Continuum Mechanics and its Metric Tensors

One can describe (Lai) the forward "flow" of a continuous blob of matter by $\mathbf{x} = \mathbf{x}(\mathbf{X}, t)$ where $\mathbf{X} = \mathbf{x}(\mathbf{X}, t_0)$. A "particle" of matter (imagine a tiny cube) that starts at location $\mathbf{X}$ at time t_0 ends up at $\mathbf{x}$ at time t . Two points in the flow 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. $\mathbf{F}$ describes how a particle starting say with a cubic shape at t_0 gets deformed into some parallelepiped (3-piped) shape at t (picture below). If we examine $d\mathbf{x} = \mathbf{F} d\mathbf{X}$ we find that $|d\mathbf{x}| \neq |d\mathbf{X}|$ since the vector $d\mathbf{X}$ typically gets rotated and stretched as $d\mathbf{X} \rightarrow d\mathbf{x}$ during the flow. [This flow is further described in Appendix K.]

The finite-time flow $\mathbf{x} = \mathbf{x}(\mathbf{X}, t)$ from time t_0 to time t can be thought of as a (generally non-linear) transformation of the form $\mathbf{x} = \mathcal{F}(\mathbf{X})$ as in Section 1 above. Recall from Section 1 that a general transformation was $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ and the linearized transformation was $d\mathbf{x}' = \mathbf{R} d\mathbf{x}$. To be compatible with Lai notation which uses symbol $\mathbf{F}$ for the deformation gradient, we have renamed the Section 1 transformation $\mathbf{F}$ to be $\mathcal{F}$, and we shall see below that $\mathbf{R}$ will in fact become deformation gradient $\mathbf{F}$.

In this flow we assume Cartesian coordinates in both $\mathbf{x}$ -space and $\mathbf{X}$ -space, so the metric tensors which determine physical distance in these spaces are both 1.

As discussed at the end of Section 5 (i), there are really *two* metric tensors in this situation called $\bar{g}'_{km}$ and $\bar{g}_{km}$. The former is the metric tensor which raises and lowers tensor indices in the general tensor formalism, while the latter is the metric tensor which determines physical distance in $\mathbf{x}'$ -space. For the usual Cartesian coordinates in $\mathbf{x}'$ -space, $\bar{g}'_{km} = \delta_{k,m}$ which is to say $\bar{g}' = 1$. The other metric tensor is given by $\bar{g}' = \mathbf{S}^T \bar{g} \mathbf{S}$ where $\bar{g}$ is the metric tensor in $\mathbf{x}$ -space. Using Cartesian coordinates there means $\bar{g} = 1$ and then $\bar{g}' = \mathbf{S}^T \mathbf{S}$.

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>continuum mechanics</u>		<u>this document</u>	(Forward Flow $\mathbf{X} \rightarrow \mathbf{x}$ [$\mathbf{x} \rightarrow \mathbf{x}'$])
$\mathbf{x}, \mathbf{X}$	$\leftrightarrow$	$\mathbf{x}', \mathbf{x}$	
$\mathbf{x} = \mathbf{x}(\mathbf{X}, t) = \mathcal{F}(\mathbf{X})$	$\leftrightarrow$	$\mathbf{x}' = \mathcal{F}(\mathbf{x})$	// Lai p70 (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 p86 (3.7.6), p105 (3.18.3)
$\mathbf{F}$	$\leftrightarrow$	$\mathbf{R}$	
$\mathbf{F}^{-1}$	$\leftrightarrow$	$\mathbf{S}$	
$\mathbf{X} = \text{Cartesian}$	$\leftrightarrow$	$\bar{g} = 1$	
$\mathbf{x} = \text{Cartesian}$	$\leftrightarrow$	$\bar{g}' = 1$	
$\mathbf{B} = \mathbf{F}\mathbf{F}^T$	$\leftrightarrow$	$\bar{g}' = \mathbf{R}\mathbf{R}^T$	// as in Section 5 (l) // Lai p121 (3.25.2)
$\mathbf{B}^{-1} = (\mathbf{F}^{-1})^T (\mathbf{F}^{-1})$	$\leftrightarrow$	$\bar{g}' = \mathbf{S}^T \mathbf{S}$	// since $\bar{g}' = \bar{g}^{-1}$

Thus, the deformation gradient $\mathbf{F}$ is just the $\mathbf{R}$ matrix of the forward transformation $\mathbf{x} = \mathbf{x}(\mathbf{X}, t) = \mathcal{F}(\mathbf{X})$. The raising/lowering metric tensor $\bar{g}' = \mathbf{R}\mathbf{R}^T$ appears as $\mathbf{B} = \mathbf{F}\mathbf{F}^T$ which is known as the *left* Cauchy-Green deformation tensor (manifestly symmetric, so a viable metric tensor).

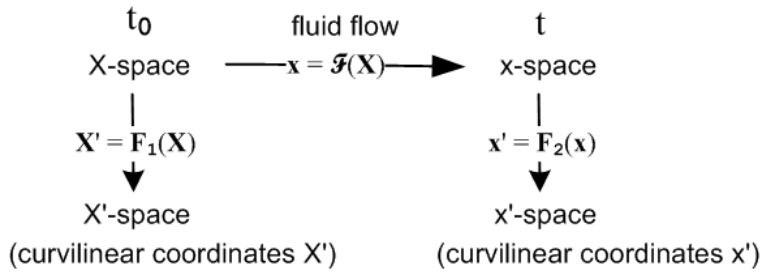
Alternately, we can consider the above flow going backwards in time and then $\mathbf{x} = \mathbf{x}$ is the starting position and $\mathbf{x}' = \mathbf{X}$ is the ending position. For this inverse flow, we let $\mathbf{F}$ have the same meaning as in the forward flow, $d\mathbf{x} = \mathbf{F} d\mathbf{X}$, and thus end up with this translation table where we now $\mathbf{x} \rightarrow \mathbf{x}$ and $\mathbf{X} \rightarrow \mathbf{x}'$:

<u>continuum mechanics</u>		<u>this document</u>	(Inverse Flow $\mathbf{x} \rightarrow \mathbf{X} [\mathbf{x} \rightarrow \mathbf{x}']$)
$\mathbf{X}, \mathbf{x}$	$\leftrightarrow$	$\mathbf{x}', \mathbf{x}$	
$\mathbf{X} = \mathbf{X}(\mathbf{x}, t) = \mathcal{F}(\mathbf{x})$	$\leftrightarrow$	$\mathbf{x}' = \mathcal{F}(\mathbf{x})$	// this $\mathcal{F}$ is the inverse of the forward flow $\mathcal{F}$
$d\mathbf{X} = \mathbf{F}^{-1} d\mathbf{x}$	$\leftrightarrow$	$d\mathbf{x}' = \mathbf{R} d\mathbf{x}$	// as in Section 2, same $\mathbf{F}$ as in forward flow
$\mathbf{F}^{-1}$	$\leftrightarrow$	$\mathbf{R}$	
$\mathbf{F}$	$\leftrightarrow$	$\mathbf{S}$	// $\mathbf{S} = \mathbf{R}^{-1}$ and $S_{ik} = (\partial x_i / \partial x'_k) \leftrightarrow F_{ik} = (\partial x_i / \partial X_k)$
$\mathbf{X} = \text{Cartesian}$	$\leftrightarrow$	$g = 1$	
$\mathbf{x} = \text{Cartesian}$	$\leftrightarrow$	$g' = 1$	
$\mathbf{C} = \mathbf{F}^T \mathbf{F}$	$\leftrightarrow$	$\bar{g}' = \mathbf{S}^T \mathbf{S}$	// as in Section 5 ($\mathcal{L}$) // Lai p114 (3.23.2)
$\mathbf{C}^{-1} = \mathbf{F}^{-1} (\mathbf{F}^{-1})^T$	$\leftrightarrow$	$g' = \mathbf{R} \mathbf{R}^T$	// since $g' = (\bar{g}')^{-1}$

To avoid confusion, in both tables g goes with flow $\mathbf{X}$ -space and g' and g' go with flow $\mathbf{x}$ -space.

For the inverse flow, the raising/lowering metric tensor $\bar{g}' = \mathbf{S}^T \mathbf{S}$ appears as $\mathbf{C} = \mathbf{F}^T \mathbf{F}$ which is the *right* Cauchy-Green deformation tensor (again manifestly symmetric, so a viable metric tensor).

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



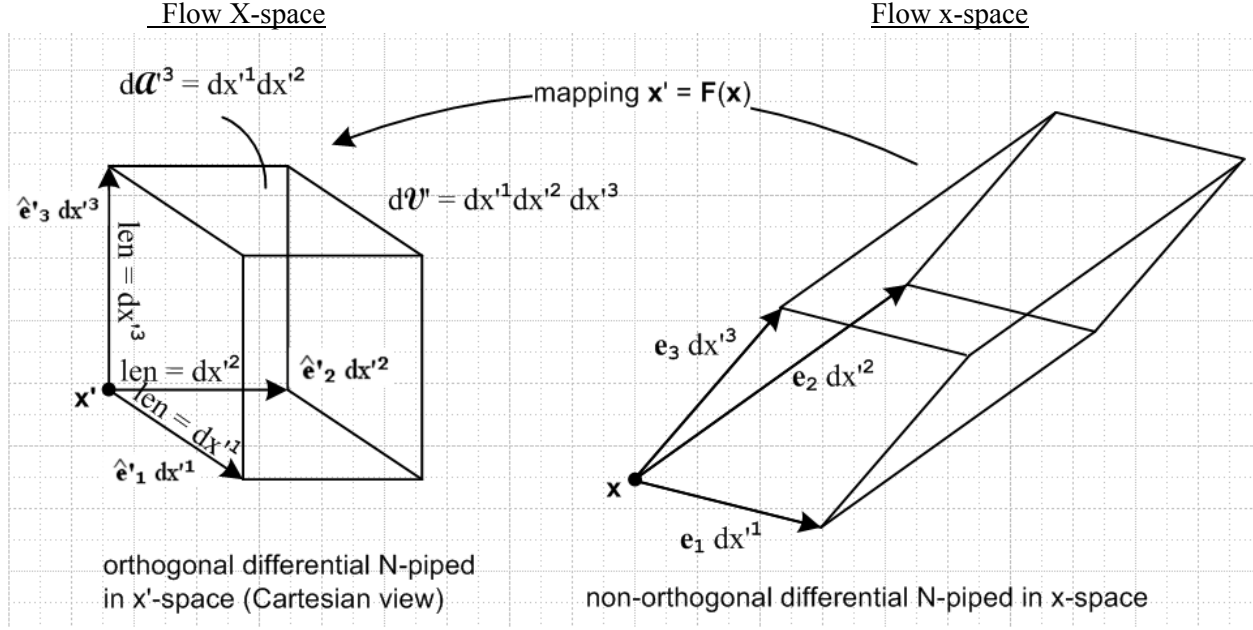
and we then have an interesting triple application of the notions of Section 1 to a real-world situation. This drawing is the implicit subject of Section 3.29 (p131) of Lai.

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

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

The $(\nabla \mathbf{x})$ notation is explained in Appendix E, and in Appendix G the object $(\nabla \mathbf{v})$ for an arbitrary vector field $\mathbf{v}(\mathbf{x})$ is expressed in general curvilinear coordinates.

Consider now this picture taken from Section 8 below,



We can identify the mapping shown in this picture with our Inverse Flow situation (table above). Section 8 discusses in much detail how length, volume and area transform under a general transformation. The length, area and volume magnitudes on the left are called $d\mathcal{L}^n = dx'^{(n)}$, $d\mathcal{A}^n$ and $d\mathcal{V}^n$, while the corresponding quantities on the right are called $dx^{(n)}$, $d\bar{\mathcal{A}}^{(n)}$ and dV , where the first two items are vectors. Cribbing the results of Section 8 (c) 7 and converting them from standard notation to developmental notation, we have

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

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

$$|dV| / d\mathcal{V}^n = |J| = g'^{1/2} \quad // \text{ where } g' \equiv \det(\bar{g}'_{ij}) = J^2, \quad \bar{g}' = S^T S$$

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

$$|dx^{(n)}| / |dX^{(n)}| = h'_n = [\bar{g}'_{nn}]^{1/2} = [(F^T F)_{nn}]^{1/2} = [C_{nn}]^{1/2} \quad // \text{ Lai p114 (3.23.6-8)}$$

$$|d\bar{\mathcal{A}}^n| / |d\mathcal{A}_0^n| = [g'_{nn} g']^{1/2} = [\text{cof}(\bar{g}'_{nn})]^{1/2} = [\text{cof}(F^T F)_{nn}]^{1/2} = [\text{cof } C_{nn}]^{1/2} \quad // \text{ Lai p129 (3.27.11)}$$

$$|dV| / |dV_0| = |J| = g'^{1/2} = [\det(\bar{g}'_{ij})]^{1/2} = [\det(F^T F)]^{1/2} = |\det(F)| \quad // \text{ Lai p 130 (3.28.3)}$$

where

	edge	area	volume	
X-space :	$dX^{(n)}$	$d\mathcal{A}_0^n$	dV_0	time t_0
x-space :	$dx^{(n)}$	$d\bar{\mathcal{A}}^n$	dV	time t

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

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

$$dA^{(n)} / dA_0^{(n)} = \det(F) | (F^{-1})^T \mathbf{u}_n | \quad \mathbf{u}_n = \text{unit base vector, } (\mathbf{u}_n)_i = \delta_{n,i}$$

which seems a far cry from our result $[\text{cof}(F^T F)_{nn}]^{1/2}$. But consider, using the Inverse Flow table,

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

so

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

where we use the Section 8 (c) 6 theorem converted to developmental notation, stating that

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

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

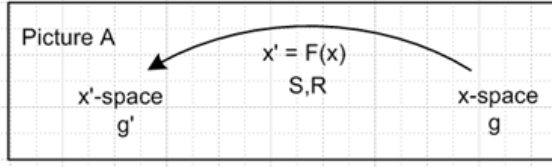
<u>Lai notation</u>	<u>Developmental notation</u>	<u>Standard Notation</u>
$d\mathbf{A}_0 = d\mathbf{X}^{(1)} \times d\mathbf{X}^{(2)} \quad (3.27.1)$	$d\bar{\mathbf{A}}_0 = d\mathbf{X}^{(1)} \times d\mathbf{X}^{(2)}$	$(dA_0)_i = \varepsilon_{ijk} [d\mathbf{X}^{(1)}]^j [d\mathbf{X}^{(2)}]^k$
$[\text{div}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

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

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

6. Reciprocal Base Vectors 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)

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

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

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

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

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

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

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

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

(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)

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

The second is

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

and the third is

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

To summarize,

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

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

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

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. Suppose the vectors $\mathbf{b}_n$ form a complete basis for x-space. Can one find a set of vectors $\mathbf{B}_n$ that has the property $\mathbf{B}_m \bullet \mathbf{b}_n = \delta_{m,n}$? As shown below, the answer is normally "yes", and the vectors $\mathbf{B}_n$ are uniquely determined by the $\mathbf{b}_n$. One says that the set $\{\mathbf{B}_n\}$ is "dual to" the set $\{\mathbf{b}_n\}$ and vice versa. If we regard $\mathbf{b}_n$ as a basis, then $\mathbf{B}_n$ is the "dual basis", and $\mathbf{B}_m \bullet \mathbf{b}_n = \delta_{m,n}$ is the "duality relation". Another terminology is that the vectors $\mathbf{B}_n$ are "reciprocal to" the vectors $\mathbf{b}_n$.

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

3. Solving for the $\mathbf{B}_n$ in terms of the $\mathbf{b}_n$. Each $\mathbf{B}_n$ has N components, so there are N^2 unknowns. The duality relation $\mathbf{B}_m \bullet \mathbf{b}_n = \delta_{m,n}$ is a set of N^2 equations. This is basically a Cramer's Rule problem in N^2 variables. Since the $\mathbf{b}_n$ form a complete basis, one can expand $\mathbf{B}_m = w_{mn}\mathbf{b}_n$. Then

$$\delta_{m,k} = \mathbf{B}_m \bullet \mathbf{b}_k = w_{mn} \mathbf{b}_n \bullet \mathbf{b}_k \quad // \quad \mathbf{b}_n \bullet \mathbf{b}_k = \bar{g}_{ij}(b_n)_i(b_k)_j$$

Define $W_{nk} \equiv \mathbf{b}_n \bullet \mathbf{b}_k$ and note that W_{nk} is symmetric. Then $\delta_{m,k} = w_{mn}W_{nk}$ or $wW = 1$. Assuming for the moment that $\det W \neq 0$, the solution is given by $w = W^{-1}$. The Section k (e) "Digression" showed that $(A^{-1})^T = (A^T)^{-1}$ for invertible A, so $w^T = (W^{-1})^T = (W^T)^{-1} = W^{-1} = w$ and therefore w is symmetric. Since w is known, the $\mathbf{B}_m = w_{mn}\mathbf{b}_n$ have been found. Finally,

$$\mathbf{B}_m \bullet \mathbf{B}_n = w_{mi}\mathbf{b}_i \bullet w_{nj}\mathbf{b}_j = w_{mi}w_{nj} \mathbf{b}_i \bullet \mathbf{b}_j = w_{mi}w_{nj}W_{ji} = w_{mi}\delta_{ni} = w_{nm}$$

4. In the case that $\mathbf{b}_n = \mathbf{e}_n$ and $\mathbf{B}_n = \mathbf{E}_n$, one finds that $W_{nk} = \mathbf{e}_n \bullet \mathbf{e}_k = \bar{g}'_{nk}$, the covariant metric tensor. Then $w = W^{-1}$ must be the contravariant metric tensor $w_{mn} = g'_{mn} = \mathbf{E}_m \bullet \mathbf{E}_n$. Then $\mathbf{B}_m = w_{mn}\mathbf{b}_n$ says $\mathbf{E}_n = g'_{mn} \mathbf{e}_n$ which agrees with our original definition of the $\mathbf{E}_n$. Moreover, $\det(W) = \det(\bar{g}'_{ab}) = g'$ of Section 5 (k), so as long as $g' \neq 0$, one has $\det W \neq 0$.

5. For the general $\mathbf{b}_n$ case, if one imagines that the $\mathbf{b}_n$ are the tangent base vectors for *some* transformation F_b with some R_b , then the issue of $\det W \neq 0$ boils down to $g' \neq 0$ for that transformation. In the case that x-space is Cartesian, we know that $J^2 = \det(\bar{g}')/\det(\bar{g}) = g'/g = g'$, and the assumed invertibility of transformation F means that $J \neq 0$ everywhere, so $g' \neq 0$ and then $\det W \neq 0$. In all cases considered in this document, $\det W \neq 0$.

6. If the $\mathbf{b}_m$ are true tensorial vectors, then the $\mathbf{B}_m = w_{mn}\mathbf{b}_n$ will be as well and then $\mathbf{B}_n \bullet \mathbf{b}_m$ is a tensorial scalar. Therefore if $\mathbf{B}_n \bullet \mathbf{b}_m = \delta_{n,m}$ in x-space, then so also $\mathbf{B}'_n \bullet \mathbf{b}'_m = \delta_{n,m}$ in x'-space, where $\mathbf{b}'_m = R\mathbf{b}_m$ and $\mathbf{B}'_n = R\mathbf{B}_n$. For example, $\mathbf{E}'_n \bullet \mathbf{e}'_m = \delta_{n,m}$ in x'-space where $\mathbf{e}'_m = R\mathbf{e}_m$ and $\mathbf{E}'_n = R\mathbf{E}_n$.

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

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

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

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

9. In standard notation (Section 7) the equations relating to $\mathbf{b}_n$ will become:

$$\mathbf{b}^m \bullet \mathbf{b}_n = \delta_{m,n} \quad \mathbf{b}_n \bullet \mathbf{b}_k = w_{nk} \quad \mathbf{b}^n \bullet \mathbf{b}^k = w^{nk} \quad \mathbf{b}^m = w^{mn} \mathbf{b}_n \quad \mathbf{b}_m = w_{mn} \mathbf{b}^n$$

(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

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

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

(d) Summary of the basic facts:

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

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

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 U'_n

Section 3 (d) introduced the *inverse* tangent base vectors called 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 U'_n according to

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

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

$$g' \leftrightarrow g \quad R \leftrightarrow S \quad e_n \rightarrow u'_n \quad e'_n \rightarrow u_n \quad E_n \rightarrow U'_n \quad E'_n \rightarrow U_n$$

Here are the key results, translated from above,

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

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

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

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

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

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

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

$$u_n \bullet u_m = \bar{g}_{nm}$$

$$U_n \bullet u_m = \delta_{n,m}$$

$$U_n \bullet U_m = g_{nm}$$

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

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	e'_n	u_n	$(e'_n)_i = \delta_{n,i}$	$(u_n)_i = \delta_{n,i}$
dual partners to the above	E'_n	U_n	$(E'_n)_i = g'_{ni}$	$(U_n)_i = g_{ni}$
tangent base vectors	u'_n	e_n	$(u'_n)_i = R_{in}$	$(e_n)_i = S_{in}$
reciprocal base vectors	U'_n	E_n	$(U'_n)_i = g'_{ia} S_{na}$ $= g_{na} R_{ia}$	$(E_n)_i = g_{ia} R_{na}$ $= g'_{na} S_{ia}$

and recall that $A_n \bullet a_m = \delta_{n,m}$ for each of the four dual pairs (two primed, two unprimed).

(f) Expanding vectors on different sets of basis vectorsx-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 $\bar{\mathbf{g}} \mathbf{V} = \sum_n \bar{V}_m \mathbf{U}_m$ 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 just momentarily as α_n ,

$$\mathbf{V} = \alpha_1 \mathbf{e}_1 + \alpha_2 \mathbf{e}_2 + \dots = \sum_n \alpha_n \mathbf{e}_n$$

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

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

Therefore, the expansion is

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

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

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

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 (script 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,

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

and here then are the x'-space expansions:

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

(g) Another way to write the $\mathbf{E}_n$

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

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

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,

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

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

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

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

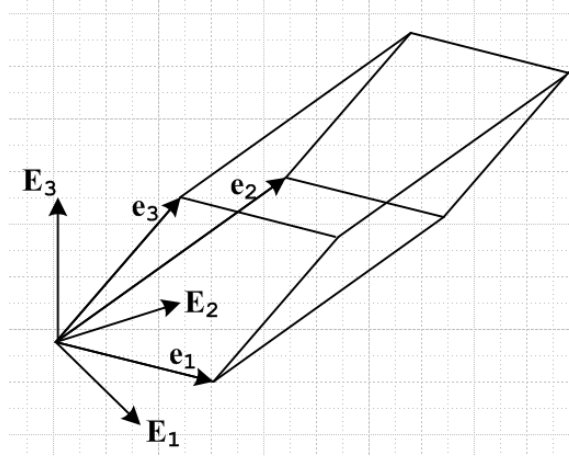
$$\mathbf{E}_1 \cdot \mathbf{e}_n = \det(\mathbf{R}) \mathbf{e}_2 \times \mathbf{e}_3 \cdot \mathbf{e}_n \quad \text{for } n = 1, 2, 3$$

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 $\mathbf{e}_2 \times \mathbf{e}_3 \cdot \mathbf{e}_1 = 1/\det(\mathbf{R}) = \det(\mathbf{S})$. But

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

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 $\det(\mathbf{R}) > 0$.



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

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

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

One could create a covariant partner to $\mathbf{e}_n$ which would be $\bar{\mathbf{e}}_n = \bar{\mathbf{g}} \mathbf{e}_n$ 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 $\bar{\mathbf{e}}_{\mathbf{n}} = \mathbf{e}_{\mathbf{n}}$ as usual, but of course $\mathbf{E}_{\mathbf{n}} \neq \mathbf{e}_{\mathbf{n}}$ 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

$$(\mathbf{e}_{\mathbf{n}})_{\mathbf{i}} \equiv g_{\mathbf{i}\mathbf{k}}(\mathbf{e}_{\mathbf{n}})^{\mathbf{k}} \quad // \text{ Standard Notation, } g \text{ lowers an index}$$

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

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

(i) Handedness of coordinate systems: the $\mathbf{e}_{\mathbf{n}}$, the sign of $\det(\mathbf{S})$, and Parity

1. Handedness of a Coordinate System. Let $\mathbf{b}_{\mathbf{n}}$ be a complete set of basis vectors in an N dimensional vector space, where the $\mathbf{b}_{\mathbf{n}}$ are not necessarily of unit length, and are not necessarily orthogonal. Consider this quantity

$$B \equiv \det(\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_N) = \varepsilon_{\mathbf{abc} \dots \mathbf{x}} (\mathbf{b}_1)_{\mathbf{a}} (\mathbf{b}_2)_{\mathbf{b}} \dots (\mathbf{b}_N)_{\mathbf{x}} = \mathbf{b}_1 \bullet [\mathbf{b}_2 \times \mathbf{b}_3 \dots \times \mathbf{b}_N]$$

where the generalized cross product is discussed in Appendix A. This basis $\mathbf{b}_{\mathbf{n}}$ defines a "coordinate system" in that we can expand a position vector as follows

$$\mathbf{x} = \sum_{\mathbf{n}} x^{(\mathbf{b})}_{\mathbf{n}} \mathbf{b}_{\mathbf{n}}$$

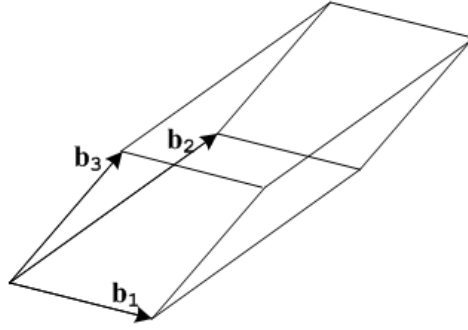
where the $x^{(\mathbf{b})}_{\mathbf{n}}$ are the "coordinates" of point $\mathbf{x}$ in this coordinate system. We make the following definition:

$$\begin{array}{lll} \text{system } \mathbf{b}_{\mathbf{n}} \text{ is a "right handed coordinate system"} & \text{iff} & B > 0 \\ \text{system } \mathbf{b}_{\mathbf{n}} \text{ is a "left handed coordinate system"} & \text{iff} & B < 0 \end{array}$$

One of course wants to show that for $N = 3$ this definition corresponds to one's intuition about right and left handed systems. For $N = 3$,

$$B \equiv \det(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3) = \varepsilon_{\mathbf{abc}} (\mathbf{b}_1)_{\mathbf{a}} (\mathbf{b}_2)_{\mathbf{b}} (\mathbf{b}_3)_{\mathbf{c}} = \mathbf{b}_1 \bullet [\mathbf{b}_2 \times \mathbf{b}_3]$$

Suppose the $\mathbf{b}_{\mathbf{n}}$ are arranged as shown in this picture, where the visual intention is that the corner nearest the label $\mathbf{b}_1$ is closest to the viewer.



With one's high-school-trained right hand, one can see that $\mathbf{b}_2 \times \mathbf{b}_3$ points in the general direction of $\mathbf{b}_1$ (certainly $\mathbf{b}_2 \times \mathbf{b}_3$ lies somewhere in the half space of the $\mathbf{b}_2, \mathbf{b}_3$ face plane which contains $\mathbf{b}_1$), and so the quantity $B = \mathbf{b}_1 \cdot [\mathbf{b}_2 \times \mathbf{b}_3] > 0$. So this is an example of a right-handed coordinate system. The figure shown is a skewed 3-piped which can be regarded as a distortion of an orthogonal 3-piped for which the $\mathbf{b}_n$ would span an orthogonal coordinate system in which one would have $\hat{\mathbf{b}}_1 = \hat{\mathbf{b}}_2 \times \hat{\mathbf{b}}_3$.

2. The x'-space $\mathbf{e}'_n$ coordinate system is always right handed. In our standard picture of x-space and x'-space, the coordinate system in x'-space is spanned by a set of basis vectors $\mathbf{e}'_n$ where $(\mathbf{e}'_n)_i = \delta_{n,i}$, as discussed in Section 3 (a). This system is "right handed" because

$$B = \varepsilon_{abc\dots x} (\mathbf{e}'_1)_a (\mathbf{e}'_2)_b \dots (\mathbf{e}'_N) = \varepsilon_{abc\dots x} \delta_{1a} \delta_{2b} \dots \delta_{xN} = \varepsilon_{123\dots N} = +1 > 0$$

where we use the standard normalization of the ε tensor as shown. Notice that this conclusion is independent of the metric tensor g' in x-space.

3. The x-space $\mathbf{u}_n$ coordinate system is always right handed. The basis $\mathbf{u}_n$ where $(\mathbf{u}_n)_i = \delta_{n,i}$ in x-space is right handed for the same reason as shown in the above paragraph, and for any g . When $g=1$ in x-space, the $\mathbf{u}_n$ form the usual Cartesian right-handed orthonormal basis in x-space. See Section 3 (c) concerning the meaning of "unit vector".

4. The x-space $\mathbf{e}_n$ coordinate system handedness is determined by the sign of $\det(S)$. Our x-space coordinate system of great interest is that spanned by the $\mathbf{e}_n$ basis vectors, where $(\mathbf{e}_n)_i = S_{in}$ as discussed in Section 3 (a). One has

$$B = \varepsilon_{abc\dots x} (\mathbf{e}_1)_a (\mathbf{e}_2)_b \dots (\mathbf{e}_N) = \varepsilon_{abc\dots x} S_{a1} S_{b2} \dots S_{xN} = \det(S)$$

Therefore, using $\sigma \equiv \text{sign}(\det S)$ and J being the Jacobian as in Section 5 (k),

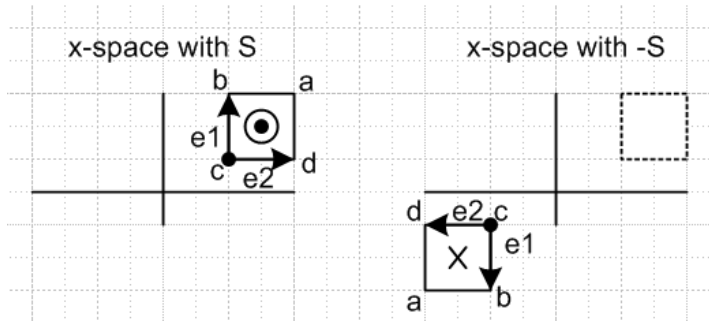
$$\begin{array}{lll} \text{system } \mathbf{e}_n \text{ is a "right handed coordinate system"} & \text{iff} & \det(S) = J > 0 \quad \text{or} \quad \sigma = +1 \\ \text{system } \mathbf{e}_n \text{ is a "left handed coordinate system"} & \text{iff} & \det(S) = J < 0 \quad \text{or} \quad \sigma = -1 \end{array}$$

5. The Parity Transformation. The identity transformation $F = 1$ results in matrix $S_I = I$ with $\det S_I = +1$. In this case the $\mathbf{e}_n$ form a right-handed coordinate system, and in fact $\mathbf{e}_n = \mathbf{u}_n$. The parity transformation $F = -1$, on the other hand, results in $S_P = -I$ with $\det(S_P) = (-1)^N$ and $\mathbf{e}_n = -\mathbf{u}_n$. When N is odd, the parity

transformation converts the right-handed $\mathbf{u}_n$ system to a left-handed $\mathbf{e}_n$ system. If S is some matrix which does not change handedness, meaning $\det S > 0$, then $S' = SS_P$ *does* change handedness for odd N , since in this case $\det S' = \det S \det S_P = (-1)^N \det S$. So given some S' that changes handedness for $N=\text{odd}$, one can regard it as "containing the parity transformation" which, if removed, would result in no handedness change. When N is even, handedness stays the same under $F = -1$.

6. N=3 Parity Inversion Example. Since $\mathbf{x}' = -\mathbf{x}$ under the parity transform $F = -1$, parity is a reflection of all position vectors through the origin. Objects sitting in x -space, such as N -pipeds, whether or not the origin lies inside the object, are "turned inside out" by the parity transformation, but the inside of the object still maps to the inside of the parity transformed object under this transformation.

Consider this crude picture which shows on the left a right-handed 3-piped in x -space where $\mathbf{e}_1$ and $\mathbf{e}_2$ happen to be perpendicular, and the back part of the 3-piped is not drawn. This 3-piped is associated with some transformation S [$(\mathbf{e}_n)_i = S_{in}$] with $\det S > 0$.



Now consider $S' = SS_P = S_P S = -S$. For this S' , the 3-piped appears as shown on the right. The two pipeds here are related by a parity transformation, all points inverting through the origin. On the right, the volume of the 3-piped lies *toward* the viewer from the plane shown. The circled dot on the left represents the out-facing normal vector of the 3-piped face area which is facing the viewer, and this normal is in the direction $-\mathbf{e}_1 \times \mathbf{e}_2$. This is called a "near face" in Appendix B since it touches the tails of the $\mathbf{e}_n$. After the parity transformation, this same face has become the *back* face on the inverted 3-piped shown on the right, with out-facing normal indicated by the X. The direction of this normal is $+\mathbf{e}_1 \times \mathbf{e}_2$. In general, an out-facing "near face" area points in the $-\mathbf{E}_n$ direction, and Appendix A shows that $\mathbf{E}_3 = \mathbf{e}_1 \times \mathbf{e}_2 / \det(S)$. On the left we have $\mathbf{E}_3 = \mathbf{e}_1 \times \mathbf{e}_2 / |\det(S)|$ so the face there just mentioned points in the $-\mathbf{E}_3 = -\mathbf{e}_1 \times \mathbf{e}_2$ direction. On the right we have $\mathbf{E}_3 = \mathbf{e}_1 \times \mathbf{e}_2 / \det(S') = -\mathbf{e}_1 \times \mathbf{e}_2 / |\det(S)|$, so the face there points in the $-\mathbf{E}_3 = +\mathbf{e}_1 \times \mathbf{e}_2$ direction.

7. The sign of $\det(S)$ in the curvilinear coordinates application. For a given ordering of the $\mathbf{x}'$ coordinates, $\det(S)$ will have a certain sign. By changing the $\mathbf{x}'_i$ ordering to any odd permutation of the original ordering (for example, swap two coordinates), $\det(S)$ will negate because two columns of $S_{ik}(\mathbf{x}') \equiv (\partial x_i / \partial x'_k)$ will be swapped. In the curvilinear coordinates application it is therefore always possible to select the ordering of the $\mathbf{x}'_i$ 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 to 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 $\varepsilon_{abc\dots}$ 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

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

or in an ordinary cross product

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

This permutation tensor has the usual properties that $\varepsilon_{123\dots N} = +1$, 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, $\varepsilon_{abc\dots}$ 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, $\varepsilon^{abc} = \varepsilon_{abc} = \varepsilon_a^b{}_c$ and so on. In a non-Cartesian x -space, however, one would say that $\varepsilon_a^b{}_c = g^{bb'} \varepsilon_{ab'c} \neq \varepsilon_{abc}$. In the Weinberg convention, one sets $\varepsilon^{123\dots N} = \varepsilon'_{123\dots N} = 1$ and $\varepsilon^{abc\dots} x = \varepsilon'_{abc\dots} x$ 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) Matrix Multiplication, the meaning of R^T , and Rotations in Standard Notation.

After reading this section, one will understand why matrix multiplication notation is often avoided in the standard notation. Since both indices of rank-2 tensors and tensor-like objects can be up or down, the meaning of matrix multiplication requires some clarification. Part of the complication is defining a reasonable transpose for the various up and down matrix forms. The smaller subsections below are presented here in a "short form" to allow for a reasonable logic flow since certain "facts" have not yet been introduced. The "long form" presentations of these facts occur in later sections of Section 7.

1. Translation of determinants. Section (f) 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).

2. Inverse of R and S. Again, section (f) 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:

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

We have silently used a certain "down tilt" matrix multiplication here which will be explained later.

3. 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

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

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

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

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:

$$\begin{aligned} M'^a{}_b &= g'^{aa'} M'_{a'}{}_b = \overset{(4)}{(R^a{}_c R^{a'}{}_d g^{cd})} \overset{(3)}{(S^e{}_{a'} S^f{}_b M_{ef})} \overset{(2)}{=} \overset{(5)}{(R^a{}_c R^{a'}{}_d g^{cd})} (S^e{}_{a'} S^f{}_b g_{ei} M^i{}_f) \\ &= [R^a{}_c R^{a'}{}_d g^{cd} S^e{}_{a'} S^f{}_b g_{ei}] M^i{}_f \\ &= [R^a{}_c (S^e{}_{a'} R^{a'}{}_d) g^{cd} S^f{}_b g_{ei}] M^i{}_f = [R^a{}_c (SR)^e{}_d g^{cd} S^f{}_b g_{ei}] M^i{}_f \\ &= [R^a{}_c \delta^e{}_d g^{cd} S^f{}_b g_{ei}] M^i{}_f = [R^a{}_c g^{cd} S^f{}_b g_{di}] M^i{}_f \quad (1) \\ &= [R^a{}_c (g^{cd} g_{di}) S^f{}_b] M^i{}_f = [R^a{}_c \delta^c{}_i S^f{}_b] M^i{}_f = [R^a{}_i S^f{}_b] M^i{}_f \quad \text{inverses} \\ &= R^a{}_{a'} S^{b'}{}_b M^{a'}{}_b, \quad \text{QED} \end{aligned}$$

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. [Later we will show that $S^i{}_j = R_j{}^i$ and this allows one to write the above equations in a manner that is easier to remember.]

4. Raising Lowering Rule:

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

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

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.

5. 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.

6. 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, and here a preview:

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

7. 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. Although these are "split level" matrices, one can see that the index c has the correct "adjacency" property to justify matrix multiplication. A second requirement is that any matrix summed index should 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 (SN means Standard Notation, dt means down-tilt) :

$$C = AB \quad // \text{ all down-tilt} \quad [C = AB]_{\text{SN}, \text{dt}} .$$

By application of suitable g tensors to the first equation above (or by lowering index a and raising index b on both sides), 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 also be represented by

$$C = AB \quad // \text{ all up-tilt} \quad [C = AB]_{\text{SN}, \text{ut}}$$

where ut means up-tilt.

It has just been shown that,

$$[C = AB]_{\text{SN}, \text{dt}} \Leftrightarrow [C = AB]_{\text{SN}, \text{ut}}$$

The conclusion 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 in section (u), 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 might wonder about the conversion from developmental notation (DN) to standard notation (SN) of the following matrix equation, where A and B are contravariant rank-2 tensors,

$$[AB = C]_{\text{DN}} \Leftrightarrow A_{ac} B_{cb} = C_{ab} \quad \rightarrow \quad A^{ac} B^{cb} = [???]^{ab} .$$

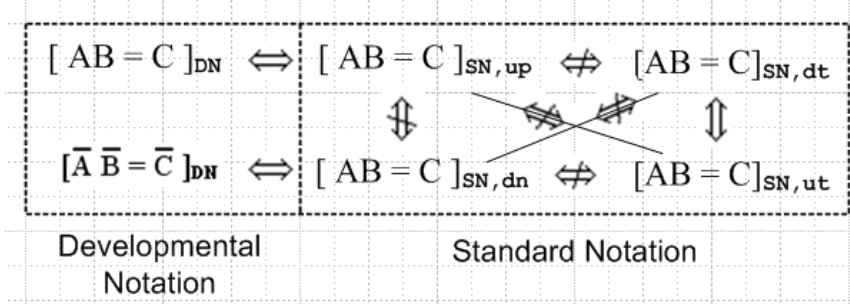
It is not hard to show entirely in developmental notation that, unless $S^T S = 1$ (meaning S is a rotation), the object C_{ab} is not a tensor. So for general underlying F, C is not a tensor, even though A and B are tensors. This fact is reflected in the LHS of the partially translated equation, where $A^{ac} B^{cb}$ is seen to not be a tensor because the c index is not a contraction (it is not tilted). The rules for translation from developmental to standard notation specify what to do with R, S, and with tensors. Since C is none of these objects, we cannot really "translate" the equation $[AB = C]_{\text{DN}}$ into standard notation. However, we could *define* $C^{ab} \equiv A^{ac} B^{cb}$ and then write

$$[AB = C]_{\text{DN}} \Leftrightarrow A_{ac} B_{cb} = C_{ab} \quad \rightarrow \quad A^{ac} B^{cb} = C^{ab} \Leftrightarrow [AB = C]_{\text{SN}, \text{up}}$$

with the understanding that C^{ab} is not a contravariant rank-2 tensor, and we now have yet a third kind of matrix multiplication in the Standard Notation, qualified by "up" meaning all indices are up on all objects. One could of course do the same thing with covariant matrices and write

$$[\bar{A} \bar{B} = \bar{C}]_{\text{DN}} \Leftrightarrow \bar{A}_{ac} \bar{B}_{cb} = \bar{C}_{ab} \quad \rightarrow \quad A_{ac} B_{cb} = C_{ab} \Leftrightarrow [AB = C]_{\text{SN}, \text{dn}}$$

The situation can be summarized in this picture:



In standard notation, there are four different kinds of "matrix multiplication" and only the down-tilt and up-tilt equations are the same equation, and only they are covariant. The moral here is that if one chooses to use matrix multiplication notation in the Standard Notation, one needs to be very clear which multiplication form one is dealing with.

What about the special case $[AB = 1]_{\text{DN}}$ where A is a contravariant tensor? In this case, it must be that $B = A^{-1}$ and then one has,

$$[AB = 1]_{\text{DN}} \Leftrightarrow A_{\text{ac}}(A^{-1})^{\text{cb}} = \delta_{\text{ab}} \quad \rightarrow \quad A^{\text{ac}}(A^{-1})^{\text{cb}} = (AA^{-1})^{\text{ab}} = (1)^{\text{ab}} = \delta^{\text{ab}}$$

that is to say: $[A(A^{-1}) = 1]_{\text{SM,up}}$

and in this case all four Standard Notation equations are valid, the other three being:

$$\begin{aligned} [AB = 1]_{\text{SM,dn}} &\Leftrightarrow A_{\text{ac}}(A^{-1})^{\text{cb}} = (AA^{-1})_{\text{ab}} = (1)_{\text{ab}} = \delta_{\text{ab}} \\ [AB = 1]_{\text{SM,ut}} &\Leftrightarrow A_{\text{a}}^{\text{c}}(A^{-1})_{\text{c}}^{\text{b}} = (AA^{-1})_{\text{a}}^{\text{b}} = (1)_{\text{a}}^{\text{b}} = \delta_{\text{a}}^{\text{b}} \\ [AB = 1]_{\text{SM,dt}} &\Leftrightarrow A^{\text{a}}_{\text{c}}(A^{-1})^{\text{c}}_{\text{b}} = (AA^{-1})^{\text{a}}_{\text{b}} = (1)^{\text{a}}_{\text{b}} = \delta^{\text{a}}_{\text{b}} \end{aligned}$$

so that

$$[AB = 1]_{\text{DN}} \Leftrightarrow [AB = 1]_{\text{SN,up}} \Leftrightarrow [AB = 1]_{\text{SN,dn}} \Leftrightarrow [AB = 1]_{\text{SN,ut}} \Leftrightarrow [AB = 1]_{\text{SN,dt}}$$

Now what about the special case $[RS = 1]_{\text{DN}}$? Since R is not a contravariant tensor the above logic does not apply in the sequence stated. But we have rules for translating R and S, so

$$[RS = 1]_{\text{DN}} \Leftrightarrow R_{\text{ac}}S_{\text{cb}} = \delta_{\text{ab}} \quad \rightarrow \quad R^{\text{a}}_{\text{c}}S^{\text{c}}_{\text{b}} = \delta^{\text{a}}_{\text{b}} \Leftrightarrow [RS = 1]_{\text{SN,dt}}$$

and since the four SN forms are equivalent ($S = R^{-1}$), we end up with the same conclusion as above,

$$[RS = 1]_{\text{DN}} \Leftrightarrow [RS = 1]_{\text{SN,up}} \Leftrightarrow [RS = 1]_{\text{SN,dn}} \Leftrightarrow [RS = 1]_{\text{SN,ut}} \Leftrightarrow [RS = 1]_{\text{SN,dt}}$$

confirming what was stated in a preliminary way in item 2 above.

Finally, what about the special case $[S^T S = 1]_{\text{DN}}$? Again using the rule that $S_{ij} \rightarrow S^i_j$,

$$[S^T S = 1]_{\text{DN}} \Leftrightarrow S_{\text{ca}}S_{\text{cb}} = \delta_{\text{ab}} \quad \rightarrow \quad S^{\text{c}}_{\text{a}}S^{\text{c}}_{\text{b}} = \delta_{\text{ab}} \Leftrightarrow ???$$

The combination $S^c_a S^c_b$ does not have the proper adjacency to be matrix multiplication. In order to achieve adjacency, we must have some notion of "transpose" in the Standard Notation. That notion is the subject of the next section, and the result is going to be this:

$$(A^T)^{ab} = A^{ba} \quad (A^T)^a_b = A_b^a \quad (A^T)_a^b = A^b_a \quad (A^T)_{ab} = A_{ba}$$

Therefore $S^c_a = (S^T)_a^c$ and we then have

$$[S^T S = 1]_{DN} \Leftrightarrow S_{ca} S_{cb} = \delta_{ab} \rightarrow (S^T)_a^c S^c_b = \delta_{ab} \Leftrightarrow ???$$

Although adjacency of the c index is now obtained, the form matches none of the Standard Notation forms for matrix multiplication defined above. Thus, no contact is being made with any of these standard forms by this translation.

Since $S^T S = 1$, one must have $S = (S^T)^{-1}$ and therefore all four of our stated SN forms are equivalent. So we then have this interesting situation:

$$[S^T S = 1]_{DN} \Leftrightarrow (S^T)_a^c S^c_b = \delta_{ab} \quad \Leftrightarrow \quad [S^T S = 1]_{SN, dt} \Leftrightarrow \text{the other three SN forms}$$

What this says is that $[S^T S = 1]_{DN}$ and $[S^T S = 1]_{SN, dt}$ are *completely different equations*. Specifically,

$$[S^T S = 1]_{DN} \Leftrightarrow S_{ca} S_{cb} = \delta_{ab} \rightarrow (S^T)_a^c S^c_b = \delta_{ab} \quad \text{or} \quad S^c_a S^c_b = \delta_{ab}$$

$$[S^T S = 1]_{SN, dt} \Leftrightarrow (S^T)^a_c S^c_b = \delta^a_b \quad \text{or} \quad S_c^a S^c_b = \delta^a_b .$$

The very last equation on the right can be written (as will be seen below) $R^a_c R_b^c = \delta_{ab}$ and this is one of the standard orthogonality conditions on R and this is always true for any underlying transformation F. Therefore, $[S^T S = 1]_{SN, dt}$ is always true, and this means that in terms of Standard Notation down-tilt matrix multiplication, the matrix S is always real orthogonal. In contrast, in terms of Developmental Notation matrix multiplication, S is only real orthogonal if it is a rotation matrix! That is to say, in the general case:

$$[S^T = S^{-1}]_{DN} \quad \Leftrightarrow \quad [S^T = S^{-1}]_{SN, dt}$$

So the very meaning of the equation $S^T = S^{-1}$ is dependent on which notion of matrix multiplication is involved. The operator S^{-1} is really defined in terms of $SS^{-1} = S^{-1}S = 1$ and this involves matrix multiplication.

One final thing to note:

$$(S_{ab})_{DN} = (S^a_b)_{SN} \neq (S_{ab})_{SN} = g_{ac}(S^c_b)_{SN} \quad // \text{ unless it happens that } g = 1$$

so clearly when one writes S_{ab} or R_{ab} one should be clear which notation is involved.

8. Transpose of a rank-2 tensor. If A is a contravariant 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 $(A^T)^a_b = A_b^a$, $(A^T)_a^b = A^b_a$, 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, $(R^T)^a_b = R_b^a$ where b is a g' type index and a is a g type index, as will be elaborated in section (o) below. Similarly $(S^T)^a_b = S_b^a$.

To summarize the situation with transposes in Standard Notation:

$$\begin{array}{lll} (A^T)^{ab} = A^{ba} & (R^T)^{ab} = R^{ba} & (S^T)^{ab} = S^{ba} \\ (A^T)^a_b = A_b^a & (R^T)^a_b = R_b^a & (S^T)^a_b = S_b^a \\ (A^T)_a^b = A^b_a & (R^T)_a^b = R^b_a & (S^T)_a^b = S^b_a \\ (A^T)_{ab} = A_{ba} & (R^T)_{ab} = R_{ba} & (S^T)_{ab} = S_{ba} \end{array}$$

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: In the Standard Notation, S is a real-orthogonal matrix.

This means all of the following: $[S^{-1} = S^T]_{SN}$ $[SS^T = 1]_{SN}$ $[S^TS = 1]_{SN}$

where $[...]_{SN}$ can mean any of the four Standard Notation matrix multiplication forms given in the previous section. This theorem does *not* apply to the matrix S in the developmental notation as also explained in the previous section. That is to say, the equations $[SS^T = 1]_{DN}$ and $[SS^T = 1]_{SN}$ are different.

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

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

Theorem 2: In the Standard Notation, R is a real-orthogonal matrix.

This means all of the following: $[R^{-1} = R^T]_{SN}$ $[RR^T = 1]_{SN}$ $[R^TR = 1]_{SN}$

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 similar 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,

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

Comment: In the developmental notation, neither R nor S is a real-orthogonal matrix, unless by accident, but in the standard notation R and S are always real-orthogonal.

Theorem 3:

- (a) $[S = R^T]_{SN}$ and $[R = S^T]_{SN}$
 (b) $S^a_b = R^b_a$ and $S_a^b = R^a_b$ (reflect indices in vertical line between them)

Proof of theorem:

(a) As shown in the previous section, $S = R^{-1}$ for any index positions in the standard notation. From Theorem 2, $R^T = R^{-1}$ in standard notation. Therefore $S = R^T$ (and so $S^T = R^{TT} = R$) .

(b) From (a) ,

$$\begin{aligned}
 S &= R^T \Rightarrow S^a_b = (R^T)^a_b = R^b_a \\
 S &= R^T \Rightarrow S^a_b = (R^T)^a_b = R^b_a
 \end{aligned}
 \quad \text{QED}$$

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 A , $\det(A) \neq \det(A^T)$.

This surprising result points out a potential hazard of using tilted matrices in the standard notation, and perhaps is an indication of why people avoid matrix notation 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)

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

In the tilted standard notation, however, one has

$$\det[A^i_j] = \varepsilon_{ab\dots} A^1_a A^1_b \dots = \varepsilon_{ab\dots} A^a_1 A^b_2 \dots$$

$$\det[(A^T)^i_j] = \varepsilon_{ab\dots} (A^T)^1_a (A^T)^1_b \dots = \varepsilon_{ab\dots} A^1_a A^2_b$$

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

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

$$\det(A^T) = \varepsilon_{ab\dots} A^1_a A^2_b = \begin{vmatrix} A^1_1 & A^2_1 \\ A^1_2 & A^2_2 \end{vmatrix} = A_1^1 A_2^2 - A_2^1 A_1^2 \neq \det(A)$$

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: $[R^T R = 1]_{\text{SN}}$ does not imply that $\det(R) = \pm 1$ for a tilted R matrix.

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 for a tilted R matrix in standard notation. Therefore, $[R^T R = 1]_{\text{SN}}$ does *not* lead us to the conclusion that the Jacobian of Section 5 (k) is forced to be ± 1 (it is always the *tilted* R and S matrix that appears in equations like $1/J = \det(R^i_j)$).

9. Orthogonality rules. The rules are easy to show:

$$\begin{aligned} [R^T R = 1]_{\text{SN}, dt} &\Rightarrow (R^T)^a_b R^b_c = \delta^a_c &\Rightarrow R_b^a R^b_c = \delta^a_c \\ [R R^T = 1]_{\text{SN}, dt} &\Rightarrow R^a_b (R^T)^b_c = \delta^a_c &\Rightarrow R^a_b R^b_c = \delta^a_c \\ [R^T R = 1]_{\text{SN}, ut} &\Rightarrow (R^T)_a^b R_b^c = \delta_a^c &\Rightarrow R_b^a R_b^c = \delta_a^c \\ [R R^T = 1]_{\text{SN}, ut} &\Rightarrow R_a^b (R^T)_b^c = \delta_a^c &\Rightarrow R_a^b R^c_b = \delta_a^c \end{aligned}$$

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

10. Rotation matrices. In developmental notation, a real orthogonal matrix R has the property that

$$R R^T = 1 \quad \Leftrightarrow \quad R^{-1} = R^T \quad \Leftrightarrow \quad R^T R = 1$$

The determinant of R must be ± 1 since

$$R R^T = 1 \quad \Rightarrow \quad (\det R)^2 = \det(1) = 1 \quad \Rightarrow \quad \det R = \pm 1$$

If $\det R = +1$, then R can be a rotation matrix (in any number of dimensions N). If $\det R = -1$, then R can be that same rotation matrix combined in any manner with an odd number of axis reflection matrices, each represented by a unit matrix in which one of the diagonal 1's is replaced by a -1. For example, for $N = 2$ one can write,

$$R = \mathcal{R} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = \text{rotation matrix} \quad \det R = +1$$

$$r_x = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = \text{reflect the x-axis}$$

$$r_y = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \text{reflect the y-axis}$$

$$R = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -\cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \mathcal{R} r_x \quad \det R = -1$$

For $N = 3$ one might write

$$R = \mathcal{R} = \text{rotation matrix (Euler angles)} \quad \det R = +1$$

$$r_x = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad r_y = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad r_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad P = r_x r_y r_z = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

$$A = R r_x \text{ or } R r_y \text{ or } r_x R \text{ or } R P \text{ etc} \quad \text{all have:} \quad \det A = -1$$

The parity P operation is included only when $N = \text{odd}$.

In general, then, a real orthogonal matrix R satisfying $RR^T = 1$ is a rotation matrix combined with any even number (including 0) of axis reflection matrices ($\det R = +1$), or it is a rotation matrix combined with any odd number of axis reflections ($\det R = -1$). If one is interested in real-orthogonal transformations that can be arrived at in a continuous morphing manner from the identity transformation, then one would include only the rotation matrices in the set of real orthogonal matrices.

So, our developmental notation characterization of a rotation matrix is this:

$$[RR^T = 1]_{DN}$$

where we now use the notation described in item 7 above.

Up to this point R has stood for a "rotation matrix", but now we shall assume that this R is also the matrix which represents the linearized F transformation as in Section 2. In standard notation one finds

$$RR^T = 1 \Leftrightarrow R_{ab}R_{cb} = \delta_{a,c} \quad \rightarrow \quad R^a{}_b R^c{}_b = \delta_{a,c}.$$

The equation $R^a{}_b R^c{}_b = \delta_{a,c}$ cannot be written in any of the four standard notation matrix multiplication forms discussed in item 7, so we just leave it "as is".

There are several other characterizations of R being rotation matrix. First consider this sequence:

$$\begin{aligned}
R^a_b R^c_b &= \delta_{a,c} && // \text{ the above} \\
R^a_b S^c_b &= \delta_{a,c} && // R^c_b = S^c_b \text{ as shown above in Theorem 3} \\
R^a_b S^c_b R^d_c &= \delta_{a,c} R^d_c && // \text{ apply } R^d_c \text{ to both sides and sum on } c \\
R^a_b (SR)_b^d &= R^d_a && // SR = 1 \\
R^a_b \delta_b^d &= R^d_a \\
R^a_d &= R^d_a \\
S_d^a &= S^d_a
\end{aligned}$$

where on each of the last two equations the two sides are related by reflection of the indices in a horizontal line running through the indices.

Since it has now been shown that for a rotation $R^a_b = R^b_a$, it follows that

$$R^a_b R^c_b = \delta_{a,c} \Rightarrow R^b_a R^b_c = \delta_{a,c}$$

The following theorem has now been proven:

Theorem 5 (Rotations): A general transformation F can be represented at a point $\mathbf{x}$ by the linearized matrices R and S as discussed in Section 2. In developmental notation if R is a rotation matrix then it can be characterized by any of these equivalent statements in developmental notation (DN) :

$$[RR^T = 1]_{DN} \quad [R^T R = 1]_{DN} \quad [R^{-1} = R^T]_{DN} .$$

In standard notation, the fact that R is a rotation matrix can be characterized by any of the following equivalent statements:

$$R^a_b R^c_b = \delta_{a,c} \quad R^b_a R^b_c = \delta_{a,c} \quad R^a_b = R^b_a .$$

Of course if R is a rotation, so is $S = R^{-1}$. Since $R^i_j = S^i_j$, the above are the same as

$$S^a_b S^c_b = \delta_{a,c} \quad S^b_a S^b_c = \delta_{a,c} \quad S^a_b = S^b_a .$$

In principle, $R = R(\mathbf{x})$ might be a different rotation matrix at different points $\mathbf{x}$ in x-space, so that F in that case would be a non-linear transformation. This would happen in 3D space if the Euler angles which parameterize a general rotation were arbitrary (but reasonable) functions of position $\mathbf{x}$.

What can be said about a rotation matrix Q which is not the R matrix which linearizes F? The translation from developmental to standard notation could be defined this way (same as for R and S)

$$\begin{aligned}
QQ^T = 1 &\Leftrightarrow Q_{ab} Q_{cb} = \delta_{a,c} && \rightarrow && Q^a_b Q^c_b = \delta_{a,c} \\
Q^T Q = 1 &\Leftrightarrow Q_{ba} Q_{bc} = \delta_{a,c} && \rightarrow && Q^b_a Q^b_c = \delta_{a,c}
\end{aligned}$$

Since $Q^a_b (Q^{-1})^b_c = \delta_{a,c}$ comparison shows that one must have $(Q^{-1})^b_c = Q^c_b$ which is a straight swap of the indices. Since we have no information regarding the possible tensor or tensor-like nature of Q, we don't know how to raise or lower its indices, and nothing much more can be said.

11. Variations on the relation between g and g'. The third subsection 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} .

In the previous section it was shown that if R is a rotation matrix, then $[RR^T = 1]_{\mathbf{DN}}$ and in standard notation on gets

$$R^a_c R^b_c = \delta^{a,b} .$$

If x-space is Cartesian so metric tensor $g = 1$, then it was just shown above that for any R,

$$g'^{ab} = R^a_c R^b_c .$$

Therefore, when R is a rotation, $g' = 1$ as well as $g = 1$, just as one would expect.

(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 $\bar{\mathcal{R}}$. 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 $\mathcal{R}_{dp}$ and shows how to expand tensors on such bases.

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

The most famous example occurs with $N=3$ and the rotation group $SU(2)$ in which case $\mathcal{R}^{(1)} \equiv \mathcal{R}$ can be decomposed according to $\mathcal{R}^{(1)} \otimes \mathcal{R}^{(1)} = \mathcal{R}^{(2)} \oplus \mathcal{R}^{(1)} \oplus \mathcal{R}^{(0)}$ where the $\oplus$ 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 $D^{(2)}_{M,M'}(\phi, \theta, \psi)$ where $M, M' = 2, 1, 0, -1, -2$, and where ϕ, θ, ψ are the "Euler angles" which serve to label a particular rotation. This $D^{(2)}$ object is the $L=2$ matrix representation of the rotation group. Taking two vectors $\mathbf{A}$ 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 symmetric 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^{abc} .

The Standard Model of elementary particle physics is chock full of direct products of this nature, where the idea of rotational symmetry is extended to other kinds of "internal" symmetry, spin and isospin being two examples. Representations of the Lie symmetry group $SU(3)$ are associated with quarks which are among of the fundamental building blocks of the Standard Model.

The group discussion above can be applied generally to quantum physics. The basic idea is that if "the physics" (the Hamiltonian or Lagrangian) describing some quantum object is invariant under a certain symmetry group (such as rotational symmetry or perhaps some discrete crystal symmetry), then the quantum states of that object can be classified according to the representations of that group. The Bohr hydrogen atom "physics" $H \sim \nabla^2 - 1/|\mathbf{r}|$ has perfect rotation group symmetry and is also symmetric about the axis (angle ψ) from center to electron (no spin). The representation functions then must have $M' = 0$, and then $D^{(L)}_{M,0}(\phi, \theta, \psi) \sim Y_{LM}(\theta, \phi)$, 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

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

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.

In our notation, the two equations above would be written

$$A'_{abc\dots} = R_a^{a'} R_b^{b'} \dots A_{a'b'c'\dots} = \sum_{\text{primed}} A_{a'b'c'\dots} R_a^{a'} R_b^{b'} \dots \quad (6)$$

$$A'^{abc\dots} = R_a^a R_b^b \dots A^{a'b'c'\dots} = \sum_{\text{primed}} A^{a'b'c'\dots} R_a^a R_b^b \dots \quad (7)$$

(k) The Contraction Tilt-Reversal Rule

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

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

When is index tilt-reversal allowed and when is it not allowed? It is always allowed when both indices of the tilted contraction are valid tensor indices. Consider these four examples to be discussed below:

$$R_a^b A_b = R_{ab} A^b \quad \text{Proof: } R_a^b A_b = R_{ac} g^{cb} g_{bd} A^d = g^{cb} g_{bd} R_{ac} A^d = \delta^c_d R_{ac} A^d = R_{ac} A^c$$

$$A^a R_a^b \neq A_a R^{ab} \quad \text{Proof: } A^a R_a^b = g^{ac} A_c R^{db} g'_{da} = g^{ac} g'_{ad} A_c R^{db} \neq A_a R^{ab}$$

$$A^a \partial_a = A_a \partial^a \quad \text{Proof: } A^a \partial_a f = g^{ac} A_c g_{ad} \partial^d f = g^{ac} g_{ad} A_c \partial^d f = \delta^c_d A_c \partial^d f = A_c \partial^c f$$

$$\begin{aligned} \partial_a A^a \neq \partial^a A_a \quad \text{Proof: } \partial_a A^a &= (g_{ac} \partial^c)(g^{ad} A_d) = g_{ac} g^{ad} (\partial^c A_d) + g_{ac} (\partial^c g^{ad}) A_d \\ &= \delta^d_c (\partial^c A_d) + g_{ac} (\partial^c g^{ad}) A_d = \partial^c A_c + (\partial_a g^{ad}) A_d \neq \partial^c A_c \end{aligned}$$

In the first example, since R_a^b is not a tensor, one is on dangerous ground doing the tilt reversal, but it happens to work because the second index is associated with metric tensor g_{ij} which is the same metric tensor that raises and lowers indices of A_b . In the second example, the tilt-reversal fails because the first index of R_a^b is associated with the x' -space metric tensor g'_{ij} . In the third example, both indices are valid tensor indices (with the same metric tensor).

The fourth example shows a failure of the tilt-reversal rule and this example is very important. The inequality becomes an equality only if the underlying transformation F is linear so that R and S and g are then constants independent of position. For general F , such as the F involved in curvilinear coordinate transformations, the object $\partial_a A^b$ is not a rank-2 tensor and so the object $\partial_a A^a$ does not represent contraction of two true tensor indices and therefore the "contraction tilt-reversal rule" does not apply. The rule of the next section also does not apply for this same reason, so $\partial_a A^a$ does not transform as a scalar under general F . Section (v) below continues this topic.

Here is one more example along the lines of the fourth example above that shows the potential danger of reversing a tilt when it is not justified. Consider the equation,

$$V^a = \varepsilon^{abc} B_{b;c} \quad (1) \quad // \text{ valid}$$

where

B_b = a tensorial vector

$B_{b;c} = \partial_c B_b - \Gamma_{bc}^n B_n$ = the covariant derivative of vector B_b = a rank-2 tensor

ε^{abc} = a rank-3 tensor density (weight -1) (the Levi-Civita tensor)
 V^a = a vector density (weight -1)

With regard to $B_{\mathbf{b};c}$: (a) in comma notation one writes $\partial_c B_{\mathbf{b}} = B_{\mathbf{b},c}$; (b) $\Gamma_{\mathbf{bc}}^{\mathbf{n}} = \Gamma_{\mathbf{cb}}^{\mathbf{n}}$.

Since all indices on equation (1) are tensor indices, one can lower index a and reverse the b and c tilts to get

$$V_{\mathbf{a}} = \varepsilon_{\mathbf{abc}} B^{\mathbf{b};c} \quad (2) \quad // \text{ valid}$$

Now go back to equation (1). Because ε^{abc} is antisymmetric on b and c, whereas $\Gamma_{\mathbf{bc}}^{\mathbf{n}}$ is symmetric on b and c, one can write $\varepsilon^{abc} B_{\mathbf{b};c} = \varepsilon^{abc} B_{\mathbf{b},c}$ since the Γ term vanishes by symmetry. Thus one gets

$$V^{\mathbf{a}} = \varepsilon^{\mathbf{abc}} B_{\mathbf{b},c} \quad (3) \quad // \text{ valid}$$

Were one to blindly lower a and reverse the b and c tilts, one would get

$$V_{\mathbf{a}} = \varepsilon_{\mathbf{abc}} B^{\mathbf{b},c} \quad (4) \quad // \text{ NOT valid}$$

The reason for "not valid" is that the reversal of the b tilt is not justified. That is,

$$\begin{aligned} V^{\mathbf{a}} &= \varepsilon^{\mathbf{abc}} B_{\mathbf{b},c} = \varepsilon^{\mathbf{abc}} \partial_c B_{\mathbf{b}} = g^{\mathbf{bb}'} \varepsilon^{\mathbf{a} \mathbf{b}',c} \partial_c (g_{\mathbf{bb}''} B^{\mathbf{b}''}) \\ \Rightarrow V_{\mathbf{a}} &= g^{\mathbf{bb}'} \varepsilon_{\mathbf{ab}',c} \partial_c (g_{\mathbf{bb}''} B^{\mathbf{b}''}) = g^{\mathbf{bb}'} \varepsilon_{\mathbf{ab}',c} \partial^c (g_{\mathbf{bb}''} B^{\mathbf{b}''}) \quad // \text{ c tilt reversal is OK} \\ &= g^{\mathbf{bb}'} g_{\mathbf{bb}''} \varepsilon_{\mathbf{ab}',c} (\partial^c B^{\mathbf{b}''}) + g^{\mathbf{bb}'} \varepsilon_{\mathbf{ab}',c} (\partial^c g_{\mathbf{bb}''}) B^{\mathbf{b}''} \\ &= \delta^{\mathbf{b}'}_{\mathbf{b}''} \varepsilon_{\mathbf{ab}',c} (\partial^c B^{\mathbf{b}''}) + g^{\mathbf{bb}'} \varepsilon_{\mathbf{ab}',c} (\partial^c g_{\mathbf{bb}''}) B^{\mathbf{b}''} \\ &= \varepsilon_{\mathbf{abc}} (\partial^c B^{\mathbf{b}}) + g^{\mathbf{bb}'} \varepsilon_{\mathbf{ab}',c} (\partial^c g_{\mathbf{bb}''}) B^{\mathbf{b}''} \\ &= \varepsilon_{\mathbf{abc}} B^{\mathbf{b},c} + g^{\mathbf{bb}'} \varepsilon_{\mathbf{ab}',c} (\partial^c g_{\mathbf{bb}''}) B^{\mathbf{b}''} = \varepsilon_{\mathbf{abc}} B^{\mathbf{b},c} + \text{extra term!} \end{aligned}$$

It is due to this extra term that (4) is not valid. Basically this is the same as the fourth example above, but the situation is embedded in a more complicated environment (extra tensors, tensor densities, comma notation, covariant derivatives, other tilted indices, etc). One way to summarize the example is this:

$$(V^{\mathbf{a}} = \varepsilon^{\mathbf{abc}} B_{\mathbf{b},c}) \Leftrightarrow (V^{\mathbf{a}} = \varepsilon^{\mathbf{abc}} B_{\mathbf{b};c}) \Leftrightarrow (V_{\mathbf{a}} = \varepsilon_{\mathbf{abc}} B^{\mathbf{b};c}) \not\Leftrightarrow (V_{\mathbf{a}} = \varepsilon_{\mathbf{abc}} B^{\mathbf{b},c})$$

(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:

$$V'^a = R^a_b V^b$$

$$V'_a = S^b_a V_b$$

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):

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

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

$$A^a Y_a = A_a Y^a = A \cdot Y$$

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:

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

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 ($g_{ab} = g_{ba}$!)

$$\begin{aligned} g_{ab}g^{bc} &= \delta_a^c = g_a^c & // \text{ start out} \\ g_a^b g_b^c &= \delta_a^c = g_a^c & // \text{ tilt reversal of the line above; just says } \delta_a^b \delta_b^c = \delta_a^c \\ g^{ab}g_{bc} &= \delta^a_c = g^a_c \\ g^a_b g^b_c &= \delta^a_c = g^a_c \end{aligned}$$

(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

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

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 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

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

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

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

(p) Repeat all the above for S

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

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

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

(q) Theorem: $S^a_b = R_b^a$ and $S_a^b = R^b_a$ (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:

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

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

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

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:

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

Notice that the above theorem says

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

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

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

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

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

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) ,

$$T'^{abc}_{de} = R^a_a, R^b_b, R^c_c, S^{d'}_d S^{e'}_e T^{a'b'c'}_{d'e'}, \quad // \text{ no pairs contracted}$$

becomes

$$T'^{abc}_{de} = R^a_a, R^b_b, R^c_c, R^{d'}_d R^{e'}_e T^{a'b'c'}_{d'e'}, \quad // \text{ no pairs contracted}$$

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 Inversion Rule, and the Cancellation Rule

The above theorem $S^a_b = R_b^a$ can be used to eliminate S in various forms of $RS = 1$:

$SR = 1$	$S^a_b R^b_c = \delta^a_c$	$R_b^a R^b_c = \delta^a_c$	$R_b^a R^b_c = \delta^a_c$	$\Sigma \text{ 1st}$
$R^T S^T = 1$	$R^b_a S^c_b = \delta^c_a$	$R^b_a R^c_b = \delta^c_a$	$R^b_a R^c_b = \delta^c_a$	$\Sigma \text{ 1st}$
$RS = 1$	$R^a_b S^b_c = \delta^a_c$	$R^a_b R^b_c = \delta^a_c$	$R^c_b R^a_b = \delta^c_a$	$\Sigma \text{ 2nd}$
$S^T R^T = 1$	$S^b_a R^c_b = \delta^c_a$	$R^b_a R^c_b = \delta^c_a$	$R^c_b R^a_b = \delta^c_a$	$\Sigma \text{ 2nd}$

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 $RR^T = R^T R = 1$.

Inversion Rule. Consider now an equation which one wants to invert for the object on the right,

$$[\text{----}^a \text{----}] = R^a_b [\text{-----}^b \text{-----}] \quad (*) \quad // \text{ before}$$

The **inversion rule** for moving R to the other side of the equation is to reflect R's two indices in the *horizontal* index plane, so the result will be

$$R_a^b [\text{----}^a \text{----}] = [\text{-----}^b \text{-----}] \quad // \text{ after}$$

Proof: Rename $b \rightarrow b'$ in (*), apply R_a^b to both sides and sum on a, then use an orthogonality rule:

$$R_a^b [\text{----}^a \text{----}] = R_a^b R^a_b, [\text{-----}^{b'} \text{-----}] = \delta^{b'}_b, [\text{-----}^{b'} \text{-----}] = [\text{-----}^b \text{-----}] \quad \text{QED}$$

Recall from the last section that reflection in the *vertical* index plane has a different implication,

$$S_{\mathbf{b}}^{\mathbf{a}} = R_{\mathbf{b}}^{\mathbf{a}} \quad \text{or} \quad (R^{-1})_{\mathbf{b}}^{\mathbf{a}} = R_{\mathbf{b}}^{\mathbf{a}} \quad \text{or} \quad (R^T)_{\mathbf{b}}^{\mathbf{a}} = R_{\mathbf{b}}^{\mathbf{a}}$$

Cancellation Rule. Next, consider a different generic equation,

$$R_{\mathbf{b}}^{\mathbf{a}} [\text{----}^{\mathbf{b}}\text{----}] = R_{\mathbf{b}}^{\mathbf{a}} [\text{-----}^{\mathbf{b}}\text{-----}] \quad (**) \quad // \text{ before}$$

The **cancellation rule** says the equation is still valid if identical contracted R factors are canceled on both sides such that the contraction index becomes a free index,

$$[\text{----}^{\mathbf{b}}\text{----}] = [\text{-----}^{\mathbf{b}}\text{-----}] \quad // \text{ after}$$

Proof: Rename $\mathbf{b} \rightarrow \mathbf{b}'$ in $(**)$, apply $R_{\mathbf{a}}^{\mathbf{b}}$ to both sides and sum on $\mathbf{a}$, then use an orthogonality rule,

$$R_{\mathbf{a}}^{\mathbf{b}} R_{\mathbf{b}'}^{\mathbf{a}} [\text{----}^{\mathbf{b}'}\text{----}] = R_{\mathbf{a}}^{\mathbf{b}} R_{\mathbf{b}'}^{\mathbf{a}} [\text{-----}^{\mathbf{b}'}\text{-----}]$$

$$\delta_{\mathbf{b}'}^{\mathbf{b}} [\text{----}^{\mathbf{b}'}\text{----}] = \delta_{\mathbf{b}'}^{\mathbf{b}} [\text{-----}^{\mathbf{b}'}\text{-----}]$$

$$[\text{----}^{\mathbf{b}}\text{----}] = [\text{-----}^{\mathbf{b}}\text{-----}] \quad \text{QED}$$

(s) The tangent and reciprocal base vectors and expansions on same

Tangent and reciprocal base vectors

Here are some basic translations:

$$(\mathbf{e}_{\mathbf{n}})_{\mathbf{i}} \rightarrow (\mathbf{e}_{\mathbf{n}})^{\mathbf{i}} \quad // \text{ contravariant index } i$$

$$(\bar{\mathbf{e}}_{\mathbf{n}})_{\mathbf{i}} \rightarrow (\mathbf{e}_{\mathbf{n}})_{\mathbf{i}} \quad // \text{ covariant index } i$$

$$(\mathbf{E}_{\mathbf{n}})_{\mathbf{i}} \rightarrow (\mathbf{e}^{\mathbf{n}})^{\mathbf{i}} \quad // \text{ contravariant index } i$$

$$(\bar{\mathbf{E}}_{\mathbf{n}})_{\mathbf{i}} \rightarrow (\mathbf{e}^{\mathbf{n}})_{\mathbf{i}} \quad // \text{ covariant index } i$$

$$(\mathbf{e}_{\mathbf{n}})_{\mathbf{i}} = S_{\mathbf{i}\mathbf{n}} \rightarrow (\mathbf{e}_{\mathbf{n}})^{\mathbf{i}} = S^{\mathbf{i}}_{\mathbf{n}} = R_{\mathbf{n}}^{\mathbf{i}} \quad // \text{ contravariant index } i$$

$$(\bar{\mathbf{E}}_{\mathbf{n}})_{\mathbf{i}} = R_{\mathbf{n}\mathbf{i}} \rightarrow (\mathbf{e}^{\mathbf{n}})_{\mathbf{i}} = R^{\mathbf{n}}_{\mathbf{i}} \quad // \text{ covariant index } i$$

$$(\mathbf{E}_{\mathbf{n}})_{\mathbf{i}} = R_{\mathbf{n}\mathbf{k}} g_{\mathbf{k}\mathbf{i}} \rightarrow (\mathbf{e}^{\mathbf{n}})^{\mathbf{i}} = R^{\mathbf{n}}_{\mathbf{k}} g^{\mathbf{k}\mathbf{i}} = R^{\mathbf{n}\mathbf{i}} \quad // \text{ contravariant index } i$$

As noted earlier, writing a vector in bold such as $\mathbf{e}_{\mathbf{n}}$ 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,

$$S = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \dots \mathbf{e}_{\mathbf{N}}] \rightarrow S^{\mathbf{i}}_{\mathbf{j}} = [(\mathbf{e}_1)^{\mathbf{k}}, (\mathbf{e}_2)^{\mathbf{k}}, (\mathbf{e}_3)^{\mathbf{k}} \dots (\mathbf{e}_{\mathbf{N}})^{\mathbf{k}}]_{\mathbf{j}}$$

$$R = [\bar{\mathbf{E}}_1, \bar{\mathbf{E}}_2, \bar{\mathbf{E}}_3 \dots \bar{\mathbf{E}}_{\mathbf{N}}]^T \rightarrow R^{\mathbf{i}}_{\mathbf{j}} = [(\mathbf{e}^1)_{\mathbf{k}}, (\mathbf{e}^2)_{\mathbf{k}}, (\mathbf{e}^3)_{\mathbf{k}} \dots (\mathbf{e}^{\mathbf{N}})_{\mathbf{k}}]^T_{\mathbf{j}}$$

The relationship between $\mathbf{e}^n$ and $\mathbf{e}_n$ is very simple,

$$\mathbf{E}_n \equiv g'_{ni} \mathbf{e}_i \quad \rightarrow \quad \mathbf{e}^n = g'^{ni} \mathbf{e}_i \quad \text{and} \quad \mathbf{e}_n = g'_{ni} \mathbf{e}^i$$

For either contravariant or covariant indices (indices are not shown), g'^{ni} raises the label on $\mathbf{e}_i$, and inverting one finds that g'_{ni} lowers the label on $\mathbf{e}^i$. This fact makes things easy to remember. Using the fact that $\mathbf{e}_i = \partial'_i \mathbf{x}$, one has $g'^{ni} \mathbf{e}_i = g'^{ni} \partial'_i \mathbf{x} = \partial'^n \mathbf{x}$ so the above line can be expressed as

$$\mathbf{E}_n \equiv g'_{ni} \mathbf{e}_i \quad \rightarrow \quad \mathbf{e}^n = \partial'^n \mathbf{x} \quad \text{and} \quad \mathbf{e}_n = \partial'_n \mathbf{x}$$

The dot products are

$$\mathbf{e}_n \bullet \mathbf{e}_m = \overline{g}'_{nm} \quad \rightarrow \quad \mathbf{e}_n \bullet \mathbf{e}_m = g'_{nm} \quad = \partial'_n \mathbf{x} \bullet \partial'_m \mathbf{x} \quad |\mathbf{e}_n| = \sqrt{g'_{nn}} = h'_n$$

$$\mathbf{E}_n \bullet \mathbf{e}_m = \delta_{n,m} \quad \rightarrow \quad \mathbf{e}^n \bullet \mathbf{e}_m = \delta^n_m \quad = \partial'^n \mathbf{x} \bullet \partial'_m \mathbf{x}$$

$$\mathbf{E}_n \bullet \mathbf{E}_m = g'_{nm} \quad \rightarrow \quad \mathbf{e}^n \bullet \mathbf{e}^m = g'^{nm} \quad = \partial'^n \mathbf{x} \bullet \partial'^m \mathbf{x} \quad |\mathbf{e}^n| = \sqrt{g'^{nn}}$$

The "labels" on the base vectors behave in this dot product structure the same way that up and down "indices" behave. This is the motivation for $\mathbf{E}_n \rightarrow \mathbf{e}^n$. Thus, the three final equations can be regarded as the same equation $\mathbf{e}_n \bullet \mathbf{e}_m = g'_{nm}$ where we can raise either or both indices/labels to get the other equations. For example, $\mathbf{e}^n \bullet \mathbf{e}_m = g'^n_m = \delta^n_m$.

The dot product at the end of Section 6 (a) becomes

$$\mathbf{E}_n \bullet \mathbf{u}_m = R_{nm} \quad \rightarrow \quad \mathbf{e}^n \bullet \mathbf{u}_m = R^n_m \quad // = \langle \mathbf{e}^n | \mathbf{u}_m \rangle \text{ in bra-ket notation (App E (g))}$$

The matrix R^n_m is a "basis change matrix" between basis $\mathbf{u}_k$ and basis $\mathbf{e}^k$.

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:

$$(\mathbf{u}'_n)_i \quad \rightarrow \quad (\mathbf{u}_n)^i \quad // \text{ contravariant index } i$$

$$(\overline{\mathbf{u}}'_n)_i \quad \rightarrow \quad (\mathbf{u}_n)_i \quad // \text{ covariant index } i$$

$$(\mathbf{U}'_n)_i \quad \rightarrow \quad (\mathbf{u}^n)^i \quad // \text{ contravariant index } i$$

$$(\overline{\mathbf{U}}'_n)_i \quad \rightarrow \quad (\mathbf{u}^n)_i \quad // \text{ covariant index } i$$

$$(\mathbf{u}'_n)_i = R_{in} \quad \rightarrow \quad (\mathbf{u}'_n)^i = R^i_n = S_n^i \quad // \text{ contravariant index } i$$

$$(\bar{U}')_i = S_{ni} \quad \rightarrow \quad (u^n)_i = S^n_i \quad // \text{ covariant index } i$$

$$(U'_n)_i = S_{nk} g'_{ki} \quad \rightarrow \quad (u^n)^i = S^n_k g'^{ki} = S^{ni} \quad // \text{ contravariant index } i$$

$$R = [u'_1, u'_2, u'_3 \dots u'_N] \quad \rightarrow \quad R^i_j = [(u'_1)^k, (u'_2)^k, (u'_3)^k \dots (u'_N)^k]$$

$$S = [\bar{U}'_1, \bar{U}'_2, \bar{U}'_3 \dots \bar{U}'_N]^T \quad \rightarrow \quad S^i_j = [(\bar{U}'^1)_k, (\bar{U}'^2)_k, (\bar{U}'^3)_k \dots (\bar{U}'^N)_k]^T$$

$$U'_n \equiv g_{ni} u'_i \quad \rightarrow \quad u'^n = g^{ni} u'_i \quad \text{and} \quad u'_n = g_{ni} u'^i$$

$$u'_n \bullet u'_m = \bar{g}_{nm} \quad \rightarrow \quad u'^n \bullet u'^m = g_{nm}$$

$$U'_n \bullet u'_m = \delta_{n,m} \quad \rightarrow \quad u'^n \bullet u'^m = \delta^n_m$$

$$U'_n \bullet U'_m = g_{nm} \quad \rightarrow \quad u'^n \bullet u'^m = g^{nm}$$

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	e'_n	u_n	$(e'_n)_i = \delta_{n,i}$	$(u_n)_i = \delta_{n,i}$
dual partners to the above	E'_n	U_n	$(E'_n)_i = g'_{ni}$	$(U_n)_i = g_{ni}$
tangent base vectors	u'_n	e_n	$(u'_n)_i = R_{in}$	$(e_n)_i = S_{in}$
reciprocal base vectors	U'_n	E_n	$(U'_n)_i = g'^{ia} S_{na}$ $= g_{na} R_{ia}$	$(E_n)_i = g_{ia} R_{na}$ $= g'^{ia} S_{ia}$

which translates into this $\rightarrow$:

	<u>x'-space</u>	<u>x-space</u>		
axis-aligned basis vectors	e'_n	u_n	$(e'_n)^i = \delta_n^i$	$(u_n)^i = \delta_n^i$
dual partners to the above	e'^n	u^n	$(e'^n)^i = g'^{ni}$	$(u^n)^i = g^{ni}$
tangent base vectors	u'_n	e_n	$(u'_n)^i = R^i_n$	$(e_n)^i = S^i_n = R_n^i$
reciprocal base vectors	U'_n	E_n	$(U'_n)^i = g'^{ia} S^n_a$ $(u'^n)_i = S^n_i$	$(e^n)^i = g^{ia} R^n_a$ $(e^n)_i = R^n_i$

x-space expansions

The x-space expansions of Section 6 (f) were

$V = V_1 u_1 + V_2 u_2 + \dots$	$= \sum_n V_n u_n$	where $U_n \bullet V = V_n$	$U_n = g_{ni} u_i$
$V = \bar{V}_1 U_1 + \bar{V}_2 U_2 + \dots$	$= \sum_n \bar{V}_n U_n$	where $u_n \bullet V = \bar{V}_n$	
$V = V'_1 e_1 + V'_2 e_2 + \dots$	$= \sum_n V'_n e_n$	where $E_n \bullet V = V'_n$	$E_n = g'^{ni} e_i$
$V = \bar{V}'_1 E_1 + \bar{V}'_2 E_2 + \dots$	$= \sum_n \bar{V}'_n E_n$	where $e_n \bullet V = \bar{V}'_n$	

and they now become $\rightarrow$:

$$\begin{array}{llll}
 \mathbf{V} = V^1 \mathbf{u}_1 + V^2 \mathbf{u}_2 + \dots & = \sum_n V^n \mathbf{u}_n & \text{where } \mathbf{u}^n \bullet \mathbf{V} = V^n & \mathbf{u}^n = g^{ni} \mathbf{u}_i \\
 \mathbf{V} = V_1 \mathbf{u}^1 + V_2 \mathbf{u}^2 + \dots & = \sum_n V_n \mathbf{u}^n & \text{where } \mathbf{u}_n \bullet \mathbf{V} = V_n & \\
 \mathbf{V} = V'^1 \mathbf{e}_1 + V'^2 \mathbf{e}_2 + \dots & = \sum_n V'^n \mathbf{e}_n & \text{where } \mathbf{e}^n \bullet \mathbf{V} = V'^n & \mathbf{e}^n = g'^{ni} \mathbf{e}_i \\
 \mathbf{V} = V'_1 \mathbf{e}'_1 + V'_2 \mathbf{e}'_2 + \dots & = \sum_n V'_n \mathbf{e}'_n & \text{where } \mathbf{e}'_n \bullet \mathbf{V} = V'_n &
 \end{array}$$

x'-space expansions

Similarly, the x'-space expansions of Section 6 (f) were

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

and they now become $\rightarrow$:

$$\begin{array}{llll}
 \mathbf{V}' = V'^1 \mathbf{e}'_1 + V'^2 \mathbf{e}'_2 + \dots & = \sum_n V'^n \mathbf{e}'_n & \text{where } \mathbf{e}'^n \bullet \mathbf{V}' = V'^n & \mathbf{e}'^n = g'^{ni} \mathbf{e}'_i \\
 \mathbf{V}' = V'_1 \mathbf{e}'^1 + V'_2 \mathbf{e}'^2 + \dots & = \sum_n V'_n \mathbf{e}'^n & \text{where } \mathbf{e}'_n \bullet \mathbf{V}' = V'_n & \\
 \mathbf{V}' = V^1 \mathbf{u}'_1 + V^2 \mathbf{u}'_2 + \dots & = \sum_n V^n \mathbf{u}'_n & \text{where } \mathbf{u}'^n \bullet \mathbf{V}' = V^n & \mathbf{u}'^n = g^{ni} \mathbf{u}'_i \\
 \mathbf{V}' = V_1 \mathbf{u}'^1 + V_2 \mathbf{u}'^2 + \dots & = \sum_n V_n \mathbf{u}'^n & \text{where } \mathbf{u}'_n \bullet \mathbf{V}' = V_n &
 \end{array}$$

Summary of all expansions:

Using implied sum notation, we can now summarize the eight expansions above, plus the unit vector expansion onto $\hat{\mathbf{e}}_n$, on just two lines :

$$\begin{array}{ll}
 \mathbf{V} = V^n \mathbf{u}_n = V_n \mathbf{u}^n = V'^n \mathbf{e}_n = V'_n \mathbf{e}'^n = \mathcal{V}^n \hat{\mathbf{e}}_n & // \text{ x-space expansions, } \mathcal{V}^n = h_n V^n \\
 \mathbf{V}' = V'^n \mathbf{e}'_n = V'_n \mathbf{e}'^n = V^n \mathbf{u}'_n = V_n \mathbf{u}'^n & // \text{ x'-space expansions}
 \end{array}$$

In all cases one sees a tilted index summation where one index is a vector index and the other is a basis vector label. Half the forms shown above can be obtained from the others by just "reversing the tilt". The power of the Standard Notation makes itself felt in relations like these.

Due to this tilt situation, sometimes a basis like $\mathbf{e}_n$ appearing in $\mathbf{V} = V'^n \mathbf{e}_n$ is called a "covariant basis" while the basis $\mathbf{e}^n$ appearing in $\mathbf{V} = V'_n \mathbf{e}^n$ is called a "contravariant basis".

Corresponding expansions of higher rank tensors are presented in section (w) below.

If $\mathbf{V}$ is a tensor density of weight W (see Appendix D and E) the rule for adjusting the above expansions is to make the replacement $V^n \rightarrow J^W V^n$ and $V'_n \rightarrow J^W V'_n$ where J is the Jacobian of Section 5 (k).

(t) Comment on Covariant versus Contravariant

Consider this expansion for a vector $\mathbf{V}$ in x -space,

$$\mathbf{V} = V^n \mathbf{b}_n \quad V^n = \mathbf{V} \bullet \mathbf{b}^n$$

where $\mathbf{b}_n$ is some basis having dual basis $\mathbf{b}^n$ where as usual $\mathbf{b}_n \bullet \mathbf{b}^m = \delta_n^m$. Imagine taking $V^n \rightarrow V'^n = R^n_m V^m$ and $\mathbf{b}_i \rightarrow \mathbf{b}'_i = Q_i^j \mathbf{b}_j$. What Q would cause the following to be true?

$$\mathbf{V} = V^n \mathbf{b}_n = V'^n \mathbf{b}'_n$$

In other words, how does one transform that basis $\mathbf{b}_n$ 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))

$$V'^n \mathbf{b}'_n = [R^n_m V^m] [R_n^j \mathbf{b}_j] = (R^n_m R_n^j) V^m \mathbf{b}_j = \delta_m^j V^m \mathbf{b}_j = V^j \mathbf{b}_j = V^n \mathbf{b}_n$$

Compare then the transformation of V^n with that of the basis $\mathbf{b}_n$:

$$\begin{aligned} V'^n &= R^n_m V^m \\ \mathbf{b}'_n &= R_n^m \mathbf{b}_m \end{aligned}$$

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 $\mathbf{b}_m$.

If one starts over with V_n components and the $\mathbf{b}^n$ "dual" (reciprocal) expansion vectors and asks for a solution to this corresponding problem,

$$\mathbf{V} = V_n \mathbf{b}^n = V'_n \mathbf{b}'^n$$

one finds not surprisingly that the dual basis must vary as $\mathbf{b}'^n = R^n_m \mathbf{b}^m$ and then one has

$$\begin{aligned} V'_n &= R_n^m V_m \\ \mathbf{b}'^n &= R^n_m \mathbf{b}^m \end{aligned}$$

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 $\mathbf{b}^m$ 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 $\mathbf{b}_n$ one has

$$\begin{aligned} V'_n &= R_n^m V_m \\ \mathbf{b}'_n &= R_n^m \mathbf{b}_m \end{aligned}$$

If the basis $\mathbf{b}_m$ is varied as shown here, then the dual basis $\mathbf{b}^m$ varies as shown above and $\mathbf{V}$ remains invariant. Comparing now the way the V_n transform with the way the basis vectors $\mathbf{b}_m$ 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 $\mathbf{b}_m$.

(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

$$Q'_a{}^d(x') = H'_{ab}(x') T'^b{}_c(x') B'^d(x')$$

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"; continuum mechanics uses the term "frame-indifferent".)

If the objects Q , H , T and B are *tensors* under F , then covariance of any valid tensor equation like the one shown above *is guaranteed!!*

The reason is that, once the contracted indices on the two sides are ignored according to the "contraction neutralization rule", the objects on the two sides of the equation have the same indices which are of the same type, so both sides are tensors of the same type, and therefore both sides transform from x -space to x' -space in the same way. If one starts, for example, with the primed equation and installs the known transformations for all the pieces, one ends up with the unprimed equation.

If this explanation is not convincing, a brute force demonstration can perhaps help out. The following is also a good exercise is using the two tilt forms of the R matrix. Recall from 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 :

$$Q'_a{}^d(x') = H'_{ab}(x') T'^b{}_c(x') B'^d(x') \quad (*) \quad // \text{ } x'\text{-space equation}$$

$$[R_a{}^{a'} R^d{}_{d'} R_c{}^{c'} Q_{a'}{}^{d'}{}_{c'}(x)] = [R_a{}^{a'} R_b{}^{b'} H_{a'}{}^{b'}(x)] [R^b{}_{b'} R_c{}^{c'} T^{b'}{}_{c'}(x)] [R^d{}_{d'} B^{d'}(x)]$$

$$= R_a^{a'} R_d^{d'} R_c^{c'} (R_b^{b'} R_{b''}^{b''}) H_{a'b'}(x) T^{b''}_{c'}(x) B^{d'}(x)$$

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, using the fact that Q is a tensor to replace Q' on the left side of (*),

$$(R_a^{a'} R_d^{d'} R_c^{c'}) Q_{a'}^{d'}(x) = (R_a^{a'} R_d^{d'} R_c^{c'}) H_{a'b'}(x) T^{b'}_{c'}(x) B^{d'}(x)$$

Now apply the Cancellation Rule of section (r) three times to conclude that

$$Q_{a'}^{d'}(x) = H_{a'b'}(x) T^{b'}_{c'}(x) B^{d'}(x)$$

and then remove all primes to get

$$Q_a^d(x) = H_{ab}(x) T^b_c(x) B^d(x) \quad // \text{ x-space equation}$$

Thus it has been shown that, if all the objects transform as tensors, the equation is covariant.

Tensor density equations are also covariant. As discussed in Appendix D, a tensor density of weight W is a generalization of a tensor which has the same transformation rule as a regular tensor, but there is an extra factor of J^{-W} on the right hand side of the rule, where J is the Jacobian $J = \det S$. For example,

$$Q'^d_{a'}(x') = J^{-W_Q} R_a^{a'} R_d^{d'} R_c^{c'} Q_{a'}^{d'}(x)$$

would indicate that Q was a tensor density of weight W_Q . If $W_Q = 0$, then Q is a regular tensor. With this definition in mind, it is easy to generalize the notion of a "covariant equation" to include tensor densities. Consider some arbitrary tensor equation which we represent by our example above,

$$Q_a^d(x) = H_{ab}(x) T^b_c(x) B^d(x)$$

Suppose all four objects Q, H, T, B are tensor densities with weights W_Q, W_H, W_T, W_B . If the four objects Q, H, T, B are tensor densities, and if the up/down free indices match on both sides (the non-contracted indices), and if $W_Q = W_H + W_T + W_B$, then this is a "valid tensor density equation" and covariance is guaranteed, so it follows that

$$Q'^d_{a'}(x') = H'_{ab}(x') T'^b_{c'}(x') B'^d(x') .$$

It is trivial to edit the above proof by just adding weight factors in the right places and then of course they cancel out on the two sides.

Examples of covariant tensor equations: 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. Maxwell's two *inhomogeneous* equations (that is, the two with sources) are, in SI units where $\epsilon_0\mu_0 = 1/c^2$,

$$\partial_\nu F^{\mu\nu} = \mu_0 J^\mu \quad \text{with } J^\mu = (c\rho, \mathbf{J})$$

while the two homogeneous equations become

$$\partial_\alpha F_{\mu\nu} + \partial_\mu F_{\nu\alpha} + \partial_\nu F_{\alpha\mu} = 0 \quad \text{or} \quad \partial_\alpha F_{\mu\nu} + \text{cyclic} = 0.$$

One can see that each of these equations involves only tensors and we expect that in x' -space these equations will take the form

$$\begin{aligned} \partial'_\nu F'^{\mu\nu} &= \mu_0 J'^\mu & \text{with } J'^\mu &= (c\rho', \mathbf{J}') \\ \partial'_\alpha F'_{\mu\nu} + \partial'_\mu F'_{\nu\alpha} + \partial'_\nu F'_{\alpha\mu} &= 0 & \text{or} \quad \partial'_\alpha F'_{\mu\nu} + \text{cyclic} &= 0 \end{aligned}$$

Objects like $\partial_\nu F^{\mu\nu}$ and $\partial_\alpha F_{\mu\nu}$ are true rank-3 tensors because the transformation F is linear.

Covariance of tensor equations involving derivatives with *non-linear* F . A tensor equation which involves derivatives of tensors is non-covariant under transformations F which are non-linear. The reason is that the derivative of a tensor is, in that case, not a tensor, as shown in the next section. Such tensor equations *can be made* covariant by replacement of all derivatives by covariant derivatives (which are indicated by a semicolon). In general relativity, this is known as the Principle of General Covariance (Weinberg p 106). A simple example is the tensor equation $g_{ab;c} = 0$ (Appendix F (i)). Examples relating to the transformation from Cartesian to curvilinear coordinates appear in Section 15.

(v) The Christoffel Business: covariant derivatives

This subject is treated in full detail in Appendix F, but here we provide some motivation. It should be noted that a normal derivative is sometimes written $\partial_a V_b = V_{b,a}$ with a comma, whereas the covariant derivative discussed below is written $V_{b;a}$ with a semicolon.

When a transformation F is non-linear, the matrix R^a_b is a function of $\mathbf{x}$. 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 $\phi'(\mathbf{x}') = \phi(\mathbf{x})$,

$$(\partial'_a \phi') = (R^d_a \partial_d) \phi = R^d_a (\partial_d \phi).$$

In special relativity, for example, where transformations are linear, $\partial_d R^c_b = 0$, there is no second term, and the object $\partial_a V_b$ transforms as a covariant rank-2 tensor,

$$(\partial'_a V'_b) = R^d_a R^c_b (\partial_d V_c) , \quad // F \text{ is a linear transformation}$$

but in the general case the second term is present, so $\partial_d V_c$ fails to transform as a rank-2 covariant tensor. In this case, one defines a certain "covariant derivative" which itself has an extra piece

$$V_{b;a} \equiv \partial_a V_b - \Gamma^k_{ab} V_k \quad \Rightarrow \quad V_{d;c} \equiv \partial_c V_d - \Gamma^k_{cd} V_k$$

where Γ^c_{ab} is a certain function of the metric tensor g . One then finds that

$$V'_{b;a} = R^d_a R^c_b V_{d;c} \quad // \text{ the notation } \nabla_c V_d \equiv V_{d;c} \text{ is also commonly used}$$

or

$$[\partial'_a V'_b - \Gamma'^k_{ab} V'_k] = R^d_a R^c_b [\partial_d V_c - \Gamma^k_{cd} V_k] \quad (*)$$

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^c_{ab} (sometimes called the "Christoffel connection") is given by

$$\Gamma^c_{ab} \equiv \{ab, c\} \equiv \left\{ \begin{matrix} c \\ ab \end{matrix} \right\} \equiv g^{cd} [ab, d] = \frac{1}{2} g^{cd} (\partial_a g_{bd} + \partial_b g_{ad} - \partial_d g_{ab}) \quad // \text{ Christoffel 2nd kind}$$

$$\Gamma_{dab} \equiv [ab, d] \equiv \frac{1}{2} (\partial_a g_{bd} + \partial_b g_{ad} - \partial_d g_{ab}) \quad // \text{ Christoffel 1st kind}$$

and this is where the various "Christoffel symbols" come into play. In general relativity and elsewhere, Γ^c_{ab} is known as the "affine connection" which represents the effect of "curved space" appearing as a force which acts on a mass (that is to say, a gravitational force), see Section 5(n).

Warning: There is a differently defined version of Γ^c_{ab} floating around in the literature. The version used above and everywhere in this document is that of Weinberg and is the most common form.

The derivative of *any* tensor field other than a scalar field shows this same complication when the underlying transformation F is non-linear. For example, $\partial_a g_{bd}(x)$ does not transform as a rank-3 tensor,

$$\partial'_a g'_{bd}(x') = (R_a^d \partial_d)(R_b^{b'} R_d^{d'} g_{b'd'}) = R_b^{b'} R_d^{d'} (\partial_d g_{b'd'}) + \text{other terms}$$

and therefore neither of the Christoffel symbols Γ_{dab} or Γ^c_{ab} transforms as a tensor in this case.

See Appendix F for more detail.

(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 **A** can be expanded in two interesting ways :

$$\mathbf{A} = \sum_n A^n \mathbf{u}_n$$

A^n are the contravariant components of **A** in x-space

$$\mathbf{A} = \sum_n A'^n \mathbf{e}_n$$

A'^n are the contravariant components of **A** in x'-space

In the first, $\mathbf{u}_n$ are axis-aligned basis vectors, and in the second $\mathbf{e}_n$ are the tangent base vectors. If **A** is instead a tensor of rank n, these expansions are replaced by

$$\mathbf{A} = \sum_{ijk\dots} A^{ijk\dots} (\mathbf{u}_i \otimes \mathbf{u}_j \otimes \mathbf{u}_k \dots) \quad A^{ijk\dots} \text{ are the contravariant components of } \mathbf{A} \text{ in x-space}$$

$$\mathbf{A} = \sum_{ijk\dots} A'^{ijk\dots} (\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \dots) \quad A'^{ijk\dots} \text{ are the contravariant components of } \mathbf{A} \text{ in x'-space}$$

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

$$\mathbf{A} = \sum_{ijk\dots} A^{ijk\dots} \mathbf{u}_i \mathbf{u}_j \mathbf{u}_k \dots \quad A^{ijk\dots} \text{ are the contravariant components of } \mathbf{A} \text{ in x-space}$$

$$\mathbf{A} = \sum_{ijk\dots} A'^{ijk\dots} \mathbf{e}_i \mathbf{e}_j \mathbf{e}_k \dots \quad A'^{ijk\dots} \text{ are the contravariant components of } \mathbf{A} \text{ in x'-space} .$$

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 Appendix 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

$$\mathbf{A} = \sum_{ij} A^{ij} \mathbf{u}_i \mathbf{u}_j^T \quad A^{ij} \text{ are the contravariant components of } \mathbf{A} \text{ in x-space}$$

$$\mathbf{A} = \sum_{ij} A'^{ij} \mathbf{e}_i \mathbf{e}_j^T \quad A'^{ij} \text{ are the contravariant components of } \mathbf{A} \text{ in x'-space} .$$

Appendix E (g) promotes the interpretation of a rank-2 tensor A as an operator in a Hilbert space, where the matrices A^{ij} and A'^{ij} are matrices *associated with* the operator A in different bases,

$$A^{nm} = \langle u^n | A | u^m \rangle = \text{the } x\text{-space components of tensor } A$$

$$A'^{nm} = \langle e^n | A | e^m \rangle = \text{the } x'\text{-space components of tensor } A$$

As Appendix E shows, these two matrices are related to each other by a similarity transformation $A' = R A R^{-1}$.

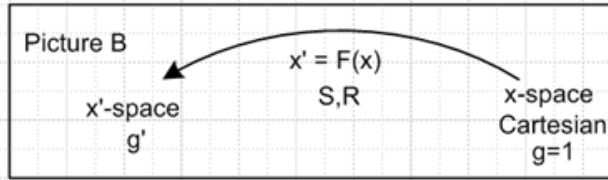
These expansion methods are used in Appendices G and H to derive curvilinear expressions for two objects that play a role in continuum mechanics, $(\nabla \mathbf{v})$ and $\text{div}(\mathbf{T})$ (where $\mathbf{T}$ is a tensor).

8. Transformation of Differential Length, Area and Volume

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 quite lengthy, a brief overview is in order:

Overview

The transformation of differential length, area, and volume is first framed in terms of the mapping of an orthogonal differential N-piped in x' -space to a skewed differential N-piped in x -space. The N-piped in x' -space has axis-aligned edges of length dx'^m , while the N-piped in x -space has edges $\mathbf{e}_n dx'^m$ where $\mathbf{e}_n$ are the tangent base vectors introduced in Section 3. We want to learn what happens to the edges, face areas and volume as one differential N-piped is mapped into the other by the curvilinear coordinate transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. After solving this problem, we go on to consider the transformations of arbitrary differential vectors, areas and volume.

Section (a): The differential N-piped mapping

The differential N-piped mapping is described and various symbols are defined.

Section (b): Properties of the finite N-piped spanned by the $\mathbf{e}_n$ in x -space

Results from Appendix B concerning *finite* N-piped geometric properties are quoted. Certain definition changes are made to make the formulas suitable for tensor analysis. The purpose of the lengthy Appendix B is to lend credence to the general formulas for elements of area and volume in N dimensions.

Section (c): Back to the differential N-piped mapping: how edges, areas and volume transform

1. Setup. The finite N-piped edges $\mathbf{e}_n$ are scaled by curvilinear coordinate variations dx'^m to create a differential N-piped in x -space having edges $(\mathbf{e}_n dx'^m)$.

2. Edge Transformation. The edges $\mathbf{e}_n$ of the x -space N-piped map into axis-aligned edges $\mathbf{e}'_n$ in x' -space.

3. Area Transformation. Tensor density notions as presented in Appendix D are used here.

4. Volume Transformation. The volume transformation is computed several different ways.

5. Covariant Magnitudes. These are $|dx'^{(n)}|$, $|dA'^m|$ and $|dV'|$ in the Curvilinear View of x' -space.

6. Two Theorems. (1) $g' / h'^2_n = g'^{nn} g' = \text{cof}(g'_{nn})$ and (2) $|\Pi^{\mathbf{x}}_{i \neq n} \mathbf{e}_i| = \sqrt{\text{cof}(g'_{nn})}$.

7. Cartesian-View Magnitude Ratios. Appropriate for the continuum mechanics application.

8. Nested Cofactor Formulas. Transformation of N-piped areas of dimension N-2 and lower.

9. Transformation of arbitrary differential vectors, areas and volume. Having built confidence with the general form of vector area and volume expressions in the N-piped case, the N-piped is jettisoned and formulas for the transformation of arbitrary vectors, areas and volume are derived.

10. Concatenation of Transformations. What happens to the transformation of vectors, areas and volumes when two transformations are concatenated? One result is that $J = J_1 J_2$.

Examples of area magnitude transformation for $N = 2, 3, 4$

Example 2: Spherical Coordinates: area patches

Section (d): Transformation of Differential Volume applied to Integration

The volume transformation obtained in Section (c) is related to the traditional notion of the Jacobian changing the "measure" in an integration. The "Jacobian Integration Rule" can then be expressed as a distributional equation.

Section (e): Interpretations of the Jacobian

(a) The differential N-piped mapping

Section 3 (a) above considered the following situation ($dx'^n > 0$) :

$$\begin{aligned} d\mathbf{x}'^{(n)} &= \mathbf{e}'_n dx'^n && \text{x'-space axis-aligned differential vector, and } (\mathbf{e}'_n)^i = \delta_n^i \\ d\mathbf{x}^{(n)} &= \mathbf{e}_n dx^n && \text{x-space mapping of the above vector under } F^{-1} \text{ or } R^{-1} \\ d\mathbf{x}'^{(n)} &= R(\mathbf{x}) d\mathbf{x}^{(n)} && \text{relation of the two differential vectors (contravariant rule)} \end{aligned}$$

A superscript (n) on the differentials makes clear there is no implied sum on n.

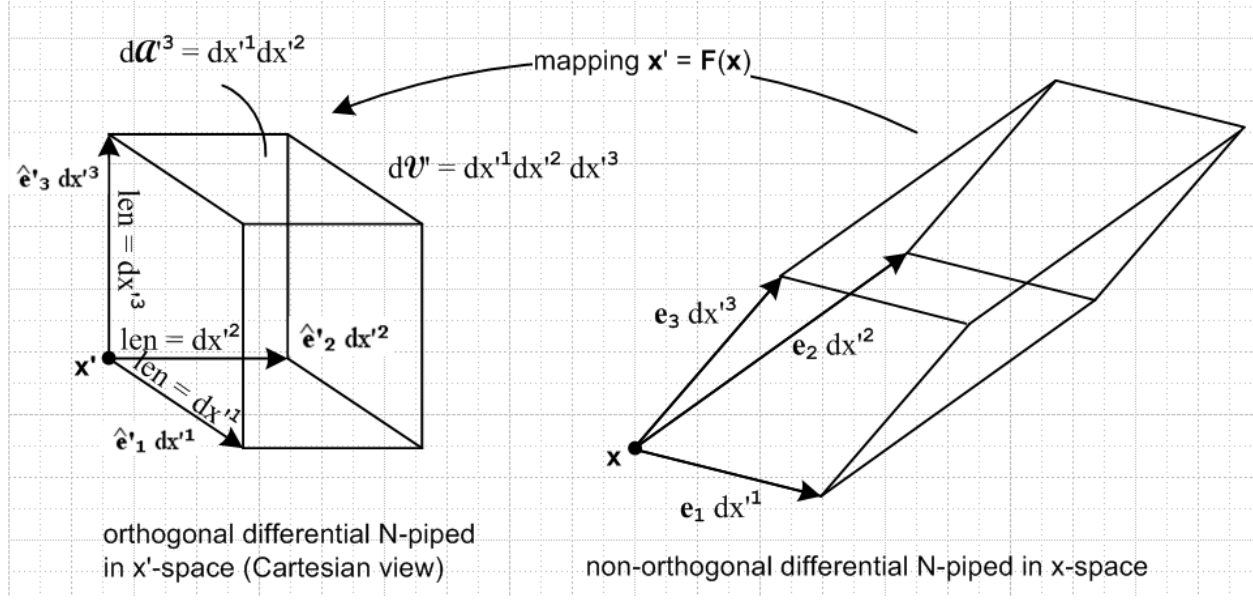
The vectors $d\mathbf{x}^{(n)}$ span a differential N-piped in x-space, while the $d\mathbf{x}'^{(n)}$ span a corresponding differential N-piped in x'-space. The two N-pipeds are related by the mapping $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. Since the regions are differentially small, this mapping is the same as the linearized mapping $d\mathbf{x}' = R d\mathbf{x}$.

The metric tensor in x-space will be taken to be $g = 1$, so it is a Cartesian space.

As discussed at the end of Section 5 (a), the x'-space N-piped can be viewed in (at least) two ways depending on how the metric tensor g' is set. For a continuum mechanics flow application, one sets $g' = 1$ and this gives the Cartesian View of the x'-space N-piped. For such flows dot products and magnitudes of vectors like $d\mathbf{x}^{(n)}$ are not invariant under the transformation. For our curvilinear coordinates application, however, we set $g' = RgR^T = RR^T$ and this causes vector dot products and magnitudes to be invariant and we can talk about such objects as being tensorial scalars. This is the Curvilinear View of x'-space.

When $g' \neq 1$, it is impossible to accurately represent the Curvilinear-View picture of the x'-space N-piped as a drawing in physical space (for $N=3$). This subject is discussed for a sample 2D system in Appendix C (e). Although the basis vectors $(\mathbf{e}'_n)^i = \delta_n^i$ in x'-space are *always* axis-aligned, they are only orthogonal for an orthogonal coordinate system, since $\mathbf{e}'_n \bullet \mathbf{e}'_m = \mathbf{e}_n \bullet \mathbf{e}_m = g'_{nm}$. Nevertheless, even for a non-orthogonal system we draw the axes as if they were orthogonal, which at least provides a representation of the notion of "axis aligned" basis vectors. For $N > 3$ one at least imagine this kind of drawing.

Due to these graphical difficulties, in the drawing below the Cartesian View of x'-space is shown. Since $g' = 1$ for this situation, the axis-aligned basis vectors $\mathbf{e}'_n$ are in fact unit vectors $\hat{\mathbf{e}}'_n$ and are orthogonal, so the picture becomes at least comprehensible:



The orthogonal Cartesian-View N-piped allows *visualization* of these curvilinear coordinate variations, all $dx'^k} > 0$,

$$\begin{aligned} d\mathcal{L}^n &\equiv dx'^n \\ d\mathcal{A}^n &\equiv \prod_{i \neq n} dx'^i \\ dV' &\equiv \prod_i dx'^i = d\mathcal{A}^n d\mathcal{L}^n \end{aligned}$$

For example, for $N=3$ one would have

$$\begin{aligned} d\mathcal{L}^1 &\equiv dx'^1 & d\mathcal{A}^1 &= dx'^2 dx'^3 & d\mathcal{A}^2 &= dx'^3 dx'^1 \\ d\mathcal{A}^3 &= dx'^1 dx'^2 & dV' &= dx'^1 dx'^2 dx'^3 \end{aligned}$$

The Cartesian-View x' -space N-piped is always orthogonal because the $(\mathbf{e}'_n)$ are orthonormal axis-aligned unit vectors (since $g'=1$). In contrast, the x -space N-piped is typically rotated and possibly skewed as well (if the coordinates x'^i describe a non-orthogonal coordinate system). The transformation F and its linearized version R map the skewed x -space N-piped into the orthogonal x' -space N-piped. As one moves around in x -space so that point $\mathbf{x}$ changes, the picture on the left above keeps its shape, just translating itself to the new point $\mathbf{x}'$, but the picture on the right changes shape and volume because the vectors $\mathbf{e}_n(\mathbf{x})$ are functions of $\mathbf{x}$.

It is our goal to write expressions for edges, areas and volumes in these two spaces and to then show how these objects transform between the two spaces. To this end, we shall rely on work done in Appendix B which is summarized in the next section. Following that, we shall add to each edge a differential associated with that edge (such as $\mathbf{e}_r \rightarrow \mathbf{e}_r dr$ in spherical coordinates), and that will bring us back to the differential N-piped picture above.

(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:

$$\mathbf{A}^n = |\det(\mathbf{S}^a_b)| \mathbf{e}^n$$

$$\mathbf{A}^n = \sigma (-1)^{n-1} \prod_{i \neq n}^x \mathbf{e}_i$$

$$\mathbf{A}^n = \sigma (-1)^{n-1} \mathbf{e}_1 \times \mathbf{e}_2 \dots \times \mathbf{e}_N \quad // \mathbf{e}_n \text{ missing} \quad \sigma \equiv \text{sign}[\det(\mathbf{S}^a_b)] = \text{sign}[\det(\mathbf{R}^a_b)]$$

$$(\mathbf{A}^n)_i = \sigma (-1)^{n-1} \varepsilon_{iabc \dots x} (\mathbf{e}_1)^a (\mathbf{e}_2)^b \dots (\mathbf{e}_N)^x \quad // \mathbf{e}_n \text{ missing}$$

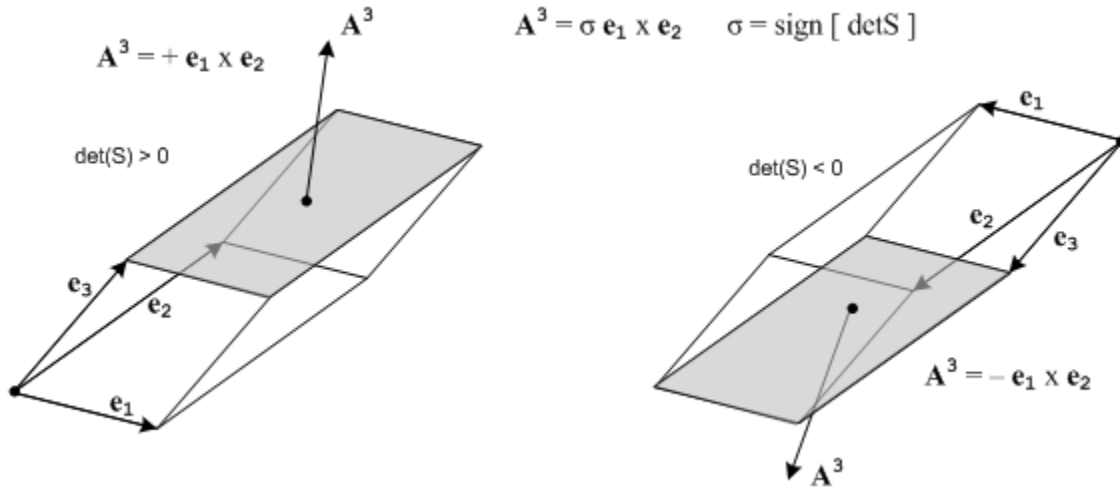
- The volume of the N-piped is given by (see Section 5 (k) concerning J)

$$V = |\det [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \dots \mathbf{e}_N]| = |\det(\mathbf{S}^a_b)| = g^{1/2} = |J|$$

The mapping picture above does not apply to a finite N-piped. The *finite* N-piped just discussed exists in x-space. One might ponder into what shape it maps in x'-space under the transformation F. In the case of spherical coordinates (Section 1 Example 2), *all* of x-space maps into a certain orthogonal "office building" in x'-space. A finite N-piped within x-space maps into some very complicated 3D shape within the office building which is bounded by curves which in general are not even coplanar. The point is that a finite N-piped in x-space does NOT map back into some nice orthogonal N-piped in x'-space and the picture drawn in the previous section does not apply. However, when differentials are added in the next section, *then*, since the mapped regions are very small, the mapping of the x-space differential N-piped is

in fact an "orthogonal" N-piped in x' -space. This is because for a tiny region near some point $\mathbf{x}$, the mapping between $d\mathbf{x}$ and $d\mathbf{x}'$ is described by matrix R .

Conventions for defining the area vector and volume. In Appendix B (and as shown above) the area vector $\mathbf{A}^n$ is defined so that the out-facing normal of the x -space N-piped's "far face n " is $\mathbf{A}^n$, regardless of the sign of $\det(S)$. This was done to simplify the computation of the flux of a vector field emerging from the N-piped in the geometric divergence calculation in Section 9. That calculation is then valid for either sign of $\det(S)$, a sign that we call σ . The following picture illustrates on the left an x -space 3-piped which is obtained by reverse-mapping the x' -space orthogonal 3-piped using an S which has $\det(S) > 0$. On the right one sees the x -space 3-piped that results for $S \rightarrow -S$. These two 3-pipeds are related by a parity inversion of all points through the origin. If the origin lies far away, these two N-pipeds lie far away from each other, a fact not illustrated in the picture:



Notice that $\mathbf{A}^3$ for the "far face 3" is out-facing in both cases.

This definition of the vector area is not suitable for the vector analysis we are about to undertake. Instead of the above situation, we will redefine $\mathbf{A}^3 = \mathbf{e}_1 \times \mathbf{e}_2$ for both pictures, and this will cause the $\mathbf{A}^3$ vector in the right picture to point up into the interior of the N-piped. This new definition allows us to interpret $\mathbf{A}^3 = \mathbf{e}_1 \times \mathbf{e}_2$ as a "valid vector equation" to which we may apply the ideas of covariance and tensor densities.

Notice that under a parity transformation, this newly defined $\mathbf{A}^3$ is a "pseudovector" which is one which does not reverse direction under a parity transformation, since $\mathbf{A}^3 = (-\mathbf{e}_1) \times (-\mathbf{e}_2)$. The subject of parity and handedness and the sign of $\det(S)$ is discussed more in Section 6 (i).

Here then are the expressions for $\mathbf{A}^n$ with this new definition, where the new forms are obtained from the previous ones by multiplying by $\sigma = \text{sign}(\det(S))$:

$$\mathbf{A}^n = \det(S^a_b) \mathbf{e}^n = J \mathbf{e}^n$$

$$\mathbf{A}^n = (-1)^{n-1} \mathbf{e}_1 \times \mathbf{e}_2 \dots \times \mathbf{e}_N \quad // \mathbf{e}_n \text{ missing}$$

$$(\mathbf{A}^n)_i = (-1)^{n-1} \varepsilon_{iabc\dots x} (e_1)^a (e_2)^b \dots (e_N)^x \quad // \mathbf{e}_n \text{ missing}$$

The second line shows that pseudovectors can only exist for an odd number of dimensions (such as $N=3$).

A similar redefinition of the volume will now be done. In Appendix B the volume is defined so as to be a positive number regardless of σ with the result $V = |\det(S)|$. We now redefine the volume by multiplication by σ , so that now $V = \det(S)$ which is of course a negative number when $\det(S) < 0$, which in turn means the $\mathbf{e}_n$ are forming a left-handed coordinate system as per Section 6 (i). So:

$$V = \det [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \dots \mathbf{e}_N] = \det(S^{\mathbf{a}}_{\mathbf{b}}) = J$$

(c) Back to the differential N-piped mapping: how edges, areas and volume transform

1. The Setup. If the edges of the finite N-piped described above are scaled by positive differentials $dx'^n > 0$, the result is a differential N-piped in x-space which has the properties listed above with the following edges and areas and volume:

$$d\mathbf{x}^{(n)} = \mathbf{e}_n dx'_n \quad // \text{ edges}$$

$$d\mathbf{A}^n = J \mathbf{e}^n (\Pi_{i \neq n} dx'^i) \quad // \text{ areas}$$

$$d\mathbf{A}^n = (-1)^{n-1} (dx'_1 \mathbf{e}_1) \times (dx'_2 \mathbf{e}_2) \dots \times (dx'_N \mathbf{e}_N) \quad // \mathbf{e}_n \text{ missing from cross product}$$

$$= (-1)^{n-1} \mathbf{e}_1 \times \mathbf{e}_2 \dots \times \mathbf{e}_N (\Pi_{i \neq n} dx'^i) \quad // \mathbf{e}_n \text{ missing from cross product}$$

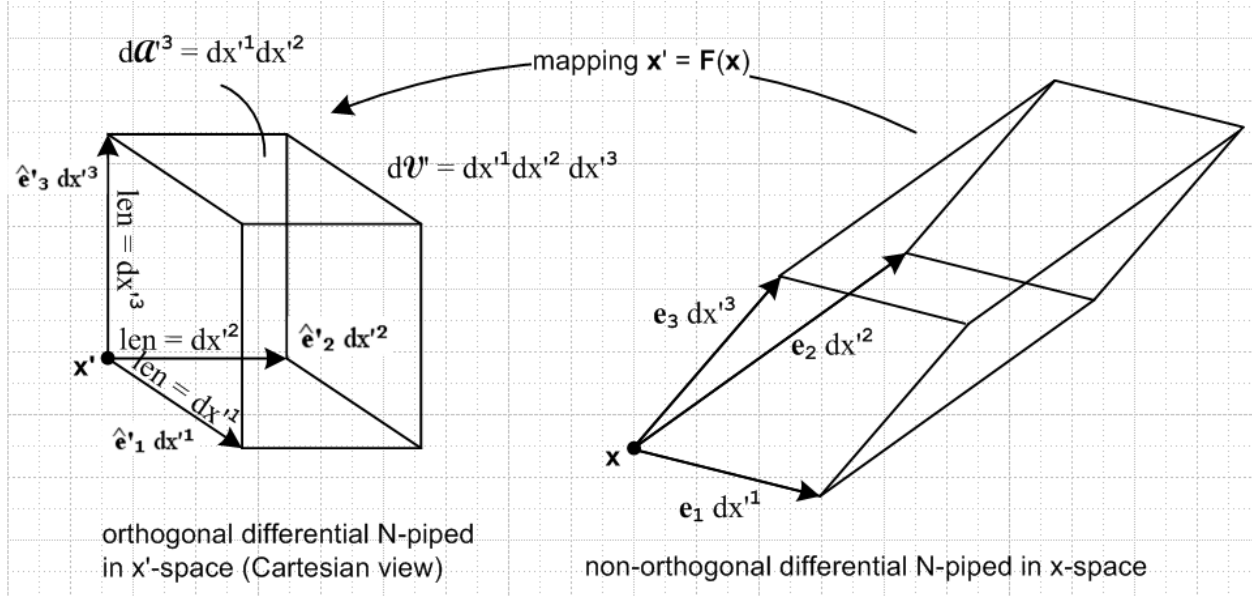
$$= (-1)^{n-1} (\Pi_{i \neq n}^{\mathbf{x}} \mathbf{e}_i) (\Pi_{i \neq n} dx'^i) \quad // \text{ shorthand of Appendix A (i)}$$

$$(d\mathbf{A}^n)_i = (-1)^{n-1} \varepsilon_{iabc \dots x} (e_1)^a (e_2)^b \dots (e_N)^x (\Pi_{i \neq n} dx'^i) \quad // \mathbf{e}_n \text{ factor and index missing}$$

$$dV = \det [dx'_1 \mathbf{e}_1, dx'_2 \mathbf{e}_2, dx'_3 \mathbf{e}_3 \dots dx'_N \mathbf{e}_N] = \det [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \dots \mathbf{e}_N] (\Pi_i dx'^i) = J \det(S^{\mathbf{a}}_{\mathbf{b}})$$

$$= \varepsilon_{abc \dots x} (dx'_1 \mathbf{e}_1)^a (dx'_2 \mathbf{e}_2)^b \dots (dx'_N \mathbf{e}_N)^x$$

where $d\mathbf{A}^n$ and dV obtained from $\mathbf{A}^n$ and V according to the new definitions described above. These equations apply to the right side of the N-piped mapping picture which is replicated here



For spherical coordinates, the N-piped on the right would be spanned by these vectors (Sec. 3, Ex. 2)

$$\mathbf{e}_r dr = \hat{r} dr, \quad \mathbf{e}_\theta d\theta = r \hat{\theta} d\theta, \quad \mathbf{e}_\phi d\phi = r \sin\theta \hat{\phi} d\phi$$

2. Edge Transformation. The edge $d\mathbf{x}^{(n)}$ we know transforms as a tensorial vector under transformation F , so

$$d\mathbf{x}'^{(n)} = R d\mathbf{x}^{(n)} \quad \text{where} \quad d\mathbf{x}^{(n)} = \mathbf{e}_n dx'^n$$

Evaluation gives

$$\begin{aligned} [d\mathbf{x}'^{(n)}]^i &= R^i_j (\mathbf{e}_n)^j dx'^n = R^i_j S^j_n dx'^n = (RS)^i_n dx'^n = \delta^i_n dx'^n \\ \Rightarrow d\mathbf{x}'^{(n)} &= \mathbf{e}'_n dx'^n \quad \text{since } (\mathbf{e}'_n)^i = \delta_n^i \end{aligned}$$

so this contravariant edge points in the n axis direction in x' -space.

3. Area Transformation. How does $d\mathbf{A}^n$ transform under F ? Looking at the component form stated above

$$(d\mathbf{A}^n)_i = (-1)^{n-1} \varepsilon_{iabc\dots x} (e_1)^a (e_2)^b \dots (e_N)^x \quad (\Pi_{i \neq n} dx'^i) \quad // \mathbf{e}_n \text{ factor and index missing}$$

it is seen that $d\mathbf{A}^n$ is a combination of tensor objects like $\varepsilon_{iabc\dots x}$ and $(e_2)^b$. As discussed in Appendix D, a vector density of weight 0 is an ordinary vector such as $\mathbf{e}_2$ or $d\mathbf{x}$. The ε tensor (rank- N) is a tensor density of weight -1. In forming more complicated tensor objects, the rule is that one adds the weights of the objects being combined. Therefore, one may conclude that the object $d\mathbf{A}^n$ is a vector density of weight -1. This has the *immediate implication* that $d\mathbf{A}^n$ transforms under F according to the rule

$$d\mathbf{A}^n = J R d\mathbf{A}^n \quad \text{or} \quad (d\mathbf{A}^n)^i = J R^i_j (d\mathbf{A}^n)^j \quad // J^{-\mathbf{w}} = J^{-(-1)} = J$$

where $J = \det(S)$ is the Jacobian of Section 5 (k). Moreover, the above equation for $(d\mathbf{A}^n)_i$ is a "valid tensor density equation" as per Section 7 (u) and is therefore covariant. This means that in x' -space the equation has the exact same form, but tensor objects are primed (the dx'^i are constants),

$$(d\mathbf{A}^n)_i = (-1)^{n-1} \varepsilon'_{iabc\dots x} (e'_1)^a (e'_2)^b \dots (e'_N)^x (\Pi_{i \neq n} dx'^i) \quad // e'_n \text{ factor and index missing}$$

Insertion of $\varepsilon'_{iabc\dots x} = J^2 \varepsilon_{iabc\dots x}$ (Appendix D (e) 3) and $(e'_n)^i = \delta_n^i$ then gives

$$\begin{aligned} (d\mathbf{A}^n)_i &= (-1)^{n-1} J^2 \varepsilon_{iabc\dots x} \delta_1^a \delta_2^b \dots \delta_N^x (\Pi_{i \neq n} dx'^i) \\ &= (-1)^{n-1} J^2 \varepsilon_{i123\dots N} (\Pi_{i \neq n} dx'^i) \quad // \text{index } n \text{ missing on } \varepsilon \\ &= \delta_n^i J^2 (\Pi_{i \neq n} dx'^i) \end{aligned}$$

The last step follows from the fact that $\varepsilon_{i123\dots N}$ with n missing must vanish if $i \neq n$, and if $i=n$ then $\varepsilon_{n123\dots N} = (-1)^n \varepsilon_{i123\dots N} = (-1)^n$. The conclusion then is that

$$d\mathbf{A}^n = J^2 (\Pi_{i \neq n} dx'^i) \mathbf{e}^n \quad \text{since } (e^n)_i = \delta_n^i \quad // \text{Section 7 (s)}$$

and the covariant vector area $d\mathbf{A}^n$ points in the n -axis direction in x' -space.

For the reader dubious of the claim that $d\mathbf{A}^n = J R d\mathbf{A}^n$, consider:

$$\begin{aligned} (d\mathbf{A}^n)_i &= J R_i^j (d\mathbf{A}^n)_j = J R_i^j \{ [(-1)^{n-1} \varepsilon_{iabc\dots x} (e_1)^a (e_2)^b \dots (e_N)^x (\Pi_{i \neq n} dx'^i)] \} \quad // e_n \text{ missing} \\ &= J R_i^j \{ [(-1)^{n-1} \varepsilon_{iabc\dots x} R_1^a R_2^b \dots R_N^x (\Pi_{i \neq n} dx'^i)] \} \quad // R_n^x \text{ missing} \\ &= J \{ [(-1)^{n-1} \varepsilon_{iabc\dots x} R_i^j R_1^a R_2^b \dots R_N^x (\Pi_{i \neq n} dx'^i)] \} \quad // R_n^x \text{ missing} \\ &= J \{ [(-1)^{n-1} \varepsilon_{iabc\dots x} S_i^j S_1^a S_2^b \dots S_N^x (\Pi_{i \neq n} dx'^i)] \} \quad // S_n^x \text{ missing} \end{aligned}$$

If $i \neq n$, one has a determinant with two columns the same since S_n^x is missing so the result is 0 so the result is proportional to δ_n^i . Continuing,

$$\begin{aligned} &= \delta_n^i J \{ (-1)^{n-1} \varepsilon_{nabc\dots x} S_n^j S_1^a S_2^b \dots S_N^x (\Pi_{i \neq n} dx'^i) \} \quad // S_n^x \text{ missing in group} \\ &= \delta_n^i J \{ (-1)^{n-1} \varepsilon_{nabc\dots x} S_1^a S_2^b \dots S_n^j \dots S_N^x (\Pi_{i \neq n} dx'^i) \} \\ &= \delta_n^i J \{ \varepsilon_{abc\dots n\dots x} S_1^a S_2^b \dots S_n^j \dots S_N^x (\Pi_{i \neq n} dx'^i) \} \\ &= \delta_n^i J \{ \det(S) (\Pi_{i \neq n} dx'^i) \} \\ &= \delta_n^i J^2 (\Pi_{i \neq n} dx'^i) \end{aligned}$$

which agrees with the result just obtained from the covariant x' -space equation.

4. Volume Transformation. What about the volume dV ? Assume for the moment that dV is correctly represented this way, where any n will do:

$$dV = d\mathbf{A}^n \bullet d\mathbf{x}^{(n)} = (d\mathbf{A}^n)_i [d\mathbf{x}^{(n)}]^i \quad // \text{ no implied sum}$$

Installing our expression for $d\mathbf{A}^n$ and $d\mathbf{x}^{(n)}$ gives

$$dV = J \mathbf{e}^n (\Pi_{i \neq n} dx'^i) \bullet (\mathbf{e}_n dx'_n) = J (\Pi_i dx'^i) \mathbf{e}^n \bullet \mathbf{e}_n = J (\Pi_i dx'^i)$$

which is seen to agree with the modified dV definition stated above. Looking at $dV = (d\mathbf{A}^n)_i [d\mathbf{x}^{(n)}]^i$, dV is seen to be a tensor combination of a vector density of weight -1 with a vector density of weight 0 (an ordinary vector $d\mathbf{x}^{(n)}$), so according to Appendix D, dV must be scalar density of weight -1. This then tells us that

$$dV' = J dV$$

Since $dV = J (\Pi_i dx'^i)$ one gets

$$dV' = J^2 (\Pi_i dx'^i) .$$

Again, one can verify this last result from the x' -space covariant form of the equation:

$$dV' = d\mathbf{A}'^n \bullet d\mathbf{x}'^{(n)} = \{ J^2 (\Pi_{i \neq n} dx'^i) \mathbf{e}'^n \} \bullet \{ \mathbf{e}'_n dx'^n \} = J^2 (\Pi_i dx'^i) \mathbf{e}'^n \bullet \mathbf{e}'_n = J^2 (\Pi_i dx'^i)$$

since $\mathbf{e}'^n \bullet \mathbf{e}'_n = \mathbf{e}^n \bullet \mathbf{e}_n = 1$.

An alternate derivation of the dV transform rule $dV' = J dV$ comes from just staring at

$$dV = \varepsilon_{abc \dots x} (dx'e_1)^a (dx'e_2)^b \dots (dx'e_N)^x$$

which by the argument above is seen directly to transform as a tensor density of weight -1. To complete the circle, we can verify for a second time the claim made above that $dV = d\mathbf{A}^n \bullet d\mathbf{x}^{(n)}$:

$$\begin{aligned} d\mathbf{A}^n \bullet d\mathbf{x}^{(n)} &= (d\mathbf{A}^n)_i [d\mathbf{x}^{(n)}]^i \\ &= \{ (-1)^{n-1} \varepsilon_{iabc \dots x} (e_1)^a (e_2)^b \dots (e_N)^x (\Pi_{k \neq n} dx'^k) \} (dx'^n e_n)^i \quad // \mathbf{e}_n \text{ missing} \dots \\ &= \varepsilon_{abc \dots i \dots x} (e_1)^a (e_2)^b \dots (e_N)^x (\Pi_k dx'^i) = \det(S) (\Pi_k dx'^i) = dV \end{aligned}$$

To summarize the above, by examining the vector density nature of our various objects, we have been able to determine exactly how edges, areas and the volume transform under transformation F :

edges	$dx'^{(n)} = R dx^{(n)}$	or	$[dx'^{(n)}]^i = R^i_j [dx^{(n)}]^j$	// ordinary vector
areas	$dA^n = J R dA^n$	or	$(dA^n)^i = J R^i_j (dA^n)^j$	// vector density $W = -1$
volume	$V' = J dV$			// scalar density $W = -1$

5. Covariant Magnitudes. The x' -space magnitudes here are the "covariant" ones which are associated with the Curvilinear View of x' -space, as discussed above. Since $dx^{(n)}$ is a vector, it follows that

$$|dx'^{(n)}|^2 = dx'^{(n)} \bullet dx'^{(n)} = dx^{(n)} \bullet dx^{(n)} = |dx^{(n)}|^2 \quad \Rightarrow \quad |dx'^{(n)}| = |dx^{(n)}|$$

Since dA^n is a vector density of weight -1, it follows that

$$|dA^n|^2 = dA^n \bullet dA^n = J^2 dA^n \bullet dA^n = J^2 |dA^n|^2 \quad \Rightarrow \quad |dA^n| = |J| |dA^n|$$

where we note that $dA^n \bullet dA^n$, being a combination of two weight -1 vector densities, is a scalar density of weight -2 and thus transforms as shown above. For completeness, we can add from the above table,

$$V' = J dV \quad \Rightarrow \quad |dV'| = |J| |dV|.$$

Moreover it has been shown above that

$$\begin{aligned} dx^{(n)} &= e_n dx'_n & \Rightarrow |dx^{(n)}| &= |e_n| dx'_n = h'_n dx'_n \\ dx'^{(n)} &= e'_n dx'_n & \Rightarrow |dx'^{(n)}| &= |e'_n| dx'_n = h'_n dx'_n \\ dA^n &= J (\Pi_{i \neq n} dx'^i) e^n & \Rightarrow |dA^n| &= |J| |e^n| (\Pi_{i \neq n} dx'^i) = (|J|/h'_n) (\Pi_{i \neq n} dx'^i) \\ dA'^n &= J^2 (\Pi_{i \neq n} dx'^i) e'^n & \Rightarrow |dA'^n| &= J^2 |e'^n| (\Pi_{i \neq n} dx'^i) = (|J|^2/h'_n) (\Pi_{i \neq n} dx'^i) \\ dV &= J (\Pi_i dx'^i) & \Rightarrow |dV| &= |J| (\Pi_i dx'^i) \\ dV' &= J^2 (\Pi_i dx'^i) & \Rightarrow |dV'| &= |J|^2 (\Pi_i dx'^i) \end{aligned}$$

which can be written more compactly using the Cartesian-View coordinate variation groupings,

$$\begin{aligned} dx^{(n)} &= e_n d\mathcal{L}^n & \Rightarrow |dx^{(n)}| &= h'_n d\mathcal{L}^n & d\mathcal{L}^n &\equiv dx^n \\ dx'^{(n)} &= e'_n d\mathcal{L}'^n & \Rightarrow |dx'^{(n)}| &= h'_n d\mathcal{L}'^n & \\ dA^n &= J d\mathcal{A}^n e^n & \Rightarrow |dA^n| &= (|J|/h'_n) d\mathcal{A}^n & d\mathcal{A}^n &\equiv \Pi_{i \neq n} dx'^i \\ dA'^n &= J^2 d\mathcal{A}'^n e'^n & \Rightarrow |dA'^n| &= (|J|^2/h'_n) d\mathcal{A}'^n & \\ dV &= J d\mathcal{V} & \Rightarrow |dV| &= |J| d\mathcal{V} & d\mathcal{V} &\equiv \Pi_i dx'^i \\ dV' &= J^2 d\mathcal{V}' & \Rightarrow |dV'| &= |J|^2 d\mathcal{V}' & \end{aligned}$$

The covariant edge magnitude is of course unchanged by the transformation since it is a scalar, while the area and volume magnitudes are magnified by $|J|$ in going from x -space to x' -space. Since $dV = dA^n \bullet dx^{(n)}$, it is clear that the area's transformation factor of $|J|$ is passed onto the volume. Notice that dV' is

always positive regardless of the sign of J , and this is because x' -space is always a right-handed coordinate system, as in Section 6 (i). In contrast, dV can have either sign depending on the sign of $J = \det(S)$, and so $dV < 0$ when the $\mathbf{e}_n$ form a left-handed coordinate system in x -space.

6. Two Theorems : $\mathbf{g}' / h'_n{}^2 = \mathbf{g}'^{nn} \mathbf{g}' = \text{cof}(\mathbf{g}'_{nn})$ and $|\Pi^{\mathbf{x}}_{i \neq n} \mathbf{e}_i| = \sqrt{\text{cof}(\mathbf{g}'_{nn})}$

We now pause to prove two small theorems which will be used below.

Theorem 1: $\mathbf{g}' / h'_n{}^2 = \mathbf{g}'^{nn} \mathbf{g}' = \text{cof}(\mathbf{g}'_{nn})$ where $g' = \det(g'_{ij})$.

The left equality is obvious since $\mathbf{g}'^{nn} = 1/h'_n{}^2$ (orthogonal coordinates). The right equality can be shown as follows:

$$(g'_{up})_{ab} \equiv g'^{ab} \quad (g'_{dn})_{ab} \equiv g'_{ab}$$

$$g'_{up} = (g'_{dn})^{-1} = \text{cof}(g'_{dn}^T) / \det(g'_{dn}) = \text{cof}(g'_{dn}) / \det(g'_{dn})$$

$$\Rightarrow (g'_{up})_{nn} = \text{cof}[(g'_{dn})_{nn}] / \det(g'_{dn})$$

or

$$g'^{nn} = \text{cof}[g'_{nn}] / g' \quad \text{QED}$$

Theorem 2: $|\Pi^{\mathbf{x}}_{i \neq n} \mathbf{e}_i| = \sqrt{\text{cof}(\mathbf{g}'_{nn})}$

The quantity on the left is this

$$|\Pi^{\mathbf{x}}_{i \neq n} \mathbf{e}_i| \equiv |\mathbf{e}_1 \times \mathbf{e}_2 \dots \times \mathbf{e}_N| \text{ where } \mathbf{e}_n \text{ is missing from cross product.}$$

We shall give three quick proofs of Theorem 2, the last being valid only for $N=3$.

• First, at the start of section (c) above, one form for $d\mathbf{A}_n$ is given by

$$d\mathbf{A}^n = (-1)^{n-1} (\Pi^{\mathbf{x}}_{i \neq n} \mathbf{e}_i) (\Pi_{i \neq n} dx^i) = (-1)^{n-1} (\Pi^{\mathbf{x}}_{i \neq n} \mathbf{e}_i) d\mathcal{A}^n$$

$$\Rightarrow |d\mathbf{A}^n| = |\Pi^{\mathbf{x}}_{i \neq n} \mathbf{e}_i| d\mathcal{A}^n$$

Comparison of this last result with the third line of the table above shows that the following must be true

$$|\Pi^{\mathbf{x}}_{i \neq n} \mathbf{e}_i| = (|J| / h'_n) = g'^{1/2} / h'_n = (g' / h'_n{}^2)^{1/2} = \sqrt{\text{cof}(\mathbf{g}'_{nn})} \quad \text{QED}$$

where the last step follows from Theorem 1.

• Here is a more direct proof:

$$|\Pi^{\mathbf{x}}_{i \neq n} (\mathbf{e}_i)|^2 = \Pi^{\mathbf{x}}_{i \neq n} (\mathbf{e}_i) \cdot \Pi^{\mathbf{x}}_{j \neq n} (\mathbf{e}_j) = [\Pi^{\mathbf{x}}_{i \neq n} (\mathbf{e}_i)]_k [\Pi^{\mathbf{x}}_{j \neq n} (\mathbf{e}_j)]_k$$

$$\begin{aligned}
&= [\varepsilon_{kabc} \dots x (e_1)^a (e_2)^b \dots (e_N)_x] [\varepsilon_{ka'b'c'} \dots x (e_1)^{a'} (e_2)^{b'} \dots (e_N)^{x'}] \quad // e_n \text{ missing in both} \\
&= \varepsilon_{kabc} \dots x \varepsilon_{ka'b'c'} \dots x \{ (e_1)^a (e_2)^b \dots (e_N)_x \} (e_1)^{a'} (e_2)^{b'} \dots (e_N)^{x'} \quad // e_n \text{ missing} \\
&= e_1 \bullet e_1 e_2 \bullet e_2 \dots e_N \bullet e_N + \text{all signed permutations of the 2nd labels} \quad // e_n \text{ missing} \\
&= g'_{11} g'_{22} \dots g'_{nn} + \text{all signed permutations of the 2nd indices} \quad // e_n \text{ missing}
\end{aligned}$$

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.

- A proof for $N=3$ uses normal vector algebra. Setting $n = 1$, for example, one needs to show that

$$|\Pi_{i \neq 1}^x e_i|^2 = |e_2 \times e_3|^2 = \text{cof}(g'_{11})$$

To this end, use the vector identity

$$(A \times B) \bullet (A \times B) = A^2 B^2 - (A \bullet B)^2$$

to show that

$$|e_2 \times e_3|^2 = (e_2 \times e_3) \bullet (e_2 \times e_3) = |e_2|^2 |e_3|^2 - (e_2 \bullet e_3)^2 = g'_{22} g'_{33} - (g'_{23})^2 = \text{cof}(g'_{11}).$$

and the cases $n = 2$ and 3 are similar.

7. Cartesian-View Magnitude Ratios. In the Cartesian View of x' -space one can write the Cartesian x' -space magnitudes as

$$|dx'^{(n)}|_c = d\mathcal{L}^n \quad |dA'^{(n)}|_c = d\mathcal{A}^n \quad |dV'|_c = d\mathcal{V}$$

Then from the three x -space equations in the above table (just above Two Theorems) one obtains the following three ratios of x -space objects divided by their corresponding Cartesian-View x' -space objects:

$$\begin{aligned}
|dx'^{(n)}| / d\mathcal{L}^n &= h'_n = [g'_{nn}]^{1/2} = \text{the scale factor for edge } dx'^{(n)} \\
|dA'^{(n)}| / d\mathcal{A}^n &= (1/h'_n) |J| = (1/h'_n) g'^{1/2} = [g'^{nn} g']^{1/2} = [\text{cof}(g'_{nn})]^{1/2} \quad // \text{Theorem 1 above} \\
|dV| / d\mathcal{V} &= |J| = g'^{1/2} \quad // g' \equiv \det(g'_{ij}) = J^2
\end{aligned}$$

It is convenient to make the definition

$$dA^n \equiv |dA'^{(n)}|$$

and then the above area magnitude ratio relation may be written

$$dA^n = \sqrt{\text{cof}(g'_{nn})} d\mathcal{A}^n$$

and $d\mathcal{A}^n = \Pi_{i \neq n} dx^{i'}$ is just a product of the appropriate curvilinear coordinate variations.

8. Nested Cofactor Formulas. The object dA^n is the "area" of a face on an N-piped. This face, which is itself an (N-1)-piped, in turn has its own "areas" which are (N-2)-pipeds, and so on, so there is a hierarchy of "areas" of dimensions N-1 all the way down. The area ratios of corresponding areas under transformation F are determined by equations similar to that above. For example, the m^{th} face of face n of an N-piped has area ratio $\sqrt{\text{cof}[\text{cof}(g'_{nn})]_{mm}}$. This at least makes some sense since the matrix $\text{cof}(g'_{nn})$ has dimension N-1, so its cofactor matrix has dimension N-2, and so on. For N=3, these faces would be line segments and one would have

$$\text{cof}(g'_{33}) = \begin{pmatrix} g'_{11} & g'_{12} \\ g'_{21} & g'_{22} \end{pmatrix} \Rightarrow \text{cof}[\text{cof}(g'_{33})]_{22} = g'_{11} = h_1'^2 \Rightarrow \sqrt{\text{cof}[\text{cof}(g'_{33})]_{22}} = h_1'$$

and h_1' is in fact the edge length ratio given above.

9. Transformation of arbitrary differential vectors, areas and volume. The above discussion is geared to the N-piped transform picture and reveals how $dx^{(n)}$, $dA^{(n)}$ and dV transform where all these quantities are directly associated with the particular differential N-piped in x-space spanned by the $(e_n dx^{(n)})$ vectors.

But suppose dx is an *arbitrary* differential vector in x-space. Certainly $dx' = R dx$, so we know how this dx transforms under F. But what about area and volume?

Based on the work in Appendix B as carried through into the N-piped transform discussion above, it seems clear that an *arbitrary* differential area in x-space dA can be *represented* as follows : (g=1)

$$dA = (dx^{[1]}) \times (dx^{[2]}) \dots \times (dx^{[N-1]})$$

$$(dA)_i = \varepsilon_{iabc...x} (dx^{[1]})^a (dx^{[2]})^b \dots (dx^{[N-1]})^x,$$

where the $dx^{[i]}$ are an arbitrary set of N-1 linearly independent differential vectors in x-space. For N=3 one would write $dA = dx^{[1]} \times dx^{[2]}$. Since ε has weight -1 and all other objects are true vectors (weight 0), one again concludes that dA transforms as a tensor density of weight -1, so

$$dA' = J R dA \quad \text{or} \quad dA'_i = J R_i^j dA_j$$

If any of these $dx^{[k]}$ were a linear combination of the N-2 others, one could say $dx^{[k]} = \sum_j \alpha_j^{(k)} dx^{[j]}$ and then the above $(dA)_i$ expression would give a sum of terms each of which vanishes by symmetry, resulting in $(dA)_i = 0$.

Finally, given the set of $dx^{[i]}$ linearly independent vectors shown above for $i = 1, 2, \dots, N-1$, we can certainly find one more such that all N are then linearly independent, so then we have a set of N arbitrary

differential vectors $dx^{[i]}$ (arbitrary as long as they are linearly independent), and these will form a volume in x-space,

$$dV = \det(dx^{[1]}, dx^{[2]}, \dots, dx^{[N]}) = \varepsilon_{abc\dots y} (dx^{[1]})^a (dx^{[2]})^b \dots (dx^{[N]})^y$$

By the argument just given, inspection shows that this dV transforms as a tensor density of weight -1 so $dV' = J dV$.

These then are the transformation rules for arbitrary differential vectors, areas and volumes transforming under F :

$$\begin{array}{llll} dx' = R dx & |dx'| = |dx| & & \\ dA' = J R dA & |dA'| = |J| |dA| & & \\ dV' = J dV & |dV'| = |J| |dV| & J = \det(S) & J^2 = g' \end{array}$$

Review and covariant form of the dA and dV equations

To review, then, in Cartesian x-space we have these expressions for area and volume

$$\begin{aligned} dA &= (dx^{[1]} \times dx^{[2]} \dots \times dx^{[N-1]}) \\ dV &= \det [dx^{[1]}, dx^{[2]}, \dots, dx^{[N]}] \end{aligned}$$

Written out in terms of covariant vector components these expressions appear as,

$$\begin{aligned} (dA)_i &= \varepsilon_{iabc\dots x} (dx^{[1]})^a (dx^{[2]})^b \dots (dx^{[N-1]})^x \\ dV &= \varepsilon_{abc\dots y} (dx^{[1]})^a (dx^{[2]})^b \dots (dx^{[N]})^y \end{aligned}$$

where ε is the usual permutation tensor involved in cross products and determinants. As noted earlier, these are both "valid tensor density equations" (end of Section 7 (u)) , so they can be written in x'-space as

$$\begin{aligned} (dA')_i &= \varepsilon'_{iabc\dots x} (dx'^{[1]})^a (dx'^{[2]})^b \dots (dx'^{[N-1]})^x \\ dV' &= \varepsilon'_{abc\dots y} (dx'^{[1]})^a (dx'^{[2]})^b \dots (dx'^{[N]})^y \end{aligned}$$

where $dx'^{[i]} = R dx^{[i]}$. From Appendix D, each ε' can be written as $\varepsilon' = J^2 \varepsilon = (g'/g) \varepsilon = g' \varepsilon$, so

$$\begin{aligned} (dA')_i &= g' \varepsilon_{iabc\dots x} (dx'^{[1]})^a (dx'^{[2]})^b \dots (dx'^{[N-1]})^x \\ dV' &= g' \varepsilon_{abc\dots y} (dx'^{[1]})^a (dx'^{[2]})^b \dots (dx'^{[N]})^y \end{aligned}$$

Since the permutation tensor now appears, these can be written in terms of cross products,

$$\begin{aligned} dA' &= g' (dx'^{[1]} \times dx'^{[2]} \dots \times dx'^{[N-1]}) \\ dV' &= g' \det [dx'^{[1]}, dx'^{[2]}, \dots, dx'^{[N]}] \end{aligned}$$

Using contravariant components $[dx'^{[i]}]^i$, this cross product and determinant are formed just as they are in x-space. These two equations then give at least some feel for "the meaning of dA' and dV' in x'-space" .

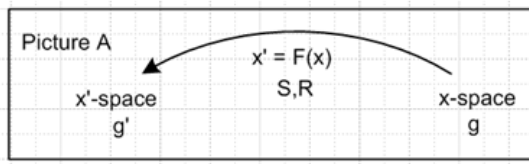
Going back to x -space, we could have written the equations there using $g = \det(g'_{ij}) = \det(\delta_{ij}) = 1$:

$$dA = g (dx^{[1]}) \times (dx^{[2]}) \dots \times (dx^{[N-1]})$$

$$dV = g \det (dx^{[1]}, dx^{[2]}, \dots dx^{[N]}) \quad .$$

This then is a useful interpretation of the covariant form of these equations. Adding primes to everything in the above two equations yields the previous two equations and only the permutation tensor ε is involved in both sets of equations.

With this understanding, the above transformation rules for differential vectors, areas and volumes can be extended from Picture B to the more general Picture A, where g is an arbitrary metric tensor,



How things look in developmental notation.

Recall that covariant tensor objects get overbars and all indices are down in the developmental notation used in Sections 1-6 of this document. Here then are some of the above equations expressed in this notation:

$$d\bar{A} = (dx^{[1]}) \times (dx^{[2]}) \dots \times (dx^{[N-1]})$$

$$dV = \det [dx^{[1]}, dx^{[2]}, \dots dx^{[N]}] \quad .$$

$$(d\bar{A})_i = \bar{\varepsilon}_{iabc\dots x} (dx^{[1]})_a (dx^{[2]})_b \dots (dx^{[N-1]})_x$$

$$dV = \bar{\varepsilon}_{abc\dots y} (dx^{[1]})_a (dx^{[2]})_b \dots (dx^{[N]})_y \quad ,$$

$$(d\bar{A}')_i = \bar{\varepsilon}'_{iabc\dots x} (dx'^{[1]})_a (dx'^{[2]})^b \dots (dx'^{[N-1]})^x$$

$$dV' = \bar{\varepsilon}'_{abc\dots y} (dx'^{[1]})^a (dx'^{[2]})^b \dots (dx'^{[N]})^y \quad .$$

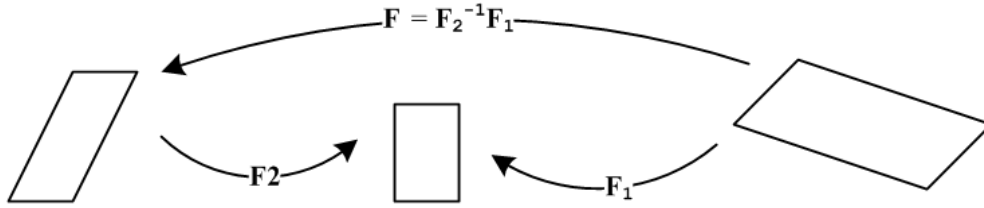
$$\begin{aligned} dx' &= R dx & |dx'| &= |dx| \\ d\bar{A}' &= J S^T d\bar{A} & |dA'| &= |J| |dA| \\ dV' &= J dV & |dV'| &= |J| |dV| \end{aligned} \quad J = \det(S) \quad J^2 = g'$$

Recall from Section 2 (c) that a covariant vector transforms as $V' = S^T V$, and a covariant vector density of weight W will then transform as $V' = J^{-W} S^T V$ and this explains the middle line of the above three. The *contravariant* differential area would transform as $dA' = J R dA$.

10. Concatenation of Transformations. Consider $x'' = F_2(x')$ and $x' = F_1(x)$ so that $x'' = F_2[F_1(x)] \equiv F(x)$. At some point x , the linearization will yield $dx'' = R_2 R_1 dx$ as the rule for vector transformation, so the R matrix associated with transformation F is $R = R_2 R_1$, and then $S = R^{-1} = R_1^{-1} R_2^{-1} = S_1 S_2$. Each

transformation will have an associated Jacobian: $J_1 = \det(S_1)$ and $J_2 = \det(S_2)$. The concatenated transformation then has $J = \det(S) = \det(S_1 S_2) = \det(S_1) \det(S_2) = J_1 J_2$. The implication is that when one concatenates two transformations in this manner, the area and volume transformations shown above still apply, where J is taken to be the product of the two underlying Jacobians.

For example, one could consider the mapping between two different skewed N-pipeds, each representing a different curvilinear coordinate system, with our orthogonal N-piped as an intermediary object,



In this case one has $F = F_2^{-1} F_1$, so $R = R_2^{-1} R_1 = S_2 R_1$ and then $S = S_1 R_2$ so $J = J_1/J_2$. This J then would be used in the above area and volume transformation rules, for example, $dA_{\text{left}} = J R dA_{\text{right}}$.

In the continuum mechanics flow application, $g = 1$ on both left and right as well as center, time t_0 is on the right, time t on the left, and the volume transformation is given by $dV_{\text{left}} = J dV_{\text{right}}$ where J is associated with the combined F . This J then characterizes the volume change between an initial and final flow particle where each is skewed in some arbitrary manner.

Examples of area magnitude transformation for $N = 2, 3, 4$

In the previous section it was shown that $dA^n = \sqrt{\text{cof}(g'_{nn})} d\mathbf{a}^n$. 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 $d\mathbf{a}^n$ are the curvilinear coordinate variations one can visualize in the Cartesian-View picture shown above.

For $N=2$ the area magnitude transformation results are (for a general non-orthogonal x' -space system)

$$\begin{aligned} dA^1 &= \sqrt{g'_{22}} d\mathbf{a}^1 = h'_2 d\mathbf{a}^1 & d\mathbf{a}^1 &= dx'^1 = d\mathcal{L}^1 \\ dA^2 &= \sqrt{g'_{11}} d\mathbf{a}^2 = h'_1 d\mathbf{a}^2 & d\mathbf{a}^2 &= dx'^2 = d\mathcal{L}^2 \end{aligned}$$

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 $[1,2,3 = r,\theta,\varphi]$

$$\begin{aligned} h'_1 &= h'_r = 1 & d\mathbf{a}^1 &= dx'^2 dx'^3 = d\theta d\varphi \\ h'_2 &= h'_\theta = r & d\mathbf{a}^2 &= dx'^3 dx'^1 = dr d\varphi \\ h'_3 &= h'_\varphi = r \sin\theta & d\mathbf{a}^3 &= dx'^1 dx'^2 = dr d\theta \end{aligned}$$

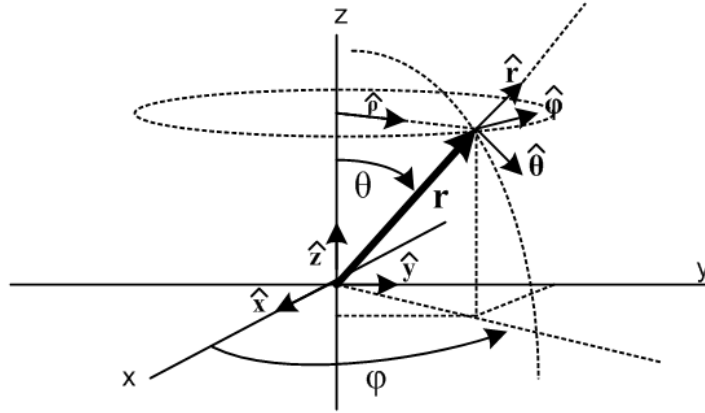
Therefore

$$\begin{aligned} dA^1 &= dA^1 \hat{\mathbf{e}}_1 \Rightarrow d\mathbf{A}^r = dA^r \hat{\mathbf{e}}_r = dA^r \hat{\mathbf{r}} \text{ with } dA^r = h'_2 h'_3 d\mathbf{a}^1 = r^2 \sin\theta d\theta d\varphi \\ dA^2 &= dA^2 \hat{\mathbf{e}}_2 \Rightarrow d\mathbf{A}^\theta = dA^\theta \hat{\mathbf{e}}_\theta = dA^\theta \hat{\boldsymbol{\theta}} \text{ with } dA^\theta = h'_3 h'_1 d\mathbf{a}^2 = r \sin\theta dr d\varphi \\ dA^3 &= dA^3 \hat{\mathbf{e}}_3 \Rightarrow d\mathbf{A}^\varphi = dA^\varphi \hat{\mathbf{e}}_\varphi = dA^\varphi \hat{\boldsymbol{\varphi}} \text{ with } dA^\varphi = h'_1 h'_2 d\mathbf{a}^3 = r dr d\theta \end{aligned}$$

so that

$$\begin{aligned} d\mathbf{A}^r &= r^2 \sin\theta d\theta d\varphi \hat{\mathbf{r}} & \rho d\varphi r d\theta & \rho = r \sin\theta \\ d\mathbf{A}^\theta &= r \sin\theta dr d\varphi \hat{\boldsymbol{\theta}} & \rho d\varphi dr & \\ d\mathbf{A}^\varphi &= r dr d\theta \hat{\boldsymbol{\varphi}} & r d\theta dr & \end{aligned}$$

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,



(d) 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 positive differential volume element in x -space which is typically not aligned with the axes and for a general transformation F is not in general orthogonal. 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,

$$\lim \sum_i dV_a(\mathbf{x}_i) f(\mathbf{x}_i) = \lim \sum_i dV_b(\mathbf{x}_i) f(\mathbf{x}_i)$$

that

$$\int_D dV_a h(\mathbf{x}) = \int_D dV_b h(\mathbf{x})$$

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

$$|J(\mathbf{x}')| d\mathcal{V}' = dV_b$$

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 on the right $h(\mathbf{x}) = h(\mathbf{x}(\mathbf{x}'))$ and region D' is the same region as D expressed in terms of the x' coordinates. Writing out the volume elements this says

$$\int_D \left(\prod_{i=1}^N dx^i \right) h(\mathbf{x}) = \int_{D'} \left(\prod_{i=1}^N dx'^i \right) |J(\mathbf{x}')| h(\mathbf{x}(\mathbf{x}'))$$

and finally, using Section 5 (k),

$$\int_D \left(\prod_{i=1}^N dx^i \right) h(\mathbf{x}) = \int_{D'} \left(\prod_{i=1}^N dx'^i \right) \left[\sqrt{\det(g'_{ab})} \right] h(\mathbf{x}(\mathbf{x}'))$$

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 $h(\mathbf{x}(r, \theta, \phi)) = h(r \sin \theta \cos \phi, r \sin \theta \sin \phi, r \cos \theta)$.

(e) 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 $|J|$:

$$J(\mathbf{x}') \equiv \det(S^{\mathbf{i}}_{\mathbf{j}}(\mathbf{x}')) = \det(\partial x^{\mathbf{i}} / \partial x'^{\mathbf{k}}) = 1 / \det(R^{\mathbf{i}}_{\mathbf{j}}(\mathbf{x}(\mathbf{x}')) = 1 / \det(\partial x'^{\mathbf{i}} / \partial x^{\mathbf{k}}) \quad // \text{Section 5 (k)}$$

$$|J(\mathbf{x}')| = \sqrt{\det(g'_{\mathbf{ab}}(\mathbf{x}'))} = \sqrt{g'(\mathbf{x}')} \quad // \text{Section 5 (k)}$$

$$|J(\mathbf{x}')| = \text{the volume of the N-piped in x-space spanned by the } \mathbf{e}_{\mathbf{n}}(\mathbf{x}), \text{ where } \mathbf{x} = \mathbf{F}^{-1}(\mathbf{x}')$$

$$|J(\mathbf{x}')| = dV_{\mathbf{N-piped}} / d\mathcal{V}' = \text{ratio of differential x-space N-piped volume} / (\prod_{\mathbf{i}=1}^{\mathbf{N}} dx'^{\mathbf{i}})$$

$$|J(\mathbf{x}')| = dV / dV' = (\prod_{\mathbf{i}=1}^{\mathbf{N}} dx^{\mathbf{i}}) / (\prod_{\mathbf{i}=1}^{\mathbf{N}} dx'^{\mathbf{i}}) \quad // \text{distributional Jacobian Integration Rule}$$

As discussed in Section 6 (i), if the curvilinear coordinates are ordered so that the $\mathbf{e}_{\mathbf{n}}$ form a right handed coordinate system, then $\det(S) > 0$, $\sigma = \text{sign}(\det(S)) = +1$, and $|J| = J$.

(f) Volume integration of a tensor field under linear transformations

In section (d) above it was shown that, with its interpretation as a distribution,

$$dV' = J^{-1} dV \quad .$$

In the language of Appendix D, this says that the volume element transforms from x-space to x'-space as a scalar density of weight +1. Since $g^{1/2}$ transforms as a scalar density of weight -1, according to App. D (b) item 3 the object $g^{1/2} dV$ then transforms as an ordinary scalar (weight 0) (see for example Weinberg p 99 (4.4.6)), Then if one were to define

$$T^{\mathbf{i} \mathbf{j} \mathbf{k} \dots} \equiv \int_{\mathbf{D}} g^{1/2} dV A^{\mathbf{i} \mathbf{j} \mathbf{k} \dots}(\mathbf{x}) \quad (*)$$

one might expect that, if $A^{\mathbf{i} \mathbf{j} \mathbf{k} \dots}(\mathbf{x})$ transforms as a tensor field, then $T^{\mathbf{i} \mathbf{j} \mathbf{k} \dots}$ might transform as tensor. To investigate this conjecture, consider the above integral in x'-space,

$$\begin{aligned} T^{\mathbf{i} \mathbf{j} \mathbf{k} \dots} &\equiv \int_{\mathbf{D}'} g'^{1/2} dV' A^{\mathbf{i} \mathbf{j} \mathbf{k} \dots}(\mathbf{x}') = \int_{\mathbf{D}} g^{1/2} dV A^{\mathbf{i} \mathbf{j} \mathbf{k} \dots}(\mathbf{x}'(\mathbf{x})) \\ &= \int_{\mathbf{D}} g^{1/2} dV R^{\mathbf{i}}_{\mathbf{i}'} R^{\mathbf{j}}_{\mathbf{j}'} R^{\mathbf{k}}_{\mathbf{k}'} \dots A^{\mathbf{i} \mathbf{j} \mathbf{k} \dots}(\mathbf{x}) \end{aligned}$$

If the underlying transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ is linear, then the $R^{\mathbf{a}}_{\mathbf{b}}$ are independent of position and one has

$$\begin{aligned}
&= R^i_{\mathbf{i}} R^j_{\mathbf{j}} \dots \int_D g^{1/2} dV A^{i'j'k' \dots}(\mathbf{x}) \\
&= R^i_{\mathbf{i}} R^j_{\mathbf{j}} \dots T^{i'j'k' \dots}
\end{aligned}$$

Therefore, for *linear* transformations $\mathbf{x}' = \mathbf{F}(\mathbf{x})$, an integral of the form (*) of a tensor field is itself a tensor of the same type.

Examples: If $g = 1$ and $\mathbf{x}' = \mathbf{R}\mathbf{x}$ where $\mathbf{R}$ is a rotation (independent of position), then $g' = 1$ as well and one may conclude that the volume integral of a tensor field of any type is a tensor of the same type. Here are two simple examples:

$$\begin{aligned}
J^i &= \int_D dV x^i \\
J^{ij} &= \int_D dV [r^2 \delta^{ij} - x^i x^j]
\end{aligned}$$

Under rotations, x^i is a true vector, and $r^2 \delta^{ij} - x^i x^j$ is a true rank-2 tensor (traceless). It follows that J^i and J^{ij} are also tensors. Under rotations, mass density ρ transforms as a scalar, so the following objects are tensors as well,

$$\begin{aligned}
I^i &= \int_D dV \rho(\mathbf{x}) x^i && // \text{ vector} \\
I^{ij} &= \int_D dV \rho(\mathbf{x}) [r^2 \delta^{ij} - x^i x^j] && // \text{ rank-2 tensor}
\end{aligned}$$

Vector components I^i are the first moments of a mass distribution, while I^{ij} is the usual inertia tensor.

Since I^{ij} is real and symmetric, it can be diagonalized by a certain rotation $\mathbf{R}$. We can think of this as a transformation from x -space to x' -space where the resulting tensor I'^{ij} is diagonal. In x' -space, the diagonal elements of the tensor I'^{ij} are then its eigenvalues ($\lambda_i = I'^{ii}$), while the eigenvectors are the axis-aligned $\mathbf{e}'_i$ of Section 3. Back in x -space, the eigenvectors of the non-diagonal I^{ij} are then $\mathbf{e}_n \equiv \mathbf{S}\mathbf{e}'_n$ where $\mathbf{S} = \mathbf{R}^{-1}$. This can be verified as follows (developmental notation)

$$I' = \mathbf{R} I \mathbf{R}^{-1} \quad \Rightarrow \quad I'_{ij} = R_{i\mathbf{i}} I_{\mathbf{i}\mathbf{j}} (R^{-1})_{\mathbf{j}j} = R_{i\mathbf{i}} R_{j\mathbf{j}} I_{\mathbf{i}\mathbf{j}} \quad // \text{ contravariant tensor}$$

$$I' \mathbf{e}'_n = \lambda_n \mathbf{e}'_n \quad // I' \text{ is diagonal}$$

so

$$I \mathbf{e}_n = [\mathbf{R}^{-1} I' \mathbf{R}] [\mathbf{S} \mathbf{e}'_n] = \mathbf{R}^{-1} I' (\mathbf{R} \mathbf{S}) \mathbf{e}'_n = \mathbf{S} I' \mathbf{e}'_n = \mathbf{S} \lambda_n \mathbf{e}'_n = \lambda_n \mathbf{S} \mathbf{e}'_n = \lambda_n \mathbf{e}_n .$$

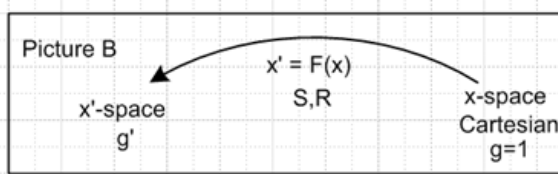
9. The Divergence in curvilinear coordinates

Note: Covariant derivations of all the curvilinear differential operator expressions appear in Section 15. In Sections 9 through 13, we provide more "physical" or "brute force" derivations which are, of necessity, much less compact. That compactness is a testament to the power of the covariant derivative formalism, which might be called "semicolon technology". The formalism is not used in Sections 9 through 13.

(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.

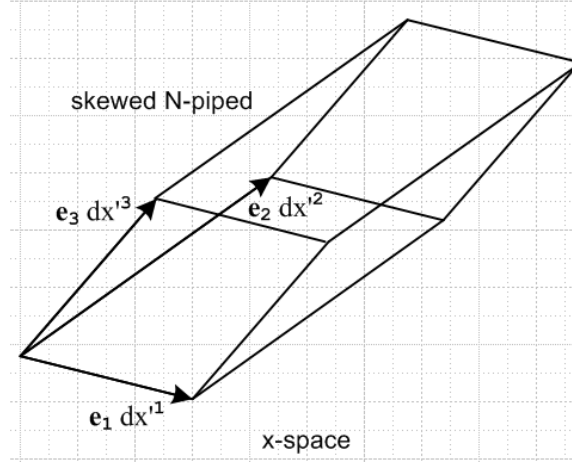
We provide here a geometric derivation (in N dimensions) of the formula for the divergence of a contravariant vector field expressed in curvilinear coordinates, which means x' -space coordinates with Picture B.



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 15 (c).

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,

$$\begin{aligned} [d\mathbf{A}^n] \bullet \mathbf{B}(\mathbf{x}) &\approx [\sqrt{g'(\mathbf{x}'_{\text{far}})} \mathbf{e}^n(\mathbf{x}_{\text{far}}) (\Pi_{i \neq n} dx^{i'})] \bullet \mathbf{B}(\mathbf{x}_{\text{far}}) \\ &= \sqrt{g'(\mathbf{x}'_{\text{far}})} (\Pi_{i \neq n} dx^{i'}) \mathbf{e}^n(\mathbf{x}_{\text{far}}) \bullet \mathbf{B}(\mathbf{x}_{\text{far}}) \end{aligned}$$

where $\mathbf{x}_{\text{far}}$ is taken to be a point at the center of far face n. Recall from Section 7 (s) that

$$\mathbf{B}(\mathbf{x}) = B'^n(\mathbf{x}') \mathbf{e}_n \quad \text{where} \quad B'^n(\mathbf{x}') = \mathbf{e}^n(\mathbf{x}) \bullet \mathbf{B}(\mathbf{x})$$

which says that, when $\mathbf{B}$ is expanded on the $\mathbf{e}_n$, the coefficients B'^n of the expansion are the contravariant components of vector $\mathbf{B}$ transformed into $\mathbf{x}'$ -space (the curvilinear coordinate space) by $\mathbf{B}' = \mathbf{R}\mathbf{B}$.

Inserting the last equation above applied at $\mathbf{x}' = \mathbf{x}'_{\text{far}}$,

$$B'^n(\mathbf{x}'_{\text{far}}) = \mathbf{e}^n(\mathbf{x}_{\text{far}}) \bullet \mathbf{B}(\mathbf{x}_{\text{far}})$$

into the previous $\approx$ equation then gives

$$d\mathbf{A}^n \bullet \mathbf{B}(\mathbf{x}_{\text{far}}) \approx \sqrt{g'(\mathbf{x}'_{\text{far}})} (\Pi_{i \neq n} dx^{i'}) B'^n(\mathbf{x}'_{\text{far}})$$

This far face n contribution to the flux integral is now expressed entirely in terms of $\mathbf{x}'$ -space objects and coordinates. A similar expression obtains for the near face n, but the sign of $d\mathbf{A}^n$ is reversed. Adding the contributions of these two faces of pair n gives

$$\int_{\text{two faces } n} d\mathbf{A} \bullet \mathbf{B}(\mathbf{x}) = \{ \sqrt{g'(\mathbf{x}'_{\text{far}})} B'^n(\mathbf{x}'_{\text{far}}) - \sqrt{g'(\mathbf{x}'_{\text{near}})} B'^n(\mathbf{x}'_{\text{near}}) \} (\Pi_{i \neq n} dx^{i'})$$

In $\mathbf{x}'$ -space, if $\mathbf{x}'_{\text{cen}}$ is a point at the center of the near face of face pair n, then

$$\begin{aligned} \mathbf{x}'_{\text{near}} &= \mathbf{x}'_{\text{cen}} \\ \mathbf{x}'_{\text{far}} &= \mathbf{x}'_{\text{cen}} + \mathbf{e}'_n dx'^n \end{aligned} \quad \text{where } \mathbf{e}'_n = \text{axis-aligned basis vector in } \mathbf{x}'\text{-space,}$$

since these two points map into the near and far face n centers in $\mathbf{x}$ -space. For any function f,

$$f(\mathbf{x}'_{\text{far}}) - f(\mathbf{x}'_{\text{near}}) = (\partial'_n f(\mathbf{x}')) dx'^n \quad // \text{ no implied sum on } n$$

where a change is made only in coordinate x'^n by amount dx'^n . Applying to $f = \sqrt{g'(\mathbf{x}')}$ yields

$$\{ \sqrt{g'(\mathbf{x}'_{\text{far}})} B'^n(\mathbf{x}'_{\text{far}}) - \sqrt{g'(\mathbf{x}'_{\text{near}})} B'^n(\mathbf{x}'_{\text{near}}) \} \approx \partial'_n [\sqrt{g'(\mathbf{x}'_{\text{cen}})} B'^n(\mathbf{x}'_{\text{cen}})] dx'^n$$

In the limit that differentials are very close to 0, replace $\mathbf{x}'_{\text{cen}}$ by $\mathbf{x}$. Then

$$\begin{aligned} \int_{\text{two faces } n} d\mathbf{A} \bullet \mathbf{B}(\mathbf{x}) &= \partial'_n [\sqrt{g'(\mathbf{x}')} B'^n(\mathbf{x}')] dx'^n (\Pi_{i \neq n} dx^{i'}) \\ &= \partial'_n [\sqrt{g'(\mathbf{x}')} B'^n(\mathbf{x}')] (\Pi_i dx^{i'}) \end{aligned}$$

where now all N differentials are present in $(\Pi_i dx'^i)$. The total flux flowing out through all N pairs of faces of the N -piped in x -space is this same result with an implicit sum on n , so

$$\text{total flux} = \int d\mathbf{A} \bullet \mathbf{B}(\mathbf{x}) = \partial'_n [\sqrt{g'(\mathbf{x}')} B'^n(\mathbf{x}')] (\Pi_i dx'^i) = \partial'_n [\sqrt{g'(\mathbf{x}')} B'^n(\mathbf{x}')] d\mathcal{V}'$$

where $d\mathcal{V}' = \Pi_i dx'^i$ 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 (c) item 7 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

$$[\text{div } \mathbf{B}](\mathbf{x}) = [1/\sqrt{g'(\mathbf{x}')}] \partial'_n [\sqrt{g'(\mathbf{x}')} B'^n(\mathbf{x}')] \quad // \text{ all } x'\text{-space coordinates and objects}$$

$$[\text{div } \mathbf{B}](\mathbf{x}) = \partial_n B^n(\mathbf{x}) \quad // \text{ all } x\text{-space coordinates and objects}$$

The added second line just shows $[\text{div } \mathbf{B}](\mathbf{x})$ expressed in terms of the Cartesian 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 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

$$\mathbf{B} = B'^n \mathbf{e}_n = B'^n h'_n \hat{\mathbf{e}}_n = \mathcal{B}'^n \hat{\mathbf{e}}_n \quad \text{where} \quad \mathcal{B}'^n \equiv B'^n h'_n$$

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:

$$B'^n = R^n_m B^m \quad (\mathcal{B}'^n / h'_n) = R^n_m (\mathcal{B}^n / 1) \quad \mathcal{B}'^n = h'_n (R^n_m \mathcal{B}^n)$$

Our Picture B results so far are these, assuming $\mathbf{B}$ is a tensorial vector,

General: $B'^n(\mathbf{x}') = R^n_m B^m(\mathbf{x}) \quad \mathbf{x} = F^{-1}(\mathbf{x}') \equiv \mathbf{x}(\mathbf{x}') \quad \mathbf{x}' = F(\mathbf{x}) \equiv \mathbf{x}'(\mathbf{x})$

$$\begin{aligned} [\text{div } \mathbf{B}](\mathbf{x}) &= [1/\sqrt{g'(\mathbf{x}')}] \partial'_n [\sqrt{g'(\mathbf{x}')} B'^n(\mathbf{x}')] & \mathbf{B} &= B'^n \mathbf{e}_n \\ [\text{div } \mathbf{B}](\mathbf{x}) &= [1/\sqrt{g'(\mathbf{x}')}] \partial'_n [\sqrt{g'(\mathbf{x}')} \mathcal{B}'^n(\mathbf{x}') / h'_n(\mathbf{x}')] & \mathbf{B} &= \mathcal{B}'^n \hat{\mathbf{e}}_n \\ [\text{div } \mathbf{B}](\mathbf{x}) &= [1/\sqrt{g'(\mathbf{x}')}] \partial'_n [\sqrt{g'(\mathbf{x}')} g'^{nm}(\mathbf{x}') B'_m(\mathbf{x}')] & \mathbf{B} &= B'_n \mathbf{e}^n \\ [\text{div } \mathbf{B}](\mathbf{x}) &= \partial_n B^n(\mathbf{x}) \quad // \quad B^n = \text{Cartesian components of } \mathbf{B} & \mathbf{B} &= B^n \hat{\mathbf{n}} \\ [\text{div } \mathbf{B}](\mathbf{x}) &= [\text{div } \mathbf{B}](\mathbf{x}') \end{aligned}$$

For orthogonal curvilinear coordinates, one has

$$g'_{ij} = h'_i{}^2 \delta_{i,j} \quad g'^{ij} = h'_i{}^{-2} \delta_{i,j} \quad \det(g'_{ij}) = \Pi_i h'_i{}^2 \quad \sqrt{g'} = (\Pi_i h'_i) = h'_1 h'_2 \dots h'_N$$

so the above expressions can be written (the arguments $\mathbf{x}'$ of the h'_n are now suppressed)

Orthogonal:

$$\begin{aligned} [\text{div } \mathbf{B}](\mathbf{x}) &= [1/(\Pi_i h'_i)] \partial'_n [(\Pi_i h'_i) B'^n(\mathbf{x}')] & \mathbf{B} &= B'^n \mathbf{e}_n \\ [\text{div } \mathbf{B}](\mathbf{x}) &= [1/(\Pi_i h'_i)] \partial'_n [(\Pi_i h'_i) \mathcal{B}'^n(\mathbf{x}') / h'_n] & \mathbf{B} &= \mathcal{B}'^n \hat{\mathbf{e}}_n \\ [\text{div } \mathbf{B}](\mathbf{x}) &= [1/(\Pi_i h'_i)] \partial'_n [(\Pi_i h'_i) B'_n(\mathbf{x}') / h'_n{}^2] & \mathbf{B} &= B'_n \mathbf{e}^n \\ [\text{div } \mathbf{B}](\mathbf{x}) &= \partial_n B^n(\mathbf{x}) \quad // \quad B^n = \text{Cartesian components of } \mathbf{B} & \mathbf{B} &= B^n \hat{\mathbf{n}} \\ [\text{div } \mathbf{B}](\mathbf{x}) &= [\text{div } \mathbf{B}](\mathbf{x}') \end{aligned}$$

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 $B'^n(\mathbf{x}') \equiv R^n_m B^m(\mathbf{x})$ and claim success. If such a definition does not result in an inconsistency, then such a $\mathbf{B}$ 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 F is non-linear so that R and $g' = RR^T$ are functions of $\mathbf{x}'$ and are not constants. Take $\mathbf{B}(\mathbf{x}) = \mathbf{x}$

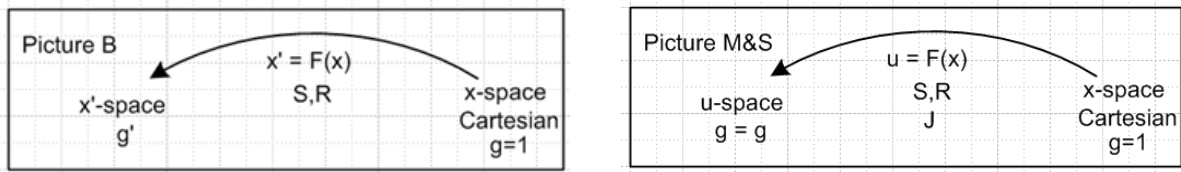
(the identity field) and try to make it be contravariant by definition, $\mathbf{x}'^n \equiv R^n_m \mathbf{x}^m$. The Cartesian divergence is then $\text{div } \mathbf{B} = \partial_n B^n(\mathbf{x}) = \partial_n x^n = 3$. But the first equation of the General block says

$$\text{div } \mathbf{B} = [1/\sqrt{g'(\mathbf{x}')}] \partial'_n [\sqrt{g'(\mathbf{x}')} x'^n] = \partial'_n x'^n + [1/\sqrt{g'(\mathbf{x}')}] x'^n \partial'_n \sqrt{g'(\mathbf{x}')} = 3 + \text{other stuff}$$

and thus the two calculations for $\text{div } \mathbf{B}$ disagree. As noted earlier, $\mathbf{x}' \equiv R\mathbf{x}$ conflicts with $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ in the case of non-linear F .

This comment can be applied to the tensorial character of the differential operators treated in later 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, $B'^n \rightarrow B^n$, $f' \rightarrow f$)

After this translation, all unprimed tensor components are functions of $\mathbf{x}$, while all italicized tensor components are functions of $\mathbf{u}$.

Here then are the translations of the two blocks above: (implied summation everywhere)

General: now $B^n(\mathbf{u}) = R^n_m B^m(\mathbf{x})$ $\mathbf{x} = F^{-1}(\mathbf{u}) \equiv \mathbf{x}(\mathbf{u})$ $\mathbf{u} = F(\mathbf{x}) \equiv \mathbf{u}(\mathbf{x})$

$$[\text{div } \mathbf{B}](\mathbf{x}) = [1/\sqrt{g}] \partial_n [\sqrt{g} B^n]$$

$$\mathbf{B} = B^n \mathbf{e}_n$$

$$[\text{div } \mathbf{B}](\mathbf{x}) = [1/\sqrt{g}] \partial_n [\sqrt{g} \mathcal{B}^n / h_n]$$

$$\mathbf{B} = \mathcal{B}^n \hat{\mathbf{e}}_n \quad // \text{ M\&S 1.06}$$

$$[\text{div } \mathbf{B}](\mathbf{x}) = [1/\sqrt{g}] \partial_n [\sqrt{g} g^{nm} B_m]$$

$$\mathbf{B} = B_n \mathbf{e}^n$$

$$[\text{div } \mathbf{B}](\mathbf{x}) = \partial_n B^n \quad // B^n = \text{Cartesian components of } \mathbf{B}$$

$$\mathbf{B} = B^n \hat{\mathbf{n}} = B_n \hat{\mathbf{n}}$$

$$[\text{div } \mathbf{B}](\mathbf{x}) = [\text{div } \mathbf{B}](\mathbf{u}) \quad // \text{ transformation (scalar)}$$

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: We use the "script MT bold" font $\mathcal{B}^n$ for components of vectors expanded onto $\hat{\mathbf{e}}_n$. It had to be something in upper case distinct from B^n and B^n . In practice one can replace $\mathcal{B}^n$ 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,θ,φ) 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 (M&F) 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

$$\operatorname{div} \mathbf{A} = \nabla \cdot \mathbf{A} = \frac{1}{h_1 h_2 h_3} \sum_n \frac{\partial}{\partial \xi_n} \left(h_1 h_2 h_3 \frac{A_n}{h_n} \right)$$

which compare to the above

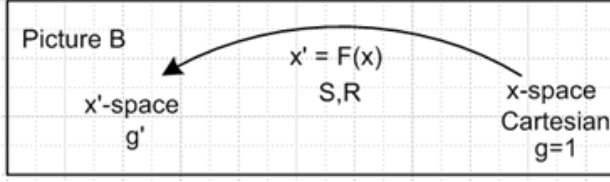
$$[\operatorname{div} \mathbf{A}](\mathbf{x}) = [div \mathbf{A}](\mathbf{u}) = [1/(h_1 h_2 h_3)] \partial_n [h_1 h_2 h_3 \mathcal{A}^n(\mathbf{u}) / h_n]$$

$$\mathbf{A} = \mathcal{A}^n \hat{\mathbf{e}}_n$$

so one should identify the M&F A_n with $\mathcal{A}^n$, the coefficient of $\hat{\mathbf{e}}_n$.

10. The Gradient in curvilinear coordinates

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

$$\mathbf{G} = G'_1 \mathbf{e}^1 + G'_2 \mathbf{e}^2 + \dots = \Sigma_{\mathbf{n}} G'_{\mathbf{n}} \mathbf{e}^{\mathbf{n}} \quad \text{where} \quad \mathbf{e}_{\mathbf{n}} \cdot \mathbf{G} = G'_{\mathbf{n}}$$

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

$$\mathcal{G}'^i(\mathbf{x}')/h'_{ij} = [\text{grad } f]^i(\mathbf{x}') = \partial'^i f'(\mathbf{x}') = g'^{ij}(\mathbf{x}') \partial'_j f'(\mathbf{x}') \quad \mathbf{G} = \text{grad } f = \mathcal{G}'^i \hat{\mathbf{e}}_i$$

Gathering up these results one gets

$$\begin{aligned} G'_i(\mathbf{x}') &= [\text{grad } f]_i(\mathbf{x}') = \partial'_i f'(\mathbf{x}') & \mathbf{G} = \text{grad } f &= G'_i(\mathbf{x}') \mathbf{e}^i \\ G'^i(\mathbf{x}') &= [\text{grad } f]^i(\mathbf{x}') = \partial'^i f'(\mathbf{x}') = g'^{ij}(\mathbf{x}') \partial'_j f'(\mathbf{x}') & \mathbf{G} = \text{grad } f &= G'^i(\mathbf{x}') \mathbf{e}_i \\ \mathcal{G}'^i(\mathbf{x}') &= h'_{ij} [\text{grad } f]^j(\mathbf{x}') = h'_{ij} \partial'^j f'(\mathbf{x}') = h'_{ij} \mathcal{G}'^j(\mathbf{x}') \partial'_j f'(\mathbf{x}') & \mathbf{G} = \text{grad } f &= \mathcal{G}'^i(\mathbf{x}') \hat{\mathbf{e}}_i \end{aligned}$$

which can be rewritten

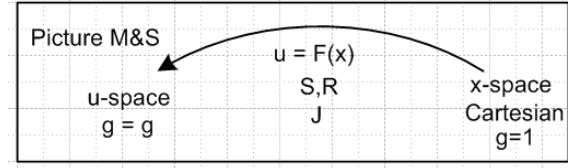
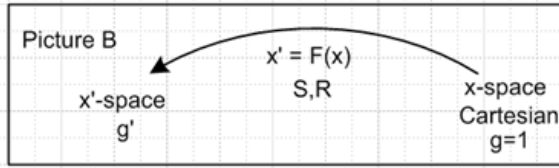
$$\begin{aligned} [\text{grad } f](\mathbf{x}) &= (\partial'_i f'(\mathbf{x}')) \mathbf{e}^i \\ [\text{grad } f](\mathbf{x}) &= (\partial'^i f'(\mathbf{x}')) \mathbf{e}_i = g'^{ij}(\mathbf{x}') (\partial'_j f'(\mathbf{x}')) \mathbf{e}_i \\ [\text{grad } f](\mathbf{x}) &= h'_{ij} (\partial'^i f'(\mathbf{x}')) \hat{\mathbf{e}}_i = h'_{ij} g'^{ij}(\mathbf{x}') (\partial'_j f'(\mathbf{x}')) \hat{\mathbf{e}}_i \\ [\text{grad } f](\mathbf{x}) &= (\partial_i f(\mathbf{x})) \hat{\mathbf{n}} = \nabla f(\mathbf{x}) \quad // \text{ Cartesian} \\ [\text{grad } f]_i(\mathbf{x}') &= R_i^j [\text{grad } f]_j(\mathbf{x}) \quad // \text{ transformation} \end{aligned} \quad f'(\mathbf{x}') = f(\mathbf{x})$$

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 $g'^{ij} = (1/h'^i{}^2) \delta_{i,j}$ so the block above becomes

$$\begin{aligned} [\text{grad } f](\mathbf{x}) &= (\partial'_i f'(\mathbf{x}')) \mathbf{e}^i \\ [\text{grad } f](\mathbf{x}) &= (\partial'^i f'(\mathbf{x}')) \mathbf{e}_i = (1/h'^i{}^2) (\partial'_i f'(\mathbf{x}')) \mathbf{e}_i \\ [\text{grad } f](\mathbf{x}) &= h'_{ij} (\partial'^i f'(\mathbf{x}')) \hat{\mathbf{e}}_i = (1/h'_{ij}) (\partial'_i f'(\mathbf{x}')) \hat{\mathbf{e}}_i \\ [\text{grad } f](\mathbf{x}) &= (\partial_i f(\mathbf{x})) \hat{\mathbf{n}} \quad // \text{ Cartesian} \\ [\text{grad } f]_i(\mathbf{x}') &= R_i^j [\text{grad } f]_j(\mathbf{x}) \quad // \text{ transformation} \end{aligned} \quad f'(\mathbf{x}') = f(\mathbf{x})$$

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

$$\text{grad } f = \mathbf{G} = [g'^{nm} (\partial'_n f')] \mathbf{e}_m = G'^m \mathbf{e}_m \quad G'^m = g'^{nm} (\partial'_n f') \quad \partial'_n f' = \partial'_n f(\mathbf{x}(x')) = \partial'_n f'(x')$$

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:

$$\begin{aligned} [\text{lap } f](\mathbf{x}) &= [1/\sqrt{g'(\mathbf{x}')}] \partial'_m [\sqrt{g'(\mathbf{x}')} g'^{nm}(\mathbf{x}') (\partial'_n f'(\mathbf{x}'))] & // \text{ implied sum on } n \text{ and } m \\ [\text{lap } f](\mathbf{x}) &= \sum_n \partial_n^2 f(\mathbf{x}) & // \text{ Cartesian} \\ [\text{lap } f](\mathbf{x}) &= [\text{lap } f](\mathbf{x}') & f'(\mathbf{x}') = f(\mathbf{x}) = f(\mathbf{x}(x')) \end{aligned}$$

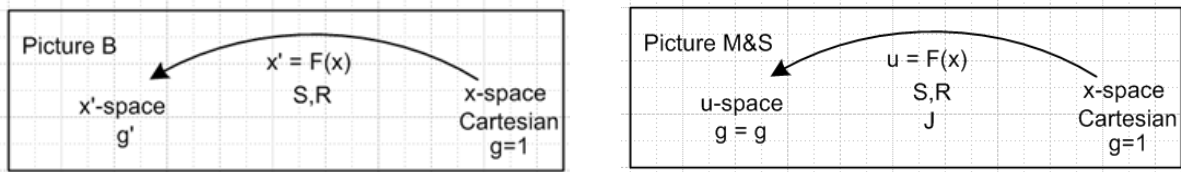
For an *orthogonal* coordinate system,

$$g'_{nm} = h'_n{}^2 \delta_{n,m} \quad g'^{nm} = (1/h'_n{}^2) \delta_{n,m} \quad \sqrt{g'} = \prod_i h'_i$$

and the first line above simplifies to

$$[\text{lap } f](\mathbf{x}) = [1/(\prod_i h'_i)] \partial'_m [(\prod_i h'_i) (1/h'_m{}^2) (\partial'_m f')]]$$

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,

$$[\text{lap } f](\mathbf{x}) = 1/(h_1 h_2 h_3) \{ \partial_1 [(h_2 h_3 / h_1) \partial_1 f] + \text{cyclic} \}$$

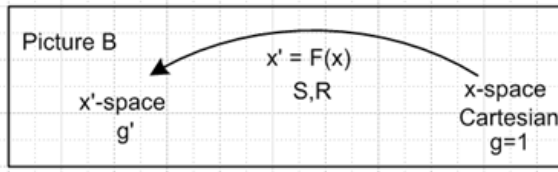
With the replacements

$$f \rightarrow \phi \quad h_i^2 \rightarrow g_{ii} \quad (\Pi_i h_i) \rightarrow g$$

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.



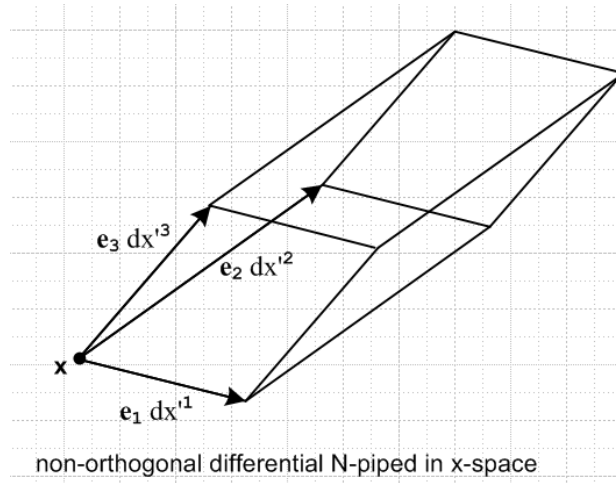
In Cartesian coordinates one writes

$$[\text{curl } \mathbf{B}]_i(\mathbf{x}) = [\nabla \times \mathbf{B}(\mathbf{x})]_i = \varepsilon_{ijk} \partial_j B_k(\mathbf{x})$$

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) item 1, the vector area associated with the face pair n is ($|J| = g^{1/2}$ when $g = 1$)

$$d\mathbf{A}^n = |\det(\mathbf{S}^i_j)| \mathbf{e}^n (\Pi_{i \neq n} dx'^i) = |J| \mathbf{e}^n (\Pi_{i \neq n} dx'^i) = g^{1/2} \mathbf{e}^n (\Pi_{i \neq n} dx'^i) \quad n = 1, 2, 3$$

where $J = \det(\mathbf{S}^i_j)$ 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 $\mathbf{e}_k$. As shown in Appendix D (h), $\mathbf{C} = \text{curl } \mathbf{B}$ is in fact a vector field *density* of weight -1. For an ordinary vector one would write the expansion as $\mathbf{C} = \sum_k C'^k \mathbf{e}_k$, but the rule for tensor densities says to replace $C'^k \rightarrow J^{-1} C'^k$, as explained toward the end of Appendix E (b). Therefore,

$$\mathbf{C} = J^{-1} \sum_{k=1}^3 C'^k \mathbf{e}_k \quad J^{-1} = (1/\sqrt{g'})$$

so that

$$\mathbf{C} \cdot d\mathbf{A}^n = J^{-1} (\sum_k C'^k \mathbf{e}_k) \cdot (g'^{1/2} \mathbf{e}^n (\prod_{i \neq n} dx'^i)) = C'^n (\prod_{i \neq n} dx'^i)$$

and then

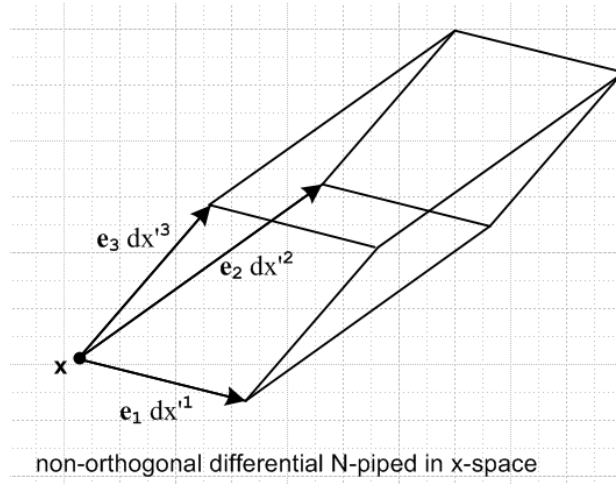
$$C'^n(\mathbf{x}') (\prod_{i \neq n} dx'^i) = (\oint \mathbf{B} \cdot d\mathbf{x})^n \quad // \text{ "curl" = circulation/area}$$

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.

Comment: This factor of J^{-1} won't appear in the final results, as will be seen, but we must include it in the analysis in order to claim that C'^k are the components of vector $\mathbf{C}$ 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 $dx^n \rightarrow 0$, 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 $f(\mathbf{x} + d\mathbf{x}') \approx f(\mathbf{x}') + \Sigma_n \partial_n f(\mathbf{x}') d\mathbf{x}'^n$ 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-Civita ε tensor. In Cartesian space, the up and down position of the indices does not matter, as for any tensor. In non-Cartesian space up and down *does* matter, as with any tensor. The only fact needed here is 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

$$C'^n(\mathbf{x}') (\Pi_{i \neq n} dx'^i) = (\oint \mathbf{B} \cdot d\mathbf{x})^n$$

Insert the section (b) result for $(\oint \mathbf{B} \cdot d\mathbf{x})^n$ to get

$$C'^n(\mathbf{x}') (\Pi_{i \neq n} dx'^i) = \varepsilon'^{nab} \partial'_a B'_b (\Pi_{i \neq n} dx'^i)$$

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

$$\begin{aligned}
 C'^n &= \varepsilon'^{nab} \partial'_a B'_b & \mathbf{C} &= J^{-1} C'^n \mathbf{e}_n & \mathbf{B} &= B'_n \mathbf{e}^n \\
 \text{curl } \mathbf{B} = \mathbf{C} &= [(1/\sqrt{g'})] \varepsilon'^{nab} \partial'_a B'_b \mathbf{e}_n
 \end{aligned}$$

The comparison between the curvilinear and Cartesian expressed curls is this:

$$[\text{curl } \mathbf{B}](\mathbf{x}) = [(1/\sqrt{g'}) \varepsilon^{nab} \partial'_a B'_b(\mathbf{x}')] \mathbf{e}_n = (1/\sqrt{g'}) \{ [\partial'_1 B'_2 - \partial'_2 B'_1] \mathbf{e}_3 + \text{cyclic} \}$$

$$[\text{curl } \mathbf{B}](\mathbf{x}) = \varepsilon^{nab} \partial_a B_b(\mathbf{x}) \hat{\mathbf{n}} = \{ [\partial_1 B_2 - \partial_2 B_1] \hat{\mathbf{3}} + \text{cyclic} \}$$

Comment 1: 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.

Comment 2: The equation $\text{curl } \mathbf{B} = \mathbf{C} = [(1/\sqrt{g'}) \varepsilon^{nab} \partial'_a B'_b] \mathbf{e}_n$ obtained above assumed that $\mathbf{B}$ was a vector and the expansion $\mathbf{B} = B'_n \mathbf{e}^n$ was used. If $\mathbf{B}$ were a vector density of weight -1, the expansion would be $\mathbf{B} = J^{-1} B'_n \mathbf{e}^n$ and the result would be $\text{curl } \mathbf{B} = \mathbf{C} = [(1/\sqrt{g'}) \varepsilon^{nab} \partial'_a (J^{-1} B'_b)] \mathbf{e}_n$. This situation will arise in consideration of the vector Laplacian in Section 13. Once again, $J = \sqrt{g'}$.

(d) Various forms of the curl

The first form is that just presented above, with $\mathbf{C} = \text{curl} \mathbf{B}$,

$$\begin{aligned} C'^n &= \varepsilon^{nab} \partial'_a B'_b & \mathbf{C} &= J^{-1} C'^n \mathbf{e}_n & \mathbf{B} &= B'_n \mathbf{e}^n \\ \text{curl } \mathbf{B} &= [(1/\sqrt{g'}) \varepsilon^{nab} \partial'_a B'_b] \mathbf{e}_n \end{aligned}$$

If it is desired to have contravariant components of $\mathbf{B}$, one gets

$$\begin{aligned} C'^n &= \varepsilon^{nab} \partial'_a (g'_{bc} B'^c) & \mathbf{C} &= J^{-1} C'^n \mathbf{e}_n & \mathbf{B} &= B'^n \mathbf{e}_n \\ \text{curl } \mathbf{B} &= [(1/\sqrt{g'}) \varepsilon^{nab} \partial'_a (g'_{bc} B'^c)] \mathbf{e}_n \end{aligned}$$

For practical applications, one usually wants both vectors expanded on the $\hat{\mathbf{e}}_n$ unit vectors in this way

$$\begin{aligned} \mathbf{C} &= J^{-1} C'^n \mathbf{e}_n = (J^{-1} C'^n h'_n) \hat{\mathbf{e}}_n \equiv \mathcal{C}'^n \hat{\mathbf{e}}_n & \mathcal{C}'^n &= h'_n J^{-1} C'^n \\ \mathbf{B} &= B'^n \mathbf{e}_n = (B'^n h'_n) \hat{\mathbf{e}}_n \equiv \mathcal{B}'^n \hat{\mathbf{e}}_n & \mathcal{B}'^n &= h'_n B'^n \end{aligned}$$

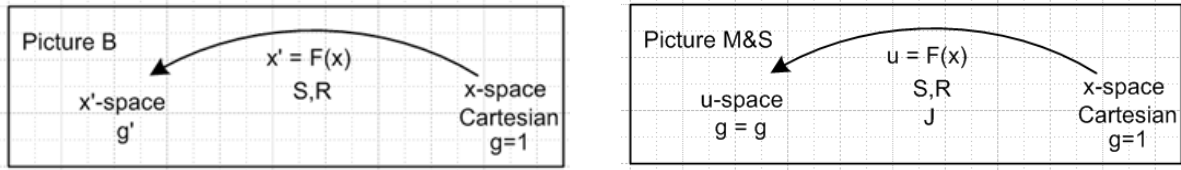
so that

$$\begin{aligned} \mathcal{C}'^n &= [(1/\sqrt{g'}) h'_n \varepsilon^{nab} \partial'_a (g'_{bc} \mathcal{B}'^c / h'_c)] & \mathbf{C} &= \mathcal{C}'^n \hat{\mathbf{e}}_n & \mathbf{B} &= \mathcal{B}'^n \hat{\mathbf{e}}_n \\ \text{curl } \mathbf{B} &= [(1/\sqrt{g'}) h'_n \varepsilon^{nab} \partial'_a (g'_{bc} \mathcal{B}'^c / h'_c)] \hat{\mathbf{e}}_n & \text{curl } \mathbf{B} &= \mathbf{C} \end{aligned}$$

To summarize: $B'^c = R^c_d B^d$ $g' = g'(x')$ etc.

$$\begin{aligned}
 [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon'^{nab} [(1/\sqrt{g'}) \partial'_a B'_b] \mathbf{e}_n & \mathbf{B} &= B'_n \mathbf{e}^n \\
 [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon'^{nab} [(1/\sqrt{g'}) \partial'_a (g'_{bc} B'^c)] \mathbf{e}_n & \mathbf{B} &= B'^n \mathbf{e}_n \\
 [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon'^{nab} [(1/\sqrt{g'}) h'_n \partial'_a (g'_{bc} \mathcal{B}'^c / h'_c)] \hat{\mathbf{e}}_n & \mathbf{B} &= \mathcal{B}'^n \hat{\mathbf{e}}_n \\
 [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} \partial_a B_b(\mathbf{x}) \hat{\mathbf{n}} & // \text{ Cartesian} & \mathbf{B} = B^n \hat{\mathbf{n}}
 \end{aligned}$$

Converting from Picture B to Picture MS (see Section 9 (c))



one gets :

$$\begin{aligned}
 [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} [(1/\sqrt{g}) \partial_a B_b] \mathbf{e}_n & \mathbf{B} &= B_n \mathbf{e}^n \\
 [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} [(1/\sqrt{g}) \partial_a (g_{bc} B^c)] \mathbf{e}_n & \mathbf{B} &= B^n \mathbf{e}_n \\
 [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} [(1/\sqrt{g}) h_n \partial_a (g_{bc} \mathcal{B}^c / h_c)] \hat{\mathbf{e}}_n & \mathbf{B} &= \mathcal{B}^n \hat{\mathbf{e}}_n \\
 [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} \partial_a B_b(\mathbf{x}) \hat{\mathbf{n}} & // \text{ Cartesian} & \mathbf{B} = B^n \hat{\mathbf{n}}
 \end{aligned}$$

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 $\det(\mathbf{Q}) \equiv \sum_i \mathbf{Q}_{1i} \text{cof}(\mathbf{Q}_{1i})$:

$$\begin{aligned}
 [\text{curl } \mathbf{B}](\mathbf{x}) &= (1/\sqrt{g}) \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ B_1 & B_2 & B_3 \end{vmatrix} & \mathbf{B} &= B_n \mathbf{e}^n \\
 [\text{curl } \mathbf{B}](\mathbf{x}) &= (1/\sqrt{g}) \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ g_{1c} B^c & g_{2c} B^c & g_{3c} B^c \end{vmatrix} & \mathbf{B} &= B^n \mathbf{e}_n \\
 [\text{curl } \mathbf{B}](\mathbf{x}) &= (1/\sqrt{g}) \begin{vmatrix} h_1 \hat{\mathbf{e}}_1 & h_2 \hat{\mathbf{e}}_2 & h_3 \hat{\mathbf{e}}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ (g_{1c}/h_c) \mathcal{B}^c & (g_{2c}/h_c) \mathcal{B}^c & (g_{3c}/h_c) \mathcal{B}^c \end{vmatrix} & \mathbf{B} &= \mathcal{B}^n \hat{\mathbf{e}}_n \quad // \text{ M\&S 1.07} \\
 [\text{curl } \mathbf{B}](\mathbf{x}) &= \begin{vmatrix} \hat{\mathbf{1}} & \hat{\mathbf{2}} & \hat{\mathbf{3}} \\ \partial_1 & \partial_2 & \partial_3 \\ B_1 & B_2 & B_3 \end{vmatrix} & // \text{ here } \partial_n = \partial/\partial x^n \text{ and } B_i = B_i(\mathbf{x}) & \mathbf{B} &= B^n \hat{\mathbf{n}}
 \end{aligned}$$

(e) The curl in orthogonal coordinate systems

For such systems $g_{ij} = \delta_{ij} h_i^2$ 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: $B^c(\mathbf{x}') = R^c_d B^d(\mathbf{x}) \quad h_i' = h_i'(\mathbf{x}') \text{ etc.}$

$$\begin{aligned} [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon'^{nab} [(1/(h_1' h_2' h_3')) \partial'_a B'_b] \mathbf{e}_n & \mathbf{B} &= B'_n \mathbf{e}'_n \\ [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon'^{nab} [(1/(h_1' h_2' h_3')) \partial'_a (h'^2_b B'^b)] \mathbf{e}_n & \mathbf{B} &= B'^n \mathbf{e}_n \\ [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon'^{nab} [(1/(h_1' h_2' h_3')) h'_n \partial'_a (h'^b_b \mathcal{B}^b)] \hat{\mathbf{e}}_n & \mathbf{B} &= \mathcal{B}'^n \hat{\mathbf{e}}_n \\ [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} \partial_a B_b(\mathbf{x}) \hat{\mathbf{n}} & // \text{ Cartesian} & \quad \mathbf{B} = B^n \hat{\mathbf{n}} \end{aligned}$$

Picture M&S: $\mathcal{B}^c(\mathbf{u}) = R^c_d B^d(\mathbf{x}) \quad h_i = h_i(\mathbf{u}) \text{ etc.}$

$$\begin{aligned} [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} [(h_1 h_2 h_3)^{-1} \partial_a B_b] \mathbf{e}_n & \mathbf{B} &= B_n \mathbf{e}^n \\ [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} [(h_1 h_2 h_3)^{-1} \partial_a (h^2_b B^b)] \mathbf{e}_n & \mathbf{B} &= B^n \mathbf{e}_n \\ [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} [(h_1 h_2 h_3)^{-1} h_n \partial_a (h^b_b \mathcal{B}^b)] \hat{\mathbf{e}}_n & \mathbf{B} &= \mathcal{B}^n \hat{\mathbf{e}}_n \end{aligned}$$

$$[\text{curl } \mathbf{B}](\mathbf{x}) = (h_1 h_2 h_3)^{-1} \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ B_1 & B_2 & B_3 \end{vmatrix} \quad \mathbf{B} = B_n \mathbf{e}^n$$

$$[\text{curl } \mathbf{B}](\mathbf{x}) = (h_1 h_2 h_3)^{-1} \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ h_1^2 B^1 & h_2^2 B^2 & h_3^2 B^3 \end{vmatrix} \quad \mathbf{B} = B^n \mathbf{e}_n$$

$$[\text{curl } \mathbf{B}](\mathbf{x}) = (h_1 h_2 h_3)^{-1} \begin{vmatrix} h_1 \hat{\mathbf{e}}_1 & h_2 \hat{\mathbf{e}}_2 & h_3 \hat{\mathbf{e}}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ h_1 \mathcal{B}^1 & h_2 \mathcal{B}^2 & h_3 \mathcal{B}^3 \end{vmatrix} \quad \mathbf{B} = \mathcal{B}^n \hat{\mathbf{e}}_n \quad // \text{ M\&S 1.07a}$$

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

$$[\text{curl } \mathbf{B}]^n(\mathbf{x}) = (1/\sqrt{g}) \epsilon^{nabc} \partial_a (\partial_b B_c) = (1/\sqrt{g}) \epsilon^{nabc} \partial_a \partial_b B_c \quad ?$$

$$[\text{curl } \mathbf{B}]^n(\mathbf{x}) = (1/\sqrt{g}) \epsilon^{nabc} \partial_a (B_b B_c) \quad ?$$

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 $B_{b;a} - B_{a;b} = \partial_a B_b - \partial_b B_a$ 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,

$$[\text{curl } \mathbf{B}]^n = \epsilon^{nab} [B_{b;a} - B_{a;b}] / 2 = \epsilon^{nab} B_{b;a} = \epsilon^{nab} [\partial_a B_b - \partial_b B_a] / 2 = \epsilon^{nab} \partial_a B_b.$$

In relativity work, since $N=4$, there *is* no vector curl, and one sees $B_{b;a} - B_{a;b}$ referred to as the covariant curl, and $\partial_a B_b - \partial_b B_a$ as the ordinary curl (Weinberg p 106).

Writing the N -dimensional contravariant curl components in this manner in Cartesian x -space,

$$[\text{curl } B]^{ij} = (B^{j;i} - B^{i;j})$$

one can then ask how this generalized curl would be expressed in terms of x' -space coordinates and objects. (This curl is a regular rank-2 tensor with weight 0, the vector curl had weight -1).

The general issue of expanding tensors is addressed in Appendix E where this general result is obtained

$$A = \Sigma_{ijk\dots} A^{ijk\dots} (\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \dots) \quad A^{ijk\dots} = \text{contravariant components of } A \text{ in } x'\text{-space}$$

Applying this to $A = [\text{curl } B]$ one gets

$$[\text{curl } B] = \Sigma_{ij} [\text{curl } B]^{ij} \mathbf{e}_i \otimes \mathbf{e}_j = \Sigma_{ij} (B^{j;i} - B^{i;j}) \mathbf{e}_i \otimes \mathbf{e}_j$$

The Cartesian components of curl B in x -space can then be expressed in terms of x' -space components and coordinates,

$$\begin{aligned} [\text{curl } B]^{ab}(\mathbf{x}) &= \Sigma_{ij} [\text{curl } B]^{ij} (\mathbf{e}_i)^a (\mathbf{e}_j)^b \\ &= \Sigma_{ij} (B^{j;i} - B^{i;j}) (\mathbf{e}_i)^a (\mathbf{e}_j)^b \end{aligned}$$

where the tangent base vectors $\mathbf{e}_n$ exist as usual in x -space, and $B^{i\dot{}} = B^{i\dot{}}(\mathbf{x}')$ where $\mathbf{x}' = \mathbf{F}(\mathbf{x})$.

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

$$\nabla^2 \mathbf{B} \equiv \text{grad}(\text{div } \mathbf{B}) - \text{curl}(\text{curl } \mathbf{B}).$$

If $\mathbf{B}$ is a vector, $\text{div } \mathbf{B}$ is a scalar, $\text{grad}(\text{div } \mathbf{B})$ is a vector, so we expect $\nabla^2 \mathbf{B}$ to be a vector. An annoying observation is that $\text{curl } \mathbf{B}$ is a vector density of weight -1 (see Appendix D (h)) and $\text{curl}(\text{curl } \mathbf{B})$ a vector density of weight -2, and so apples are being subtracted from oranges to make $\nabla^2 \mathbf{B}$. This tensor conundrum is resolved in Section 15 (g) so we ignore it for now and proceed undaunted.

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 (a) 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

$$\begin{aligned} \Rightarrow \mathbf{G} &= \partial'_k \{ (1/\sqrt{g'}) \partial'_i (\sqrt{g'} B'^i) \} \mathbf{e}^k \\ &= \partial'^n \{ (1/\sqrt{g'}) \partial'_i (\sqrt{g'} B'^i) \} \mathbf{e}_n \end{aligned}$$

The second term $\mathbf{V}$ is little more complicated. First, from Section 12 (d),

$$\begin{aligned} \mathbf{C} = \text{curl } \mathbf{B} &= \varepsilon'^{nab} [(1/\sqrt{g'}) \partial'_a \{ B'_b \}] \mathbf{e}_n = J^{-1} C'^n \mathbf{e}_n \quad J = \sqrt{g'} \\ \mathbf{V} = \text{curl } \mathbf{C} &= \varepsilon'^{ncd} [(1/\sqrt{g'}) \partial'_c \{ J^{-1} C'_d \}] \mathbf{e}_n = J^{-1} V'^n \mathbf{e}_n \quad (*) \end{aligned}$$

where recall from Appendix D (d) that $\varepsilon'^{abc\dots} = \varepsilon^{abc\dots}$ = the usual permutation tensor, but written up and primed so as to be in covariant form. The reason for the factor J^{-1} in $\partial'_c \{ J^{-1} C'_d \}$ was explained in Comment 2 at the end of Section 12 (c) (namely, $\mathbf{C}$ is a vector density of weight -1). The above lines can be rewritten as

$$\begin{aligned} C'^n &= \varepsilon'^{nab} \partial'_a \{ B'_b \} \\ V'^n &= \varepsilon'^{ncd} \partial'_c \{ J^{-1} C'_d \} \quad // J = \sqrt{g'} \end{aligned}$$

Then write

$$C'_d = g'_{de} C'^e = g'_{de} \varepsilon'^{eab} (\partial'_a B'_b)$$

and insert this into the V'_n equation to get

$$\begin{aligned} V'^n &= \varepsilon'^{ncd} \partial'_c \{ (1/\sqrt{g'}) g'_{de} \varepsilon'^{eab} (\partial'_a B'_b) \} \\ &= \varepsilon'^{ncd} \varepsilon'^{eab} \partial'_c \{ (1/\sqrt{g'}) g'_{de} (\partial'_a B'_b) \} \end{aligned}$$

so

$$\mathbf{V} = \text{curl } \mathbf{C} = J^{-1} V'^n \mathbf{e}_n = [(1/\sqrt{g'}) \varepsilon'^{ncd} \varepsilon'^{eab} \partial'_c \{ (1/\sqrt{g'}) g'_{de} (\partial'_a B'_b) \}] \mathbf{e}_n$$

We can then summarize our results:

$$\begin{aligned} \star \mathbf{B} &= \mathbf{G} - \mathbf{V} \\ \mathbf{G} &= \partial'^n \{ (1/\sqrt{g'}) \partial'_i (\sqrt{g'} B'^i) \} \mathbf{e}_n & \mathbf{B} &= B'^n \mathbf{e}_n \\ \mathbf{V} &= [(1/\sqrt{g'}) \varepsilon'^{ncd} \varepsilon'^{eab} \partial'_c \{ (1/\sqrt{g'}) g'_{de} (\partial'_a B'_b) \}] \mathbf{e}_n & \mathbf{B} &= B'_n \mathbf{e}^n \\ &= [(1/\sqrt{g'}) \varepsilon'^{ncd} \varepsilon'^{eab} \partial'_c \{ (1/\sqrt{g'}) g'_{de} (\partial'_a [g'_{bf} B'^f]) \}] \mathbf{e}_n & \mathbf{B} &= B'^n \mathbf{e}_n \end{aligned}$$

Often one wants vectors expanded onto the unit vectors $\hat{\mathbf{e}}_n$

$$\mathbf{B} = B'^n \mathbf{e}_n = B'^n h'_n \hat{\mathbf{e}}_n = \mathcal{B}'^n \hat{\mathbf{e}}_n \quad \Rightarrow \quad B'^n = \mathcal{B}'^n / h'_n$$

and the above set of equations becomes

$$\star \mathbf{B} = \mathbf{G} - \mathbf{V}$$

$$\mathbf{G} = h'_n \partial^n \{ (1/\sqrt{g'}) \partial'_i (\sqrt{g'} \mathcal{B}^{i'} / h'_{i'}) \} \hat{\mathbf{e}}_n$$

$$\mathbf{B} = \mathcal{B}^n \hat{\mathbf{e}}_n$$

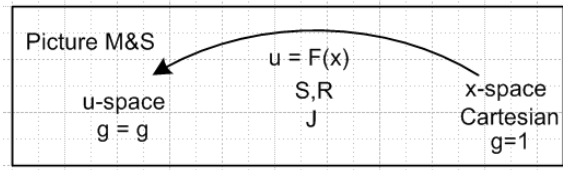
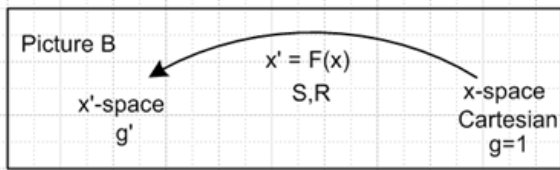
$$\mathbf{V} = h'_n [(1/\sqrt{g'}) \varepsilon^{ncd} \varepsilon^{eab} \partial'_c \{ (1/\sqrt{g'}) g'_{de} (\partial'_a [g'_{bf} \mathcal{B}^{f'} / h'_{f'}]) \}] \hat{\mathbf{e}}_n$$

and to this list we can add the Cartesian form

$$\star \mathbf{B} = [\nabla^2 B^n] \hat{\mathbf{n}} \quad // \text{ Cartesian, } \hat{\mathbf{n}} = \mathbf{u}_n$$

$$\mathbf{B} = B^n \hat{\mathbf{n}}$$

Converting from Picture B to Picture MS gives (see Section 9 (c))



$$\star \mathbf{B} = \mathbf{G} - \mathbf{V}$$

$$\mathbf{G} = \partial^n \{ (1/\sqrt{g}) \partial_i (\sqrt{g} B^i) \} \mathbf{e}_n$$

$$\mathbf{B} = B^n \mathbf{e}_n$$

$$\mathbf{V} = [(1/\sqrt{g}) \varepsilon^{ncd} \varepsilon^{eab} \partial_c \{ (1/\sqrt{g}) g_{de} (\partial_a [g_{bf} B^f]) \}] \mathbf{e}_n$$

$$\mathbf{G} = h_n \partial^n \{ (1/\sqrt{g}) \partial_i (\sqrt{g} \mathcal{B}^i / h_i) \} \hat{\mathbf{e}}_n$$

$$\mathbf{B} = \mathcal{B}^n \hat{\mathbf{e}}_n$$

$$\mathbf{V} = h_n [(1/\sqrt{g}) \varepsilon^{ncd} \varepsilon^{eab} \partial_c \{ (1/\sqrt{g}) g_{de} (\partial_a [g_{bf} \mathcal{B}^f / h_f]) \}] \hat{\mathbf{e}}_n$$

$$\star \mathbf{B} = [\nabla^2 B^n] \hat{\mathbf{n}} \quad // \text{ Cartesian, } \hat{\mathbf{n}} = \mathbf{u}_n$$

$$\mathbf{B} = B^n \hat{\mathbf{n}}$$

In the 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&S and shall use only the $\hat{\mathbf{e}}_n$ expansion. Setting $g_{ij} = h_i^2 \delta_{i,j}$ and $g^{ij} = (1/h_i)^2 \delta_{i,j}$ things simplify somewhat,

$$\begin{aligned} \star \mathbf{B} &= h_n \hat{\mathbf{e}}_n [g^{nk} \partial_k \{ (1/\sqrt{g}) \partial_i (\sqrt{g} \mathcal{B}^i(\mathbf{u}) / h_i) \} \\ &\quad - (1/\sqrt{g}) \varepsilon^{ncd} \varepsilon^{eab} \partial_c \{ (1/\sqrt{g}) g_{de} (\partial_a [g_{bf} \mathcal{B}^f(\mathbf{u}) / h_f]) \}] \\ &= h_n \hat{\mathbf{e}}_n [\delta_{k,n} \partial_k \{ (1/\sqrt{g}) \partial_i (\sqrt{g} \mathcal{B}^i(\mathbf{u}) / h_i) \} (1/h_n)^2 \\ &\quad - (1/\sqrt{g}) \varepsilon^{ncd} \varepsilon^{eab} \partial_c \{ (1/\sqrt{g}) h_d^2 \delta_{d,e} (\partial_a [h_b^2 \delta_{b,f} \mathcal{B}^f(\mathbf{u}) / h_f]) \}] \\ &= h_n \hat{\mathbf{e}}_n [\partial_n \{ (1/\sqrt{g}) \partial_i (\sqrt{g} \mathcal{B}^i(\mathbf{u}) / h_i) \} (1/h_n)^2 \end{aligned}$$

$$\begin{aligned}
& - (1/\sqrt{g}) \varepsilon^{ncd} \varepsilon^{dab} \partial_c \{ (1/\sqrt{g}) h_d^2 (\partial_a [h_b \mathcal{B}^b(u)]) \}] \\
& = \hat{\mathbf{e}}_n [(1/h_n) \partial_n \{ (1/\sqrt{g}) \partial_i (\sqrt{g} \mathcal{B}^i/h_i) \} - (h_n/\sqrt{g}) \varepsilon^{ncd} \varepsilon^{dab} \partial_c \{ (1/\sqrt{g}) h_d^2 (\partial_a [h_b \mathcal{B}^b]) \}]
\end{aligned}$$

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

$$\begin{aligned}
& \varepsilon^{1cd} \varepsilon^{dab} \partial_c \{ (1/\sqrt{g}) h_d^2 (\partial_a [h_b \mathcal{B}^b]) \} \\
& = \varepsilon^{3ab} \partial_2 \{ (1/\sqrt{g}) h_3^2 (\partial_a [h_b \mathcal{B}^b]) \} - \varepsilon^{2ab} \partial_3 \{ (1/\sqrt{g}) h_2^2 (\partial_a [h_b \mathcal{B}^b]) \} \\
& \quad \quad \quad c=2 \quad d=3 \quad \quad \quad c=3 \quad d=2 \\
& = [\quad \partial_2 \{ (1/\sqrt{g}) h_3^2 (\partial_1 [h_2 \mathcal{B}^2]) \} \quad - \quad \partial_2 \{ (1/\sqrt{g}) h_3^2 (\partial_2 [h_1 \mathcal{B}^1]) \} \quad] \\
& \quad \quad \quad a=1 \quad b=2 \quad \quad \quad a=2 \quad b=1 \\
& - [\quad \partial_3 \{ (1/\sqrt{g}) h_2^2 (\partial_3 [h_1 \mathcal{B}^1]) \} \quad - \quad \partial_3 \{ (1/\sqrt{g}) h_2^2 (\partial_1 [h_3 \mathcal{B}^3]) \} \quad] \\
& \quad \quad \quad a=3 \quad b=1 \quad \quad \quad a=1 \quad b=3 \\
& = \partial_2 \{ (1/\sqrt{g}) h_3^2 (\partial_1 [h_2 \mathcal{B}^2] - \partial_2 [h_1 \mathcal{B}^1]) \} \\
& - \partial_3 \{ (1/\sqrt{g}) h_2^2 (\partial_3 [h_1 \mathcal{B}^1] - \partial_1 [h_3 \mathcal{B}^3]) \}
\end{aligned}$$

Define now Γ_n in cyclic fashion.

$$\begin{aligned}
\Gamma_1 & \equiv (1/\sqrt{g}) h_1^2 (\partial_2 [h_3 \mathcal{B}^3] - \partial_3 [h_2 \mathcal{B}^2]) \\
\Gamma_2 & \equiv (1/\sqrt{g}) h_2^2 (\partial_3 [h_1 \mathcal{B}^1] - \partial_1 [h_3 \mathcal{B}^3]) \\
\Gamma_3 & \equiv (1/\sqrt{g}) h_3^2 (\partial_1 [h_2 \mathcal{B}^2] - \partial_2 [h_1 \mathcal{B}^1])
\end{aligned}$$

and then we have shown that

$$\begin{aligned}
\text{2nd term (n=1)} & = -(h_1/\sqrt{g}) \varepsilon^{1cd} \varepsilon^{dab} \partial_c \{ (1/\sqrt{g}) h_d^2 (\partial_a [h_b \mathcal{B}^b]) \} \hat{\mathbf{e}}_1 \\
& = -(h_1/\sqrt{g}) (\partial_2 \Gamma_3 - \partial_3 \Gamma_2) \hat{\mathbf{e}}_1 = + (h_1/\sqrt{g}) (\partial_3 \Gamma_2 - \partial_2 \Gamma_3) \hat{\mathbf{e}}_1
\end{aligned}$$

Therefore the entire first term (n=1) of $\star \mathbf{B}$ is given by

$$\star \mathbf{B} \text{ (first term)} = [(1/h_1) \partial_1 T + (h_1/\sqrt{g}) (\partial_3 \Gamma_2 - \partial_2 \Gamma_3)] \hat{\mathbf{e}}_1$$

The other two terms are obtained by cyclic permutation so the final result is then

$$\begin{aligned}
 [\star \mathbf{B}](\mathbf{x}) &= [(1/h_1) \partial_1 T + (h_1/\sqrt{g}) (\partial_3 \Gamma_2 - \partial_2 \Gamma_3)] \hat{\mathbf{e}}_1 + \text{cyclic} \\
 &= [(1/h_1) \partial_1 T + (h_1/\sqrt{g}) (\partial_3 \Gamma_2 - \partial_2 \Gamma_3)] \hat{\mathbf{e}}_1 \\
 &\quad + [(1/h_2) \partial_2 T + (h_2/\sqrt{g}) (\partial_1 \Gamma_3 - \partial_3 \Gamma_1)] \hat{\mathbf{e}}_2 \\
 &\quad + [(1/h_3) \partial_3 T + (h_3/\sqrt{g}) (\partial_2 \Gamma_1 - \partial_1 \Gamma_2)] \hat{\mathbf{e}}_3 \quad // \text{M\&S 1.11}
 \end{aligned}$$

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

$$\mathbf{B} \rightarrow \mathbf{E} \quad \mathcal{B}^n \rightarrow E_n \quad \hat{\mathbf{e}}_n \rightarrow \mathbf{a}_n \quad h_i \rightarrow \sqrt{g_{ii}} \quad T \rightarrow Y$$

the result agrees with M&S p 3 (1.11). A more compact summary is this:

$$\begin{aligned}
 \star \mathbf{B} &= [(1/h_n) \partial_n T - (h_n/\sqrt{g}) \varepsilon^{nab} \partial_a \Gamma_b] \hat{\mathbf{e}}_n \\
 T &= (1/\sqrt{g}) \partial_i (\sqrt{g} \mathcal{B}^i/h_i) \\
 \Gamma_b &= (1/\sqrt{g}) h_b^2 \varepsilon^{bcd} (\partial_c [h_d \mathcal{B}^d])
 \end{aligned}$$

(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,

$$\begin{aligned}
 h_i &= 1 \quad g = 1 \quad \mathbf{u} = \mathbf{F}(\mathbf{x}) = \mathbf{x} \\
 (\hat{\mathbf{e}}_n)^i &= S_n^i = \delta_n^i \quad \Rightarrow \quad \hat{\mathbf{e}}_n = \hat{\mathbf{n}}, \\
 \mathbf{B} &= \mathcal{B}^n \hat{\mathbf{e}}_n = B^n \hat{\mathbf{n}} \quad \Rightarrow \quad \mathcal{B}^n = B^n
 \end{aligned}$$

so the above 3-line equation block becomes

$$\begin{aligned}
 \star \mathbf{B} &= [\partial_n T - \varepsilon^{nab} \partial_a \Gamma_b] \hat{\mathbf{n}} \\
 T &= \partial_i (B^i) \\
 \Gamma_b &= \varepsilon^{bcd} (\partial_c B^d)
 \end{aligned}$$

or

$$\begin{aligned}
\star \mathbf{B} &= [\partial_{\mathbf{n}} \{ \partial_{\mathbf{i}} B^{\mathbf{i}} \} - \varepsilon^{\mathbf{nab}} \partial_{\mathbf{a}} \{ \varepsilon^{\mathbf{bcd}} (\partial_{\mathbf{c}} B^{\mathbf{d}}) \}] \hat{\mathbf{n}} & (*) \\
&= [\partial_{\mathbf{n}} \{ \text{div } \mathbf{B} \} - \varepsilon^{\mathbf{nab}} \partial_{\mathbf{a}} \{ (\text{curl } \mathbf{B})_{\mathbf{b}} \}] \hat{\mathbf{n}} \\
&= [\partial_{\mathbf{n}} \{ \text{div } \mathbf{B} \} - [\text{curl } (\text{curl } \mathbf{B})]_{\mathbf{n}}] \hat{\mathbf{n}} \\
&= \text{grad} \{ \text{div } \mathbf{B} \} - [\text{curl } (\text{curl } \mathbf{B})] & \text{QED}
\end{aligned}$$

Second, one can verify the claim made earlier that in Cartesian coordinates

$$[\star \mathbf{B}]_{\mathbf{n}} = \nabla^2 B_{\mathbf{n}} .$$

To show this, it is necessary to show that (left side from (*) above)

$$\begin{aligned}
\partial_{\mathbf{n}} \partial_{\mathbf{i}} B_{\mathbf{i}} - \varepsilon^{\mathbf{nab}} \partial_{\mathbf{a}} \varepsilon^{\mathbf{bcd}} (\partial_{\mathbf{c}} B_{\mathbf{d}}) &= \partial_{\mathbf{i}}^2 B_{\mathbf{n}} & ? \\
\varepsilon^{\mathbf{nab}} \varepsilon^{\mathbf{bcd}} \partial_{\mathbf{a}} (\partial_{\mathbf{c}} B_{\mathbf{d}}) &= \partial_{\mathbf{n}} \partial_{\mathbf{i}} B_{\mathbf{i}} - \partial_{\mathbf{i}}^2 B_{\mathbf{n}} & ? \\
\varepsilon^{\mathbf{bna}} \varepsilon^{\mathbf{bcd}} \partial_{\mathbf{a}} (\partial_{\mathbf{c}} B_{\mathbf{d}}) &= \partial_{\mathbf{n}} \partial_{\mathbf{i}} B_{\mathbf{i}} - \partial_{\mathbf{i}}^2 B_{\mathbf{n}} & ?
\end{aligned}$$

But since index b now appears only in the ε 's, use (up and down indices same in Cartesian x-space)

$$\varepsilon_{\mathbf{bna}} \varepsilon_{\mathbf{bcd}} = \delta_{\mathbf{nc}} \delta_{\mathbf{ad}} - \delta_{\mathbf{nd}} \delta_{\mathbf{ac}} \quad // \text{ Appendix D (j) item 4}$$

so

$$\begin{aligned}
(\delta_{\mathbf{nc}} \delta_{\mathbf{ad}} - \delta_{\mathbf{nd}} \delta_{\mathbf{ac}}) \partial_{\mathbf{a}} (\partial_{\mathbf{c}} B_{\mathbf{d}}) &= \partial_{\mathbf{n}} \partial_{\mathbf{i}} B_{\mathbf{i}} - \partial_{\mathbf{i}}^2 B_{\mathbf{n}} & ? \\
\delta_{\mathbf{nc}} \delta_{\mathbf{ad}} \partial_{\mathbf{a}} (\partial_{\mathbf{c}} B_{\mathbf{d}}) - \delta_{\mathbf{nd}} \delta_{\mathbf{ac}} \partial_{\mathbf{a}} (\partial_{\mathbf{c}} B_{\mathbf{d}}) &= \partial_{\mathbf{n}} \partial_{\mathbf{i}} B_{\mathbf{i}} - \partial_{\mathbf{i}}^2 B_{\mathbf{n}} & ? \\
\partial_{\mathbf{a}} (\partial_{\mathbf{n}} B_{\mathbf{a}}) - \partial_{\mathbf{a}} (\partial_{\mathbf{a}} B_{\mathbf{n}}) &= \partial_{\mathbf{n}} \partial_{\mathbf{i}} B_{\mathbf{i}} - \partial_{\mathbf{i}}^2 B_{\mathbf{n}} & ? \\
\partial_{\mathbf{n}} (\partial_{\mathbf{a}} B_{\mathbf{a}}) - \partial_{\mathbf{a}}^2 B_{\mathbf{n}} &= \partial_{\mathbf{n}} (\partial_{\mathbf{i}} B_{\mathbf{i}}) - \partial_{\mathbf{i}}^2 B_{\mathbf{n}} & ?
\end{aligned}$$

Since this last equation is true on inspection, QED.

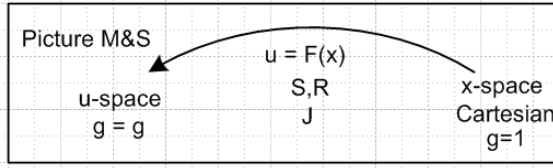
Thus it has been shown that, in Cartesian coordinates,

$$[\star \mathbf{B}]_{\mathbf{n}} = \nabla^2 (B_{\mathbf{n}}) = [\text{grad}(\text{div } \mathbf{B}) - \text{curl } (\text{curl } \mathbf{B})]_{\mathbf{n}}$$

It is this second (and rather complicated) form $[\text{grad}(\text{div } \mathbf{B}) - \text{curl } (\text{curl } \mathbf{B})]_{\mathbf{n}}$ which allowed us to obtain an expression for $[\star \mathbf{B}]$ in general curvilinear coordinates using methods described in this document.

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:

$$\begin{array}{ll} [\text{grad } f](\mathbf{x}) = (\partial_i f) \mathbf{e}^i & \\ [\text{grad } f](\mathbf{x}) = (\partial^i f) \mathbf{e}_i & = g^{ij} (\partial_j f) \mathbf{e}_i \\ [\text{grad } f](\mathbf{x}) = h_i (\partial^i f) \hat{\mathbf{e}}_i & = h_i g^{ij} (\partial_j f) \hat{\mathbf{e}}_i \\ [\text{grad } f](\mathbf{x}) = (\partial_i f) \hat{\mathbf{n}} & // \text{Cartesian} \\ [\text{grad } f]_i(\mathbf{u}) = R_i^j [\text{grad } f]_j(\mathbf{x}) & // \text{transformation (vector)} & f(\mathbf{u}) = f(\mathbf{x}) = f(\mathbf{x}(\mathbf{u})) \end{array}$$

gradient orthogonal:

$$\begin{array}{ll} [\text{grad } f](\mathbf{x}) = (\partial_i f) \mathbf{e}^i & \\ [\text{grad } f](\mathbf{x}) = (\partial^i f) \mathbf{e}_i & = (1/h_i^2) (\partial_i f) \mathbf{e}_i \\ [\text{grad } f](\mathbf{x}) = h_i (\partial^i f) \hat{\mathbf{e}}_i & = (1/h_i) (\partial_i f) \hat{\mathbf{e}}_i \quad // \text{M\&S 1.05} \end{array}$$

gradient dotted with a vector:

$$\begin{aligned}
 [\text{grad } f](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) &= (\partial^n f) B_n & \mathbf{B} &= B_n \mathbf{e}_n \\
 [\text{grad } f](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) &= (\partial_n f) B^n & \mathbf{B} &= B^n \mathbf{e}_n \\
 [\text{grad } f](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) &= (\partial_n f) \mathcal{B}^n / h_n & \mathbf{B} &= \mathcal{B}^n \hat{\mathbf{e}}_n \\
 [\text{grad } f](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) &= (\partial_n f) B^n & // \text{ Cartesian} & \mathbf{B} = B^n \hat{\mathbf{n}} \\
 [\text{grad } f](\mathbf{x}) \bullet \mathbf{B}(\mathbf{x}) &= [\text{grad } f](\mathbf{u}) \bullet \mathbf{B}(\mathbf{u}) & // \text{ transformation (scalar)} &
 \end{aligned}$$

(c) Laplacian

Laplacian general:

$$\begin{aligned}
 [\text{lap } f](\mathbf{x}) &= [1/\sqrt{g}] \partial_m [\sqrt{g} g^{nm} (\partial_n f)] \\
 [\text{lap } f](\mathbf{x}) &= \partial_n^2 f(\mathbf{x}) & // \text{ Cartesian} & f(\mathbf{u}) = f(\mathbf{x}) = f(\mathbf{x}(\mathbf{u})) \\
 [\text{lap } f](\mathbf{x}) &= [\text{lap } f](\mathbf{u}) & // \text{ transformation (scalar)} &
 \end{aligned}$$

Laplacian orthogonal:

$$[\text{lap } f](\mathbf{x}) = [1/(\Pi_i h_i)] \partial_m [(\Pi_i h_i) (1/h_m^2) (\partial_m f)] \quad // \text{ orthogonal} \quad // \text{ M\&S 1.09}$$

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

$$\begin{aligned}
 [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} [(1/\sqrt{g}) \partial_a B_b] \mathbf{e}_n & \mathbf{B} &= B_n \mathbf{e}_n \\
 [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} [(1/\sqrt{g}) \partial_a (g_{bc} B^c)] \mathbf{e}_n & \mathbf{B} &= B^n \mathbf{e}_n \\
 [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} [(1/\sqrt{g}) h_n \partial_a (g_{bc} \mathcal{B}^c / h_c)] \hat{\mathbf{e}}_n & \mathbf{B} &= \mathcal{B}^n \hat{\mathbf{e}}_n \\
 [\text{curl } \mathbf{B}](\mathbf{x}) &= \varepsilon^{nab} \partial_a B_b(\mathbf{x}) \hat{\mathbf{n}} & // \text{ Cartesian} & \mathbf{B} = B^n \hat{\mathbf{n}}
 \end{aligned}$$

$$[\text{curl } \mathbf{B}](\mathbf{x}) = (1/\sqrt{g}) \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ B_1 & B_2 & B_3 \end{vmatrix} \quad \mathbf{B} = B_n \mathbf{e}_n$$

$$[\text{curl } \mathbf{B}](\mathbf{x}) = (1/\sqrt{g}) \begin{vmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ g_{1c} B^c & g_{2c} B^c & g_{3c} B^c \end{vmatrix} \quad \mathbf{B} = B^n \mathbf{e}_n$$

$$[\text{curl } \mathbf{B}](\mathbf{x}) = (1/\sqrt{g}) \begin{vmatrix} h_1 \hat{\mathbf{e}}_1 & h_2 \hat{\mathbf{e}}_2 & h_3 \hat{\mathbf{e}}_3 \\ \partial_1 & \partial_2 & \partial_3 \\ (g_{1c}/h_c) \mathcal{B}^c & (g_{2c}/h_c) \mathcal{B}^c & (g_{3c}/h_c) \mathcal{B}^c \end{vmatrix} \quad \mathbf{B} = \mathcal{B}^n \hat{\mathbf{e}}_n \quad // \text{M\&S 1.07}$$

$$[\text{curl } \mathbf{B}](\mathbf{x}) = \begin{vmatrix} \hat{1} & \hat{2} & \hat{3} \\ \partial_1 & \partial_2 & \partial_3 \\ B_1 & B_2 & B_3 \end{vmatrix} \quad // \text{here } \partial_n = \partial/\partial x^n \text{ and } B_i = B_i(\mathbf{x}) \quad \mathbf{B} = B^n \hat{\mathbf{n}}$$

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

$$[\star \mathbf{B}](\mathbf{x}) = \mathbf{G} - \mathbf{V}$$

$$\begin{aligned} \mathbf{G} &= \partial^n \{ (1/\sqrt{g}) \partial_i (\sqrt{g} B^i) \} \mathbf{e}_n & \mathbf{B} &= B^n \mathbf{e}_n \\ \mathbf{V} &= [(1/\sqrt{g}) \varepsilon^{ncd} \varepsilon^{eab} \partial_c \{ (1/\sqrt{g}) g_{de} (\partial_a [g_{bf} B^f]) \}] \mathbf{e}_n \end{aligned}$$

$$\begin{aligned} \mathbf{G} &= h_n \partial^n \{ (1/\sqrt{g}) \partial_i (\sqrt{g} \mathcal{B}^i / h_i) \} \hat{\mathbf{e}}_n & \mathbf{B} &= \mathcal{B}^n \hat{\mathbf{e}}_n \\ \mathbf{V} &= h_n [(1/\sqrt{g}) \varepsilon^{ncd} \varepsilon^{eab} \partial_c \{ (1/\sqrt{g}) g_{de} (\partial_a [g_{bf} \mathcal{B}^f / h_f]) \}] \hat{\mathbf{e}}_n \end{aligned}$$

$$\star \mathbf{B} = [\nabla^2 B^n] \hat{\mathbf{n}} \quad // \text{Cartesian, } \hat{\mathbf{n}} = \mathbf{u}_n \quad \mathbf{B} = B^n \hat{\mathbf{n}}$$

vector Laplacian orthogonal:

$$\begin{aligned}
[\star \mathbf{B}](\mathbf{x}) &= [(1/h_1) \partial_1 T + (h_1/\sqrt{g}) (\partial_3 \Gamma_2 - \partial_2 \Gamma_3)] \hat{\mathbf{e}}_1 \\
&+ [(1/h_2) \partial_2 T + (h_2/\sqrt{g}) (\partial_1 \Gamma_3 - \partial_3 \Gamma_1)] \hat{\mathbf{e}}_2 \\
&+ [(1/h_3) \partial_3 T + (h_3/\sqrt{g}) (\partial_2 \Gamma_1 - \partial_1 \Gamma_2)] \hat{\mathbf{e}}_3
\end{aligned}
\quad // \text{ M\&S 1.11}$$

$$\begin{aligned}
T &= (1/\sqrt{g}) \partial_i (\sqrt{g} \mathcal{B}^i / h_i) \\
\Gamma_1 &= (1/\sqrt{g}) h_1^2 (\partial_2 [h_3 \mathcal{B}^3] - \partial_3 [h_2 \mathcal{B}^2]) \\
\Gamma_2 &= (1/\sqrt{g}) h_2^2 (\partial_3 [h_1 \mathcal{B}^1] - \partial_1 [h_3 \mathcal{B}^3]) \\
\Gamma_3 &= (1/\sqrt{g}) h_3^2 (\partial_1 [h_2 \mathcal{B}^2] - \partial_2 [h_1 \mathcal{B}^1])
\end{aligned}$$

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

$$\begin{array}{llll}
\text{general:} & g \equiv \det(g_{ab}) & h_n^2 \equiv g_{nn} & \partial^i = g^{ij} \partial_j \\
\text{orthogonal:} & \sqrt{g} = (\Pi_i h_i) = h_1 h_2 \dots h_N & g_{nm} = h_n^2 \delta_{n,m} & \mathcal{B}^n = h_n B^n \quad \mathbf{e}_n = h_n \hat{\mathbf{e}}_n \\
& & & g^{nm} = h_n^{-2} \delta_{n,m}
\end{array}$$

$$\begin{aligned}
\mathbf{e}_1 = r(-\sin\theta, \cos\theta) &= \mathbf{e}_\theta = r \hat{\mathbf{e}}_\theta & // = r \hat{\boldsymbol{\theta}} \\
\mathbf{e}_2 = (\cos\theta, \sin\theta) &= \mathbf{e}_r = \hat{\mathbf{e}}_r & // = \hat{\mathbf{r}}
\end{aligned}$$

$$g_{ij} = \begin{pmatrix} \theta & r \\ r^2 & 0 \\ 0 & 1 \end{pmatrix} \begin{matrix} \theta \\ r \end{matrix} \quad R = \begin{pmatrix} x & y \\ -\sin\theta/r & \cos\theta/r \\ \cos(\theta) & \sin\theta \end{pmatrix} \begin{matrix} \theta \\ r \end{matrix} \quad u_1 = \theta \quad u_2 = r$$

$$h_1 = h_\theta = \sqrt{g_{\theta\theta}} = r \quad h_2 = h_r = \sqrt{g_{rr}} = 1$$

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.

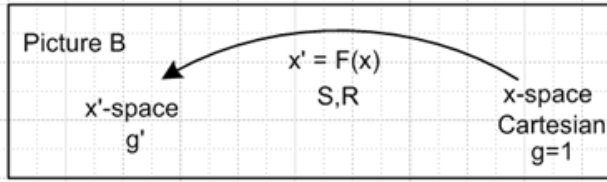
15. Covariant derivation of all curvilinear differential operator expressions

(a) Review of Sections 9 through 13

Let us review the discussion of Sections 9 through 13 concerning the expression of differential operators

div, grad, lap, curl and $\star$

in curvilinear coordinates. The underlying framework was provided by Picture B,



where the coordinates x'^n of x' -space were the curvilinear coordinates of interest, while the coordinates of x -space were the Cartesian coordinates. The transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ provided the connection between the curvilinear coordinates and the Cartesian ones.

In **Section 9** the object $\text{div } \mathbf{B}$ was treated as the $\mathbf{B}$ flux emitting from an N-piped in x -space divided by the volume of that N-piped in the limit the N-piped shrank to a point. Using Appendix B formulas for the N-piped area, and using the key result (Section 7 (s)),

$$\mathbf{B} = B'^n \mathbf{e}_n,$$

where the $\mathbf{e}_n$ are tangent base vectors in x -space, while the B'^n are components of $\mathbf{B}$ in x' -space, we obtained a way to write the Cartesian-space $\text{div } \mathbf{B}$ in terms of curvilinear $\mathbf{x}'$ coordinates and x' -space objects, namely, $B'^n(\mathbf{x}')$ and the gradient operator $\partial'_i = \partial/\partial x'^i$.

In **Section 10** the object $\text{grad } f$ was expanded using another Section 7 (s) expansion which is that shown above with the index tilt reversed,

$$\text{grad } f = [\text{grad } f]'_n \mathbf{e}^n = (\partial'_n f') \mathbf{e}^n$$

where $\mathbf{e}^n$ are the reciprocal base vectors. This result was expressed in various ways, and $\text{grad } f \bullet \mathbf{B}$ was also treated such that $\text{grad } f \bullet \mathbf{B} = (\partial'_n f') \mathbf{e}^n \bullet B'^m \mathbf{e}_m = (\partial'_n f') B'^n$. [Recall that $f'(\mathbf{x}') = f(\mathbf{x})$.]

In **Section 11** the Laplacian was written as $\text{lap } f = \text{div } (\text{grad } f)$ and then the results of Sections 9 and 10 for div and grad were used to express $\text{lap } f$ in terms of curvilinear coordinates. In traditional notation, one writes $\nabla^2 f = \nabla \bullet [\nabla f] = \text{div } (\text{grad } f)$.

In **Section 12** the object $[\text{curl } \mathbf{B}]^n$ was treated as the circulation line integral around near face n of the same N-piped used in Section 9 for divergence, divided by the area of that face, again in the limit the N-

piped shrank to a point. With the same expansion for $\mathbf{B}$ shown above, this led to an expression of curl $\mathbf{B}$ once again in terms of the curvilinear x' coordinates and the components B'^n .

In **Section 13** the vector Laplacian was treated using the definition

$$\star \mathbf{B} \equiv \text{grad}(\text{div } \mathbf{B}) - \text{curl}(\text{curl } \mathbf{B})$$

and then the results of Sections 9, 10 and 12 on div, grad and curl were recruited to produce an expression for $\star \mathbf{B}$ in terms of $B'^n(x')$ and x' -space coordinates.

(b) The Covariant Method

We shall now repeat all of the above work using "the covariant method" which, as will be shown, gets most results extremely quickly, but at a cost of requiring knowledge of tensor densities and covariant differentiation, as described in Appendices D, E and F.

Imagine that one has a Cartesian x -space expression for some tensor object of interest Q ,

$$Q\text{----} = [\text{Cartesian form with various up and down indices and various derivatives}]$$

where $Q\text{----}$ means that Q has some number of up and down indices. The following four steps should then be carried out:

1. First, arrange any summed index pairs within [...] so they are "tilted". Then when [...] is "tensorized" those tilted pairs will be true contractions.

Note that doing this does not affect the Cartesian value of this object [...] since up and down indices are the same in Cartesian space.

2. Second, replace all derivatives with covariant derivatives. Thus is done by first writing all derivatives in comma notation, and then by replacing those commas with semicolons. For example:

$$\partial_\alpha B_a = B_{a,\alpha} \rightarrow B_{a;\alpha} \quad // \quad B_{a;\alpha} = \partial_\alpha B_a - \Gamma_{a\alpha}^n B_n$$

Objects like $\partial_\alpha B_a$ which are not tensors become objects like $B_{a;\alpha}$ which *are* tensors (see Appendix F (g) through (i)).

Note again that doing this does not affect the Cartesian value of this object [...], because in Cartesian x -space $\Gamma_{a\alpha}^n = 0$. This is so because, as implied by Appendix F (a) (where x -space is called ξ -space),

$$\Gamma_{ab}^c \equiv \mathbf{u}^c \bullet (\partial_a \mathbf{u}_b)$$

and the basis vectors $\mathbf{u}_n$ are constants in Cartesian space so $(\partial_a \mathbf{u}_b) = 0$ and then $\Gamma_{ab}^c = 0$.

3. Third, insert appropriate powers of $g^{1/2}$ as needed to achieve appropriate tensor density weights so that all terms in [...] have the same weight and this weight equals the weight of $Q\text{----}$. Tensor densities are discussed in Appendix D, and this weight adjustment idea appears in Appendix D (b) item 3 Example 2.

Note again that doing this does not affect the Cartesian value of this object [...] because $g = 1$ in Cartesian space ($g = \det(g_{ij})$).

4. At this point one has

$$Q\text{----} = \{ \text{tensorized form of the Cartesian expression of interest} \}$$

Since this is then a "true tensor equation" (it could be a true tensor density equation), according to the rule of covariance described in Section 7 (u) the above equation can be converted to x' -space by simply priming all objects within $\{ \}$. One then has

$$Q'\text{----} = \{ \text{tensorized form of the Cartesian expression of interest} \}'$$

where the notation $\{...\}'$ means that everything inside $\{...\}$ is primed.

Perhaps at this point one does some simplifications of the resulting $\{...\}'$ object. The result then *is* the expression of $Q\text{----}$ in general curvilinear coordinates if the underlying transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ describes the transformation from Cartesian to curvilinear coordinates. (For a vector, recall that $\mathbf{Q} = Q^a \mathbf{e}_a$ so the "curvilinear coordinates" expression of $\mathbf{Q}$ involves the the x' -space components $Q'^a(x')$ but the tangent base vectors $\mathbf{e}_a$ are in x -space. A similar statement can be made for Q being any tensor, as outlined in Appendix (a) and (b)).

This then is "the covariant method". The method of course applies to tensor analysis applications other than "curvilinear coordinates".

Uniqueness. One might obtain two tensorizations of a Cartesian expression that look different. Both these tensorizations must lead to the same $Q'\text{----}$ in x' -space. For example, suppose one finds two tensorizations of a vector Q^a , call them $\{...\}_1^a$ and $\{...\}_2^a$ which, when evaluated in Cartesian x -space, are equal

$$\{...\}_1^a = \{...\}_2^a \quad \text{when both sides are evaluated in Cartesian } x\text{-space}$$

One can then transform both objects to x' -space in the usual manner,

$$\{...\}_1'^a = R^a_b \{...\}_1^b \quad \text{and} \quad \{...\}_2'^a = R^a_b \{...\}_2^b$$

Therefore, since $\{...\}_1^a = \{...\}_2^a$ in x -space, one must have $\{...\}_1'^a = \{...\}_2'^a$ in x' -space.

Example: In Section (g) below we will establish the following tensorization of $(\star B)^n$ (the vector Laplacian),

$$(\star B)^n = \{.....\}_1^n = (B^j_{;j})^{;n} - g^{-1/2} \epsilon^{nab} (g^{-1/2} g_{bc} \epsilon^{cde} B_{e;d})_{;a} \quad .$$

Since in Cartesian x -space one can write $(\star B)^n = \nabla^2(B^n) = \partial_j \partial^j B^n = B^{n,j}_{;j}$, another tensorization of $(\star B)^n$ is given by

$$(\star B)^n = \{\dots\}_2^n = B^{n;j}_{;j} \quad .$$

Therefore, the following must be true in x' -space

$$B^{n;j}_{;j} = (B^{j;j})^{;n} - g'^{-1/2} \varepsilon^{nab} (g'^{-1/2} g'_{bc} \varepsilon^{cde} B'_{e;d})_{;a}$$

Verifying this equality is a non-trivial exercise. Both forms are used in Appendix I to compute $\star B$ in spherical and cylindrical coordinates.

In the following sections, then, we shall in short order derive the curvilinear expressions for the basic differential operators using the covariant method outlined above. This method is used as well (along with parallel brute force methods) to obtain expressions for objects $\nabla \mathbf{v}$, $\text{div} \mathbf{T}$ and ∇T in Appendices G, H, K.

(c) divergence (Section 9)

In Cartesian coordinates, $\text{div} B = \partial_i B^i = B^i_{;i}$. The true tensorial tensor which matches this is $\text{div} B \equiv B^i_{;i}$ since $\Gamma = 0$ for Cartesian coordinates. Since $B^i_{;i}$ is a scalar, it is the same in x -space as in x' -space. According to the first $J = 1$ example of Appendix F (i), $B^a_{;\alpha} = \partial_\alpha B^a + \Gamma^a_{\alpha n} B^n$, so

$$[\text{div} B] = [\text{div} B]' = B^i_{;i} = \partial'_i B'^i + \Gamma'^i_{in} B'^n$$

where B'^n are the contravariant components of vector $\mathbf{B}$ in x' -space. As shown in Appendix F (d), $\Gamma'^i_{in} = (1/\sqrt{g'}) \partial'_n (\sqrt{g'})$ and therefore

$$[\text{div} B] = \partial'_n B'^n + (1/\sqrt{g'}) \partial'_n (\sqrt{g'}) B'^n = (1/\sqrt{g'}) \partial'_n (\sqrt{g'} B'^n).$$

This matches the result obtained in Section 9 by geometric methods.

(d) gradient and gradient dot vector (Section 10)

If f is a scalar, then $\mathbf{G} = \text{grad } f$ transforms as an ordinary vector, so it is *already* a tensor. According to Section 7 (s) or Appendix E (b), the expansion of such a vector may be written

$$\mathbf{G} = \text{grad } f = G'_i \mathbf{e}^i = (\partial'_i f') \mathbf{e}^i = (\partial'^i f') \mathbf{e}_i$$

Also, $\text{grad } f \bullet \mathbf{B}$ is a scalar, again already a tensor. Thus

$$[\text{grad } f]' \bullet \mathbf{B}' = \text{grad } f \bullet \mathbf{B} = (\partial'_i f') B'^i$$

Both these results match those of Section 10. Since f is a scalar, $f'(\mathbf{x}') = f(\mathbf{x})$.

(e) Laplacian (Section 11)

In Cartesian x -space the Laplacian of a scalar function f can be written as

$$[\text{lap } f]_{\text{cart}} = \partial^i \partial_i f = f'^i{}_{,i}$$

The tensorized version of the Laplacian is then taken to be the following, using the $\rightarrow$ rule,

$$\text{lap } f = f'^i{}_{,i}$$

Since $f'^i{}_{,i}$ is the contraction of a rank-2 tensor, it is a scalar, and therefore

$$\text{lap } f = f'^i{}_{,i} = f'^i{}_{;i} = \partial'_i f'^i + \Gamma'^i{}_{in} f'^n$$

where the example $B'^a{}_{;\alpha} = \partial'_\alpha B'^a + \Gamma'^a{}_{\alpha n} B'^n$ from Appendix F (i) is used. Another example shows that for scalar B , $B'^\alpha = \partial'^\alpha B$. Applying this to f'^i and f'^n , the above becomes

$$\text{lap } f = \partial'_i \partial'^i f' + \Gamma'^i{}_{in} \partial'^n f'$$

Again, as shown in Appendix F (d), $\Gamma'^i{}_{in} = (1/\sqrt{g'}) \partial'_n (\sqrt{g'})$ and therefore

$$\text{lap } f = \partial'_n \partial'^n f' + (1/\sqrt{g'}) \partial'_n (\sqrt{g'}) (\partial'^n f') = [1/\sqrt{g'}] \partial'_n [\sqrt{g'} (\partial'^n f')]$$

which matches the result of Section 11.

(f) curl (Section 12)

As shown in Appendix D (h), $\mathbf{C} \equiv \text{curl } \mathbf{B}$ is a vector density of weight -1 because it contains the ε tensor which has this weight. According to Appendix E (b), the expansion of such a vector density is given by $\mathbf{C} = J^{-1} C'^i \mathbf{e}_i$. Therefore,

$$\mathbf{C} = \text{curl } \mathbf{B} = J^{-1} C'^i \mathbf{e}_i = J^{-1} \varepsilon'^{ijk} B'_{k,j} \mathbf{e}_i = (1/\sqrt{g'}) \varepsilon'^{ijk} B'_{k,j} \mathbf{e}_i$$

so

$$[\text{curl } \mathbf{B}^{(\mathbf{e})}]^i = (1/\sqrt{g'}) \varepsilon'^{ijk} B'_{k,j}$$

This matches the result of Section 12.

(g) vector Laplacian (Section 13)

This differential operator is going to take a bit more work. In Section 13 the vector Laplacian is written $\star$ (instead of ∇^2) and in Cartesian coordinates is defined by

$$\star \mathbf{B} \equiv \text{grad}(\text{div } \mathbf{B}) - \text{curl}(\text{curl } \mathbf{B})$$

Recall that in Cartesian space (as shown in Section 13 (c))

$$[\text{grad}(\text{div } \mathbf{B}) - \text{curl}(\text{curl } \mathbf{B})]^n = \nabla^2 (B^n) = (\star \mathbf{B})^n \quad // \text{ Cartesian x-space}$$

so we are allowed to use the LHS here to find the curvilinear expression of $\star B$. In components,

$$\begin{aligned} (\star B)^n &= \partial^n (\partial_j B^j) - \varepsilon^{nab} \partial_a [\text{curl } B]_b \\ &= \partial^n (\partial_j B^j) - \varepsilon^{nab} \partial_a (\varepsilon_{bde} \partial^d B^e) \end{aligned}$$

In Cartesian space up and down index position does not matter and we are just jockeying the indices in search of a tensorized form with contracted indices where possible. Writing the above in comma notation gives

$$(\star B)^n = (B^j_{;j})^{;n} - \varepsilon^{nab} (\varepsilon_{bde} B^{e;d})_{;a}$$

and then we are free to replace commas with semicolons since $\Gamma = 0$ in Cartesian coordinates (see Appendix F sections (d) and (i)) :

$$(\star B)^n = (B^j_{;j})^{;n} - \varepsilon^{nab} (\varepsilon_{bde} B^{e;d})_{;a}$$

There are several "technical difficulties" visible here. The first term $(B^j_{;j})^{;n}$ is a true vector since it is the contraction of a rank-3 tensor on two tilted indices. In the second term, in order to neutralize the weight of each ε tensor density (each has weight -1), we shall now add a benign factor of $g^{-1/2}$, using the idea of Appendix D (b) item 3 Example 2 (that is, $g^{-1/2}$ is a scalar density of weight +1). (The factor is benign since $g = 1$ in Cartesian coordinates.)

$$(\star B)^n = (B^j_{;j})^{;n} - g^{-1/2} \varepsilon^{nab} (g^{-1/2} \varepsilon_{bde} B^{e;d})_{;a}$$

Now both objects $T_b \equiv (g^{-1/2} \varepsilon_{bde} B^{e;d})$ and $V^n \equiv g^{-1/2} \varepsilon^{nab} T_{b;a}$ are normal vectors, so the above equation says that a normal vector is the difference between two other normal vectors.

Anticipating a few steps ahead when our ship lands in x' -space, we know that $\varepsilon^{bde} B_{e;d} = \varepsilon^{bde} B_{e,d}$ due to the $e \leftrightarrow d$ symmetry of the Γ term in $B_{e;d}$ (see Appendix F (i) Examples with $J=1$), and this *motivates us* to perform a "tilt change" in this product. Such a tilt-change also gets the $;d$ index "down" which is a nicer place for it to be since it will soon become ∂_d . The tilt operation is done as follows ,

$$\varepsilon_{bde} B^{e;d} = \varepsilon_b^d B^e_{;d} = \varepsilon_b^{de} B_{e;d} = g_{bc} \varepsilon^{cde} B_{e;d} \quad .$$

Recall that tilt changes like this are only allowed with true tensor indices (Section 7 (k)) and that is why the e tilt reversal can only be done after the semicolon in $B_{e;d}$ is installed. The result is then

$$(\star B)^n = (B^j_{;j})^{;n} - g^{-1/2} \varepsilon^{nab} (g^{-1/2} g_{bc} \varepsilon^{cde} B_{e;d})_{;a} \quad .$$

Although ε_b^{de} is a fine tensor, we really prefer ε^{cde} since *this* is the permutation tensor, and that is why the g_{bc} factor is added.

At this point every index is a tensor index, so we have a "true tensor equation" in the sense of Section 7 (u). The above equation agrees exactly with the Cartesian expression of $(\star B)^n$ since $g = 1$, $g_{bc} = \delta_{b,c}$, and the semicolons are commas since $\Gamma = 0$. The point is that we have found a way to write the Cartesian expression of $(\star B)^n$ such that both terms are true tensors. The equation is therefore now "covariant" and the equation in x' -space can be obtained simply by priming all objects:

$$(\star B)^n = (B'^j{}_{;j})'^n - g'^{-1/2} \epsilon'^{nab} (g'^{-1/2} g'_{bc} \epsilon'^{cde} B'_{e;d})_{;a}$$

The semicolons were installed to obtain a true tensor equation, but now that we have successfully arrived in x' -space, we want to remove as many semicolons as possible since they imply "extra terms". Consider for example, this grouping which appears in the above,

$$\epsilon'^{nab} (g'^{-1/2} g'_{bc} \epsilon'^{cde} B'_{e;d})_{;a} = \epsilon'^{nab} T'_{b;a} \quad T'_b \equiv (g'^{-1/2} g'_{bc} \epsilon'^{cde} B'_{e;d})$$

Since $T_{b;a} = T_{b,a} - \Gamma^n_{ba} T_n$ (Appendix F (i)) , the above expression is equal to $\epsilon'^{nab} T_{b,a}$ because ϵ'^{nab} is antisymmetric on a,b while Γ^n_{ba} is symmetric. Thus,

$$(\star B)^n = (B'^j{}_{;j})'^n - g'^{-1/2} \epsilon'^{nab} (g'^{-1/2} g'_{bc} \epsilon'^{cde} B'_{e;d})_{;a} \quad \uparrow$$

Now we activate the fact that $\epsilon'^{cde} B'_{e;d} = \epsilon'^{cde} B'_{e,d}$ for the exact same symmetry reason to get

$$(\star B)^n = (B'^j{}_{;j})'^n - g'^{-1/2} \epsilon'^{nab} (g'^{-1/2} g'_{bc} \epsilon'^{cde} B'_{e,d})_{;a} \quad \uparrow$$

Meanwhile, an Appendix F (i) example that says $D'^n = D'^n$ so that $(B'^j{}_{;j})'^n = (B'^j{}_{;j})'^n$, since the object $B'^j{}_{;j}$ is a scalar like D . Therefore,

$$(\star B)^n = (B'^j{}_{;j})'^n - g'^{-1/2} \epsilon'^{nab} (g'^{-1/2} g'_{bc} \epsilon'^{cde} B'_{e,d})_{;a} \quad \uparrow$$

Since ϵ'^{cde} is a constant (the permutation tensor, see Appendix D (d)) , we *now* pull it through the ∂'_a derivative implied by $_{;a}$ and then showing all derivatives the above becomes

$$(\star B)^n = \partial'^n (B'^j{}_{;j}) - g'^{-1/2} \epsilon'^{nab} \epsilon'^{cde} \partial'_a (g'^{-1/2} g'_{bc} \partial'_d B'_e) .$$

The object $(B'^j{}_{;j})$ we know from section (c) above

$$[\text{div} B] = B'^j{}_{;j} = (1/\sqrt{g'}) \partial'_j (\sqrt{g'} B'^j)$$

so that

$$(\star B)^n = \partial'^n \{ (1/\sqrt{g'}) \partial'_j (\sqrt{g'} B'^j) \} - g'^{-1/2} \epsilon'^{nab} \epsilon'^{cde} \partial'_a (g'^{-1/2} g'_{bc} \partial'_d B'_e)$$

In order to compare this result to that of Section 13, it is necessary to process the indices in the second term as follows,

$$\begin{aligned} & g'^{-1/2} \varepsilon^{nab} \varepsilon'^{cde} \partial'_a (g'^{-1/2} g'_{bc} \partial'_d B'_e) \\ & g'^{-1/2} \varepsilon^{nAB} \varepsilon'^{CDE} \partial'_A (g'^{-1/2} g'_{BC} \partial'_D B'_E) \end{aligned}$$

Then take $A \rightarrow c, B \rightarrow d, C \rightarrow e, D \rightarrow a, E \rightarrow b$ to get

$$g'^{-1/2} \varepsilon^{ncd} \varepsilon'^{eab} \partial'_c (g'^{-1/2} g'_{de} \partial'_a B'_b) .$$

Then also changing $j \rightarrow i$ in the first term we have,

$$(\star B)^{in} = \partial^{in} \{ (1/\sqrt{g'}) \partial'_i (\sqrt{g'} B'^i) \} - g'^{-1/2} \varepsilon^{ncd} \varepsilon'^{eab} \partial'_c (g'^{-1/2} g'_{de} \partial'_a B'_b) .$$

This may now be compared with the Section 13 (a) result which we quote :

$$\begin{aligned} \star B &= G - V \\ G &= \partial^{in} \{ (1/\sqrt{g'}) \partial'_i (\sqrt{g'} B'^i) \} \mathbf{e}_n & \mathbf{B} &= B'^n \mathbf{e}_n \\ V &= [(1/\sqrt{g'}) \varepsilon^{ncd} \varepsilon'^{eab} \partial'_c \{ (1/\sqrt{g'}) g'_{de} (\partial'_a B'_b) \}] \mathbf{e}_n & \mathbf{B} &= B'_n \mathbf{e}^n \end{aligned}$$

The results agree, and further processing of this result is given in Section 13.

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

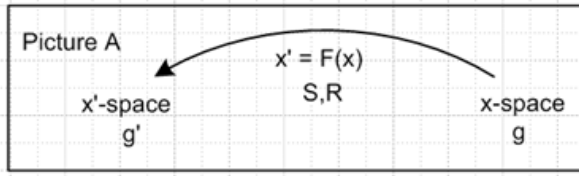
$$\mathbf{E}_n \rightarrow \mathbf{e}^n \quad R_{ij} \rightarrow R^i_j \quad S_{ij} \rightarrow S^i_j \quad \bar{g}'_{nm} \rightarrow g'_{nm} \quad g'_{nm} \rightarrow g'^{nm} .$$

Introduction

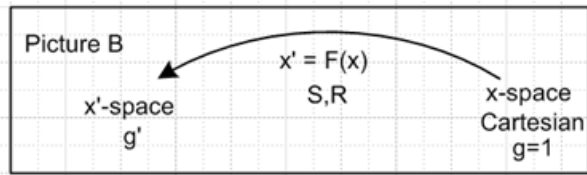
In Section 6 of the main text the reciprocal base vectors are defined as $\mathbf{E}_n \equiv g'_{ni} \mathbf{e}_i$, and the results given in that section,

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

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. One application of the reciprocal base vectors is in the study of curvilinear coordinates where one always takes $g = 1$, and g' is then the curvilinear coordinates metric tensor of interest. Therefore, the reader should think of this Appendix in the context of Picture B



(a) 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

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

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

$$(\mathbf{E}_k)_\alpha \equiv \det(R) \varepsilon_{abc\dots \alpha \dots x} (\mathbf{e}_1)_a (\mathbf{e}_2)_b \dots (\mathbf{e}_N)_x \quad // (\mathbf{e}_k)_\kappa \text{ is missing, } \alpha \text{ in } \kappa \text{ position (the } k^{\text{th}})$$

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:

$$Q_a \equiv \varepsilon_{abc\dots x} B_b C_c D_d \dots X_x$$

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),

$$\begin{aligned} \mathbf{E}_1 &= \det(R) (-1)^{1-1} \mathbf{e}_2 \times \hat{\mathbf{3}} = \det(R) \mathbf{e}_2 \times \hat{\mathbf{3}} & \text{or} & & (\mathbf{E}_1)_k &= \det(R) \varepsilon_{ka} (\mathbf{e}_2)_a \\ \mathbf{E}_2 &= \det(R) (-1)^{2-1} \mathbf{e}_1 \times \hat{\mathbf{3}} = -\det(R) \mathbf{e}_1 \times \hat{\mathbf{3}} & \text{or} & & (\mathbf{E}_2)_k &= -\det(R) \varepsilon_{ka} (\mathbf{e}_1)_a \end{aligned}$$

One can combine these two lines into one as follows (eg, $k = 1$, then $3-1 = 2$, etc)

$$\mathbf{E}_k = \det(R) (-1)^{k-1} \mathbf{e}_{3-k} \times \hat{\mathbf{3}} = \det(R) \mathbf{e}_{3-k} \times \hat{\mathbf{3}} \quad \text{or} \quad (\mathbf{E}_1)_k = \det(R) (-1)^{k-1} \varepsilon_{ka} (\mathbf{e}_{3-k})_a$$

The vector "trick" notation shows that $\mathbf{E}_1 \bullet \mathbf{e}_2 = 0$ and $\mathbf{E}_2 \bullet \mathbf{e}_1 = 0$,

$$\begin{aligned} \mathbf{E}_1 \bullet \mathbf{e}_2 &= \det(R) \mathbf{e}_2 \times \hat{\mathbf{3}} \bullet \mathbf{e}_2 = 0 \\ \mathbf{E}_2 \bullet \mathbf{e}_1 &= -\det(R) \mathbf{e}_1 \times \hat{\mathbf{3}} \bullet \mathbf{e}_1 = 0 \end{aligned}$$

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 ,

$$\mathbf{E}_k \bullet \mathbf{e}_k = (\mathbf{E}_k)_\alpha (\mathbf{e}_k)_\alpha = \{ \det(R) (-1)^{k-1} \varepsilon_{\alpha abc \dots x} (\mathbf{e}_1)_a (\mathbf{e}_2)_b \dots (\mathbf{e}_N)_x \} (\mathbf{e}_k)_\alpha$$

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

$$\mathbf{Q} = \Pi_{i=1}^{N-1} \mathbf{A}^{(i)} = \Pi_{\mathbf{i}} \mathbf{A}^{(\mathbf{i})}$$

where in the second form it is understood that $\mathbf{i}$ takes on all values $\mathbf{i} = 1$ to $N-1$. The superscript $\mathbf{x}$ means that this is not a regular product, it is our generalized cross product. This $\Pi^{\mathbf{x}}$ 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 $\Pi^{\mathbf{x}}$ 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

$$\mathbf{Q} = \Pi_{i=1 \dots N, i \neq n}^{\mathbf{x}} \mathbf{A}^{(i)} \equiv \Pi_{\mathbf{i} \neq \mathbf{n}}^{\mathbf{x}} \mathbf{A}^{(\mathbf{i})}$$

And of course this idea can be applied to the expression for $\mathbf{E}_k$

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

$$\mathbf{E}_k = \det(\mathbf{R}) (-1)^{k-1} \Pi_{\mathbf{i} \neq \mathbf{k}}^{\mathbf{x}} \mathbf{e}_i$$

Once again, for $N=2$ the $\Pi^{\mathbf{x}}$ symbol implies that ($\mathbf{e}_k$ = "missing", $\mathbf{e}_{3-k}$ = the one not missing)

$$\Pi_{\mathbf{i} \neq \mathbf{k}}^{\mathbf{x}} \mathbf{e}_i = \Pi_{i=1 \dots 2, i \neq k}^{\mathbf{x}} \mathbf{e}_i = \mathbf{e}_{3-k} \times \hat{\mathbf{3}}$$

$$\mathbf{E}_k = \det(\mathbf{R}) (-1)^{k-1} \mathbf{e}_{3-k} \times \hat{\mathbf{3}}$$

which is the "trick" notation of section (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.

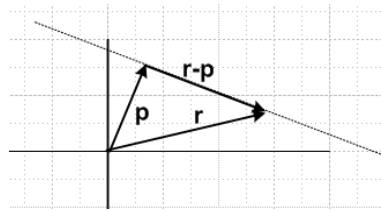
Section (d) focuses on the area and volume of N-pipeds in various dimensions, and simple expressions for the volume and vector areas of the faces of an N-piped are obtained.

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.. 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

$$\mathbf{p} \cdot (\mathbf{r}-\mathbf{p}) = 0 \quad \Rightarrow \quad \mathbf{r} \cdot \mathbf{p} = p^2 \quad \Rightarrow \quad \mathbf{r} \cdot \hat{\mathbf{p}} = p$$

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.

$$N=1 \quad \mathbf{r}_{\text{volume}1} = \alpha_1 \mathbf{e}_1 \quad 0 \leq \alpha_1 \leq 1$$

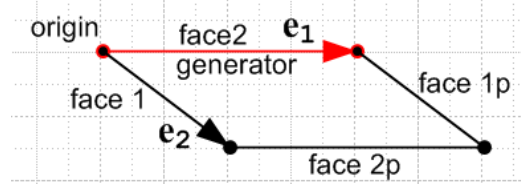
This piped has two vertices located at $\mathbf{v}_1 = 0$ and $\mathbf{v}_2 = \mathbf{e}_1$. These two vertices are also the "faces" of this 1-piped, so there are two faces (one pair of faces). These faces are 0 dimensional and therefore don't point in any direction (they are the endpoints of the segment). 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,

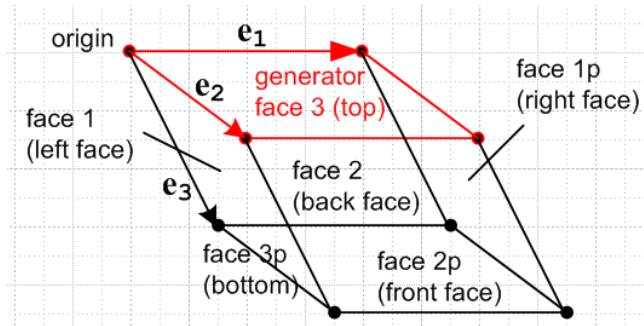
$$\begin{aligned}
 \mathbf{E}_1 \cdot \mathbf{r}_{\text{face1}} &= \mathbf{E}_1 \cdot \alpha_2 \mathbf{e}_2 = 0 & \Rightarrow & \hat{\mathbf{E}}_1 \cdot \mathbf{r}_{\text{face1}} = 0 \\
 \mathbf{E}_1 \cdot \mathbf{r}_{\text{face1p}} &= \mathbf{E}_1 \cdot [\alpha_2 \mathbf{e}_2 + \mathbf{e}_1] = 1 & \Rightarrow & \hat{\mathbf{E}}_1 \cdot \mathbf{r}_{\text{face1p}} = 1/E_1
 \end{aligned}$$

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 $\mathbb{R}^3$ 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

$$\begin{aligned}
 \mathbf{r}_{\text{volume3}} &= \alpha_1 \mathbf{e}_1 + \alpha_2 \mathbf{e}_2 + \alpha_3 \mathbf{e}_3 & 0 \leq \alpha_1, \alpha_2, \alpha_3 \leq 1 \\
 \mathbf{r}_{\text{face3}} &= \alpha_1 \mathbf{e}_1 + \alpha_2 \mathbf{e}_2 & 0 \leq \alpha_1, \alpha_2 \leq 1 & \text{// the generator face} \\
 \mathbf{r}_{\text{face3p}} &= \alpha_1 \mathbf{e}_1 + \alpha_2 \mathbf{e}_2 + \mathbf{e}_3 & 0 \leq \alpha_1, \alpha_2 \leq 1 & \text{// partner face to the above} \\
 \mathbf{r}_{\text{face2}} &= \alpha_1 \mathbf{e}_1 + \alpha_3 \mathbf{e}_3 & 0 \leq \alpha_1, \alpha_3 \leq 1 & \text{// the generator face} \\
 \mathbf{r}_{\text{face2p}} &= \alpha_1 \mathbf{e}_1 + \alpha_3 \mathbf{e}_3 + \mathbf{e}_2 & 0 \leq \alpha_1, \alpha_3 \leq 1 & \text{// partner face to the above} \\
 \mathbf{r}_{\text{face1}} &= \alpha_2 \mathbf{e}_2 + \alpha_3 \mathbf{e}_3 & 0 \leq \alpha_2, \alpha_3 \leq 1 & \text{// the generator face} \\
 \mathbf{r}_{\text{face1p}} &= \alpha_2 \mathbf{e}_2 + \alpha_3 \mathbf{e}_3 + \mathbf{e}_1 & 0 \leq \alpha_2, \alpha_3 \leq 1 & \text{// partner face to the above}
 \end{aligned}$$

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:

$$\begin{aligned} \mathbf{E}_1 \bullet \mathbf{r}_{\text{face}1} &= \mathbf{E}_1 \bullet [\alpha_2 \mathbf{e}_2 + \alpha_3 \mathbf{e}_3] = 0 & \Rightarrow & \hat{\mathbf{E}}_1 \bullet \mathbf{r}_{\text{face}1} = 0 \\ \mathbf{E}_1 \bullet \mathbf{r}_{\text{face}1p} &= \mathbf{E}_1 \bullet [\alpha_2 \mathbf{e}_2 + \alpha_3 \mathbf{e}_3 + \mathbf{e}_1] = 1 & \Rightarrow & \hat{\mathbf{E}}_1 \bullet \mathbf{r}_{\text{face}1p} = 1/|\mathbf{E}_1| \end{aligned}$$

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:

$$\begin{aligned} \mathbf{r}_{\text{volume}N} &= \sum_n \alpha_n \mathbf{e}_n & 0 \leq \alpha_n \leq 1 \\ \mathbf{r}_{\text{face}(i)} &= \sum_{n \neq i} \alpha_n \mathbf{e}_n & 0 \leq \alpha_n \leq 1 & i = 1, 2, \dots, N \\ \mathbf{r}_{\text{face}(ip)} &= \sum_{n \neq i} \alpha_n \mathbf{e}_n + \mathbf{e}_i & 0 \leq \alpha_n \leq 1 & i = 1, 2, \dots, N \end{aligned}$$

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:

$$\begin{aligned} \mathbf{E}_i \bullet \mathbf{r}_{\text{face}(i)} &= \mathbf{E}_i \bullet [\sum_{n \neq i} \alpha_n \mathbf{e}_n] = 0 & \Rightarrow & \hat{\mathbf{E}}_i \bullet \mathbf{r}_{\text{face}(i)} = 0 \\ \mathbf{E}_i \bullet \mathbf{r}_{\text{face}(ip)} &= \mathbf{E}_i \bullet [\sum_{n \neq i} \alpha_n \mathbf{e}_n + \mathbf{e}_i] = 1 & \Rightarrow & \hat{\mathbf{E}}_i \bullet \mathbf{r}_{\text{face}(ip)} = 1/E_i \end{aligned}$$

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 ip (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,

$$A_1 = |\mathbf{e}_2|$$

$$A_2 = |\mathbf{e}_1|$$

where here we maintain the plan of labeling an area by the index of the spanning vector which is omitted in making the area. The *vector* areas can be written, based on the work above,

$$\mathbf{A}_1 = |\mathbf{e}_2| \hat{\mathbf{E}}_1 \quad // \quad \mathbf{E}_k = \sum_i g'_{ki} \mathbf{e}_i, \text{ see Appendix A (g)}$$

$$\mathbf{A}_2 = |\mathbf{e}_1| \hat{\mathbf{E}}_2$$

and these vectors are out-facing for faces 1p and 2p. We claim that both these results can be expressed in a single formula

$$\mathbf{A}_n = |\det(S)| \mathbf{E}_n .$$

One can see that the direction is correct for $n = 1, 2$, so it is just a question of verifying the magnitude. One must show that

$$|\det(S)| |\mathbf{E}_1| = |\mathbf{e}_2| \quad \text{and} \quad |\det(S)| |\mathbf{E}_2| = |\mathbf{e}_1|$$

or

$$|\mathbf{E}_k| = |\det(R)| |\mathbf{e}_{3-k}| \quad k=1,2 \quad // \text{ RS} = 1$$

Using the $N=2$ trick notation from Appendix A (c),

$$\mathbf{E}_k = \det(R) (-1)^{k-1} \mathbf{e}_{3-k} \times \hat{\mathbf{3}}$$

so that

$$|\mathbf{E}_k| = |\det(R)| |\mathbf{e}_{3-k} \times \hat{\mathbf{3}}| = |\det(R)| |\mathbf{e}_{3-k}| \quad k = 1, 2$$

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):

$$\begin{aligned} \mathbf{A}_n &= |\det(\mathbf{S})| \mathbf{E}_n = |\det(\mathbf{S})| \det(\mathbf{R}) (-1)^{n-1} \prod_{i \neq n} \mathbf{e}_i \\ &= \sigma (-1)^{n-1} \prod_{i \neq n} \mathbf{e}_i \quad \sigma \equiv \text{sign}(\det(\mathbf{S})) = \text{sign}(\det(\mathbf{R})) \end{aligned}$$

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,

$$\text{volume}(2) = | \mathbf{e}_1 \times \mathbf{e}_2 | = | \varepsilon_{ab} (\mathbf{e}_1)_a (\mathbf{e}_2)_b | = | \det [\mathbf{e}_1, \mathbf{e}_2] | = | \det(\mathbf{S}) |$$

where $\mathbf{S}$ 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

$$\begin{aligned}
 \mathbf{E}_n &= \det(\mathbf{R}) (-1)^{n-1} \mathbf{e}_1 \times \dots \times \mathbf{e}_3 && // \mathbf{e}_n \text{ missing} \\
 \text{so} \\
 |\mathbf{E}_n| &= |\det(\mathbf{R})| |\mathbf{e}_1 \times \dots \times \mathbf{e}_3| && // \mathbf{e}_n \text{ missing}
 \end{aligned}$$

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

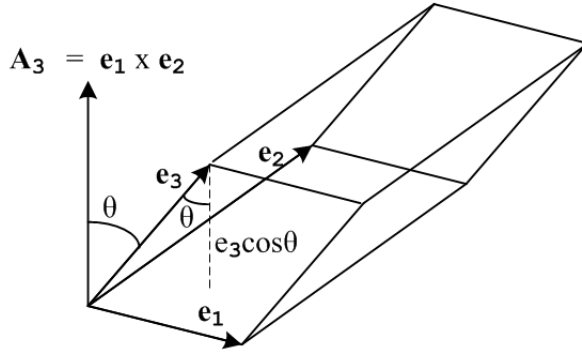
$$\begin{aligned}
 \mathbf{A}_n &= |\det(\mathbf{S})| \mathbf{E}_n = \sigma (-1)^{n-1} \mathbf{e}_1 \times \dots \times \mathbf{e}_3 && // \mathbf{e}_n \text{ missing} \\
 &= \sigma (-1)^{n-1} \prod_{i \neq n} \mathbf{e}_i
 \end{aligned}$$

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

$$\begin{aligned}
 \text{volume}(3) &= |\mathbf{A}_1 \bullet \mathbf{e}_1| = |\mathbf{A}_2 \bullet \mathbf{e}_2| = |\mathbf{A}_3 \bullet \mathbf{e}_3| \\
 \text{or} \\
 \text{volume}(3) &= |\mathbf{e}_2 \times \mathbf{e}_3 \bullet \mathbf{e}_1| = |\mathbf{e}_3 \times \mathbf{e}_1 \bullet \mathbf{e}_2| = |\mathbf{e}_1 \times \mathbf{e}_2 \bullet \mathbf{e}_3|
 \end{aligned}$$

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

$$\begin{aligned} \mathbf{e}_3 \bullet \mathbf{e}_1 \times \mathbf{e}_2 &= (\mathbf{e}_3)_i \varepsilon_{ijk} (\mathbf{e}_1)_j (\mathbf{e}_2)_k = \varepsilon_{ijk} (\mathbf{e}_1)_j (\mathbf{e}_2)_k (\mathbf{e}_3)_i = \varepsilon_{jki} (\mathbf{e}_1)_j (\mathbf{e}_2)_k (\mathbf{e}_3)_i \\ &= \det [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3] = \det(S) \end{aligned}$$

so that

$$\text{volume}(3) = |\mathbf{e}_3 \bullet \mathbf{e}_1 \times \mathbf{e}_2| = |\varepsilon_{abc} (\mathbf{e}_1)_a (\mathbf{e}_2)_b (\mathbf{e}_3)_c| = |\det [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]| = |\det(S)|$$

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

$$\text{volume}(3) = |\varepsilon_{abc} (\mathbf{e}_1)_a (\mathbf{e}_2)_b (\mathbf{e}_3)_c|$$

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 $|\varepsilon_{abc4} (\mathbf{e}_1)_a (\mathbf{e}_2)_b (\mathbf{e}_3)_c|$, where ε is now a 4D ε tensor. The conclusion is that

$$A_4 = |\varepsilon_{abc4} (\mathbf{e}_1)_a (\mathbf{e}_2)_b (\mathbf{e}_3)_c|$$

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

$$A_1 = |\varepsilon_{1abc} (\mathbf{e}_2)_a (\mathbf{e}_3)_b (\mathbf{e}_4)_c|$$

$$A_2 = |\varepsilon_{a2bc} (\mathbf{e}_3)_a (\mathbf{e}_4)_b (\mathbf{e}_1)_c|$$

$$A_3 = |\varepsilon_{ab3c} (\mathbf{e}_4)_a (\mathbf{e}_1)_b (\mathbf{e}_2)_c|$$

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

$$\begin{aligned} \mathbf{A}_n &= |\det(S)| \mathbf{E}_n = \sigma (-1)^{n-1} \mathbf{e}_1 \times \dots \times \mathbf{e}_4 \quad // \mathbf{e}_n \text{ missing} \\ &= \sigma (-1)^{n-1} \Pi_{i \neq n}^{\times} \mathbf{e}_i \quad \sigma \equiv \text{sign}[\det(S)] = \text{sign}[\det(R)] \end{aligned}$$

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 $\mathbf{e}_i$ vectors, while the Nth components are all 0. This is so because face N lies in a plane defined by this Nth component being 0. The volume(N-1) formula expressed in terms of these new ND vectors is therefore $|\varepsilon_{abc \dots xN} (\mathbf{e}_1)_a (\mathbf{e}_2)_b \dots (\mathbf{e}_{N-1})_x|$, 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),

$$(\mathbf{A}_i)_\alpha = \sigma (-1)^{i-1} \varepsilon_{\alpha \mathbf{a} \mathbf{b} \mathbf{c} \dots \mathbf{x}} (\mathbf{e}_1)_\mathbf{a} (\mathbf{e}_2)_\mathbf{b} \dots (\mathbf{e}_N)_\mathbf{x} \quad // \text{where } (\mathbf{e}_i)_i \text{ and } i \text{ are missing}$$

For $N=2$ the last two expressions for $\mathbf{A}_i$ are interpreted as shown in Appendix A (c)

$$\mathbf{A}_i = \sigma (-1)^{i-1} \mathbf{e}_{3-i} \times \hat{\mathbf{z}} \quad i = 1, 2$$

8. The volume of an N-piped spanned by $\mathbf{e}_1 \dots \mathbf{e}_N$ is given by

$$\text{volume}(N) = |\det(\mathbf{S})| = |\det [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \dots, \mathbf{e}_N]| = |\varepsilon_{\mathbf{a} \mathbf{b} \mathbf{c} \dots \mathbf{x}} (\mathbf{e}_1)_\mathbf{a} (\mathbf{e}_2)_\mathbf{b} (\mathbf{e}_3)_\mathbf{c} \dots (\mathbf{e}_N)_\mathbf{x}|$$

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$	$\mathbf{e}_\rho = \mathbf{e}_\rho \hat{\mathbf{e}}_\rho$

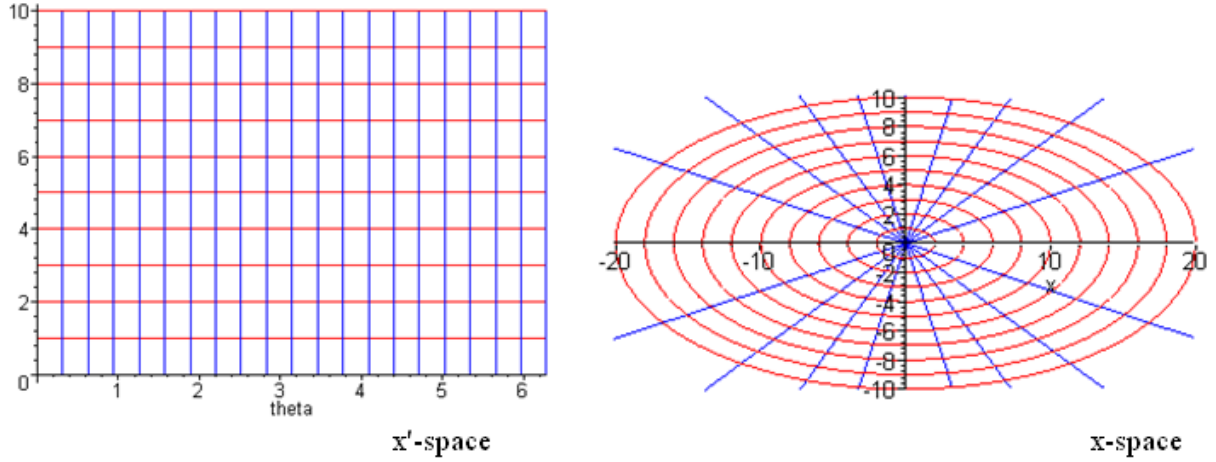
The covariant metric tensor is,

$$\bar{g}' = S^T S = \begin{pmatrix} \rho^2 \{a^2 \sin^2(\theta) + b^2 \cos^2(\theta)\} & [b^2 - a^2] \rho \sin(\theta) \cos(\theta) \\ [b^2 - a^2] \rho \sin(\theta) \cos(\theta) & a^2 \cos^2(\theta) + b^2 \sin^2(\theta) \end{pmatrix} = \begin{pmatrix} e_1 \bullet e_1 & e_1 \bullet e_2 \\ e_2 \bullet e_1 & e_2 \bullet e_2 \end{pmatrix}$$

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 $b = 1$)



where ρ is the x' -space vertical axis. 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

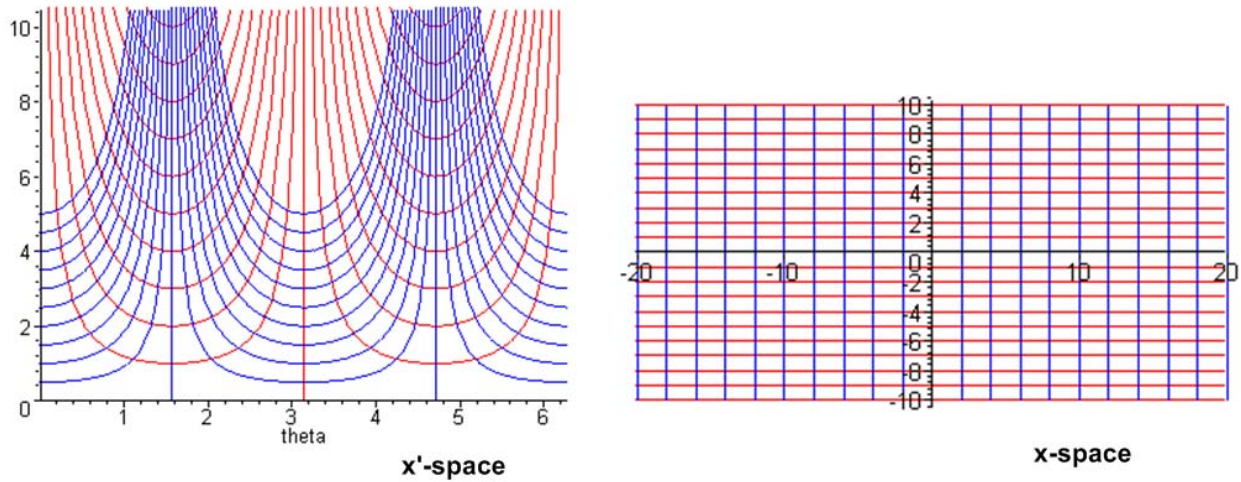
which are obtained from the forward transformation equations

$$\rho^2 = x^2/a^2 + y^2/b^2$$

$$\tan \theta = y/x$$

(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

$$\begin{array}{l} x = a\rho \cos \theta \\ y = b\rho \sin \theta \end{array}$$

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 V'_n .

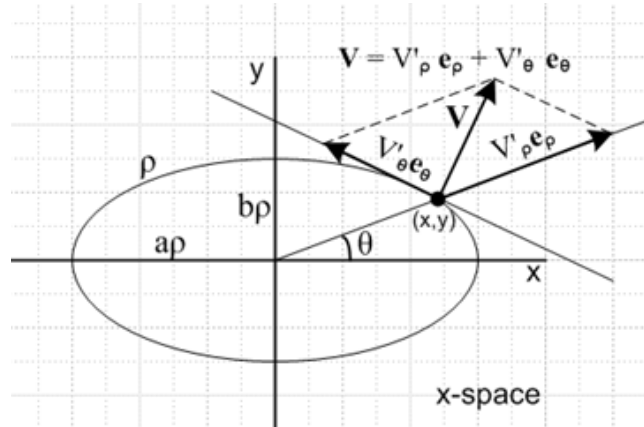
A contravariant vector field $\mathbf{V}(\mathbf{x})$ can be expanded in these two ways (Section 6 (f))

$$\mathbf{V} = V_1(\mathbf{x}) \hat{\mathbf{1}} + V_2(\mathbf{x}) \hat{\mathbf{2}} = V_x(\mathbf{x}) \hat{\mathbf{x}} + V_y(\mathbf{x}) \hat{\mathbf{y}} \quad // \quad \mathbf{u}_n = \hat{\mathbf{n}} \text{ for Cartesian}$$

$$\mathbf{V} = V'_1(\mathbf{x}') \mathbf{e}_1 + V'_2(\mathbf{x}') \mathbf{e}_2 = V'_\theta(\mathbf{x}') \mathbf{e}_\theta + V'_\rho(\mathbf{x}') \mathbf{e}_\rho \quad \mathbf{V}'(\mathbf{x}') = \mathbf{R}(\mathbf{x}) \mathbf{V}(\mathbf{x})$$

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" $\mathbf{x}'$. 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 $\mathbf{V}'$, the transformed vector in x' -space, when it is expanded on the axis-aligned vectors $\mathbf{e}'_n$ in x' -space.

Here is a graphical representation of this vector $\mathbf{V}$ in x-space:

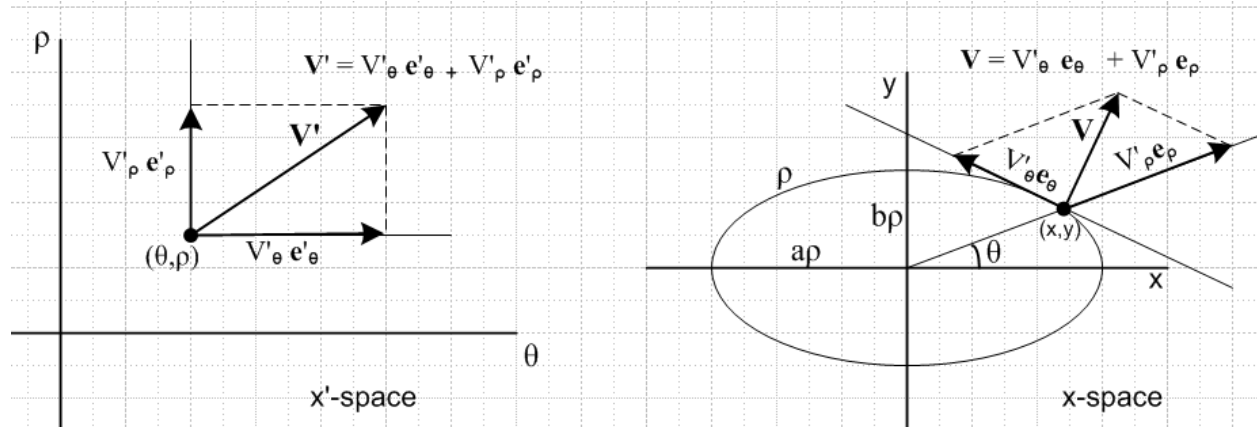


As advertised, the tangent base vectors are not at right angles. The $\mathbf{V}$ parallelogram accurately illustrates the equation $\mathbf{V} = V'_\theta \mathbf{e}_\theta + V'_\rho \mathbf{e}_\rho$. Since x-space is Cartesian, there is no distinction between Cartesian length (graphical length) and 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 R and use $V'_i = R_{ij} V_j$.

(e) Drawing a contravariant vector $\mathbf{V}'$ in x'-space: 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 axis-aligned 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 equation $\mathbf{V} = V'_\theta \mathbf{e}_\theta + V'_\rho \mathbf{e}_\rho$ one gets the expansion noted above,

$$\mathbf{V}' = V'_\theta \mathbf{e}'_\theta + V'_\rho \mathbf{e}'_\rho$$

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 $|\mathbf{A}|^2 = \delta_{ij} A_i A_j = \sum A_i^2$, so one has, since $(\mathbf{e}'_n)_i = \delta_{n,i}$,

$$\bar{g}' = 1 \quad |\mathbf{e}'_n| = 1 \quad \mathbf{e}'_n = \hat{\mathbf{e}}'_n = \hat{\mathbf{h}}'_n = \text{the usual axis-aligned unit vectors in } x'\text{-space}$$

$$\mathbf{V}' = V'_\theta \hat{\boldsymbol{\theta}} + V'_\rho \hat{\boldsymbol{\rho}} \quad \hat{\boldsymbol{\theta}} = \hat{\mathbf{e}}'_1 \quad \hat{\boldsymbol{\rho}} = \hat{\mathbf{e}}'_2$$

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

$$\hat{\boldsymbol{\theta}} \cdot \hat{\boldsymbol{\theta}} = 1$$

$$\hat{\boldsymbol{\rho}} \cdot \hat{\boldsymbol{\rho}} = 1$$

$$\hat{\boldsymbol{\theta}} \cdot \hat{\boldsymbol{\rho}} = 0$$

This Cartesian View, which is x' -space with $\bar{g}' = 1$, is appropriate in applications in which it is not required or desired that norms and dot products be tensorial scalars, as discussed at the end of Section 5 (a). For example, when $\bar{g}'$ is set to 1, one has $|\mathbf{V}'| \neq |\mathbf{V}|$ in the above picture.

Another use of this view involves integration as will be seen below.

In the Curvilinear View of x' -space, one assumes that $\bar{g}'$ takes a value which enforces the scalarity of norms and dot products between x -space and x' -space, which is to say, one takes $\bar{g}' = S^T \bar{g} S$ where $\bar{g}$ is the x -space metric tensor for x -space. Normally $\bar{g}=1$ (Cartesian x -space), so $\bar{g}' = S^T S$. In this Curvilinear View, then, the length $|\mathbf{A}'|$ of a contravariant vector $\mathbf{A}'$ is determined by $|\mathbf{A}'|^2 = \bar{g}'_{ij} A'_i A'_j$ where

$$\bar{g}' = S^T S \neq 1 \quad |\mathbf{e}'_n| = |\mathbf{e}_n| = h'_n \equiv \sqrt{\bar{g}'_{nn}} \quad n = 1, 2 \text{ for } \theta, \rho$$

$$\mathbf{V}' = V'_\theta \mathbf{e}'_\theta + V'_\rho \mathbf{e}'_\rho = V'_\theta h'_\theta \hat{\mathbf{e}}'_\theta + V'_\rho h'_\rho \hat{\mathbf{e}}'_\rho \quad \hat{\mathbf{e}}'_n \equiv \mathbf{e}'_n / |\mathbf{e}_n| = \mathbf{e}'_n / h'_n$$

$$|\mathbf{V}'|^2 = |\mathbf{V}|^2 = \mathbf{V}_{\mathbf{x}}^2 + \mathbf{V}_{\mathbf{y}}^2 \neq (\mathbf{V}'_{\theta} h'_{\theta})^2 + (\mathbf{V}'_{\rho} h'_{\rho})^2 \quad // \text{ unless } \mathbf{x}'_{\mathbf{i}} \text{ are orthogonal coordinates}$$

This last inequality says that in the Curvilinear View the Pythagorean Theorem is invalid. In fact

$$\begin{aligned} |\mathbf{V}'|^2 &= \bar{g}'_{ij} V'_i V'_j = \bar{g}'_{\theta\theta} V'^2_{\theta} + \bar{g}'_{\rho\rho} V'^2_{\rho} + 2 \bar{g}'_{\theta\rho} V'_{\theta} V'_{\rho} \\ &= (\mathbf{V}'_{\theta} h'_{\theta})^2 + (\mathbf{V}'_{\rho} h'_{\rho})^2 + 2 \bar{g}'_{\theta\rho} V'_{\theta} V'_{\rho} \end{aligned}$$

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,

$$\begin{aligned} \hat{\mathbf{e}}'_{\theta} \bullet \hat{\mathbf{e}}'_{\theta} &= \mathbf{e}'_{\theta} \bullet \mathbf{e}'_{\theta} / (h'^2_{\theta}) = \mathbf{e}_{\theta} \bullet \mathbf{e}_{\theta} / (h'^2_{\theta}) = \bar{g}'_{\theta\theta} / (h'^2_{\theta}) = 1 \\ \hat{\mathbf{e}}'_{\rho} \bullet \hat{\mathbf{e}}'_{\rho} &= \mathbf{e}'_{\rho} \bullet \mathbf{e}'_{\rho} / (h'^2_{\rho}) = \mathbf{e}_{\rho} \bullet \mathbf{e}_{\rho} / (h'^2_{\rho}) = \bar{g}'_{\rho\rho} / (h'^2_{\rho}) = 1 \\ \hat{\mathbf{e}}'_{\theta} \bullet \hat{\mathbf{e}}'_{\rho} &= \mathbf{e}'_{\theta} \bullet \mathbf{e}'_{\rho} / (h'_{\theta} h'_{\rho}) = \mathbf{e}_{\theta} \bullet \mathbf{e}_{\rho} / (h'_{\theta} h'_{\rho}) = \bar{g}'_{\theta\rho} / (h'_{\theta} h'_{\rho}) \neq 0 \quad <= !! \end{aligned}$$

so that the $\hat{\mathbf{e}}'_n$ are unit vectors having unit covariant length, but $\hat{\mathbf{e}}'_{\theta} \bullet \hat{\mathbf{e}}'_{\rho} \neq 0$ despite the fact that these vectors are drawn at right angles in the $\mathbf{x}'$ -space graph above, $\mathbf{e}_n = h'_n \hat{\mathbf{e}}'_n$. One might imagine trying to slant the lines of the $\mathbf{x}'$ -space graph to cause all intersection points to have angles which match the metric tensor, which is to say, at each intersection point one would need an angle ψ where $\hat{\mathbf{e}}'_{\theta} \bullet \hat{\mathbf{e}}'_{\rho} = \cos\psi$. But in general $\hat{\mathbf{e}}'_n \bullet \hat{\mathbf{e}}'_m = \bar{g}'_{nm} / (h'_n h'_m)$ has a different value at every point, so such a graph would be quite complex.

The upshot is that for a *non-orthogonal* system, the axes in $\mathbf{x}'$ -space are still drawn at right angles and the purpose of the graph is mainly to "locate" all the points $\mathbf{x}'$ which correspond to points $\mathbf{x}$ in $\mathbf{x}$ -space according to $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. The graph does successfully represent the idea that $\mathbf{V}' = V'_{\rho} \mathbf{e}'_{\rho} + V'_{\theta} \mathbf{e}'_{\theta}$, but one must give up on Euclidean geometry for this vector sum triangle. It might be imagined that the $\mathbf{x}'$ -space graph is the projection onto the plane of paper of some vectors drawn on a curved surface emerging from the plane of paper, and that is then why Pythagoras is wrong.

In the case of an *orthogonal* coordinate system (diagonal $\bar{g}'$), the 90 degree angles between the axes in $\mathbf{x}'$ -space are accurate representations of the fact that $\hat{\mathbf{e}}'_n \bullet \hat{\mathbf{e}}'_m = 0$ when $n \neq m$. And since scalars are preserved, one has in the Curvilinear View,

$$|\mathbf{V}'|^2 = \sum_n (h'_n V'_n)^2 = \sum_n \mathbf{v}'_n{}^2 = |\mathbf{V}|^2 = \sum_n V_n{}^2 \quad // \text{ orthogonal only}$$

where

$$\begin{aligned} \mathbf{v}'_n &\equiv h'_n V'_n \quad \text{and} \quad \mathbf{V}' = \sum_n \mathbf{v}'_n \hat{\mathbf{e}}'_n \\ \text{and} \quad \mathbf{V} &= \sum_n \mathbf{v}_n \hat{\mathbf{e}}_n \end{aligned}$$

One can then still apply regular Euclidean geometry to the vector addition N-piped in $\mathbf{x}'$ -space in the sense that $|\mathbf{V}'|^2 = \sum_n (h'_n V'_n)^2$.

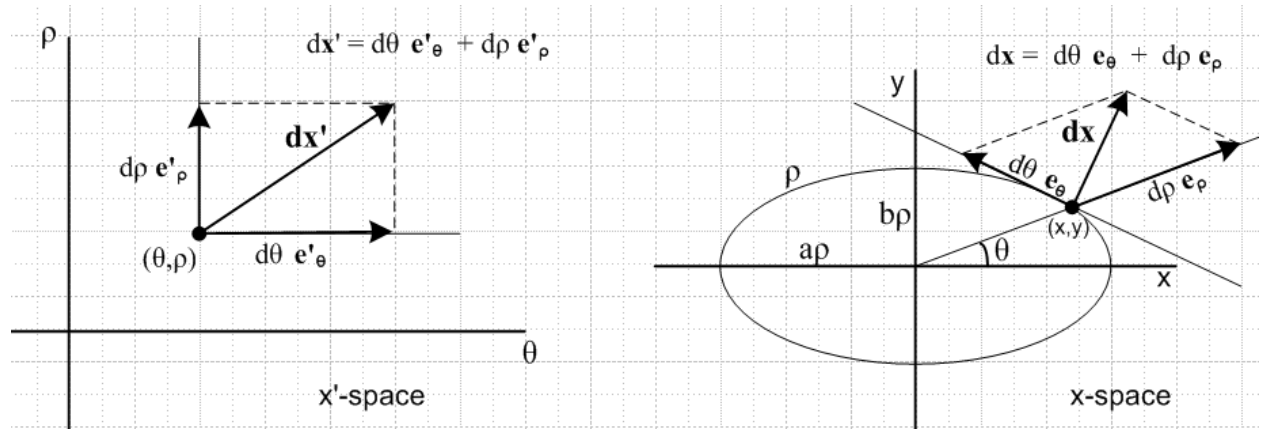
(f) Drawing the specific contravariant vector dx in x-space and x' -space

Since dx is the primordial contravariant vector, everything stated in the last two sections applies with $V \rightarrow dx$ 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 dx and dx' are,

$$dx = d\theta \mathbf{e}_\theta + d\rho \mathbf{e}_\rho \quad // \text{ in x-space}$$

$$dx' = d\theta \mathbf{e}'_\theta + d\rho \mathbf{e}'_\rho \quad // \text{ in x'-space}$$

For $V = dx$ the picture above becomes



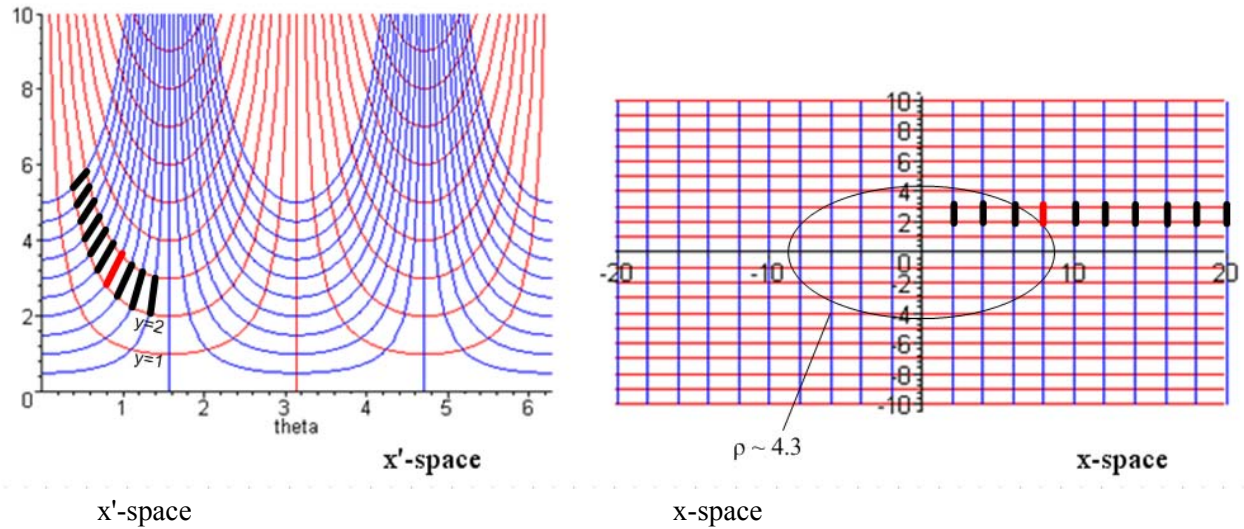
It must be understood that now the vector arrows like dx are highly magnified and in reality are very small compared to, say, the curvature of the ellipse. From above,

$$\begin{aligned} \mathbf{e}'_\theta \cdot \mathbf{e}'_\theta &= \bar{g}'_{\theta\theta} & \hat{\mathbf{e}}'_\theta \cdot \hat{\mathbf{e}}'_\theta &= 1 \\ \mathbf{e}'_\rho \cdot \mathbf{e}'_\rho &= \bar{g}'_{\rho\rho} & \hat{\mathbf{e}}'_\rho \cdot \hat{\mathbf{e}}'_\rho &= 1 \\ \mathbf{e}'_\rho \cdot \mathbf{e}'_\theta &= \bar{g}'_{\rho\theta} & \hat{\mathbf{e}}'_\theta \cdot \hat{\mathbf{e}}'_\rho &= \bar{g}'_{\theta\rho} / (h'_\theta h'_\rho) \end{aligned}$$

and once again the "right angle" in the x' -space picture is deceptive.

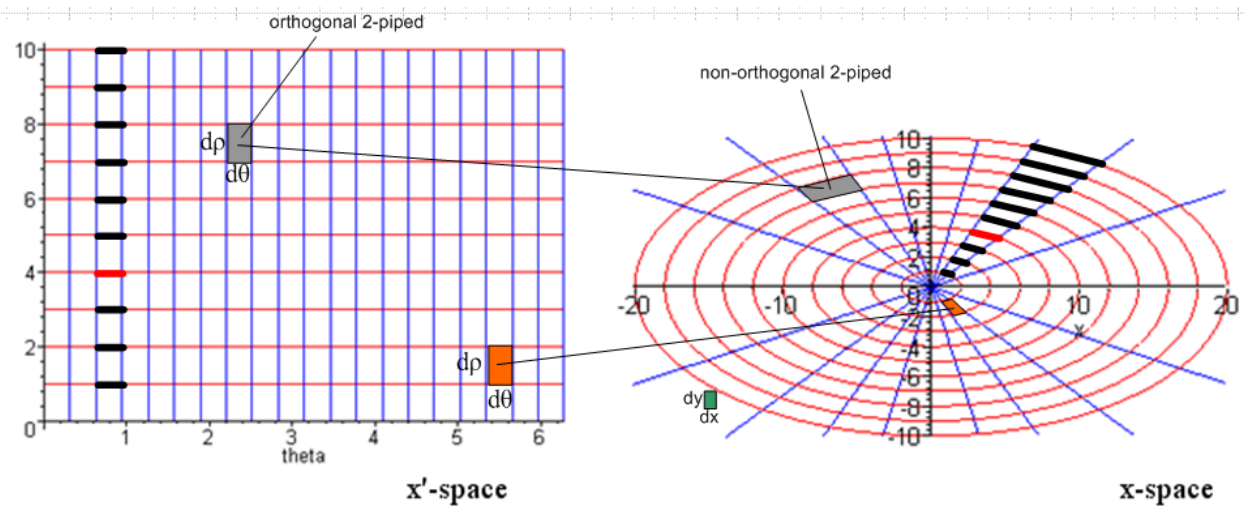
(g) Study of how dx transforms in the mapping between x-space and x' -space

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 $d\mathbf{x}$ bar on the right and operationally apply *the previous picture*. First determine the tangent base vectors $\mathbf{e}_\theta$ and $\mathbf{e}_\rho$ at the location of the red bar. Then setting $d\mathbf{x} = d\theta \mathbf{e}_\theta + d\rho \mathbf{e}_\rho$, consider the value of the two numbers $d\theta$ and $d\rho$ for this red bar. Graphically, knowing which way $\mathbf{e}_\theta$ and $\mathbf{e}_\rho$ point at the bottom of the red $d\mathbf{x}$, one expects $d\theta > 0$ and $d\rho > 0$. The red $d\mathbf{x}'$ bar on the left has these Cartesian values $d\theta$ and $d\rho$, and has a Cartesian-view length of $|d\mathbf{x}|^2 = (d\theta)^2 + (d\rho)^2$. One can see from the picture that these Cartesian lengths *vary* for the 10 bars shown, though the lengths are all the same in x-space. The Curvilinear-view lengths of the x'-space bars are all the same, and are equal to the Cartesian length of those bars in x-space since $d\mathbf{x}' \cdot d\mathbf{x}' = d\mathbf{x} \cdot d\mathbf{x}$.

Consider now some bar mapping in the other direction:



Now the $d\mathbf{x}$ bars on the right all have different lengths. Those on the left have the same Cartesian length, which is what the drawing shows, but each one's Curvilinear-view length matches that of its corresponding bar on the right. The ratio of the length of a bar on the right to the Cartesian length of the corresponding bar on the left is the scale factor h_θ which recall is a function of location in space:

$$\begin{aligned} \text{bar on right} &= d\mathbf{x}^{(1)} = \mathbf{e}_1 dx'^1 = \hat{\mathbf{e}}_1 h'_1 dx'^1 = \mathbf{e}_\theta d\theta = \hat{\mathbf{e}}_\theta h_\theta d\theta & \text{graph length} &= h_\theta d\theta \\ \text{bar on left (Cartesian view)} &= d\mathbf{x}'^{(1)} = \mathbf{e}'_1 dx'^1 = \hat{\mathbf{e}}'_1 dx'^1 = \hat{\mathbf{e}}_\theta d\theta & \text{graph length} &= d\theta \end{aligned}$$

$$\Rightarrow \text{right bar length} / \text{left bar length} = h_\theta = \rho \sqrt{a^2 \sin^2 \theta + b^2 \cos^2 \theta} \quad (\text{increases with } \rho)$$

If $a=b$, then $h_\theta = \rho$ and the bar length on the right is then $\rho d\theta$ as is obvious in polar coordinates.

(h) Derivation of the Jacobian Integration Rule

Consider now an integral $\int d\theta d\rho f(\theta, \rho)$. The tiny rectangles of area $d\theta d\rho$, like the specific gray and orange ones highlighted on the left above, are regarded for the purposes of integration as being in the Cartesian view of $\mathbf{x}'$ -space. One then writes [dA' is called $d\mathcal{V}'$ in Section 8]

$$dA' \equiv d\rho d\theta = \text{the area of a differential patch in Cartesian-view } \mathbf{x}'\text{-space}$$

This is the graphical area one sees in the picture. There is no need to define or consider any Curvilinear-view area in $\mathbf{x}'$ -space because the Cartesian-view area is being used.

In the limiting process which defines the integration, each $d\theta d\rho$ patch on the left has the same area $d\rho d\theta$. The interior of each patch on the left maps into some parallelogram patch on the right. One is not surprised to see that the patch areas on the right are different, though they map into patches on the left of the same Cartesian-view area. As shown in Section 8 (e), the ratio of the two patch areas is the absolute value of the Jacobian $|J(\mathbf{x}')|$,

$$(\text{area of skewed patch on the right at location } \mathbf{x}) = |J(\mathbf{x}')| dA' = |J(\mathbf{x}')| d\rho d\theta$$

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) $\mathbf{x}'$ -space mapping into a non-orthogonal differential N -piped in $\mathbf{x}$ -space.

Now back to the integration issue. There are two ways an integration can be done in Cartesian $\mathbf{x}$ -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 $|J(\mathbf{x}')|dA' = |J(\mathbf{x}')|d\theta d\rho$ and the areas are different because the Jacobian is a function of $\mathbf{x} = \mathbf{x}(\mathbf{x}')$.

In the second integral, every patch $dA_2(\mathbf{x}_i)$ has the same area $dxdy$, so really $dA_2(\mathbf{x}_i)$ does not depend on $\mathbf{x}_i$ in this form of the integration. One such $dxdy$ patch is shown in green above. The coverage of the dA_2 patches is of course also "perfect coverage" in the calculus limit.

Since both integrals cover the same area perfectly, they both give the same result in the limiting process that defines the integral. This point is sometimes misunderstood. One is not just "replacing" a parallelogram patch such as the orange one on the right with some $dxdy$ patch that approximates it in area, like the green patch. The statement is about an integration. Thus one has

$$\lim \sum_i dA_1(\mathbf{x}_i) f(\mathbf{x}_i) = \lim \sum_i dA_2(\mathbf{x}_i) f(\mathbf{x}_i)$$

or

$$\int [|J(\mathbf{x}')| d\theta d\rho] f(\mathbf{x}(\mathbf{x}')) = \int [dxdy] f(\mathbf{x})$$

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

$$|J(\theta, \rho)| d\rho d\theta = dxdy$$

where the meaning of this symbolic equality is the integral statement above,

$$\int_D dxdy f(\mathbf{x}) = \int_{D'} d\theta d\rho |J(\mathbf{x}')| f(\mathbf{x}(\mathbf{x}')) ,$$

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

$$dA \equiv dxdy$$

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

$dV = dx_1 dx_2 \dots dx_N$ = 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 that that mapping which involved a non-orthogonal N-piped in x -space.

To finish off our sample $N=2$ case, recall from earlier that for our polar elliptical coordinate system

$$|J'(x')| = |\det(S)| = ab\rho$$

and therefore

$$\int dx dy f(x,y) = \int d\theta d\rho |J(\mathbf{x}')| f(a\rho \cos\theta, b\rho \sin\theta) = ab \int d\theta d\rho \rho f(a\rho \cos\theta, b\rho \sin\theta)$$

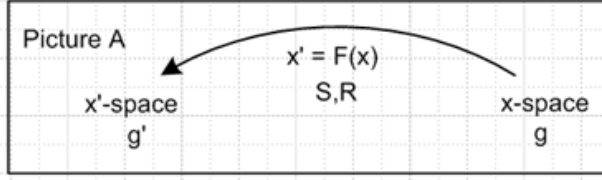
In the limit of regular polar coordinates, one then has $a = b = 1$ and $\rho = r$ so

$$\int dx dy f(x,y) = \int r dr d\theta f(r \cos\theta, r \sin\theta)$$

which is the familiar result.

Appendix D: Tensor Densities and the ϵ tensor

Picture A is used in this Appendix along with Standard Notation.



(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, we allow for the possibility of $J < 0$.

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 a regular standard-issue tensorial tensor.

Now suppose instead that the object T were to transform like this, with J being the Jacobian noted above,

$$T'^{abc}_{de} = J^{-W} R^a_a R^b_b R^c_c R_d^{d'} R_e^{e'} T^{a'b'c'}_{d'e'}$$

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

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

This is the scalar density mentioned in Section 5 (k) of weight -2 . Notice that from g one can construct other scalar densities of other weights, for example

$$g'^{-1} = J^{-(2)} g^{-1} \Rightarrow \text{weight}(g^{-1}) = +2 \quad g'^{-1} \text{ is a scalar density of weight } +2$$

(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 density does not alter its weight W . If indices a and d are contracted in the example above, one gets

$$\begin{aligned} T'^{abc}_{ae} &= J^{-W} R^a_a R^b_b R^c_c R^{d'}_a R^{e'}_e T^{a'b'c'}_{d'e'} \\ &= J^{-W} (R^a_a R^{d'}_a) R^b_b R^c_c R^{e'}_e T^{a'b'c'}_{d'e'} \\ &= J^{-W} \delta_a^{d'} R^b_b R^c_c R^{e'}_e T^{a'b'c'}_{d'e'} \\ &= J^{-W} R^b_b R^c_c R^{e'}_e T^{a'b'c'}_{a'e'} \end{aligned}$$

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.

Example 1:

$$\begin{aligned} A'^a &= J^{-W1} R^a_a A^{a'} \\ B'^c_d &= J^{-W2} R^c_c R^{d'}_d B^{c'}_{d'} \\ \Rightarrow (A'^a B'^c_d) &= J^{-(W1+W2)} R^a_a R^c_c R^{d'}_d (A^{a'} B^{c'}_{d'}) \end{aligned}$$

Example 2: Suppose in the outer product the first factor is the scalar density $g^{-W1/2}$ of weight $W1$:

$$\begin{aligned} g'^{-W1/2} &= J^{-W1} g^{-W1/2} \quad // \text{ since } g' = J^2 g \text{ from Section 5 (k), no } R \text{ factors since scalar} \\ B'^c_d &= J^{-W2} R^c_c R^{d'}_d B^{c'}_{d'} \quad // \text{ same as in previous example} \\ \Rightarrow g'^{-W1/2} B'^c_d &= J^{-(W1+W2)} R^a_a R^c_c R^{d'}_d (g^{-W1/2} B^{c'}_{d'}) \end{aligned}$$

If one selects $W_1 = -W_2$, the added factor *neutralizes* the weight of the tensor density to which it is prefixed, generating thereby a regular tensor (weight 0). So if tensor density B has weight W ,

$$(g^{W/2} B^c_d) = R^a_a R^c_c R_d^{d'} (g^{W/2} B^{c'}_{d'})$$

and then $(g^{W/2} B^i_j)$ transforms under F as a regular tensor. (One should always keep in mind the fact that there is an underlying transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ upon which the House of Tensor is built).

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,

$$T^{abc}_{de} = J^{-W} R^a_a R^b_b R^c_c R_d^{d'} R_e^{e'} T^{a'b'c'}_{d'e'}$$

$$T^{abc}_d{}^e = g^{ex} T^{abc}_{dx} \quad // \text{raise last index in } x\text{-space}$$

$$T^{a'b'c'}_{d'e'} = g_{e'e''} T^{a'b'c'}_{d,e''} \quad // \text{lower last index in } x\text{-space}$$

Therefore

$$\begin{aligned} T^{abc}_d{}^e &= g^{ex} [J^{-W} R^a_a R^b_b R^c_c R_d^{d'} R_x^{e'} T^{a'b'c'}_{d'e'}] \\ &= g^{ex} [J^{-W} R^a_a R^b_b R^c_c R_d^{d'} R_x^{e'} (g_{e'e''} T^{a'b'c'}_{d,e''})] \\ &= J^W R^a_a R^b_b R^c_c R_d^{d'} (g^{ex} R_x^{e'} g_{e'e''}) T^{a'b'c'}_{d,e''} \\ &= J^{-W} R^a_a R^b_b R^c_c R_d^{d'} (R^{e''}_{e'}) T^{a'b'c'}_{d,e''} \quad // \text{Section 7 (o) facts about } R \end{aligned}$$

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 $\mathbf{A}$ and $\mathbf{B}$ of weights W and w is a scalar density of weight $W + w$ and therefore $\mathbf{A} \bullet \mathbf{B}' = J^{-(W+w)} \mathbf{A} \bullet \mathbf{B}$.

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 $\mathbf{A}$ 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. As $J \rightarrow 1$, tensor densities become true tensors. One could imagine some limiting/morphing process on an underlying transformation F such that the linearized transformation matrix R approaches a rotation matrix at all points in space ($RR^T = 1$ and $\det R = 1$) and then $J = \det S \rightarrow 1$. In this case $J^{-W} \rightarrow 1^{-W} = 1$ and therefore any tensor density, regardless of its weight W , becomes an ordinary tensor. Perhaps we should restrict this comment to underlying transformations F having $\det S > 0$ since passing through $\det S = 0$ is problematical.

An example: the cross product considered in section (g) below of $N-1$ contravariant vectors becomes in this limit an ordinary covariant vector. If $g=1$ in x -space, then $g' = RR^T = 1$ in x' -space and then that resulting vector can be considered either contravariant or covariant since both spaces are then Cartesian. This is the case with $\mathbf{A} = \mathbf{B} \times \mathbf{C}$ under rotations in 3D space. One can think of the ϵ_{abc} as moving in this limit from a tensor density of weight -1 to an ordinary tensor.

(c) Theorem about Totally Antisymmetric Tensors: there is really only one: $\epsilon^{abc\dots}$

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

$$\begin{aligned} r^{abc\dots} &= p(f \epsilon^{123\dots}) \\ \epsilon^{abc\dots} &= p \epsilon^{123\dots} \end{aligned}$$

and dividing these two equations one finds,

$$r^{abc\dots} = f \epsilon^{abc\dots}$$

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

$$\varepsilon'^{abc\dots} = J^{-W} R^a_a R^b_b \dots \varepsilon^{a'b'c'\dots} \quad (*)$$

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

$$\begin{aligned} \varepsilon'^{bac\dots} &= J^{-W} R^b_a R^a_b \dots \varepsilon^{a'b'c'\dots} = J^{-W} R^b_b R^a_a \dots \varepsilon^{b'a'c'\dots} \\ &= J^{-W} R^a_a R^b_b \dots (-\varepsilon^{a'b'c'\dots}) = -\varepsilon'^{abc\dots} \end{aligned}$$

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

$$\varepsilon'^{abc\dots} = K \varepsilon^{abc\dots}$$

where K is some scalar function, perhaps just a constant. Equation (*) above then reads

$$K \varepsilon^{abc\dots} = J^{-W} R^a_a R^b_b \dots \varepsilon^{a'b'c'\dots} \quad (**)$$

Setting in the standard order, one finds that

$$K \varepsilon^{123\dots} = J^{-W} R^1_a R^2_b \dots \varepsilon^{a'b'c'\dots}$$

or

$$K = J^{-W} \det(R^i_j) = J^{-W} (J)^{-1} = J^{-(W+1)}$$

so now

$$\varepsilon'^{abc\dots} = K \varepsilon^{abc\dots} = J^{-(W+1)} \varepsilon^{abc\dots}$$

A second assumption is now made: that $\varepsilon^{abc\dots}$ (contravariant!) has the same value structure in any frame of reference, which is to say it is the *same* in x' -space as it is in x -space,

$$\varepsilon'^{abc\dots} = \varepsilon^{abc\dots}$$

This assumption is consistent with taking $W = -1$ in the previous equation.

Again, this follows the convention of Weinberg p 99. Some authors instead arrange for the above equation to be true for the covariant ε tensors, and use then $\varepsilon_{123\dots N} = \varepsilon'_{123\dots N} = +1$, but we shall follow Weinberg.

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,

$$\epsilon'^{abc\dots} = J R^a_a R^b_b \dots \epsilon^{a'b'c'\dots} \quad // \text{ this is (*) above with } W = -1$$

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 $abc\dots$ (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 indices 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

$$\epsilon'_{abc\dots} = [\det(g'_{ij}) / \det(g_{ij})] \epsilon_{abc\dots} = (g'/g) \epsilon_{abc\dots}$$

From Section 5 (k) one has $(g'/g) = J^2$ so the conclusions regarding ϵ are these:

$$\begin{aligned} \epsilon'_{abc\dots} &= J^2 \epsilon_{abc\dots} = (g'/g) \epsilon_{abc\dots} & \epsilon'^{abc\dots} &= \epsilon^{abc\dots} = \text{permutation tensor} \quad // \text{ general} \\ & & \epsilon_{abc\dots} &= \text{permutation tensor if } g=1 \end{aligned}$$

These conclusions are valid for the "added sign s convention" as well since $\det(g)$ and $\det(g')$ always have the same sign as shown in Section 5 (k).

Two comments:

- Although we set $\epsilon'^{abc\dots} = \epsilon^{abc\dots}$ by fiat, we cannot similarly set $\epsilon'_{abc\dots} = \epsilon_{abc\dots}$ by fiat. This latter result comes out being $\epsilon'_{abc\dots} = J^2 \epsilon_{abc\dots}$ as just shown.
- The fact that $\epsilon'_{abc\dots} = J^2 \epsilon_{abc\dots}$ 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 : repeat section (d) as if its weight were not known

According to section (b) 5, lowering indices does not change the weight of a tensor density. Section (d) showed that $\epsilon^{abc\dots}$ is a tensor density of weight -1, so we know right away that $\epsilon_{abc\dots}$ is also a tensor density of weight -1. Nevertheless, it is interesting to see what happens when the same method used in section (d) for $\epsilon^{abc\dots}$ is applied to $\epsilon_{abc\dots}$.

We start by assuming $\epsilon_{abc\dots}$ is a tensor density of some unknown weight W ,

$$\epsilon'_{abc\dots} = J^{-W} [R_a^{a'} R_b^{b'} \dots \epsilon_{a'b'c'\dots}] \quad (*)$$

Section (e) 2 noted that $\epsilon_{a'b'c'\dots}$ 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 $\text{RHS}(*) = K \epsilon_{abc\dots}$, so (*) then says

$$K \epsilon_{abc\dots} = J^{-W} [R_a^{a'} R_b^{b'} \dots \epsilon_{a'b'c'\dots}] \quad (**)$$

Use section (e) 2 to set $\epsilon_{a'b'c'\dots} = \det(g_{ij}) \epsilon^{a'b'c'\dots}$ inside the bracket,

$$K \epsilon_{abc\dots} = J^{-W} [R_a^{a'} R_b^{b'} \dots \det(g_{ij}) \epsilon^{a'b'c'\dots}] ,$$

and then install the reference sequence on both sides

$$K \epsilon_{123\dots} = J^{-W} [R_1^{a'} R_2^{b'} \dots \det(g_{ij}) \epsilon^{a'b'c'\dots}]$$

But section (e) 1 says $\epsilon_{123\dots} = \det(g_{ij})$, so cancel $\det(g_{ij})$ on both sides to get

$$K = J^{-W} [R_1^{a'} R_2^{b'} \dots \epsilon^{a'b'c'\dots}] = J^{-W} \det(R_i^j) = J^{-W} \det(S^j_i) = J^{-W} J = J^{-(W-1)}$$

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

$$\epsilon'_{abc\dots} = K \epsilon_{abc\dots} = J^{-(W-1)} \epsilon_{abc\dots}$$

But section (e) 3 said that $\epsilon'_{abc\dots} = J^2 \epsilon_{abc\dots}$ and therefore one gets $W = -1$.

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 vector 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

It has just been shown that $\mathbf{C} = \mathbf{A} \times \mathbf{B}$ is a vector density of weight -1, this being a special case of the generalized cross product discussion above. As noted in section (a) 7, if R happens to be a (global) rotation, then $\mathbf{C}$ is in fact a tensorial vector. One might conjecture that $\mathbf{C} = \nabla \times \mathbf{B}$ is also a vector density of weight -1, and that conjecture is correct as is now shown. Consider

$$C^n = \varepsilon^{nab} \partial_a B_b$$

where $\mathbf{B}$ is assumed to be an tensorial vector. It is helpful to write this equation in the following manner

$$C^n = \varepsilon^{nab} [\partial_a B_b - \partial_b B_a] / 2$$

where the second term is the same as the first term, since

$$-\varepsilon^{nab} \partial_b B_a = \varepsilon^{nba} \partial_b B_a = \varepsilon^{nab} \partial_a B_b.$$

Recall now from Section 7 (v) that the covariant derivative of a vector is given by

$$B_{b;a} = \partial_a B_b - \Gamma_{ab}^c B_c$$

where the affine connection Γ_{ab}^c is symmetric under $a \leftrightarrow b$. Therefore

$$B_{b;a} - B_{a;b} = [\partial_a B_b - \Gamma_{ab}^c B_c] - [\partial_b B_a - \Gamma_{ba}^c B_c] = \partial_a B_b - \partial_b B_a$$

Therefore C^n can be expressed as

$$C^n = \varepsilon^{nab} [B_{b;a} - B_{a;b}] / 2$$

so by the same ε anti-symmetry noted above the final result is

$$C^n = \varepsilon^{nab} B_{b;a} \quad .$$

The major feature of $B_{b;a}$ -- as discussed in Section 7 (v) -- is that it is a rank-2 tensor if $\mathbf{B}$ is a tensorial vector. The weight addition rule of section (b) 3 can then be applied to $\varepsilon^{nab} B_{b;a}$. Since ε^{nab} has weight -1 and $B_{b;a}$ has weight 0, the conclusion is that C^n is a vector density of weight -1.

Thus, $\mathbf{C} = \text{curl } \mathbf{B}$ is a vector density of weight -1. If the underlying transformation F is a rotation, $\mathbf{C}$ becomes an ordinary vector as per section (a) 7.

(i) Tensor $\mathcal{E}$ as a weight 0 version of ε : three conventions

1. 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

$$\begin{aligned} \varepsilon_{abc\dots} &= g \varepsilon^{abc\dots} & \varepsilon^{123\dots} &= +1 & \varepsilon_{123\dots} &= g \\ \varepsilon'_{abc\dots} &= g' \varepsilon'^{abc\dots} & \varepsilon'^{123\dots} &= +1 & \varepsilon'_{123\dots} &= g' \\ \varepsilon'_{abc\dots} &= J^2 \varepsilon_{abc\dots} = (g'/g) \varepsilon_{abc\dots} & & & \varepsilon & \text{ is a rank-N tensor of weight } W = -1 \end{aligned}$$

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))

$$\begin{aligned} \mathcal{E}_{abc\dots} &\equiv |g|^{-1/2} \varepsilon_{abc\dots} & \Rightarrow & \mathcal{E}_{123\dots} = |g|^{-1/2} g = |g|^{-1/2} s |g| = s |g|^{1/2} \\ \mathcal{E}'_{abc\dots} &\equiv |g'|^{-1/2} \varepsilon'_{abc\dots} & \Rightarrow & \mathcal{E}'_{123\dots} = |g'|^{-1/2} g' = |g'|^{-1/2} s |g'| = s |g'|^{1/2} \end{aligned}$$

It was shown in Section 5 (k) that $g' = J^2 g$ so that g transforms as a scalar density of weight -2. Since the sign of g and g' are the same, it follows that $(sg) = |g|$ is also a scalar density of weight -2, and then the quantity $|g|^{-1/2}$ transforms as a scalar density of weight +1, since $|g'|^{-1/2} = J^{-1} |g|^{-1/2}$. Looking at the equation $\mathcal{E}_{abc\dots} \equiv |g|^{-1/2} \varepsilon_{abc\dots}$ above, and using the weight summation rule of section (b) 3, one

concludes at that $\mathcal{E}_{abc\dots}$ transforms as a tensor of weight $(+1) + (-1) = 0$, and so $\mathcal{E}_{abc\dots}$ is an ordinary tensor. This is the motivation of the above definitions. It was shown at the end of Section 7 (u) that a tensor density equation with matching weights is "covariant", so one is not surprised to see the second line above being the same as the first line but everything is primed ($s = s'$).

Raising indices on both sides gives

$$\begin{aligned}\mathcal{E}^{abc\dots} &\equiv |g|^{-1/2} \varepsilon^{abc\dots} \Rightarrow & \mathcal{E}^{123\dots} &= |g|^{-1/2} \\ \mathcal{E}'^{abc\dots} &\equiv |g'|^{-1/2} \varepsilon'^{abc\dots} \Rightarrow & \mathcal{E}'^{123\dots} &= |g'|^{-1/2}\end{aligned}$$

To compare $\mathcal{E}_{abc\dots}$ and $\mathcal{E}^{abc\dots}$,

$$\begin{aligned}\mathcal{E}_{abc\dots} &= |g|^{-1/2} \varepsilon_{abc\dots} = |g|^{-1/2} g \varepsilon^{abc\dots} = s |g|^{-1/2} |g| \varepsilon^{abc\dots} = s |g|^{1/2} \varepsilon^{abc\dots} \\ \mathcal{E}'_{abc\dots} &= |g'|^{-1/2} \varepsilon'_{abc\dots}\end{aligned}$$

so that,

$$\mathcal{E}_{abc\dots} = s |g| \mathcal{E}^{abc\dots} = g \mathcal{E}^{abc\dots}$$

Summarizing,

$$\begin{aligned}\mathcal{E}^{123\dots} &= |g|^{-1/2} & \mathcal{E}_{123\dots} &= s |g|^{+1/2} & \mathcal{E}_{abc\dots} &= g \mathcal{E}^{abc\dots} = s |g| \mathcal{E}^{abc\dots} \\ \mathcal{E}'^{123\dots} &= |g'|^{-1/2} & \mathcal{E}'_{123\dots} &= s |g'|^{+1/2} \\ \mathcal{E}^{abc\dots} \mathcal{E}_{ABC\dots} &= |g|^{-1} \varepsilon^{abc\dots} \varepsilon_{ABC\dots} \\ \mathcal{E}'^{abc\dots} \mathcal{E}'_{ABC\dots} &= |g'|^{-1} \varepsilon'^{abc\dots} \varepsilon'_{ABC\dots}\end{aligned}$$

Since $|g|^{-1}$ is a scalar density of weight +2 and each ε has weight -1 and each $\mathcal{E}$ has weight 0, one is happy to see the weights balance of the two sides of this pair of covariant equations.

2. Equations in the "added sign s convention"

The previous section shows how things work out using the "Weinberg convention" noted at the start of section (e). Here is the previous section redone in the "added sign s convention":

In section (e) it was established that ($g = \det(g_{ij})$)

$$\begin{aligned}\varepsilon_{abc\dots} &= s g \varepsilon^{abc\dots} & \varepsilon^{123\dots} &= +1 & \varepsilon_{123\dots} &= s g = |g| & // g \rightarrow s g \\ \varepsilon'_{abc\dots} &= s g' \varepsilon'^{abc\dots} & \varepsilon'^{123\dots} &= +1 & \varepsilon'_{123\dots} &= s g' = |g'| & // g' \rightarrow s g' \\ \varepsilon'_{abc\dots} &= |J|^2 \varepsilon_{abc\dots} = (g'/g) \varepsilon_{abc\dots} & & & \varepsilon & \text{ is rank-N tensor of weight } W = -1 & // \text{ same}\end{aligned}$$

Consider the following new objects defined by ($sg = |g|$, $s = \text{sign}(g) = \text{sign}(g')$) as in Section 5 (k))

$$\begin{aligned}\mathcal{E}_{abc\dots} &\equiv |g|^{-1/2} \varepsilon_{abc\dots} \Rightarrow & \mathcal{E}_{123\dots} &= |g|^{-1/2} |g| = |g|^{1/2} \\ \mathcal{E}'_{abc\dots} &\equiv |g'|^{-1/2} \varepsilon'_{abc\dots} \Rightarrow & \mathcal{E}'_{123\dots} &= |g'|^{-1/2} |g'| = |g'|^{1/2}\end{aligned}$$

But the same argument given above, $\mathcal{E}_{abc\dots}$ is an ordinary covariant tensor (ie, weight = 0). However, the indices cannot be raised by g^{ij} . In this convention then one must make independent definitions of the contravariant components as follows,

$$\begin{aligned}\mathcal{E}^{abc\dots} &\equiv |g|^{-1/2} \varepsilon^{abc\dots} \Rightarrow & \mathcal{E}^{123\dots} &= |g|^{-1/2} \\ \mathcal{E}'^{abc\dots} &\equiv |g'|^{-1/2} \varepsilon'^{abc\dots} \Rightarrow & \mathcal{E}'^{123\dots} &= |g'|^{-1/2}\end{aligned}$$

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.

3. 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,

$$\begin{aligned}\mathcal{E}^{123\dots} &= \sigma |g|^{-1/2} & \mathcal{E}_{123\dots} &= \sigma |g|^{+1/2} & \mathcal{E}_{abc\dots} &= |g| \mathcal{E}^{abc\dots} \\ \mathcal{E}'^{123\dots} &= \sigma |g'|^{-1/2} & \mathcal{E}'_{123\dots} &= \sigma |g'|^{+1/2} & \mathcal{E}'_{abc\dots} &= |g'| \mathcal{E}'^{abc\dots}\end{aligned}$$

$$\begin{aligned}\epsilon^{abc\dots} \epsilon_{ABC\dots} &= |g|^{-1} \epsilon^{abc\dots} \epsilon_{ABC\dots} \\ \epsilon'^{abc\dots} \epsilon'_{ABC\dots} &= |g'|^{-1} \epsilon'^{abc\dots} \epsilon'_{ABC\dots}\end{aligned}$$

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 , $\epsilon\epsilon$ and contracted $\epsilon\epsilon$ 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 $123\dots N$ 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 $K \epsilon_{abc\dots x}$:

$$\sum_P p P_2(M_{a1}M_{b2} M_{c3}\dots M_{xN}) = K \epsilon_{abc\dots x}$$

Setting $abc\dots x$ to $123\dots N$, 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

$$\sum_P p P_2(M_{11}M_{22} M_{33}\dots M_{NN}) = \sum_{abc\dots x} \epsilon_{abc\dots x} M_{1a}M_{2b} M_{3c}\dots M_{Nx}$$

because $p = \epsilon_{abc\dots x}$ 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

$$\sum_P p P_2(M_{a1}M_{b2} M_{c3}\dots M_{xN}) = \det(M) \epsilon_{abc\dots x}$$

Consider now the following matrix where $abc\dots x$ is some permutation of $123\dots x$,

$$\text{where } M^{(abc\dots)} = \begin{matrix} M_{a1} & M_{a2} & M_{a3} & \dots & M_{aN} \\ M_{b1} & M_{b2} & M_{b3} & \dots & M_{bN} \\ M_{c1} & M_{c2} & M_{c3} & \dots & M_{cN} \\ \dots & & & & \\ M_{x1} & M_{x2} & M_{x3} & \dots & M_{xN} \end{matrix}$$

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 $\epsilon_{abc\dots x}$. Therefore

$$\det(M^{(abc\dots)}) = \epsilon_{abc\dots x} \det(M)$$

and therefore

$$\sum_p P_2(M_{a1}M_{b2}M_{c3}\dots M_{xN}) = \det(M^{(abc\dots)}) = \det(M) \epsilon_{abc\dots x}$$

The permutation sum is thus just the determinant of matrix $M^{(abc\dots)}$. 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

$$\sum_p P_2(\delta_{a,1}\delta_{b,2}\delta_{c,3}\dots\delta_{x,N}) = \det[\delta^{(abc\dots)}] = \epsilon_{abc\dots x}$$

Thus is obtained the famous representation of $\epsilon_{abc\dots x}$ as a certain determinant of Kronecker deltas,

$$\epsilon_{abc\dots x} = \det[\delta^{(abc\dots)}]$$

$$\text{where } \delta^{(abc\dots)} = \begin{matrix} \delta_{a,1} & \delta_{a,2} & \delta_{a,3} & \dots & \delta_{a,N} & = \mathbf{R}_a \\ \delta_{b,1} & \delta_{b,2} & \delta_{b,3} & \dots & \delta_{b,N} & = \mathbf{R}_b \\ \delta_{c,1} & \delta_{c,2} & \delta_{c,3} & \dots & \delta_{c,N} & = \mathbf{R}_c \\ \dots & & & & & \\ \delta_{x,1} & \delta_{x,2} & \delta_{x,3} & \dots & \delta_{x,N} & = \mathbf{R}_x \end{matrix}$$

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

$$\varepsilon_{\mathbf{abc} \dots \mathbf{x}} = \begin{vmatrix} \delta_{\mathbf{a},1} & \delta_{\mathbf{a},2} & \delta_{\mathbf{a},3} & \dots & \delta_{\mathbf{a},N} \\ \delta_{\mathbf{b},1} & \delta_{\mathbf{b},2} & \delta_{\mathbf{b},3} & \dots & \delta_{\mathbf{b},N} \\ \delta_{\mathbf{c},1} & \delta_{\mathbf{c},2} & \delta_{\mathbf{c},3} & \dots & \delta_{\mathbf{c},N} \\ \dots & \dots & \dots & \dots & \dots \\ \delta_{\mathbf{x},1} & \delta_{\mathbf{x},2} & \delta_{\mathbf{x},3} & \dots & \delta_{\mathbf{x},N} \end{vmatrix} = |\delta^{(\mathbf{abc} \dots)}|$$

which is the same as

$$\varepsilon_{\mathbf{abc} \dots \mathbf{x}} = \sum_{\mathbf{p}} \mathbf{p} \, P_2(\delta_{\mathbf{a},1} \delta_{\mathbf{b},2} \delta_{\mathbf{c},3} \dots \delta_{\mathbf{x},N}) .$$

3. Outer product of two ε tensors.

Consider now

$$\varepsilon_{\mathbf{abc} \dots \mathbf{x}} = \det \begin{pmatrix} \mathbf{R}_{\mathbf{a}} \\ \mathbf{R}_{\mathbf{b}} \\ \dots \\ \mathbf{R}_{\mathbf{x}} \end{pmatrix} \quad \text{and} \quad \varepsilon_{\mathbf{a}'\mathbf{b}'\mathbf{c}' \dots \mathbf{x}'} = \det \begin{pmatrix} \mathbf{R}_{\mathbf{a}'} \\ \mathbf{R}_{\mathbf{b}'} \\ \dots \\ \mathbf{R}_{\mathbf{x}'} \end{pmatrix}$$

Then

$$\begin{aligned} \varepsilon_{\mathbf{abc} \dots \mathbf{x}} \varepsilon_{\mathbf{a}'\mathbf{b}'\mathbf{c}' \dots \mathbf{x}'} &= \det \begin{pmatrix} \mathbf{R}_{\mathbf{a}} \\ \mathbf{R}_{\mathbf{b}} \\ \dots \\ \mathbf{R}_{\mathbf{x}} \end{pmatrix} \det \begin{pmatrix} \mathbf{R}_{\mathbf{a}'} \\ \mathbf{R}_{\mathbf{b}'} \\ \dots \\ \mathbf{R}_{\mathbf{x}'} \end{pmatrix} = \det \begin{pmatrix} \mathbf{R}_{\mathbf{a}} \\ \mathbf{R}_{\mathbf{b}} \\ \dots \\ \mathbf{R}_{\mathbf{x}} \end{pmatrix} \det (\mathbf{R}_{\mathbf{a}'}, \mathbf{R}_{\mathbf{b}'}, \dots, \mathbf{R}_{\mathbf{x}'}) \\ &= \det \left\{ \begin{pmatrix} \mathbf{R}_{\mathbf{a}} \\ \mathbf{R}_{\mathbf{b}} \\ \dots \\ \mathbf{R}_{\mathbf{x}} \end{pmatrix} (\mathbf{R}_{\mathbf{a}'}, \mathbf{R}_{\mathbf{b}'}, \dots, \mathbf{R}_{\mathbf{x}'}) \right\} \end{aligned}$$

which is the determinant of this matrix

$$\begin{array}{cccc} \mathbf{R}_{\mathbf{a}} \bullet \mathbf{R}_{\mathbf{a}'} & \mathbf{R}_{\mathbf{a}} \bullet \mathbf{R}_{\mathbf{b}'} & \mathbf{R}_{\mathbf{a}} \bullet \mathbf{R}_{\mathbf{c}'} & \dots \mathbf{R}_{\mathbf{a}} \bullet \mathbf{R}_{\mathbf{x}'} \\ \mathbf{R}_{\mathbf{b}} \bullet \mathbf{R}_{\mathbf{a}'} & \mathbf{R}_{\mathbf{b}} \bullet \mathbf{R}_{\mathbf{b}'} & \mathbf{R}_{\mathbf{b}} \bullet \mathbf{R}_{\mathbf{c}'} & \dots \mathbf{R}_{\mathbf{b}} \bullet \mathbf{R}_{\mathbf{x}'} \\ \mathbf{R}_{\mathbf{c}} \bullet \mathbf{R}_{\mathbf{a}'} & \mathbf{R}_{\mathbf{c}} \bullet \mathbf{R}_{\mathbf{b}'} & \mathbf{R}_{\mathbf{c}} \bullet \mathbf{R}_{\mathbf{c}'} & \dots \mathbf{R}_{\mathbf{c}} \bullet \mathbf{R}_{\mathbf{x}'} \\ \dots & \dots & \dots & \dots \\ \mathbf{R}_{\mathbf{x}} \bullet \mathbf{R}_{\mathbf{a}'} & \mathbf{R}_{\mathbf{x}} \bullet \mathbf{R}_{\mathbf{b}'} & \mathbf{R}_{\mathbf{x}} \bullet \mathbf{R}_{\mathbf{c}'} & \dots \mathbf{R}_{\mathbf{x}} \bullet \mathbf{R}_{\mathbf{x}'} \end{array}$$

A typical element of this matrix is given by

$$\mathbf{R}_{\mathbf{c}} \bullet \mathbf{R}_{\mathbf{b}'} = (\mathbf{R}_{\mathbf{c}})_i (\mathbf{R}_{\mathbf{b}'})_i = \delta_{\mathbf{c},i} \delta_{\mathbf{b}',i} = \delta_{\mathbf{c},\mathbf{b}'}$$

so that matrix can be written as

$$\delta_{\mathbf{a},\mathbf{a}'} \quad \delta_{\mathbf{a},\mathbf{b}'} \quad \delta_{\mathbf{a},\mathbf{c}'} \quad \dots \delta_{\mathbf{a},\mathbf{x}'}$$

$$\begin{array}{ccccccc}
\delta_{b,a'} & \delta_{b,b'} & \delta_{b,c'} & \dots & \delta_{b,x'} \\
\delta_{c,a'} & \delta_{c,b'} & \delta_{c,c'} & \dots & \delta_{c,x'} \\
\dots & & & & \\
\delta_{x,a'} & \delta_{x,b'} & \delta_{x,c'} & \dots & \delta_{x,x'}
\end{array} \equiv \delta^{(abc\dots x; a'b'c'\dots x')}$$

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

$$\begin{aligned}
\sum_a \varepsilon_{abc\dots x} \varepsilon_{ab'c'\dots x'} &= \varepsilon_{Abc\dots x} \varepsilon_{Ab'c'\dots x'} \\
&= \sum_P p P_2(\delta_{A,A} \delta_{b,b'} \delta_{c,c'} \dots \delta_{x,x'}) = \sum_P p P_2(\delta_{b,b'} \delta_{c,c'} \dots \delta_{x,x'}) \quad (*)
\end{aligned}$$

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,

$$\sum_a \varepsilon_{abc} \varepsilon_{ab'c'} = \sum_P p P_2(\delta_{b,b'} \delta_{c,c'}) = \delta_{b,b'} \delta_{c,c'} - \delta_{b,c'} \delta_{c,b'}$$

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,

$$\sum_a \varepsilon_{abc\dots x} \varepsilon_{ab'c'\dots x'} = [\text{cof } \delta^{(abc\dots x; a'b'c'\dots x')}]_{aa}$$

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)

$$\varepsilon_{abc\dots x} \varepsilon_{ab'c'\dots x'} = \begin{vmatrix} \delta_{b,b'} & \delta_{b,c'} & \dots & \delta_{b,x'} \\ \delta_{c,b'} & \delta_{c,c'} & \dots & \delta_{c,x'} \\ \delta_{x,b'} & \delta_{x,c'} & \dots & \delta_{x,x'} \end{vmatrix} = [\text{cof } \delta^{(abc\dots x; a'b'c'\dots x')}]_{aa}$$

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 $\varepsilon\varepsilon$ product:

$$\sum_{a,b} \varepsilon_{abcd\dots x} \varepsilon_{abc'd'\dots x'}$$

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

$$\begin{aligned} \sum_{a,b} \varepsilon_{abcd\dots x} \varepsilon_{abc'd'\dots x'} &= 2! \varepsilon_{ABc\dots x} \varepsilon_{ABc'\dots x'} \\ &= 2! \sum_P p P_2(\delta_{A,A} \delta_{B,B'} \delta_{c,c'} \delta_{d,d'} \dots \delta_{x,x'}) = 2! \sum_P p P_2(\delta_{c,c'} \delta_{d,d'} \dots \delta_{x,x'}) \end{aligned} \quad (*)$$

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,

$$\sum_{a,b} \varepsilon_{abc} \varepsilon_{abc'} = 2! \sum_P p P_2(\delta_{c,c'}) = 2 \delta_{c,c'}$$

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)

$$\varepsilon_{abcd\dots x} \varepsilon_{abc'd'\dots x'} = 2! \begin{vmatrix} \delta_{c,c'} & \dots & \delta_{c,x'} \\ \dots & & \\ \delta_{x,c'} & \dots & \delta_{x,x'} \end{vmatrix} = 2! \{ [\text{cof } \delta^{(abc\dots x; a'b'c'\dots x')}]_{aa}\}_{bb}$$

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)

$$\varepsilon_{abcd\dots x} \varepsilon_{abcd'\dots x'} = 3! \begin{vmatrix} \delta_{a,a'} & \dots & \delta_{a,x'} \\ \dots & & \\ \delta_{x,a'} & \dots & \delta_{x,x'} \end{vmatrix} = 3! \{ \{ [\text{cof } \delta^{(abc\dots x; a'b'c'\dots x')}]_{aa}\}_{bb} \}_{cc}$$

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 ($\varepsilon_{a'b'} \rightarrow s \varepsilon_{a'b'}$)

$$\mathcal{E}^{ab}\mathcal{E}_{a'b'} = |g|^{-1} \varepsilon^{ab}\varepsilon_{a'b'} = s \begin{vmatrix} g^a_{a'} & g^a_{b'} \\ g^b_{a'} & g^b_{b'} \end{vmatrix} = s (g^a_{a'} g^b_{b'} - g^a_{b'} g^b_{a'})$$

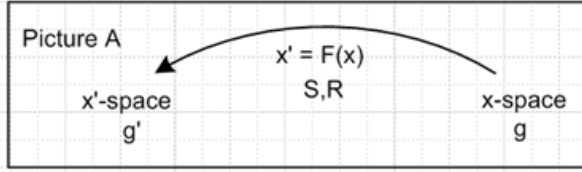
The second equation is *undefined* (when $s=-1$), and the third equation is

$$\mathcal{E}_{ab} \mathcal{E}_{a'b'} = |g|^{-1} \varepsilon_{ab} \varepsilon_{a'b'} = \begin{vmatrix} g_{aa'} & g_{ab'} \\ g_{ba'} & g_{bb'} \end{vmatrix} = g_{aa'} g_{bb'} - g_{ab'} g_{ba'}$$

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,

$$\begin{aligned} (A \otimes B \otimes C \dots)^{abc \dots} &\equiv A^a B^b C^c \dots \\ (A \otimes B \otimes C \dots)^a_{b}{}^c \dots &\equiv A^a B_b C^c \dots \quad \text{etc} \end{aligned}$$

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* ("reciprocal") 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 (unpriming indices)

$$\alpha^{ijk} = A \bullet (\mathbf{b}^i \otimes \mathbf{b}^j \otimes \mathbf{b}^k) = A^{abc} (\mathbf{b}^i \otimes \mathbf{b}^j \otimes \mathbf{b}^k)_{abc} = A^{abc} (b^i)_a (b^j)_b (b^k)_c . \quad (*)$$

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,

$$A = \sum_{ijk} A'^{ijk} (e_i \otimes e_j \otimes e_k) \quad .$$

Expansions like the above are the generalizations to tensors of any rank of these vector expansions stated in Section 7 (s),

$$\begin{aligned}A &= \sum_i \alpha^i \mathbf{b}_i & \alpha^i &= \mathbf{b}^i \bullet A & // \text{arbitrary basis} \\ A &= \sum_i A^i \mathbf{u}_i & & & // \text{axis aligned unit vectors} \\ A &= \sum_i A^i \mathbf{e}_i & & & // \text{tangent base vectors}\end{aligned}$$

where we continue to write rank-1 tensors (vectors) in bold font: **A**.

To summarize, here is the general rank-n tensor expansion for an arbitrary basis, and then for the two specific bases just discussed:

$$\begin{aligned}A &= \sum_{ijk\dots} \alpha^{ijk\dots} (\mathbf{b}_i \otimes \mathbf{b}_j \otimes \mathbf{b}_k \dots) & \alpha^{ijk\dots} &= A^{abc\dots} (b^i)_a (b^j)_b (b^k)_c \dots \\ A &= \sum_{ijk\dots} A^{ijk\dots} (\mathbf{u}_i \otimes \mathbf{u}_j \otimes \mathbf{u}_k \dots) & A^{ijk\dots} &= \text{contravariant components of } A \text{ in } x\text{-space} \\ A &= \sum_{ijk\dots} A'^{ijk\dots} (\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \dots) & A'^{ijk\dots} &= \text{contravariant components of } A \text{ in } x'\text{-space}\end{aligned}$$

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 $g'_{ij} = \mathbf{e}_i \bullet \mathbf{e}_j = \delta_{ij}$, 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 $g'_{ab} = h'^2_a \delta_{a,b}$. One then has $\hat{\mathbf{e}}_i \bullet \hat{\mathbf{e}}_j = \delta_{i,j}$ since

$$\hat{\mathbf{e}}_i \bullet \hat{\mathbf{e}}_j = \mathbf{e}_i \bullet \mathbf{e}_j / (h'_i h'_j) = g'_{ij} / (h'_i h'_j) = h'^2_{ij} / (h'_i h'_j) = \delta_{i,j} \quad .$$

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. Expansions on the unit versions of the tangent base vectors $\hat{\mathbf{e}}_i$ are discussed more in section (h) below.

Tensor density. If A is a tensor density of weight W, the general rule is to make this replacement:

$$A^{ijk} \dots \rightarrow J^W A^{ijk} \dots$$

so the third general expansion above would be written

$$A = J^W \Sigma_{ijk} \dots A^{ijk} \dots (\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \dots) \quad A^{ijk} \dots = \text{contravariant components of A in } x'\text{-space}$$

As justification for this rule, start with a regular tensor transformation for A,

$$A^{ijk} \dots = R^i_i R^j_j R^k_k \dots A^{i'j'k'} \dots$$

The rule gives

$$J^W A^{ijk} \dots = R^i_i R^j_j R^k_k \dots A^{i'j'k'} \dots$$

or

$$A^{ijk} \dots = J^{-W} R^i_i R^j_j R^k_k \dots A^{i'j'k'} \dots$$

which is the correct form for the transformation of a tensor density of weight W (Appendix D).

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:

$$A = \Sigma_{ijk} \dots \alpha^{ijk} \dots (\mathbf{b}_i \otimes \mathbf{b}_j \otimes \mathbf{b}_k \dots) \quad \alpha^{ijk} \dots = A^{abc} \dots (b^i)_a (b^j)_b (b^k)_c \dots$$

$$A = \Sigma_{ijk} \dots A^{ijk} \dots (\mathbf{u}_i \otimes \mathbf{u}_j \otimes \mathbf{u}_k \dots) \quad A^{ijk} \dots = \text{components of A in } x\text{-space}$$

$$A = \Sigma_{ijk} \dots \mathcal{A}^{ijk} \dots (\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \dots) \quad \mathcal{A}^{ijk} \dots = R^i_a R^j_b R^k_c A^{abc}$$

Since A is not a tensor, in this case $\mathcal{A}^{ijk} \dots$ are not the contravariant components of tensor A in x' -space relative to the transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$.

(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

$$\begin{aligned} A &= \sum_{i j k \dots} \alpha^{i j k \dots} (\mathbf{b}_i \otimes \mathbf{b}_j \otimes \mathbf{b}_k \dots) \\ &= \sum_{i j k \dots} \alpha^{i j k \dots} (\mathbf{b}_i \mathbf{b}_j \mathbf{b}_k \dots) \\ &= \sum_{i j k \dots} \alpha^{i j k \dots} (\mathbf{i} \mathbf{j} \mathbf{k} \dots) \end{aligned}$$

where

$$\begin{aligned} \alpha^{i j k \dots} &= A \bullet (\mathbf{b}^i \otimes \mathbf{b}^j \otimes \mathbf{b}^k \dots) \\ &= A \bullet (\mathbf{b}^i \mathbf{b}^j \mathbf{b}^k \dots) \\ &= A \bullet (\mathbf{i}^d \mathbf{j}^d \mathbf{k}^d \dots) \\ &= A^{abc \dots} (b^i)_a (b^j)_b (b^k)_c \dots \end{aligned}$$

where we have just made up a notation $\mathbf{i}^d$ to stand for the dual vector $\mathbf{b}^i$.

One can find further discussion of polyadic notation for example in Backus.

(d) Dyadic Products

When two vectors $\mathbf{A}$ and $\mathbf{B}$ are combined in polyadic notation, the result is called a dyadic product ($\mathbf{AB}$) [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

$$A = \sum_{i j} \alpha^{ij} (\mathbf{b}_i \mathbf{b}_j) \quad \alpha^{ij} = A^{ab} (b^i)_a (b^j)_b = A^{ab} (\mathbf{b}^i \mathbf{b}^j)_{ab} .$$

Notice that the dyadic product $(\mathbf{AB})$ 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 $\mathbf{A}$ and $\mathbf{B}$ (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* $(\mathbf{AB})$ in the $\mathbf{u}_n$ basis so $(\mathbf{AB})^{ij} = \langle \mathbf{u}^i | (\mathbf{AB}) | \mathbf{u}^j \rangle$, 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 $\mathbf{V}$, certainly $V^i = (V^T)^i$, meaning the object in the i^{th} row of $\mathbf{V}$ is the same as the object in the i^{th} column 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 $\mathbf{ab}$ 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 $[\mathbf{ab}]$ is needed to indicate that $\mathbf{ab}$ 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

$$(\mathbf{ab})_{ij} = (\mathbf{ab}^T)_{ij} = a_i b^T_j = a_i b_j .$$

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 dyadic 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 (same symbol for matrix), and if $\mathbf{c}$ and $\mathbf{d}$ are vectors, then one defines:

$$\begin{aligned} A \bullet \mathbf{c} &\equiv A\mathbf{c} = \text{a column vector} &\Rightarrow (A \bullet \mathbf{c})^i &= (A\mathbf{c})^i = A^{ij}c_j &\Rightarrow A \bullet \mathbf{c} &= \sum_{ij} A^{ij}c_j \mathbf{u}_i \\ \mathbf{c} \bullet A &\equiv \mathbf{c}^T A = \text{a row vector} &\Rightarrow (\mathbf{c} \bullet A)^i &= (\mathbf{c}^T A)^i = c_j A^{ji} &\Rightarrow \mathbf{c} \bullet A &= \sum_{ij} c_j A^{ji} \mathbf{u}_i \\ \mathbf{d} \bullet A \bullet \mathbf{c} &= \mathbf{d}^T A \mathbf{c} = \text{a number} &&= d_i A^{ij} c_j. \end{aligned}$$

It then follows that, for the particular dyadic $A = \mathbf{ab}$,

$$\begin{aligned} (\mathbf{ab}) \bullet \mathbf{c} &\equiv (\mathbf{ab}) \mathbf{c} = (\mathbf{ab}^T) \mathbf{c} = \mathbf{a}(\mathbf{b}^T \mathbf{c}) = \mathbf{a}(\mathbf{b} \bullet \mathbf{c}) = (\mathbf{b} \bullet \mathbf{c}) \mathbf{a} = \text{a column vector} \\ \mathbf{c} \bullet (\mathbf{ab}) &\equiv \mathbf{c}^T (\mathbf{ab}) = \mathbf{c}^T (\mathbf{ab}^T) = (\mathbf{c}^T \mathbf{a}) \mathbf{b}^T = (\mathbf{c} \bullet \mathbf{a}) \mathbf{b}^T = \text{a row vector} \\ \mathbf{d} \bullet (\mathbf{ab}) \bullet \mathbf{c} &= \mathbf{d}^T (\mathbf{ab}) \mathbf{c} = \mathbf{d}^T \mathbf{ab}^T \mathbf{c} = (\mathbf{d}^T \mathbf{a})(\mathbf{b}^T \mathbf{c}) = (\mathbf{d} \bullet \mathbf{a})(\mathbf{b} \bullet \mathbf{c}) = \text{a number}. \end{aligned}$$

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

$$A \bullet B \equiv AB$$

Regarding this small size dot $\bullet$: (1) from a matrix algebra point of view, it is completely superfluous except in the case $\mathbf{c} \bullet A \equiv \mathbf{c}^T A$; (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 5 (i); (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 $\mathbf{b}_i$,

$$\mathbf{V} = \sum_i [\mathbf{V}^{(b)}]^i \mathbf{b}_i \quad // [\mathbf{V}^{(b)}]^i \text{ are the coefficients of this expansion .}$$

Moreover, one can regard a rank-2 tensor A as an "operator" in this Hilbert space,

$$A = \sum_{i,j} [A^{(b)}]^{ij} \mathbf{b}_i \mathbf{b}_j^T$$

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

$$[A^{(b)}]^{nm} = (\mathbf{b}^T)^n A \mathbf{b}^m$$

Here, one regards A as an operator in the x Hilbert space, whereas $[A^{(b)}]^{nm}$ is a "matrix" which is *associated* with the operator A in the particular $\mathbf{b}^n$ basis. 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 $(\mathbf{b}^T)^n A \mathbf{b}^m$ has a meaning distinct from our interpretation of it in terms of the matrix combination of three objects. In the matrix interpretation, one writes $(\mathbf{b}^T)^n A \mathbf{b}^m = [(\mathbf{b}^T)^n]_i A^{ij} [\mathbf{b}^m]_j = [\mathbf{b}^T]_i A^{ij} [\mathbf{b}^m]_j$ 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 $\mathbf{b}_i \rightarrow |b_i\rangle$, $\mathbf{b}_i^T \rightarrow \langle b_i|$, so that the above equations become

$$\begin{aligned} |V\rangle &= \sum_i [\mathbf{V}^{(b)}]^i |b_i\rangle & \langle b_j | b^i \rangle &= \delta_j^i & \text{orthogonality of the basis} \\ [\mathbf{V}^{(b)}]^i &= \langle b^i | V \rangle & 1 &= \sum_i |b_i\rangle \langle b^i| & \text{completeness of the basis} \\ A &= \sum_{i,j} [A^{(b)}]^{ij} |b_i\rangle \langle b_j| & \langle U | V \rangle &= \mathbf{U} \bullet \mathbf{V} = \text{scalar product} \\ [A^{(b)}]^{ij} &= \langle b^i | A | b^j \rangle \end{aligned}$$

In this notation, the N $|b_i\rangle$ are a set of basis vectors which span an N-dimensional real Hilbert Space, while $\langle b_i|$ span the so-called adjoint (or transpose in our case) Hilbert Space. One then refers to $[A^{(b)}]^{ij}$ as the "matrix element of the operator A in the $\mathbf{b}^i$ basis ". In general, $|b_i\rangle$ and $|b^i\rangle$ are different vectors. In this notation, based on what was presented earlier, one can write,

$A^{nm} = \langle u^n | A | u^m \rangle$ = the x-space components of tensor A (basis u_n) raise/lower with g

$A'^{nm} = \langle e^n | A | e^m \rangle$ = the x'-space components of tensor A (basis e_n) raise/lower with g'

$[A^{(b)}]^{nm} = \langle b^n | A | b^m \rangle$ = the matrix of A in the b_n basis raise/lower with w

In the first of these three lines, one can raise and lower indices with g^{ab} and g_{ab} on both sides of the equation. On the second line this can be done with g'^{ab} and g'_{ab} . It was shown in Section 6 (b) that $b^n = w^{nm} b'_m$ and conversely $b'_n = w_{nm} b^n$ where w_{nm} is the metric tensor g' one would get for some underlying transformation F_b which causes b_n to be its tangent base vectors. Thus, on the third line above we can raise and lower indices on each side with w^{ab} and w_{ab} where $w_{nm} = b_n \bullet b_m$.

Notice in the last three equations that the operator A between the vertical bars is *the exact same operator* in each case. The matrices are different not because the operator has changed, but because the basis vectors are different.

$[A^{(b)}]^{nm}$ are the components of a rank-2 tensor in only two cases -- those shown in the first pair of equations above. In the first case A^{nm} are components of a tensor in x-space, and in the second case the A'^{nm} are components of a tensor in x'-space.

In a consistent notation one might write $A^{nm} = [A^{(u)}]^{nm}$ and $A'^{nm} = [A^{(e)}]^{nm}$.

Bases are related by a transformation. Consider again,

$$[A^{(b)}]^{nm} = \langle b^n | A | b^m \rangle = (b^n)^T A b^m = [b^n]_i A^{ij} [b^m]_j = \langle b^n | u_i \rangle \langle u^i | A | u^j \rangle \langle u_j | b^m \rangle.$$

We lower index m on both sides (using w_{ab} as noted above) and reverse the j tilt to get

$$[A^{(b)}]^{n}_{\ m} = \langle b^n | A | b_m \rangle = (b^n)^T A b_m = [b^n]_i A^{ij} [b_m]^j = \langle b^n | u_i \rangle \langle u^i | A | u_j \rangle \langle u^j | b_m \rangle.$$

One could then define the following tensor-like object,

$$B^n_i \equiv [b^n]_i.$$

The first index on B is raised and lowered by w, while the second is raised and lowered by g, so this object is a bit like R and S in its non-tensor nature. Lowering n and raising i then gives

$$B^n_i = [b^n]_i = (B^T)^i_n,$$

where we use the notion of the transpose of a tilted matrix described in Section 7 (i) item 8. One then has

$$[A^{(b)}]^{n}_{\ m} = B^n_i A^{ij} (B^T)^j_m.$$

Since all the matrices are tilted the same way and summed indices are contractions, this is one of the "legal" Standard Notation matrix forms and we then write,

$$A^{(b)} = BAB^T \quad \text{or more precisely} \quad [A^{(b)}]_{\text{SN}, \text{dt}} = BAB^T$$

where SN,dt means Standard Notation, down-tilt, as described in Section 7 (i) item 7. The matrix equation $A^{(\mathbf{b})} = BAB^T$ shows that the $[A^{(\mathbf{b})}]^n_m$ are related to the A^i_j by a "congruence transformation" with a matrix $B_n^i = [\mathbf{b}_n]^i$ whose rows are the basis vectors $\mathbf{b}^m$. When $\mathbf{b}_m = \mathbf{u}_m$, matrix B is the identity matrix, and when $\mathbf{b}_m = \mathbf{e}_m$ one has $B_n^i = [\mathbf{e}_n]^i = R_n^i$, so that $B = R$ in this case. It was shown in Section 7 (i) that in standard notation R is real orthogonal, so in fact one has for the $\mathbf{b}_m = \mathbf{e}_m$ basis,

$$A^{(\mathbf{e})} = BAB^T = R A R^T = R A R^{-1} = R A S.$$

Specifically in this case,

$$[A^{(\mathbf{e})}]^n_m = R^n_i A^i_j S^j_m = R^n_i R_m^j A^i_j = A^n_m.$$

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

$$|A \mathbf{c} \rangle = \text{a new Hilbert space vector which results when } A \text{ is applied to } |\mathbf{c}\rangle, A|\mathbf{c}\rangle$$

$$\langle A^T \mathbf{d} | = \text{a new transpose Hilbert space vector which results when } A \text{ is applied to } \langle \mathbf{d}|, \langle \mathbf{d}|A.$$

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 some vector $\langle \mathbf{f}|$ in the transpose Hilbert space. The distinction between A as an abstract operator in the Hilbert space, and the A in (Ac) and $(A^T \mathbf{d}) = (\mathbf{d}^T A)^T$ 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:

$$\begin{array}{llll} A \bullet \mathbf{c} = A \mathbf{c} & \mathbf{d} \bullet A = (A^T \mathbf{d})^T = \mathbf{d}^T A & \mathbf{d} \bullet A \bullet \mathbf{c} = \mathbf{d}^T A \mathbf{c} & A \bullet B \mathbf{c} \\ A | \mathbf{c} \rangle = |A \mathbf{c} \rangle & \langle \mathbf{d} | A = \langle A^T \mathbf{d} | & \langle \mathbf{d} | A | \mathbf{c} \rangle = \langle \mathbf{d} | A \mathbf{c} \rangle & AB | \mathbf{c} \rangle \end{array}$$

In the rightmost column operator B is applied first to $|c\rangle$ to get vector $|(Bc)\rangle$, and then operator A is applied to $|(Bc)\rangle$ to give yet another vector $|(A(Bc))\rangle$. In bra-ket notation the product of two abstract operators is given just as AB , but in dyadic notation it is written $A \cdot B$.

Dyadics as operators. According to the above discussion, one can regard a dyadic $(\mathbf{AB})$, 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 $g \neq 1$)

$$\begin{aligned} T^{nm} &= (\mathbf{AB})^{nm} = \langle \mathbf{u}^n | (\mathbf{AB}) | \mathbf{u}^m \rangle = (\mathbf{u}^n)^T \mathbf{A} \mathbf{B}^T \mathbf{u}^m \\ &= [(\mathbf{u}^n)^T]_a A^a (B^T)_b [\mathbf{u}^m]_b = \delta^n_a A^a B^b \delta^m_b = A^n B^m. \end{aligned}$$

In the generic $\mathbf{b}_n$ basis one has

$$[(\mathbf{AB})^{(b)}]^{nm} = \langle \mathbf{b}^n | (\mathbf{AB}) | \mathbf{b}^m \rangle.$$

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 $A \cdot \mathbf{c} = \sum_{ij} A^{ij} c_j \mathbf{u}_i$ from the start of section (f).)

$$\mathfrak{U} \cdot \mathbf{B} = \sum_{mn} a_m A_{mn} B_n; \quad \mathbf{B} \cdot \mathfrak{U} = \sum_{mn} B_m A_{mn} a_n \quad (1.6.5)$$

$$\begin{aligned} \mathfrak{U} &= a_1 A_{11} a_1 + a_1 A_{12} a_2 + a_1 A_{13} a_3 + a_2 A_{21} a_1 + a_2 A_{22} a_2 + a_2 A_{23} a_3 \\ &\quad + a_3 A_{31} a_1 + a_3 A_{32} a_2 + a_3 A_{33} a_3 \quad (1.6.6) \end{aligned}$$

$$(\mathfrak{U}^{-1}) \cdot \mathfrak{U} = \mathfrak{U} \cdot (\mathfrak{U}^{-1}) = \mathfrak{I}$$

There is, of course, a zero dyadic $\mathfrak{O}$ and a unity dyadic $\mathfrak{I}$ called the *idemfactor*:

$$\mathfrak{O} \cdot \mathbf{F} = \mathbf{0}; \quad \mathfrak{I} \cdot \mathbf{F} = \mathbf{F}; \quad \mathfrak{I} = a_1 a_1 + a_2 a_2 + a_3 a_3$$

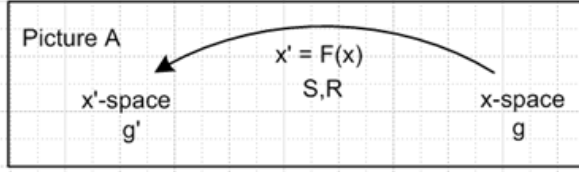
where $\mathbf{F}$ is any vector.

$$\mathfrak{U}_s = A_x \mathbf{ji} + B_x \mathbf{ij} + B_y \mathbf{ik} + B_z \mathbf{ji} + A_y \mathbf{jj} + B_x \mathbf{jk} + B_y \mathbf{ki} + B_z \mathbf{kj} + A_z \mathbf{kk}$$

Notice the impressive name "idemfactor" for the identity operator $1 = \sum_i | a_i \rangle \langle a_i | = \sum_i \mathbf{a}_i \mathbf{a}_i^T = \sum_i \mathbf{a}_i \mathbf{a}_i$.

(h) Expansions of tensors on unit tangent base vectors

We start with the general Picture A (and later specialize to orthogonal coordinates),



In section (b) above it was established that one can expand a tensor A on the tangent base vectors $\mathbf{e}_n$ as

$$A = \sum_{ijk\dots} A'^{ijk\dots} (\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k \dots) \quad A'^{ijk\dots} = \text{contravariant components of A in x'-space}$$

$$A'^{ijk\dots} = R^i{}_j R^j{}_k R^k{}_{k'} \dots A'^{i'j'k'\dots}$$

Since $\mathbf{e}_n = h'_n \hat{\mathbf{e}}_n$, this same expansion for tensor A can be written

$$A = \sum_{ijk\dots} h'_i h'_j h'_k \dots A'^{ijk\dots} (\hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j \otimes \hat{\mathbf{e}}_k \dots)$$

$$= \sum_{ijk\dots} [A^{(\hat{\mathbf{e}})}]^{ijk\dots} (\hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j \otimes \hat{\mathbf{e}}_k \dots)$$

where the unit-vector expansion coefficients are given by

$$\begin{aligned} [A^{(\hat{\mathbf{e}})}]^{ijk\dots} &= h'_i h'_j h'_k \dots A'^{ijk\dots} \\ &= h'_i h'_j h'_k \dots R^i{}_j R^j{}_k R^k{}_{k'} \dots A'^{i'j'k'\dots} \\ &= (h'_i R^i{}_j) (h'_j R^j{}_k) (h'_k R^k{}_{k'}) \dots A'^{i'j'k'\dots} \end{aligned}$$

Coefficient notation. In a curvilinear coordinates application of these expansions, the expansion coefficients are usually written in the following manner,

$$[A^{(\hat{\mathbf{e}})}]^{ijk\dots} = A_{\mathbf{x}'_i \mathbf{x}'_j \mathbf{x}'_k \dots}$$

where the $\mathbf{x}'_n$ are the names of the coordinates. For example, for a rank-4 tensor in spherical coordinates with coordinates $\mathbf{x}'_1 = r$, $\mathbf{x}'_2 = \theta$ and $\mathbf{x}'_3 = \phi$ one might write

$$[A^{(\hat{\mathbf{e}})}]^{2213} = A_{\theta\theta r\phi}$$

Since $[A^{(\hat{\mathbf{e}})}]^{ijk\dots}$ is not a tensor (with respect to F), there is no particular reason to put the indices "up" and for that reason they are usually written down, as in $A_{\theta\theta r\phi}$.

Matrices M and N. It is convenient now to define

$$M^a_b \equiv h'_a R^a_b$$

so that then

$$[A^{(\hat{e})}]^{ijk\dots} = M^i_i, M^j_j, M^k_k, \dots, A^{i'j'k'\dots}$$

Defining N^a_b to be the inverse of M^a_b , one has

$$N^a_b \equiv h'^{-1}_b S^a_b = h'^{-1}_b R^a_b$$

so

$$A^{ijk\dots} = N^i_i, N^j_j, N^k_k, \dots, [A^{(\hat{e})}]^{i'j'k'\dots}$$

To verify that this N is the correct inverse of M,

$$M^a_k N^k_c = (h'_a R^a_k)(h'^{-1}_c R^k_c) = (h'_a / h'_c) R^a_k R^k_c = (h'_a / h'_c) \delta^a_c = \delta^a_c$$

making use of the orthogonality rule of Section 7 (r), $R^a_k R^k_c = \delta^a_c$. M and N can be written in terms of the tangent base vectors as follows:

$$M^n_i \equiv h'_n R^n_i = h'_n (e^n)_i$$

$$N^i_n = h'^{-1}_n R^n_i = h'^{-1}_n (e^n)_i = (\hat{e}_n)^i$$

which says that $N^i_n = \{ \hat{e}_1, \hat{e}_2, \dots \}$ -- the columns of N^i_n are the unit tangent base vectors.

Rank-1 tensors. For a vector, the above coefficient relation is written

$$[A^{(\hat{e})}]^i = M^i_j A^j \quad \text{or} \quad A^{(\hat{e})} = M A \quad \text{and} \quad A = N A^{(\hat{e})}$$

where A has these two familiar expansions,

$$A = A^n u_n = [A^{(\hat{e})}]^n \hat{e}_n \quad [A^{(\hat{e})}]^n = A_{x'_n} \quad \text{for example } [A^{(\hat{e})}]^1 = A_x$$

Rank-2 tensors. Here the coefficient relation is

$$[A^{(\hat{e})}]^{ij} = M^i_i, M^j_j, A^{i'j'} = M^i_i, A^{i'j'} M^j_j,$$

which can be written

$$[A^{(\hat{e})}]^{nm} = M^n_i A^{ij} M^m_j = M^n_i A^{ij} (M^T)^j_m \quad M^n_i = h'_n (e^n)_i$$

Defining $\mathbf{b}^n = h'_n \mathbf{e}^n$, then $[\mathbf{b}^n]_i = h'_n (\mathbf{e}^n)_i = M^n_i = B^n_i$ of section (g). Meanwhile, from Section 6 (b) we know that $w^{nk} = \mathbf{b}^n \bullet \mathbf{b}^k = h'_n h'_k (\mathbf{e}^n \bullet \mathbf{e}^k)$, and then $w_{nk} = (w^{-1})^{nk}$. In any event, whatever w_{nk} is, it is the object which can lower the n or m indices on both sides of the above equation. Lowering just the m index and then reversing the j tilt gives

$$[A^{(\hat{e})}]^n_m = M^n_i A^{ij} M_{mj} = M^n_i A^{ij} (M^T)_{jm} = M^n_i A^{ij} (M^T)^j_m$$

and we replicate the section (g) result with $B = M$:

$$A^{(\hat{e})} = M A M^T \quad // \quad [A^{(\hat{e})}]_{\text{SN}, \text{dt}} = M A M^T$$

Matrix H and x''-space. In the discussion above one has x-space with basis vectors $\mathbf{u}_n$ and x'-space with basis vectors $\mathbf{e}_n$ (the tangent base vectors). It is useful then to define x''-space as the space whose basis vectors are the $\hat{\mathbf{e}}_n$ unit vectors, which are generally not orthogonal. The relation $M^a_b \equiv h'_a R^a_b$ given above can be written in down-tilt form as

$$M = HR \quad \text{where} \quad H^i_j \equiv \text{diag}(h'_1, h'_2, \dots).$$

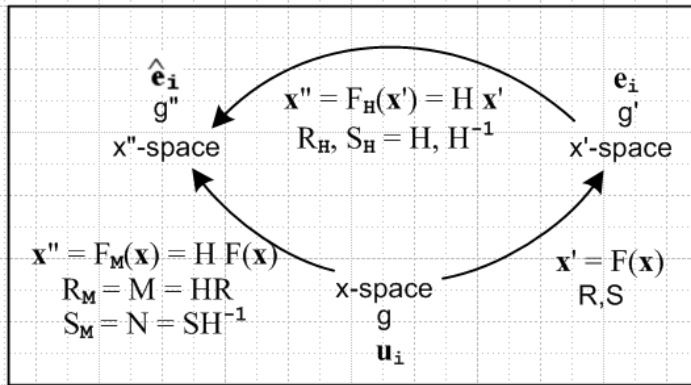
It then follow that

$$N = M^{-1} = R^{-1}H^{-1} = SH^{-1}$$

Finally, note that

$$\mathbf{x} = x^i \mathbf{u}_i = x^i \mathbf{e}_i = x^i (h'_i \hat{\mathbf{e}}_i) = x''^i \hat{\mathbf{e}}_i \Rightarrow x''^i = h'_i x^i \quad \text{or} \quad \mathbf{x}'' = H \mathbf{x}'$$

which shows that the transformation from x'-space to x''-space is linear with matrix H. We can now show all three spaces in the same picture as follows:



This picture shows that the transformation directly from x-space to x''-space is

$$\mathbf{F}_M(\mathbf{x}) = H F(\mathbf{x}) .$$

Here $\mathbf{F}(\mathbf{x})$ is a (generally) non-linear transformation assumed to connect x' -space to x -space. This is then concatenated with linear transformation H to get non-linear transformation $\mathbf{F}_M(\mathbf{x})$.

We can now write $A^{(\hat{\mathbf{e}})}$ as A'' and restate equations above as

$$A = \Sigma_{ijk\dots} A''^{ijk\dots} (\hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j \otimes \hat{\mathbf{e}}_k \dots) \quad // \text{rank-}n \text{ tensor expanded on } \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j \otimes \hat{\mathbf{e}}_k \dots$$

$$A''^{ijk\dots} = M^i_i, M^j_j, M^k_k, \dots A'^{i'j'k'\dots} \quad // \text{rank-}n \text{ tensor transformation}$$

$$A'^{ijk\dots} = N^i_i, N^j_j, N^k_k, \dots A''^{i'j'k'\dots} \quad // \text{inverse of the above}$$

$$A''^i = M^i_j A^j \quad // \text{rank-1 tensor}$$

$$A''^{ij} = M^i_i, M^j_j, A'^{i'j'} \quad // \text{rank-2 tensor}$$

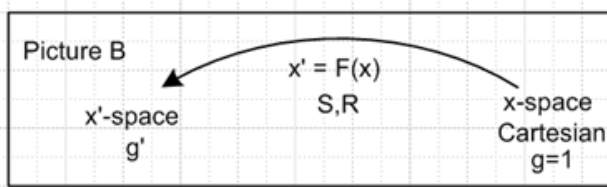
$$A'' = M A M^T \quad // \text{rank-2 tensor, matrix notation}$$

The word "tensor" suddenly has a new meaning in the above equations. The equations indicate objects being tensors with respect to this non-linear transformation $\mathbf{F}_M(\mathbf{x})$ whose linearized-at-a-point matrix is $\mathbf{R}_M = \mathbf{M} = \mathbf{H}\mathbf{R}$, where $\mathbf{R}$ is the linearized-at-a-point matrix version of $\mathbf{F}(\mathbf{x})$, and $\mathbf{H}$ is the diagonal matrix of scale factors h'_i which are associated with g'_{ij} in x' -space. The $A'^{ijk\dots}$ are contravariant components of tensor A in x -space, while $A''^{ijk\dots}$ are the corresponding contravariant components of A in x' -space, all with respect to $\mathbf{F}_M(\mathbf{x})$ and its matrix $\mathbf{R}_M = \mathbf{M}$ where, for example, $d\mathbf{x}'' = \mathbf{M} d\mathbf{x}$.

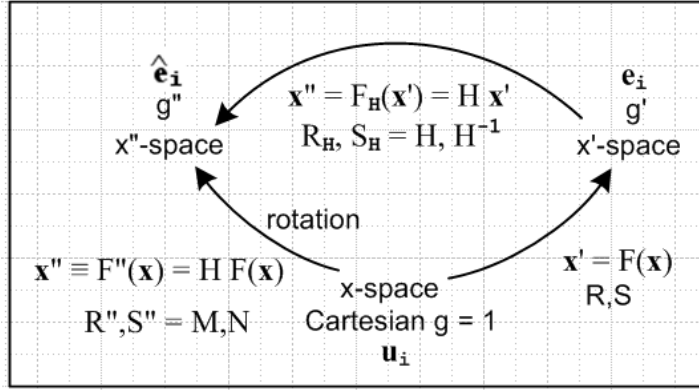
At this point the spaces are completely general, and none of $\mathbf{R}$, $\mathbf{R}_H = \mathbf{H}$, $\mathbf{R}_M = \mathbf{M}$ is a rotation matrix. In the next section, we shall specialize the above picture so that x -space is Cartesian with $g = 1$, and x' -space is the space of a set of *orthogonal* curvilinear coordinates $\mathbf{x}'$. In this scenario, $\mathbf{F}(\mathbf{x})$ is non-linear and so then is $\mathbf{F}_M(\mathbf{x}) = \mathbf{H} \mathbf{F}(\mathbf{x})$. Since the $\hat{\mathbf{e}}_i$ now form a frame of *orthonormal* vectors, and since $\mathbf{u}_i$ also form such a frame, one will not be surprised to find that $\mathbf{M}$ is now a rotation which relates these two frame sets.

Orthogonal curvilinear coordinates application

We now switch to picture B ($g=1$) and assume that the x'_i are orthogonal coordinates,



and our three-frame picture above then becomes



In this situation, x-space is Cartesian with $g_{ab} = g^{ab} = \delta_{ab}$ and $g'_{ab} = h'_a{}^2 \delta_{ab}$ and $g'^{ab} = h'_a{}^{-2} \delta_{ab}$. Since the metric tensors are diagonal, one could write $H = \sqrt{g'_{ab}}$, but we continue to use H.

In what follows, the plan is simply to exercise both the developmental and standard notations with regard to the M and N matrices. To this end, we first collect the following facts from Section 7 (o),

$$\begin{aligned} R^{ab} &= R^a{}_b, \quad g^{b'b} = R^a{}_b \\ R_a{}^b &= g'_{aa'} R^{a'}{}_b, \quad g^{b'b} = g'_{aa'} R^{a'}{}_b = h'_a{}^2 R^{ab} = h'_a{}^2 R^a{}_b \\ R_{ab} &= g'_{aa'} R^{a'}{}_b = h'_a{}^2 R^a{}_b \end{aligned}$$

or

$$R^{ab} = R^a{}_b \quad R_a{}^b = R_{ab} = h'_a{}^2 R^a{}_b \quad R^a{}_b = h'_a{}^{-2} R_a{}^b .$$

Also from Section 7 (11),

$$\begin{aligned} g'^{ab} &= R^a{}_a' R^b{}_b', \quad g^{a'b'} = R^a{}_a' R^b{}_b' = R^a{}_c R^b{}_c \\ g'_{ab} &= S^{a'}{}_a S^{b'}{}_b g_{a'b'} = R_a{}^{a'} R_b{}^{b'} g_{a'b'} = R_a{}^{a'} R_b{}^{a'} = R_a{}^c R_b{}^c \end{aligned}$$

or

$$g'^{ab} = R^a{}_c R^b{}_c \quad \text{and} \quad g'_{ab} = R_a{}^c R_b{}^c .$$

In this scenario, regardless of what R and S are, M is a "rotation" (verified below), where we include in this term possible axis reflections. What we really mean is that in developmental notation M is a real-orthogonal matrix, $MM^T = 1$. Since $N=M^{-1}$, N is then also a rotation.

To prove that M is a rotation in developmental notation, the standard notation equation $M^a{}_b \equiv h'_a R^a{}_b$ can be reverse-translated to $M_{ab} = h'_a R_{ab}$. Then (now $g' = RR^T$ from Section 5 (ℓ))

$$\begin{aligned} (MM^T)_{ac} &= M_{ab} M_{cb} = h'_a R_{ab} h'_c R_{cb} = h'_a h'_c R_{ab} R^T_{bc} = h'_a h'_c (RR^T)_{ac} \\ &= h'_a h'_c g'_{ac} = h'_a h'_c [h'_a{}^{-2} \delta_{a,c}] = \delta_{a,c} \Rightarrow MM^T = 1 . \end{aligned}$$

Proving the same thing directly in standard notation requires showing that $M^a{}_b M^c{}_b = \delta_{a,c}$ (see the end of Section 7 (i))

$$M^a{}_b M^c{}_b = h'_a R^a{}_b h'_c R^c{}_b = h'_a h'_c R^a{}_b R^c{}_b = h'_a h'_c R^a{}_b (h'_c{}^{-2} R^c{}_b) = (h'_a/h'_c) (R^a{}_b R^c{}_b)$$

$$= (h'_a/h'_c)\delta^a_c = \delta^a_c = \delta_{a,c}$$

where use was again made of the orthogonality rule of Section 7 (r), $R^a_b R^b_c = \delta^a_c$.

Relation between M and N. Looking at

$$M^a_b M^c_b = \delta_{a,c}$$

and knowing that

$$M^a_b (M^{-1})^b_c = \delta^a_c = \delta_{a,c}$$

one concludes that $(M^{-1})^b_c = M^c_b$. But $(M^{-1})^b_c = N^b_c$ so

$$N^b_c = M^c_b \quad // \text{ reminder: this does not say that } N = M^T \text{ in standard notation}$$

which can be verified from the above expressions for N and M.

Interpretation of N and M. Since $\mathbf{e}_n = S \mathbf{u}_n$ (Section 3 (a) with $\mathbf{e}'_n = \mathbf{u}_n$) and since $\mathbf{e}_n = h'_n \hat{\mathbf{e}}_n$, it follows that

$$\hat{\mathbf{e}}_n = h'^{-1}_n S \mathbf{u}_n$$

or

$$(\hat{\mathbf{e}}_n)^a = h'^{-1}_n S^a_b (\mathbf{u}_n)^b = h'^{-1}_n S^a_b \delta_n^b = h'^{-1}_n S^a_n = N^a_n = N^a_b \delta_n^b = N^a_b (\mathbf{u}_n)^b$$

or

$$\hat{\mathbf{e}}_n = N \mathbf{u}_n \quad \text{and} \quad (\hat{\mathbf{e}}_n)^a = N^a_n. \quad // \Rightarrow \mathbf{u}_n = M \hat{\mathbf{e}}_n$$

Since the $\mathbf{u}_n$ are the Cartesian unit vectors, it seems intuitively obvious that the transformation that moves this frame of orthonormal unit vectors $\{\mathbf{u}_n\}$ into the orthonormal frame $\{\hat{\mathbf{e}}_n\}$ must be a "rotation".

Above it was shown that $N^a_n = M^n_a$, therefore

$$M^n_a = N^a_n = (\hat{\mathbf{e}}_n)^a$$

The rotation matrix $N^a_n = (\hat{\mathbf{e}}_n)^a$ has the orthogonal basis vectors $\hat{\mathbf{e}}_n$ as its columns, while the rotation matrix $M^n_a = (\hat{\mathbf{e}}_n)^a$ has the orthogonal basis vectors $\hat{\mathbf{e}}_n$ as its rows. From this point of view, it seems pretty reasonable that $MN = 1$.

It might be noted that, in our situation with Cartesian x-space and orthogonal coordinates, $\hat{\mathbf{e}}^n = \hat{\mathbf{e}}_n$:

$$\hat{\mathbf{e}}^n = \mathbf{e}^n / |\mathbf{e}^n| \quad |\mathbf{e}^n|^2 = \mathbf{e}^n \bullet \mathbf{e}^n = g^{nn} = h'^{-2}_n$$

so

$$\hat{\mathbf{e}}^n = \mathbf{e}^n h'_n = h'_n g^{nn} \mathbf{e}_n = h'_n h'^{-2}_n \mathbf{e}_n = h'^{-1}_n \mathbf{e}_n = \hat{\mathbf{e}}_n .$$

The relation $\mathbf{u}_n = M \hat{\mathbf{e}}_n$ can be written $\mathbf{u}_n = M(\mathbf{x}) \hat{\mathbf{e}}_n(\mathbf{x})$ to emphasize that the rotation $M(\mathbf{x}) = R_M(\mathbf{x})$ is really a different rotation at every point $\mathbf{x}$, since the $\hat{\mathbf{e}}_n(\mathbf{x})$ vary with $\mathbf{x}$. This is very different from a global rotation which is the same at all points. For a global rotation R , $F = R = \text{linear}$ and $\partial_i u_j$ is a tensor. For R_M being a rotation which varies from point to point, F_M is non-linear just as F defining the curvilinear coordinates is non-linear, and $\partial_i u_j$ fails to be a tensor under either F or F_M .

One implication of the above picture relates to tensor equations being covariant, as discussed in Section 7 (u). If one has a tensor field equation in $\mathbf{x}$ -space,

$$Q_a^d(\mathbf{x}) = H_{ab}(\mathbf{x}) T^b_c(\mathbf{x}) B^d(\mathbf{x}),$$

in which all the objects transform as tensors with respect to the underlying $\mathbf{x}'' = F_M(\mathbf{x})$ (and its linear approximation $M(\mathbf{x})$ as in $d\mathbf{x}'' = M(\mathbf{x}) d\mathbf{x}$), then the equation is covariant and takes the same form in $\mathbf{x}''$ -space,

$$Q''^d_a(\mathbf{x}'') = H''_{ab}(\mathbf{x}'') T''^b_c(\mathbf{x}'') B''^d(\mathbf{x}'') .$$

An $\mathbf{x}$ -space observer has axes $\mathbf{u}_n$ (Frame S) while an $\mathbf{x}''$ -space observer has axes $\hat{\mathbf{e}}_n(\mathbf{x})$ (Frame S''), and these two sets of observation axes are related by $\mathbf{u}_n = M(\mathbf{x}) \hat{\mathbf{e}}_n(\mathbf{x})$ where $M(\mathbf{x})$ is a rotation. If the first equation describes something at location $\mathbf{x}$ in the realm of Newtonian mechanics, we expect the equation to have the same form in both Frame S and Frame S'' which are related by this local rotation $M(\mathbf{x})$. In other words, rotations are an invariance of Newtonian mechanics, and this means equations are covariant with respect to rotations. The above example, which might apply to fluid dynamics, has this covariance at each point $\mathbf{x}$ in the fluid, and it happens that the rotation is a different rotation at different points $\mathbf{x}$, but it is always a rotation.

Example: Polar Coordinates. In polar coordinates now with ordering $r, \theta = 1, 2$ one has

$$\begin{aligned} S^1_1 &= (\partial x / \partial r) = \cos \theta & x &= r \cos \theta \\ S^1_2 &= (\partial x / \partial \theta) = -r \sin \theta & y &= r \sin \theta \\ S^2_1 &= (\partial y / \partial r) = \sin \theta \\ S^2_2 &= (\partial y / \partial \theta) = r \cos \theta \end{aligned}$$

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

$$[g' = R R^T]_{\mathbf{D}\mathbf{N}} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta / r & \cos \theta / r \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta / r \\ \sin \theta & \cos \theta / r \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1/r^2 \end{pmatrix} \rightarrow g'^{ab} = \begin{pmatrix} h'^{-2}_{\mathbf{r}} & 0 \\ 0 & h'^{-2}_{\theta} \end{pmatrix}$$

so $h_{\mathbf{r}} = 1$ and $h_{\theta} = r$.

The N and M matrices may be computed as follows:

$$M^a_b \equiv h'^a_{\mathbf{a}} R^a_b = \begin{pmatrix} 1 & 0 \\ 0 & r \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta / r & \cos \theta / r \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} = R_z(-\theta)$$

$$N^{\mathbf{a}}_{\mathbf{b}} \equiv S^{\mathbf{a}}_{\mathbf{b}} h'_{\mathbf{b}}{}^{-1} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1/r \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = R_{\mathbf{z}}(\theta)$$

Therefore, the relation between a rank-2 tensor's $\hat{\mathbf{e}}_{\mathbf{n}}$ -expanded form components and the Cartesian form components is given by the expression stated above for rank-2 tensors,

$$A^{(\hat{\mathbf{e}})} = M A M^T$$

or

$$\begin{pmatrix} A_{\mathbf{r}\mathbf{r}} & A_{\mathbf{r}\theta} \\ A_{\theta\mathbf{r}} & A_{\theta\theta} \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \quad // \text{Lai p 316 Problem 5.71}$$

Airy functions. In isotropic elastic stress analysis for states of plane stress and plane strain, the Cartesian stress tensor T_{ij} has a simple form in which the upper left four components can be represented as derivatives of a potential-like function called an Airy function ϕ , so that $T_{11} = \partial_2^2 \phi$, $T_{22} = \partial_1^2 \phi$, and $T_{12} = T_{21} = -\partial_1 \partial_2 \phi$. In this case, the above equation becomes

$$\begin{pmatrix} T_{\mathbf{r}\mathbf{r}} & T_{\mathbf{r}\theta} \\ T_{\theta\mathbf{r}} & T_{\theta\theta} \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \partial_2^2 \phi & -\partial_1 \partial_2 \phi \\ -\partial_1 \partial_2 \phi & \partial_1^2 \phi \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \quad (*)$$

where

$$\begin{aligned} \partial_1 &= \cos\theta \partial_{\mathbf{r}} - (\sin\theta/r) \partial_{\theta} & \partial_1 &= \partial/\partial x_1 \\ \partial_2 &= \sin\theta \partial_{\mathbf{r}} + (\cos\theta/r) \partial_{\theta} & \partial_2 &= \partial/\partial x_2 \end{aligned}$$

Using Maple's dchange function, one can have Maple compute $\begin{pmatrix} T_{\mathbf{r}\mathbf{r}} & T_{\mathbf{r}\theta} \\ T_{\theta\mathbf{r}} & T_{\theta\theta} \end{pmatrix}$ from (*) to be

$$\begin{bmatrix} \frac{\frac{\partial}{\partial r} \phi(r, \theta)}{r} + \frac{\frac{\partial^2}{\partial \theta^2} \phi(r, \theta)}{r^2} & -\frac{\frac{\partial^2}{\partial \theta \partial r} \phi(r, \theta)}{r} + \frac{\frac{\partial}{\partial \theta} \phi(r, \theta)}{r^2} \\ -\frac{\frac{\partial^2}{\partial \theta \partial r} \phi(r, \theta)}{r} + \frac{\frac{\partial}{\partial \theta} \phi(r, \theta)}{r^2} & \frac{\partial^2}{\partial r^2} \phi(r, \theta) \end{bmatrix}$$

// Lai p 264 (5.27.3)

This then is a real-world example of using a rank-2 tensor in curvilinear coordinates expanded on the unit tangent base vectors. The mentioned plane of strain or stress has Cartesian coordinates x_1, x_2 which are converted to polar coordinates r, θ . The third Cartesian coordinate x_3 is more or less ignored.

The relation between the stress tensor T_{ij} and the infinitesimal strain tensor E_{ij} for an isotropic material is stated in Cartesian coordinate x -space as (a form of Hooke's Law generalizing $F = -kx$),

$$T_{ij} = \lambda \operatorname{tr}(E) \delta_{ij} + 2\mu E_{ij} \quad \text{or the same thing} \quad T^{ij} = \lambda \operatorname{tr}(E) \delta^{ij} + 2\mu E^{ij}$$

where λ and μ are Lamé's constants. With respect to transformation F_M , this is a "true tensor equation" ($\text{tr}(E) = E_{kk}$ is scalar under rotations), so according to Section 7 (u) it is "covariant" and in x -space may be written

$$T^{ij} = \lambda \text{tr}(E) \delta^{ij} + 2\mu E^{ij}$$

or

$$[T^{(\hat{e})}]^{ij} = \lambda [T^{(\hat{e})}]^{kk} \delta^{ij} + 2\mu [E^{(\hat{e})}]^{ij} .$$

For example, using the notation convention described above,

$$\begin{aligned} T_{rr} &= \lambda [T_{rr} + T_{\theta\theta}] + 2\mu E_{rr} \\ T_{r\theta} &= 2\mu E_{r\theta} . \end{aligned}$$

Notice that $\delta^{ij} = \delta^{ij}$ under transformation F_M , since $\delta^{ij} = M^i_a M^j_b \delta^{ab} = M^i_a M^j_a = \delta^{ij}$, whereas under transformation F one has $\delta^{ij} = R^i_a R^j_b \delta^{ab} = R^i_a R^j_a = g^{ij}$.

By way of contrast, the Cartesian-coordinates equation $E_{ij} = (\partial_i u_j + \partial_j u_i)/2$, which relates strain tensor E_{ij} to the vector displacement $\mathbf{u}$ of a continuum particle, is not a "true tensor equation", so $E_{r\theta} \neq (\partial_r u_\theta + \partial_\theta u_r)/2$. In fact, this relation is $E = [(\nabla \mathbf{u})^T + (\nabla \mathbf{u})]/2$ and $(\nabla \mathbf{u})$ for polar coordinates is computed in Appendix G and one ends up with $E_{r\theta} = (\partial_r u_\theta + (1/r) \partial_\theta u_r - u_\theta/r)/2$.

(i) Tensor expansions in a mixed basis

Recall the rank- n tensor expansion from section (b) above,

$$A = \sum_{ijk\dots} \alpha^{ijk\dots} (\mathbf{b}_i \otimes \mathbf{b}_j \otimes \mathbf{b}_k \dots) \quad \alpha^{ijk} = A \bullet (\mathbf{b}^i \otimes \mathbf{b}^j \otimes \mathbf{b}^k)$$

where $\alpha^{ijk\dots}$ are the coefficients of the expansion of A on the direct product basis shown. A might be a tensor, or it might be a tensor-like object. To make explicit the fact that the coefficients depend on the choice of basis, one might write (one b for each index, number of b 's is the rank of the tensor),

$$\alpha^{ijk\dots} = [A^{(b,b,b\dots)}]^{ijk\dots} .$$

The fact that the indices $ijk\dots$ are "up" indicates that the b label stands for the $\mathbf{b}_i$ basis and not $\mathbf{b}^i$. So here is an example showing the expansion of a rank-3 tensor,

$$A = \sum_{ijk} [A^{(b,b,b)}]^{ijk} (\mathbf{b}_i \otimes \mathbf{b}_j \otimes \mathbf{b}_k) \quad [A^{(b,b,b)}]^{ijk} = A \bullet (\mathbf{b}^i \otimes \mathbf{b}^j \otimes \mathbf{b}^k) .$$

Earlier we used the simpler notation $[A^{(b)}]^{ijk}$ for the above coefficient, but now we want to show all the basis elements because now we want to consider a "mixed basis expansion" such as

$$A = \sum_{ijk} [A^{(b,e,u)}]^{ijk} (\mathbf{b}_i \otimes \mathbf{e}_j \otimes \mathbf{u}_k) \quad [A^{(b,e,u)}]^{ijk} = A \bullet (\mathbf{b}^i \otimes \mathbf{e}^j \otimes \mathbf{u}^k) .$$

This is a completely viable expansion since the b , e and u basis vectors are each a complete set within their part of the direct-product space. To verify the validity of this expansion, consider :

$$\begin{aligned}
 [A^{(\mathbf{b}, \mathbf{e}, \mathbf{u})}]^{ijk} &= \{A\} \bullet (\mathbf{b}^i \otimes \mathbf{e}^j \otimes \mathbf{u}^k) \\
 &= \{ \sum_{i', j', k'} [A^{(\mathbf{b}, \mathbf{e}, \mathbf{u})}]^{i' j' k'} (\mathbf{b}_{i'} \otimes \mathbf{e}_{j'} \otimes \mathbf{u}_{k'}) \} \bullet (\mathbf{b}^i \otimes \mathbf{e}^j \otimes \mathbf{u}^k) \\
 &= \{ \sum_{i', j', k'} [A^{(\mathbf{b}, \mathbf{e}, \mathbf{u})}]^{i' j' k'} (\mathbf{b}^{i'} \bullet \mathbf{b}_i) (\mathbf{e}^{j'} \bullet \mathbf{e}_j) (\mathbf{u}^{k'} \bullet \mathbf{u}_k) \} \\
 &= \{ \sum_{i', j', k'} [A^{(\mathbf{b}, \mathbf{e}, \mathbf{u})}]^{i' j' k'} \delta^{i'}_i \delta^{j'}_j \delta^{k'}_k \} \\
 &= [A^{(\mathbf{b}, \mathbf{e}, \mathbf{u})}]^{ijk} .
 \end{aligned}$$

Alternatively, one could relate this coefficient to the coefficients expanded on $\mathbf{u}_i \otimes \mathbf{u}_j \otimes \mathbf{u}_k$,

$$[A^{(\mathbf{b}, \mathbf{e}, \mathbf{u})}]^{ijk} = A \bullet (\mathbf{b}^i \otimes \mathbf{e}^j \otimes \mathbf{u}^k) = A^{abc} (\mathbf{b}^i)_a (\mathbf{e}^j)_b (\mathbf{u}^k)_c$$

as was shown near the start of section (b).

In the case of a rank-2 tensor, one has the option of using the other notations discussed above,

$$\begin{aligned}
 A &= \sum_{ij} [A^{(\mathbf{b}, \mathbf{u})}]^{ij} (\mathbf{b}_i \otimes \mathbf{u}_j) = \sum_{ij} [A^{(\mathbf{b}, \mathbf{u})}]^{ij} (\mathbf{b}_i \mathbf{u}_j) \\
 &\quad \text{direct product} \qquad \qquad \text{dyadic} \\
 &= \sum_{ij} [A^{(\mathbf{b}, \mathbf{u})}]^{ij} (\mathbf{b}_i \mathbf{u}_j^T) = \sum_{ij} [A^{(\mathbf{b}, \mathbf{u})}]^{ij} |b_i\rangle\langle u_j| \\
 &\quad \text{matrix} \qquad \qquad \text{bra-ket}
 \end{aligned}$$

where

$$[A^{(\mathbf{b}, \mathbf{u})}]^{ij} = A \bullet (\mathbf{b}^i \otimes \mathbf{u}^j) = A^{ab} (\mathbf{b}^i)_a (\mathbf{u}^j)_b = (\mathbf{b}^i)^T A (\mathbf{u}^j) = \langle b^i | A | u^j \rangle .$$

One can of course use unit versions of the $\mathbf{e}_n$ basis vectors, $\hat{\mathbf{e}}_n$, and then one might write for example

$$A = \sum_{ijk} [A^{(\hat{\mathbf{e}}, \hat{\mathbf{e}}, \mathbf{u})}]^{ijk} (\hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j \otimes \mathbf{u}_k)$$

where the hats are replicated into the superscript tensor label.

Example of a mixed-basis expansion of a tensor

The identity tensor can be expanded this way, since $\mathbf{u}^i \bullet \mathbf{u}_j = (\mathbf{u}^i)^T \mathbf{u}_j = \delta^i_j = \langle \mathbf{u}^i | \mathbf{u}_j \rangle$,

$$\begin{aligned}
 1 &= \sum_j \mathbf{u}_j \otimes \mathbf{u}^j = \sum_j \mathbf{u}_j (\mathbf{u}^j)^T = \sum_j \mathbf{u}_j \mathbf{u}^j = \sum_j | \mathbf{u}_j \rangle \langle \mathbf{u}^j | . \\
 &\quad \text{direct product} \quad \text{matrix} \quad \text{dyadic} \quad \text{bra-ket}
 \end{aligned}$$

The $\mathbf{u}_i$ are related to the $\mathbf{e}_i$ according to

$$\mathbf{u}_j = R^i_j \mathbf{e}_i .$$

Proof: $(\mathbf{u}_j)^{\mathbf{a}} = R^{\mathbf{i}}_{\mathbf{j}} (\mathbf{e}_i)^{\mathbf{a}} \Rightarrow \delta_j^{\mathbf{a}} = R^{\mathbf{i}}_{\mathbf{j}} R_i^{\mathbf{a}}$ which is an orthogonality rule of Section 7 (r).

It follows that

$$1 = \underbrace{\sum_{i,j} R^{\mathbf{i}}_{\mathbf{j}} \mathbf{e}_i \otimes \mathbf{u}_j}_{\text{direct product}} = \underbrace{\sum_{i,j} R^{\mathbf{i}}_{\mathbf{j}} \mathbf{e}_i (\mathbf{u}_j)^{\mathbf{T}}}_{\text{matrix}} = \underbrace{\sum_{i,j} R^{\mathbf{i}}_{\mathbf{j}} \mathbf{e}_i \mathbf{u}_j^{\mathbf{T}}}_{\text{dyadic}} = \underbrace{\sum_{i,j} R^{\mathbf{i}}_{\mathbf{j}} |\mathbf{e}_i\rangle \langle \mathbf{u}_j|}_{\text{bra-ket}}$$

where

$$R^{\mathbf{i}}_{\mathbf{j}} = \underbrace{(\mathbf{e}^{\mathbf{i}})^{\mathbf{T}}}_{\text{matrix}} \underbrace{1}_{\text{dot}} \underbrace{\mathbf{u}_j}_{\text{dyadic}} = \mathbf{e}^{\mathbf{i}} \bullet \mathbf{u}_j = \mathbf{e}^{\mathbf{i}} \mathbf{u}_j = \langle \mathbf{e}^{\mathbf{i}} | \mathbf{u}_j \rangle \quad // \text{ as in Section 7 (s)}$$

This then is a mixed-basis expansion of the identity tensor. Another form would be

$$1 = \sum_{i,j} R^{\mathbf{i}j} \mathbf{e}_i \otimes \mathbf{u}_j = \sum_{i,j} [1^{(\mathbf{e}, \mathbf{u})}]^{\mathbf{i}j} \mathbf{e}_i \otimes \mathbf{u}_j$$

following the notation discussed above, leading to this rather obscure way of writing $R^{\mathbf{i}j}$,

$$R^{\mathbf{i}j} = [1^{(\mathbf{e}, \mathbf{u})}]^{\mathbf{i}j}.$$

(j) What is a tensor?

We are now in a better position to examine some possible answers to this question.

(1) A tensor is an *operator* like A which lives inside a direct product Hilbert Space.

We can associate with this tensor A a large variety of up-indexed objects such as $[A^{(\mathbf{b}, \mathbf{e}, \mathbf{u})}]^{\mathbf{i}j\mathbf{k}}$ in the example above. If one has at hand K different bases of interest, then for a tensor of rank n there would be K^n possible up-indexed objects. In the case of rank 2, these K^2 different indexed objects are matrices. Given the $\mathbf{u}_n$ and $\mathbf{e}_n$ bases used throughout this document, there are two *special* up-indexed objects of rank 3: $[A^{(\mathbf{u}, \mathbf{u}, \mathbf{u})}]^{\mathbf{i}j\mathbf{k}}$ and $[A^{(\mathbf{e}, \mathbf{e}, \mathbf{e})}]^{\mathbf{i}j\mathbf{k}}$ which we abbreviate as $[A^{(\mathbf{u})}]^{\mathbf{i}j\mathbf{k}}$ and $[A^{(\mathbf{e})}]^{\mathbf{i}j\mathbf{k}}$ or as $[A]^{\mathbf{i}j\mathbf{k}}$ and $[A']^{\mathbf{i}j\mathbf{k}}$. The indexed object $[A]^{\mathbf{i}j\mathbf{k}}$ is a set of N^3 contravariant components of a rank-3 tensor in x -space, and $[A']^{\mathbf{i}j\mathbf{k}}$ is a set of N^3 contravariant components of the same rank-3 tensor in x' -space, where these two spaces are linked by a transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$ which has a linearized form $d\mathbf{x}' = \mathbf{R} d\mathbf{x}$ at a point $\mathbf{x}$. The two sets of contravariant components are related by

$$[A']^{\mathbf{i}j\mathbf{k}} = R^{\mathbf{i}}_{\mathbf{i}'} R^{\mathbf{j}}_{\mathbf{j}'} R^{\mathbf{k}}_{\mathbf{k}'} [A]^{\mathbf{i}'\mathbf{j}'\mathbf{k}'}.$$

These sets of components "transform as a rank-3 tensor with respect to the underlying transformation $\mathbf{x}' = \mathbf{F}(\mathbf{x})$. As outlined in Section 7, the all-up indexed tensor of rank n is just one of a family of 2^n tensors where the indices take all possibly up and down positions, and each such tensor has a corresponding transformation rule, such as

$$[A']^{\mathbf{i}j\mathbf{k}} = R^{\mathbf{i}}_{\mathbf{i}'} R^{\mathbf{j}}_{\mathbf{j}'} R^{\mathbf{k}}_{\mathbf{k}'} [A]^{\mathbf{i}'\mathbf{j}'\mathbf{k}'} \quad \text{for} \quad A = \sum_{i,j,k} [A']^{\mathbf{i}j\mathbf{k}} (\mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k).$$

For this definition of "tensor" as an abstract operator A , there are many possible set pairs of components (indexed objects where the values of all indices are set in all possible ways) which do NOT transform as a

rank-3 tensor with respect to F as just described. The component sets that do transform as tensors are those for which the components are coefficients of an expansion of A on a direct product basis where the individual basis vectors are selected from $\mathbf{e}_i$ or $\mathbf{e}^i$, or are selected from $\mathbf{u}_i$ or $\mathbf{u}^i$. If some other generic basis vector $\mathbf{b}_i$ appears, then the component set does not transform as a tensor under F .

Definition (1) is basically the definition of "tensor" used in this document.

(2) Another definition of tensor might be: a tensor is any of the indexed objects mentioned in (1) above. The set of components like $[A^{(\mathbf{b}, \mathbf{e}, \mathbf{u})}]^{ijk}$ is called a "tensor" because it is a possible coefficient set that can be obtained by expanding the "tensor operator" A on suitable basis vectors. Just as a matrix is sometimes written without its indices, so this tensor object might be represented just as $A^{(\mathbf{b}, \mathbf{e}, \mathbf{u})}$. The components of this particular indexed object do not transform as either end of the transformation rule stated above, so this tensor is a tensor, but does not transform as a tensor.

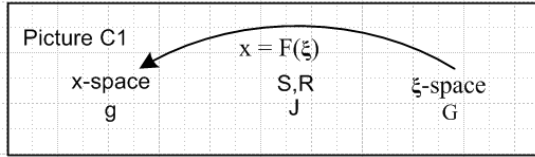
(3) A third possible definition: a tensor is any indexed object each of whose indices ranges from 1 to N where N is the dimension of one's space of interest. By this definition, any $N \times N$ matrix A_{ij} would be a tensor of rank 2. It would be very unlikely that a random matrix like this would be part of the transformation rule $[A']^{ij} = R^i_k R^j_l [A]^{kl}$, so this matrix A_{ij} is then a tensor, but it probably doesn't transform as a tensor.

In fields of physics involving relativity, the first definition is normally used, and an indexed object is called a tensor only if it transforms the way a tensor should transform with respect to a transformation of interest. For example, the affine connection Γ^c_{ab} is never called a tensor. In most other areas of physics definition (3) seems more common, where any matrix is a rank-2 tensor, also known as a second-order tensor. The whole subject of second order tensors is then identified with linear algebra where the operators are matrices. It may turn out that a particular matrix is in fact a tensor by definition (1) with respect to rotations. This is the case for basic matrices involved in equations which must be covariant. In continuum mechanics, which generally uses definition (3), there are many tensors which do not transform as tensors under rotations or other transformations. In that field, when a tensor in fact transforms as a tensor, it is called an objective tensor (sometimes an indifferent tensor). Equations which are covariant in the sense of Section 7 (u) are called "frame indifferent". See Appendix K for examples.

Appendix F: The Affine Connection Γ^c_{ab} and Covariant Derivatives

(a) Definition and Interpretation of Γ : $\Gamma^c_{ab} = e^c \bullet (\partial_a e_b) = R^c_i (\partial_a R^i_b)$

Context can be confusing in a discussion of the affine connection Γ , so we start with a modified Picture C in which the quasi-Cartesian space on the right is called ξ -space instead of $x^{(0)}$ -space as in Picture C. The notation ξ^i for the coordinates of ξ -space seems traditional in general relativity writing, an application in which the Γ object appears frequently.



Recall from Section 1 that the metric tensor G is a diagonal matrix whose elements are independently $+1$ or -1 . And recall from Section 5 (b) that in x -space, $g_{ab} = R^i_a R^j_b G_{ij}$. If $G = 1$, then the x^i coordinates of x -space are the "curvilinear coordinates" and the ξ^i are the "Cartesian coordinates".

In Picture C1 the tangent base vectors e_n exist in ξ -space and the components of e_n are given by $(e_n)^i = R^i_n$, as in Example 1 of Section 3 (polar coordinates). Since matrix R is the linearization of transformation F at ξ and since $x = F(\xi)$, one can regard R as a function either of ξ or x . We choose x as the variable and write $[e_n(x)]^i = R^i_n(x)$. For example, Example 2 of Section 3 (spherical coordinates) showed that $e_\phi(x) = r \sin\theta \hat{\phi}(r, \theta, \phi)$.

In any event, $e_n(x)$ varies with x , and one wonders just *how* e_n varies with x . For a small variation dx in the x -space coordinates, one has,

$$d(e_n)^i = \partial_j (e_n)^i dx^j \quad \partial_j \equiv \partial / \partial x^j$$

or

$$(de_n)^i = (\partial_j e_n)^i dx^j .$$

Since e_n and $(\partial_j e_n)$ are both vectors in ξ -space, and since the e_k are known to form a complete basis in ξ -space, it must be possible to expand $(\partial_j e_n)$ on the e_k with some appropriate coefficients, call them Γ^k_{jn} :

$$(\partial_j e_n) = \Gamma^k_{jn} e_k .$$

Dotting this equation into e^k gives

$$\Gamma^k_{jn} = e^k \bullet (\partial_j e_n) = R^k_i (\partial_j R^i_n)$$

where we recall from Section 7 (s) that $(e^k)_i = R^k_i$ and $(e_n)^i = R^i_n$. These coefficients $\Gamma^k_{jn}(x)$ comprise a tensor-like field called the **affine connection**, so we shall regard the above line as the definition of Γ^k_{jn} in the context of Picture C1 above. With more standard index names, the above becomes

$$\Gamma^c_{ab} = e^c \bullet (\partial_a e_b) = R^c_i (\partial_a R^i_b) .$$

From Section 7 (q) we know that, for Picture C1,

$$\begin{aligned} R^c_i &= (\partial x^c / \partial \xi^i) \\ R_b^i &= (\partial \xi^i / \partial x^b) \\ (\partial_a R_b^i) &= (\partial^2 \xi^i / \partial x^a \partial x^b) = (\partial_b R_a^i) \end{aligned}$$

Therefore one can write

$$\Gamma^c_{ab} = R^c_i (\partial_a R_b^i) = (\partial x^c / \partial \xi^i) (\partial^2 \xi^i / \partial x^a \partial x^b) = \frac{\partial x^c}{\partial \xi^i} \frac{\partial^2 x^i}{\partial \xi^a \partial \xi^b},$$

which form appears in Weinberg p 100 (4.5.1). Notice again that $(\partial_b R_c^n) = (\partial_c R_b^n)$ and that Γ^c_{bc} is symmetric on the lower two indices. In the next section an alternate form for Γ is derived, so both forms will be stated here:

$$\begin{aligned} \Gamma^c_{ab} &= R^c_i (\partial_a R_b^i) & // \partial_a &= \partial / \partial x^a \\ \Gamma^c_{ab} &= -R_b^i (\partial_a R^c_i) \end{aligned}$$

(b) Identities of the form $(\partial_a R^d_n) = -R^e_n R^d_m (\partial_a R^e_m)$

The identities are :

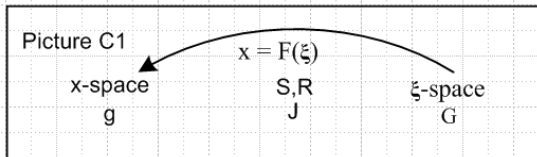
$$\begin{aligned} R^d_m (\partial_a R^e_m) &= -R^e_m (\partial_a R^d_m) & 1 \\ (\partial_a R^d_n) &= -R^e_n R^d_m (\partial_a R^e_m) & 2 \quad \partial_a \equiv \partial / \partial x^a \\ (\partial_a R^d_n) &= -R^e_n R^d_m (\partial_a R^e_m) & 3 \end{aligned}$$

The second two lines are just restatements of the first line, but all are derived below.

Corollary: The first identity above allows an alternate form for the affine connection in terms of R:

$$\begin{aligned} \Gamma^c_{ab} &= R^c_i (\partial_a R_b^i) & // \text{as in section (a) above} \\ \Gamma^c_{ab} &= -R_b^i (\partial_a R^c_i) & // \text{alternate form} \end{aligned}$$

Our context is:



Proof: These identities are a simple consequence of the fact that $RS = 1$ which in standard notation is written $\delta^c_b = R^c_\alpha R_b^\alpha$ (one of the orthogonality rules). So,

$$0 = \partial_a(\delta^d_e) = \partial_a(R^d_m R_e^m) = R^d_m (\partial_a R_e^m) + R_e^m (\partial_a R^d_m) \quad \text{QED 1} \quad (*)$$

Apply $\Sigma_e R^n$ to both sides of (*) to get (or, just use the Inversion Rule of Section 7 (r))

$$\begin{aligned} 0 &= R^n_e R^d_m (\partial_a R_e^m) + (R^n_e R_e^m) (\partial_a R^d_m) = R^n_e R^d_m (\partial_a R_e^m) + \delta_n^m (\partial_a R^d_m) \\ &= R^n_e R^d_m (\partial_a R_e^m) + (\partial_a R^d_n) \end{aligned}$$

$$\Rightarrow (\partial_a R^d_n) = - R^n_e R^d_m (\partial_a R_e^m) \quad \text{QED 2}$$

Alternatively, apply $\Sigma_d R^d$ to both sides of (*) to get

$$\begin{aligned} 0 &= (R^d_n R^d_m) (\partial_a R_e^m) + R^d_n R_e^m (\partial_a R^d_m) = \delta_n^m (\partial_a R_e^m) + R^d_n R_e^m (\partial_a R^d_m) \\ &= (\partial_a R_e^n) + R^d_n R_e^m (\partial_a R^d_m) \end{aligned}$$

$$\Rightarrow (\partial_a R_e^n) = - R^d_n R_e^m (\partial_a R^d_m) \quad \text{now swap d and e:}$$

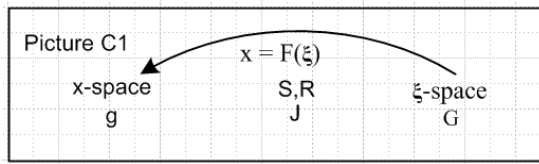
$$\Rightarrow (\partial_a R^d_n) = - R^n_e R^d_m (\partial_a R_e^m) \quad \text{QED 3}$$

(c) Identities of the form $(\partial_c g^{ab}) = -[g^{an} \Gamma^b_{cn} + g^{bn} \Gamma^a_{cn}]$

The derivatives of the metric tensor are given by $(\partial_c = \partial/\partial x^c)$

$$(\partial_c g^{ab}) = -[g^{an} \Gamma^b_{cn} + g^{bn} \Gamma^a_{cn}] \quad 1$$

$$(\partial_c g_{ab}) = +[g_{an} \Gamma^n_{cb} + g_{bn} \Gamma^n_{ca}] \quad 2$$



Proof of 1: $(\partial_c g^{ab}) = -[g^{an} \Gamma^b_{cn} + g^{bn} \Gamma^a_{cn}]$

The LHS is given by

$$\text{LHS} = (\partial_c g^{ab}) = \partial_c (R^a_i R^b_i) G^{ii} = R^a_i (\partial_c R^b_i) G^{ii} + R^b_i (\partial_c R^a_i) G^{ii}$$

For the RHS, the Γ objects can be replaced by their alternate definitions from section (a)

$$\begin{aligned} \Gamma^c_{ab} &= -R^i_b (\partial_a R^c_i) & // \text{ from section (a)} \\ \Gamma^b_{cn} &= -R^n_i (\partial_c R^b_i) & // b \rightarrow n \text{ then } c \rightarrow b \text{ then } a \rightarrow c \\ \Gamma^a_{cn} &= -R^n_i (\partial_c R^a_i) \end{aligned}$$

The RHS of the claimed identity may then be written

$$\begin{aligned}
 \text{RHS} &= -g^{an} \Gamma_{cn}^b - g^{bn} \Gamma_{cn}^a \\
 &= \{R_{\mathbf{k}}^a R_{\mathbf{k}}^n G^{\mathbf{k}\mathbf{k}}\} \{R_{\mathbf{n}}^i (\partial_c R_{\mathbf{i}}^b)\} + \{R_{\mathbf{k}}^b R_{\mathbf{k}}^n G^{\mathbf{k}\mathbf{k}}\} \{R_{\mathbf{n}}^i (\partial_c R_{\mathbf{i}}^a)\} \\
 &= R_{\mathbf{k}}^a (R_{\mathbf{n}}^n R_{\mathbf{n}}^i) (\partial_c R_{\mathbf{i}}^b) G^{\mathbf{k}\mathbf{k}} + R_{\mathbf{k}}^b (R_{\mathbf{n}}^n R_{\mathbf{n}}^i) (\partial_c R_{\mathbf{i}}^a) G^{\mathbf{k}\mathbf{k}} \\
 &= R_{\mathbf{k}}^a \delta_{\mathbf{k}}^i (\partial_c R_{\mathbf{i}}^b) G^{\mathbf{k}\mathbf{k}} + R_{\mathbf{k}}^b \delta_{\mathbf{k}}^i (\partial_c R_{\mathbf{i}}^a) G^{\mathbf{k}\mathbf{k}} \\
 &= R_{\mathbf{i}}^a (\partial_c R_{\mathbf{i}}^b) G^{\mathbf{i}\mathbf{i}} + R_{\mathbf{i}}^b (\partial_c R_{\mathbf{i}}^a) G^{\mathbf{i}\mathbf{i}} = \text{LHS} \quad \text{QED}
 \end{aligned}$$

Proof of 2: $(\partial_c g_{ab}) = +[g_{an} \Gamma_{cb}^n + g_{bn} \Gamma_{ca}^n]$

The LHS is given by

$$\text{LHS} = (\partial_c g_{ab}) = \partial_c (R_{\mathbf{a}}^i R_{\mathbf{b}}^i) G_{\mathbf{i}\mathbf{i}} = R_{\mathbf{a}}^i (\partial_c R_{\mathbf{b}}^i) G_{\mathbf{i}\mathbf{i}} + R_{\mathbf{b}}^i (\partial_c R_{\mathbf{a}}^i) G_{\mathbf{i}\mathbf{i}}$$

For the RHS, the Γ objects can be replaced by their primary definitions from section (a)

$$\begin{aligned}
 \Gamma_{\mathbf{ab}}^{\mathbf{c}} &= R_{\mathbf{i}}^{\mathbf{c}} (\partial_{\mathbf{a}} R_{\mathbf{b}}^{\mathbf{i}}) \\
 \Gamma_{\mathbf{cb}}^{\mathbf{n}} &= R_{\mathbf{i}}^{\mathbf{n}} (\partial_{\mathbf{c}} R_{\mathbf{b}}^{\mathbf{i}}) \quad // \mathbf{c} \rightarrow \mathbf{n} \text{ then } \mathbf{a} \rightarrow \mathbf{c} \text{ then } \mathbf{i} \rightarrow \mathbf{k} \\
 \Gamma_{\mathbf{ca}}^{\mathbf{n}} &= R_{\mathbf{i}}^{\mathbf{n}} (\partial_{\mathbf{c}} R_{\mathbf{a}}^{\mathbf{i}})
 \end{aligned}$$

The RHS of the claimed identity may then be written

$$\begin{aligned}
 \text{RHS} &= g_{an} \Gamma_{cb}^n + g_{bn} \Gamma_{ca}^n \\
 &= \{R_{\mathbf{a}}^{\mathbf{k}} R_{\mathbf{n}}^{\mathbf{k}} G_{\mathbf{k}\mathbf{k}}\} \{R_{\mathbf{i}}^{\mathbf{n}} (\partial_{\mathbf{c}} R_{\mathbf{b}}^{\mathbf{i}})\} + \{R_{\mathbf{b}}^{\mathbf{k}} R_{\mathbf{n}}^{\mathbf{k}} G_{\mathbf{k}\mathbf{k}}\} \{R_{\mathbf{i}}^{\mathbf{n}} (\partial_{\mathbf{c}} R_{\mathbf{a}}^{\mathbf{i}})\} \\
 &= R_{\mathbf{a}}^{\mathbf{k}} (R_{\mathbf{n}}^{\mathbf{k}} R_{\mathbf{n}}^i) (\partial_{\mathbf{c}} R_{\mathbf{b}}^i) G_{\mathbf{k}\mathbf{k}} + R_{\mathbf{b}}^{\mathbf{k}} (R_{\mathbf{n}}^{\mathbf{k}} R_{\mathbf{n}}^i) (\partial_{\mathbf{c}} R_{\mathbf{a}}^i) G_{\mathbf{k}\mathbf{k}} \\
 &= R_{\mathbf{a}}^{\mathbf{k}} \delta_{\mathbf{i}}^i (\partial_{\mathbf{c}} R_{\mathbf{b}}^i) G_{\mathbf{k}\mathbf{k}} + R_{\mathbf{b}}^{\mathbf{k}} \delta_{\mathbf{i}}^i (\partial_{\mathbf{c}} R_{\mathbf{a}}^i) G_{\mathbf{k}\mathbf{k}} \\
 &= R_{\mathbf{i}}^i (\partial_{\mathbf{c}} R_{\mathbf{b}}^i) G_{\mathbf{i}\mathbf{i}} + R_{\mathbf{i}}^i (\partial_{\mathbf{c}} R_{\mathbf{a}}^i) G_{\mathbf{i}\mathbf{i}} = \text{LHS} \quad \text{QED}
 \end{aligned}$$

(d) Identity: $\Gamma_{\mathbf{ab}}^{\mathbf{d}} = (1/2) g^{\mathbf{dc}} [\partial_{\mathbf{a}} g_{\mathbf{bc}} + \partial_{\mathbf{b}} g_{\mathbf{ca}} - \partial_{\mathbf{c}} g_{\mathbf{ab}}]$

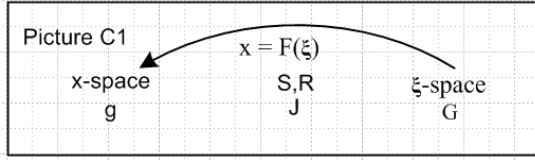
The identity states that $\Gamma_{\mathbf{ab}}^{\mathbf{d}}$ may be expressed entirely in terms of the metric tensor,

$$\Gamma_{\mathbf{ab}}^{\mathbf{d}} = (1/2) g^{\mathbf{dc}} [\partial_{\mathbf{a}} g_{\mathbf{bc}} + \partial_{\mathbf{b}} g_{\mathbf{ca}} - \partial_{\mathbf{c}} g_{\mathbf{ab}}] .$$

Recall in our definition above, $\Gamma_{ab}^d = R_{\mathbf{k}}^d(\partial_a R_b^{\mathbf{k}})$, that Γ was given in terms of R matrices.

The following corollary (derived at the end of this section) concerns contraction of the upper Γ index with a lower one:

$$\Gamma_{an}^a = (1/2) g^{ad} \partial_n g_{ad} = (1/2)(1/g) \partial_n g = (1/\sqrt{g}) \partial_n (\sqrt{g}) .$$



Proof: $\Gamma_{ab}^d = (1/2) g^{dc} [\partial_a g_{bc} + \partial_b g_{ca} - \partial_c g_{ab}]$

We know that

$$\begin{aligned} g_{ab} &= R_a^e R_b^e G_{ee} & // \text{ sum on } e \\ g^{dc} &= R_i^d R_i^c G^{ii} & // \text{ sum on } i \end{aligned}$$

The first line below is computed from the first line in the pair above, then the next two lines below are obtained by doing forward cyclic permutations of the first line :

$$\begin{aligned} \partial_c g_{ab} &= [R_b^e (\partial_c R_a^e) + R_a^e (\partial_c R_b^e)] G_{ee} \\ \partial_a g_{bc} &= [R_c^e (\partial_a R_b^e) + R_b^e (\partial_a R_c^e)] G_{ee} \\ \partial_b g_{ca} &= [R_a^e (\partial_b R_c^e) + R_c^e (\partial_b R_a^e)] G_{ee} . \end{aligned}$$

The last four lines can be inserted into the Right Hand Side of our desired identity to obtain

$$\begin{aligned} (\text{RHS})_{ab}^d &= (1/2) g^{dc} [\partial_a g_{bc} + \partial_b g_{ca} - \partial_c g_{ab}] \\ &= (1/2) (R_i^d R_i^c G^{ii}) G_{ee} * \\ &\quad [\underset{1}{R_c^e (\partial_a R_b^e)} + \underset{2}{R_b^e (\partial_a R_c^e)} + \underset{3}{R_a^e (\partial_b R_c^e)} + \underset{4}{R_c^e (\partial_b R_a^e)} - \underset{5}{R_b^e (\partial_c R_a^e)} - \underset{6}{R_a^e (\partial_c R_b^e)}] \end{aligned}$$

Due to the symmetry noted above in section (a), $(\partial_i R_j^e) = (\partial_j R_i^e)$, terms 2 and 5 cancel as do terms 3 and 6, while terms 1 and 4 are equal. Therefore,

$$\begin{aligned} (\text{RHS})_{ab}^d &= (1/2) R_i^d R_i^c G^{ii} G_{ee} * 2 R_c^e (\partial_a R_b^e) \\ &= R_i^d (R_c^e R_i^c) G^{ii} G_{ee} (\partial_a R_b^e) = R_i^d \delta_i^e G^{ii} G_{ee} (\partial_a R_b^e) \\ &= R_e^d G^{ee} G_{ee} (\partial_a R_b^e) = R_e^d (\partial_a R_b^e) = \Gamma_{ab}^d \quad \text{QED} \end{aligned}$$

Proof of Corollary: The abovementioned corollary is this ($g \equiv \det(g_{ab})$)

$$\Gamma^a_{an} = (1/2) g^{ad} (\partial_a g_{nd} + \partial_n g_{ad} - \partial_d g_{an}) = (1/2) g^{ad} \partial_n g_{ad} = (1/2) (1/g) \partial_n g = (1/\sqrt{g}) \partial_n (\sqrt{g})$$

The first and third terms of the second expression cancel due to symmetry

$$g^{ad} \partial_a g_{nd} - g^{ad} \partial_d g_{an} = g^{ad} \partial_a g_{nd} - g^{da} \partial_a g_{dn} = g^{ad} \partial_a g_{nd} - g^{ad} \partial_a g_{nd} = 0$$

and the fact that $(1/g) \partial_n g = g^{ab} (\partial_n g_{ab})$ is proved as follows:

$$(1) g^{ab} = (g^{-1})_{ab} = \text{cof}(g_{ab})^T / \det(g_{ab}) = \text{cof}(g_{ab}) / g \Rightarrow \text{cof}(g_{ab}) = g g^{ab}$$

$$(2) g = \det(g_{ab}) = \Sigma_{ab} g_{ab} \text{cof}(g_{ab}) \Rightarrow \partial g / \partial g_{ab} = \text{cof}(g_{ab}) = g g^{ab}$$

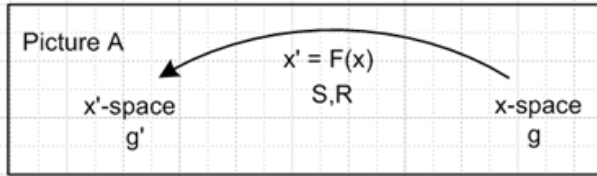
$$(3) \partial_n g = \partial g / \partial x^n = (\partial g / \partial g_{ab}) (\partial g_{ab} / \partial x^n) = g g^{ab} (\partial_n g_{ab}) \Rightarrow (1/g) \partial_n g = g^{ab} (\partial_n g_{ab}).$$

The final form shown is just calculus :

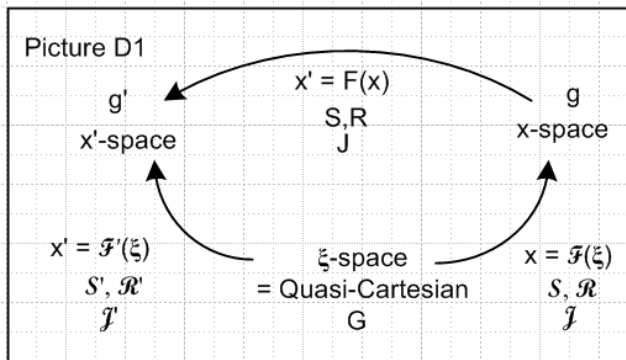
$$g^{-1/2} \partial_n (g^{1/2}) = g^{-1/2} (1/2) g^{-1/2} \partial_n g = (1/2) (1/g) (\partial_n g).$$

(e) Picture D1 Context

Our main context of interest is Picture A,



For the proof given in the next section below, it is useful to think of Picture A as the top part of this Picture D1,



The relationship between the R's and S's are these

$$\begin{aligned} R &= \mathcal{R}' \mathcal{R}^{-1} = \mathcal{R}' S & \Rightarrow & \quad \mathcal{R}' = R \mathcal{R} \\ S &= \mathcal{R} \mathcal{R}'^{-1} = \mathcal{R} S' \end{aligned}$$

The tangent and reciprocal base vectors in ξ -space associated with transformations $\mathcal{F}$ and $\mathcal{F}'$ are these:

$$\begin{aligned} (\mathbf{e}_n)^i &= \mathcal{R}_n^i & (\mathbf{e}^n)_i &= \mathcal{R}^n_i & \text{x-space} \\ (\mathbf{e}'_n)^i &= \mathcal{R}'_n^i & (\mathbf{e}'^n)_i &= \mathcal{R}'^n_i & \text{x'-space} \end{aligned}$$

Now there are *two* affine connections,

$$\begin{aligned} \Gamma^c_{ab} &\equiv (\partial x^c / \partial \xi^n) (\partial^2 \xi^n / \partial x^a \partial x^b) = \mathcal{R}^c_n \partial_a (\partial \xi^n / \partial x^b) = \mathcal{R}^c_n (\partial_a \mathcal{R}_b^n) & \partial_a &= \partial / \partial x^a \\ &= [\mathbf{e}^c]_i \bullet (\partial_a [\mathbf{e}_b]^i) = \mathbf{e}^c \bullet (\partial_a \mathbf{e}_b) \\ \Gamma'^c_{ab} &\equiv (\partial x'^c / \partial \xi^n) (\partial^2 \xi^n / \partial x'^a \partial x'^b) = \mathcal{R}'^c_n \partial'_a (\partial \xi^n / \partial x'^b) = \mathcal{R}'^c_n (\partial'_a \mathcal{R}'_b^n) & \partial'_a &= \partial / \partial x'^a \\ &= [\mathbf{e}'^c]_i \bullet (\partial'_a [\mathbf{e}'_b]^i) = \mathbf{e}'^c \bullet (\partial'_a \mathbf{e}'_b) \end{aligned}$$

(f) Relations between Γ and Γ'

The claimed relations are the following in the context of Picture A shown above,

$$\Gamma'^c_{ab} = R^c_d R_a^\alpha R_b^\beta \Gamma^d_{\alpha\beta} + R^c_\alpha (\partial'_a R_b^\alpha) \quad // \text{Weinberg p 100 (4.5.2)}$$

$$\Gamma'^c_{ab} = R^c_d R_a^\alpha R_b^\beta \Gamma^d_{\alpha\beta} - R_b^\beta (\partial'_a R^c_\beta)$$

$$\Gamma'^c_{ab} = R^c_d R_a^\alpha R_b^\beta \Gamma^d_{\alpha\beta} - R_a^\alpha R_b^\beta (\partial_\alpha R^c_\beta) \quad // \text{Weinberg p 102 (4.5.8)}$$

If the second term were not present, the relation would state that $\Gamma^d_{\alpha\beta}$ transforms as a mixed rank-3 tensor in the usual manner (Section 7 (j)). Since the second term *is* present, $\Gamma^d_{\alpha\beta}$ is not a tensor.

Proof of the first relation : We now make use of Picture D1 shown above. Start with the Γ' definition given above,

$$\begin{aligned} \Gamma'^c_{ab} &\equiv \mathcal{R}'^c_n (\partial'_a \mathcal{R}'_b^n) = (R \mathcal{R})^c_n \partial'_a (R \mathcal{R})_b^n = R^c_d \mathcal{R}^d_n \partial'_a (R_b^\beta \mathcal{R}_\beta^n) & // \mathcal{R}' &= R \mathcal{R} \\ &= R^c_d \mathcal{R}^d_n R_b^\beta (\partial'_a \mathcal{R}_\beta^n) + R^c_d (\mathcal{R}^d_n \mathcal{R}_\beta^n) (\partial'_a R_b^\beta) \\ &= R^c_d \mathcal{R}^d_n R_b^\beta ([R_a^\alpha \partial_\alpha] \mathcal{R}_\beta^n) + R^c_d (\delta^d_\beta) (\partial'_a R_b^\beta) & // \partial'_a &= R_a^\alpha \partial_\alpha \text{ in first term only} \\ &= R^c_d R_a^\alpha R_b^\beta \mathcal{R}^d_n (\partial_\alpha \mathcal{R}_\beta^n) + R^c_\beta (\partial'_a R_b^\beta) \\ &= R^c_d R_a^\alpha R_b^\beta \Gamma^d_{\alpha\beta} + R^c_\alpha (\partial'_a R_b^\alpha) & // \Gamma^d_{\alpha\beta} &\equiv \mathcal{R}^d_n (\partial_\alpha \mathcal{R}_\beta^n) \quad \text{QED} \end{aligned}$$

Magically, all the $\mathcal{R}$'s have gone away.

The second term in the above relation can be written a different manner as follows. Consider,

$$\begin{aligned} 0 &= \partial'_a (\delta^c_b) = \partial'_a (R^c_\alpha R_b^\alpha) = R^c_\alpha (\partial'_a R_b^\alpha) + R_b^\alpha (\partial'_a R^c_\alpha) \\ \Rightarrow R^c_\alpha (\partial'_a R_b^\alpha) &= -R_b^\alpha (\partial'_a R^c_\alpha) = -R_b^\beta (\partial'_a R^c_\beta) \\ &= -R_b^\beta R_a^\alpha (\partial_\alpha R^c_\beta) \end{aligned}$$

and this gives the other two relations stated above.

(g) Statement and Proof of the Covariant Derivative Theorem

Many examples of this theorem will be given later. In this proof it is assumed that the tensor density of interest is purely covariant (all indices "down"). In the next section it will then be shown how to adjust the theorem if one or more of the tensor density indices is "up".

Covariant Derivative Theorem: The covariant derivative ($B_{abc...x;\alpha}$ as defined below) of a covariant tensor density of rank n and weight W ($B_{abc...x}$) transforms as a covariant tensor density of rank $n+1$ and weight W .

The implication is that all the indices including α of $B_{abc...x;\alpha}$ can be treated as ordinary tensor indices with respect to raising, lowering, contraction, and so on. The first term in $B_{abc...x;\alpha}$ is the regular derivative $\partial_\alpha B_{abc...x}$, often written as $B_{abc...x,\alpha}$ (comma, not semicolon), and this first term is *not* a tensor. Only when all the "correction terms" are included does the object become a tensor.

The covariant derivative in x -space and then in x' -space is defined as follows: (Weinberg p 104 4.6.12)

$$\begin{aligned} B_{abc...x;\alpha} &\equiv \underset{\text{del}}{\partial_\alpha B_{abc...x}} - \Gamma_{a\alpha}^n B_{nbc...x} - \Gamma_{b\alpha}^n B_{anc...x} - \dots - \Gamma_{x\alpha}^n B_{abc...n} \quad // \text{ x-space} \\ &\quad + (W/2g) (\partial_\alpha g) B_{abc...x} \\ B'_{abc...x;\alpha} &\equiv \underset{\text{del}}{\partial'_\alpha B'_{abc...x}} - \Gamma'^n_{a\alpha} B'_{nbc...x} - \Gamma'^n_{b\alpha} B'_{anc...x} - \dots - \Gamma'^n_{x\alpha} B'_{abc...n} \quad // \text{ x'-space} \\ &\quad + (W/2g') (\partial'_\alpha g') B'_{abc...x} \end{aligned}$$

The two definitions are the same except everything is primed in x' -space (except constant weight W). This is as one would expect if the x -space equation were a "true tensor equation" as discussed in Section 7 (u) and were therefore "covariant". Although the Γ objects are not tensors themselves, the combination of terms shown in the definition of $B_{abc...x;\alpha}$ is a rank $n+1$ tensor (as will be demonstrated).

As was shown in section (d), $(1/2)(1/g)\partial_\alpha g = \Gamma^{\mathbf{x}}_{\mathbf{x}\alpha}$ so the W terms could be written as $W \Gamma^{\mathbf{x}}_{\mathbf{x}\alpha} B_{abc...x}$ and $W \Gamma'^{\mathbf{x}}_{\mathbf{x}\alpha} B'_{abc...x}$, and this form is commonly seen in the literature on this subject.

A proof of the theorem must then show that $B_{abc...x;\alpha}$ as defined above in fact transforms as a tensor density of rank $n+1$ and weight W , which is to say, one must show that

$$B'_{abc...x;\alpha} = J^{-W} R_A^{a'} R_B^{b'} \dots R_X^{x'} R_\alpha^{\alpha'} B_{a'b'c'...x';\alpha'}$$

or

$$\begin{aligned} B'_{ABC...X;\alpha} \\ = J^{-W} R_\alpha^{\alpha'} \{R_A^{a'} R_B^{b'} \dots R_X^{x'}\} * \\ B_{a'b'c'...x';\alpha'} \end{aligned}$$

or, in gory detail,

$$\begin{aligned} \partial'_\alpha B'_{ABC...X} - \Gamma_{\alpha A}^n B'_{nBC...X} - \Gamma_{\alpha B}^n B'_{AnC...X} - \dots - \Gamma_{\alpha X}^n B'_{ABC...n} \quad // \text{ LHS} \\ \text{del}' \quad \quad \quad \text{a'-term} \quad \quad \quad \text{b'-term} \quad \quad \quad \text{x'-term} \\ + (W/2g') (\partial'_\alpha g') B'_{ABC...X} \\ \\ = J^{-W} R_\alpha^{\alpha'} \{R_A^{a'} R_B^{b'} \dots R_X^{x'}\} * \quad // \text{ RHS} \\ \{ \partial_\alpha B_{a'b'c'...x'} - \Gamma_{\alpha a'}^n B_{nb'c'...x'} - \Gamma_{\alpha b'}^n B_{a'nc'...x'} - \dots - \Gamma_{\alpha x'}^n B_{a'b'c'...n} \\ \text{del} \quad \quad \quad \text{a-term} \quad \quad \quad \text{b-term} \quad \quad \quad \text{x-term} \\ + (W/2g) (\partial_\alpha g) B_{a'b'c'...x'} \} \end{aligned}$$

Proof:

1. Expand the LHS del' term and show del'-del matches the RHS del term.

$$\begin{aligned} \partial'_\alpha B'_{ABC...X} &= (R_\alpha^\beta \partial_\beta) (J^{-W} R_A^a R_B^b \dots R_X^x B_{abc...x}) & \text{del}' \\ &= R_\alpha^\beta (\partial_\beta J^{-W}) R_A^a R_B^b \dots R_X^x B_{abc...x} & \text{del'-J} \\ &= R_\alpha^\beta J^{-W} (\partial_\beta R_A^a) R_B^b \dots R_X^x B_{abc...x} & \text{del'-a} \\ &+ R_\alpha^\beta J^{-W} R_A^a (\partial_\beta R_B^b) \dots R_X^x B_{abc...x} & \text{del'-b} \\ &\dots & \\ &+ R_\alpha^\beta J^{-W} R_A^a R_B^b \dots (\partial_\beta R_X^x) B_{abc...x} & \text{del'-x} \\ &+ R_\alpha^\beta J^{-W} R_A^a R_B^b \dots R_X^x (\partial_\beta B_{abc...x}) & \text{del'-del} \end{aligned}$$

We first claim that this del'-del term of the LHS matches the del term on the RHS:

$$\begin{aligned} \text{del'-del LHS} &= R_\alpha^\beta J^{-W} \{R_A^a R_B^b \dots R_X^x\} (\partial_\beta B_{abc...x}) \\ \text{del RHS} &= J^{-W} R_\alpha^{\alpha'} \{R_A^{a'} R_B^{b'} \dots R_X^{x'}\} \{ \partial_{\alpha'} B_{a'b'c'...x'} \} \end{aligned}$$

Unpriming all the Latin indices and setting $\alpha' = \beta$ shows that these two terms indeed match.

2. Show that the a-related terms balance. The a'-term on the LHS is this

$$- \Gamma_{\mathbf{A}\alpha}^{\mathbf{n}} B'_{\mathbf{n}BC} \dots \mathbf{x}$$

The connection between Γ' and Γ given in section (f) says

$$\Gamma'^c_{ab} = R^c_d R_a^{\mathbf{x}} R_b^{\sigma} \Gamma^d_{\mathbf{x}\sigma} - R_a^{\mathbf{x}} R_b^{\sigma} (\partial_{\mathbf{x}} R^c_{\sigma})$$

or

$$\Gamma'^{\mathbf{n}}_{\mathbf{A}\alpha} = R^n_d R_{\mathbf{A}}^{\mathbf{x}} R_{\alpha}^{\sigma} \Gamma^d_{\mathbf{x}\sigma} - R_{\mathbf{A}}^{\mathbf{x}} R_{\alpha}^{\sigma} (\partial_{\mathbf{x}} R^n_{\sigma}) .$$

Inserting the above for $\Gamma'^{\mathbf{n}}_{\mathbf{A}\alpha}$ and the transformation rule for B' , the LHS a'-term becomes

$$- \{ R^n_d R_{\mathbf{A}}^{\mathbf{x}} R_{\alpha}^{\sigma} \Gamma^d_{\mathbf{x}\sigma} - R_{\mathbf{A}}^{\mathbf{x}} R_{\alpha}^{\sigma} (\partial_{\mathbf{x}} R^n_{\sigma}) \} \{ J^{-\mathbf{w}} R_n^{\mathbf{n}'} R_B^{\mathbf{b}} R_C^{\mathbf{c}} \dots R_{\mathbf{x}}^{\mathbf{x}} B_{\mathbf{n}'}^{\mathbf{b}c} \dots \mathbf{x} \}$$

and to this we must add the contribution called del'-a above.

Meanwhile, the RHS a-term is this

$$\begin{aligned} & J^{-\mathbf{w}} R_{\alpha}^{\alpha'} \{ R_{\mathbf{A}}^{\mathbf{a}'} R_B^{\mathbf{b}'} \dots R_{\mathbf{x}}^{\mathbf{x}'} \} \{ - \Gamma_{\mathbf{a}\alpha'}^{\mathbf{n}} B_{\mathbf{n}b'c'} \dots \mathbf{x} \} \quad // \text{ remove Latin primes, then } n \rightarrow n' \\ & = - \{ J^{-\mathbf{w}} R_{\alpha}^{\alpha'} \{ R_{\mathbf{A}}^{\mathbf{a}} R_B^{\mathbf{b}} \dots R_{\mathbf{x}}^{\mathbf{x}} \} \{ \Gamma_{\mathbf{a}\alpha'}^{\mathbf{n}'} B_{\mathbf{n}'}^{\mathbf{b}c} \dots \mathbf{x} \} \end{aligned}$$

so we have to show that

$$\begin{aligned} & - \{ R^n_d R_{\mathbf{A}}^{\mathbf{x}} R_{\alpha}^{\sigma} \Gamma^d_{\mathbf{x}\sigma} - R_{\mathbf{A}}^{\mathbf{x}} R_{\alpha}^{\sigma} (\partial_{\mathbf{x}} R^n_{\sigma}) \} \{ J^{-\mathbf{w}} R_n^{\mathbf{n}'} R_B^{\mathbf{b}} R_C^{\mathbf{c}} \dots R_{\mathbf{x}}^{\mathbf{x}} B_{\mathbf{n}'}^{\mathbf{b}c} \dots \mathbf{x} \} \\ & \quad + R_{\alpha}^{\beta} J^{-\mathbf{w}} (\partial_{\beta} R_{\mathbf{A}}^{\mathbf{n}'}) R_B^{\mathbf{b}} \dots R_{\mathbf{x}}^{\mathbf{x}} B_{\mathbf{n}'}^{\mathbf{b}c} \dots \mathbf{x} \quad // a \rightarrow n', \text{ this is the del'-a term} \\ & = - \{ J^{-\mathbf{w}} R_{\alpha}^{\alpha'} \{ R_{\mathbf{A}}^{\mathbf{a}} R_B^{\mathbf{b}} \dots R_{\mathbf{x}}^{\mathbf{x}} \} \{ \Gamma_{\mathbf{a}\alpha'}^{\mathbf{n}'} B_{\mathbf{n}'}^{\mathbf{b}c} \dots \mathbf{x} \} \} ? \end{aligned}$$

where now the del'-a term has been added in to the LHS, and in so doing index $a \rightarrow n'$. We can see that the factors $J^{-\mathbf{w}} R_B^{\mathbf{b}} R_C^{\mathbf{c}} \dots R_{\mathbf{x}}^{\mathbf{x}} B_{\mathbf{n}'}^{\mathbf{b}c} \dots \mathbf{x}$ are the same on both sides so they can be removed to give a simpler relation which we must show is valid:

$$\begin{aligned} & - \{ R^n_d R_{\mathbf{A}}^{\mathbf{x}} R_{\alpha}^{\sigma} \Gamma^d_{\mathbf{x}\sigma} - R_{\mathbf{A}}^{\mathbf{x}} R_{\alpha}^{\sigma} (\partial_{\mathbf{x}} R^n_{\sigma}) \} \{ R_n^{\mathbf{n}'} \} \\ & \quad + R_{\alpha}^{\beta} (\partial_{\beta} R_{\mathbf{A}}^{\mathbf{n}'}) \\ & = - \{ R_{\alpha}^{\alpha'} \{ R_{\mathbf{A}}^{\mathbf{a}} \} \{ \Gamma_{\mathbf{a}\alpha'}^{\mathbf{n}'} \} \end{aligned}$$

In the first term one sees $R^n_d R_n^{\mathbf{n}'} = \delta_d^{\mathbf{n}'}$ which pins d to n' in that term only, so the above becomes

$$- R_{\mathbf{A}}^{\mathbf{x}} R_{\alpha}^{\sigma} \Gamma^{\mathbf{n}'}_{\mathbf{x}\sigma} + R_n^{\mathbf{n}'} R_{\mathbf{A}}^{\mathbf{x}} R_{\alpha}^{\sigma} (\partial_{\mathbf{x}} R^n_{\sigma}) + R_{\alpha}^{\beta} (\partial_{\beta} R_{\mathbf{A}}^{\mathbf{n}'}) = - R_{\alpha}^{\sigma} R_{\mathbf{A}}^{\mathbf{x}} \Gamma^{\mathbf{n}'}_{\mathbf{x}\sigma} .$$

The first term on the left cancels the term on the right, and do $\beta \rightarrow \sigma$ in the third term to get

$$R_{\mathbf{A}}^{\mathbf{x}} R_{\alpha}^{\sigma} \{ R_n^{\mathbf{n}'} (\partial_{\mathbf{x}} R^n_{\sigma}) \} + R_{\alpha}^{\sigma} (\partial_{\sigma} R_{\mathbf{A}}^{\mathbf{n}'}) = 0 .$$

Then cancel the common R_α^σ factor and use the symmetry $(\partial_\kappa R_\sigma^n) = (\partial_\sigma R_\kappa^n)$ to get

$$(\partial_\sigma R_A^{n'}) = -R_n^{n'} R_A^{\kappa} (\partial_\sigma R_\kappa^n)$$

Now in this order do $\sigma \rightarrow a$, $A \rightarrow d$, $n \rightarrow e$, $n' \rightarrow n$, $\kappa \rightarrow m$ to get

$$(\partial_a R_d^n) = -R_e^n R_d^m (\partial_a R_m^e).$$

But this is seen to be the third identity of section (b)! By reversing the above sequence of steps, one shows that the three a-related terms in the above LHS = RHS equation balance:

a'-term + del'-a = a-term.

or

$$\begin{aligned} & - \{ R_d^n R_A^{\kappa} R_\alpha^\sigma \Gamma_{\kappa\sigma}^d - R_A^{\kappa} R_\alpha^\sigma (\partial_\kappa R_\sigma^n) \} \{ R_n^{n'} R_B^b R_C^c \dots R_X^{\kappa} B_{n'b c \dots x} \} \\ & + R_\alpha^{\beta} (\partial_\beta R_A^{n'}) R_B^b \dots R_X^{\kappa} B_{n'b c \dots x} \\ & = - \{ R_\alpha^{\alpha'} \{ R_A^a R_B^b \dots R_X^{\kappa} \} \{ \Gamma_{\alpha\alpha'}^{n'} B_{n'b c \dots x} \} \} \end{aligned}$$

In reversing the sequence, one of course adds back in the deleted common factors as well as the various implied index sums. It seems clearer to state the proof this way rather than start with the last equality with no justification and artificially thread backwards to the desired equation. This method is used as well in the next section.

3. Show that the b-related terms balance. In the previous equation, which was shown true, do $A, a \leftrightarrow B, b$ (indices) and $B_{n'b c \dots x} \rightarrow B_{an'c \dots x}$ to get the following known-valid equation :

$$\begin{aligned} & - \{ R_d^n R_B^{\kappa} R_\alpha^\sigma \Gamma_{\kappa\sigma}^d - R_B^{\kappa} R_\alpha^\sigma (\partial_\kappa R_\sigma^n) \} \{ R_n^{n'} R_A^a R_C^c \dots R_X^{\kappa} B_{an'c \dots x} \} \\ & + R_\alpha^{\beta} (\partial_\beta R_B^{n'}) R_A^a \dots R_X^{\kappa} B_{an'c \dots x} \\ & = - \{ R_\alpha^{\alpha'} \{ R_B^b R_A^a \dots R_X^{\kappa} \} \{ \Gamma_{\alpha\alpha'}^{n'} B_{an'c \dots x} \} \} . \end{aligned}$$

The first line is in fact the b'-term $(-\Gamma_{B\alpha}^{n'} B'_{An c \dots x})$, the second line is del'-b, and the RHS is the b-term. This shows that the three b-related terms match LHS = RHS.

4. Similarly, the c,d,...x terms match. Just repeat item 3 above for each extra index.

5. It remains to show that the three terms so far neglected match as well. These terms are

LHS:	$R_\alpha^{\beta} (\partial_\beta J^{-W}) R_A^a R_B^b \dots R_X^{\kappa} B_{abc \dots x}$	// the del'-J term
	$+ (W/2g') (\partial'_\alpha g') B'_{ABC \dots x}$	// the LHS W term
RHS:	$J^{-W} R_\alpha^{\alpha'} \{ R_A^a R_B^b \dots R_X^{\kappa} \} (W/2g) (\partial_\alpha g) B_{a'b'c' \dots x'}$	// the RHS W term

That is, one must show that

$$\begin{aligned}
& R_{\alpha}^{\beta} (\partial_{\beta} J^{-\mathbf{W}}) R_{\mathbf{A}}^{\mathbf{a}} R_{\mathbf{B}}^{\mathbf{b}} \dots R_{\mathbf{X}}^{\mathbf{x}} B_{\mathbf{abc} \dots \mathbf{x}} + (W/2g') (\partial'_{\alpha} g') B'_{\mathbf{ABC} \dots \mathbf{x}} \\
& = J^{-\mathbf{W}} R_{\alpha}^{\alpha'} \{ R_{\mathbf{A}}^{\mathbf{a}'} R_{\mathbf{B}}^{\mathbf{b}'} \dots R_{\mathbf{X}}^{\mathbf{x}'} \} (W/2g) (\partial_{\alpha'} g) B_{\mathbf{a}' \mathbf{b}' \mathbf{c}' \dots \mathbf{x}'} \quad .
\end{aligned}$$

Expanding $B'_{\mathbf{ABC} \dots \mathbf{x}}$ in the second term gives

$$\begin{aligned}
& R_{\alpha}^{\beta} (\partial_{\beta} J^{-\mathbf{W}}) R_{\mathbf{A}}^{\mathbf{a}} R_{\mathbf{B}}^{\mathbf{b}} \dots R_{\mathbf{X}}^{\mathbf{x}} B_{\mathbf{abc} \dots \mathbf{x}} + (W/2g') (\partial'_{\alpha} g') J^{-\mathbf{W}} \{ R_{\mathbf{A}}^{\mathbf{a}'} R_{\mathbf{B}}^{\mathbf{b}'} \dots R_{\mathbf{X}}^{\mathbf{x}'} \} * B_{\mathbf{a}' \mathbf{b}' \mathbf{c}' \dots \mathbf{x}'} \\
& = J^{-\mathbf{W}} R_{\alpha}^{\alpha'} \{ R_{\mathbf{A}}^{\mathbf{a}'} R_{\mathbf{B}}^{\mathbf{b}'} \dots R_{\mathbf{X}}^{\mathbf{x}'} \} (W/2g) (\partial_{\alpha'} g) B_{\mathbf{a}' \mathbf{b}' \mathbf{c}' \dots \mathbf{x}'} \quad .
\end{aligned}$$

Now unprime the primed Latin indices and set $\beta = \alpha'$ in the first term to get

$$\begin{aligned}
& R_{\alpha}^{\alpha'} (\partial_{\alpha'} J^{-\mathbf{W}}) R_{\mathbf{A}}^{\mathbf{a}} R_{\mathbf{B}}^{\mathbf{b}} \dots R_{\mathbf{X}}^{\mathbf{x}} B_{\mathbf{abc} \dots \mathbf{x}} + (W/2g') (\partial'_{\alpha} g') J^{-\mathbf{W}} \{ R_{\mathbf{A}}^{\mathbf{a}} R_{\mathbf{B}}^{\mathbf{b}} \dots R_{\mathbf{X}}^{\mathbf{x}} \} * B_{\mathbf{abc} \dots \mathbf{x}} \\
& = J^{-\mathbf{W}} R_{\alpha}^{\alpha'} \{ R_{\mathbf{A}}^{\mathbf{a}} R_{\mathbf{B}}^{\mathbf{b}} \dots R_{\mathbf{X}}^{\mathbf{x}} \} (W/2g) (\partial_{\alpha'} g) B_{\mathbf{abc} \dots \mathbf{x}} \quad .
\end{aligned}$$

Next, remove all common factors (and associated implied sums) to get

$$R_{\alpha}^{\alpha'} (\partial_{\alpha'} J^{-\mathbf{W}}) + (W/2g') (\partial'_{\alpha} g') J^{-\mathbf{W}} = J^{-\mathbf{W}} R_{\alpha}^{\alpha'} (W/2g) (\partial_{\alpha'} g) \quad .$$

Since $\partial'_{\alpha} = R_{\alpha}^{\alpha'} \partial_{\alpha'}$, this becomes

$$(\partial'_{\alpha} J^{-\mathbf{W}}) + (W/2g') (\partial'_{\alpha} g') J^{-\mathbf{W}} = J^{-\mathbf{W}} (W/2g) (\partial'_{\alpha} g) \quad .$$

Move the second term to the RHS and do the derivative to get

$$(-W) J^{-\mathbf{W}-1} (\partial'_{\alpha} J) = J^{-\mathbf{W}} (W/2) [(\partial'_{\alpha} g)/g - (\partial'_{\alpha} g')/g']$$

so it remains then to show that

$$J^{-1} (\partial'_{\alpha} J) = (1/2) [(\partial'_{\alpha} g')/g' - (\partial'_{\alpha} g)/g]$$

From Section 5 (k) one has $J^2 = g'/g$ and $J = (g'/g)^{1/2}$ so we need to show that

$$\partial'_{\alpha} (g'/g)^{1/2} = (1/2) (g'/g)^{1/2} [(\partial'_{\alpha} g')/g' - (\partial'_{\alpha} g)/g]$$

or

$$(1/2) (g'/g)^{-1/2} \partial'_{\alpha} (g'/g) = (1/2) (g'/g)^{1/2} [(\partial'_{\alpha} g')/g' - (\partial'_{\alpha} g)/g]$$

or

$$\partial'_{\alpha} (g'/g) = (g'/g) [(\partial'_{\alpha} g')/g' - (\partial'_{\alpha} g)/g] = [g(\partial'_{\alpha} g') - g'(\partial'_{\alpha} g)]/g^2. \quad (*)$$

But (finally) evaluation of the LHS of (*) gives

$$[g(\partial'_{\alpha} g') - g'(\partial'_{\alpha} g)]/g^2$$

which shows that equation (*) is valid. Once again, by reversing the above sequence of steps, one shows that the remaining three terms are in balance.

QED

(h) Rule for raising any index on a covariant derivative of a covariant tensor density.

The Rule is stated at the end of this section before the examples.

Consider the general form given in section (g) for a covariant derivative

$$B_{abc...x;\alpha} \equiv \underset{\text{del}}{\partial_\alpha B_{abc...x}} - \underset{\text{a-term}}{\Gamma_{\alpha}^n B_{nbc...x}} - \underset{\text{b-term}}{\Gamma_{\alpha}^n B_{abn...x}} - \dots - \underset{\text{x-term}}{\Gamma_{\alpha}^n B_{abc...n}} \quad // \text{ x-space} \\ + (W/2g) (\partial_\alpha g) B_{abc...x}$$

Notice that there are n indices on $B_{abc...x}$ and there are n corresponding terms on the RHS in addition to the del and W terms. Each index of $B_{abc...x}$ thus has its own "correction term". What happens if one of the indices (say b) on $B_{abc...x}$ is raised? To find out, apply $g^{\beta b}$ to both sides. The effect of doing this is trivial for all terms *except* the del term and the b-term, since b is a regular tensor index on all such terms, so we get

$$B_a^\beta{}_{c...x;\alpha} \equiv g^{\beta b} \underset{\text{del}}{\partial_\alpha B_{abc...x}} - \underset{\text{a-term}}{\Gamma_{\alpha}^n B_{nbc...x}} - \underset{\text{b-term}}{g^{\beta b} \Gamma_{\alpha}^n B_{abn...x}} - \dots - \underset{\text{x-term}}{\Gamma_{\alpha}^n B_{abc...n}} \\ + (W/2g) (\partial_\alpha g) B_a^\beta{}_{c...x}$$

The del term can be written

$$g^{\beta b} (\partial_\alpha B_{abc...x}) = \partial_\alpha (g^{\beta b} B_{abc...x}) - (\partial_\alpha g^{\beta b}) B_{abc...x} \\ = \partial_\alpha B_a^\beta{}_{c...x} - (\partial_\alpha g^{\beta b}) B_{abc...x} .$$

The second term here can be combined with the b-term to give

$$\text{del-extra} + \text{b-term} = -(\partial_\alpha g^{\beta b}) B_{abc...x} - g^{\beta b} \Gamma_{\alpha}^n B_{abn...x} \\ = -(\partial_\alpha g^{\beta n}) B_{anc...x} - g^{\beta b} \Gamma_{\alpha}^n B_{abn...x} \quad // \text{ b} \rightarrow \text{n in first term only} \\ = [-(\partial_\alpha g^{\beta n}) - g^{\beta b} \Gamma_{\alpha}^n] B_{anc...x} .$$

The first identity of section (c) reads

$$(\partial_c g^{ab}) = -[g^{ai} \Gamma_{ci}^b + g^{bi} \Gamma_{ci}^a] \\ \text{or} \\ [-(\partial_c g^{ab}) - g^{ai} \Gamma_{ci}^b] = g^{bi} \Gamma_{ci}^a \quad // \text{ now do } c \rightarrow \alpha, b \rightarrow n, a \rightarrow \beta \\ \text{or} \\ [-(\partial_\alpha g^{\beta n}) - g^{\beta i} \Gamma_{\alpha i}^n] = g^{ni} \Gamma_{\alpha i}^\beta \\ \text{or} \\ [-(\partial_\alpha g^{\beta n}) - g^{\beta b} \Gamma_{\alpha b}^n] = g^{ni} \Gamma_{\alpha i}^\beta .$$

Therefore,

$$\text{del-extra} + \text{b-term} = g^{ni} \Gamma_{\alpha i}^{\beta} B_{anc...x} = \Gamma_{\alpha i}^{\beta} B_a^i c...x = \Gamma_{\alpha n}^{\beta} B_a^n c...x$$

so it has been shown that

$$B_a^{\beta} c...x;_{\alpha} \equiv \partial_{\alpha} B_a^{\beta} c...x - \Gamma_{a\alpha}^n B_n^{\beta} c...x + \Gamma_{\alpha n}^{\beta} B_a^n c...x - \dots - \Gamma_{x\alpha}^n B_a^{\beta} c...n + W \Gamma_{x\alpha}^{\kappa} B_a^{\beta} c...x$$

new b term

Here then is a comparison

$$\begin{aligned} B_{abc...x;_{\alpha}} &\equiv \partial_{\alpha} B_{abc...x} - \Gamma_{a\alpha}^n B_{nbc...x} - \Gamma_{b\alpha}^n B_{anc...x} - \dots - \Gamma_{x\alpha}^n B_{abc...n} + W \Gamma_{x\alpha}^{\kappa} B_{abc...x} \\ B_a^b c...x;_{\alpha} &\equiv \partial_{\alpha} B_a^b c...x - \Gamma_{a\alpha}^n B_n^b c...x + \Gamma_{\alpha n}^b B_a^n c...x - \dots - \Gamma_{x\alpha}^n B_a^b c...n + W \Gamma_{x\alpha}^{\kappa} B_a^b c...x \end{aligned}$$

del a-term b-term x-term W-term

Rule for raising some non-last index q:

- (1) In all terms, raise the corresponding B index.
- (2) in the q correction term, make the replacement $-\Gamma_{q\alpha}^n \rightarrow +\Gamma_{n\alpha}^q (= \Gamma_{\alpha n}^q)$

Corollary: In order to construct the covariant derivative of *any* rank-n tensor density B----- (indices in any positions), write out the above form and use a covariant correction term for each covariant index (such as $-\Gamma_{a\alpha}^n B_n^b c...x$ for the a index above) and use a contravariant correction term for each contravariant index (such as $+\Gamma_{\alpha n}^b B_a^n c...x$ for the b index above).

Rule for raising the last index α : Raising the $;\alpha$ index must be done "manually" so the first term will have $g^{\alpha\alpha'} \partial_{\alpha} = \partial^{\alpha'}$ and all remaining terms will have explicit $g^{\alpha\alpha'}$ factors.

(i) Examples of covariant derivative expressions

It will be assumed that all B objects in the examples are true tensors unless otherwise specified.

Example J = 0 (covariant derivative of a scalar B)

// J is the rank of the B tensor

$$\begin{aligned} B_{;\alpha} &= \partial_{\alpha} B && \text{covariant vector} && // = B_{,\alpha} \\ B^{;\alpha} &= \partial^{\alpha} B && \text{contravariant vector} && // = B^{,\alpha} \end{aligned}$$

The above examples also apply if B is a scalar density of any weight W. Since $\partial_{\alpha} B$ is then a vector density of weight W (the addition rule), one will have $(\partial'_{\alpha} B) = J^{-W} R_{\alpha}^{\alpha'} (\partial_{\alpha} B)$ and no "correction terms" are required.

Example J=1: (covariant derivative of a vector **B**)

$$B_{a;\alpha} = \partial_{\alpha} B_a - \Gamma_{a\alpha}^n B_n \quad \text{covariant rank-2 tensor} \quad // \text{2nd term is symmetric on } a \leftrightarrow \alpha$$

$$B^a{}_{;\alpha} = \partial_{\alpha} B^a + \Gamma_{\alpha n}^a B^n \quad \text{mixed rank-2 tensor}$$

To obtain the other two possibilities, it is necessary to apply the metric tensor $g^{\alpha\beta}$ and $g^{\alpha\beta}\partial_{\beta} = \partial^{\alpha}$,

$$B_a{}^{;\alpha} = \partial^{\alpha} B_a - g^{\alpha\beta}\Gamma_{a\beta}^n B_n \quad \text{mixed rank-2 tensor}$$

$$B^a{}^{;\alpha} = \partial^{\alpha} B^a + g^{\alpha\beta}\Gamma_{\beta n}^a B^n \quad \text{contravariant rank-2 tensor}$$

Example J=2: (covariant derivative of a rank-2 tensor)

$$B_{ab;\alpha} \equiv \partial_{\alpha} B_{ab} - \Gamma_{a\alpha}^n B_{nb} - \Gamma_{b\alpha}^n B_{an} \quad \text{covariant rank-3 tensor}$$

$$B^a{}_{b;\alpha} \equiv \partial_{\alpha} B^a{}_b + \Gamma_{\alpha n}^a B^n{}_b - \Gamma_{b\alpha}^n B^a{}_n \quad \text{etc.}$$

$$B_a{}^b{}_{;\alpha} \equiv \partial_{\alpha} B_a{}^b - \Gamma_{a\alpha}^n B_n{}^b + \Gamma_{\alpha n}^b B_a{}^n$$

$$B^{ab}{}_{;\alpha} \equiv \partial_{\alpha} B^{ab} + \Gamma_{\alpha n}^a B^{nb} + \Gamma_{\alpha n}^b B^{an}.$$

Again, application of $g^{\alpha\beta}$ would give expressions for the other four possibilities with $;\alpha$ being "up". These "other possibilities" are always present, but we shall no longer mention them in the following examples.

Example J=3: (covariant derivative of a rank-3 tensor)

$$B_{abc;\alpha} \equiv \partial_{\alpha} B_{abc} - \Gamma_{a\alpha}^n B_{nbc} - \Gamma_{b\alpha}^n B_{anc} - \Gamma_{c\alpha}^n B_{abn}$$

$$B^a{}_{bc;\alpha} \equiv \partial_{\alpha} B^a{}_{bc} + \Gamma_{\alpha n}^a B^n{}_{bc} - \Gamma_{b\alpha}^n B^a{}_{nc} - \Gamma_{c\alpha}^n B^a{}_{bn}$$

...

$$B^{abc}{}_{;\alpha} \equiv \partial_{\alpha} B^{abc} + \Gamma_{n\alpha}^a B^{nbc} + \Gamma_{n\alpha}^b B^{anc} + \Gamma_{n\alpha}^c B^{abn}$$

Special J=2 application to the metric tensor:

$$g_{ab;\alpha} \equiv \partial_{\alpha} g_{ab} - \Gamma_{a\alpha}^n g_{nb} - \Gamma_{b\alpha}^n g_{an} = 0 \quad \text{by section (c) identity 2}$$

$$g^a{}_{b;\alpha} \equiv \partial_{\alpha} g^a{}_b + \Gamma_{\alpha n}^a g^n{}_b - \Gamma_{b\alpha}^n g^a{}_n = 0 + \Gamma_{\alpha b}^a - \Gamma_{b\alpha}^a = 0$$

$$g_a{}^b{}_{;\alpha} \equiv \partial_{\alpha} g_a{}^b - \Gamma_{a\alpha}^n g_n{}^b + \Gamma_{\alpha n}^b g_a{}^n = 0 - \Gamma_{a\alpha}^b + \Gamma_{\alpha a}^b = 0$$

$$g^{ab}{}_{;\alpha} \equiv \partial_{\alpha} g^{ab} + \Gamma_{\alpha n}^a g^{nb} + \Gamma_{\alpha n}^b g^{an} = 0 \quad \text{by section (c) identity 1}$$

The middle lines use the fact that $g^i_j = \delta^i_j$. Since $g_{ab;\alpha}$ is a tensor, knowing that any one of the above vanishes implies that all four lines vanish! The net result is

$$g_{ab;\alpha} = g^a_b{}_{;\alpha} = g^b_a{}_{;\alpha} = g^{ab}{}_{;\alpha} = 0 \quad . \quad // \text{ Weinberg p 105 (4.6.16,17,18)}$$

The covariant derivative of any form of the metric tensor vanishes. As Weinberg points out, one knows that in a quasi-Cartesian x-space $g_{ab;\alpha} = 0$ since $g_{ab} = G_{aa}\delta_{a,b}$ and $\Gamma = 0$. Then in any x'-space $g'_{ab;\alpha} = 0$ as well since $g'_{ab;\alpha} = R_a{}^{a'} R_b{}^{b'} R_\alpha{}^{\alpha'} g_{a'b'}{}_{;\alpha'}$.

Example J=2: (double covariant derivatives)

Consider again the J=1 examples from above

$$\begin{aligned} B_{a;\alpha} &= \partial_\alpha B_a - \Gamma_{a\alpha}^n B_n \\ B^a{}_{;\alpha} &= \partial_\alpha B^a + \Gamma_{\alpha n}^a B^n \end{aligned} \quad .$$

This applies to *any* vector B_a . As the J=0 example shows, $B_{;a}$ and $B^{;a}$ are bona-fide vectors (covariant and contravariant components of the same vector) and therefore

$$\begin{aligned} B_{;a;\alpha} &= \partial_\alpha B_{;a} - \Gamma_{a\alpha}^n B_{;n} \\ B^{;a}{}_{;\alpha} &= \partial_\alpha B^{;a} + \Gamma_{\alpha n}^a B^{;n} \end{aligned}$$

where we have simply inserted a semicolon in each term.

Consider again the J=2 examples from above,

$$\begin{aligned} B_{ab;\alpha} &= \partial_\alpha B_{ab} - \Gamma_{a\alpha}^n B_{nb} - \Gamma_{b\alpha}^n B_{an} \\ B^a{}_{b;\alpha} &= \partial_\alpha B^a{}_b + \Gamma_{\alpha n}^a B^n{}_b - \Gamma_{b\alpha}^n B^a{}_n \end{aligned}$$

This applies to any rank-2 tensors B_{ab} or $B^a{}_b$. According to sections (i) and (h) above, $B_{a;b}$ and $B^a{}_{;b}$ are bona-fide rank-2 tensors, and therefore

$$\begin{aligned} B_{a;b;\alpha} &= \partial_\alpha B_{a;b} - \Gamma_{a\alpha}^n B_{n;b} - \Gamma_{b\alpha}^n B_{a;n} \\ B^a{}_{;b;\alpha} &= \partial_\alpha B^a{}_{;b} + \Gamma_{\alpha n}^a B^n{}_{;b} - \Gamma_{b\alpha}^n B^a{}_{;n} \end{aligned}$$

In a similar manner one can derive expressions for triple covariant derivatives and beyond. For example

$$B_{a;b;c;\alpha} = \partial_\alpha B_{a;b;c} - \Gamma_{a\alpha}^n B_{n;b;c} - \Gamma_{b\alpha}^n B_{a;n;c} - \Gamma_{c\alpha}^n B_{a;b;n} \quad .$$

The next examples are for **tensor densities**: (The trivial J=0 cases were already considered above.)

Example J = 1 (vector density of weight W)

// recall $\Gamma^{\kappa}_{\kappa\alpha} = (1/2g)\partial_{\alpha}g$

$$\begin{aligned} B_{a;\alpha} &= \partial_{\alpha} B_a - \Gamma^{\mathbf{n}}_{a\alpha} B_{\mathbf{n}} + W \Gamma^{\kappa}_{\kappa\alpha} B_a \\ B^{\mathbf{a}}_{;\alpha} &= \partial_{\alpha} B^{\mathbf{a}} + \Gamma^{\mathbf{a}}_{\alpha\mathbf{n}} B^{\mathbf{n}} + W \Gamma^{\kappa}_{\kappa\alpha} B^{\mathbf{a}} \end{aligned}$$

covariant rank-2 tensor density

Example J = 2 (rank-2 tensor density of weight W)

$$\begin{aligned} B_{ab;\alpha} &= \partial_{\alpha} B_{ab} - \Gamma^{\mathbf{n}}_{a\alpha} B_{\mathbf{n}b} - \Gamma^{\mathbf{n}}_{b\alpha} B_{a\mathbf{n}} + W \Gamma^{\kappa}_{\kappa\alpha} B_{ab} \\ B^{\mathbf{a}}_{\mathbf{b};\alpha} &= \partial_{\alpha} B^{\mathbf{a}}_{\mathbf{b}} + \Gamma^{\mathbf{a}}_{\alpha\mathbf{n}} B^{\mathbf{n}}_{\mathbf{b}} - \Gamma^{\mathbf{n}}_{b\alpha} B^{\mathbf{a}}_{\mathbf{n}} + W \Gamma^{\kappa}_{\kappa\alpha} B^{\mathbf{a}}_{\mathbf{b}} \\ B^{\mathbf{b}}_{\mathbf{a};\alpha} &= \partial_{\alpha} B^{\mathbf{b}}_{\mathbf{a}} - \Gamma^{\mathbf{n}}_{a\alpha} B^{\mathbf{b}}_{\mathbf{n}} + \Gamma^{\mathbf{b}}_{\alpha\mathbf{n}} B^{\mathbf{n}}_{\mathbf{a}} + W \Gamma^{\kappa}_{\kappa\alpha} B^{\mathbf{b}}_{\mathbf{a}} \\ B^{ab;\alpha} &= \partial_{\alpha} B^{ab} + \Gamma^{\mathbf{a}}_{\alpha\mathbf{n}} B^{nb} + \Gamma^{\mathbf{b}}_{\alpha\mathbf{n}} B^{an} + W \Gamma^{\kappa}_{\kappa\alpha} B^{ab} \end{aligned}$$

covariant rank-3 tensor density

As noted in Appendix D (b) item 3 Example 2, adding a factor $g^{W/2}$ to a tensor density of weight W neutralizes the weight, and the result is a regular tensor. Here then are a few examples in which this is done. Since the product is a tensor, there are no W correction terms.

Example J = 0 (covariant derivative of a scalar density B of weight W)

$$\begin{aligned} (g^{W/2}B)_{;\alpha} &= \partial_{\alpha}(g^{W/2}B) && \text{covariant vector} \\ (g^{W/2}B)^{;\alpha} &= \partial^{\alpha}(g^{W/2}B) && \text{contravariant vector} \end{aligned}$$

Example J=1: (covariant derivative of a vector density B of weight W)

$$\begin{aligned} (g^{W/2}B_a)_{;\alpha} &= \partial_{\alpha}(g^{W/2}B_a) - g^{W/2} \Gamma^{\mathbf{n}}_{a\alpha} B_{\mathbf{n}} && \text{covariant rank-2 tensor} \\ (g^{W/2}B^{\mathbf{a}})_{;\alpha} &= \partial_{\alpha}(g^{W/2}B^{\mathbf{a}}) + g^{W/2} \Gamma^{\mathbf{a}}_{\alpha\mathbf{n}} B^{\mathbf{n}} && \text{mixed rank-2 tensor} \end{aligned}$$

Example J=2: (covariant derivative of a tensor density B of weight W)

$$\begin{aligned} (g^{W/2}B_{ab})_{;\alpha} &= \partial_{\alpha}(g^{W/2}B_{ab}) - \Gamma^{\mathbf{n}}_{a\alpha}(g^{W/2}B_{\mathbf{n}b}) - \Gamma^{\mathbf{n}}_{b\alpha}(g^{W/2}B_{a\mathbf{n}}) && \text{covariant rank-3 tensor} \\ (g^{W/2}B^{\mathbf{a}}_{\mathbf{b}})_{;\alpha} &= \partial_{\alpha}(g^{W/2}B^{\mathbf{a}}_{\mathbf{b}}) + \Gamma^{\mathbf{a}}_{\alpha\mathbf{n}}(g^{W/2}B^{\mathbf{n}}_{\mathbf{b}}) - \Gamma^{\mathbf{n}}_{b\alpha}(g^{W/2}B^{\mathbf{a}}_{\mathbf{n}}) && \text{etc.} \\ (g^{W/2}B^{\mathbf{b}}_{\mathbf{a}})_{;\alpha} &= \partial_{\alpha}(g^{W/2}B^{\mathbf{b}}_{\mathbf{a}}) - \Gamma^{\mathbf{n}}_{a\alpha}(g^{W/2}B^{\mathbf{b}}_{\mathbf{n}}) + \Gamma^{\mathbf{b}}_{\alpha\mathbf{n}}(g^{W/2}B^{\mathbf{n}}_{\mathbf{a}}) \\ (g^{W/2}B^{ab})_{;\alpha} &= \partial_{\alpha}(g^{W/2}B^{ab}) + \Gamma^{\mathbf{a}}_{\alpha\mathbf{n}}(g^{W/2}B^{nb}) + \Gamma^{\mathbf{b}}_{\alpha\mathbf{n}}(g^{W/2}B^{an}) \end{aligned}$$

(j) The Leibniz rule for the covariant derivative of the product of two tensor densities

If A and B are arbitrary tensor densities each with an arbitrary set of up and down indices and arbitrary weight, then the *claim* of the product rule is this:

$$(A^{----}B^{----})_{;\alpha} \equiv A^{----}_{;\alpha} B^{----} + A^{----} B^{----}_{;\alpha} \quad // \text{Weinberg p 105 (4.6.14)} \quad (*)$$

Recall from the Covariant Derivative Theorem of section (g) above that A^{----} and $A^{----}_{;\alpha}$ have the same weight, call it W_A . Similarly, B^{----} and $B^{----}_{;\alpha}$ have the same weight W_B . According to the outer product rule of Appendix D (b) item 3, both terms on the RHS above have weight $W_A + W_B$ and therefore this sum is the weight of the LHS $(A^{----}B^{----})_{;\alpha}$ as well.

Proof: Start with

$$A^{----};_{\alpha} = A^{----},_{\alpha} + (\text{A index correction terms}) + W_A \Gamma^{\kappa}_{\kappa\alpha} A^{----}$$

$$B^{----};_{\alpha} = B^{----},_{\alpha} + (\text{B index correction terms}) + W_B \Gamma^{\kappa}_{\kappa\alpha} B^{----}$$

The "index correction terms" are those Γ terms discussed in the previous sections. One can then write out the two terms on the RHS of (*) above as:

$$A^{----};_{\alpha} B^{----} = A^{----},_{\alpha} B^{----} + (\text{A index correction terms}) B^{----} + [W_A \Gamma^{\kappa}_{\kappa\alpha} A^{----}] B^{----}$$

$$A^{----} B^{----};_{\alpha} = A^{----} B^{----},_{\alpha} + A^{----} (\text{B index correction terms}) + A^{----} [W_B \Gamma^{\kappa}_{\kappa\alpha} B^{----}]$$

Meanwhile, the LHS of (*) can be written as

$$(A^{----}B^{----});_{\alpha} = (A^{----}B^{----}),_{\alpha} + (\text{all index correction terms}) + (W_A + W_B) \Gamma^{\kappa}_{\kappa\alpha} (A^{----}B^{----})$$

Momentarily ignoring the index correction terms, it is clear that the other terms match between LHS and RHS. The W terms match by visual inspection, while the regular derivative terms match due to the "regular" Leibniz rule for the derivative of a product

$$(A^{----}B^{----}),_{\alpha} = A^{----},_{\alpha} B^{----} + A^{----} B^{----},_{\alpha}$$

which is to say

$$\partial_{\alpha}(A^{----}B^{----}) = (\partial_{\alpha}A^{----})B^{----} + A^{----} (\partial_{\alpha} B^{----}) .$$

Consider now the LHS terms called (all index correction terms) above. This set of terms can be partitioned into two groups,

$$(\text{all index correction terms}) = (\text{terms involving A indices}) + (\text{terms involving B indices}) .$$

Let us pause to look at a simple example where ICT means we just show the index correction terms,

$$(A_{ab}B^{cd});_{\alpha}|^{ICT} = - \Gamma^n_{a\alpha}(A_{nb}B^{cd}) - \Gamma^n_{b\alpha}(A_{an}B^{cd}) + \Gamma^c_{\alpha n}(A_{ab}B^{nd}) + \Gamma^d_{\alpha n}(A_{ab}B^{cn}) .$$

The correction terms can be reordered in this way

$$\begin{aligned} &= \{ - \Gamma^n_{a\alpha}(A_{nb}) - \Gamma^n_{b\alpha}(A_{an}) \} B^{cd} + A_{ab} \{ \Gamma^c_{\alpha n}B^{nd} + \Gamma^d_{\alpha n}B^{cn} \} \\ &= \{ \text{index correction terms for } A_{ab} \} B^{cd} + A_{ab} \{ \text{index correction terms for } B^{cd} \} \end{aligned}$$

Just so, in the general case one has

$$(A----B----)_{;\alpha}^{ICT} = (\text{all index correction terms})$$

$$= \{ \text{index correction terms for } A---- \} B---- + A---- \{ \text{index correction terms for } B---- \}$$

and this then shows that the index correction terms on the two sides of (*) do in fact match. **QED**

Once the above product rule is verified, it is then easy to generalize just as for regular derivatives:

$$(A----B----C----)_{;\alpha} \equiv A----_{;\alpha} B---- C---- + A---- B----_{;\alpha} C---- + A---- B---- C----_{;\alpha}$$

Examples with two vectors:

$$(A_a B_b)_{;\alpha} = A_{a;\alpha} B_b + A_a B_{b;\alpha}$$

$$(A^a B_b)_{;\alpha} = A^a_{;\alpha} B_b + A^a B_{b;\alpha}$$

$$(A^a B^b)_{;\alpha} = A^a_{;\alpha} B^b + A^a B^{b;\alpha}$$

$$(A_a B^b)_{;\alpha} = A_{a;\alpha} B^b + A_a B^{b;\alpha}$$

Example with a scalar function A and a vector B:

$$(AB_b)_{;\alpha} = A_{;\alpha} B_b + AB_{b;\alpha} = A_{,\alpha} B_b + AB_{b;\alpha} \quad // \text{ for a scalar } A, A_{;\alpha} = A_{,\alpha}$$

Example with a scalar constant A and a vector B:

$$(AB_b)_{;\alpha} = A(B_{b;\alpha}) \quad // \text{ since } A_{;\alpha} = A_{,\alpha} = 0$$

so a scalar constant can always be extracted from $(AB_b)_{;n}$ to give $A(B_{b;n})$. A *tensor* constant like ϵ^{abc} cannot be extracted in this manner since $\epsilon^{abc}_{;\alpha} \neq 0$. In fact

$$\begin{aligned} \epsilon^{abc}_{;\alpha} &= \Gamma^a_{n\alpha} \epsilon^{nbc} + \Gamma^b_{n\alpha} \epsilon^{anc} + \Gamma^c_{n\alpha} \epsilon^{abn} \\ \epsilon^{123}_{;\alpha} &= \Gamma^1_{n\alpha} \epsilon^{n23} + \Gamma^2_{n\alpha} \epsilon^{1n3} + \Gamma^3_{n\alpha} \epsilon^{12n} = \Gamma^1_{1\alpha} + \Gamma^2_{2\alpha} + \Gamma^3_{3\alpha} \end{aligned}$$

A more general example:

$$(A_{ab}{}^c B_{de})_{;\alpha} \equiv A_{ab}{}^c_{;\alpha} B_{de} + A_{ab}{}^c B_{de;\alpha}$$

An example with the metric tensor:

$$(g^{ab} B_{----b----})_{;\alpha} = g^{ab}_{;\alpha} B_{----b----} + g^{ab} B_{----b----;\alpha}$$

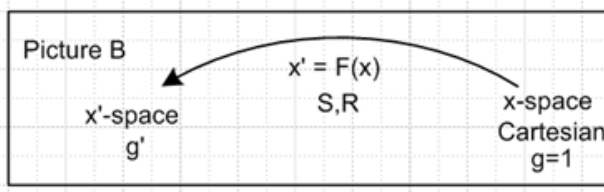
But $g^{ab}_{;\alpha} = 0$ as shown at the end of section (h). Therefore,

$$(B_{----a----})_{;\alpha} = g^{ab} B_{----b----;\alpha}$$

which says that raising an index "commutes" with covariant differentiation -- one can raise an index ignoring the fact that α is sitting there. But we already know this must be true because we know that the object $B^{a\alpha}_{b\gamma}$ is a true tensor, and g^{ab} can raise any index on a true tensor.

Appendix G: Expansion of $(\nabla \mathbf{v})$ in curvilinear coordinates ($\mathbf{v}$ = vector)

This appendix assumes the usual curvilinear coordinates context, Picture B



(a) Continuum Mechanics motivation

Although the polyadic notation is regarded as archaic by some writers (eg, Wolfram), it is well embedded into the literature of continuum mechanics, a field awash in tensors.

In the literature one sometimes sees, in Cartesian coordinates,

$$(\nabla \mathbf{A})_{ij} \equiv \partial_j A_i \quad \equiv A_{i,j}$$

where the indices are the reverse of the normal dyadic definition of Appendix E,

$$(\mathbf{B}\mathbf{A})_{ij} \equiv B_i A_j .$$

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)),

$$D\mathbf{A}_i/Dt = \partial_t A_i + \nabla \mathbf{A}_i \bullet \mathbf{v} = \partial_t A_i + (\partial_j A_i) v_j = \partial_t A_i + (\nabla \mathbf{A})_{ij} v_j = \partial_t A_i + [(\nabla \mathbf{A}) \mathbf{v}]_i$$

$$\Rightarrow D\mathbf{A}/Dt = \partial_t \mathbf{A} + (\nabla \mathbf{A}) \mathbf{v} \quad // = \partial_t \mathbf{A} + (\mathbf{v} \bullet \nabla) \mathbf{A} = (\partial_t + \mathbf{v} \bullet \nabla) \mathbf{A} = D/Dt (\mathbf{A}).$$

$D\mathbf{A}_i/Dt$ is a historical notation for the total derivative $d\mathbf{A}_i(\mathbf{x}, t)/dt$. Here $\mathbf{v}(\mathbf{x}, t)$ is the velocity field of the moving matter blob. An example is acceleration $\mathbf{a}$, where $\mathbf{A} = \mathbf{v}$,

$$\mathbf{a} = D\mathbf{v}/Dt = \partial_t \mathbf{v} + (\nabla \mathbf{v}) \mathbf{v} \quad // = \partial_t \mathbf{v} + (\mathbf{v} \bullet \nabla) \mathbf{v} = (\partial_t + \mathbf{v} \bullet \nabla) \mathbf{v} = D/Dt (\mathbf{v}).$$

The object $(\nabla \mathbf{v})$, called the velocity gradient, is of great interest in fluid mechanics. Correspondingly, the object $(\nabla \mathbf{u})$ is of great interest in the theory of elastic solids, where $\mathbf{u}$ is the displacement field.

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. This is done in the following sections, where we replace the generic vector $\mathbf{A}$ by generic vector $\mathbf{v}$ and this $\mathbf{v}$ has nothing to do with the velocity $\mathbf{v}$ mentioned above.

(b) Expansion of $\nabla \mathbf{v}$ on $\mathbf{e}^i \otimes \mathbf{e}^j$ by Method 1: Use fact that $\mathbf{v}_{b;a}$ is a tensor.

The covariant derivative $v_{\mathbf{b};\mathbf{a}}$ is discussed in Appendix F (g,h,i). We shall define a true tensor object $(\nabla \mathbf{v})$ as $v_{\mathbf{b};\mathbf{a}}$ so that

$$\begin{aligned} (\nabla \mathbf{v})_{\mathbf{ba}} &\equiv v_{\mathbf{b};\mathbf{a}} = [v_{\mathbf{b},\mathbf{a}} - \Gamma_{\mathbf{ab}}^{\mathbf{c}} v_{\mathbf{c}}] & // v_{\mathbf{b},\mathbf{a}} &\equiv \partial_{\mathbf{a}} v_{\mathbf{b}} \\ (\nabla \mathbf{v})'_{\mathbf{ba}} &\equiv v'_{\mathbf{b};\mathbf{a}} = [v'_{\mathbf{b},\mathbf{a}} - \Gamma'^{\mathbf{c}}_{\mathbf{ab}} v'_{\mathbf{c}}] & // v'_{\mathbf{b},\mathbf{a}} &\equiv \partial'_{\mathbf{a}} v'_{\mathbf{b}} \end{aligned}$$

In Cartesian space $\Gamma = 0$ so one has

$$(\nabla \mathbf{v})_{\mathbf{ba}} = v_{\mathbf{b},\mathbf{a}} = \partial_{\mathbf{a}} v_{\mathbf{b}}$$

so $(\nabla \mathbf{v})_{\mathbf{ba}}$ aligns with our object of interest in Cartesian space.

As shown in Appendix E (b) one can expand the rank-2 tensor $v_{\mathbf{b};\mathbf{a}}$ in either of these ways,

$$\begin{aligned} \nabla \mathbf{v} &= \sum_{\mathbf{i},\mathbf{j}} v_{\mathbf{i};\mathbf{j}} \mathbf{u}^{\mathbf{i}} \otimes \mathbf{u}^{\mathbf{j}} & v_{\mathbf{i};\mathbf{j}} &= [v_{\mathbf{i},\mathbf{j}} - \Gamma_{\mathbf{ij}}^{\mathbf{c}} v_{\mathbf{c}}] = v_{\mathbf{i},\mathbf{j}} = \partial_{\mathbf{j}} v_{\mathbf{i}} \\ \nabla \mathbf{v} &= \sum_{\mathbf{i},\mathbf{j}} v'_{\mathbf{i};\mathbf{j}} \mathbf{e}^{\mathbf{i}} \otimes \mathbf{e}^{\mathbf{j}} & v'_{\mathbf{i};\mathbf{j}} &= [v'_{\mathbf{i},\mathbf{j}} - \Gamma'^{\mathbf{c}}_{\mathbf{ij}} v'_{\mathbf{c}}] \end{aligned}$$

where the $\mathbf{u}^{\mathbf{i}}$ are the Cartesian basis vectors in x-space, while $\mathbf{e}^{\mathbf{i}}$ are the reciprocal base vectors.

According to Appendix F (e), the affine connection Γ' in x'-space is given by $\Gamma'^{\mathbf{c}}_{\mathbf{ab}} = \mathcal{R}'^{\mathbf{c}}_{\mathbf{n}} (\partial'_{\mathbf{a}} \mathcal{R}'^{\mathbf{n}}_{\mathbf{b}})$.

When x-space is Cartesian, $\mathcal{R} = 1$ and then $\mathcal{R}' = \mathbf{R}$, so

$$\begin{aligned} \Gamma'^{\mathbf{c}}_{\mathbf{ab}} &= \mathbf{R}^{\mathbf{c}}_{\mathbf{i}} (\partial'_{\mathbf{a}} \mathbf{R}^{\mathbf{i}}_{\mathbf{b}}) \\ \Gamma'^{\mathbf{c}}_{\mathbf{ij}} &= \mathbf{R}^{\mathbf{c}}_{\mathbf{k}} (\partial'_{\mathbf{j}} \mathbf{R}^{\mathbf{k}}_{\mathbf{i}}) . \end{aligned}$$

Inserting this into our expansions above gives

$$\begin{aligned} \nabla \mathbf{v} &= \sum_{\mathbf{i},\mathbf{j}} v'_{\mathbf{i};\mathbf{j}} \mathbf{e}^{\mathbf{i}} \otimes \mathbf{e}^{\mathbf{j}} & v'_{\mathbf{i};\mathbf{j}} &= [(\partial'_{\mathbf{j}} v'_{\mathbf{i}}) - \mathbf{R}^{\mathbf{c}}_{\mathbf{k}} (\partial'_{\mathbf{j}} \mathbf{R}^{\mathbf{k}}_{\mathbf{i}}) v'_{\mathbf{c}}] \\ \text{or} & & & \\ \nabla \mathbf{v} &= \sum_{\mathbf{i},\mathbf{j}} [(\nabla \mathbf{v})^{(\mathbf{e})}]_{\mathbf{ij}} \mathbf{e}^{\mathbf{i}} \otimes \mathbf{e}^{\mathbf{j}} & [(\nabla \mathbf{v})^{(\mathbf{e})}]_{\mathbf{ij}} &= [(\partial'_{\mathbf{j}} v'_{\mathbf{i}}) - \mathbf{R}^{\mathbf{c}}_{\mathbf{k}} (\partial'_{\mathbf{j}} \mathbf{R}^{\mathbf{k}}_{\mathbf{i}}) v'_{\mathbf{c}}] \end{aligned}$$

Note that the expression shown contains only x'-space coordinates and objects. The (e) superscript tells us that this matrix element of operator $(\nabla \mathbf{v})$ is taken in the $\mathbf{e}_{\mathbf{n}}$ basis, see Appendix E (g):

$$[(\nabla \mathbf{v})^{(\mathbf{e})}]_{\mathbf{ij}} = \langle \mathbf{e}_{\mathbf{i}} | \nabla \mathbf{v} | \mathbf{e}_{\mathbf{j}} \rangle = \underbrace{(\mathbf{e}_{\mathbf{i}})^{\mathbf{T}}}_{\text{bra-ket}} \underbrace{(\nabla \mathbf{v})}_{\text{matrix}} \underbrace{\mathbf{e}_{\mathbf{j}}}_{\text{dot of vectors}} = \mathbf{e}_{\mathbf{i}} \bullet \underbrace{(\nabla \mathbf{v}) \bullet \mathbf{e}_{\mathbf{j}}}_{\text{dyadic}}$$

An alternate form is obtained using the identities of Appendix F (b) (adjusted from Picture C1 to Picture D1 shown in that Appendix),

$$\begin{aligned} (\partial'_{\mathbf{a}} \mathbf{R}^{\mathbf{n}}_{\mathbf{d}}) &= -\mathbf{R}^{\mathbf{n}}_{\mathbf{e}} \mathbf{R}^{\mathbf{m}}_{\mathbf{d}} (\partial'_{\mathbf{a}} \mathbf{R}^{\mathbf{e}}_{\mathbf{m}}) \\ (\partial'_{\mathbf{j}} \mathbf{R}^{\mathbf{k}}_{\mathbf{i}}) &= -\mathbf{R}^{\mathbf{k}}_{\mathbf{e}} \mathbf{R}^{\mathbf{m}}_{\mathbf{i}} (\partial'_{\mathbf{j}} \mathbf{R}^{\mathbf{e}}_{\mathbf{m}}) \end{aligned}$$

Then

$$\mathbf{R}^{\mathbf{c}}_{\mathbf{k}} (\partial'_{\mathbf{j}} \mathbf{R}^{\mathbf{k}}_{\mathbf{i}}) = -\mathbf{R}^{\mathbf{c}}_{\mathbf{k}} \mathbf{R}^{\mathbf{k}}_{\mathbf{e}} \mathbf{R}^{\mathbf{m}}_{\mathbf{i}} (\partial'_{\mathbf{j}} \mathbf{R}^{\mathbf{e}}_{\mathbf{m}}) = -\delta^{\mathbf{c}}_{\mathbf{e}} \mathbf{R}^{\mathbf{m}}_{\mathbf{i}} (\partial'_{\mathbf{j}} \mathbf{R}^{\mathbf{e}}_{\mathbf{m}}) = -\mathbf{R}^{\mathbf{m}}_{\mathbf{i}} (\partial'_{\mathbf{j}} \mathbf{R}^{\mathbf{c}}_{\mathbf{m}})$$

so that

$$\nabla \mathbf{v} = \Sigma_{ij} [(\nabla \mathbf{v})^{(e)}]_{ij} \mathbf{e}^i \otimes \mathbf{e}^j \quad [(\nabla \mathbf{v})^{(e)}]_{ij} = v'_{i;j} = [(\partial'_j v^i) + R^m_{ij} (\partial'_j R^c_m) v^c] .$$

We shall use this second form for $[(\nabla \mathbf{v})^{(e)}]_{ij}$ below.

Comment 1: A variation of the above development would be to start this way

$$\begin{aligned} \nabla \mathbf{v} &= \Sigma_{ij} v^i_{;j} \mathbf{u}_i \otimes \mathbf{u}^j & v^i_{;j} &= [v^i_{,j} + \Gamma^i_{jc} v^c] = v^i_{,j} = \partial_j v^i \\ \nabla \mathbf{v} &= \Sigma_{ij} v'^i_{;j} \mathbf{e}_i \otimes \mathbf{e}^j & v'^i_{;j} &= [v'^i_{,j} + \Gamma'^i_{jc} v'^c] \end{aligned}$$

which quickly leads to a result similar to the above, where

$$\nabla \mathbf{v} = \Sigma_{ij} [(\nabla \mathbf{v})^{(e)}]^i_j \mathbf{e}_i \otimes \mathbf{e}^j \quad [(\nabla \mathbf{v})^{(e)}]^i_j = v'^i_{;j} = [(\partial'_j v'^i) - R^m_{jc} (\partial'_j R^i_m) v'^c] .$$

Comment 2: The idea that $[(\nabla \mathbf{v})^{(e)}]^i_j = [\partial'_j v'^i + \Gamma'^i_{jc} v'^c]$ can be reached by this alternate path:

$$dv^i = (\partial_j v^i) dx^j \text{ (chain rule)} \Rightarrow d\mathbf{v} = (\nabla \mathbf{v}) d\mathbf{x}$$

$$\text{Expand: } d\mathbf{x} = dx'^i \mathbf{e}_i \text{ and } \mathbf{v} = v'^n \mathbf{e}_n \Rightarrow d\mathbf{v} = dv'^n \mathbf{e}_n + v'^n d\mathbf{e}_n ,$$

but

$$d\mathbf{e}_n = (\partial'_j \mathbf{e}_n) dx'^j = \Gamma'^i_{jn} \mathbf{e}_i dx'^j \quad // \text{ from definition of } \Gamma \text{ in Appendix F (a), but Picture A}$$

so

$$d\mathbf{v} = dv'^n \mathbf{e}_n + v'^n \Gamma'^i_{jn} dx'^j \mathbf{e}_i = dv'^i \mathbf{e}_i + v'^n \Gamma'^i_{jn} dx'^j \mathbf{e}_i = (dv'^i + v'^a \Gamma'^i_{ja} dx'^j) \mathbf{e}_i .$$

But

$$dv'^i = (\partial v'^i / \partial x'^j) dx'^j = (\partial'_j v'^i) dx'^j$$

so

$$d\mathbf{v} = (\partial'_j v'^i) dx'^j + v'^a \Gamma'^i_{ja} dx'^j \mathbf{e}_i = [\partial'_j v'^i + v'^a \Gamma'^i_{ja}] dx'^j \mathbf{e}_i .$$

Then

$$d\mathbf{v} = (\nabla \mathbf{v}) d\mathbf{x} \Rightarrow [\partial'_j v'^i + v'^a \Gamma'^i_{ja}] dx'^j \mathbf{e}_i = (\nabla \mathbf{v}) dx'^j \mathbf{e}_j$$

or

$$[\partial'_j v'^i + v'^a \Gamma'^i_{ja}] \mathbf{e}_i = (\nabla \mathbf{v}) \mathbf{e}_j$$

$$\Rightarrow \mathbf{e}^i \bullet (\nabla \mathbf{v}) \mathbf{e}_j = \mathbf{e}^i \bullet [\partial'_j v'^i + v'^a \Gamma'^i_{ja}] \mathbf{e}_i = (\partial'_j v'^i + v'^a \Gamma'^i_{ja})$$

$$\Rightarrow [(\nabla \mathbf{v})^{(e)}]^i_j = \mathbf{e}^i \bullet (\nabla \mathbf{v}) \mathbf{e}_j = [\partial'_j v'^i + \Gamma'^i_{jc} v'^c]$$

(c) Expansion of $\nabla \mathbf{v}$ on $\mathbf{e}^i \otimes \mathbf{e}^j$ by Method 2: Use brute force.

Method 1 is perhaps elegant in that it makes use of tensor transformations and the affine connection Γ . But a simple brute force method is really just as simple and does not require knowledge of Γ and covariant differentiation. Instead of using the $\mathbf{e}^i \otimes \mathbf{e}^j$ notation for basis vectors, here we use the alternate

notation $\mathbf{e}^i(\mathbf{e}^j)^T$ which works for rank-2 tensor expansions. Recall that $\mathbf{e}^i(\mathbf{e}^j)^T$ is an $N \times N$ matrix as discussed in Appendix E (e).

In this brute force method, start with the Cartesian space expansion of Appendix E,

$$(\nabla \mathbf{v}) = \Sigma_{cd} (\nabla \mathbf{v})_{dc} \mathbf{u}^d \mathbf{u}^{cT} = \Sigma_{cd} (\partial_c v_d) \mathbf{u}^d \mathbf{u}^{cT} \quad // \text{ note that } \mathbf{u}_d = \mathbf{u}^d \text{ in Cartesian space}$$

To express things in x' -coordinates, first write

$$\partial_c v_d = (R^i_c \partial'_i) (R^j_d v'_j) = R^i_c [(\partial'_i R^j_d) v'_j + R^j_d (\partial'_i v'_j)]$$

The next step is to express the matrix $\mathbf{u}^d \mathbf{u}^{cT}$ as a linear combination of $\mathbf{e}^e \mathbf{e}^{fT}$. In Section 3 (b) it was shown that the tangent base vectors transform in this way

$$\mathbf{e}'_n = \Sigma_i R^i_n \mathbf{e}_i \quad \text{where} \quad (\mathbf{e}'_n)^i = \delta_n^i \quad (*)$$

For present purposes, we write this as

$$\mathbf{u}_d = \Sigma_e R^e_d \mathbf{e}_e \quad \text{where} \quad (\mathbf{u}_d)^e = \delta_d^e$$

Therefore

$$\mathbf{u}^d = \Sigma_e R_e^d \mathbf{e}^e$$

Then we can change this column vector to a row vector,

$$\mathbf{u}^{cT} = \Sigma_f R_f^c \mathbf{e}_f^T$$

and so

$$\mathbf{u}^d \mathbf{u}^{cT} = \Sigma_{ef} R_e^d R_f^c \mathbf{e}^e \mathbf{e}^{fT} = R_e^d R_f^c \mathbf{e}^e \mathbf{e}^{fT}$$

Inserting these two results gives

$$\begin{aligned} (\nabla \mathbf{v}) &= \Sigma_{cd} (\nabla \mathbf{v})_{dc} \mathbf{u}_d \mathbf{u}_c^T = \Sigma_{cd} (\partial_c v_d) \mathbf{u}_d \mathbf{u}_c^T \\ &= \{ R^i_c [(\partial'_i R^j_d) v'_j + R^j_d (\partial'_i v'_j)] \} \{ R_e^d R_f^c \mathbf{e}^e \mathbf{e}^{fT} \} \\ &= \{ R^i_c [(\partial'_i R^j_d) v'_j + R^j_d (\partial'_i v'_j)] \} \{ R_e^d R_f^c \mathbf{e}^e \mathbf{e}^{fT} \} \\ &= \{ (R_f^c R^i_c) [R_e^d (\partial'_i R^j_d) v'_j + (R_e^d R^j_d) (\partial'_i v'_j)] \} \mathbf{e}^e \mathbf{e}^{fT} \\ &= \delta_f^i [R_e^d (\partial'_i R^j_d) v'_j + \delta_e^j (\partial'_i v'_j)] \mathbf{e}^e \mathbf{e}^{fT} \\ &= [R_e^d (\partial'_f R^j_d) v'_j + (\partial'_f v'_e)] \mathbf{e}^e \mathbf{e}^{fT} \end{aligned}$$

$$= \Sigma_{\mathbf{ef}} [(\nabla \mathbf{v})^{(\mathbf{e})}]_{\mathbf{ef}} \mathbf{e}^{\mathbf{e}} \mathbf{e}^{\mathbf{fT}}$$

where

$$[(\nabla \mathbf{v})^{(\mathbf{e})}]_{\mathbf{ef}} = (\partial'_{\mathbf{f}} v'_{\mathbf{e}}) + R_{\mathbf{e}}^{\mathbf{d}} (\partial'_{\mathbf{f}} R^{\mathbf{j}}_{\mathbf{d}}) v'_{\mathbf{j}}$$

or

$$[(\nabla \mathbf{v})^{(\mathbf{e})}]_{\mathbf{ij}} = (\partial'_{\mathbf{j}} v'_{\mathbf{i}}) + R_{\mathbf{i}}^{\mathbf{m}} (\partial'_{\mathbf{j}} R^{\mathbf{j}}_{\mathbf{m}}) v'_{\mathbf{j}}$$

and this agrees with the second form obtained by Method 1.

(d) Expansion on $\mathbf{e}_i \otimes \mathbf{e}_j$ and $\hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j$

Reversing the covariant tilts in the above expansion one gets

$$\nabla \mathbf{v} = \Sigma_{\mathbf{ij}} v'^{\mathbf{i};\mathbf{j}} \mathbf{e}_i \otimes \mathbf{e}_j$$

$$\text{where } v'^{\mathbf{i};\mathbf{j}} = g'^{\mathbf{ia}} g'^{\mathbf{jb}} v'_{\mathbf{a};\mathbf{b}} = g'^{\mathbf{ia}} g'^{\mathbf{jb}} [(\partial'_{\mathbf{b}} v'_{\mathbf{a}}) + R_{\mathbf{a}}^{\mathbf{m}} (\partial'_{\mathbf{b}} R^{\mathbf{c}}_{\mathbf{m}}) v'_{\mathbf{c}}]$$

Then since $\mathbf{e}_i = h'_i \hat{\mathbf{e}}_i$ one gets yet another expansion (more generally see Appendix E (h)),

$$\nabla \mathbf{v} = \Sigma_{\mathbf{ij}} (v'^{\mathbf{i};\mathbf{j}} h'_i h'_j) \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j = \Sigma_{\mathbf{ij}} [(\nabla \mathbf{v})^{(\hat{\mathbf{e}})}]_{\mathbf{ij}} \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j$$

where

$$[(\nabla \mathbf{v})^{(\hat{\mathbf{e}})}]_{\mathbf{ij}} = h'_i h'_j v'^{\mathbf{i};\mathbf{j}} = h'_i h'_j g'^{\mathbf{ia}} g'^{\mathbf{jb}} [(\partial'_{\mathbf{b}} v'_{\mathbf{a}}) + R_{\mathbf{a}}^{\mathbf{m}} (\partial'_{\mathbf{b}} R^{\mathbf{c}}_{\mathbf{m}}) v'_{\mathbf{c}}] .$$

Here is a summary of results so far

$$\nabla \mathbf{v} = \Sigma_{\mathbf{ij}} [(\nabla \mathbf{v})^{(\mathbf{e})}]_{\mathbf{ij}} \mathbf{e}^{\mathbf{i}} \otimes \mathbf{e}^{\mathbf{j}} \quad [(\nabla \mathbf{v})^{(\mathbf{e})}]_{\mathbf{ij}} = [(\partial'_{\mathbf{j}} v'_{\mathbf{i}}) + R_{\mathbf{i}}^{\mathbf{m}} (\partial'_{\mathbf{j}} R^{\mathbf{c}}_{\mathbf{m}}) v'_{\mathbf{c}}]$$

$$\nabla \mathbf{v} = \Sigma_{\mathbf{ij}} [(\nabla \mathbf{v})^{(\mathbf{e})}]^{\mathbf{i}}_{\mathbf{j}} \mathbf{e}_i \otimes \mathbf{e}^{\mathbf{j}} \quad [(\nabla \mathbf{v})^{(\mathbf{e})}]^{\mathbf{i}}_{\mathbf{j}} = [(\partial'_{\mathbf{j}} v'^{\mathbf{i}}) - R_{\mathbf{c}}^{\mathbf{m}} (\partial'_{\mathbf{j}} R^{\mathbf{i}}_{\mathbf{m}}) v'^{\mathbf{c}}] .$$

$$\nabla \mathbf{v} = \Sigma_{\mathbf{ij}} [(\nabla \mathbf{v})^{(\mathbf{e})}]^{\mathbf{ij}} \mathbf{e}_i \otimes \mathbf{e}_j \quad [(\nabla \mathbf{v})^{(\mathbf{e})}]^{\mathbf{ij}} = g'^{\mathbf{ia}} g'^{\mathbf{jb}} [(\partial'_{\mathbf{b}} v'_{\mathbf{a}}) + R_{\mathbf{a}}^{\mathbf{m}} (\partial'_{\mathbf{b}} R^{\mathbf{c}}_{\mathbf{m}}) v'_{\mathbf{c}}]$$

$$\nabla \mathbf{v} = \Sigma_{\mathbf{ij}} [(\nabla \mathbf{v})^{(\hat{\mathbf{e}})}]^{\mathbf{ij}} \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j \quad [(\nabla \mathbf{v})^{(\hat{\mathbf{e}})}]^{\mathbf{ij}} = h'_i h'_j g'^{\mathbf{ia}} g'^{\mathbf{jb}} [(\partial'_{\mathbf{b}} v'_{\mathbf{a}}) + R_{\mathbf{a}}^{\mathbf{m}} (\partial'_{\mathbf{b}} R^{\mathbf{c}}_{\mathbf{m}}) v'_{\mathbf{c}}]$$

One could replace $v'_{\mathbf{a}} = g'_{\mathbf{ad}} v'^{\mathbf{d}}$ in any of the above results. For example, the last object becomes

$$[(\nabla \mathbf{v})^{(\hat{\mathbf{e}})}]^{\mathbf{ij}} = h'_i h'_j g'^{\mathbf{ia}} g'^{\mathbf{jb}} [(\partial'_{\mathbf{b}} [g'_{\mathbf{ad}} v'^{\mathbf{d}}]) + R_{\mathbf{a}}^{\mathbf{m}} (\partial'_{\mathbf{b}} R^{\mathbf{c}}_{\mathbf{m}}) (g'_{\mathbf{cd}} v'^{\mathbf{d}})] .$$

Another choice is to use the $\mathbf{u}^{\mathbf{n}}$ components of $\mathbf{v}$ obtained when expanding $\mathbf{v}$ on the $\hat{\mathbf{e}}_{\mathbf{n}}$,

$$\mathbf{v} = \sum_n v^n \mathbf{e}_n = \sum_n v^n (h'_n \hat{\mathbf{e}}_n) = \sum_n (h'_n v^n) \hat{\mathbf{e}}_n = \sum_n \mathbf{u}^n \hat{\mathbf{e}}_n \quad \Rightarrow \quad \mathbf{u}^n \equiv h'_n v^n$$

so one then replaces $v^n = h'_n{}^{-1} \mathbf{u}^n$. The same object above then becomes

$$[(\nabla \mathbf{v})^{(\hat{\mathbf{e}})}]^{ij} = h'_i h'_j g'^{ia} g'^{jb} [(\partial'_b [g'_{ad} h'^d{}_d \mathbf{u}^d]) + R^m{}_b (\partial'_b R^c{}_m) (g'_{cd} h'^d{}_d \mathbf{u}^d)]$$

As discussed in Section 14 Example 1, the components $\mathbf{u}^n$ 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 $\mathbf{u}^n$ and just write $\mathbf{u}^n = v_{x', (n)}$. For example, in spherical coordinates r, θ, φ :

$$\mathbf{u}^1 = v_r \quad \mathbf{u}^2 = v_\theta \quad \mathbf{u}^3 = v_\varphi \quad \mathbf{v} = \sum_n \mathbf{u}^n \hat{\mathbf{e}}_n = v_r \hat{\mathbf{r}} + v_\theta \hat{\boldsymbol{\theta}} + v_\varphi \hat{\boldsymbol{\phi}}.$$

(e) Orthogonal coordinate systems

In this case $g'_{ab} = h'_a{}^2 \delta_{a,b}$ and $g'^{ab} = h'_a{}^{-2} \delta^{a,b}$ and things simplify. The above four expansions become (there is no change in the first two expansions)

$$\begin{aligned} \nabla \mathbf{v} &= \sum_{ij} [(\nabla \mathbf{v})^{(\mathbf{e})}]_{ij} \mathbf{e}^i \otimes \mathbf{e}^j & [(\nabla \mathbf{v})^{(\mathbf{e})}]_{ij} &= [(\partial'_j v^i) + R^m{}_i (\partial'_j R^c{}_m) v^c] \\ \nabla \mathbf{v} &= \sum_{ij} [(\nabla \mathbf{v})^{(\mathbf{e})}]^{ij} \mathbf{e}_i \otimes \mathbf{e}_j & [(\nabla \mathbf{v})^{(\mathbf{e})}]^{ij} &= [(\partial'_j v^i) - R^m{}_c (\partial'_j R^i{}_m) v^c] . \\ \nabla \mathbf{v} &= \sum_{ij} [(\nabla \mathbf{v})^{(\mathbf{e})}]^{ij} \mathbf{e}_i \otimes \mathbf{e}_j & [(\nabla \mathbf{v})^{(\mathbf{e})}]^{ij} &= h'_i{}^{-2} h'_j{}^{-2} [(\partial'_j v^i) + R^m{}_i (\partial'_j R^c{}_m) v^c] \\ \nabla \mathbf{v} &= \sum_{ij} [(\nabla \mathbf{v})^{(\hat{\mathbf{e}})}]^{ij} \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j & [(\nabla \mathbf{v})^{(\hat{\mathbf{e}})}]^{ij} &= h'_i{}^{-1} h'_j{}^{-1} [(\partial'_j v^i) + R^m{}_i (\partial'_j R^c{}_m) v^c] \end{aligned}$$

and the last object noted above becomes ($m \rightarrow d$ and $c \rightarrow b$)

$$\begin{aligned} P^{ij} &\equiv [(\nabla \mathbf{v})^{(\hat{\mathbf{e}})}]^{ij} = h'_i{}^{-1} h'_j{}^{-1} [(\partial'_j [h'_i \mathbf{u}^i]) + R^m{}_i (\partial'_j R^c{}_m) (h'_c \mathbf{u}^c)] \\ &= h'_i{}^{-1} h'_j{}^{-1} [(\partial'_j [h'_i \mathbf{u}^i]) + R^d{}_i (\partial'_j R^b{}_d) (h'_b \mathbf{u}^b)] \quad // m \rightarrow d, c \rightarrow b \\ &= h'_i{}^{-1} h'_j{}^{-1} [\underset{T2}{(\partial'_j h'_i)} \mathbf{u}^i + \underset{T3}{h'_i (\partial'_j \mathbf{u}^i)} + \underset{T1}{h'_i{}^2 R^i{}_d (\partial'_j R^b{}_d) (h'_b \mathbf{u}^b)}] \end{aligned}$$

where $R^i{}_d = g'_{ia} R^a{}_b g'^{bd} = h'_i{}^2 R^i{}_d$ is used to put all R's into their down-tilt form.

(f) Maple evaluation of $(\nabla \mathbf{v})$ in several coordinate systems

This object P^{ij} shown above can be computed in Maple by a simple program which is easily modified for other orthogonal curvilinear systems. The first task is to obtain the R matrix from the inverse transformation equations,


```

> xp[1] := r;
xp1 := r
> xp[2] := theta;
xp2 := θ
> xp[3] := phi;
xp3 := ϕ
> assume(r>0,theta>0,theta<Pi);
>
> x[1] := r*sin(theta)*cos(phi);
x1 := r sin(θ) cos(ϕ)
> x[2] := r*sin(theta)*sin(phi);
x2 := r sin(θ) sin(ϕ)
> x[3] := r*cos(theta);
x3 := r cos(θ)
> Vp := vector( [v[xp[1]],v[xp[2]],v[xp[3]] 1 ] );
Vp := [vr, vθ, vϕ]
> S_ := (i,j) -> diff(x[i],xp[j]);
S_ := (i,j) → ∂x_i / ∂xp_j
> S := matrix(N,N,S_);
S := ⎡ sin(θ) cos(ϕ)  r cos(θ) cos(ϕ)  -r sin(θ) sin(ϕ) ⎤
    ⎢ sin(θ) sin(ϕ)  r cos(θ) sin(ϕ)  r sin(θ) cos(ϕ) ⎥
    ⎣ cos(θ)        -r sin(θ)        0                ⎦
> R := simplify(inverse(S));
R := ⎡ sin(θ) cos(ϕ)  sin(θ) sin(ϕ)  cos(θ) ⎤
    ⎢ cos(ϕ) cos(θ)  sin(ϕ) cos(θ)  -sin(θ) ⎥
    ⎣ -sin(ϕ) / (r sin(θ))  cos(ϕ) / (r sin(θ))  0 ⎦

```

Then one needs the scale factors h_n ' from the metric tensor $\bar{g} = S^T S$ (dev notation Section 5 (ℓ))

```

gcov := simplify(evalm( transpose(S) &* S));
gcov := ⎡ 1  0  0 ⎤
        ⎢ 0  r^2  0 ⎥
        ⎣ 0  0  r^2 - r^2 cos(θ)^2 ⎦
for n from 1 to N do hp[n] := simplify(sqrt(gcov[n,n])) od;
hp1 := 1
hp2 := r
hp3 := r sin(θ)

```

The terms T1,T2 and T3 shown above are then entered,

$$\mathbf{p}^{ij} \equiv [(\nabla \mathbf{v})^{(\hat{\mathbf{e}})}]^{ij} = \mathbf{h}'_i{}^{-1} \mathbf{h}'_j{}^{-1} \left[\underbrace{(\partial'_j \mathbf{h}'_i)}_{T2} \mathbf{u}^i + \mathbf{h}'_i (\underbrace{\partial'_j \mathbf{u}^i}_{T3}) + \mathbf{h}'_i{}^2 \underbrace{R^i_d (\partial'_j R^b_d)}_{T1} (\mathbf{h}'_b \mathbf{u}^b) \right]$$

```
[> T1_ := (i,j) -> (hp[i])^2 * sum(
    sum(R[i,d]*Diff(R[b,d],xp[j])*hp[b]*Vp[b],d=1..N),b=1..N) ;
    T1_=(i,j) -> hp_i^2 * sum_{b=1}^N sum_{d=1}^N R_{i,d} * (d/dxp_j R_{b,d}) * hp_b * Vp_b
[> T2_ := (i,j) -> Vp[i] *diff(hp[i],xp[j]) ;
    T2_=(i,j) -> Vp_i * (d/dxp_j hp_i)
[> T3_ := (i,j) -> hp[i] * Diff(Vp[i],xp[j]) ;
    T3_=(i,j) -> hp_i * (d/dxp_j Vp_i)
```

The terms are then added and simplified and out pops the result,

```
[> P_ := (i,j) -> (1/hp[i])*(1/hp[j])*(value(T1_(i,j)) + (T2_(i,j)) + T3_(i,j));
    P_=(i,j) -> (value(T1_(i,j)) + T2_(i,j) + T3_(i,j)) / (hp_i * hp_j)
[> P :=matrix(N,N):
[> for n from 1 to N do
    for m from 1 to N do
        P[n,m] := expand(simplify(P_(n,m)));
    od;
od;
[> evalm(P);
>
```

$$\begin{bmatrix} \frac{\partial}{\partial r} v_r & -\frac{v_\theta}{r} + \frac{\partial}{\partial \theta} v_r & -\frac{v_\phi}{r} + \frac{\partial}{\partial \phi} v_r \\ \frac{\partial}{\partial r} v_\theta & \frac{v_r}{r} + \frac{\partial}{\partial \theta} v_\theta & -\frac{\cos(\theta) v_\phi}{r \sin(\theta)} + \frac{\partial}{\partial \phi} v_\theta \\ \frac{\partial}{\partial r} v_\phi & \frac{\partial}{\partial \theta} v_\phi & \frac{v_r}{r} + \frac{\cos(\theta) v_\theta}{r \sin(\theta)} + \frac{\partial}{\partial \phi} v_\phi \end{bmatrix}$$

Below are some sample results (including the above):

$$(\nabla \mathbf{v}) = \Sigma_{ij} P^{ij} \hat{\mathbf{e}}_i \hat{\mathbf{e}}_j^T \quad P^{ij} = [(\nabla \mathbf{v})^{(\hat{\mathbf{e}})}]^{ij}$$

- P^{ij} in **polar** coordinates (where 1,2 = r,θ) :

$$\begin{bmatrix} \frac{\partial}{\partial r} v_r & -\frac{v_\theta}{r} + \frac{\partial}{\partial \theta} \frac{v_r}{r} \\ \frac{\partial}{\partial r} v_\theta & \frac{v_r}{r} + \frac{\partial}{\partial \theta} \frac{v_\theta}{r} \end{bmatrix}$$

// agrees with Lai (2.23.23)

- P^{ij} in **cylindrical** coordinates (where 1,2,3 = r,θ,z) :

$$\begin{bmatrix} \frac{\partial}{\partial r} v_r & -\frac{v_\theta}{r} + \frac{\partial}{\partial \theta} \frac{v_r}{r} & \frac{\partial}{\partial z} v_r \\ \frac{\partial}{\partial r} v_\theta & \frac{v_r}{r} + \frac{\partial}{\partial \theta} \frac{v_\theta}{r} & \frac{\partial}{\partial z} v_\theta \\ \frac{\partial}{\partial r} v_z & \frac{\partial}{\partial \theta} \frac{v_z}{r} & \frac{\partial}{\partial z} v_z \end{bmatrix}$$

// 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) :

$$\begin{bmatrix} \frac{\partial}{\partial r} v_r & -\frac{v_\theta}{r} + \frac{\partial}{\partial \theta} \frac{v_r}{r} & -\frac{v_\phi}{r} + \frac{\partial}{\partial \phi} \frac{v_r}{r \sin(\theta)} \\ \frac{\partial}{\partial r} v_\theta & \frac{v_r}{r} + \frac{\partial}{\partial \theta} \frac{v_\theta}{r} & -\frac{\cos(\theta) v_\phi}{r \sin(\theta)} + \frac{\partial}{\partial \phi} \frac{v_\theta}{r \sin(\theta)} \\ \frac{\partial}{\partial r} v_\phi & \frac{\partial}{\partial \theta} \frac{v_\phi}{r} & \frac{v_r}{r} + \frac{\cos(\theta) v_\theta}{r \sin(\theta)} + \frac{\partial}{\partial \phi} \frac{v_\phi}{r \sin(\theta)} \end{bmatrix}$$

// agrees with Lai (2.35.25)

By entering the usual inverse transformation equations $\mathbf{x}' = \mathbf{F}^{-1}(\mathbf{x})$, and with suitable small alterations, the above Maple code can compute $\nabla \mathbf{v}$ in any orthogonal curvilinear coordinate system in any number of dimensions N.

Appendix H: Expansion of div(T) in curvilinear coordinates (T = rank-2 tensor)

The object of attention in this Appendix, **div**T, is expressed this way in Cartesian coordinates,

$$(\text{div}T)^i = \partial_j T^{ij} ,$$

where T^{ij} is a rank-2 tensor. One can regard the above equation as describing the normal divergence of the "vector" which forms the i^{th} row of the matrix T_{ij} . In general (that is to say, under general transformations F), the rows of T_{ij} are not tensorial vectors, and that is why $(\text{div}T)^i$ is not a tensorial scalar. Nor, in fact, are the $(\text{div}T)^i$ as defined above the components of a tensorial vector! As shown in section (b) below, **div**T will be redefined as a true tensorial vector which equals $\partial_j T^{ij}$ in Cartesian coordinates.

(a) Continuum Mechanics motivation

The vector **div**T arises for example when Newton's 2nd Law $\mathbf{F} = m\mathbf{a}$ is applied to a particle of continuous matter, in which case this law is known as Cauchy's Equation of Motion,

$$\mathbf{div}T + \rho\mathbf{B} = \rho\mathbf{a} \quad // \text{Lai p 169 (4.7.4)}$$

In this equation T is known as the Cauchy stress tensor, ρ is the mass density of the particle, **B** is any action-at-a-distance force (body force) per unit mass (such as gravity), and of course **a** is the acceleration of the particle.

Comment. Under rotations and translations the quantity $(\text{div}T)^i = \partial_j T^{ij}$ is a tensorial vector, ρ is a tensorial scalar, and **a** and **B** are tensorial vectors. As one would expect, $\mathbf{div}T + \rho\mathbf{B} = \rho\mathbf{a}$ is covariant in the sense of Section 7 (u) under these kinds of transformations.

As in the case of $\nabla\mathbf{v}$, one is interested in expressing **div**T in general curvilinear coordinates.

(b) Expansion of divT on e_n by Method 1: Use fact that $T^{ab}{}_{;\alpha}$ is a tensor.

As shown in the examples of Appendix F (i), the following object transforms as a rank-3 tensor,

$$\begin{aligned} T^{ab}{}_{;\alpha} &\equiv \partial_\alpha T^{ab} + \Gamma^a_{\alpha n} T^{nb} + \Gamma^b_{\alpha n} T^{an} = \partial_\alpha T^{ab} & // \text{x-space} & \quad \text{where } g=1, \Gamma=0 \\ T'^{ab}{}_{;\alpha} &\equiv \partial'_\alpha T'^{ab} + \Gamma'^a_{\alpha n} T'^{nb} + \Gamma'^b_{\alpha n} T'^{an} & // \text{x'-space} \end{aligned}$$

Contracting b with α yields the following tensorial vector equations,

$$\begin{aligned} (\text{div}T)^a &= T^{ab}{}_{;b} = \partial_b T^{ab} & // \text{x-space} & \quad \text{where } \Gamma = 0 \\ (\text{div}T)^a &= T'^{ab}{}_{;b} \equiv \partial'_b T'^{ab} + \Gamma'^a_{bn} T'^{nb} + \Gamma'^b_{bn} T'^{an} & // \text{x'-space} \end{aligned}$$

Note that the Cartesian space statement $(\text{div}T)^a = \partial_b T^{ab}$ is obtained, as noted in the opening comments above. According to Section 7 (s) or Appendix E (b), the vector **div**T can be expanded as

$$\mathbf{divT} = \Sigma_a (\mathbf{divT})^a \mathbf{e}_a \quad (\mathbf{divT})^a = [(\mathbf{divT})^{(e)}]^a \quad // \mathbf{divT} \text{ in the } \mathbf{e}_n \text{ basis}$$

From Appendix F (a), adjusted from Picture C to Picture A context (primes on ∂),

$$\begin{aligned} \Gamma^c_{ab} &= -R_b^i (\partial'_a R^c_i) \\ \Gamma^a_{bn} &= -R_n^i (\partial'_b R^a_i) \quad // b \rightarrow n \text{ then } a \rightarrow b \text{ then } c \rightarrow a \\ \Gamma^b_{bn} &= -R_n^i (\partial'_b R^b_i) \end{aligned}$$

$$\begin{aligned} \Rightarrow (\mathbf{divT})^a &= \partial'_b T^{ab} + \Gamma^a_{bn} T^{nb} + \Gamma^b_{bn} T^{an} \\ &= \partial'_b T^{ab} - R_n^i (\partial'_b R^a_i) T^{nb} - R_n^i (\partial'_b R^b_i) T^{an} \end{aligned}$$

The conclusion is that

$$\begin{aligned} \mathbf{divT} &= \Sigma_a [(\mathbf{divT})^{(e)}]^a \mathbf{e}_a \\ [(\mathbf{divT})^{(e)}]^a &= \partial'_b T^{ab} - R_n^i (\partial'_b R^a_i) T^{nb} - R_n^i (\partial'_b R^b_i) T^{an} \quad // \text{sum on } n \text{ and } b \end{aligned}$$

The vector $\mathbf{divT} = \Sigma_a (\partial_b T^{ab}) \mathbf{u}_a$ has thus been expressed in terms of x'-space coordinates and objects, and as usual the $\mathbf{e}_n$ are the tangent base vectors in x-space.

(c) Expansion of $\mathbf{divT}$ on $\mathbf{e}_n$ by Method 2: Use brute force

Start with the known expansion of $\mathbf{divT}$ in Cartesian x-space

$$\mathbf{divT} = \Sigma_i [\mathbf{divT}]^i \mathbf{u}_i = \Sigma_i (\partial_j T^{ij}) \mathbf{u}_i$$

Compute

$$\begin{aligned} (\partial_j T^{ij}) &= (R^a_j \partial'_a) (R_b^i R_c^j T^{b'c'}) \\ &= R^a_j R_b^i R_c^j (\partial'_a T^{b'c'}) + R^a_j R_b^i (\partial'_a R_c^j) T^{b'c'} + R^a_j R_c^j (\partial'_a R_b^i) T^{b'c'} \\ &= (R^a_j R_c^j) R_b^i (\partial'_a T^{b'c'}) + R^a_j R_b^i (\partial'_a R_c^j) T^{b'c'} + (R^a_j R_c^j) (\partial'_a R_b^i) T^{b'c'} \\ &= \delta^a_{c'} R_b^i (\partial'_a T^{b'c'}) + R^a_j R_b^i (\partial'_a R_c^j) T^{b'c'} + \delta^a_{c'} (\partial'_a R_b^i) T^{b'c'} \\ &= R_b^i (\partial'_a T^{b'a}) + R^a_j R_b^i (\partial'_a R_c^j) T^{b'c'} + (\partial'_a R_b^i) T^{b'a} \end{aligned}$$

Then use the same $\mathbf{u}_i$ expansion as in Appendix G (c),

$$\mathbf{u}_i = \Sigma_e R^n_i \mathbf{e}_n$$

Combining one gets,

$$\begin{aligned}
\mathbf{div} \mathbf{T} &= \Sigma_i (\partial_j T^{ij}) \mathbf{u}_i \\
&= \{ (R^n_i R_b'^i) (\partial'_a T^{b'a}) + R^a_j (R^n_i R_b'^i) (\partial'_a R_c'^j) T^{b'c'} + R^n_i (\partial'_a R_b'^i) T^{b'a} \} \mathbf{e}_n \\
&= \{ \delta^n_b (\partial'_a T^{b'a}) + R^a_j \delta^n_b (\partial'_a R_c'^j) T^{b'c'} + R^n_i (\partial'_a R_b'^i) T^{b'a} \} \mathbf{e}_n \\
&= \{ (\partial'_a T^{na}) + R^a_j (\partial'_a R_c'^j) T^{nc'} + R^n_i (\partial'_a R_b'^i) T^{b'a} \} \mathbf{e}_n \\
&= \Sigma_n [(\mathbf{div} \mathbf{T})^{(e)}]^n \mathbf{e}_n
\end{aligned}$$

where

$$[(\mathbf{div} \mathbf{T})^{(e)}]^n = (\partial'_a T^{na}) + R^a_j (\partial'_a R_c'^j) T^{nc'} + R^n_i (\partial'_a R_b'^i) T^{b'a}$$

From Appendix F (b) item 3 one has (Picture C1 $\rightarrow$ Picture A so primes on ∂ 's),

$$\begin{aligned}
(\partial'_a R_d^n) &= -R_e^n R_d^m (\partial'_a R_m^e) \\
(\partial'_a R_c'^j) &= -R_e^j R_c'^m (\partial'_a R_m^e) \quad // d \rightarrow c', n \rightarrow j \\
(\partial'_a R_b'^i) &= -R_e^i R_b'^m (\partial'_a R_m^e)
\end{aligned}$$

so then

$$\begin{aligned}
[(\mathbf{div} \mathbf{T})^{(e)}]^n &= (\partial'_a T^{na}) + R^a_j (\partial'_a R_c'^j) T^{nc'} + R^n_i (\partial'_a R_b'^i) T^{b'a} \\
&= (\partial'_a T^{na}) - R^a_j R_e^j R_c'^m (\partial'_a R_m^e) T^{nc'} - R^n_i R_e^i R_b'^m (\partial'_a R_m^e) T^{b'a} \\
&= (\partial'_a T^{na}) - \delta^a_e R_c'^m (\partial'_a R_m^e) T^{nc'} - \delta^n_e R_b'^m (\partial'_a R_m^e) T^{b'a} \\
&= (\partial'_a T^{na}) - R_c'^m (\partial'_a R_m^a) T^{nc'} - R_b'^m (\partial'_a R_m^n) T^{b'a}
\end{aligned}$$

Reverse the order of the last two terms

$$[(\mathbf{div} \mathbf{T})^{(e)}]^n = (\partial'_a T^{na}) - R_b'^m (\partial'_a R_m^n) T^{b'a} - R_c'^m (\partial'_a R_m^a) T^{nc'}.$$

Replace summation indices $m \rightarrow i$, $a \rightarrow b$,

$$= (\partial'_b T^{nb}) - R_b'^i (\partial'_b R_i^n) T^{b'a} - R_c'^i (\partial'_b R_i^b) T^{nc'}.$$

Finally, change $n \rightarrow a$ and then $c' \rightarrow n$ and $b' \rightarrow n$,

$$[(\mathbf{div} \mathbf{T})^{(e)}]^a = (\partial'_b T^{ab}) - R_n^i (\partial'_b R_i^a) T^{nb} - R_n^i (\partial'_b R_i^b) T^{an}. \quad // \text{sum on } n \text{ and } b$$

This is seen to match the result of Method 1 of the previous section. The brute force method is in fact not too bad and requires no explicit use of the affine connection Γ .

Technical Note: In the above expression one can write for example $R_n^i(\partial'_b R^a_i) = g'_{nm} R^m_i(\partial'_b R^a_i)$. Then using the fact that $R^m_i R^a_i = g^{ma}$ one finds $R^m_i(\partial'_b R^a_i) + R^a_i(\partial'_b R^m_i) = (\partial'_b g^{ma})$ which then allows the replacement $R_n^i(\partial'_b R^a_i) = g'_{nm} (\partial'_b g^{ma}) - g'_{nm} R^a_i(\partial'_b R^m_i)$. This kind of transformation leads to an alternate form for **div**T shown above, still with all down-tilt R matrices, but the index structure is different. This other form appears when one does the "brute force" method starting with $(\partial_j T^{ij})$ instead of with $(\partial_j T^{ij})$. Of course in Cartesian space these two objects must be the same.

(d) Adjustment for T expanded on $(\hat{e}_i \otimes \hat{e}_j)$ and **div**T expanded on $\hat{e}_a$

In the above, it has been assumed that T^{ij} is a rank-2 tensor so that, in the notation of Appendix E,

$$T = \Sigma_{ij} T^{ij} (\mathbf{u}_i \otimes \mathbf{u}_j) = \Sigma_{ij} T^{ij} (\mathbf{e}_i \otimes \mathbf{e}_j)$$

If one is interested in an expansion of T on the unit vectors $\hat{e}_i$ where $\mathbf{e}_i = h'_i \hat{e}_i$ this becomes

$$T = \Sigma_{ij} [T^{ij} h'_i h'_j] (\hat{e}_i \otimes \hat{e}_j) = \Sigma_{ij} [T^{(\hat{e})}]^{ij} (\hat{e}_i \otimes \hat{e}_j)$$

One then has

$$[T^{(\hat{e})}]^{ij} = h'_i h'_j T^{ij} \Rightarrow T^{ij} = h'^{-1}_i h'^{-1}_j [T^{(\hat{e})}]^{ij}$$

If one is interested in this form of the T matrix elements, then one is likely also interested in this expansion for **div**T,

$$\mathbf{div}T = \Sigma_a [(\mathbf{div}T)^{(e)}]^a \mathbf{e}_a = \Sigma_n \{ [(\mathbf{div}T)^{(e)}]^a h'_a \} \hat{e}_a \equiv [(\mathbf{div}T)^{(\hat{e})}]^a \hat{e}_a$$

where then

$$\begin{aligned} [(\mathbf{div}T)^{(\hat{e})}]^a &= h'_a [(\mathbf{div}T)^{(e)}]^a \\ &= h'_a [(\partial'_b T^{ab}) - R_n^i (\partial'_b R^a_i) T^{nb} - R_n^i (\partial'_b R^b_i) T^{an}] \\ &= h'_a * \{ \partial'_b (h'^{-1}_a h'^{-1}_b [T^{(\hat{e})}]^{ab}) - R_n^i (\partial'_b R^a_i) h'^{-1}_n h'^{-1}_b [T^{(\hat{e})}]^{nb} \\ &\quad - R_n^i (\partial'_b R^b_i) h'^{-1}_a h'^{-1}_n [T^{(\hat{e})}]^{an} \} \\ &= h'_a * \{ \partial'_b (h'^{-1}_a h'^{-1}_b) [T^{(\hat{e})}]^{ab} - R_n^i (\partial'_b R^a_i) h'^{-1}_n h'^{-1}_b [T^{(\hat{e})}]^{nb} \\ &\quad + h'^{-1}_a h'^{-1}_b (\partial'_b [T^{(\hat{e})}]^{ab}) - R_n^i (\partial'_b R^b_i) h'^{-1}_a h'^{-1}_n [T^{(\hat{e})}]^{an} \} \end{aligned}$$

Now

$$\partial'_b (h'^{-1}_a h'^{-1}_b) = \partial'_b (h'_a h'_b)^{-1} = - (h'_a h'_b)^{-2} \partial'_b (h'_a h'_b) = - h'^{-2}_a h'^{-2}_b \partial'_b (h'_a h'_b)$$

so the above sequence for $[(\mathbf{divT})^{(\hat{e})}]^a$ continues,

$$\begin{aligned}
 &= h'_a * \{ -h'_a{}^{-2} h'_b{}^{-2} \partial'_b (h'_a h'_b) [T^{(\hat{e})}]^{ab} - R_n^i (\partial'_b R^a_i) h'_n{}^{-1} h'_b{}^{-1} [T^{(\hat{e})}]^{nb} \\
 &\quad + h'_a{}^{-1} h'_b{}^{-1} (\partial'_b [T^{(\hat{e})}]^{ab}) - R_n^i (\partial'_b R^b_i) h'_a{}^{-1} h'_n{}^{-1} [T^{(\hat{e})}]^{an} \} \\
 &= \{ -h'_a{}^{-1} h'_b{}^{-2} \partial'_b (h'_a h'_b) [T^{(\hat{e})}]^{ab} - h'_a h'_n{}^{-1} h'_b{}^{-1} g'_{nm} R^m_i (\partial'_b R^a_i) [T^{(\hat{e})}]^{nb} \\
 &\quad + h'_b{}^{-1} (\partial'_b [T^{(\hat{e})}]^{ab}) - h'_n{}^{-1} g'_{nm} R^m_i (\partial'_b R^b_i) [T^{(\hat{e})}]^{an} \}
 \end{aligned}$$

where recall that $R^n_i = R^{ni}$ since x-space is Cartesian. In the last form only the down-tilt R matrix appears which simplifies calculation with Maple. This and all other results above are valid for general curvilinear coordinates, orthogonal as well as non-orthogonal.

At this point, we will specialize to **orthogonal** systems so the last form above becomes

$$\begin{aligned}
 [(\mathbf{divT})^{(\hat{e})}]^a &= \{ \overset{T1}{-h'_a{}^{-1} h'_b{}^{-2} \partial'_b (h'_a h'_b) [T^{(\hat{e})}]^{ab}} - \overset{T3}{h'_a h'_b{}^{-1} h'_n R^n_i (\partial'_b R^a_i) [T^{(\hat{e})}]^{nb}} \\
 &\quad + \overset{T2}{h'_b{}^{-1} (\partial'_b [T^{(\hat{e})}]^{ab})} - \overset{T4}{h'_n R^n_i (\partial'_b R^b_i) [T^{(\hat{e})}]^{an}} \}
 \end{aligned}$$

Even for an orthogonal curvilinear coordinate system, the form of this tensor divergence is amazingly complicated. Here it is expressed in terms of the down-tilt R matrix and the scale factors and as usual repeated indices are summed.

(e) Maple: $\mathbf{divT}$ in cylindrical and spherical coordinates

The above object $[(\mathbf{divT})^{(\hat{e})}]^a$ can be evaluated by Maple code very similar to that shown in Appendix G for $(\nabla \mathbf{v})$. The main difference is the set of entry lines for the terms,

```

> T1_ := (a) -> sum( -(1/hp[a])*(1/hp[b]^2)*Diff(hp[a]*hp[b],xp[b])*Te[a,b],b=1..N);

T1_ = a -> sum_{b=1}^N \left( - \frac{\left( \frac{\partial}{\partial x_b} h p_a h p_b \right) T e_{a,b}}{h p_a h p_b^2} \right)

> T2_ := (a) -> sum((1/hp[b])*Diff(Te[a,b],xp[b]),b=1..N);

T2_ = a -> sum_{b=1}^N \frac{\frac{\partial}{\partial x_b} T e_{a,b}}{h p_b}

> T3_ := (a) -> sum( sum( sum(
-hp[a]*(1/hp[b])*hp[n]*R[n,i]*Diff(R[a,i],xp[b])*Te[n,b],b=1..3),i=1..3),n=1..3);

T3_ = a -> sum_{n=1}^3 \left( \sum_{i=1}^3 \left( \sum_{b=1}^3 \left( - \frac{h p_a h p_n R_{n,i} \left( \frac{\partial}{\partial x_b} R_{a,i} \right) T e_{n,b}}{h p_b} \right) \right) \right)

> T4_ := (a) -> sum( sum( sum(
-hp[n]*R[n,i]*Diff(R[b,i],xp[b])*Te[a,n],b=1..N),i=1..N),n=1..N);

T4_ = a -> sum_{n=1}^N \left( \sum_{i=1}^N \left( \sum_{b=1}^N \left( - h p_n R_{n,i} \left( \frac{\partial}{\partial x_b} R_{b,i} \right) T e_{a,n} \right) \right) \right)
    
```


Here are some results for $[(\text{divT})^{(\hat{\mathbf{e}})}]_{\mathbf{a}} = \text{"divT}_{\mathbf{a}}\text{"}$:

- **cylindrical** coordinates (where 1,2,3 = r,θ,z) :

$$\text{divT}_1 := \left(\frac{\partial}{\partial r} T_{r,r} \right) + \frac{\frac{\partial}{\partial \theta} T_{r,\theta}}{r} + \left(\frac{\partial}{\partial z} T_{r,z} \right) - \frac{T_{\theta,\theta}}{r} + \frac{T_{r,r}}{r}$$

$$\text{divT}_2 := \frac{T_{\theta,r}}{r} + \left(\frac{\partial}{\partial r} T_{\theta,r} \right) + \frac{\frac{\partial}{\partial \theta} T_{\theta,\theta}}{r} + \left(\frac{\partial}{\partial z} T_{\theta,z} \right) + \frac{T_{r,\theta}}{r}$$

$$\text{divT}_3 := \left(\frac{\partial}{\partial r} T_{z,r} \right) + \frac{\frac{\partial}{\partial \theta} T_{z,\theta}}{r} + \left(\frac{\partial}{\partial z} T_{z,z} \right) + \frac{T_{z,r}}{r}$$

// agrees with Lai p 60 (2.34.8,9,10)

- For **polar** coordinates P^1 and P^2 are given by the first two lines above with the last terms set to 0. These polar results then agree with Lai p58 (2.33.32,33).

- **spherical** coordinates (where 1,2,3 = r,θ,φ) :

$$\text{divT}_1 := \left(\frac{\partial}{\partial r} T_{r,r} \right) + \frac{\frac{\partial}{\partial \theta} T_{r,\theta}}{r} + \frac{\frac{\partial}{\partial \phi} T_{r,\phi}}{r \sin(\theta)} - \frac{T_{\theta,\theta}}{r} - \frac{T_{\phi,\phi}}{r} + 2 \frac{T_{r,r}}{r} + \frac{\cos(\theta) T_{r,\theta}}{r \sin(\theta)}$$

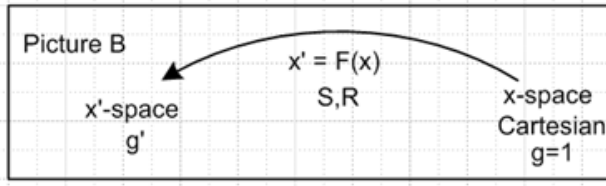
$$\text{divT}_2 := \frac{\frac{\partial}{\partial \phi} T_{\theta,\phi}}{r \sin(\theta)} + \frac{\frac{\partial}{\partial \theta} T_{\theta,\theta}}{r} + 2 \frac{T_{\theta,r}}{r} + \frac{\cos(\theta) T_{\theta,\theta}}{r \sin(\theta)} + \left(\frac{\partial}{\partial r} T_{\theta,r} \right) - \frac{\cos(\theta) T_{\phi,\phi}}{r \sin(\theta)} + \frac{T_{r,\theta}}{r}$$

$$\text{divT}_3 := 2 \frac{T_{\phi,r}}{r} + \frac{\cos(\theta) T_{\theta,\phi}}{r \sin(\theta)} + \left(\frac{\partial}{\partial r} T_{\phi,r} \right) + \frac{\frac{\partial}{\partial \phi} T_{\phi,\phi}}{r \sin(\theta)} + \frac{\frac{\partial}{\partial \theta} T_{\phi,\theta}}{r} + \frac{\cos(\theta) T_{\phi,\theta}}{r \sin(\theta)} + \frac{T_{r,\phi}}{r}$$

The expressions above agree with Lai p 65 (2.35.33,34,35).

Appendix I : The Vector Laplacian in Spherical and Cylindrical Coordinates

This appendix assumes the usual curvilinear coordinates context, Picture B,



In this section the vector Laplacian is computed in two different ways, each associated with a particular "tensorization" of its Cartesian form. The second method, though less pleasant than the first, gives insight into why the vector Laplacian always includes the scalar Laplacian of the field components. It is in this inclusive form that results are usually stated in the literature.

In passing, it should be noted that the vector Laplacian is not just an idle mathematical curiosity. It shows up for example in the wave equations for electric and magnetic fields in a vacuum,

$$(\nabla^2 + k^2)\mathbf{E}(\mathbf{x}) = 0 \quad (\nabla^2 + k^2)\mathbf{B}(\mathbf{x}) = 0 \quad k = \omega/c \quad \mathbf{E}, \mathbf{B}(\mathbf{x}, t) = \mathbf{E}, \mathbf{B}(\mathbf{x})e^{-i\omega t}$$

In continuum mechanics, it appears for example in the Navier/Cauchy equation which describes the small vector displacement $\mathbf{u}$ field in an isotropic elastic solid,

$$\rho_0 \partial_t^2 \mathbf{u} = \rho_0 \mathbf{B} + (\lambda + \mu) \nabla e + \mu \nabla^2 \mathbf{u} \quad e = \text{div } \mathbf{u} = \text{dilatation} \quad // \text{Lai p 216 (5.6.9)}$$

Here ρ_0 is the unperturbed mass density, $\mathbf{B}$ the body force, and λ and μ are the two Lamé constants which describe an isotropic elastic medium. The vector Laplacian makes another appearance in the better-known Navier-Stokes equation which describes the vector velocity field $\mathbf{v}$ in an incompressible Newtonian fluid,

$$\rho [\partial_t \mathbf{v} + (\nabla \mathbf{v}) \mathbf{v}] = \rho \mathbf{B} - \nabla p + \mu \nabla^2 \mathbf{v} \quad // \text{Lai p 361 (6.7.6)}$$

where ρ is the mass density and p is pressure.

(a) The first method : a review

In Section 13 (c) it was shown that, in Cartesian coordinates,

$$\begin{aligned} \nabla^2 (\mathbf{B}_n) &= [\text{grad}(\text{div } \mathbf{B}) - \text{curl}(\text{curl } \mathbf{B})]_n & \nabla^2 &= \partial_j \partial^j \\ \text{or} \\ \nabla \bullet \nabla (\mathbf{B}_n) &= [\nabla(\nabla \bullet \mathbf{B}) - \nabla \times (\nabla \times \mathbf{B})]_n \end{aligned}$$

When all components are considered in a single equation, one could write

$$\begin{aligned} \nabla^2 (\mathbf{B}) &= \text{grad}(\text{div } \mathbf{B}) - \text{curl}(\text{curl } \mathbf{B}) \\ \text{or} \\ \nabla \bullet \nabla (\mathbf{B}) &= \nabla(\nabla \bullet \mathbf{B}) - \nabla \times (\nabla \times \mathbf{B}) \end{aligned}$$

and this is frequently done (see examples cited above). Since $\nabla^2(\mathbf{B}) \neq \partial_j \partial^j \mathbf{B}$ in curvilinear coordinates, it seems notionally safer in our current context to use a different symbol for the vector Laplacian operator, and following Moon and Spencer we use $\star$ so the second last equation above then says:

$$\star \mathbf{B} = \text{grad}(\text{div } \mathbf{B}) - \text{curl}(\text{curl } \mathbf{B}) .$$

Since $[\star \mathbf{B}]_n$ agrees with $\nabla^2(\mathbf{B}_n)$ in Cartesian coordinates, and since we know how to write div, grad and curl in curvilinear coordinates (Sections 9,10,12 or Section 15), the right hand side of the above equation provides our "first method" of writing $\star \mathbf{B}$ in curvilinear coordinates. In Section 15 (g) it was shown that the proper "tensorization" of the above equation is given by

$$(\star \mathbf{B})^n = (B^j{}_{;j})^{;n} - g^{-1/2} \varepsilon^{nab} (g^{-1/2} \varepsilon_{bde} B^{e;d})_{;a} \quad \star \mathbf{B} = (\star B)^n \mathbf{u}_n$$

and therefore, in x' -space (x' are the curvilinear coordinates of interest) ,

$$(\star \mathbf{B})'^n = (B'^j{}_{;j})'^n - g'^{-1/2} \varepsilon'^{nab} (g'^{-1/2} \varepsilon'_{bde} B'^{e;d})_{;a} \quad \star \mathbf{B} = (\star B)'^n \mathbf{e}_n$$

where $(B'^j{}_{;j}) = \text{div } \mathbf{B}$. The object $\star \mathbf{B}$ is a normal vector (weight 0), assuming $\mathbf{B}$ is a normal vector. Doing various simplifying steps, we then arrived at the following expression which lends itself to calculation:

$$\begin{aligned} (\star \mathbf{B})'^n &= \partial'^n \{ (1/\sqrt{g'}) \partial'_{i'} (\sqrt{g'} B'^i{}_{;i'}) \} \\ &- (1/\sqrt{g'}) \varepsilon'^{ncd} \varepsilon'^{eab} \partial'_{c'} \{ (1/\sqrt{g'}) g'_{de} (\partial'_{a'} [g'_{bf} B'^f{}_{;f'}]) \} \end{aligned} \quad \star \mathbf{B} = (\star B)'^n \mathbf{e}_n$$

where ε'^{ncd} is the usual permutation tensor ($\varepsilon'^{ncd} = \varepsilon^{ncd}$). In this notation, $B'^i{}_{;i'}$ is an official x' -space contravariant vector component.

We are often interested in working with vectors which are expanded on unit vector versions of the tangent base vectors. In such a unit vector expansion of a vector, italic font has been used for the components. As shown in various places, one has

$$\mathbf{B}'^n = \mathcal{B}'^n / h'_n \quad \mathbf{e}_n = h'_n \hat{\mathbf{e}}_n \quad \mathbf{B} = B'_n \mathbf{e}_n = \mathcal{B}'^n \hat{\mathbf{e}}_n \quad h'_n = \text{scale factor}$$

and this then gives

$$\begin{aligned} (\star \mathbf{B})'^n &= \partial'^n \{ (1/\sqrt{g'}) \partial'_{i'} (\sqrt{g'} \mathcal{B}'^i{}_{;i'} / h'_i) \} \\ &- (1/\sqrt{g'}) \varepsilon'^{ncd} \varepsilon'^{eab} \partial'_{c'} \{ (1/\sqrt{g'}) g'_{de} (\partial'_{a'} [g'_{bf} \mathcal{B}'^f{}_{;f'} / h'_f]) \} \end{aligned} \quad \star \mathbf{B} = (\star \mathcal{B})'^n \mathbf{e}_n$$

or, as we will use it with unit vectors,

$$\begin{aligned} (\star \mathcal{B})'^n &= h'_n \partial'^n \{ (1/\sqrt{g'}) \partial'_{i'} (\sqrt{g'} \mathcal{B}'^i{}_{;i'} / h'_i) \} \\ &- h'_n (1/\sqrt{g'}) \varepsilon'^{ncd} \varepsilon'^{eab} \partial'_{c'} \{ (1/\sqrt{g'}) g'_{de} (\partial'_{a'} [g'_{bf} \mathcal{B}'^f{}_{;f'} / h'_f]) \} . \end{aligned} \quad \star \mathbf{B} = (\star \mathcal{B})'^n \hat{\mathbf{e}}_n$$

Specializing to orthogonal coordinates yields the form we shall use for computation in Maple,

$$(\star \mathcal{B})'^n = (1/h'_n) \partial'_n \{ (1/\sqrt{g'}) \partial'_i (\sqrt{g'} \mathcal{B}'^i / h'_i) \} - (h'_n / \sqrt{g'}) \varepsilon'^{ncd} \varepsilon'^{dab} \partial'_c \{ (1/\sqrt{g'}) h'_d{}^2 (\partial'_a [h'_b \mathcal{B}'^b]) \} \quad \star \mathbf{B} = (\star \mathcal{B})'^n \hat{\mathbf{e}}_n$$

The second term could be written as two terms by reducing the product $\varepsilon'^{ncd} \varepsilon'^{dab}$ into $\delta\delta - \delta\delta$ in the usual manner, but Maple is happy to just "do it" as stated. And for non-orthogonal coordinates, the reduction of $\varepsilon'^{ncd} \varepsilon'^{eab}$ is much uglier (Appendix D (j) item 3) and then one would want to use the $\varepsilon\varepsilon$ product as is.

(b) The first method in spherical coordinates: Maple speaks

In this section, the following notation is used:

$$\mathcal{B}'^1 = B_r \quad \mathcal{B}'^2 = B_\theta \quad \mathcal{B}'^3 = B_\phi$$

$$\mathbf{B} = \mathcal{B}'^n \hat{\mathbf{e}}_n = B_r \hat{\mathbf{e}}_r + B_\theta \hat{\mathbf{e}}_\theta + B_\phi \hat{\mathbf{e}}_\phi = B_r \hat{\mathbf{r}} + B_\theta \hat{\boldsymbol{\theta}} + B_\phi \hat{\boldsymbol{\phi}} .$$

Maple begins in the same manner as shown earlier in Appendix G (f), the idea being that this code could be used for any coordinate system,

```

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);

Bp := vector( [B[xp[1]], B[xp[2]], B[xp[3]]] );

S_ := (i,j) -> diff(x[i], xp[j]);

S := matrix(N,N,S_);

```

$$N=3$$

$$xp_1 := r$$

$$xp_2 := \theta$$

$$xp_3 := \phi$$

$$x_1 := r \sin(\theta) \cos(\phi)$$

$$x_2 := r \sin(\theta) \sin(\phi)$$

$$x_3 := r \cos(\theta)$$

$$Bp := [B_r, B_\theta, B_\phi]$$

$$S_ = (i,j) \rightarrow \frac{\partial}{\partial xp_j} x_i$$

$$S = \begin{bmatrix} \sin(\theta) \cos(\phi) & r \cos(\theta) \cos(\phi) & -r \sin(\theta) \sin(\phi) \\ \sin(\theta) \sin(\phi) & r \cos(\theta) \sin(\phi) & r \sin(\theta) \cos(\phi) \\ \cos(\theta) & -r \sin(\theta) & 0 \end{bmatrix}$$

```

gcov := simplify(evalm( transpose(S) &* S));

```

$$gcov := \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 - r^2 \cos(\theta)^2 \end{bmatrix}$$

```

for m from 1 to N do hp[m] := simplify(sqrt(gcov[m,m])) od;

```

$$hp_1 := 1$$

$$hp_2 := r$$

$$hp_3 := r \sin(\theta)$$

```

sq_gp := hp[1]*hp[2]*hp[3];

```

$$sq_gp := r^2 \sin(\theta)$$

The last object is $\sqrt{g'}$, where $hp[n] = h'_n$. The next chunk of code computes the *scalar* Laplacian of an unspecified function f , and this expression will be used below in parsing the vector Laplacian results :

Here is the formula to be used for computing the regular Laplacian:

$$[\text{lap } f](\mathbf{x}) = [1/\sqrt{g'(\mathbf{x})}] \partial'_n [\sqrt{g'(\mathbf{x})} h'_n{}^{-2}(\mathbf{x}) (\partial'_n f'(\mathbf{x}))]$$

```

lapf := (1/sq_gp)*sum(Diff(sq_gp *
hp[n]^(-2)*Diff(f(xp[1],xp[2],xp[3]),xp[n]),xp[n]),n=1..N);

```

$$\text{lapf} = \frac{\left(\frac{\partial}{\partial r} r^2 \sin(\theta) \left(\frac{\partial}{\partial r} f(r, \theta, \phi) \right) \right) + \left(\frac{\partial}{\partial \theta} \sin(\theta) \left(\frac{\partial}{\partial \theta} f(r, \theta, \phi) \right) \right) + \left(\frac{\partial}{\partial \phi} \frac{\partial}{\partial \phi} f(r, \theta, \phi) \right)}{r^2 \sin(\theta)}$$

```

subs(f(xp[1],xp[2],xp[3])=op(0,f(xp[1],xp[2],xp[3])),value(expand(%)));

```

$$2 \frac{\partial}{\partial r} f + \left(\frac{\partial^2}{\partial r^2} f \right) + \frac{\cos(\theta) \left(\frac{\partial}{\partial \theta} f \right)}{r^2 \sin(\theta)} + \frac{\frac{\partial^2}{\partial \theta^2} f}{r^2} + \frac{\frac{\partial^2}{\partial \phi^2} f}{r^2 \sin(\theta)^2}$$

The subs command used here (and more intensely below) removes the arguments of the function f for purely cosmetic reasons (see A Maple User's Guide nearby for details, operands section). The arguments are added in the first place to prevent Maple from thinking the unspecified function f is a constant.

Next comes a low-budget implementation of the permutation tensor $\text{eps}(a,b,c) = \epsilon^{abc}$,

```

eps := proc(a,b,c)
if type(a,numeric) and type(b,numeric) and type(c,numeric) then
  if (a=1 and b=2 and c=3) then 1;
  elif (a=1 and b=3 and c=2) then -1;
  elif (a=2 and b=1 and c=3) then -1;
  elif (a=2 and b=3 and c=1) then 1;
  elif (a=3 and b=1 and c=2) then 1;
  elif (a=3 and b=2 and c=1) then -1;
  else 0;
  fi;
else 'eps(a,b,c)';
fi;
end;

```

The two terms of the above ($\star B$)ⁿ expression shown at the end of section (a) are then duly entered,

$$(\star B)^n = T1^n + T2^n$$

$$T1^n = (1/h'_n) \partial'_n \{ (1/\sqrt{g'}) \partial'_i (\sqrt{g'} B'^i / h'_i) \}$$

$$T2^n = - (h'_n / \sqrt{g'}) \epsilon^{mcd} \epsilon^{dab} \partial'_c \{ (1/\sqrt{g'}) h'_a{}^2 (\partial'_a [h'_b B'^b]) \}$$

```
T1 := (n) -> (1/hp[n])*sum(Diff((1/sq_gp)*
Diff(sq_gp*Bp[i](xp[1],xp[2],xp[3])/hp[i],xp[i]),xp[n]),i=1..N);
```

$$T1 := n \rightarrow \frac{\sum_{i=1}^N \left(\frac{\partial}{\partial x p_n} \frac{\frac{\partial}{\partial x p_i} \frac{sq_gp B p_i(x p_1, x p_2, x p_3)}{hp_i}}{sq_gp} \right)}{hp_n}$$

```
T2 := (n) -> -(hp[n]/sq_gp)*sum(sum(sum(sum(eps(n,c,d)*eps(d,a,b) *Diff((1/sq_gp) *
hp[d]^2*Diff(hp[b]*Bp[b](xp[1],xp[2],xp[3]),xp[a]),xp[c]),a=1..N),b=1..N),c=1..N),d=1..N);
```

$$T2 := n \rightarrow - \frac{hp_n \left(\sum_{d=1}^N \left(\sum_{c=1}^N \left(\sum_{b=1}^N \left(\sum_{a=1}^N \epsilon ps(n,c,d) \epsilon ps(d,a,b) \left(\frac{\partial}{\partial x p_c} \frac{hp_d^2 \left(\frac{\partial}{\partial x p_a} hp_b B p_b(x p_1, x p_2, x p_3) \right)}{sq_gp} \right) \right) \right) \right) \right)}{sq_gp}$$

```
starB:= (n) -> T1(n) + T2(n):
```

In the following code, we use (in line with the notation shown at the start of this section)

$$q1_ = (\star B)^1 = (\star B)_r$$

$$q2_ = (\star B)^2 = (\star B)_\theta$$

$$q3_ = (\star B)^3 = (\star B)_\phi$$

```
q1 := expand(value(starB(1))):
'subs(Bp[1](xp[1],xp[2],xp[3]) = op(0,Bp[1](xp[1],xp[2],xp[3])),q1) ':
'subs(Bp[2](xp[1],xp[2],xp[3]) = op(0,Bp[2](xp[1],xp[2],xp[3])),%) ':
q1_ := 'subs(Bp[3](xp[1],xp[2],xp[3]) = op(0,Bp[3](xp[1],xp[2],xp[3])),%) ':
q1_;
```

$$-2 \frac{B_r}{r^2} + 2 \frac{\frac{\partial}{\partial r} B_r}{r} + \left(\frac{\partial^2}{\partial r^2} B_r \right) - 2 \frac{\cos(\theta) B_\theta}{r^2 \sin(\theta)} - 2 \frac{\frac{\partial}{\partial \theta} B_\theta}{r^2} - 2 \frac{\frac{\partial}{\partial \phi} B_\phi}{r^2 \sin(\theta)} + \frac{\frac{\partial^2}{\partial \phi^2} B_r}{r^2 \sin(\theta)^2} + \frac{\cos(\theta) \left(\frac{\partial}{\partial \theta} B_r \right)}{r^2 \sin(\theta)} + \frac{\frac{\partial^2}{\partial \theta^2} B_r}{r^2}$$

```
q2 := expand(value(starB(2))):
'subs(Bp[1](xp[1],xp[2],xp[3]) = op(0,Bp[1](xp[1],xp[2],xp[3])),q2) ':
'subs(Bp[2](xp[1],xp[2],xp[3]) = op(0,Bp[2](xp[1],xp[2],xp[3])),%) ':
q2_ := 'subs(Bp[3](xp[1],xp[2],xp[3]) = op(0,Bp[3](xp[1],xp[2],xp[3])),%) ':
q2_;
```

$$2 \frac{\frac{\partial}{\partial \theta} B_r}{r^2} - \frac{\cos(\theta)^2 B_\theta}{r^2 \sin(\theta)^2} + \frac{\cos(\theta) \left(\frac{\partial}{\partial \theta} B_\theta \right)}{r^2 \sin(\theta)} - \frac{B_\theta}{r^2} + \frac{\frac{\partial^2}{\partial \theta^2} B_\theta}{r^2} - 2 \frac{\cos(\theta) \left(\frac{\partial}{\partial \phi} B_\phi \right)}{r^2 \sin(\theta)^2} + \frac{\frac{\partial^2}{\partial \phi^2} B_\theta}{r^2 \sin(\theta)^2} + 2 \frac{\frac{\partial}{\partial r} B_\theta}{r} + \left(\frac{\partial^2}{\partial r^2} B_\theta \right)$$

$$(\star B)_\theta = \nabla^2(B_\theta) - (1/r^2) [-2 \partial_\theta B_r + \cot^2 \theta B_\theta + B_\theta + 2 \cot \theta \csc \theta \partial_\varphi B_\varphi]$$

$$= \nabla^2(B_\theta) - (1/r^2) [\csc^2\theta B_\theta - 2 \partial_\theta B_r + 2 \cot\theta \csc\theta \partial_\phi B_\phi]$$

q3_;

$$2 \frac{\frac{\partial}{\partial \phi} B_r}{r^2 \sin(\theta)} + 2 \frac{\cos(\theta) \left(\frac{\partial}{\partial \phi} B_\theta \right)}{r^2 \sin(\theta)^2} + \frac{\frac{\partial^2}{\partial \phi^2} B_\phi}{r^2 \sin(\theta)^2} - \frac{\cos(\theta)^2 B_\phi}{r^2 \sin(\theta)^2} + \frac{\cos(\theta) \left(\frac{\partial}{\partial \theta} B_\phi \right)}{r^2 \sin(\theta)} - \frac{B_\phi}{r^2} + \frac{\frac{\partial^2}{\partial \theta^2} B_\phi}{r^2} + 2 \frac{\frac{\partial}{\partial r} B_\phi}{r} + \left(\frac{\partial^2}{\partial r^2} B_\phi \right)$$

5
3
4
1
2

$$(\star B)_\varphi = \nabla^2(B_\varphi) - (1/r^2) [\csc^2\theta B_\varphi - 2\csc\theta\partial_\varphi B_r - 2\cot\theta\csc\theta\partial_\varphi B_\theta]$$

In this manner, we end up with the components of the vector Laplacian expressed in the traditional manner,

$$(\star B)_r = \nabla^2(B_r) - (2/r^2) [B_r + \cot\theta B_\theta + \partial_\theta B_\theta + \csc\theta \partial_\phi B_\phi]$$

$$(\nabla^2 B)_\theta = \nabla^2(B_\theta) - (1/r^2) [\csc^2\theta B_\theta - 2 \partial_\theta B_r + 2 \cot\theta \csc\theta \partial_\phi B_\phi]$$

$$(\star B)_\varphi = \nabla^2(B_\varphi) - (1/r^2) [\csc^2\theta B_\varphi - 2\csc\theta\partial_\varphi B_r - 2\cot\theta\csc\theta\partial_\varphi B_\theta]$$

where $\nabla^2 f = (1/r^2)\partial_{\mathbf{r}}(r^2\partial_{\mathbf{r}}f) + (1/r^2\sin\theta)\partial_{\theta}(\sin\theta\partial_{\theta}f) + (1/r^2\sin^2\theta)\partial_{\varphi}^2 f$

$$= (2/r)\partial_{\mathbf{r}}\mathbf{f} + \partial_{\mathbf{r}}^2\mathbf{f} + (\cot\theta/r^2)\partial_{\theta}\mathbf{f} + (1/r^2)\partial_{\theta}^2\mathbf{f} + (1/r^2\sin^2\theta)\partial_{\phi}^2\mathbf{f}$$

Once again, written out in gory detail, these three expressions are (in the same order r, θ, ϕ) :

$$\begin{aligned}
& -2 \frac{B_r}{r^2} + 2 \frac{\frac{\partial}{\partial r} B_r}{r} + \left(\frac{\partial^2}{\partial r^2} B_r \right) - 2 \frac{\cos(\theta) B_\theta}{r^2 \sin(\theta)} - 2 \frac{\frac{\partial}{\partial \theta} B_\theta}{r^2} - 2 \frac{\frac{\partial}{\partial \phi} B_\phi}{r^2 \sin(\theta)} + \frac{\frac{\partial^2}{\partial \phi^2} B_r}{r^2 \sin(\theta)^2} + \frac{\cos(\theta) \left(\frac{\partial}{\partial \theta} B_r \right)}{r^2 \sin(\theta)} + \frac{\frac{\partial^2}{\partial \theta^2} B_r}{r^2} \\
& 2 \frac{\frac{\partial}{\partial \theta} B_r}{r^2} - \frac{\cos(\theta)^2 B_\theta}{r^2 \sin(\theta)^2} + \frac{\cos(\theta) \left(\frac{\partial}{\partial \theta} B_\theta \right)}{r^2 \sin(\theta)} - \frac{B_\theta}{r^2} + \frac{\frac{\partial^2}{\partial \theta^2} B_\theta}{r^2} - 2 \frac{\cos(\theta) \left(\frac{\partial}{\partial \phi} B_\phi \right)}{r^2 \sin(\theta)^2} + \frac{\frac{\partial^2}{\partial \phi^2} B_\theta}{r^2 \sin(\theta)^2} + 2 \frac{\frac{\partial}{\partial r} B_\theta}{r} + \left(\frac{\partial^2}{\partial r^2} B_\theta \right) \\
& 2 \frac{\frac{\partial}{\partial \phi} B_r}{\sin(\theta) r^2} + 2 \frac{\cos(\theta) \left(\frac{\partial}{\partial \phi} B_\theta \right)}{\sin(\theta)^2 r^2} + \frac{\frac{\partial^2}{\partial \phi^2} B_\phi}{r^2 \sin(\theta)^2} - \frac{\cos(\theta)^2 B_\phi}{r^2 \sin(\theta)^2} + \frac{\cos(\theta) \left(\frac{\partial}{\partial \theta} B_\phi \right)}{r^2 \sin(\theta)} - \frac{B_\phi}{r^2} + \frac{\frac{\partial^2}{\partial \theta^2} B_\phi}{r^2} + 2 \frac{\frac{\partial}{\partial r} B_\phi}{r} + \left(\frac{\partial^2}{\partial r^2} B_\phi \right)
\end{aligned}$$

and each expression contains nine terms.

The curious reader might wonder *why*, in each case, five of the nine terms of the vector Laplacian components can be represented by the scalar Laplacian acting on the component. This question is answered in the following section.

(d) The second method : Part I

In the first method, described in section (a) above, we used this tensorization of the vector Laplacian,

$$(\star B)^n = (B'^j{}_{;j})^{;n} - g'^{-1/2} \epsilon'^{nab} (g'^{-1/2} g'_{bc} \epsilon'^{cde} B'_e{}_{;d})_{;a} .$$

At the end of Section 15 (b) it was noted that there is an alternative tensorization , namely

$$(\star B)^n = B'^n{}_{;j}{}^{;j} .$$

These two tensors must be the same since the tensorization of a Cartesian form equation is unique, but it is not so easy to show. Be that as it may, our "second method" is to use this $B'^n{}_{;j}{}^{;j}$ tensorization to compute once again the components of the vector Laplacian.

Appendix F (i) has among its examples the following covariant derivative of a rank 2 tensor,

$$B^{ab}{}_{;\alpha} \equiv \partial_\alpha B^{ab} + \Gamma^a_{\alpha n} B^{nb} + \Gamma^b_{\alpha n} B^{an}$$

and since $B^{a;b}$ is a rank-2 tensor one can write (indices are substituted in the second line)

$$\begin{aligned} B^{a;b}{}_{;\alpha} &\equiv \partial_\alpha B^{a;b} + \Gamma^a_{\alpha k} B^{k;b} + \Gamma^b_{\alpha k} B^{a;k} \\ B^{n;j}{}_{;j} &\equiv \partial_j B^{n;j} + \Gamma^n_{jk} B^{k;j} + \Gamma^j_{jk} B^{n;k} \quad // \Gamma^j_{jk} = (1/\sqrt{g}) \partial_k(\sqrt{g}) \text{ as in App F (d)} . \end{aligned}$$

Another example in that Appendix shows that (the lower three lines are index substitutions of the first)

$$\begin{aligned} B^{a;\alpha} &= \partial^\alpha B^a + g^{ab} \Gamma^a_{bs} B^s \\ B^{n;j} &= \partial^j B^n + g^{jb} \Gamma^n_{bs} B^s \\ B^{k;j} &= \partial^j B^k + g^{jb} \Gamma^k_{bs} B^s \\ B^{n;k} &= \partial^k B^n + g^{kb} \Gamma^n_{bs} B^s \end{aligned}$$

Therefore,

$$B^{n;j}{}_{;j} = \partial_j [\partial^j B^n + g^{jb} \Gamma^n_{bs} B^s] + \Gamma^n_{jk} [\partial^j B^k + g^{jb} \Gamma^k_{bs} B^s] + \Gamma^j_{jk} [\partial^k B^n + g^{kb} \Gamma^n_{bs} B^s]$$

Combining the second last term with the first gives

$$\begin{aligned} B^{n;j}{}_{;j} &= [\partial_j \partial^j B^n + \Gamma^j_{jk} (\partial^k B^n)] \\ &\quad + \partial_j (g^{jb} \Gamma^n_{bs} B^s) + \Gamma^n_{jk} [\partial^j B^k + g^{jb} \Gamma^k_{bs} B^s] + \Gamma^j_{jk} [g^{kb} \Gamma^n_{bs} B^s] \end{aligned}$$

Since no terms have been dropped, we are still "covariant" and in x' -space everything gets primed,

$$\begin{aligned} B'^n{}_{;j}{}^{;j} &= [\partial'_j \partial'^j B'^n + \Gamma'^j_{jk} (\partial'^k B'^n)] \\ &\quad + \partial'_j (g'^{jb} \Gamma'^n_{bs} B'^s) + \Gamma'^n_{jk} [\partial'^j B'^k + g'^{jb} \Gamma'^k_{bs} B'^s] + \Gamma'^j_{jk} [g'^{kb} \Gamma'^n_{bs} B'^s] \end{aligned}$$

The first two terms can be written this way,

$$[\partial'_j \partial'^j B'^n + \Gamma'^j_{jk} (\partial'^k B'^n)] = [\partial'_j \partial'^j B'^n + (1/\sqrt{g'}) \partial'_k (\sqrt{g'}) (\partial'^k B'^n)] = \text{lap} (B'^n)$$

in the sense that we earlier wrote (Section 15 (e)),

$$\partial'_j \partial'^j f' + (1/\sqrt{g'}) \partial'_k (\sqrt{g'}) \partial'^k f = (1/\sqrt{g'}) \partial'_k [\sqrt{g'} (\partial'^k f)] = \text{lap} (f)$$

Therefore

$$\begin{aligned} B'^n{}_{;j}{}^j &= \text{lap} (B'^n) + \partial'_j (g'^{jb} \Gamma'^n_{bs} B'^s) + \Gamma'^n_{jk} [\partial'^j B'^k + g'^{jb} \Gamma'^k_{bs} B'^s] + \Gamma'^j_{jk} [g'^{kb} \Gamma'^n_{bs} B'^s] \\ &= \text{lap} (B'^n) + \text{Extra Terms} \end{aligned}$$

So we begin to see why the scalar Laplacian of a B component appears in $(\star B)^n$.

(e) The second method : Part II

Unfortunately, we don't want to see $\text{lap} (B'^n)$, we want to see $\text{lap}(\mathcal{B}'^n)$! Consider then,

$$\text{lap} (B'^n) = \text{lap} (\mathcal{B}'^n/h_n) = [\partial'_j \partial'^j (\mathcal{B}'^n/h_n) + (1/\sqrt{g'}) \partial'_k (\sqrt{g'}) (\partial'^k (\mathcal{B}'^n/h_n))]]$$

and one computes the pieces as follows:

$$\begin{aligned} \partial'_j \partial'^j (\mathcal{B}'^n/h_n) &= \partial'_j \partial'^j (\mathcal{B}'^n h^{-1}_n) = \partial'_j [(\partial'^j \mathcal{B}'^n) h^{-1}_n + \mathcal{B}'^n (\partial'^j h^{-1}_n)] \\ &= (\partial'_j \partial'^j \mathcal{B}'^n) h^{-1}_n + (\partial'^j \mathcal{B}'^n) (\partial'_j h^{-1}_n) + (\partial'_j \mathcal{B}'^n) (\partial'^j h^{-1}_n) + \mathcal{B}'^n (\partial'_j \partial'^j h^{-1}_n) \\ &= (\partial'_j \partial'^j \mathcal{B}'^n) h^{-1}_n + 2(\partial'^j \mathcal{B}'^n) (\partial'_j h^{-1}_n) + \mathcal{B}'^n (\partial'_j \partial'^j h^{-1}_n) \end{aligned}$$

$$\partial'^k (\mathcal{B}'^n/h_n) = [(\partial'^k \mathcal{B}'^n) h^{-1}_n + \mathcal{B}'^n (\partial'^k h^{-1}_n)]$$

so that

$$\begin{aligned} \text{lap} (B'^n) &= (\partial'_j \partial'^j \mathcal{B}'^n) h^{-1}_n + 2(\partial'^j \mathcal{B}'^n) (\partial'_j h^{-1}_n) + \mathcal{B}'^n (\partial'_j \partial'^j h^{-1}_n) \\ &\quad + (1/\sqrt{g'}) \partial'_k (\sqrt{g'}) [(\partial'^k \mathcal{B}'^n) h^{-1}_n + \mathcal{B}'^n (\partial'^k h^{-1}_n)] \\ &= h^{-1}_n \{ (\partial'_j \partial'^j \mathcal{B}'^n) + (1/\sqrt{g'}) \partial'_k (\sqrt{g'}) (\partial'^k \mathcal{B}'^n) \} \\ &\quad + 2(\partial'^j \mathcal{B}'^n) (\partial'_j h^{-1}_n) + \mathcal{B}'^n (\partial'_j \partial'^j h^{-1}_n) + (1/\sqrt{g'}) \partial'_k (\sqrt{g'}) \mathcal{B}'^n (\partial'^k h^{-1}_n) \\ &= h^{-1}_n \text{lap} (\mathcal{B}'^n) \\ &\quad + 2(\partial'^j \mathcal{B}'^n) (\partial'_j h^{-1}_n) + \mathcal{B}'^n (\partial'_j \partial'^j h^{-1}_n) + (1/\sqrt{g'}) \partial'_k (\sqrt{g'}) \mathcal{B}'^n (\partial'^k h^{-1}_n) \\ &= h^{-1}_n \text{lap} (\mathcal{B}'^n) + \text{Other Terms} \end{aligned}$$

The conclusion so far is that

$$\begin{aligned} (\star \mathbf{B})^n &= B^{n,j}_{,j} = \text{lap}(\mathbf{B}^n) + \text{Extra Terms} \\ &= h^{-1}_n \text{lap}(\mathcal{B}^n) + \text{Other Terms} + \text{Extra Terms} \end{aligned}$$

As before, our interest is with the expansion $\star \mathbf{B} = (\star \mathcal{B})^n \hat{\mathbf{e}}_n$ where then

$$(\star \mathcal{B})^n = \text{lap}(\mathcal{B}^n) + h'_n [\text{Other Terms} + \text{Extra Terms}]$$

This result applies to any x' -space curvilinear coordinate system, orthogonal or otherwise. Thus we have demonstrated why it is that $\text{lap}(\mathcal{B}^n)$ always appears as part of the vector Laplacian component $(\star \mathcal{B})^n$.

The Terms shown are these, just quoting from above,

$$\text{Extra Terms} = \partial'_j (g'^{jb} \Gamma^n_{bs} B'^s) + \Gamma^n_{jk} [\partial'^j B'^k + g'^{jb} \Gamma'^k_{bs} B'^s] + \Gamma'^j_{jk} [g'^{kb} \Gamma^n_{bs} B'^s]$$

$$\text{Other Terms} = 2(\partial'^j \mathcal{B}^n) (\partial'_j h^{-1}_n) + \mathcal{B}^n (\partial'^j \partial'_j h^{-1}_n) + (1/\sqrt{g'}) \partial'_k (\sqrt{g'}) \mathcal{B}^n (\partial'^k h^{-1}_n)$$

Rather than attempt algebraic simplification of the above terms, we will just throw them into Maple as is. Replacing $B'^s = \mathcal{B}^s/h'_s$ and "lowering" the differential operators appropriately, one gets

Extra Terms =

$$\begin{aligned} \partial'_j (g'^{jb} \Gamma^n_{bs} \mathcal{B}^s/h'_s) &+ \Gamma^n_{jk} g'^{js} \partial'_s (\mathcal{B}^k/h'_k) &+ \Gamma^n_{jk} g'^{jb} \Gamma'^k_{bs} \mathcal{B}^s/h'_s &+ \Gamma'^j_{jk} g'^{kb} \Gamma^n_{bs} \mathcal{B}^s/h'_s \\ \text{ET1} &\qquad\qquad\qquad \text{ET2} &\qquad\qquad\qquad \text{ET3} &\qquad\qquad\qquad \text{ET4} \end{aligned}$$

Other Terms =

$$\begin{aligned} 2(g'^{js} \partial'_s \mathcal{B}^n) (\partial'_j h_n^{-1}) &+ \mathcal{B}^n (\partial'_j \partial'^j h_n^{-1}) &+ (1/\sqrt{g'}) \partial'_k (\sqrt{g'}) \mathcal{B}^n (\partial'^k h_n^{-1}) \\ \text{OT1} &\qquad\qquad\qquad \text{OT2} &\qquad\qquad\qquad \text{OT3} \end{aligned}$$

(f) The second method in spherical coordinates: Maple speaks again

In method 1, there was no need in the Maple calculation for the affine connection object ,

$$\Gamma'^d_{ab} = (1/2) g'^{dc} [\partial'_a g'_{bc} + \partial'_b g'_{ca} - \partial'_c g'_{ab}]$$

which in orthogonal coordinates simplifies to

$$\Gamma'^d_{ab} = (1/2) h'^d_d{}^{-2} [\delta^d_b \partial'_a (h'^2_b) + \delta^d_a \partial'_b (h'^2_a) - \delta_{ab} \partial'_d (h'^2_a)] .$$

This last form shows that $\Gamma'^d_{ab} = 0$ unless two indices match, in which case it *might* not vanish. For coordinate systems like spherical and cylindrical coordinates, Γ'^d_{ab} is very **sparse**, and this explains why we are not going to be simply swamped by all those Extra Terms shown above.

For spherical coordinates, only 9 of the 27 elements of the object Γ'^d_{ab} are non-zero :

$$\begin{aligned}
 \Gamma^1_{22} &= -r & \Gamma^2_{12} &= \Gamma^2_{21} = 1/r & // \text{ notation: } & \Gamma^1_{22} = \Gamma^r_{\theta\theta} \\
 \Gamma^1_{33} &= -r \sin^2\theta & \Gamma^3_{13} &= \Gamma^3_{31} = 1/r \\
 \Gamma^2_{33} &= -\cos\theta\sin\theta & \Gamma^3_{23} &= \Gamma^3_{32} = \cot\theta & // 1,2,3 &= r,\theta,\varphi = \text{radius, polar, azimuthal}
 \end{aligned}$$

$$\Gamma^r = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -r & 0 \\ 0 & 0 & -r\sin^2\theta \end{bmatrix} \quad \Gamma^\theta = \begin{bmatrix} 0 & 1/r & 0 \\ 1/r & 0 & 0 \\ 0 & 0 & -\sin\theta\cos\theta \end{bmatrix} \quad \Gamma^\varphi = \begin{bmatrix} 0 & 0 & 1/r \\ 0 & 0 & \cot\theta \\ 1/r & \cot\theta & 0 \end{bmatrix}$$

To maintain generality, however, we let Maple compute $\Gamma^d_{ab} = G(d,a,b)$ from the first equation above, so

$$\Gamma^d_{ab} = (1/2) g^{dc} [\partial_a g_{bc} + \partial_b g_{ca} - \partial_c g_{ab}]$$

Compute the affine connection.

gcontra := inverse(gcov);

$$gcontra = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{r^2} & 0 \\ 0 & 0 & -\frac{1}{r^2(-1+\cos(\theta)^2)} \end{bmatrix}$$

G := (d,a,b) -> (1/2)*sum(gcontra[d,c]*(Diff(gcov[b,c],xp[a]) + Diff(gcov[c,a],xp[b]) - Diff(gcov[a,b],xp[c])),c=1..N);

$$G := (d, a, b) \rightarrow \frac{1}{2} \left(\sum_{c=1}^N gcontra_{d,c} \left(\left(\frac{\partial}{\partial xp_a} gcov_{b,c} \right) + \left(\frac{\partial}{\partial xp_b} gcov_{c,a} \right) - \left(\frac{\partial}{\partial xp_c} gcov_{a,b} \right) \right) \right)$$

value(G(3,3,1)): simplify(%);

$$\frac{1}{r}$$

The Extra Terms are then entered, as shown above :

(continued on next page)

Extra Terms =

$$\begin{array}{cccc} \partial'_j (g'^{jb} \Gamma_{bs}^n \mathcal{B}'^s / h'_s) & + \Gamma_{jk}^n g'^{js} \partial'_s (\mathcal{B}'^k / h'_k) & + \Gamma_{jk}^n g'^{jb} \Gamma_{bs}^k \mathcal{B}'^s / h'_s & + \Gamma'^j_{jk} g'^{kb} \Gamma_{bs}^n \mathcal{B}'^s / h'_s \\ \text{ET1} & \text{ET2} & \text{ET3} & \text{ET4} \end{array}$$

Compute ET = "the extra terms"

```
ET1 := (n) ->
sum(sum(sum(Diff(gcontra[j,b]*G(n,b,s)*Bp[s](xp[1],xp[2],xp[3])/hp[s],xp[j]),b=1..N),j=1..N),
s=1..N);
```

$$ET1 = n \rightarrow \sum_{s=1}^N \left(\sum_{j=1}^N \left(\sum_{b=1}^N \left(\frac{\partial}{\partial x_j} \frac{g_{contra,j,b} G(n,b,s) B_p(s)(x_1, x_2, x_3)}{h_p(s)} \right) \right) \right)$$

```
ET2 := (n) ->
```

```
sum(sum(sum(G(n,j,k)*gcontra[j,s]*Diff(Bp[k](xp[1],xp[2],xp[3])/hp[k],xp[s]),j=1..N),k=1..N),
s=1..N);
```

$$ET2 = n \rightarrow \sum_{s=1}^N \left(\sum_{k=1}^N \left(\sum_{j=1}^N G(n,j,k) g_{contra,j,s} \left(\frac{\partial}{\partial x_s} \frac{B_p(k)(x_1, x_2, x_3)}{h_p(k)} \right) \right) \right)$$

```
ET3 := (n) -> sum(sum(sum(sum(
```

```
G(n,j,k)*gcontra[j,b]*G(k,b,s)*Bp[s](xp[1],xp[2],xp[3])/hp[s],b=1..N),j=1..N),s=1..N),k=1..N)
;
```

$$ET3 = n \rightarrow \sum_{k=1}^N \left(\sum_{s=1}^N \left(\sum_{j=1}^N \left(\sum_{b=1}^N \frac{G(n,j,k) g_{contra,j,b} G(k,b,s) B_p(s)(x_1, x_2, x_3)}{h_p(s)} \right) \right) \right)$$

```
ET4 := (n) -> sum(sum(sum(sum(
```

```
G(j,j,k)*gcontra[k,b]*G(n,b,s)*Bp[s](xp[1],xp[2],xp[3])/hp[s],b=1..N),j=1..N),s=1..N),k=1..N)
;
```

$$ET4 = n \rightarrow \sum_{k=1}^N \left(\sum_{s=1}^N \left(\sum_{j=1}^N \left(\sum_{b=1}^N \frac{G(j,j,k) g_{contra,k,b} G(n,b,s) B_p(s)(x_1, x_2, x_3)}{h_p(s)} \right) \right) \right)$$

And then the Other Terms are entered as well,

Other Terms =

$$\begin{array}{cccc} 2(g'^{js} \partial'_s \mathcal{B}'^n) (\partial'_j h'_n)^{-1} & + \mathcal{B}'^n (\partial'_j \partial'^j h'_n)^{-1} & + (1/\sqrt{g'}) \partial'_k (\sqrt{g'}) \mathcal{B}'^n (\partial'^k h'_n)^{-1} & \\ \text{OT1} & \text{OT2} & \text{OT3} & \dots \end{array}$$

```
OT1 := (n) ->
```

```
sum(sum(2*Diff(1/hp[n],xp[j])*gcontra[j,s]*Diff(Bp[n](xp[1],xp[2],xp[3]),xp[s]),j=1..N),s=1..N);
```

$$OT1 = n \rightarrow \sum_{s=1}^N \left(\sum_{j=1}^N \left(2 \left(\frac{\partial}{\partial x_j} \frac{1}{h_p(n)} \right) g_{contra,j,s} \left(\frac{\partial}{\partial x_s} B_p(n)(x_1, x_2, x_3) \right) \right) \right)$$

```
OT2 := (n) -> sum(sum(gcontra[j,s]*
```

```
Diff(Diff(1/hp[n],xp[j]),xp[s])*Bp[n](xp[1],xp[2],xp[3]),j=1..N),s=1..N);
```

$$OT2 = n \rightarrow \sum_{s=1}^N \left(\sum_{j=1}^N g_{contra,j,s} \left(\frac{\partial}{\partial x_s} \left(\frac{\partial}{\partial x_j} \frac{1}{h_p(n)} \right) \right) B_p(n)(x_1, x_2, x_3) \right)$$

```
OT3 := (n) ->
```

```
sum(sum((1/sq_gp)*Diff(sq_gp,xp[j])*gcontra[j,s]*Diff(1/hp[n],xp[s])*Bp[n](xp[1],xp[2],xp[3]),j=1..N),s=1..N);
```

$$OT3 = n \rightarrow \sum_{s=1}^N \left(\sum_{j=1}^N \frac{\left(\frac{\partial}{\partial x_j} \frac{sq_gp}{h_p(n)} \right) g_{contra,j,s} \left(\frac{\partial}{\partial x_s} \frac{1}{h_p(n)} \right) B_p(n)(x_1, x_2, x_3)}{sq_gp} \right)$$

The final results are then generated :

```

starBx := (n) -> hp[n]*(ET1(n)+ET2(n)+ET3(n)+ET4(n)+OT1(n)+OT2(n)+OT3(n));
starBx = n -> hp_n(ET1(n)+ET2(n)+ET3(n)+ET4(n)+OT1(n)+OT2(n)+OT3(n))
value(starBx(1)):simplify(%):expand(%);

```

$$-2 \frac{\cos(\theta) B_{\theta}(r, \theta, \phi)}{\sin(\theta) r^2} - 2 \frac{\frac{\partial}{\partial \theta} B_{\theta}(r, \theta, \phi)}{r^2} - 2 \frac{\frac{\partial}{\partial \phi} B_{\phi}(r, \theta, \phi)}{\sin(\theta) r^2} - 2 \frac{B_r(r, \theta, \phi)}{r^2}$$

```

value(starBx(2)):expand(%):simplify(%):subs(cos(theta)^2=1-sin(theta)^2, %):expand(%);

```

$$2 \frac{\frac{\partial}{\partial \theta} B_r(r, \theta, \phi)}{r^2} - 2 \frac{\cos(\theta) \left(\frac{\partial}{\partial \phi} B_{\phi}(r, \theta, \phi) \right)}{r^2 \sin(\theta)^2} - \frac{B_{\theta}(r, \theta, \phi)}{r^2 \sin(\theta)^2}$$

```

value(starBx(3)):expand(%):simplify(%):subs(cos(theta)^2=1-sin(theta)^2, %):expand(%);

```

$$2 \frac{\cos(\theta) \left(\frac{\partial}{\partial \phi} B_{\theta}(r, \theta, \phi) \right)}{\sin(\theta)^2 r^2} - \frac{B_{\phi}(r, \theta, \phi)}{\sin(\theta)^2 r^2} + 2 \frac{\frac{\partial}{\partial \phi} B_r(r, \theta, \phi)}{\sin(\theta) r^2}$$

These are seen to match the unnumbered terms in the q1_,q2_,q3_ expressions of section (c) above, those terms which get added to the scalar Laplacian contribution. Quoting from method 1 above,

$$\begin{aligned}
 (\star B)_{\mathbf{r}} &= \nabla^2(B_{\mathbf{r}}) - (2/r^2) [B_{\mathbf{r}} + \cot\theta B_{\theta} + \partial_{\theta} B_{\theta} + \csc\theta \partial_{\phi} B_{\phi}] \\
 (\star B)_{\theta} &= \nabla^2(B_{\theta}) - (1/r^2) [\csc^2\theta B_{\theta} - 2 \partial_{\theta} B_{\mathbf{r}} + 2 \cot\theta \csc\theta \partial_{\phi} B_{\phi}] \\
 (\star B)_{\phi} &= \nabla^2(B_{\phi}) - (1/r^2) [\csc^2\theta B_{\phi} - 2 \csc\theta \partial_{\phi} B_{\mathbf{r}} - 2 \cot\theta \csc\theta \partial_{\theta} B_{\theta}]
 \end{aligned}$$

(g) Results for Cylindrical Coordinates from both methods

The Maple program was easily modified for this system. Here are the results:

$$\begin{aligned}
 \mathcal{B}^1 &= B_{\mathbf{r}} & \mathcal{B}^2 &= B_{\phi} & \mathcal{B}^3 &= B_z \\
 \mathbf{B} &= \mathcal{B}^n \hat{\mathbf{e}}_n = B_{\mathbf{r}} \hat{\mathbf{e}}_{\mathbf{r}} + B_{\phi} \hat{\mathbf{e}}_{\phi} + B_z \hat{\mathbf{e}}_z = B_{\mathbf{r}} \hat{\mathbf{r}} + B_{\theta} \hat{\boldsymbol{\phi}} + B_z \hat{\mathbf{z}} .
 \end{aligned}$$

The scalar Laplacian of an unspecified function f :

$$\begin{aligned}
 \nabla^2 f &= (1/r) \partial_{\mathbf{r}}(r f) + (1/r^2) \partial_{\phi}^2 f + \partial_z^2 f = \\
 &\quad \underset{1}{\frac{\partial}{\partial r}} + \underset{2}{\frac{\partial^2}{\partial r^2}} + \underset{3}{\frac{\partial^2}{\partial \phi^2}} + \underset{4}{\frac{\partial^2}{\partial z^2}} \\
 &\quad \underset{1}{\frac{\partial}{\partial r}} + \underset{2}{\left(\frac{\partial^2}{\partial r^2} \right)} + \underset{3}{\frac{\partial^2}{\partial \phi^2}} + \underset{4}{\left(\frac{\partial^2}{\partial z^2} \right)}
 \end{aligned}$$

The components of the vector Laplacian are found by the "first method" to be

q1_;

$$-\frac{B_r}{r^2} + \frac{\frac{\partial}{\partial r} B_r}{r} + \left(\frac{\partial^2}{\partial r^2} B_r \right) - 2 \frac{\frac{\partial}{\partial \phi} B_\phi}{r^2} + \left(\frac{\partial^2}{\partial z^2} B_r \right) + \frac{\frac{\partial^2}{\partial \phi^2} B_r}{r^2}$$

1 2 4 3

Therefore,

$$(\star \mathbf{B})_{\mathbf{r}} = \nabla^2(B_{\mathbf{r}}) - B_{\mathbf{r}}/r^2 - 2\partial_\phi B_\phi/r^2.$$

q2_;

$$2 \frac{\frac{\partial}{\partial \phi} B_r}{r^2} + \frac{\frac{\partial^2}{\partial \phi^2} B_\phi}{r^2} + \left(\frac{\partial^2}{\partial z^2} B_\phi \right) - \frac{B_\phi}{r^2} + \frac{\frac{\partial}{\partial r} B_\phi}{r} + \left(\frac{\partial^2}{\partial r^2} B_\phi \right)$$

3 4 1 2

Therefore,

$$(\star \mathbf{B})_\phi = \nabla^2(B_\phi) - B_\phi/r^2 + 2\partial_\phi B_{\mathbf{r}}/r^2.$$

q3_;

$$\left(\frac{\partial^2}{\partial z^2} B_z \right) + \frac{\frac{\partial^2}{\partial \phi^2} B_z}{r^2} + \frac{\frac{\partial}{\partial r} B_z}{r} + \left(\frac{\partial^2}{\partial r^2} B_z \right)$$

4 3 1 2

Therefore,

$$(\star \mathbf{B})_{\mathbf{z}} = \nabla^2(B_{\mathbf{z}})$$

as befits a component which is Cartesian. In this manner, we end up with the components of the vector Laplacian expressed in the traditional manner,

$$(\star \mathbf{B})_{\mathbf{r}} = \nabla^2(B_{\mathbf{r}}) - (1/r^2)[B_{\mathbf{r}} - 2\partial_\phi B_\phi]$$

$$(\star \mathbf{B})_\phi = \nabla^2(B_\phi) - (1/r^2)[B_\phi + 2\partial_\phi B_{\mathbf{r}}]$$

$$(\star \mathbf{B})_{\mathbf{z}} = \nabla^2(B_{\mathbf{z}})$$

$$\text{where } \nabla^2 f = (1/r)\partial_{\mathbf{r}}(rf) + (1/r^2)\partial_\phi^2 f + \partial_{\mathbf{z}}^2 f$$

$$= \partial_{\mathbf{r}}^2 f + (1/r)\partial_{\mathbf{r}} f + (1/r^2)\partial_\phi^2 f + \partial_{\mathbf{z}}^2 f$$

Once again, written out in gory detail, these three expressions are (in the same order r, ϕ , z) :

$$\begin{aligned}
 & -\frac{B_r}{r^2} + \frac{\frac{\partial}{\partial r} B_r}{r} + \left(\frac{\partial^2}{\partial r^2} B_r \right) - 2 \frac{\frac{\partial}{\partial \phi} B_\phi}{r^2} + \left(\frac{\partial^2}{\partial z^2} B_r \right) + \frac{\frac{\partial^2}{\partial \phi^2} B_r}{r^2} \\
 & 2 \frac{\frac{\partial}{\partial \phi} B_r}{r^2} + \frac{\frac{\partial^2}{\partial \phi^2} B_\phi}{r^2} + \left(\frac{\partial^2}{\partial z^2} B_\phi \right) - \frac{B_\phi}{r^2} + \frac{\frac{\partial}{\partial r} B_\phi}{r} + \left(\frac{\partial^2}{\partial r^2} B_\phi \right) \\
 & \left(\frac{\partial^2}{\partial z^2} B_z \right) + \frac{\frac{\partial^2}{\partial \phi^2} B_z}{r^2} + \frac{\frac{\partial}{\partial r} B_z}{r} + \left(\frac{\partial^2}{\partial r^2} B_z \right)
 \end{aligned}$$

The "second method" produces these results for the terms which are added to the scalar Laplacian,

```

starBx := (n) ->
hp[n]*(ET1(n)+ET2(n)+ET3(n)+ET4(n)+OT1(n)+OT2(n)+OT3(n));
starBx := n -> hp_n(ET1(n)+ET2(n)+ET3(n)+ET4(n)+OT1(n)+OT2(n)+OT3(n))
value(starBx(1)):simplify(%):expand(%);

```

$$-2 \frac{\frac{\partial}{\partial \phi} B_\phi(r, \phi, z)}{r^2} - \frac{B_r(r, \phi, z)}{r^2}$$

```

value(starBx(2)):expand(%):simplify(%):subs(cos(phi)^2=1-sin(phi)^2, %):expand(%);

```

$$2 \frac{\frac{\partial}{\partial \phi} B_r(r, \phi, z)}{r^2} - \frac{B_\phi(r, \phi, z)}{r^2}$$

```

value(starBx(3)):expand(%):simplify(%):subs(cos(phi)^2=1-sin(phi)^2, %):expand(%);

```

0

and these are seen to agree with the unnumbered terms in q1_, q2_ and q3_ shown above.

In cylindrical coordinates the Γ object is even sparser than in spherical coordinates. One has

$$\Gamma'^{\mathbf{d}}_{\mathbf{ab}} = (1/2) h'^{\mathbf{d}}_{\mathbf{d}}{}^{-2} [\delta'^{\mathbf{d}}_{\mathbf{b}} \partial'_{\mathbf{a}}(h'^{\mathbf{b}}_{\mathbf{b}}{}^2) + \delta'^{\mathbf{d}}_{\mathbf{a}} \partial'_{\mathbf{b}}(h'^{\mathbf{a}}_{\mathbf{a}}{}^2) - \delta_{\mathbf{ab}} \partial'_{\mathbf{d}}(h'^{\mathbf{a}}_{\mathbf{a}}{}^2)] .$$

$$\Gamma'^1_{22} = -r \quad // \text{ notation: } \Gamma'^1_{22} = \Gamma^{\mathbf{r}}_{\varphi\varphi}$$

$$\Gamma'^2_{12} = 1/r$$

$$\Gamma'^2_{21} = 1/r \quad // 1,2,3 = r,\varphi,z$$

$$\Gamma^{\mathbf{r}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -r & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \Gamma^{\varphi} = \begin{bmatrix} 0 & 1/r & 0 \\ 1/r & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \Gamma^{\mathbf{z}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

so only 3 of 27 components are non-vanishing. Recall from Appendix F (a) that Γ just describes how the tangent base vectors move as $\mathbf{x}'$ moves, $(\partial'_j \mathbf{e}_n(\mathbf{x}')) = \Gamma'^{\mathbf{k}}_{jn}(\mathbf{x}') \mathbf{e}_k(\mathbf{x}')$, and cylindrical coordinates is just polar coordinates with z tacked on so the tangent base vectors don't move much, $\mathbf{e}_3 = \hat{\mathbf{z}}$ not at all.

$$d\mathbf{T}^{ij}(\mathbf{x},t)/dt = \partial \mathbf{T}^{ij}/\partial t + (\partial \mathbf{T}^{ij}/\partial x^{\mathbf{k}}) (\partial x^{\mathbf{k}}/dt) = \partial \mathbf{T}^{ij}/\partial t + (\partial \mathbf{T}^{ij}/\partial x^{\mathbf{k}}) v^{\mathbf{k}}$$

or

$$d_{\mathbf{t}} \mathbf{T}^{ij} = \partial_{\mathbf{t}} \mathbf{T}^{ij} + (\partial_{\mathbf{k}} \mathbf{T}^{ij}) v^{\mathbf{k}} . \quad // d_{\mathbf{t}} \equiv d/dt, \partial_{\mathbf{t}} \equiv \partial/\partial t, \partial_{\mathbf{k}} \equiv \partial/\partial x^{\mathbf{k}}$$

We take this as a useful prototype equation to work with for two reasons. First, it contains our object of interest, which is the gradient of a rank-2 tensor. Second, this equation plays a role in the continuum mechanics of non-Newtonian fluids as discussed below in section (f).

One can define, in Cartesian coordinates, a $(\nabla \mathbf{T})$ object :

$$(\nabla \mathbf{T})^{ij}_{\mathbf{k}} \equiv \partial_{\mathbf{k}} \mathbf{T}^{ij}$$

so that

$$d_{\mathbf{t}} \mathbf{T}^{ij} = \partial_{\mathbf{t}} \mathbf{T}^{ij} + (\nabla \mathbf{T})^{ij}_{\mathbf{k}} v^{\mathbf{k}} .$$

Comment: Our convention has been to bold vectors and not to bold other tensors. In line with this idea, we shall write $\nabla \mathbf{T}$ where the grad is bolded and the $\mathbf{T}$ is not bolded.

In order to express the above equation in curvilinear coordinates, it must be "tensorized" in the sense of Section 15 (b) so that the equation is covariant. Thus, $(\nabla \mathbf{T})^{ij}_{\mathbf{k}}$ must be regarded as components of a (mixed) rank-3 tensor which, in Cartesian coordinates, are equal to $\partial_{\mathbf{k}} \mathbf{T}^{ij}$. Since $v^{\mathbf{k}}$ are the components of a tensorial vector, $(\nabla \mathbf{T})^{ij}_{\mathbf{k}} v^{\mathbf{k}}$ transforms as a rank-2 tensor, and then all terms in the above equation are rank-2 tensors and the equation is then covariant and therefore appears this way in x' -space,

$$d_{\mathbf{t}} \mathbf{T}'^{ij} = \partial_{\mathbf{t}} \mathbf{T}'^{ij} + (\nabla \mathbf{T})'^{ij}_{\mathbf{k}} v'^{\mathbf{k}} .$$

In our usual formalism, x' -space is the space of some generic curvilinear coordinates x'^n (not necessarily orthogonal) and then the above equation tells us the form taken by the time derivative equation in curvilinear coordinates, and it remains only to compute the objects $(\nabla T)^{ij}{}_{\mathbf{k}}$.

For convenience, we can lower the tensorial ij indices on the above tensor equations to get

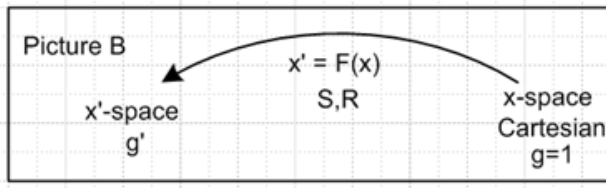
$$d_{\mathbf{t}} T_{ij} = \partial_{\mathbf{t}} T_{ij} + (\nabla T)_{ij\mathbf{k}} v^{\mathbf{k}} \quad (\nabla T)_{ij\mathbf{k}} \equiv \partial_{\mathbf{k}} T_{ij} \quad // \text{ Cartesian coordinates}$$

$$d_{\mathbf{t}} T'_{ij} = \partial_{\mathbf{t}} T'_{ij} + (\nabla T)'_{ij\mathbf{k}} v'^{\mathbf{k}} \quad // \text{ curvilinear coordinates}$$

and then we can deal with the pure covariant tensor components $(\nabla T)'_{ij\mathbf{k}}$.

Appendix J: Expansion of (∇T) in curvilinear coordinates ($T = \text{rank-2 tensor}$)

This appendix assumes the usual curvilinear coordinates context, Picture B

**(a) Total time derivative as prototype equation**

The total time derivative of a contravariant rank-2 tensor field $T^{ij}(\mathbf{x}, t)$ can be written as

$$dT^{ij}(\mathbf{x}, t)/dt = \partial T^{ij}/\partial t + (\partial T^{ij}/\partial x^k) (\partial x^k/dt) = \partial T^{ij}/\partial t + (\partial T^{ij}/\partial x^k) v^k$$

or

$$d_t T^{ij} = \partial_t T^{ij} + (\partial_k T^{ij}) v^k \quad // \quad d_t \equiv d/dt, \quad \partial_t \equiv \partial/\partial t, \quad \partial_k \equiv \partial/\partial x^k$$

We take this as a useful prototype equation to work with for two reasons. First, it contains our object of interest, which is the gradient of a rank-2 tensor. Second, this equation plays a role in the continuum mechanics of non-Newtonian fluids as discussed below in section (f).

One can define, in Cartesian coordinates, a (∇T) object :

$$(\nabla T)^{ij}_k \equiv \partial_k T^{ij}$$

so that

$$d_t T^{ij} = \partial_t T^{ij} + (\nabla T)^{ij}_k v^k.$$

Comment: Our convention has been to bold vectors and not to bold other tensors. In line with this idea, we shall write ∇T where the grad is bolded and the T is not bolded.

In order to express the above equation in curvilinear coordinates, it must be "tensorized" in the sense of Section 15 (b) so that the equation is covariant. Thus, $(\nabla T)^{ij}_k$ must be regarded as components of a (mixed) rank-3 tensor which, in Cartesian coordinates, are equal to $\partial_k T^{ij}$. Since v^k are the components of a tensorial vector, $(\nabla T)^{ij}_k v^k$ transforms as a rank-2 tensor, and then all terms in the above equation are rank-2 tensors and the equation is then covariant and therefore appears this way in x' -space,

$$d_t T'^{ij} = \partial'_t T'^{ij} + (\nabla' T')^{ij}_k v'^k.$$

In our usual formalism, x' -space is the space of some generic curvilinear coordinates x'^n (not necessarily orthogonal) and then the above equation tells us the form taken by the time derivative equation in curvilinear coordinates, and it remains only to compute the objects $(\nabla' T')^{ij}_k$.

For convenience, we can lower the tensorial ij indices on the above tensor equations to get

$$d_t T_{ij} = \partial_t T_{ij} + (\nabla T)_{ijk} v^k \quad (\nabla T)_{ijk} \equiv \partial_k T_{ij} \quad // \text{ Cartesian coordinates}$$

$$d_t T'_{ij} = \partial_t T'_{ij} + (\nabla T)'_{ijk} v'^k \quad // \text{ curvilinear coordinates}$$

and then we can deal with the pure covariant tensor components $(\nabla T)'_{ijk}$.

(b) Computation of components $(\nabla T)'_{ijk}$

The covariant derivative is discussed in Appendix F (g,h,i). We shall define a true tensor object $(\nabla T)_{abc}$ as $T_{ab;\alpha}$ so that (see App F, J=2 example with $B_{ab;\alpha}$, take $B \rightarrow T$ and $\alpha \rightarrow c$)

$$(\nabla T)_{abc} = T_{ab;c} \equiv T_{ab,c} - \Gamma_{ac}^n T_{nb} - \Gamma_{bc}^n T_{an} = T_{ab,c} = \partial_c T_{ab} \quad // \text{ x-space}$$

$$(\nabla T)'_{abc} = T'_{ab;c} \equiv T'_{ab,c} - \Gamma'^n_{ac} T'_{nb} - \Gamma'^n_{bc} T'_{an} \quad // \text{ x'-space}$$

Recall that $T_{ab,c}$ is a shorthand for $\partial_c T_{ab}$ and that Γ^c_{ab} is the affine connection which tells how the basis vectors change as one moves around in space. In Cartesian x-space (first line above) the basis vectors are the fixed $\mathbf{u}_n$ which don't change, so $\Gamma^c_{ab} \equiv 0$. In curvilinear x'-space, $\Gamma'^c_{ab} \neq 0$. Since $T_{ab;c}$ transforms as a true rank-3 tensor, its defining equation is "covariant" (Section 7 (u)) so that in x'-space the equation has exactly the same form but everything is primed.

The above two lines characterize the process of "tensorization": we find a true "tensorial tensor" $T_{ab;c}$ which agrees with $(\nabla T)_{abc} = \partial_c T_{ab}$ in Cartesian space. The tensorized version of $T_{ab,c}$ is unique, and it is $T_{ab;c}$. Since this tensor is given by $T'_{ab;c}$ in x'-space, we use the second line above to compute the components of the tensor (∇T) object in x'-space, which is to say, in curvilinear coordinates.

It is a simple matter to have Maple compute Γ'^c_{ab} for any curvilinear coordinate system, and then the second line above reports out the components $(\nabla T)'_{abc}$:

$$(\nabla T)'_{abc} = \partial'_c T'_{ab} - \Gamma'^n_{ac} T'_{nb} - \Gamma'^n_{bc} T'_{an} \quad (*)$$

$$\Gamma'^d_{ab} = (1/2) g'^{dc} [\partial'_a g'_{bc} + \partial'_b g'_{ca} - \partial'_c g'_{ab}] \quad // \text{ Appendix F (d), } g' = \text{metric tensor}$$

and then we know how to write the time derivative equation in any curvilinear coordinates

$$d_t T'_{ij} = \partial_t T'_{ij} + (\nabla T)'_{ijk} v'^k$$

or

$$d_t T'^{ij} = \partial_t T'^{ij} + (\nabla T)^{ij}{}_k v'^k \quad (\nabla T)^{ij}{}_k = g'^{ii'} g'^{jj'} (\nabla T)'_{i'j'k}$$

In Appendix I (b) we show Maple code to compute the covariant metric tensor g'_{ij} from the curvilinear coordinates defining equations (such as $x = r \sin \theta \cos \phi$ for sphericals). Then Appendix I (f) shows the extra code for computing g'^{ij} and Γ'^d_{ab} . The reader can then add a few extra lines to have Maple compute the

desired $(\nabla T)'_{abc}$ using the equation (*) above. Later in this appendix we shall use Maple to compute certain related quantities which we can then verify against a known source.

(c) Tensor expansions of ∇T on the $\mathbf{u}_n$ and $\mathbf{e}_n$ base vectors

It is useful at this point to write out the tensor expansions for the rank-3 tensor (∇T) to show exactly where the components $(\nabla T)'_{abc}$ appear. As shown in Appendix E (b) one can expand the rank-3 tensor $T_{ab;c}$ in various ways. The first line below shows expansions on x-space basis vectors, while the second line shows expansion on the reciprocal and tangent base vectors (which also exist in x-space) :

$$\begin{aligned}\nabla T &= \Sigma_{ijk} T_{ij;k} \mathbf{u}^i \otimes \mathbf{u}^j \otimes \mathbf{u}^k = \Sigma_{ijk} [(\nabla T)^{(u)}]_{ijk} \mathbf{u}^i \otimes \mathbf{u}^j \otimes \mathbf{u}^k = \Sigma_{ijk} [(\nabla T)^{(u)}]^{ijk} \mathbf{u}_i \otimes \mathbf{u}_j \otimes \mathbf{u}_k \\ \nabla T &= \Sigma_{ijk} T'_{ij;k} \mathbf{e}^i \otimes \mathbf{e}^j \otimes \mathbf{e}^k = \Sigma_{ijk} [(\nabla T)^{(e)}]_{ijk} \mathbf{e}^i \otimes \mathbf{e}^j \otimes \mathbf{e}^k = \Sigma_{ijk} [(\nabla T)^{(e)}]^{ijk} \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k\end{aligned}$$

As usual, up and down index positions are the same for the first line in Cartesian space, but are significant on the second line. The superscripts on the $(\nabla T)^{(u)}$ and $(\nabla T)^{(e)}$ components indicate which basis vectors are being expanded upon. One normally just writes $(\nabla T)^{(u)} = (\nabla T)$, and $(\nabla T)^{(e)} = (\nabla T)'$ where the prime indicates the curvilinear x'-space. So, we now have three different notations for the expansion components:

$$T_{ij;k} = [(\nabla T)^{(u)}]_{ijk} = (\nabla T)_{ijk}$$

$$T'_{ij;k} = [(\nabla T)^{(e)}]_{ijk} = (\nabla T)'_{ijk} .$$

Just to fill things out, here are the corresponding expansions for rank-2 tensor T ,

$$\begin{aligned}T &= \Sigma_{ij} T_{ij} \mathbf{u}^i \otimes \mathbf{u}^j = \Sigma_{ij} [T^{(u)}]_{ij} \mathbf{u}^i \otimes \mathbf{u}^j = \Sigma_{ij} [T^{(u)}]^{ij} \mathbf{u}_i \otimes \mathbf{u}_j \\ T &= \Sigma_{ij} T'_{ij} \mathbf{e}^i \otimes \mathbf{e}^j = \Sigma_{ij} [T^{(e)}]_{ij} \mathbf{e}^i \otimes \mathbf{e}^j = \Sigma_{ij} [T^{(e)}]^{ij} \mathbf{e}_i \otimes \mathbf{e}_j\end{aligned}$$

$$T_{ij} = [T^{(u)}]_{ij}$$

$$T'_{ij} = [T^{(e)}]_{ij}$$

and then for a rank-1 tensor $\mathbf{v}$,

$$\begin{aligned}\mathbf{v} &= \Sigma_i v_i \mathbf{u}^i = \Sigma_i [v^{(u)}]_i \mathbf{u}^i = \Sigma_i [v^{(u)}]^i \mathbf{u}_i \\ \mathbf{v} &= \Sigma_i v'_i \mathbf{e}^i = \Sigma_i [v^{(e)}]_i \mathbf{e}^i = \Sigma_i [v^{(e)}]^i \mathbf{e}_i\end{aligned}$$

$$v_i = [v^{(u)}]_i$$

$$v'_i = [v^{(e)}]_i$$

(d) Tensor expansions of ∇T on the $\hat{\mathbf{e}}_n$ base vectors

In practical applications, it is sometimes useful to deal with components of tensors which are expanded on the *unit versions* of the tangent base vectors, $\hat{\mathbf{e}}_n \equiv \mathbf{e}_n / |\mathbf{e}_n| = \mathbf{e}_n / h'_n$. On the one hand, this introduces

major complications (see below) since such components are non-covariant (neither contravariant nor covariant nor mixed). On the other hand, since the unit vectors are all dimensionless, all tensor components have the same dimensions, which is very useful in any practical engineering work. We shall refer to such tensor components as "unit-base-vector components".

This subject is addressed in Appendix E (h). We start with the expansion given above,

$$\nabla T = \Sigma_{ijk} [(\nabla T)^{(e)}]^{ijk} \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k$$

and process it in this manner ,

$$\begin{aligned} &= \Sigma_{ijk} [(\nabla T)^{(e)}]^{ijk} (h'_i \hat{\mathbf{e}}_i) \otimes (h'_j \hat{\mathbf{e}}_j) \otimes (h'_k \hat{\mathbf{e}}_k) \\ &= \Sigma_{ijk} \{ h'_i h'_j h'_k [(\nabla T)^{(e)}]^{ijk} \} \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j \otimes \hat{\mathbf{e}}_k \\ &= \Sigma_{ijk} [(\nabla T)^{(\hat{e})}]_{ijk} \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j \otimes \hat{\mathbf{e}}_k \quad \text{where} \quad [(\nabla T)^{(\hat{e})}]_{ijk} = h'_i h'_j h'_k [(\nabla T)^{(e)}]^{ijk} . \end{aligned}$$

Since the $\hat{\mathbf{e}}_n$ don't transform as vectors, and since the $[(\nabla T)^{(\hat{e})}]_{ijk}$ are not tensorial components, we let the indices on $[(\nabla T)^{(\hat{e})}]_{ijk}$ arbitrarily be "down". Putting them up gives a false impression of a covariant matching index tilt.

Now for convenience later, we make one more definition

$$(\nabla \mathcal{T})'_{ijk} \equiv [(\nabla T)^{(\hat{e})}]_{ijk} = h'_i h'_j h'_k [(\nabla T)^{(e)}]^{ijk} = h'_i h'_j h'_k (\nabla T)^{ijk}$$

where we follow our convention that the components of unit-base-vector tensors are written in script. (The symbol $\mathcal{T}$ is a script T , not a "tau" τ .)

Similarly, we can write the unit-base-vector expansion for a rank-2 tensor T

$$T = \Sigma_{ij} [T^{(\hat{e})}]_{ij} \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j \quad \text{where} \quad [T^{(\hat{e})}]_{ij} = h'_i h'_j [T^{(e)}]^{ij}$$

with the definition

$$\mathcal{T}'_{ij} \equiv [T^{(\hat{e})}]_{ij} = h'_i h'_j [T^{(e)}]^{ij} = h'_i h'_j T^{ij} .$$

Finally, for a vector $\mathbf{v}$,

$$\mathbf{v} = \Sigma_i [v^{(\hat{e})}]_i \hat{\mathbf{e}}_i \quad \text{where} \quad [v^{(\hat{e})}]_i = h'_i [v^{(e)}]^i$$

with the definition

$$v'_i \equiv [v^{(\hat{e})}]_i = h'_i [v^{(e)}]^i = h'_i v^i .$$

In terms of the scripted unit-base-vector components, our three expansions are:

$$\nabla T = \Sigma_{ijk} (\nabla \mathcal{T})'_{ijk} \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j \otimes \hat{\mathbf{e}}_k \quad (\nabla \mathcal{T})'_{ijk} = h'_i h'_j h'_k (\nabla T)^{ijk}$$

$$T = \Sigma_{ij} \mathcal{T}'_{ij} \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j \quad \mathcal{T}'_{ij} = h'_i h'_j T^{ij}$$

$$\mathbf{v} = \Sigma_i \mathbf{u}'_i \hat{\mathbf{e}}_i \quad \mathbf{u}'_i = h'_i v^i$$

The scripted tensor components have indices which are integers, $i = 1, 2, \dots, N$. Once we actually *select* a particular curvilinear coordinate system, one can replace the indices with curvilinear coordinate names and then, since those names *indicate* that one is talking about x' -space components, and since we just assume we are dealing with unit-base-vector components, both the script and the prime can be dropped. For example, in spherical coordinates with $1, 2, 3 = r, \theta, \phi$ one can write

$$\begin{aligned} (\nabla \mathcal{T})'_{123} &= (\nabla T)_{r\theta\phi} \\ \mathcal{T}'_{12} &= T_{r\theta} & \mathcal{T}'_{11} &= T_{rr} \\ \mathbf{u}'_3 &= v_\phi & \mathbf{u}'_1 &= v_r \end{aligned}$$

Notice that the scripted forms are always necessary when summations like Σ_{ijk} are involved, unless one is willing to write out all the terms in the sum, which is a bit clumsy. The continuum mechanics book of Lai, which we shall refer to below, avoids summations in curvilinear coordinates and thus has no need for our scripted components.

(e) Total time derivative equation written in unit-base-vector curvilinear components

Recall from section (a) our prototype time derivative equation of interest in x' -space

$$d_t T^{ij} = \partial_t T^{ij} + (\nabla T)^{ij}_k v^k$$

How does one write this equation in terms of unit-base-vector tensor components? From section (d) we had (no implied sums here)

$$\mathcal{T}'_{ij} = h'_i h'_j T^{ij} \quad \mathbf{u}'_k = h'_k v^k$$

so the above d_t equation can be processed as follows:

$$d_t T^{ij} = \partial_t T^{ij} + (\nabla T)^{ij}_k v^k$$

$$h'_i h'_j d_t T^{ij} = h'_i h'_j \partial_t T^{ij} + h'_i h'_j (\nabla T)^{ij}_k h'^{-1}_k h'_k v^k$$

$$d_t (h'_i h'_j T^{ij}) = \partial_t (h'_i h'_j T^{ij}) + h'_i h'_j h'^{-1}_k (\nabla T)^{ij}_k (h'_k v^k) \quad // h'_n = h'_n(\mathbf{x}'), \text{ no } t$$

$$d_t \mathcal{T}'_{ij} = \partial_t \mathcal{T}'_{ij} + [h'_i h'_j h'^{-1}_k (\nabla T)^{ij}_k] \mathbf{u}'_k$$

$$d_t \mathcal{T}'_{ij} = \partial_t \mathcal{T}'_{ij} + Q'_{ijk} \mathbf{u}'_k \quad Q'_{ijk} \equiv h'_i h'_j h'^{-1}_k (\nabla T)^{ij}_k \quad (*)$$

Comment: Note that $d_t h'_n(\mathbf{x}) = \partial_t h'_n(\mathbf{x}) = 0$ and not $d_t h'_n(\mathbf{x}) = \partial_t h'_n(\mathbf{x}) + (\nabla h'_n) \cdot \mathbf{v}$. The reason is that the field $h'_n(\mathbf{x})$ is not an Eulerian fluid property (see application below), it is a property of space at point $\mathbf{x}$.

Recall that no assumption has been made that the curvilinear coordinates are orthogonal. Section (b) showed how the $(\nabla T)^{ij}_{\mathbf{k}}$ can be computed for any curvilinear coordinate system, and thus one can compute the Q'_{ijk} shown above. Our question is now answered: (*) shows how one writes the total time derivative equation in arbitrary curvilinear coordinates.

If we now assume the coordinates are orthogonal, then

$$(\nabla T)^{ij}_{\mathbf{k}} = g_{\mathbf{k}\mathbf{k}'} (\nabla T)^{ij\mathbf{k}'} = h_{\mathbf{k}}^2 \delta_{\mathbf{k},\mathbf{k}'} (\nabla T)^{ij\mathbf{k}'} = h_{\mathbf{k}}^2 (\nabla T)^{ij\mathbf{k}}$$

and then

$$Q'_{ijk} \equiv h'_i h'_j h'^{-1}_{\mathbf{k}} (\nabla T)^{ij}_{\mathbf{k}} = h'_i h'_j h'^{-1}_{\mathbf{k}} h_{\mathbf{k}}^2 (\nabla T)^{ij\mathbf{k}} = h'_i h'_j h'_{\mathbf{k}} (\nabla T)^{ij\mathbf{k}} = (\nabla \mathcal{T})_{ijk}$$

and so for orthogonal coordinates (*) may be written

$$d_t \mathcal{T}'_{ij} = \partial_t \mathcal{T}'_{ij} + (\nabla \mathcal{T})'_{ijk} u'_{\mathbf{k}} \quad (\nabla \mathcal{T})'_{ijk} \equiv h'_i h'_j h'_{\mathbf{k}} (\nabla T)^{ij\mathbf{k}} \quad // \text{ orthogonal}$$

We shall assume from now on that the curvilinear coordinates are orthogonal.

As a reminder, here is what the above equation says if $i = j = 1$, using the notation scheme just described above :

$$d_t \mathcal{T}'_{11} = \partial_t \mathcal{T}'_{11} + (\nabla \mathcal{T})'_{11\mathbf{k}} u'_{\mathbf{k}}$$

$$d_t T_{\mathbf{r}\mathbf{r}} = \partial_t T_{\mathbf{r}\mathbf{r}} + (\nabla T)_{\mathbf{r}\mathbf{r}\mathbf{r}} v_{\mathbf{r}} + (\nabla T)_{\mathbf{r}\mathbf{r}\theta} v_{\theta} + (\nabla T)_{\mathbf{r}\mathbf{r}\phi} v_{\phi}$$

(f) Shorthand notations and a continuum mechanics application

Consider our original Cartesian time derivative equation,

$$d_t T^{ij} = \partial_t T^{ij} + (\nabla T)^{ij}_{\mathbf{k}} v^{\mathbf{k}} \quad .$$

One could *regard* $(\nabla T)^{ij}_{\mathbf{k}}$ as the components of a vector $(\nabla T)^{ij}$ labeled by fixed values i and j , so that

$$[(\nabla T)^{ij}]_{\mathbf{k}} = (\nabla T)^{ij}_{\mathbf{k}} \quad .$$

Then one can say

$$d_t T^{ij} = \partial_t T^{ij} + (\nabla T)^{ij} \bullet \mathbf{v} \quad .$$

The next step is to suppress the ij labels, since the equation above is true for any i and j ,

$$d_t T = \partial_t T + (\nabla T) \bullet \mathbf{v} \quad .$$

This is *only a shorthand notation*, no precision justification is required. We can do the same thing to the equation written in curvilinear coordinates

$$d_t \mathcal{T}'_{ij} = \partial_t \mathcal{T}'_{ij} + (\nabla \mathcal{T})'_{ijk} \mathbf{u}'_k$$

$$[(\nabla \mathcal{T})'_{ij}]_k = (\nabla \mathcal{T})'_{ijk}$$

$$d_t \mathcal{T}'_{ij} = \partial_t \mathcal{T}'_{ij} + (\nabla \mathcal{T})'_{ij} \bullet \mathbf{u}'_k$$

$$d_t \mathcal{T}' = \partial_t \mathcal{T}' + (\nabla \mathcal{T})' \bullet \mathbf{u}'$$

and then we have

$$d_t T = \partial_t T + (\nabla T) \bullet \mathbf{v}$$

$$d_t \mathcal{T}' = \partial_t \mathcal{T}' + (\nabla \mathcal{T})' \bullet \mathbf{u}'$$

which gives the nice impression that the equation is written in a coordinate-independent manner. In the book of Lai, the $\bullet$ is omitted and the shorthand notations are written

$$\begin{aligned} d_t T &= \partial_t T + (\nabla T) \mathbf{v} \\ d_t \mathcal{T}' &= \partial_t \mathcal{T}' + (\nabla \mathcal{T})' \mathbf{u}' \end{aligned} \quad (*)$$

where one imagines that (∇T) and $(\nabla \mathcal{T})'$ are 3-index operators which act on a 1-index object to generate a 2-index object. Again, it is just a shorthand. The real meaning of these equations is

$$\begin{aligned} d_t T^{ij} &= \partial_t T^{ij} + (\nabla T)^{ijk} v^k & \text{or} & & d_t T_{ij} &= \partial_t T_{ij} + (\nabla T)_{ijk} v^k \\ d_t \mathcal{T}'_{ij} &= \partial_t \mathcal{T}'_{ij} + (\nabla \mathcal{T})'_{ijk} \mathbf{u}'_k \end{aligned}$$

An **application** of our total time derivative equation appears in the first line of Lai p 470 ,

$$DA_1/Dt = \partial A_1/\partial_t + (\nabla A_1) \mathbf{v}$$

where D/Dt is the way a total time derivative is expressed in continuum mechanics ($D/Dt = d/dt$). It is the convective or material derivative for Eulerian-picture functions (like $A_1(\mathbf{x}, t)$), meaning that the second term registers a time change in a fluid property at the fixed point $\mathbf{x}$ due to "new fluid" with velocity $\mathbf{v}$ passing through that point. In the notation described above, this would be written

$$d_t \mathcal{A}'_1 = \partial_t \mathcal{A}'_1 + (\nabla \mathcal{A}_1)' \mathbf{u}'$$

which is just an example of equation (*) above. The detailed meaning is

$$d_t (\mathcal{A}'_1)_{ij} = \partial_t (\mathcal{A}'_1)_{ij} + (\nabla \mathcal{A}_1)'_{ijk} \mathbf{u}'_k$$

and for $i = j = 1$ this says (spherical coordinates) ,

$$d_t[(A_1)_{rr}] = \partial_t[(A_1)_{rr}] + (\nabla A_1)_{rrr} v_r + (\nabla A_1)_{rr\theta} v_\theta + (\nabla A_1)_{rr\phi} v_\phi .$$

The tensor A_1 is the first "Rivlin-Ericksen tensor" associated with the flow of non-Newtonian fluids (rheology). The A_i tensors, briefly mentioned in Appendix K (d), are derivatives of a certain deformation tensor called C_t , and computation of the A_i is done in the following iterative manner,

$$A_{i+1} = D_t A_i + A_i(\nabla \mathbf{v}) + (\nabla \mathbf{v})^T A_i \quad // \text{ Lai p 468 (8.11.3)}$$

where $(\nabla \mathbf{v})$ is the gradient-of-vector object treated in Appendix G, $\mathbf{v}$ being the fluid velocity field. In particular, $A_2 = D_t A_1 + A_1(\nabla \mathbf{v}) + (\nabla \mathbf{v})^T A_1$ which involves our object of interest $D_t A_1 = d_t A_1$. So, in order to compute A_2 in curvilinear coordinates, one needs $d_t(\mathcal{A}_1)_{ij}$ which involves the $(\nabla \mathcal{A}_1)'_{ijk}$. In order to use the above equation in practice, one has to *know* for example that $(\nabla A_1)_{rr\theta} = [\partial_\theta(A_1)_{rr} - (A_1)_{\theta r} - (A_1)_{r\theta}]/r$, a fact that is certainly not immediately obvious (see table in next section).

So, our next task is to compute the $(\nabla \mathcal{T})'_{ijk}$ which appears in our generic equation above,

$$\begin{aligned} d_t \mathcal{T}' &= \partial_t \mathcal{T}' + (\nabla \mathcal{T})' u' \\ d_t \mathcal{T}'_{ij} &= \partial_t \mathcal{T}'_{ij} + (\nabla \mathcal{T})'_{ijk} u'_k \end{aligned}$$

(g) Maple computation of the $(\nabla \mathcal{T})'_{ijk}$ components in spherical coordinates

From section (e) we have, for arbitrary curvilinear coordinates,

$$(\nabla \mathcal{T})'_{ijk} \equiv h'_i h'_j h'_k (\nabla T)^{ijk}$$

and from section (b) ,

$$(\nabla T)'_{abc} = \partial'_c T'_{ab} - \Gamma'^n_{ac} T'_{nb} - \Gamma'^n_{bc} T'_{an}$$

$$\Gamma'^d_{ab} = (1/2) g'^{dc} [\partial'_a g'_{bc} + \partial'_b g'_{ca} - \partial'_c g'_{ab}] \quad . \quad g' = \text{metric tensor for } x'\text{-space}$$

But we are now assuming only orthogonal coordinates, so

$$\begin{aligned} g'_{ab} &= \delta_{a,b} h_a^2 & g'^{ab} &= \delta_{a,b} h_a^{-2} \\ \Rightarrow (\nabla T)^{ijk} &= g'^{ii'} g'^{jj'} g'^{kk'} (\nabla T)'_{i'j'k'} = (h'_i h'_j h'_k)^{-2} (\nabla T)'_{ijk} \end{aligned}$$

and then

$$(\nabla \mathcal{T})'_{ijk} = (h'_i h'_j h'_k)^{-1} (\nabla T)'_{ijk} .$$

We want the result expressed in terms of the $\mathcal{T}'_{ij}$ and not the T'_{ij} , so recall

$$\mathcal{T}'_{ij} = (h'_i h'_j) T'^{ij} = (h'_i h'_j) g'^{ii'} g'^{jj'} T'_{i'j'} = (h'_i h'_j)^{-1} T'_{ij}$$

$$\Rightarrow T'_{ij} = (h'_i h'_j \mathcal{T}'_{ij})$$

Then

$$(\nabla T)'_{abc} = \partial'_c T'_{ab} - \Gamma'^n_{ac} T'_{nb} - \Gamma'^n_{bc} T'_{an}$$

$$(\nabla T)'_{ijk} = \partial'_k T'_{ij} - \Gamma'^n_{ik} T'_{nj} - \Gamma'^n_{jk} T'_{in}$$

$$(\nabla T)'_{ijk} = \partial'_k (h'_i h'_j \mathcal{T}'_{ij}) - \Gamma'^n_{ik} (h'_n h'_j \mathcal{T}'_{nj}) - \Gamma'^n_{jk} (h'_i h'_n \mathcal{T}'_{in})$$

$$= h'_i h'_j (\partial'_k \mathcal{T}'_{ij}) + \partial'_k (h'_i h'_j) \mathcal{T}'_{ij} - \Gamma'^n_{ik} (h'_n h'_j \mathcal{T}'_{nj}) - \Gamma'^n_{jk} (h'_i h'_n \mathcal{T}'_{in})$$

where we break the ∂'_k term in two pieces for Maple technical reasons.

So the Maple program will compute

$$(\nabla \mathcal{T})'_{ijk} = (h'_i h'_j h'_k)^{-1} (\nabla T)'_{ijk}$$

$$\text{where } (\nabla T)'_{ijk} = \underbrace{h'_i h'_j (\partial'_k \mathcal{T}'_{ij})}_{T1} + \underbrace{\partial'_k (h'_i h'_j) \mathcal{T}'_{ij}}_{T2} - \underbrace{\Gamma'^n_{ik} (h'_n h'_j \mathcal{T}'_{nj})}_{T3} - \underbrace{\Gamma'^n_{jk} (h'_i h'_n \mathcal{T}'_{in})}_{T4}$$

The Maple code is similar to that reported in Appendix I (b) and (f). The code computes the affine connection from the metric tensor,

Affine Connection $\Gamma'^d_{ab} = (1/2) g'^{dc} [\partial'_a g'_{bc} + \partial'_b g'_{ca} - \partial'_c g'_{ab}]$

```
> G := (d, a, b) -> (1/2)*sum(gcontra[d,c]*(Diff(gcov[b,c],xp[a]) + Diff(gcov[c,a],xp[b]) -
    Diff(gcov[a,b],xp[c])), c=1..N);
```

$$G := (d, a, b) \rightarrow \frac{1}{2} \left(\sum_{c=1}^N gcontra_{d,c} \left(\left(\frac{\partial}{\partial xp_a} gcov_{b,c} \right) + \left(\frac{\partial}{\partial xp_b} gcov_{c,a} \right) - \left(\frac{\partial}{\partial xp_c} gcov_{a,b} \right) \right) \right)$$

Then the terms shown above are entered ($\mathcal{T}'_{ij} = Te[i,j]$, $h'_i = hp[i]$, $\Gamma'^d_{ab} = G(d,a,b)$, etc.)

Compute grad T

```

· T1 := (i,j,k) -> hp[i]*hp[j]*Diff(Te[i,j],xp[k]);

$$T1 := (i,j,k) \rightarrow hp_i hp_j \left( \frac{\partial}{\partial xp_k} Te_{i,j} \right)$$

· T2 := (i,j,k) -> Diff(hp[i]*hp[j],xp[k])*Te[i,j];

$$T2 := (i,j,k) \rightarrow \left( \frac{\partial}{\partial xp_k} hp_i hp_j \right) Te_{i,j}$$

· T3 := (i,j,k) -> - sum(G(n,i,k)*hp[n]*hp[j]*Te[n,j],n=1..3);

$$T3 := (i,j,k) \rightarrow - \left( \sum_{n=1}^3 G(n,i,k) hp_n hp_j Te_{n,j} \right)$$

· T4 := (i,j,k) -> - sum(G(n,j,k)*hp[i]*hp[n]*Te[i,n],n=1..3);

$$T4 := (i,j,k) \rightarrow - \left( \sum_{n=1}^3 G(n,j,k) hp_i hp_n Te_{i,n} \right)$$


```

The terms are then added, multiplied by $(h'_i h'_j h'_k)^{-1}$, and then displayed,

```

> DT := (i,j,k) -> (hp[i]*hp[j]*hp[k])^(-1)*( T1(i,j,k) + value(T2(i,j,k)) +
value(T3(i,j,k))+ value(T4(i,j,k))) ;

$$DT := (i,j,k) \rightarrow \frac{T1(i,j,k) + \text{value}(T2(i,j,k)) + \text{value}(T3(i,j,k)) + \text{value}(T4(i,j,k))}{hp_i hp_j hp_k}$$

> for i from 1 to 3 do
  for j from 1 to 3 do
    for k from 1 to 3 do
      f := DT(i,j,k);
      g := subs(cos(theta)^2= 1-sin(theta)^2,f);
      h[k] := subs(cos(theta) = cot(theta)*sin(theta), expand(g));
    od;
    print(xp[i],xp[j],xp[1]," = ",h[1],"      ",xp[i],xp[j],xp[2]," = ",h[2],"
",xp[i],xp[j],xp[3]," = ",h[3]);
  od
od;

```

and here are the resulting values for $(\nabla \mathcal{T})'_{ijk}$ (for example, $(\nabla \mathcal{T})'_{112} = (\nabla T)_{rr\theta}$)

$$\begin{aligned}
 r, r, r, " &= \frac{\partial}{\partial r} T_{r, r} & r, r, \theta, " &= \frac{\partial}{\partial \theta} T_{r, r} - \frac{T_{\theta, r}}{r} - \frac{T_{r, \theta}}{r}, & r, r, \phi, " &= \frac{\partial}{\partial \phi} T_{r, r} - \frac{T_{\phi, r}}{r} - \frac{T_{r, \phi}}{r} \\
 r, \theta, r, " &= \frac{\partial}{\partial r} T_{r, \theta} & r, \theta, \theta, " &= \frac{\partial}{\partial \theta} T_{r, \theta} - \frac{T_{\theta, \theta}}{r} + \frac{T_{r, r}}{r}, & r, \theta, \phi, " &= \frac{\partial}{\partial \phi} T_{r, \theta} - \frac{T_{\phi, \theta}}{r} - \frac{\cot(\theta) T_{r, \phi}}{r} \\
 r, \phi, r, " &= \frac{\partial}{\partial r} T_{r, \phi} & r, \phi, \theta, " &= \frac{\partial}{\partial \theta} T_{r, \phi} - \frac{T_{\theta, \phi}}{r}, & r, \phi, \phi, " &= \frac{\partial}{\partial \phi} T_{r, \phi} - \frac{T_{\phi, \phi}}{r} + \frac{T_{r, r}}{r} + \frac{\cot(\theta) T_{r, \theta}}{r} \\
 \theta, r, r, " &= \frac{\partial}{\partial r} T_{\theta, r} & \theta, r, \theta, " &= \frac{\partial}{\partial \theta} T_{\theta, r} + \frac{T_{r, r}}{r} - \frac{T_{\theta, \theta}}{r}, & \theta, r, \phi, " &= \frac{\partial}{\partial \phi} T_{\theta, r} - \frac{\cot(\theta) T_{\phi, r}}{r} - \frac{T_{\theta, \phi}}{r} \\
 \theta, \theta, r, " &= \frac{\partial}{\partial r} T_{\theta, \theta} & \theta, \theta, \theta, " &= \frac{\partial}{\partial \theta} T_{\theta, \theta} + \frac{T_{r, \theta}}{r} + \frac{T_{\theta, r}}{r}, & \theta, \theta, \phi, " &= \frac{\partial}{\partial \phi} T_{\theta, \theta} - \frac{\cot(\theta) T_{\phi, \theta}}{r} - \frac{\cot(\theta) T_{\theta, \phi}}{r} \\
 \theta, \phi, r, " &= \frac{\partial}{\partial r} T_{\theta, \phi} & \theta, \phi, \theta, " &= \frac{\partial}{\partial \theta} T_{\theta, \phi} + \frac{T_{r, \phi}}{r}, & \theta, \phi, \phi, " &= \frac{\partial}{\partial \phi} T_{\theta, \phi} - \frac{\cot(\theta) T_{\phi, \phi}}{r} + \frac{T_{\theta, r}}{r} + \frac{\cot(\theta) T_{\theta, \theta}}{r} \\
 \phi, r, r, " &= \frac{\partial}{\partial r} T_{\phi, r} & \phi, r, \theta, " &= \frac{\partial}{\partial \theta} T_{\phi, r} - \frac{T_{\phi, \theta}}{r}, & \phi, r, \phi, " &= \frac{\partial}{\partial \phi} T_{\phi, r} + \frac{T_{r, r}}{r} + \frac{\cot(\theta) T_{\theta, r}}{r} - \frac{T_{\phi, \phi}}{r} \\
 \phi, \theta, r, " &= \frac{\partial}{\partial r} T_{\phi, \theta} & \phi, \theta, \theta, " &= \frac{\partial}{\partial \theta} T_{\phi, \theta} + \frac{T_{\phi, r}}{r}, & \phi, \theta, \phi, " &= \frac{\partial}{\partial \phi} T_{\phi, \theta} + \frac{T_{r, \theta}}{r} + \frac{\cot(\theta) T_{\theta, \theta}}{r} - \frac{\cot(\theta) T_{\phi, \phi}}{r} \\
 \phi, \phi, r, " &= \frac{\partial}{\partial r} T_{\phi, \phi} & \phi, \phi, \theta, " &= \frac{\partial}{\partial \theta} T_{\phi, \phi}, & \phi, \phi, \phi, " &= \frac{\partial}{\partial \phi} T_{\phi, \phi} + \frac{T_{r, \phi}}{r} + \frac{\cot(\theta) T_{\theta, \phi}}{r} + \frac{T_{\phi, r}}{r} + \frac{\cot(\theta) T_{\phi, \theta}}{r}
 \end{aligned}$$

The strange " symbols in the above table should be ignored, just a Maple artifact.

These results agree with the spherical coordinates table given in Lai p 505.

The divT results of Appendix H can be verified from the above table using

$$(\text{div} \mathcal{T})'_i = \partial'_j \mathcal{T}'_{ij} = (\nabla \mathcal{T})'_{ijj}$$

For example,

$$\begin{aligned}
 (\text{div} T)_{\mathbf{r}} &= (\nabla T)_{\mathbf{r}\mathbf{r}\mathbf{r}} + (\nabla T)_{\mathbf{r}\theta\theta} + (\nabla T)_{\mathbf{r}\phi\phi} \\
 &= \partial_{\mathbf{r}} T_{\mathbf{r}\mathbf{r}} + (1/r)[\partial_{\theta} T_{\mathbf{r}\theta} - T_{\theta\theta} + T_{\mathbf{r}\mathbf{r}}] + (1/r)[\csc\theta \partial_{\phi} T_{\mathbf{r}\phi} - T_{\phi\phi} + T_{\mathbf{r}\mathbf{r}} + \cot\theta T_{\mathbf{r}\theta}] \\
 &= \partial_{\mathbf{r}} T_{\mathbf{r}\mathbf{r}} + (1/r)\partial_{\theta} T_{\mathbf{r}\theta} - T_{\theta\theta}/r + (2/r)T_{\mathbf{r}\mathbf{r}} + (1/r\sin\theta) \partial_{\phi} T_{\mathbf{r}\phi} - T_{\phi\phi}/r + \cot\theta T_{\mathbf{r}\theta}/r
 \end{aligned}$$

and we quote from Appendix H

$$\text{div} T_1 = \left(\frac{\partial}{\partial r} T_{r,r} \right) + \frac{\frac{\partial}{\partial \theta} T_{r,\theta}}{r} + \frac{\frac{\partial}{\partial \phi} T_{r,\phi}}{r \sin(\theta)} - \frac{T_{\theta,\theta}}{r} - \frac{T_{\phi,\phi}}{r} + 2 \frac{T_{r,r}}{r} + \frac{\cos(\theta) T_{r,\theta}}{r \sin(\theta)}$$

The **Lai method** of computing $(\nabla \mathcal{T})'_{ijk}$ is different from ours. Lai uses an "affine connection" which is geared to the unit tangent base vectors, which in our notation would be written

$$\partial'_j \hat{\mathbf{e}}_i = \Gamma'^{(\text{Lai})}_{ijk} \hat{\mathbf{e}}_k, \quad // \text{ Lai p 502 (8A.12)}$$

whereas "the true" affine connection measures the change of the full tangent base vectors,

$$(\partial'_j \mathbf{e}_n) = \Gamma^{nk}_{jn} \mathbf{e}_k \quad // \text{ Appendix F (a)}$$

It is not hard to show that the connection between these two affine connections is,

$$\Gamma'^{(\text{Lai})}_{ijk} = h'^{-1}_i [h'_k \Gamma^{nk}_{ji} - \partial'_j (h'_i) \delta_{i,k}] = h'^{-1}_k [-h'_i \Gamma^{ik}_{jk} + \delta_{i,k} (\partial'_j h'_i)]$$

where the second form can be obtained from the first using this identity from Appendix F (c)

$$(\partial_c g^{ab}) = -[g^{an} \Gamma^{bn}_{cn} + g^{bn} \Gamma^{an}_{cn}]$$

The object Γ^{d}_{ab} in spherical coordinates has 9 of its 27 components non-zero, as shown at the start of Appendix I (f), but $\Gamma^{d}_{ab} = \Gamma^{d}_{ba}$ means there are only 6 distinct non-vanishing components. Correspondingly, the $\Gamma'^{(\text{Lai})}_{ijk}$ object has 6 of its 27 components non-zero (Lai p 503 (8A.14)).

Lai uses $M_{ijk} = (\nabla \mathcal{T})'_{ijk}$ and in Lai notation the expansion is $\nabla T = \Sigma_{ijk} M_{ijk} \hat{\mathbf{e}}_i \hat{\mathbf{e}}_j \hat{\mathbf{e}}_k$ (p 501) which provides an example of the polyadic notation described in our Appendix E (c) ($\hat{\mathbf{e}}_i \hat{\mathbf{e}}_j \hat{\mathbf{e}}_k = \hat{\mathbf{e}}_i \otimes \hat{\mathbf{e}}_j \otimes \hat{\mathbf{e}}_k$).

(h) Maple computation of the $(\nabla \mathcal{T})'_{ijk}$ components in cylindrical coordinates

Making four small edits to the spherical coordinates Maple program converts it to a cylindrical coordinates program. Here are the resulting values for $(\nabla \mathcal{T})'_{ijk}$ (for example, $(\nabla \mathcal{T})'_{123} = (\nabla T)_{r\theta z}$)

$$\begin{aligned}
r, r, r, &= \frac{\partial}{\partial r} T_{r, r} \quad r, r, \theta, = \frac{\partial}{\partial \theta} T_{r, r} - \frac{T_{\theta, r}}{r} - \frac{T_{r, \theta}}{r}, \quad r, r, z, = \frac{\partial}{\partial z} T_{r, r} \\
r, \theta, r, &= \frac{\partial}{\partial r} T_{r, \theta} \quad r, \theta, \theta, = \frac{\partial}{\partial \theta} T_{r, \theta} - \frac{T_{\theta, \theta}}{r} + \frac{T_{r, r}}{r}, \quad r, \theta, z, = \frac{\partial}{\partial z} T_{r, \theta} \\
r, z, r, &= \frac{\partial}{\partial r} T_{r, z} \quad r, z, \theta, = \frac{\partial}{\partial \theta} T_{r, z} - \frac{T_{\theta, z}}{r}, \quad r, z, z, = \frac{\partial}{\partial z} T_{r, z} \\
\theta, r, r, &= \frac{\partial}{\partial r} T_{\theta, r} \quad \theta, r, \theta, = \frac{\partial}{\partial \theta} T_{\theta, r} + \frac{T_{r, r}}{r} - \frac{T_{\theta, \theta}}{r}, \quad \theta, r, z, = \frac{\partial}{\partial z} T_{\theta, r} \\
\theta, \theta, r, &= \frac{\partial}{\partial r} T_{\theta, \theta} \quad \theta, \theta, \theta, = \frac{\partial}{\partial \theta} T_{\theta, \theta} + \frac{T_{r, \theta}}{r} + \frac{T_{\theta, r}}{r}, \quad \theta, \theta, z, = \frac{\partial}{\partial z} T_{\theta, \theta} \\
\theta, z, r, &= \frac{\partial}{\partial r} T_{\theta, z} \quad \theta, z, \theta, = \frac{\partial}{\partial \theta} T_{\theta, z} + \frac{T_{r, z}}{r}, \quad \theta, z, z, = \frac{\partial}{\partial z} T_{\theta, z} \\
z, r, r, &= \frac{\partial}{\partial r} T_{z, r} \quad z, r, \theta, = \frac{\partial}{\partial \theta} T_{z, r} - \frac{T_{z, \theta}}{r}, \quad z, r, z, = \frac{\partial}{\partial z} T_{z, r} \\
z, \theta, r, &= \frac{\partial}{\partial r} T_{z, \theta} \quad z, \theta, \theta, = \frac{\partial}{\partial \theta} T_{z, \theta} + \frac{T_{z, r}}{r}, \quad z, \theta, z, = \frac{\partial}{\partial z} T_{z, \theta} \\
z, z, r, &= \frac{\partial}{\partial r} T_{z, z} \quad z, z, \theta, = \frac{\partial}{\partial \theta} T_{z, z}, \quad z, z, z, = \frac{\partial}{\partial z} T_{z, z}
\end{aligned}$$

and these expressions agree with those in the table on page 504 of Lai. The same comment made about $\text{div} T$ at the end of the previous section applies here as well. The object $\Gamma^{\mathbf{d}}_{\mathbf{ab}}$ has 3 of 27 components non zero as shown at the end of Appendix I (g), only 2 of which are distinct. Correspondingly $\Gamma^{(\mathbf{Lai})}_{\mathbf{ijk}}$ has only 2 of 27 non zero as shown in Lai p 502 (8A.13).

It should be emphasized that this same simple Maple code can be used to compute the $(\nabla \mathcal{T})'_{\mathbf{ijk}}$ for *any* system of orthogonal curvilinear coordinates in any number of dimensions N . And the code implied at the end of section (b) and in the middle of section (e) with $Q'_{\mathbf{ijk}}$ computes $(\nabla T)'_{\mathbf{abc}}$ and $(\nabla \mathcal{T})'_{\mathbf{ijk}}$ for non-orthogonal as well as orthogonal coordinates.

Appendix K: Deformation Tensors in Continuum Mechanics

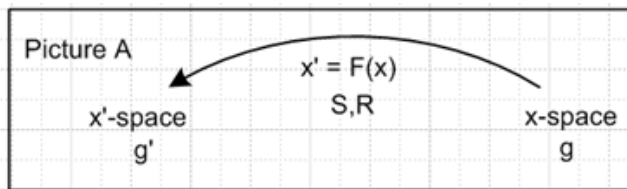
Tensor-like objects appear everywhere in continuum mechanics. As was noted in Section 2 (k), and later in Appendix E (i), continuum mechanics texts generally refer to all objects having indices as being "tensors", whether or not these objects actually transform as tensors with respect to some underlying transformation. The most commonly appearing tensors have two indices and are just 3×3 matrices associated with 3D space. In this category, there are several kinds of stress tensors, and many kinds of strain and deformation tensors which describe how a tiny volume of continuous matter (perhaps a tiny cube near some point $\mathbf{x}$) changes shape in response to some applied stress. For a fluid, a fixed stress pattern can cause a continuous ongoing change of shape which is measured then by a "rate of deformation tensor" often called $\mathbf{D}$.

An equation relating stress to strain/deformation is called a constitutive equation and describes the "response" of some physical system to "stimulus". The constitutive equation for a spring is $\mathbf{F} = -k\Delta\mathbf{x}$, for example, which is distinct from the equation of motion for a mass on a spring which is $\mathbf{F} = m\mathbf{a}$. For the spring, the stimulus is the force $\mathbf{F}$ ("stress"), and the response is the spring stretch $\Delta\mathbf{x}$ ("strain").

In this section we shall study the tensor aspects of several kinds of deformation tensors appearing in continuum mechanics. In the book of Lai *et. al.* this material is spread over several chapters, but here it will all be put in one place with Lai references provided. Section (c) below considers the form of a candidate constitutive equation for a continuous solid whose form is determined by the requirement that the equation be "covariant" as discussed in Section 7 (u). Similarly, sections (d) and (e) consider covariant constitutive equations for fluids

(a) A Preliminary Deformation Flow Picture

Recall the general nature of the Pictures appearing through this document, such as



The arrow represents an underlying generally non-linear transformation between $\mathbf{x}$ -space and $\mathbf{x}'$ -space given by $\mathbf{x}' = \mathbf{F}(\mathbf{x})$, while $\mathbf{R}(\mathbf{x})$ ($\mathbf{S} = \mathbf{R}^{-1}$) is the linearized-at-point- $\mathbf{x}$ version of the transformation which defines the notion of a vector as in $d\mathbf{x}' = \mathbf{R} d\mathbf{x}$ (see Section 2). Since $\mathbf{F}$ will have another meaning below, we shall change the transformation name so that $\mathbf{x}' = \mathcal{F}(\mathbf{x})$.

Now consider the picture below in which appear two sequential transformations $\mathcal{F}_{t_0}$ and $\mathcal{F}_t$. There are three spaces called $\mathbf{X}$ -space on the bottom, $\mathbf{x}$ -space in the middle, and $\mathbf{x}'$ -space on the top. The spaces are associated with frames of reference $\mathbf{S}_0$, $\mathbf{S}$ and $\mathbf{S}'$. Each frame has some set of basis vectors to be discussed below. The linearized $\mathbf{R}$ matrix objects associated with the two transformations are shown to the right and are given the names $\mathbf{R}_{t_0} = \mathbf{F}$ on the bottom and $\mathbf{R}_t = \mathbf{F}_t$ on the top. The transformation $\mathbf{x}' = \mathcal{F}_t(\mathbf{x}, \tau)$ is a spatial coordinate transformation only, the time coordinate τ is a parameter. In the picture below, time increases in the upward direction, so $\tau > t > t_0$.

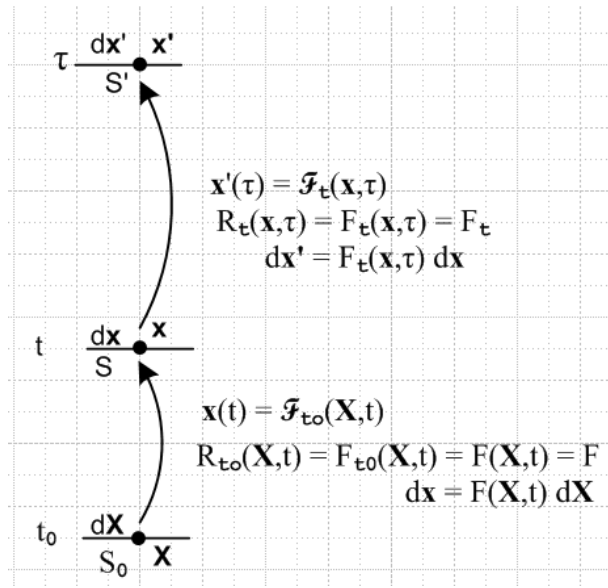


Fig 1

Consider for the moment just the bottom transformation. The transformation $\mathcal{F}_{t_0} \equiv \mathcal{F}$ describes the deformation of a particle of continuous matter which starts at position $\mathbf{X}$ and time t_0 and ends up at position $\mathbf{x}$ at time t . If we look at a large cube of continuous matter, we might find that it deforms in a very complicated manner as determined by the non-linear transformation $\mathcal{F}$ applied to all the particles within this large cube. The cube gets stirred up and is probably no longer recognizable. But if instead we consider a differentially small starting cube at $\mathbf{X}$ and t_0 , we shall find that at time t that cube is at location $\mathbf{x}$ but has been transformed into a tiny rotated parallelepiped whose axes are no longer orthogonal. It is assumed that the flow is reasonable and smooth, we are not considering some kind of "explosion" here. We use the words flow and fluid, but the deformation concept applies to elastic solids as well as fluids since these deform in some way when they are stressed (think jello or even steel).

Rather than think of the flow in terms of the edges of this tiny cube, one can instead consider two very closely spaced points in the fluid close to $\mathbf{X}$ which are separated by spacing $d\mathbf{X}$ at time t_0 , which we think of as "a little dumbbell". At time t , if one carefully tracks the "pathlines" of the ends of the dumbbell, one finds that the dumbbell tumbles and stretches and ends up as $d\mathbf{x}$ at time t and location $\mathbf{x}$,

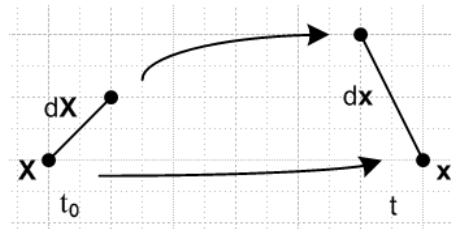


Fig 2

This differential dumbbell can be regarded as a mathematical "probe" embedded in the continuous medium. The relationship between $d\mathbf{x}$ and $d\mathbf{X}$ is given by

$$d\mathbf{x} = \mathbf{F}(\mathbf{X}, t) d\mathbf{X}$$

$$dx_i = F_{ij} dX_j$$

$$// \text{Lai p 105 (3.18.13)}$$

where the matrix F_{ij} is called "the deformation gradient". It is also known as "the deformation gradient tensor" even though it is not a "tensorial tensor" with respect to any identifiable transformation. In Section 5 (o) the above equation was identified with $d\mathbf{x}' = \mathbf{R}(\mathbf{x})d\mathbf{x}$ with $\mathbf{R}(\mathbf{x})$ here being $\mathbf{F}(\mathbf{x},t)$. The picture above is then the second figure shown in Section 2 (left-right switched). Thus, the deformation gradient $\mathbf{F}$ is the linearized version (at point $\mathbf{x}$) of some fancy non-linear (and unknown) "flow transformation" $\mathcal{F}$. Since one can write

$$dx_i = (\partial x_i / \partial X_j) dX_j ,$$

one finds that

$$F_{ij} = (\partial x_i / \partial X_j) \equiv \partial_j X_i \quad \text{or} \quad \mathbf{F} = (\nabla \mathbf{x}) \quad // \text{ Lai p 105 (3.18.4)}$$

where the gradient ∇ is with respect to $\mathbf{X}$, so it is really $\nabla = \nabla^{(\mathbf{x})}$. Thus the name "deformation gradient". [Notice that $(\nabla \mathbf{x})$ is a matrix. In Appendix G the form of $(\nabla \mathbf{v})$ for arbitrary vector $\mathbf{v}$ is found in arbitrary curvilinear coordinates. The index order reversal $F_{ij} = \partial_j X_i$ is mentioned there as well.]

The deformation gradient $\mathbf{F}(\mathbf{X},t)$ depends implicitly on the time t_0 . At $t = t_0 + \varepsilon$ (with a very small ε) no flow has yet taken place, so $dx_i = dX_i$ and then $\mathbf{F}(\mathbf{X},t_0) = \mathbf{1}$. Time t_0 is called the reference time and one could display it by writing $\mathbf{F}(\mathbf{X},t) = \mathbf{F}_{t_0}(\mathbf{X},t)$, but normally this t_0 label is suppressed.

Now consider the upper flow in Fig 1 above. It is entirely analogous to the lower flow, but names are changed. We get from the lower flow to the upper flow by making these replacements :

$$\begin{array}{llllll} t_0 \rightarrow t & d\mathbf{X} \rightarrow d\mathbf{x} & \mathcal{F}_{t_0}(\mathbf{X},t) \rightarrow \mathcal{F}_t(\mathbf{x},\tau) & S_0 \rightarrow S & \mathcal{F}_{t_0}(\mathbf{X},t_0) = 1 & \rightarrow \mathcal{F}_t(\mathbf{x},t) = 1 \\ t \rightarrow \tau & d\mathbf{x} \rightarrow d\mathbf{x}' & \mathbf{F}_{t_0}(\mathbf{X},t) \rightarrow \mathbf{F}_t(\mathbf{x},\tau) & S \rightarrow S' & \mathbf{F}_{t_0}(\mathbf{X},t_0) = 1 & \rightarrow \mathbf{F}_t(\mathbf{x},t) = 1 \end{array}$$

The equations corresponding to those shown above are

$$d\mathbf{x}' = \mathbf{F}_t(\mathbf{x},\tau) d\mathbf{x} \quad dx'_i = (F_t)_{ij} dx_j \quad // \text{ Lai p 457 (8.7.2)}$$

Since one can write

$$dx'_i = (\partial x'_i / \partial x_j) dx_j ,$$

one finds that

$$(F_t)_{ij} = (\partial x'_i / \partial x_j) \quad \text{or} \quad \mathbf{F}_t = (\nabla \mathbf{x}') \quad // \text{ Lai p 457 (8.7.3)}$$

where the gradient ∇ is with respect to $\mathbf{x}$, so it is really $\nabla = \nabla^{(\mathbf{x})}$.

In the lower flow, t_0 is the reference time, and t is the "current time". In the upper flow, the current time t is the reference time, and τ is some time $\tau > t$. The upper flow is relative to current time t as reference, and for that reason the word "relative" is pre-pended to the names of all related tensors. Thus, $\mathbf{F}_t$ is called the "relative deformation gradient", whereas $\mathbf{F}$ is just the "deformation gradient".

Why are relative tensors useful?

The main motivation for use of the relative tensors concerns differentiation with respect to time in the vicinity of the current time t . One can write

$$d_{\mathbf{t}}F_{\mathbf{t}}(\mathbf{x},t) \equiv [\partial_{\tau}F_{\mathbf{t}}(\mathbf{x},\tau)]^{\tau=t} \quad // \mathbf{x} \text{ fixed (for example, } \mathbf{x} = \mathbf{x}_1)$$

where

$$\partial_{\tau}F_{\mathbf{t}}(\mathbf{x},\tau) \approx [F_{\mathbf{t}}(\mathbf{x},\tau+d\tau) - F_{\mathbf{t}}(\mathbf{x},\tau)] / d\tau .$$

Here $d_{\mathbf{t}} = D_{\mathbf{t}} = d/dt = D/Dt$ = the total time derivative, and $\partial_{\mathbf{t}} = \partial/\partial t$ = the partial time derivative. Since $\mathbf{x}$ is fixed, $d\mathbf{x} = 0$ so $d_{\mathbf{t}} = \partial_{\mathbf{t}}$. We want to know the *rate* of deformation at some fixed current time t , and it is the τ argument of the function $F_{\mathbf{t}}(\mathbf{x},\tau)$ that lets this derivative be computed.

One could in theory carry out this same differentiation using the "non-relative" tensors by doing d/dt_0 with t_0 near t :

$$d_{\mathbf{t}}F_{\mathbf{t}}(\mathbf{X},t) \equiv [\partial_{t_0}F_{t_0}(\mathbf{X},t)]^{t_0=t} \quad // \mathbf{X} \text{ fixed}$$

where

$$\partial_{t_0}F_{t_0}(\mathbf{X},t) \approx [F_{t_0}(\mathbf{X},t) - F_{t_0-dt_0}(\mathbf{X},t)] / dt_0 ,$$

but this goes against the grain of the idea that $\mathbf{X}$ is a material coordinate at initial time t_0 which is earlier than t . And in the above, we end up with a statement about Lagrangian functions $f(\mathbf{X},t)$ rather than Eulerian functions $f(\mathbf{x},t)$, though one might argue that as $t_0 \rightarrow t$, one has $\mathbf{X} \rightarrow \mathbf{x}$. The relative tensor approach makes the differentiation process clearer, and will be used below for that purpose.

The frames of reference and the cameraman

Each of the three frames of reference S_0 , S and S' in Fig 1 has its own set of orthonormal basis vectors which we are completely free to set in any manner. As a construct it is helpful to imagine that, as the flow proceeds, it is observed by a cameraman who flies around on a camera platform which translates and rotates in some arbitrary manner. Since our main concern will be with the dumbbells like $d\mathbf{X}$, $d\mathbf{x}$ and $d\mathbf{x}'$, the translational part of the camera platform motion is irrelevant since $d\mathbf{X}$ is invariant under translations. We allow the cameraman's arbitrary orientation at times t_0 , t and τ to determine the axes of the three frames S_0 , S and S' . The cameraman is an "observer".

The *values* of the deformation gradient matrix elements F_{ij} depend on the choice of basis vectors in frames S_0 and S , so they in fact are dependent on how the cameraman flies his platform. Consider,

$$d\mathbf{x} = dx_1 \hat{\mathbf{e}}_1^{(S)} + dx_2 \hat{\mathbf{e}}_2^{(S)} + dx_3 \hat{\mathbf{e}}_3^{(S)}$$

$$d\mathbf{X} = dX_1 \hat{\mathbf{e}}_1^{(S_0)} + dX_2 \hat{\mathbf{e}}_2^{(S_0)} + dX_3 \hat{\mathbf{e}}_3^{(S_0)}$$

Once the axes are chosen, the value of the F_{ij} are determined, for example,

$$F_{12} \approx (dx_1)/(dX_2) .$$

If we were to rotate the basis vectors in frame S, for example, dx_1 would change, dX_2 would stay the same, and F_{12} would change.

Tensor expansions of the deformation gradients

These two expansions won't be used below, they are just inserted here "for interest". The expansions use the basis vectors just defined above which are $\hat{\mathbf{e}}_n^{(S_0)}$ for frame S_0 and $\hat{\mathbf{e}}_n^{(S)}$ for frame S.

- Consider this candidate expansion for the deformation gradient F,

$$F = \sum_{ij} F_{ij} \hat{\mathbf{e}}_i^{(S)} \otimes \hat{\mathbf{e}}_j^{(S_0)} = \sum_{ij} F_{ij} [\hat{\mathbf{e}}_i^{(S)}] [\hat{\mathbf{e}}_j^{(S_0)}]^T$$

$$\text{where } F_{ij} = [F^{(S, S_0)}]_{ij} = [\hat{\mathbf{e}}_i^{(S)}]^T F [\hat{\mathbf{e}}_j^{(S_0)}] = (\partial x_i / \partial X_j),$$

where we write the expansion in both direct product and matrix form as in Appendix E. This is a "mixed basis expansion" as discussed in Appendix E (i). Consider the application of this expansion to $d\mathbf{X}$:

$$\begin{aligned} & \{ \sum_{ij} F_{ij} [\hat{\mathbf{e}}_i^{(S)}] [\hat{\mathbf{e}}_j^{(S_0)}]^T \} d\mathbf{X} \\ &= \sum_{ij} F_{ij} [\hat{\mathbf{e}}_i^{(S)}] [\hat{\mathbf{e}}_j^{(S_0)}]^T \{ \sum_k dX_k \hat{\mathbf{e}}_k^{(S_0)} \} \\ &= \sum_{ij} F_{ij} \sum_k dX_k [\hat{\mathbf{e}}_i^{(S)}] [\hat{\mathbf{e}}_j^{(S_0)}]^T [\hat{\mathbf{e}}_k^{(S_0)}] \\ &= \sum_{ij} F_{ij} \sum_k dX_k [\hat{\mathbf{e}}_i^{(S)}] \delta_{j,k} \\ &= [\sum_{ij} F_{ij} dX_j] \hat{\mathbf{e}}_i^{(S)} \\ &= [dx_i] \hat{\mathbf{e}}_i^{(S)} \quad // \text{ using the fact that } F d\mathbf{X} = d\mathbf{x} \\ &= d\mathbf{x} \end{aligned}$$

Since $\{\text{expansion}\} d\mathbf{X} = d\mathbf{x}$ and since $(\nabla^{(\mathbf{x})} \mathbf{x}) d\mathbf{X} = d\mathbf{x}$ by the chain rule, it seems reasonable to conclude that $\{\text{expansion}\} = (\nabla^{(\mathbf{x})} \mathbf{x}) = F$.

- If we agree to use the $\hat{\mathbf{e}}_i^{(S)}$ for both $d\mathbf{x}$ and $d\mathbf{x}'$, so that $d\mathbf{x} = dx_i \hat{\mathbf{e}}_i^{(S)}$ and $d\mathbf{x}' = dx'_i \hat{\mathbf{e}}_i^{(S)}$, then the following is a viable expansion for the relative deformation gradient F_t :

$$F_t = \sum_{ij} (F_t)_{ij} \hat{\mathbf{e}}_i^{(S)} \otimes \hat{\mathbf{e}}_j^{(S)} = \sum_{ij} (F_t)_{ij} [\hat{\mathbf{e}}_i^{(S)}] [\hat{\mathbf{e}}_j^{(S)}]^T$$

$$\text{where } (F_t)_{ij} = [F_t^{(S, S)}]_{ij} = [\hat{\mathbf{e}}_i^{(S)}]^T F_t [\hat{\mathbf{e}}_j^{(S)}] = (\partial x'_i / \partial x_j)$$

The verification is similar to the above,

$$\begin{aligned}
 & \{ \Sigma_{ij} (F_t)_{ij} [\hat{e}_i^{(s)}] [\hat{e}_j^{(s)}]^T \} dx \\
 &= \Sigma_{ij} (F_t)_{ij} [\hat{e}_i^{(s)}] [\hat{e}_j^{(s)}]^T \{ \Sigma_k dx_k \hat{e}_k^{(s)} \} \\
 &= \Sigma_{ij} (F_t)_{ij} \Sigma_k dx_k [\hat{e}_i^{(s)}] [\hat{e}_j^{(s)}]^T [\hat{e}_k^{(s)}] \\
 &= \Sigma_{ij} (F_t)_{ij} \Sigma_k dx_k [\hat{e}_i^{(s)}] \delta_{j,k} \\
 &= [\Sigma_{ij} (F_t)_{ij} dx_j] \hat{e}_i^{(s)} \\
 &= [dx'_i] \hat{e}_i^{(s)} \quad // \text{ using the fact that } F_t dx = dx' \\
 &= dx' .
 \end{aligned}$$

Since {expansion} $dx = dx'$ and since $(\nabla^{(x)} x') dx = dx'$ by the chain rule, it seems reasonable to conclude that {expansion} $= (\nabla^{(x)} x') = F_t$.

(b) A More Complicated Deformation Flow Picture

Consider now this flow picture,

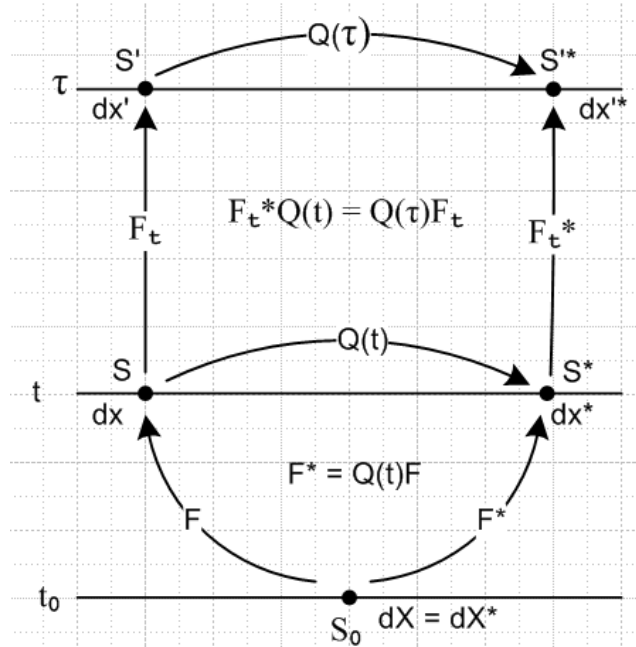


Fig 3

There is much to be said about this drawing.

The left side is the same as in Fig 1 shown above.

The picture is simplified in that it shows only the linearized R-type transformations like F and not the full transformations like $\mathcal{F}$, and these R-type matrices are now labeled right on the transformation arrows.

We only care about these F matrices because we only care about the dumbbells like $d\mathbf{x}$. For example, on the lower left we have $d\mathbf{x} = F d\mathbf{X}$.

The two sides of the picture represent observations of *the same flow* by two independent flying cameraman observers, call them C and C^* . On each side the basis vectors of the various frames are set by the motions of these cameramen. The two cameramen have agreed to start off at time t_0 with their camera platforms in exact alignment, so there is no need for a frame S_0^* .

The two frames of reference S and S^* are related by some Galilean transformation (rotation plus translation) which brings the two independent camera platforms into alignment at time t :

$$\mathbf{x}^* = Q(t) (\mathbf{x} - \mathbf{x}_0) + \mathbf{c}(t) \quad \Rightarrow \quad d\mathbf{x}^* = Q(t) d\mathbf{x}$$

Here $\mathbf{x}_0$ is a randomly selected center for the rotation $Q(t)$, and $\mathbf{c}(t)$ the corresponding translation. As above, we only care about the $Q(t)$ part of this transformation, so in terms of $d\mathbf{x}$ objects, the frames S and S^* are in effect related by the rotation $Q(t)$.

The arrows in the picture correctly describe the transformations of the $d\mathbf{x}$ type objects in moving between frames. In the lower part, for example, we have

$$d\mathbf{x}^* = Q(t) d\mathbf{x} \quad d\mathbf{x} = F d\mathbf{X} \quad d\mathbf{x}^* = F^* d\mathbf{X}$$

Comparing the left and right equations one has $Q(t) d\mathbf{x} = F^* d\mathbf{X}$ and then the center equation can be used on the left side to get $Q(t) F d\mathbf{X} = F^* d\mathbf{X}$. Since this has to be true for any $d\mathbf{X}$, we have $Q(t) F = F^*$ as shown in the drawing. This result is trivially obtained just by looking at the alternate arrow paths from frame S_0 to frame S^* . So:

$$F^* = Q(t) F \quad \text{or} \quad F^*(\mathbf{x}^*, t^*) = Q(t) F(\mathbf{x}, t) \quad t^* = t$$

At time τ we have a similar situation, but there are four arrows instead of three. Comparing the arrow paths from frame S to frame S'^* one finds

$$F_t^* Q(t) = Q(\tau) F_t \quad \Rightarrow \quad F_t^* = Q(\tau) F_t Q(t)^T \quad \text{or} \quad F_t^*(\mathbf{x}^*, \tau^*) = Q(\tau) F_t(\mathbf{x}, \tau) Q(t)^T \quad t^* = t$$

The two Q 's are rotations (reflections included) and are therefore orthogonal so $Q^{-1} = Q^T$.

Does F transform as a tensor with respect to rotation $Q(t)$?

We can think of $F(\mathbf{x}, t)$ as a property of the continuous material at location $\mathbf{x}$ and current time t . F describes the "state of deformation". If F transformed as a tensor with respect to $Q(t)$, one would need this to be true,

$$F^* = Q(t) F Q(t)^T \quad // \text{ not true!}$$

which is the matrix form for the transformation of a rank-2 tensor as shown, for example, in Section 5 (f). But we have just seen that $F^* = Q(t) F$ so the required $Q(t)^T$ on the right is missing. We conclude

therefore that in fact F , although it is called a tensor, does not transform as a tensor under $Q(t)$. One then says that F is a non-objective tensor with respect to $Q(t)$. Equation $F^* = Q(t) F$ in fact says that the columns of matrix F transform as vectors under $Q(t)$, which is very different from saying F transforms as a rank-2 tensor under $Q(t)$.

If one is trying to construct a phenomenological equation modeling a continuous material at point $\mathbf{x}$ and time t , one must make sure that equation is "covariant" (frame-indifferent) with respect to rotation $Q(t)$. The observers (cameramen) in frame S and frame S^* must see equations which have exactly the same form, which means the elements in the equations must be objective with respect to $Q(t)$. See Section 7 (u) for a general discussion of "covariance". Since F is non-objective, it is not directly useful in the construction of covariant model equations.

Do any of the usual "derived tensors" transform as tensors with respect to $Q(t)$?

By "the usual derived tensors" we mean B , C , U , V and associated R all defined as follows:

$B = FF^T$ = the left Cauchy-Green deformation tensor = the Piola deformation tensor

$C = F^T F$ = the right Cauchy-Green deformation tensor = the Finger deformation tensor

$F = RU = VR$ R = rotation U, V = symmetric positive definite

Tensors B and C are defined simply as shown, and both are therefore symmetric tensors. The last line is a statement of the **polar decomposition** theorem which says that any (real) non-singular matrix ($\det \neq 0$) can be *uniquely* written in these two ways (we apply this theorem to the deformation tensor F)

$$F = RU = VR \quad \Rightarrow \quad U = R^T VR \quad \text{and} \quad V = RUR^T \quad // \text{Lai p 114 (3.21.1,2,4)}$$

where R is a rotation matrix and V and U are *symmetric* positive definite matrices (meaning the eigenvalues are all positive) known as the left and right stretch tensors. Note that R is the same matrix in both the RU and VR forms. The idea is that the R matrix takes into account the rotational part of the deformation F , while U or V take into account the stretch component of the deformation. If the deformation is a pure rotation, $U = V = 1$, whereas if the deformation is a pure stretch then $R = 1$. A general deformation is a rotation/stretch/shear affair and one will find that none of R , U , V are unity.

One can combine the three equations above to find that

$$B = FF^T = (VR)(VR)^T = VRR^T V^T = VV^T = VV = V^2 \quad // \text{Lai p 125 (3.25.1)}$$

$$C = F^T F = (RU)^T (RU) = U^T R^T RU = U^T U = U^2 \quad // \text{Lai p 115 (3.22.1,2)}$$

So our task is to discover whether any of these derived tensors transform as a tensor relative to $Q(t)$. If they do transform as tensors (if they are objective), then they are candidates for use in constructing model equations for the continuous material.

We start with B and C :

$$B^* = F^* F^{*T} = (QF)(QF)^T = QF F^T Q^T = QBQ^T \quad \Rightarrow \quad B^* = Q(t)BQ(t)^T.$$

$$C^* = F^{*\mathbf{T}}F^* = (QF)^{\mathbf{T}}(QF) = F^{\mathbf{T}}Q^{\mathbf{T}}QF = F^{\mathbf{T}}F = C \quad \Rightarrow \quad C^* = C$$

Thus, the left Cauchy-Green deformation tensor B actually does transform as a rank-2 tensor with respect to $Q(t)$, so it is a tensorial tensor, it is "objective". In contrast, since $C^* = C$, the right Cauchy-Green deformation tensor does not transform as a rank-2 tensor. In fact each element of matrix C transforms as a tensorial scalar with respect to $Q(t)$.

What about V and U as defined above, the left and right stretch tensors?

$$F = RU = VR \quad F^* = R^*U^* = V^*R^*$$

Consider,

$$F^* = QF = Q(RU) = (QR)(U) = R^*U^*$$

Since U is positive definite symmetric, and since QR is a rotation, and since the polar decomposition is unique, it must be that

$$R^* = QR \quad \text{and} \quad U^* = U \quad .$$

Next write

$$F^* = QF = Q(VR) = (QVQ^{\mathbf{T}})(QR) = V^*R^*$$

Since the eigenvalues of symmetric V are determined by $\det(V - \lambda I) = 0$, and since this is the same as the equation $\det(QVQ^{\mathbf{T}} - \lambda I) = 0$, $QVQ^{\mathbf{T}}$ has the same eigenvalues as V and so $(QVQ^{\mathbf{T}})$ is symmetric and positive definite. Due to this fact and the fact that QR is a rotation, and the fact that the polar decomposition is unique, it must be that

$$V^* = QVQ^{\mathbf{T}} \quad \text{and} \quad QR = R^* \quad .$$

So here is a summary for our tensors of interest. Only two of the five deformation tensors actually transform as tensors. The references are to Lai page 336-337 :

$$\begin{aligned} F^* &= Q(t)F & // \text{ Lai (5.56.21)} \\ B^* &= Q(t)BQ(t)^{\mathbf{T}} & // \text{ rank-2 tensor with respect to } Q(t) \text{ so objective} \quad // \text{ Lai (5.56.31)} \\ C^* &= C & // \text{ Lai (5.56.28)} \\ U^* &= U \\ V^* &= Q(t)VQ(t)^{\mathbf{T}} & // \text{ rank-2 tensor with respect to } Q(t) \text{ so objective} \\ R^* &= Q(t)R \end{aligned}$$

Comment: Recall that $F = F(\mathbf{x}, t)$ has a hidden parameter t_0 so in fact $F = F_{t_0}(\mathbf{x}, t)$. Similarly, all derived tensors have this same hidden parameter. Thus, for example, one could write the transformation of B as

$$B_{t_0}^*(\mathbf{x}^*, t^*) = Q(t) B_{t_0}(\mathbf{x}, t) Q(t)^{\mathbf{T}} \quad t^* = t \quad \mathbf{x}^* = Q(t)(\mathbf{x} - \mathbf{x}_0) + \mathbf{c}(t) \quad d\mathbf{x}^* = Q(t) d\mathbf{x} \quad .$$

The parameter t_0 is treated as a fixed constant here and plays no role in the question of whether or not B transforms as a rank-2 tensor. The important time argument of B is the current time t , and the main idea is that $B^*(t) = Q(t) B(t) Q(t)^T$ so that $B(t)$ is objective with respect to the rotation $Q(t)$. The transformation is valid for any value of t_0 . In the limit that $t_0 \rightarrow t$, the equation says $1 = Q(t) 1 Q(t)^T$ which of course is true since rotation $Q(t)$ is orthogonal.

Do any of the usual *relative* derived tensors transform as tensors with respect to $Q(t)$?

Again we think of a relative tensor W_t as being a property of the continuous material at current time t , a measure of the state of deformation. Such a tensor is objective only if $W_t^* = Q(t) W_t Q(t)^T$. With regard to the above Comment, in this new situation it is the t of W_t which is the time variable of interest (the current time), and time τ is regarded as a fixed parameter, as was t_0 in the Comment. It just happens that the notational positions of the current time t and the parameter time τ are swapped in this case relative to the last, so now we have

$$W_t^*(x^*, \tau^*) = Q(t) W_t(x, \tau) Q(t)^T \quad \tau^* = \tau \quad x^* = Q(t) (x - x_0) + c(t) \quad dx^* = Q(t) dx \quad .$$

A tensor W_t which transforms as a rank-2 tensor (is objective) with respect to rotation $Q(t)$ must satisfy the rule above, where the arguments of both Q rotations are t . As in the Comment above, this transformation is valid for any value of parameter τ , and as $\tau \rightarrow t$, the equation says $1 = Q(t) 1 Q(t)^T$.

Our study of the transformation properties of the relative tensors proceeds in a manner similar to that used for the regular tensors above. We start with $B_t \equiv F_t F_t^T$:

$$\begin{aligned} B_t^* &= F_t^* F_t^{*T} = [Q(\tau) F_t Q^T(t)] [Q(\tau) F_t Q^T(t)]^T = Q(\tau) F_t Q^T(t) Q(t) F_t^T Q(\tau)^T \\ &= Q(\tau) F_t F_t^T Q(\tau)^T = Q(\tau) B_t Q(\tau)^T \quad // \text{ not a rank-2 tensor since } t \neq \tau \end{aligned}$$

Next comes $C_t \equiv F_t^T F_t$:

$$\begin{aligned} C_t^* &= F_t^{*T} F_t^* = [Q(\tau) F_t Q^T(t)]^T [Q(\tau) F_t Q^T(t)] = Q(t) F_t^T Q(\tau)^T Q(\tau) F_t Q^T(t) \\ &= Q(t) F_t^T F_t Q^T(t) = Q(t) C_t Q^T(t) \quad // \text{ yes a rank-2 tensor with respect to } Q(t) \end{aligned}$$

What about the left and right relative stretch tensors V_t and U_t ?

$$F_t = R_t U_t = V_t R_t \quad F_t^* = R_t^* U_t^* = V_t^* R_t^*$$

Consider,

$$F_t^* = Q(\tau) F_t Q^T(t) = Q(\tau) R_t U_t Q^T(t) = [Q(\tau) R_t Q^T(t)] [Q(t) U_t Q^T(t)] = R_t^* U_t^* \quad .$$

Since $[Q(\tau) R_t Q^T(t)]$ is a rotation and since $[Q(t) U_t Q^T(t)]$ is a symmetric positive definite matrix by the argument given in the previous section, and since the polar decomposition is unique, it must be that

$$R_t^* = Q(\tau) R_t Q^T(t) \quad \text{and} \quad U_t^* = Q(t) U_t Q^T(\tau) \quad // U_t \text{ is a rank-2 tensor}$$

Finally, write

$$F_t^* = Q(\tau) F_t Q^T(t) = Q(\tau) V_t R_t Q^T(t) = [Q(\tau) V_t Q^T(\tau)] [Q(\tau) R_t Q^T(t)] = V_t^* R_t^*$$

By the same argument used several times above, we conclude that

$$R_t^* = Q(\tau) R_t Q^T(t) \quad \text{and} \quad V_t^* = Q(\tau) V_t Q^T(\tau) \quad // V_t \text{ is not a rank-2 tensor, } \tau \neq t$$

The rule for transforming R_t is the same as found a few lines above.

Here then are the conclusions, with references to Lai page 472:

$$\begin{aligned} F_t^* &= Q(\tau) F_t Q^T(t) && // \text{Lai (8.13.6)} \\ B_t^* &= Q(\tau) B_t Q^T(\tau) && // \text{Lai (8.13.12)} \\ C_t^* &= Q(t) C_t Q^T(t) && // \text{rank-2 tensor with respect to } Q(t) \text{ so objective} \quad // \text{Lai (8.13.10)} \\ U_t^* &= Q(t) U_t Q^T(t) && // \text{rank-2 tensor with respect to } Q(t) \text{ so objective} \quad // \text{Lai (8.13.9)} \\ V_t^* &= Q(\tau) V_t Q^T(\tau) && // \text{Lai (8.13.12)} \\ R_t^* &= Q(\tau) R_t Q^T(t) && // \text{Lai (8.13.8)} \end{aligned}$$

Notice that among the "normal" tensors, B and V are objective, whereas among the "relative tensors" it is C_t and U_t that are objective. All the other tensors are "non-objective".

(c) Form of a solid constitutive equation involving the deformation tensor

For a solid continuous material in frame S one can *consider* a stress/deformation relationship of the form $T = f(B)$, where T is the Cauchy stress tensor, B is the left Cauchy-Green deformation tensor mentioned in section (b) above, and f is "some function".

In frame S^* , there will be some covariant version of the equation $T^* = f^*(B^*)$. If the medium is isotropic (rotationally invariant in its properties), then $f^* = f$ and one will have $T^* = f(B^*)$ in Frame S^* . Two observers of the same system in frames related by a rotation cannot observe different functions $f \neq f^*$ if the material is isotropic. Notice that there are two separate issues here: (1) equation must be covariant under rotations to be viable; (2) isotropic implies $f = f^*$.

If f is a polynomial, or a function which can be approximated by one (f is smooth), then $T = f(B)$ with polynomial coefficients which are rotational scalars (with respect to Q) is a viable equation form for the following reason: since B is a rank-2 tensor, so is any power of B ,

$$B^{*2} = [QBQ^T][QBQ^T] = Q B^2 Q^T \quad \text{etc.}$$

and if the polynomial coefficients are scalars, then $f(B)$ is a rank-2 tensor.

Just as a particle force $\mathbf{F}$ transforms as a rank-1 tensor under rotations, the Cauchy stress tensor T transforms as a rank-2 tensor under rotations, and then both sides of $T = f(B)$ transform in the same way -- as rank-2 tensors. Any candidate equation between T and a deformation tensor which did not have both sides transforming the same way would be invalid from the get-go (except perhaps as an approximation).

The scalar coefficients must be functions of the B_{ij} and there are three such scalars known as the principal scalar invariants of B (Lai p 40), one of which is $\det(B)$, so the scalar coefficients can be any functions of these three scalar invariants. Furthermore, one can use the fact that $B = FF^T$ is symmetric along with the Cayley-Hamilton theorem (symmetric matrix B satisfies its own secular equation, whose coefficients by the way are those scalar invariants) to show that any powers of B in polynomial $f(B)$ larger than degree 2 can be expressed as a linear combination of I , B and B^2 . One ends up then with $T = aI + bB + cB^2$ where a, b, c are functions of the three scalar invariants of tensor B .

Since both sides of $T = f(B)$ transform in the same way (rank-2 tensors), the equation $T = f(B)$ is "covariant" as discussed in Section 7 (u), meaning it has the same form in frame S^* as it has in S .

The equation $T = f(B)$ is a relation between stress and strain in the form of deformation, and as such is called a constitutive equation for the continuous material. One wants such equations to be covariant between frames of reference related by any Galilean transformation (rotation + translation), even if one or both of these frames are non-inertial. This is an extension of Hooke's Law for a spring, $\mathbf{F} = -k \Delta \mathbf{x}$, which is covariant under rotations and translations.

In contrast, equations of motion are only covariant if both frame S and S^* are inertial frames.

Notice that this entire discussion falls apart completely if one tries $T = f(F)$ or $T = f(C)$ as a candidate constitutive relation, since then the two sides of the equation don't transform the same way.

This subject is discussed in Lai pp 334-342 and p 40 for the scalar invariants. The requirement of covariance for an isotropic material and the fact that B is symmetric and transforms as a tensor puts a severe restriction on the form of the constitutive equation and we end up with $T = aI + bB + cB^2$. Since one can replace $B^3 = \alpha B^2 + \beta B + \gamma I$, if B is invertible ($\det B \neq 0$) one has $B^2 = \alpha B + \beta I + \gamma B^{-1}$ and this allows the alternate form $T = a'I + b'B + c'B^{-1}$. This last equation is used to model large deformations of an isotropic elastic material. An example is the Mooney-Rivlin theory for rubber.

(d) Some fluid constitutive equations

It was noted in section (b) that the *relative* deformation tensors are appropriate when one is interested in time derivatives of the tensors. It was also noted that the relative deformation tensor C_t is objective. One can expand $C_t(\mathbf{x}, \tau)$ in a Taylor series about current time t in this manner ($\partial_\tau \equiv \partial/\partial\tau$),

$$C_t(\mathbf{x}, \tau) = \sum_{n=0}^{\infty} [\partial_\tau^n C_t(\mathbf{x}, \tau)]^{\tau=t} (\tau-t)^n/n! = \sum_{n=0}^{\infty} A_n(\mathbf{x}, t) (\tau-t)^n/n! \quad // \text{Lai p 463 (8.10.1)}$$

$$A_n(\mathbf{x}, t) \equiv [\partial_\tau^n C_t(\mathbf{x}, \tau)]^{\tau=t},$$

where the coefficient derivatives are given the names $A_n(\mathbf{x}, t)$ called Rivlin-Ericksen tensors. Each of these coefficient tensors is in fact objective, just as is C_t , since (as usual, $t = t^*$, $\tau = \tau^*$)

$$Q(t) [\partial_\tau^n C_t(\mathbf{x}, \tau)]^{\tau=t} Q^T(t) = \{ \partial_\tau^n [Q(t) C_t(\mathbf{x}, \tau) Q^T(t)] \}^{\tau=t} = \{ \partial_{\tau^*}^n C_t^*(\mathbf{x}^*, \tau^*) \}^{\tau^*=t}$$

$$\Rightarrow Q(t) A_n(\mathbf{x}, t) Q^T(t) = A_n^*(\mathbf{x}^*, t)$$

These $A_n(\mathbf{x}, t)$ tensors appear in various models of "non-Newtonian" fluid behavior, the general study of which is called rheology, based on the Greek word for a current flow (a rheostat controls electric current),

rheo- ('ri:ə, ri:'b), also **reo-**,
used as comb. form of Gr. $\rho\acute{\epsilon}\omicron\varsigma$ stream, current,

// OED2

Here are a few covariant constitutive equations and the names assigned to them (Lai p 481). Note that for any normal fluid, there is always a $-pI$ tensor term in the expression for stress T , where p is the fluid pressure and I is the identity matrix. The diagonal elements of matrix $-pI$ are the equal normal stresses of the surroundings of a tiny cube of fluid pulling out on the cube faces, hence the $-p$ ($p > 0$) since we know the fluid actually pushes in on the cube.

$$T = -pI + \text{functional of } C_t(\tau), \tau \leq t \quad // \text{ "simple" fluid, since } \nabla^n F_t \text{ not involved } (C_t = F_t^T F_t)$$

$$T = -pI + \int_{-\infty}^t d\tau f_1(\tau) C_t(\tau) \quad // \text{ single-integral simple fluid. } f_1(\tau) = \text{a memory weight function}$$

$$T = -pI + f(A_1, A_2, \dots, A_N) \quad // \text{ Rivlin-Ericksen incompressible fluid of complexity } N$$

$$T = -pI + f(A_1, A_2) \quad // \text{ viscometric flow fluid (there are conditions on } A_1 \text{ and } A_2)$$

$$T = -pI + \mu_1 A_1 + \mu_2 A_1^2 + \mu_3 A_2 \quad // \text{ second order fluid (paint, blood, polymers)}$$

$$T = -pI + \mu A_1 \quad // \text{ incompressible Newtonian fluid (fluids like water)}$$

It turns out that $A_1 = 2D$ where $D = [(\nabla \mathbf{v}) + (\nabla \mathbf{v})^T] / 2 \equiv (\nabla \mathbf{v})^{\text{sym}}$, so A_1 is twice the rate of deformation tensor D . The other A_n can then be found from this recursion relation,

$$A_{n+1} = d_t A_n + A_n (\nabla \mathbf{v}) + (\nabla \mathbf{v})^T A_n \quad // \text{ Lai p 468 (8.11.2)}$$

Here $\mathbf{v}$ is the fluid velocity vector and $d_t = d/dt = D/Dt$. Again, $(\nabla \mathbf{v})$ is the subject of Appendix G.

(e) Corotational and other objective time derivatives of the Cauchy stress tensor

The Cauchy stress tensor T transforms as a tensor under $Q(t)$; it is objective. One can write therefore,

$$T^*(\mathbf{x}^*, t^*) = Q(t) T(\mathbf{x}, t) Q(t)^T \quad t^* = t \quad \mathbf{x}^* = Q(t) (\mathbf{x} - \mathbf{x}_0) + \mathbf{c}(t) \quad d\mathbf{x}^* = Q(t) d\mathbf{x}$$

Clarification of the above equation

One can think of the above equation $T^* = QTQ^T$ as involving operators in Hilbert Space, as outlined in Appendix E (g). In the upper part of Fig 3 above we show four different frames of reference called S , S^* , S' and S'^* each of which has its own set of basis vectors we might call $\mathbf{u}_n$, $\mathbf{u}_n^*$, $\mathbf{u}'_n$ and $\mathbf{u}'_n^*$. It happens that the picture refers to S and S^* at time t , and S' and S'^* at time τ , but any basis vectors can be "used" at any time one wants. For example, here are four expansions of the operator $T(\mathbf{x}, t)$

$$\begin{aligned} T(\mathbf{x}, t) &= \sum_{ab} T_{ab}(\mathbf{x}, t) \mathbf{u}_a \otimes \mathbf{u}_b = \sum_{ab} T^*_{ab}(\mathbf{x}^*, t) \mathbf{u}_a^* \otimes \mathbf{u}_b^* \\ &= \sum_{ab} T'_{ab}(\mathbf{x}', t) \mathbf{u}'_a \otimes \mathbf{u}'_b = \sum_{ab} T'^*_{ab}(\mathbf{x}'^*, t) \mathbf{u}'_a{}^* \otimes \mathbf{u}'_b{}^* \end{aligned}$$

in which we see four different kinds of components T_{ab} , T^*_{ab} , T'_{ab} , T'^*_{ab} . The spatial arguments of each component are written as appropriate for that frame of reference and of course all "correspond" to each other (for example, $\mathbf{x}' = \mathcal{F}_t(\mathbf{x}, \tau)$). Recall from Section 2 (i) the notion of the transformation of a contravariant vector field in developmental notation

$$\mathbf{V}'(\mathbf{x}') = \mathbf{R} \mathbf{V}(\mathbf{x}) \quad \text{contravariant} \quad R_{ik}(\mathbf{x}) \equiv (\partial x'_i / \partial x_k) \quad \mathbf{R} = \mathbf{S}^{-1}$$

where the argument is appropriate to the space of interest.

The time argument t in the above four expansions of T can be set to any arbitrary value. The stress tensor at a point $\mathbf{x}$ is in general a function of time t . One could for example set $t = \tau$ in all the expansions.

Having said this, we now decide that only the frame S basis vectors $\mathbf{u}_n$ shall be used in our expansions and components. Then for example (these $\mathbf{u}_n$ were called $\hat{\mathbf{e}}_n^{(S)}$ earlier)

$$T(\mathbf{x}, t) = \sum_{ab} T_{ab}(\mathbf{x}, t) \mathbf{u}_a \otimes \mathbf{u}_b$$

$$T^*(\mathbf{x}^*, t) = \sum_{ab} T^*_{ab}(\mathbf{x}^*, t) \mathbf{u}_a \otimes \mathbf{u}_b$$

$$Q(t) = \sum_{ab} Q_{ab}(t) \mathbf{u}_a \otimes \mathbf{u}_b .$$

Our operator statement of objectivity then becomes the following when expressed in components,

$$T^*(\mathbf{x}^*, t^*)_{ij} = Q(t)_{ia} T(\mathbf{x}, t)_{ab} Q(t)^T_{bj} \quad t^* = t$$

Thus, there should be no confusion about the following two equations which we express back in operator notation with the position arguments suppressed (but shown on the right)

$$T^*(t) = Q(t) T(t) Q(t)^T \quad // \quad T^*(\mathbf{x}^*, t) = Q(t) T(\mathbf{x}, t) Q(t)^T$$

$$T^*(\tau) = Q(\tau) T(\tau) Q(\tau)^T \quad // \quad T^*(\mathbf{x}^*, \tau) = Q(\tau) T(\mathbf{x}, \tau) Q(\tau)^T$$

Problem: The tensor dT/dt fails to transform as a rank-2 tensor, even though T does so transform.

If one tries to construct covariant constitutive equations involving dT/dt , a problem arises because dT/dt is non-objective,

$$T^*(t) = Q(t) T(t) Q(t)^T$$

$$(dT^*/dt) = Q (dT/dt) Q^T + [(dQ/dt) T Q^T + Q T (dQ/dt)^T] ,$$

so there are two extra unwanted terms. Just as $\mathbf{B} = \mathbf{F}\mathbf{F}^T$ is constructed to provide an objective derived tensor from non-objective $\mathbf{F}$, one can construct a derived version of (dT/dt) which is objective. In the next three sections, three different derived versions are described.

The corotational/Jaumann derivatives

The first step is to define an adjusted stress tensor $J_{\mathbf{t}}(\tau)$ at time τ according to (see Lai p 483, (8.19.3); Lai does not have a t subscript on J).

$$J_{\mathbf{t}}(\tau) \equiv R_{\mathbf{t}}^T(\tau) T(\tau) R_{\mathbf{t}}(\tau) \quad // \quad J_{\mathbf{t}}(\mathbf{x}, \tau) \equiv R_{\mathbf{t}}^T(\mathbf{x}, \tau) T(\mathbf{x}, \tau) R_{\mathbf{t}}(\mathbf{x}, \tau)$$

where $R_{\mathbf{t}}(\tau)$ is the rotation which appears above in section (b), where we had (showing τ arguments),

$$R_{\mathbf{t}}^*(\tau) = Q(\tau) R_{\mathbf{t}}(\tau) Q^T(t) \quad .$$

The tensor $R_{\mathbf{t}}(\tau)$ is non-objective due to appearance of $Q(\tau)$ instead of $Q(t)$ on the left (see comments on $W_{\mathbf{t}}$ above). Recall that this rotation $R_{\mathbf{t}}(\tau)$ is unique and is determined from the deformation tensor by the polar decomposition $F_{\mathbf{t}}(\tau) = R_{\mathbf{t}}(\tau) U_{\mathbf{t}}(\tau) = V_{\mathbf{t}}(\tau) R_{\mathbf{t}}(\tau)$. Thus, in some sense $J_{\mathbf{t}}(\tau)$ knows about the stress tensor $T(\tau)$, *and* it knows something about the deformation tensor through $R_{\mathbf{t}}(\tau)$. [The meaning of the term "corotational" is explained far below.]

The **claim** now is that the time derivative of this corotating stress tensor $J_{\mathbf{t}}$ is objective, meaning that tensor $d_{\mathbf{t}} J_{\mathbf{t}}$ transforms as a rank-2 tensor under the rotation $Q(t)$. Here is a proof :

We first assemble the following facts,

$$\begin{aligned} T^*(\tau) &= Q(\tau) T(\tau) Q^T(\tau) & // \text{ transformation of stress tensor } T \text{ at time } \tau \text{ (see prev section)} \\ R_{\mathbf{t}}^*(\tau) &= Q(\tau) R_{\mathbf{t}}(\tau) Q^T(t) & // \text{ how } R_{\mathbf{t}}(\tau) \text{ transforms, where } F_{\mathbf{t}}(\tau) = R_{\mathbf{t}}(\tau) U_{\mathbf{t}}(\tau) \\ J_{\mathbf{t}}(\tau) &\equiv R_{\mathbf{t}}^T(\tau) T(\tau) R_{\mathbf{t}}(\tau) & // \text{ definition of } J_{\mathbf{t}}(\tau) \text{ in frame } S \\ J_{\mathbf{t}}^*(\tau) &\equiv R_{\mathbf{t}}^{*T}(\tau) T^*(\tau) R_{\mathbf{t}}^*(\tau) & // \text{ corresponding } J_{\mathbf{t}}^* \text{ in frame } S^* \end{aligned}$$

and then we combine these ingredients to obtain a transformation rule for $J_{\mathbf{t}}$:

$$\begin{aligned} J_{\mathbf{t}}^*(\tau) &\equiv R_{\mathbf{t}}^{*T}(\tau) T^*(\tau) R_{\mathbf{t}}^*(\tau) = [Q(\tau) R_{\mathbf{t}}(\tau) Q^T(t)]^T [Q(\tau) T(\tau) Q^T(\tau)] [Q(\tau) R_{\mathbf{t}}(\tau) Q^T(t)] \\ &= [Q(t) R_{\mathbf{t}}^T(\tau) Q^T(\tau)] [Q(\tau) T(\tau) Q^T(\tau)] [Q(\tau) R_{\mathbf{t}}(\tau) Q^T(t)] \\ &= Q(t) R_{\mathbf{t}}^T(\tau) [Q^T(\tau) Q(\tau)] T(\tau) [Q(\tau)^T Q(\tau)] R_{\mathbf{t}}(\tau) Q^T(t) \\ &= Q(t) [R_{\mathbf{t}}^T(\tau) T(\tau) R_{\mathbf{t}}(\tau)] Q^T(t) \\ &= Q(t) J_{\mathbf{t}}(\tau) Q^T(t) . \end{aligned} \quad (*)$$

Since this equation $J_{\mathbf{t}}^*(\tau) = Q(t) J_{\mathbf{t}}(\tau) Q^T(t)$ fulfills the condition described earlier for $W_{\mathbf{t}}$ to be objective, we conclude that the corotating stress transforms as a rank-2 tensor, where τ is treated as a parameter.

Consider now the limit of (*) as $\tau \rightarrow t$. One finds,

$$\begin{aligned} J_t(\tau) &\equiv R_t^T(\tau) T(\tau) R_t(\tau) \\ J_t(t) &\equiv R_t^T(t) T(t) R_t(t) = 1 T(t) 1 = T(t) \end{aligned}$$

and

$$\begin{aligned} J_t^*(\tau) &\equiv R_t^{*T}(\tau) T^*(\tau) R_t^*(\tau) \\ J_t^*(t) &\equiv R_t^{*T}(t) T^*(t) R_t^*(t) = 1 T^*(t) 1 = T^*(t) . \end{aligned}$$

In this limit, the corotation $R_t^{-1}(\tau)$ has come to a halt, and (*) becomes a statement that T is objective.

More interestingly, we can apply $\partial_\tau^n = \partial^n / \partial \tau^n$ to both sides of (*) to get

$$d_\tau^n J_t^*(\tau) = Q(t) [d_\tau^n J_t(\tau)] Q^T(t) .$$

Taking the limit $\tau \rightarrow t$ then gives

$$[d_\tau^n J_t^*](t) = Q(t) [d_\tau^n J_t](t) Q^T(t)$$

which says that $d_\tau^n J_t$ are all objective tensors. And in particular, for $n = 1$,

$$(d_\tau J_t)^* = Q(t) (d_\tau J_t) Q^T(t) ,$$

and this concludes our proof that dJ_t/dt is objective, whereas dT/dt is not objective.

The above objective tensor time derivatives are sometimes written using the following strange notation

$$\overset{O}{T}_n \equiv [d^n J_t(t)/dt^n], n = 1, 2, 3, \dots \quad \overset{O}{T} \equiv \overset{O}{T}_1 \quad J_t(\tau) \equiv R_t^T(\tau) T(\tau) R_t(\tau)$$

and these are called corotational or Jaumann derivatives (Lai p 484) [Jaumann-Zaremba]. It can be shown that

$$\overset{O}{T} = d_\tau T + TW - WT \quad \text{where} \quad W = [(\nabla \mathbf{v}) - (\nabla \mathbf{v})^T]/2 = \text{"the spin tensor"} \quad // \text{Lai p 484 (8.19.10)}$$

The Oldroyd Lower convected derivatives

An alternative solution to the same problem uses a different adjusted stress tensor,

$$J_L(\tau) \equiv F_\tau^T(\tau) T(\tau) F_\tau(\tau) \quad // \text{Lai p 484 (8.19.12)}$$

We suppress the t subscript on J_L just to avoid having to write $(J_L)_t(\tau)$. In the table at the end of section (b) one sees that F_t transforms the same way R_t does, so one can repeat the above analysis to conclude that the derivatives $[d^n J_L(t)/dt^n]$ are all objective tensors (just replace $R_t \rightarrow F_t$ everywhere), so

$$\overset{\Delta}{T}_n \equiv [d^n J_L(t)/dt^n], n = 1, 2, 3, \dots \quad \overset{\Delta}{T} \equiv \overset{\Delta}{T}_1, \quad J_L(\tau) \equiv F_\tau^T(\tau) T(\tau) F_\tau(\tau) \quad // = \overset{\cup}{T}_n$$

and these are the "Oldroyd *lower* convected derivatives" (Lai p 485, called $\overset{\circ}{T}_n$). $\overset{\circ}{T}$ is sometimes called the Cotter-Rivlin stress rate. It can be shown that

$$\overset{\circ}{T} = \overset{\circ}{T} = d_t T + T(\nabla \mathbf{v}) + (\nabla \mathbf{v})^T T \quad // \text{Lai p 484 (8.19.21)}$$

The Oldroyd Upper convected derivatives

Finally, consider again the non-objective way that F_t transforms (table end of section (b))

$$F_t^*(\tau) = Q(\tau) F_t(\tau) Q^T(\tau)$$

$$\Rightarrow (F_t^{-1})^*(\tau) = Q(\tau) (F_t^{-1}(\tau)) Q^T(\tau) \quad // \text{inverted}$$

$$\Rightarrow (F_t^{-1})^{*T}(\tau) = Q(\tau) (F_t^{-1})^T(\tau) Q^T(\tau) \quad // \text{then transposed}$$

This object $F_t^{-1,T}$ therefore transforms the same way R_t and F_t transform, so we obtain a *third* set of objective time derivatives called the Oldroyd *upper* convected derivatives (Lai p 486 uses $\hat{T}_n$)

$$\overset{\nabla}{T}_n \equiv [d^n J_{\mathbf{v}}(t)/dt^n], n = 1, 2, 3, \dots \quad \overset{\nabla}{T} \equiv \overset{\nabla}{T}_1, \quad J_{\mathbf{v}}(\tau) \equiv F_t^{-1}(\tau) T(\tau) F_t^{-1,T}(\tau) \quad // = \hat{T}_n$$

The meaning of the term "convected" is explained below. It can be shown that

$$\overset{\nabla}{T} = \hat{T} = d_t T - (\nabla \mathbf{v}) T - T(\nabla \mathbf{v})^T \quad // \text{Lai p 486 (8.19.26)}$$

Covariant constitutive equations

Constitutive equations involving an objective time derivative of the stress tensor are called "rate type constitutive equations". Here are some models for incompressible fluids :

$$T = -pI + S \quad \text{where } S + \lambda \overset{\circ}{S} = 2\mu D \quad // \text{a convected Maxwell fluid}$$

$$T = -pI + S \quad \text{where } S + \lambda(\partial S / \partial t) = 2\mu D \quad // \text{linear Maxwell fluid, see below (non-covariant)}$$

$$T = -pI + S \quad \text{where } S = 2\mu D \quad // \text{Newtonian fluid}$$

$$T = -pI + S \quad \text{where } S + \lambda_1 \overset{\circ}{S} = 2\mu(D + \lambda_2 \overset{\circ}{D}) \quad // \text{a corotational Jeffrey fluid}$$

$$T = -pI + S \quad \text{where } S + \lambda_1 \overset{\nabla}{T} = 2\mu(D + \lambda_2 \overset{\nabla}{T}) \quad // \text{Oldroyd fluid A}$$

Fluids with stress time derivatives in their constitutive equations exhibit both elastic and viscous behavior at the same time. Pull on a chunk of such a fluid and the pull is initially resisted by an elastic force, but after a while the internal stress field damps out (molasses, honey) and that elastic force goes away, as if the fluid were microscopically constructed of little springs and dragging dashpots. When a constitutive equation includes a time derivative of stress, the "response" (in this case $\mathbf{D} = [(\nabla \mathbf{v}) + (\nabla \mathbf{v})^T]/2$) to the "stimulus" (T or S) includes factors of the form $e^{-t/c}$ where the c are decay time constants which are functions of the fluid parameters λ_i . In this case, the fluid has memory of its past over a time period less than these time constants, as with the honey example. For flow that is very slow relative to these time constants, the time derivative term may be neglected. In the moderately slow flow case, it can be shown that the distinction between the corotational time derivative $\overset{\circ}{S}$ and (dS/dt) can be neglected and then the convected Maxwell fluid shown above becomes the traditional linear Maxwell fluid which is modeled on those springs and dashpots with $S + \lambda (\partial S/\partial t) = 2\mu D$. (The time derivatives here are meant to act only on the second argument of $S(\mathbf{x}, \tau)$ so may be regarded as partial derivatives.) If $\lambda = 0$, the linear Maxwell fluid becomes an (incompressible) Newtonian fluid like water which has no memory.

Comment: The linear Maxwell fluid equation $S + \lambda (\partial S/\partial t) = 2\mu D$ can be solved for S using the standard Green's Function method and the solution is $S(t) = 2 \int_{-\infty}^t dt' [(\mu/\lambda)e^{-(t-t')/\lambda}] D(t')$ where the bracketed quantity (the Green's Function or kernel) is called the stress relaxation function $\phi(t-t')$. One can see in this solution the notion of memory (history) with time constant λ : the stress of the present is a function of the rate of deformation D going on in the entire past history. This solution fits into the "simple fluid" form shown earlier, where recall that $D = (1/2)A_1$ and $A_1 = [\partial_\tau C_t(\mathbf{x}, \tau)]^{\tau=t}$.

Our main point is to demonstrate the construction of constitutive equations which are covariant with respect to rotations, and which therefore can contain only tensors which in fact transform as tensors under rotations. In continuum mechanics, such tensors are said to be objective tensors.

Interpretation of the adjusted stress tensors discussed above.

In Section 2 we discuss the notion of the transformation of a contravariant vector $\mathbf{V}' = R\mathbf{V}$ in developmental notation. In x' -space, the vector components are $V'_i = R_{ij}V_j$ where V_j are the components in x -space. If R is a rotation matrix, then the unit basis vectors in the two spaces can be taken as Cartesian, call them $\mathbf{u}'_n$ in x' -space and $\mathbf{u}_n$ in x -space. We have these two expansions of $\mathbf{V}$:

$$\mathbf{V} = \sum_n V_n \mathbf{u}_n = \sum_n V'_n \mathbf{u}'_n \quad \text{where} \quad V_n = \mathbf{V} \bullet \mathbf{u}_n \quad V'_n = \mathbf{V} \bullet \mathbf{u}'_n$$

In the "active view" of things, we can think of $\mathbf{V}' = R\mathbf{V}$ as creating a new vector $\mathbf{V}'$ in x -space from the old vector $\mathbf{V}$ created by rotating the vector $\mathbf{V}$ by R. In the "passive view", we think of the V'_i as the components of the original vector $\mathbf{V}$ projected onto the backwards-rotated basis vectors $\mathbf{u}'_n = R^{-1} \mathbf{u}_n$. To verify this relation between the basis vectors, we can write

$$V'_n = \mathbf{V} \bullet \mathbf{u}'_n = \mathbf{V} \bullet R^{-1} \mathbf{u}_n = R\mathbf{V} \bullet R R^{-1} \mathbf{u}_n = R\mathbf{V} \bullet \mathbf{u}_n = \mathbf{V}' \bullet \mathbf{u}_n = V'_n.$$

So one can think *either* of $\mathbf{V}$ being rotated forward in x -space into $\mathbf{V}'$ where $\mathbf{V}'$ has x -space components V'_n , *or* one can think of the V'_n as the components of $\mathbf{V}$ one measures in frame that is backwards rotated by R^{-1} , that is, $\mathbf{u}'_n = R^{-1} \mathbf{u}_n$.

Consider then a rank-2 tensor that transforms as in Section 5 (f) according to $M' = R M R^T$. The passive interpretation is that the components M'_{ij} are those one observes in a frame of reference whose basis vectors are rotated by R^{-1} relative to the basis vectors of the unprimed frame, just as in the vector case of the last paragraph. If we now set $R = \mathcal{R}^{-1}$, then $M' = \mathcal{R}^{-1} M (\mathcal{R}^{-1})^T$ tells us that the components M'_{ij} of tensor M are those measured in a frame whose basis vectors are rotated forward by $\mathcal{R}$ relative to the unprimed frame. If it happens that $\mathcal{R} = \mathcal{R}(t)$, we would say that M'_{ij} are the components of M which are observed in a frame of reference which is rotating by $\mathcal{R}(t)$ relative to the frame of the unprimed components M_{ij} . The basis vectors of the primed frame are then $\mathbf{u}'_n = \mathcal{R} \mathbf{u}_n$.

With this long-winded introduction, we now consider the corotational stress tensor $J_t(\tau)$ from above,

$$J_t(\tau) \equiv R_t^T(\tau) T(\tau) R_t(\tau)$$

Since $R_t(\tau)$ is a rotation, $R_t^T(\tau) = R_t^{-1}(\tau)$, so we have, suppressing τ ,

$$J_t = R_t^{-1} T R_t.$$

Therefore, we can regard $(J_t)_{ij}$ as T'_{ij} , the components of stress T measured in a frame which is rotating by R_t relative to the frame in which the T_{ij} are measured. Since this primed frame rotates by R_t relative to the unprimed frame, it is called a corotating frame, and J_t is then called the corotational stress, and its time derivative is called the corotational stress rate.

We next consider the upper Oldroyd stress given above a

$$J_U \equiv F_t^{-1} T (F_t^{-1})^T$$

In analogy with the above discussion, we can regard $(J_U)_{ij}$ as T'_{ij} , the components of stress T measured in a frame which is *deforming* by F_t relative to the unprimed frame. That is to say, the basis vectors of the primed frame are given by $\mathbf{u}'_n = F_t \mathbf{u}_n$. In this case, since F_t is not a rotation, if the $\mathbf{u}_n$ start as unit vectors, then the $\mathbf{u}'_n$ are not unit vectors. One can think of each basis vector $\mathbf{u}_n$ as being aligned with its own dumbbell $d\mathbf{x}^{(n)}$ and then we have $d\mathbf{x}'^{(n)} = F_t d\mathbf{x}^{(n)}$. What this says is that the basis vectors are embedded in the fluid which deforms as it flows according to F_t . The basis vectors "convect" with the fluid, so this upper Oldroyd stress is called the upper convective stress tensor.

For the lower Oldroyd stress we have

$$J_L \equiv F_t^T T F_t$$

and this cannot be written in the form $J_L = \mathcal{R}^{-1} M (\mathcal{R}^{-1})^T$ so this does not fit into our interpretive

template. But in the next section we show that J_L is the covariant partner to the contravariant tensor J_U so they are both the same animal and we are happy to have the interpretation above for J_U .

The Oldroyd convected stresses in developmental and standard notation

In the previous section two adjusted stress tensors were introduced,

$$\begin{aligned} J_U &\equiv F_t^{-1} T F_t^{-1T} && // \text{ upper} \\ J_L &\equiv F_t^T T F_t && // \text{ lower} \end{aligned}$$

In developmental notation, a covariant tensor gets an overbar while a contravariant one does not. If we assume that frame S is Cartesian, then $T = \bar{T}$ as per Section 5 (h) for vectors. We can interpret the above two equations in this manner

$$\begin{aligned} J &\equiv F_t^{-1} T F_t^{-1T} && // \text{ upper} \\ \bar{J} &\equiv F_t^T \bar{T} F_t && // \text{ lower} \end{aligned}$$

which we compare with the first equations in Section 5 (f) where we change generic matrix name M to T ,

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

Setting $R = F_t^{-1}$ and $S = R^{-1} = F_t$ gives

$$\begin{aligned} T' &= F_t^{-1} T F_t^{-1T} && // \text{ contravariant rank-2 tensor} \\ \bar{T}' &= F_t^T \bar{T} F_t && // \text{ covariant rank-2 tensor} \end{aligned}$$

Therefore we identify

$$\begin{aligned} J_U &= J = T' = \text{contravariant stress tensor } T \text{ viewed from a frame convecting at } F_t^{-1} \\ J_L &= \bar{J} = \bar{T}' = \text{covariant stress tensor } T \text{ viewed from a frame convecting at } F_t^{-1} \end{aligned}$$

In standard notation the two equations

$$\begin{aligned} J &\equiv F_t^{-1} T F_t^{-1T} && // \text{ upper} \\ \bar{J} &\equiv F_t^T \bar{T} F_t && // \text{ lower} \end{aligned}$$

become

$$\begin{aligned} J^{ij} &= (F_t^{-1})^i_a (F_t^{-1})^j_b T^{ab} && (J_U)_{ij} = J^{ij} = \text{contravariant} \\ J_{ij} &= (F_t^{-1})^a_i (F_t^{-1})^b_j T_{ab} && (J_L)_{ij} = J_{ij} = \text{covariant} \end{aligned}$$

and this explains the meaning of the words "upper" and "lower" in respect to the Oldroyd objects. See footnote on Lai page 485.

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).

W.M. Lai, D. Rubin and E. Krempf, *Introduction to Continuum Mechanics*, 4th Ed. (Butterworth-Heinemann/Elsevier, Amsterdam, 2010). This book has had the same three authors since its first edition in 1974, and it is with apologies to the last two authors that we reference the book just as Lai, which seems more comprehensible than LRK.

P. Lucht. The most recent version of the document you are reading and related documents are downloadable at <http://user.xmission.com/~rimrock>. If not there, search on "Phil Lucht Documents" or the document title.

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)