LECTURES ON VECTORS, TENSORS AND RELATED TOPICS IN APPLIED MECHANICS

4th Edition, Fall 2007

by

Matthew R. Begley
Associate Professor
Department of Mechanical and Aerospace Engineering
& Department of Materials Science and Engineering
University of Virginia
(434) 243 8728
begley@virginia.edu

331 MEC Building
122 Engineer's Way
Charlottesville, VA 22904

ACKNOWLEDGEMENT

These ***notes*** are meant to serve as the guide for a series of student lectures; my attempt was to make a packet of notes which covers the basic and somewhat trickier aspects of vectors and tensors, from an *engineering* perspective. The intent was to provide a "nuts-and-bolts" explanation of the material, with an emphasis on physical interpretations of the math, to illustrate the utility of the material to engineering problems. **In so doing, I freely and frequently pulled direct quotations, derivations and even style-of-presentation from some of the following references. Thus, the notes are not original, though the compilation is**. Wherever possible, I have included the proper reference, including page numbers. I do not doubt there are instances where, in transcribing my notes, I failed to recollect that I followed (or even quoted) the works below - for these oversights I apologize.

REFERENCES

An Introduction to Vector and Tensor Analysis, Robert C. Wrede, Dover Publications, 1973, New York

Vector and Tensor Analysis with Applications, A.I. Borisenko and I.E Tarapov, Dover Publications, 1968, New York

Handbook of Applied Mathematics: Selected Results and Methods, ed. by Carl E. Pearson, Von Nostrand Reinhold Co., 1974, New York. (Chapter 4: Tensors, by B.Budiansky, pg. 179)

Introduction to the Mechanics of a Continuous Medium, L.E. Malvern, Prentice Hall, Inc., 1969, Englewood Cliffs, N.J. (Appendix I: Tensors, pg. 569)

Tensor Calculus, J.L. Synge and A. Schild, Dover Publications, 1949, New York

Applications of Tensor Analysis, A.J. McConnell, Dover Publications, 1957, New York

Vector and Tensor Analysis, G.E. Hay, Dover Publications, 1953, New York

Mathematics in Science and Engineering, Vol. I: Concepts from Tensor Analysis and Differential Geometry, Tracey Y. Thomas, Academic Press, 1961, New York

CONTENTS

I. INTRODUCTION

I.1 The goal of vector and tensor mathematics

Borisenko and Tarapov, pg. 11:
"The great merit of vectors in applied problems is that equations describing physical phenomena can be formulated without reference to any particular coordinate system[1]. However, in actually carrying out the calculations needed to solve a given problem, one must eventually cast the problem into a form involving scalars[2]. This is done by introducing a suitable coordinate system[3], and then replacing the given vector (or tensor) equations by an equivalent system of scalar equations involving only numbers obeying the ordinary rules of arithmetic. The key step is to expand the vectors (or tensors) with respect to a suitable basis, corresponding to the chosen system of coordinates[4]."

1. For example, Newton's second law of motion, when expressed in terms of vectors, is valid regardless of whether or not you choose a Cartesian, cylindrical, spherical, etc. coordinate system. Or, a material's state of stress is the state of stress, regardless of coordinate system - though the components (scalars) will necessarily differ from one coordinate system to the next.

2. Scalars are only a function of position; i.e. they do not depend on the "scale" of the coordinate system; invariants are things that have the same form from one coordinate system to the next, such as the magnitude of a vector.

3. Suitable coordinate systems vary from problem to problem; rectangular Cartesian coordinate systems are the most common, but prove unbelievably cumbersome in some instances (e.g., in analyzing motion constrained to lie on a sphere) and completely inadequate in others (e.g. general relativity).

4. Expanding vectors (and tensors) with respect to a suitable basis is the crux of the matter (or at least one of them); you may know the components of one vector (or tensor) with respect to one basis, and the components of another in a second basis, so....how do you get your quantities in the same basis?

- *Vector space*: A system of mathematical objects which have an additive operation producing a group structure and can be multiplied by elements from a field in a manner similar to contraction or magnification of directed line segments in Euclidean space.

- *Tensor*: 1. An object relative to a locally Euclidean space which possesses a specified system of components for every coordinate system and which changes under a transformation of coordinates. 2. A multi-linear function on the Cartesian product of several copies of a vector space and the dual of the vector space to the field of scalars on the vector space.

I.2 Coordinate systems and base vectors

I.2.1 Coordinate systems

Coordinate axes: One of a set of lines or curves used to define a coordinate system; the value of one of the coordinates uniquely determines the location of a point on the axis, while the values of the other coordinates vanish on the axis.*

Cartesian:	(x,y,z)	:	(x_1=x, x_2=y, x_3=z)
Polar:	(r,θ)	:	(x_1=r, x_2=θ)
Cylindrical:	(r,θ,z)	:	(x_1=r, x_2=θ, x_3=z)
Spherical:	(r,θ,φ)	:	(x_1=r, x_2=θ, x_3=φ)
Surface coordinates:		considered later	

Generally, the coordinates of a point can be expressed in terms of any three quantities; that is, the location of a given point is expressed as the vector defined as originating at the origin and extending to the point of interest:

$$\underline{x} = \left(f_1(\xi^1), f_2(\xi^2), f_3(\xi^3)\right) \text{ where } \xi^{\,i} \text{ are the generalized coordinates.}$$

I.2.2. Vectors

A vector is a directed line segment, that has magnitude, direction and sense. Vectors can also be thought of as ordered n-tuples (x_1,x_2,...,x_n) and as such, are not limited to three dimensional space. Most applications will involve vectors with components expressed in Cartesian coordinate systems. Ordered n-tuples must obey laws of addition and transformation to qualify as vectors. Note that vectors may be expressed in a variety (sometimes infinite) of coordinate systems; such expressions will necessarily be different, *but represent the same vector*.

From Wrede's text, pg. 51:
"A Cartesian vector is a collection {U^1,U^2,U^3} of ordered triples, each associated with a rectangular Cartesian coordinate system and such that any two satisfy the transformation law

$$U^j = \frac{\partial X^j}{\partial \overline{X}^i}\overline{U}^k$$

The triples (U^1, U^2, U^3) etc. are called components of the Cartesian vector in the respective coordinate systems." That is, the barred components ($\overline{U}^1, \overline{U}^2, \overline{U}^3$) correspond to the barred coordinate system ($\overline{X}^1, \overline{X}^2, \overline{X}^3$).

- *Free vectors*: can be displaced parallel to itself and applied at any point
- *Sliding vector*: can only be displaced along a line containing the vector, as in the force attached to a mass fastened at a fixed point

* *McGraw-Hill Dictionary of Mathematics*, pg. 52

- *Bound vectors*: a vector referred to a fixed point - the "value" of a vector field at a fixed point in space; e.g. the wind velocity at a given point

An example of an ordered triple that is not a vector is a directed line segment representation of a rotation. Suppose a rotation applied to a body is described by a line aligned with the axis of rotation, with a line length equal to the rotation angle (in radians). Intuitively, we might think of this as a vector, but it is not – it does not obey vector addition or vector transformation rules.

I.2.3. Base vectors

Basis: A set of linearly **independent** vectors in a vector space such that any vector in the space is a linear combination of vectors from the set. (D of M,pg. 18.)

Coordinate basis: a basis for tensors on a manifold (surface) induced by a set of local coordinates. (D of M, pg. 52.)

Any three independent non-coplanar vectors can serve as a *basis* for representing a vector or tensor. If all base vectors are of unit length and mutually orthogonal (perpendicular), they are said to be *orthonormal*. If the magnitude and direction of base vectors are independent of position and aligned with the coordinate axes, the system is said to be *rectangular*. In general however, base vectors do not need to orthogonal; additionally, the magnitude and direction of base vectors may dependent on position.

Budiansky, pg. 191:
"Until now, no hint has been given concerning an appropriate motivation for choosing any particular set of general base vectors. The choice, in applications of tensor analysis, is almost always tied in a special way to the general system of coordinates that is used to locate points in space. (The choice of coordinate system, in turn, is guided by such things as the shape of the region under consideration and the technique of solution to be used for solving boundary-value problems.)"

Motivation for using unusually-shaped coordinate systems is illustrated by Malvern, pg. 569:
"Another important reason is that the constitutive equations defining the material behavior may be easier to formulate relative to a coordinate net of particles than to a coordinate net of geometric points in space. When the material undergoes deformation, what started out as a rectangular Cartesian net of material particles will be deformed into a (not necessarily orthogonal) curvilinear net, which may be useful as the reference system for the next increment in deformation."

For example, in analyzing the deformation of a pressurized cylinder, the most likely choice for a coordinate system would be cylindrical; bases vectors would be orthonormal vectors oriented in the radial direction and orthogonal direction in a given z-plane. In solving for the stress fields around an elliptical hole, the solution is easiest found using generalized polar coordinates. These examples are excellent, as the choice of coordinate systems and base vectors highlight the governing equations.

Base vectors are most generally defined according to:

$$\underline{e}_i = \frac{\partial \underline{x}}{\partial \xi^i}$$

Basis vector found in this way are referred to as **natural basis vectors**. Other types of base vectors are: *orthonormal* (rectangular Cartesian), *orthogonal* (cylindrical), *general* (generalized polar). Base vectors, their transformation, and expansion of a vector about a given set of base vectors will be more extensively examined in Lecture III, §3.

I.2.4 Other terms in differential geometry

<u>Other terms</u>: (from the **McGraw-Hill Dictionary of Mathematics**)

euclidean space: A space consisting of all ordered sets $(x_1,x_2,...,x_n)$ of n numbers with the distance between $(x_1,x_2,...,x_n)$ and $(y_1,y_2,...,y_n)$ being given by

$$\left[\sum_{i=1}^{n}(x_i - y_i)^2\right]^{1/2} ;$$

the number n is called the dimension of the space.

- *manifold*: A topological space which is locally Euclidean; there are four types: topological, piecewise linear, differentiable, and complex, depending on whether the local coordinate systems are obtained from continuous, piecewise linear, differentiable or complex analytic functions of those in euclidean space; intuitively, a surface

- *elliptic geometry*: the geometry obtained from euclidean geometry by replacing the parallel line postulate with the postulate that no line may be drawn through a given point, parallel to a given line; also known as Riemannian geometry

- *Riemannian manifold*: a differentiable manifold where the tangent vectors about each point have an inner product so defined as to allow a generalized study of distance and orthogonality.

- *Riemann space*: A Riemannian manifold or subset of euclidean space where tensors can be defined to allow a general study of distance, angle and curvature.

I.3 Indicial notation

I.3.1. The summation convention[1]

A repeated index, subscript or superscript, in one term of an expression (i.e. not separated by a plus or minus) implies a summation on that index over the appropriate range. (E.g., i = 1,2 for 2-D space, i=1,2,3 for 3-D space, and i=1,2,3,...,n for n-dimensional space.) The convention is applied unless there is an explicit note to the contrary. To avoid ambiguity, the index is never repeated more than twice, unless of course, the summation convention is removed *for that expression* with a note next to the expression.

[1] The write-up here paraphrases J.R. Rice's notes on solid mechanics; see the Encyclopedia Brittanica entry for "mechanics".

For example, the linear equation $a_1x_1 + a_2x_2 + a_3x_3 = b$ is written as:

$$a_i x_i = b = \sum_{i=1}^{3} a_i x_i$$

Extending this example to a system of linear equations:

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{aligned}\right\} = a_{ij}x_j = b_i$$

Multiple summations can be succinctly written:

$$\begin{aligned} b_1 = \quad & a_{111}x_1x_1 + a_{121}x_2x_1 + a_{131}x_3x_1 \\ + \quad & a_{112}x_1x_2 + a_{122}x_2x_2 + a_{132}x_3x_2 \quad \text{is the same as } b_1 = a_{1ij}x_ix_j \\ + \quad & a_{113}x_1x_3 + a_{123}x_2x_3 + a_{133}x_3x_3 \end{aligned}$$

Hence, the equation $a_{ijk}x_jx_k = b_i$ implies three (non-linear) equations of the type shown above, with each equation containing a different subscript on b, with the same subscript on the first index of a.

Dummy index: an index which has no mathematical significance and is used to facilitate notation; it is usually summed over. In the above example, j and k are dummy indices.

Free index: the index which represents the remaining component when the expression is expanded; the free index takes on all values in the range 1,n, usually resulting in a different equation or term for each value. In the above example, i is the only free index.

I.3.2. Kronecker delta

The Kronecker delta is denoted as δ_{ij} and defined as:

$$\delta_{11} = \delta_{22} = \delta_{33} = 1$$

and

$$\delta_{12} = \delta_{21} = \delta_{31} = \delta_{31} = \delta_{32} = \delta_{23} = 0$$

Thus, the K. delta is one when the subscripts are the same, and zero otherwise. It can readily be shown that the Kronecker delta can be used to change subscripts, or to represent the scalar value three:

$$\delta_{ij}a_i = a_j \quad \text{and} \quad \delta_{ii} = \delta_{11} + \delta_{22} + \delta_{33} = 3$$

I.3.3. The alternating symbol (permutation symbol)

The alternating symbol is denoted as ε_{ijk} and defined as follows:

$$\begin{gathered} \varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = 1 \\ \varepsilon_{321} = \varepsilon_{213} = \varepsilon_{132} = -1 \\ \varepsilon_{iik} = \varepsilon_{iki} = \varepsilon_{kii} = 0 \text{ (no summation)} \end{gathered}$$

Thus, the alternating symbol equals one if the indices are an even permutation with no two equal, equals negative one if the indices are an odd permutation with no two equal, and equals zero if any two of the indices are equal. Later, when the cross product is explicitly defined, the following relation can be shown for *orthonormal bases:*

$$\underset{\sim}{e}_i \times \underset{\sim}{e}_j \times \underset{\sim}{e}_k = \varepsilon_{ijk}$$

In a related equation, the determinant of a 3x3 matrix can be written concisely as a function of it's row entries and the alternating symbol:

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \varepsilon_{ijk}a_ib_jc_k$$

Finally, there is a well known relationship between summation of two alternating symbols and Kronecker deltas, often referred to as the **$\varepsilon-\delta$ relation**:

$$\varepsilon_{ijk}\varepsilon_{irs} = \delta_{jr}\delta_{ks} - \delta_{js}\delta_{kr}$$

I.3.4. Substitutions

Suppose the following were given: $\pi = a_{ij}y_ix_j$ and $y_i = b_{ij}x_j$; the goal is to find a form of π which only depends on x_k; straight substitution leads to a meaningless form: $\pi = a_{ij}b_{ij}x_jx_j$ (**Why is this meaningless**?) The general procedure for indicial substitution is to:

1. Identify the dummy indices that appear in both expressions: j for the example above
2. Substitute for the dummy index in one expression with an index that does not appear in either form: in the above, write $y_i = b_{ir}x_r$
3. Make the substitution with the new form: $\pi = a_{ij}b_{ir}x_rx_j$

I.3.5. Some relationships to note

Quantities are said to be *symmetric* if two indices may be arbitrarily exchanged without changing the result: for instance, the Kronecker delta is symmetric, as $\delta_{ij} = \delta_{ji}$. Note also:

$$\frac{\partial x_i}{\partial x_j} = \delta_{ij} \text{ (where } x_i \text{ are coordinates)}$$

$$a_{ij}\left(x_i + y_j\right) \neq a_{ij}x_i + a_{ij}y_j,$$
$$\text{but } a_{ij}\left(x_j + y_j\right) = a_{ij}x_j + a_{ij}y_j.$$
$$a_{ij}x_i y_j \neq a_{ij}y_i x_j,$$
$$\text{but } a_{ij}x_i y_j = a_{ij}y_j x_i.$$

I.4 Notation conventions

Typically in texts, scalars are represented with normal type, vectors are given as bold-face, lower case letters and tensors are given as bold-face upper case letters. The components of the vector will be the same letter with normal type and either subscripts or superscripts to indicate the component, similarly for tensors.

Many texts will use x_i (or x^i) for Cartesian coordinates, and ξ^i (or ξ_i) for curvilinear (or general) coordinates. Likewise, the base vectors in Cartesian coordinates are typically represented by $\underline{i}_j$, whereas general base vectors are typically given as $\underline{e}_j$. Moreover, some texts distinguish between two and three-dimensional space by using Greek and alphabetic subscripts, respectively. That is, $\underline{e}_\beta$ and v_β implies subscripts in the range $\alpha,\beta = 1,2$; while $\underline{e}_j$ or v_i implies i,j,k = 1,2,3. Finally, though the distinction will only be apparent for some cases, ***contravariant*** components are denoted with a superscript, v^i, while ***covariant*** components are denoted with a subscript, v_i.

For the most part, this notation will be followed, with the following exceptions. An underscore will be used in place of **bold-face** to represent vector and tensor quantities (since boldface is difficult to represent by hand): that is, **u** becomes $\underline{u}$ (vectors) and **T** becomes $\underline{T}$ (tensors). Additionally, base vectors, *regardless of coordinate system*, will generally be represented as $\underline{e}_i$. Lastly, it will be assumed that the range of subscripts is implied by the problem at hand, regardless of whether or not Greek or traditional subscripts are used.

Clearly, adopting a consistent notation is important, particularly when learning vector and tensor mathematics. It should be emphasized that any ambiguity should be removed with explicit written notes accompanying the expression. ***Particular attention should be placed on the conventions distinguishing between scalars and tensor quantities, using lowercase to denote vectors, uppercase to denote tensors***. When learning tensor and vector calculus one will readily discover that different texts cover some subjects well, and others not so well; the serious student is encouraged to cover the material by reading multiple sources. Consistent notation will be used on a subject-by-subject basis; the reader should take care to examine the notation *in context*.

I.5 Linear systems of equations and matrices

I.5.1. Matrix operations

Equality: two matrices are equal if and only if all their entries are equal; hence, two matrices of different dimensions can never be equal. [A] and [B] are equal if $a_{ij} = b_{ij}$ for all i,j.

Addition, subtraction: addition and subtraction are performed on all individual locations in the matrix: i.e, [C] = [A]+[B], $a_{ij} + b_{ij} = c_{ij}$

Multiplication by a scalar: multiplication by a scalar merely multiplies each individual entry in the matrix by the scalar: k[A] = [kA] = k a_{ij}.

Multiplication of matrices: two matrices can be multiplied if and only if the number of columns in the second matches the number of rows in the first; in other words, j = i in the equation $[C]_{k \times m} = [A]_{k \times i}[B]_{j \times m}$. Note that process of matrix multiplication is **not** commutative, so that $[A][B] \neq [B][A]$. Furthermore, $[A][B] = 0$ does **not** necessarily imply either [A] or [B] is zero - in generally both could be non-zero.

Using indicial notation, for $[C]_{k \times m} = [A]_{k \times i}[B]_{j \times m}$:

$$c_{km} = a_{kj}b_{jm} \text{ or } c_{km} = b_{jm}a_{kj}.$$

Note that flipping indices changes the order in which the matrices are multiplied; that is....

$$[D] = [B][A] = d_{km} = b_{mj}a_{jk} = a_{jk}b_{mj}$$

It is important that the order of matrix multiplication be strictly adhered to - in general, switching the order of multiplication changes the expression: non-commutative property of matrix multiplication is often not obvious! Removal of matrices (cancellation on two sides of an equation) is performed with inverses - this preserves order.

Transpose: the interchange of rows and columns. The transpose of matrix [A] is denoted $[A]^T$. Note the following identities:

$$\left([A]^T\right)^T = [A]$$
$$([A]+[B])^T = [A]^T + [B]^T$$
$$(k[A])^T = k[A]^T$$
$$([A][B])^T = [B]^T[A]^T$$

Inverse of square matrices: the inverse of a square matrix [A] is denoted $[A]^{-1}$, and is defined as the matrix which upon multiplying by A produces the identity matrix:

$$[A][A]^{-1} = [A]^{-1}[A] = [I]$$

A product of invertible matrices (e.g. [A][B] where $[A]^{-1}$ and $[B]^{-1}$ exist) is always invertible, and the inverse of the product is the product of the inverses in reverse order:

$$([A][B])^{-1} = [B]^{-1}[A]^{-1}$$

How is this derived? As follows, noting that the order of matrix multiplication is strictly observed. By definition of the inverse...

$$([A][B])([A][B])^{-1} = [I]$$

Pre-multiply by the inverse of [A]....

$$[A]^{-1}([A][B])([A][B])^{-1} = [A]^{-1}[I] = [A]^{-1}$$

Then again, by the inverse of [B]....

$$[B]^{-1}([B])([A][B])^{-1} = [B]^{-1}[A]^{-1}$$

I.5.2. Types of matrices

Diagonal, triangular matrices: a diagonal matrix has zero entries excepted for the main diagonal; one possible indicial notation for this matrix would be....

$$[A]_{diagonal} = a_{ij}\delta_{ij} \text{ (no sum)}$$

triangular matrices have zeros for all row numbers less than the column numbers (upper triangular) or zeros for all row numbers greater than the column numbers (lower triangular).

The identity matrix: A square matrix whose entries are zero except for the main diagonal, where the entries are 1, denoted [I]. The identity matrix times any square matrix results in the original square matrix; that is, [A][I] = [I][A] = [A]. The index notation for [I] is Kroneker's delta: [I] = δ_{ij}.

Symmetric, skew (or anti-symmetric) matrices: a matrix is said to be symmetric if the entry in the ith row and jth column equals the entry in the jth row and ith column, i.e. $a_{ij} = a_{ji}$. A skew matrix has the negative of the entry in the ith row and jth column in the jth row and ith column, i.e. $a_{ij} = - a_{ji}$. (Symmetric and skew tensors, which can be represented by a matrices, play a very important role in mechanics.)

Orthogonal: The square matrix Q is said to be orthogonal if and only if the follow holds true: $[Q]^T[Q] = [I] = [Q][Q]^T$, so therefore $[Q]^T = [Q]^{-1}$.

Linear mapping: a square matrix [A] can be thought of as a linear mapping or transformation of one vector into another. In other words for [A][x] = [y], the square matrix [A] *maps* or *transforms* the vector [x] into the vector [y]. Since the transformation is linear, [A] will map [x] into only one vector.

I.5.3. Determinants

The determinant is the sum of all signed elementary products from a matrix [A]. For a 3x3 matrix:

$$\det\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{23} & a_{33} \end{bmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{23} & a_{33} \end{vmatrix} =$$

$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} + a_{12}a_{21}a_{33} + a_{11}a_{213}a_{32}$$

In indicial notation...

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \varepsilon_{ijk} a_i b_j c_k$$

Thus, the determinant expanded by rows is:

$$\det[A] = \varepsilon_{ijk} a_{1i} a_{2j} a_{3k} .$$

Or, conversely, expanded by columns is:

$$\det[A] = \varepsilon_{ijk} a_{i1} a_{j2} a_{k3} .$$

Interchanging rows or columns changes the sign of the determinant, but not its absolute value. If any two rows or columns are equal, the determinant is zero - left to an exercise as to why the above equations imply this.

Singular and non-singular matrices: a singular matrix has determinant zero; any system of equations with a singular coefficient matrix will not have a solution.

Positive-definite and semi-definite matrices: positive definite matrices have the property that:

$$[x]^T[A][x] > 0$$

for any nx1 matrix [x]. Semi-positive definite matrices have the property that:

$$[x]^T[A][x] \geq 0$$

Most matrices encountered in engineering are at least semi-positive definite; for example, elemental stiffness matrices are semi-positive definite, and global stiffness matrices (formed once boundary conditions have been applied) are positive definite.

I.5.4. Properties of the determinant function

a. If [A] contains a row of zeros, then det[A] = 0.
b. If [A] is triangular, then det[A] is the product of all entries on the main diagonal.
c. If [A'] results when a single row of [A] is multiplied by k, then det[A'] = k det [A]
d. Interchanging rows changes the sign the sign of the determinant: det[A'] = - det[A], when [A'] results from interchanging rows of [A].
e. det[A] = det($[A]^t$).

f. det([A][B]) = det[A]det[B]: As Howard Anton writes: "***The elegant simplicity of this result contrasted with the complex nature of both matrix multiplication and determinant definition is both refreshing and surprising. We shall omit the proof***."[2]
g. $\det([A]^{-1}) = 1/\det[A]$ - A matrix is invertible if and only if its determinant is non zero.
h. If any two rows or columns are linearly dependent, the matrix is singular (and hence, det[A] = 0) and not invertible.

I.5.5. Condition number

The condition number is defined as $\kappa(A) \equiv \|A\| \cdot \|A^{-1}\|$, where $\|A\| \equiv$ matrix norm. While there are many types of norms, the **row sum norm** is

$$\|A\|_\infty \equiv \max_{1 \le i \le n} \sum_{j=1}^{n} a_{ij},$$

while the **column sum norm** is

$$\|A\|_1 \equiv \max_{1 \le j \le n} \sum_{i=1}^{n} a_{ij}.$$

Thus, the norm is largest of n numbers, when the magnitudes of the entries of n rows or n columns are added. The matrix is said to be "ill-conditioned" when the condition number is large. Basically, this means that the matrix is highly sensitive to perturbations; the matrix A may map [x] into [y], but maps [x+dx] (where dx is a small perturbation) into something very different. The best possible condition number is one, as $\kappa(A) \ge 1$. An ill-conditioned matrix is *nearly* singular. **Note well, however, that there is no useful relationship between the determinant and the condition number.**

I.6 Matrix algebra

I.6.1. Cramer's rule

Given the system of three equations and three unknowns (x_i):

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= y_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= y_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= y_3 \end{aligned}$$

The solution can be found according to Cramer's rule. Let $D = \det[A]$ (where $[A] = a_{ij}$) then x_i is given by:

[2] *Elementary Linear Algebra*, H. Anton, John Wiley and Sons, 1984, New York, pg. 72.

$$x_1 = \frac{1}{D}\begin{vmatrix} y_1 & a_{12} & a_{13} \\ y_2 & a_{22} & a_{23} \\ y_3 & a_{23} & a_{33} \end{vmatrix}$$

which is expanded as:

$$x_1 = \frac{y_1}{D}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - \frac{y_2}{D}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + \frac{y_3}{D}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}$$

Generalizing to any of the x_i's, the result in indicial notation is:

$$x_m = \frac{S^{nm} y_n}{D}$$

where S^{nm} is the *cofactor* (signed subdeterminant) to the element A_{mn}. The nm cofactor is the determinant of the (n-1) x (n-1) matrix that is left after striking the nth row and mth column, mulitiplied by the factor $(-1)^{m+n}$.

In matrix notation:

$$[x] = \left[\frac{S^{mn}}{D}\right]^T [y] \rightarrow [A]^{-1} = \left[\frac{S^{mn}}{D}\right]^T$$

I.6.2. Eigenvalues and eigenvectors

The matrix [A] can be thought of as "mapping" point P=[x] into point Q=[y]; since it is a homogeneous mapping, it associates *any* point r[x] to a point r[y]. In other words, [A] maps the line OP to the line OQ. This is an important concept when tensors are being used - the strain tensor is exactly this type of mapping function.

Are there any lines not affected by the transformation?

Are there any lines that transform into themselves? Note that we ask about lines, not necessarily directed line segments, or vectors; i.e. the length of the "line" may change in the transformation. In other words, are there any lines (x_1, x_2, x_3) transform into $(\lambda x_1, \lambda x_2, \lambda x_3)$?

Yes, and such lines are called *eigenvectors* (proper vectors, characteristic vectors or latent vectors).

$$[M][x] = \lambda[x] = \lambda[I][x]$$

$$\text{Eigenvalue problem: } [[M] - \lambda[I]][x] = 0$$

A non-trivial solution exists only if the determinant of this matrix is zero. Thus the eigenvalue problem solution is found by setting the determinant equal to zero:

$$\left|[M]-\lambda[I]\right|=0$$

When the determinant is expanded, a characteristic equation results, with n characteristic roots (for a general nxn matrix), referred to as the *eigenvalues*.

Substituting one of the eigenvalues into the problem (above) determines the corresponding eigenvector (*to within a scalar multiple*); that is, it retrieves the line (though not the magnitude). Normalizing the vectors to have unit length determines the lines to within a factor of +/- 1.

- Real symmetric matrices have only real eigenvalues
- **Distinct eigenvectors are orthogonal**
- If all eigenvalues are equal, *every* direction is an eigenvector.
- If the eigenvectors are placed as columns in a matrix [L], then L is an orthogonal matrix; $[L]^T[M][L]$ is a diagonal matrix with $m_{ii} = l_i$ (no sum).

I.6.3. Examples

Poorly conditioned matrices:[3]
Consider the following matrix, which will be shown to be poorly conditioned.

$$[A]=\begin{bmatrix}1000 & 999\\ 999 & 998\end{bmatrix};\ [A]^{-1}=\begin{bmatrix}-998 & 999\\ 999 & -1000\end{bmatrix}$$

Thus, the condition numbers are:

Row sum norm $\kappa_\infty(A)=\|A\|\cdot\|A^{-1}\|=(1999)^2\approx 4x10^6$

Col. sum norm $\kappa_1(A)=\|A\|\cdot\|A^{-1}\|=(1999)^2\approx 4x10^6$

Thus, this matrix is poorly conditioned. What does that mean? Consider the solution to the following set of equations:

$$\begin{bmatrix}1000 & 999\\ 999 & 998\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}=\begin{bmatrix}1999\\ 1997\end{bmatrix}$$

The unique solution is readily found, i.e. $[x]^T=[1 \quad 1]$. However, consider the following nearly identical problem:

$$\begin{bmatrix}1000 & 999\\ 999 & 998\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}=\begin{bmatrix}1998.99\\ 1997.01\end{bmatrix}$$

[3] Taken nearly verbatim from Fundamentals of Matrix Computations, by D.S. Watkins, John Wiley and Sons, New York, 1991. (Paraphrased slightly).

One would expect that such a small perturbation of the right-hand side would have minor consequences: not so! The solution to this "perturbed" problem is $[x]^T=[20.97 \quad -18.99]$.

It should be pointed out that this example was constructed in a very special way; i.e. the "base" values of the rhs were chosen in the direction of maximum magnification of the vector [x] by the matrix [A]. The perturbation was chosen in the direction of maximum magnification of the vector [x] by the matrix $[A]^{-1}$. If the righthand side and perturbation were not chose so carefully, the results would not have been as impressive.

c. Eigenvalues and eigenvectors:

Consider n springs and n lumped masses arranged in series (one after the other). Defining T as the kinetic energy and V as the potential, equations of motion for the system can be derived using Lagrange's equation:[4]

$$\frac{d}{dt}\left(\frac{\partial T}{\partial x_i}\right)-\frac{\partial T}{\partial x_i}+\frac{\partial V}{\partial x_i}=F_i$$

The kinetic energy of the system is written as

$$T=\frac{1}{2}m_1(\dot{x}_1)^2+\frac{1}{2}m_2(\dot{x}_2)^2+\ldots+\frac{1}{2}m_n(\dot{x}_n)^2$$

In matrix notation this would be:

$$T=\begin{bmatrix}\frac{1}{2}m_1\dot{x}_1 & \frac{1}{2}m_2\dot{x}_2 & \ldots\frac{1}{2}m_n\dot{x}_n\end{bmatrix}_{1xn}\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_n\end{bmatrix}_{nx1}$$

$$T=\frac{1}{2}\begin{bmatrix}\dot{x}_1 & \dot{x}_2 & \ldots\dot{x}_n\end{bmatrix}_{1xn}\begin{bmatrix}m_1 & & \\ & m_2 & \\ & & m_n\end{bmatrix}_{nxn}\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_n\end{bmatrix}_{nx1}$$

$$=\frac{1}{2}\left([\dot{x}]^T\right)_{1xn}[m]_{nxn}[\dot{x}]_{nx1}$$

Derivative is easiest taken in indicial notation:

$$T=\frac{1}{2}\dot{x}_j m_{jk}\dot{x}_k$$

[4] See *Mechanical Vibrations*, Rao, pg. 280.

$$\frac{\partial T}{\partial \dot{x}_i} = \frac{1}{2}\frac{\partial \dot{x}_j}{\partial \dot{x}_i} m_{jk}\dot{x}_k + \frac{1}{2}\dot{x}_j m_{jk}\frac{\partial \dot{x}_k}{\partial \dot{x}_i}$$
$$= \frac{1}{2}\delta_{ji} m_{jk}\dot{x}_k + \frac{1}{2}\dot{x}_j m_{jk}\delta_{ki}$$
$$= \frac{1}{2} m_{ik}\dot{x}_k + \frac{1}{2}\dot{x}_j m_{ji}$$

Noting that the mass matrix is symmetric, $m_{ij} = m_{ji}$:

$$\frac{\partial T}{\partial \dot{x}_i} = m_{ik}\dot{x}_k = [m]_{nxn}[\dot{x}]_{nx1}$$

Taking the time derivative yields:

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{x}_i}\right) = m_{ik}\ddot{x}_k = [m]_{nxn}[\ddot{x}]_{nx1}$$

Similarly, the potential energy V can be written:

$$V = \frac{1}{2}k_1(x_1)^2 + \frac{1}{2}k_2(x_2 - x_1)^2 + \ldots + \frac{1}{2}k_n(x_n - x_{n-1})^2$$

or, using indicial notation

$$V = \frac{1}{2}x_j k_{jk} x_j$$

where the matrix k_{ij} = [k] is found by expansion of the original form. Thus, we have essentially the same matrix equation as the kinetic energy, i.e.:

$$V = \frac{1}{2}\left([x]^T\right)_{1xn}[k]_{nxn}[x]_{nx1}$$

Taking the derivative follows the same procedure. Putting together the pieces of Lagrange's equations, we have the following results:

$$[m]_{nxn}[\ddot{x}]_{nx1} + [k]_{nxn}[x]_{nx1} = [0]_{nx1}$$

$$m_{ij}\ddot{x}_j + k_{ij}x_j = 0$$

If all the springs have the same stiffness, k, and all the masses are the same, m, the mass and compliance matrices (inverse of the stiffness) are:

$$[m] = m\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} = m[I]$$

$$[a] = [k]^{-1} = \frac{1}{k}\begin{bmatrix}1 & 1 & 1\\ 1 & 2 & 2\\ 1 & 2 & 3\end{bmatrix}$$

The natural mode shapes are found for an unforced system, i.e. F_i = 0. Assume a solution of the form $x_i(t) = X_i\, T(t)$;

$$[m][x]\ddot{T}(t) + [k][x]T(t) = [0]$$

This implies

$$[k][x] = \lambda[m][x]$$

This is an eigenvalue problem; the eigenvalues l_i are related to the natural frequencies of the system, and eigenvectors will be the natural mode shapes.

$$[k]^{-1}[k][x] = \lambda[k]^{-1}[m][x] = \lambda[D][x]$$

$$\left|[D] - \lambda[I]\right| = 0$$

$$\left|\begin{bmatrix}\lambda & 0 & 0\\ 0 & \lambda & 0\\ 0 & 0 & \lambda\end{bmatrix} - \frac{m}{k}\begin{bmatrix}1 & 1 & 1\\ 1 & 2 & 2\\ 1 & 2 & 3\end{bmatrix}\right| = 0$$

$$\left(\frac{m}{k\lambda}\right)^3 - 5\left(\frac{m}{k\lambda}\right)^2 + 6\left(\frac{m}{k\lambda}\right) - 1 = 0$$

The solution is:

$$a_1 = \frac{m}{k\lambda_1} = 0.1986$$
$$a_2 = 1.5553$$
$$a_3 = 3.2490$$

The eigenvectors determined from: $\left[\lambda_i[I] - [D]\right][X]^i = 0$

$$\left[\begin{bmatrix}\lambda_i & 0 & 0\\ 0 & \lambda_i & 0\\ 0 & 0 & \lambda_i\end{bmatrix} - \frac{m}{k}\begin{bmatrix}1 & 1 & 1\\ 1 & 2 & 2\\ 1 & 2 & 3\end{bmatrix}\right]\begin{bmatrix}X_1\\ X_2\\ X_3\end{bmatrix}^i = 0$$

The eigenvectors are as follows, expressed in the form of the modal matrix

$$[X] = \left[X^1 \quad X^2 \quad X^3\right] = \begin{bmatrix} 1.0000 & 1.0000 & 1.0000 \\ 1.8019 & 0.4450 & -1.2468 \\ 2.2470 & -0.8020 & 0.5544 \end{bmatrix}$$

II: VECTORS

II.1 Introductory remarks

II.1.1. Definitions

Scalars: quantities completely defined by magnitude and sign. Temperature, concentration, work, etc.

Magnitude, direction, sense: As will be discussed in detail later (pg. 19.) the magnitude of a vector is generally taken to be its length, given in most general terms as $|\underline{v}| = \sqrt{v^i v_i}$ (Euclidean space); the direction is the orientation of the vector with regards to a fixed coordinate system, while the sense gives the relative direction with regards to the fixed coordinates.

Equal vectors: must have same magnitude, sense and direction

Co-linear vectors: have the same direction, may be of different magnitude and sense.

Ordered n-tuple: a vector may be represented in terms of an ordered set of n numbers; useful when considering spaces with n>3, which cannot be visualized; vector operations are the same for this notation: From Wrede's text: a general n dimensional vector is a collection of ordered sets of n numbers $(v_1, v_2, v_3, \ldots v_n)$, each associated with an n-dimensional coordinate system, such that any two satisfy the transformation law

$$v^i = \frac{\partial x^i}{\partial \bar{x}^j} \bar{v}^j$$

Euclidean vector space: A space of n dimensions over the real numbers for which there is a defined scalar product, a rule of composition having the properties below, whichs associates a real number $\bar{u} \cdot \bar{v}$ with every pair of elements $\bar{u}$ and $\bar{v}$. A proper Euclidean vector space is such that $\bar{u} \cdot \bar{u} > 0$ for all non-zero $\bar{u}$.*

II.1.2. Addition, subtraction, projection and multiplication by a scalar

Vectors obey the following laws of arithmetic:

$$\underline{a} + \underline{b} = \underline{b} + \underline{a}$$

$$(\underline{a} + \underline{b}) + \underline{c} = \underline{a} + (\underline{b} + \underline{c})$$

$$\underline{a} + \underline{0} = \underline{a}$$

$$\underline{a} + (-\underline{a}) = \underline{0}$$

$$|\underline{b}| = |m||\underline{a}|; \underline{b} = m\underline{a}$$

* Malvern, pg. 572

II.1.3. Linear independence

Borisenko and Tarapov, pg. 7: We say that n vectors $\underline{v}_1, \underline{v}_2, \underline{v}_3, \ldots \underline{v}_n$ are **linearly dependent** if there exist n scalars $a_1, a_2, a_3, \ldots a_n$ (not all equal to zero) such that

$$a_1 \underline{v}_1 + a_2 \underline{v}_2 + a_3 \underline{v}_3 + \ldots a_n \underline{v}_n = 0,$$

that is, if some (nontrivial) linear combination of vectors equals zero.

- **If one vector can be expressed in terms of another, or another set, the set of vectors including said vectors is said to be linearly dependent.**
- **Two linearly dependent vectors are collinear; the ratio of the two magnitudes of the vectors is given by the ratio of the a's given above.**
- **Three linearly dependent vectors lie in the same plane**

How do you determine $(x_1, x_2, x_3, \ldots x_n)$, $(y_1, y_2, y_3, \ldots y_n)$ and so forth are linearly independent? (Or, how do you know if they lie in the same plane?) Use the determinant of the matrix formed by n vectors (of n dimension): if the determinant is not zero, the vectors are linearly *in*dependent. For example, consider the vectors v_1=(1,0,0), v_2=(4,5,6) and v_3=(0,5/2,3) - are they linearly independent?

$$\begin{vmatrix} 1 & 0 & 0 \\ 4 & 5 & 6 \\ 0 & 5/2 & 3 \end{vmatrix} = 1[5 \cdot 3 - 5 \cdot 6/2] = 0$$

Thus, these vectors are linearly *dependent*. Why? Because $v_3 = (1/2)v_2 - 2v_1$!

II.1.4. Base vectors

Any n linearly independent vectors can form a basis for an n-dimensional system. Any vector can be *expanded* about a given basis: that is, expressed as a sum of coefficients and base vectors.

- **Base vectors do *not* need to be of unit length**
- **Base vectors do *not* need to be orthogonal**
- **Base vectors *can* depend on position**

Base vectors are expressed as ordered n-tuples with respect to a given coordinate system; you can transform base vectors to another coordinate system, and, you can transform vectors expressed in one basis to vectors expressed in another basis.

Orthogonal bases: all bases vectors are mutuallly orthogonal.
Orthonormal bases: all base vectors are mutually orthogonal and have unit length.
Oblique bases: base vectors are not orthogonal and may or may not be of unit length.

Natural covariant basis vectors are derived from the form: $\underline{e}_i = \frac{\partial \underline{x}}{\partial x_i}$

II.2 Vectors in rectangular Cartesian coordinate systems

II.2.1. Dot product

The dot or "inner" product is defined most generally as:

$$\underline{p}\cdot\underline{q} = p^i q^j(\underline{e}_i\cdot\underline{e}_j) = p^i q^j g_{ij}$$

where g_{ij} is the ***metric tensor*** for the given basis. This expression holds for any set of basis vectors; for rectangular Cartesian coordinate systems, it simplifies further.

For **rectangular Cartesian coordinate systems**, the metric tensor components are given by the Kronecker delta, *and the covariant and contravariant bases* (to be defined later) and components of vectors are identical: hence, for rectangular Cartesian coordinate systems (RCCS):

$$\underline{p}\cdot\underline{q} = p^i q^j \delta_{ij} = p^i q^i = p_i q_i \quad \text{(RCCS)}$$

The dot product is thus a ***contraction***, as it lowers the order of the tensors: i.e., the dot product of two vectors (1st order tensors) is a scalar (0th order tensor).

Some properties of the dot product (***valid for all coordinate systems***):

a.) $\underline{p}\cdot\underline{q} = \underline{q}\cdot\underline{p}$
b.) $\alpha\underline{p}\cdot\underline{q} = \underline{p}\cdot\alpha\underline{q} = \alpha(\underline{p}\cdot\underline{q})$
c.) $\underline{p}\cdot(\underline{q}+\underline{r}) = \underline{p}\cdot\underline{q} + \underline{p}\cdot\underline{r}$
d.) $\underline{p}\cdot\underline{p} \geq 0$; $\underline{p}\cdot\underline{p} = 0$ iff $\underline{p} = \underline{0}$ Positive definite property
e.) $|\underline{p}\cdot\underline{p}| = p = (\underline{p}\cdot\underline{p})^{1/2} = (p^k p_k)^{1/2}$ **Magnitude of vector p**

For rectangular Cartesian coordinates (and some others), the dot product has the geometrical interpretation of relating the angle between two vectors, in the plane containing the vectors...

$$\underline{p}\cdot\underline{q} = |\underline{p}||\underline{q}|\cos(\theta)$$

Thus, the dot product of two vectors can also be thought of as the *projection* of one vector onto another: the dot product of **p** and **q** is the magnitude of **p** times of projection of **q** onto **p**.

Clearly, in orthogonal coordinate systems, the dot product of two vectors that are perpendicular is zero. That is,

$$\underline{p}\cdot\underline{q} = 0 \text{ then } \underline{p} \text{ is perpendicular to } \underline{q} \quad (\underline{p}\perp\underline{q})$$

Equation of a <u>plane</u> using the dot product: given a point located by the position vector $\underline{x}_o$ and a vector normal the desired plane, $\underline{n}$, any vector in the plane $\underline{x}$ will obey the following equation:

$$\underline{n}\cdot(\underline{x}-\underline{x}_o) = 0$$

Equation of a <u>sphere</u> using the dot product: given a point located by the position vector $\underline{x}_o$, a sphere of radius a centered at that point is governed by the equation below, where $\underline{x}$ is the position vector of a point on the sphere (with origin defined by $\underline{x}_o$):

$$(\underline{x}-\underline{x}_o)\cdot(\underline{x}-\underline{x}_o) = a^2$$

II.2.2. Cross product

The cross product is defined as a binary relation between two vectors given by:

$$\underline{p}\times\underline{q} = \underline{e}_r \varepsilon_{rst} p^s q^t$$

where e is the alternating symbol.

If $\underline{p}$ and $\underline{q}$ are ***Cartesian*** vectors, then the components of $\underline{p}$ x $\underline{q}$ transform as vectors with respect to *orthogonal* groups of transformations; that is, ***the cross product results in a vector***. This is equivalent to the symbolic determinant expressed as:

$$\underline{p}\times\underline{q} = \begin{vmatrix} \underline{e}_1 & \underline{e}_2 & \underline{e}_3 \\ p^1 & p^2 & p^3 \\ q^1 & q^2 & q^3 \end{vmatrix}$$

<u>Rules of the cross product:</u>

a.) $\underline{p}\times\underline{q} = -\underline{q}\times\underline{p}$
b.) $\underline{p}\times(\underline{q}+\underline{r}) = \underline{p}\times\underline{q} + \underline{p}\times\underline{r}$
c.) $\alpha(\underline{p}\times\underline{q}) = \alpha\underline{p}\times\underline{q} = \underline{p}\times\alpha\underline{q}$
d.) $\underline{p}\times\underline{q} = |\underline{p}||\underline{q}|\sin\theta\,\underline{n}$ where q is the angle between $\underline{p}$ and $\underline{q}$ and $\underline{n}$ is the unit normal to the plane containing $\underline{p}$ and $\underline{q}$.
e.) **Triple scalar product:** (not invariant with respect to an inversion):
$$\underline{p}\cdot\underline{q}\times\underline{r} = \varepsilon_{ijk} p^i q^j r^k = \underline{p}\times\underline{q}\cdot\underline{r}$$
d) **Triple vector cross:** $\underline{p}\times\underline{q}\times\underline{r} = \underline{q}(\underline{p}\cdot\underline{r}) - \underline{r}(\underline{p}\cdot\underline{q})$

<u>Orthonormal base vectors:</u> $\underline{e}_i\cdot\underline{e}_k = \delta_{ij}$ and $\underline{e}_i\times\underline{e}_j\cdot\underline{e}_k = \varepsilon_{ijk}$

II.2.3. Transformations (Cartesian components with no change in unit distance along the coordinate)

Direct transformation: $\bar{\underline{e}}_i = a_{ij}\underline{e}_j$, where a_{ij} are the coefficients of the direct transformation. a_{ij} are the direction cosines for the direct transformation, *taken from the unbarred to the barred axis*.

Inverse transformation: $\underline{e}_i = \bar{a}_{ij}\bar{\underline{e}}_j$, where $\bar{a}_{ij}$ are the coefficients of the direct transformation.

Orthogonal transformations: rotations of coordinate axes.

$$a_{jp}a_{kp} = \delta_{jk} \qquad a_{pj}a_{pk} = \delta_{jk}$$
$$\bar{a}_{jp}\bar{a}_{kp} = \delta_{jk} \qquad \bar{a}_{pj}\bar{a}_{pk} = \delta_{jk}$$
$$\bar{a}_{ij} = a_{ji} \qquad \text{(orthogonal transformations)}$$
$$\bar{a}_{jp}a_{pk} = \delta_{jk} \qquad \bar{a}_{jp}a_{pk} = \delta_{jk}$$

II.3. General bases

II.3.1. Introduction to general bases

Vectors in general coordinate systems can be expanded about a variety of bases, and the notation used for the vector components highlights the basis being used. Contravariant components correspond to the **covariant** basis, which is generally found via the relationship:

$$\underline{e}_i = \frac{\partial \underline{x}}{\partial x_i}$$

where x_i are the general coordinates and $\underline{x}$ is the position vector. Basis vectors found via this relationship are referred to as the ***natural covariant basis vectors***. The **subscript** denotes the **covariant** quantity. **Superscripts** are used to denote **contravariant** quantities: since the contravariant components go with the covariant basis, and vice versa, superscripts and subscripts appear in pairs, as below:

$$\underline{v} = v^i\underline{e}_i = v_i\underline{e}^i$$

Typically, the type of component dictates the name of the representation: i.e., the first summation above is the contravariant representation (even though the basis vectors are said to be covariant), and the second is the covariant representation (even though the basis vectors are said to contravariant.) Malvern writes:

"Some authors call the set of components v_i a covariant vector and the set of components v^j a contravariant vector, ***but there is just one vector***, represented in terms of contravariant components with respect to the given base vectors, or convariant components with respect to the reciprocal base."

Covariant and contravariant bases are said to be reciprocal; this will be defined in subsequent sections.

II.3.2. The metric tensor, the dot product and cross product for general bases

When general base vectors are used, **the dot product** is defined as:

$$\underline{p}\cdot\underline{q} = p^r q^s \underline{e}_s \cdot \underline{e}_r = g_{rs}p^r q^s$$

where g_{rs} is the **metric tensor** for the basis: the metric tensor relate the dot products of the basis vectors for the space - that is,

$$\underline{e}_s \cdot \underline{e}_r = g_{rs} \qquad \text{and} \qquad \underline{e}^s \cdot \underline{e}^r = g^{rs}$$

where g_{rs} are the covariant components of the tensor, and g^{rs} are the contravariant components of the tensor. **"The metric tensor describes the fundamental geometric characteristics of a space arithmetized using $\underline{e}_i$ and x_i."** (Borisenko and Tarapov, pg. 32) One of the ways this can be illustrated is by considering how an element of arc length is defined in a general coordinate system.

The arc length between two neighboring points, x_i and $x_i + dx_i$, can be written as

$$(ds)^2 = d\underline{x}\cdot d\underline{x}$$

where $d\underline{x}$ is the infinitessimal displacement vector which points from x_i to the location $x_i + dx_i$. The vectors are written as the sum of the change in the position vector with respect to a coordinate and the incremental change in that coordinate: that is,

$$(ds)^2 = \frac{\partial \underline{x}}{\partial x_i}dx_i \cdot \frac{\partial \underline{x}}{\partial x_j}dx_j$$

Then, with the natural basis vectors defined as above, this is the same as...

$$(ds)^2 = \underline{e}_i dx_i \cdot \underline{e}_j dx_j$$

And finally,

$$(ds)^2 = g_{ij}dx_i dx_j$$

Clearly, with basis vectors chose as $\underline{e}_j = \frac{\partial \underline{x}}{\partial x_j}$, **$g_{ij}$ provides the set of factors for converting increments in coordinates to increments in length. Thus, the metric tensor controls how incremental length is defined in the space "arithmetized" by the chosen coordinates (and basis vectors). Because of this control over incremental length, the metric tensor will be shown to play a critical role in differentiation in general coordinate systems.**

You have used the metric tensor countless times before! What is the change in distance in polar coordinates when the radius is held fixed and the angle is perturbed?

$$(ds)^2 = (rd\theta)^2$$

So, clearly, $g_{22} = g_{rr} = r^2$! We'll come back to this example in more detail....

Finally, the cross product is defined as follows

$$\underline{p} \times \underline{q} = \sigma\sqrt{g}\varepsilon_{rst}p^r q^s \underline{e}^t$$

where s equals positive one if the covariant basis vectors are right-handed, and negative one otherwise, and g is the determinant of the matrix formed by the metric tensor g_{rs}.

II.3.3 Reciprocal bases

The **natural basis** has been defined as

$$\underline{e}_i = \frac{\partial \underline{x}}{\partial x_i}$$

and is often referred to as a "covariant" basis. The question arises is there a related basis - one that may be referred to as contravariant? Reciprocal bases are defined as having the following relationship:

$$\underline{e}_i \cdot \underline{e}^j = \delta_{ij}$$

Thus, the "contravariant" basis corresponding to the natural basis can be found using this relationship.

Moveover, it is seen that the metric tensor "maps" one set of basis vectors into the other:

$$\underline{e}_i = g_{ij}\underline{e}^j$$

Or, using the contravariant components of the metric tensor,

$$\underline{e}^i = g^{ij}\underline{e}_j$$

It follows from this definition that if the given basis is rectangular Cartesian, the reciprocal basis is indentical to the given basis: or, the contravariant and covariant basis vectors are the same.

The reciprocal basis can be found using the natural basis and the cross product (though the cross product in general coordinate systems has yet to be defined). Since $\underline{e}^1$ perpendicular to $\underline{e}_2$ and $\underline{e}_3$ (according to definition above), the $\underline{e}^1$ has to be in the direction of $\underline{e}_2 \times \underline{e}_3$;

$$E\underline{e}^1 = \underline{e}_2 \times \underline{e}_3$$

Dotting both sides with $\underline{e}_1$ will determine the scalar factor E;

$$E\underline{e}^1 \cdot \underline{e}_1 = \underline{e}_2 \times \underline{e}_3 \cdot \underline{e}_1 = E$$

Thus, the reciprocal basis are given by:

$$\underline{e}^i = \frac{\underline{e}_j \times \underline{e}_k}{\underline{e}_i \cdot \underline{e}_j \times \underline{e}_k}$$

where i,j,k should be an even permutation. A simple calculation using this relationship illustrates that for **orthonormal** basis vectors, there is no distinction between covariant and contravariant basis vectors.

When give bases are orthogonal but **not** orthonormal, the following is true - note that the superscript and subscript indicate different quantities:

$$\left|\underline{e}^i\right| = \frac{1}{\left|\underline{e}_i\right|}$$

This reciprocal relationship is not satisfied when the given basis is not orthogonal. (It is true for orthogonal curvilinear coordinate systems.)

III.3.4. Invariance

It should be noted that the dot product, defined as above, is *invariant* upon coordinate change. That is, a change of coordinates does not affect the scalar produced by the dot product.

The invariant product $\underline{p} \cdot \underline{q}$ can be written a variety of ways....

$$\underline{p} \cdot \underline{q} = g_{rs}p^r q^s = g^{rs}p_r q_s = g_r^s p_s q^r$$

but is generally written as follows:

$$\underline{p} \cdot \underline{q} = p^r q_r = p_r q^r$$

with one superscript (contravariant component) and one subscript (covariant component). This form is used to prove the invariance of the dot product under coordinate change.

Consider the scalar given by:

$$E = \underline{p} \cdot \underline{q} = p^r q_r$$

After transforming the vector $\underline{p}$ and $\underline{q}$ into another coordinate system, the following series of manipulations proves the scalar E will be the same in each coordinate system....

$$\overline{p}^i = p^r \frac{\partial \overline{x}_i}{\partial x_r} \quad \text{and} \quad \overline{q}_i = q_s \frac{\partial x_s}{\partial \overline{x}_i}$$

Then

$$E = \underline{\bar{p}} \cdot \underline{\bar{q}} = \bar{p}^r \bar{q}_r$$

$$E = \bar{p}^r \bar{q}_r = p^i \frac{\partial \bar{x}_r}{\partial x_i} q_j \frac{\partial x_j}{\partial \bar{x}_r}$$

$$E = p^i q_j \frac{\partial \bar{x}_r}{\partial x_i} \frac{\partial x_j}{\partial \bar{x}_s} = p^i q_j \frac{\partial x_j}{\partial x_i} = p^i q_j \delta_{ij} = p^i q_i$$

II.3.5. Covariant and contravariant representations and the Jacobian

Contravariant components are the vector components that correspond to expansion about the *covariant basis*, that is:

$$\underline{v} = v^i \underline{e}_i$$

Or, the same vector can be expressed in terms of covariant components and the reciprocal basis to the covariant basis (sometimes referred to as the *contravariant basis*):

$$\underline{v} = v_i \underline{e}^i$$

So, do the two sets of components imply two different vectors? NO! Malvern sums it best:

"Some authors call the set of components v_j a covariant vector and the set v^k a contravariant vector, but from our point of view there is just one vector, name v, which, for a given basis, may be represented in terms of contravariant components with respect to the give base vectors, or alternatively, in terms of covariant components with respect to the reciprocal basis of the given one"[1].

The prototype of the covariant representation of a vector is the gradient of a scalar, with the components given by $\frac{\partial f}{\partial x_j}$; it will be shown later that these quantities behave as covariant components of a vector. The term "covariant" derives from the fact that these components transform in the same manner as the covariant base vectors; *that is, covariant components transform by the same coefficients used to transform the given base vectors*. By contrast, the contravariant components transform by the coefficients used in the backward change from the reciprocal basis to the given basis.

This is expressed mathematically as follows. Suppose the base vectors transform according to the following relationships:

$$\underline{\bar{e}}_j = a_j^p \underline{e}_p \quad \text{and} \quad \underline{e}^j = b_j^p \underline{\bar{e}}^p$$

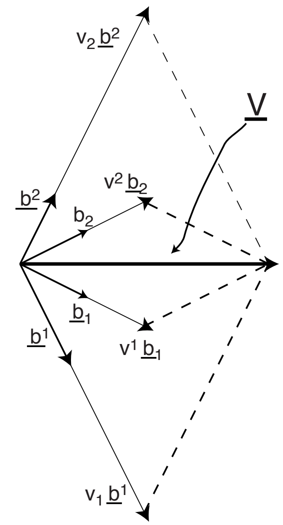
Then the covariant components will transform via:

[1] *Introduction to the Mechanics of a Continuous Medium*, Malvern, pg. 575.

$$\bar{v}_j = a_j^p v_p$$

And the contravariant components tranform via:

$$\bar{v}^j = b_j^p v^p$$



If you don't buy the coefficient explanation, you apparently have excellent company in Prof. Emeritus B. Budiansky at Harvard, who writes in his classical chapter on tensors:

"For historical reasons (now unimportant) the v_i are call covariant components of $\underline{v}$ and the v^i are contravariant."[2]

$\underline{\mathbf{b}}^1$ and $\underline{\mathbf{b}}^2$ are reciprocal base vectors for $\underline{\mathbf{b}}_1$ and $\underline{\mathbf{b}}_2$.

$\underline{\mathbf{v}} = v_j \underline{\mathbf{b}}^j$; v_j are covariant components of $\underline{\mathbf{v}}$

$\underline{\mathbf{v}} = v^j \underline{\mathbf{b}}_j$; v^j are contravariant components of $\underline{\mathbf{v}}$.

The coefficients of transformation are determined from the Jacobian matrix, which relates how the coordinate systems map into one another. The Jacobian matrix is:

$$J_{ij} = \frac{\partial \bar{x}_i}{\partial x_j}$$

It will be noted that the inverse of the Jacobian matrix provides the coefficients of the reverse transformation. That is,

$$J_{ij}^{-1} = \frac{\partial x_j}{\partial \bar{x}_i}$$

The following identity should be noted:

$$\frac{\partial x_i}{\partial \bar{x}_j} \frac{\partial \bar{x}_j}{\partial x_k} = \delta_{ik}$$

[2] B. Budiansky, *Chapter 4: Tensors*, *Handbook of Applied Mathematics: Selected results and Methods*, ed. by Carl E. Pearson, von Nostrand Reinhold Co., 1974, New York.

For orthogonal transformations, such as rotations of Cartesian coordinate systems, $J^{-1} = J^T$, so contravariant and covariant tensors transform identically; that is, there is no distinction between covariant and contravariant components in Cartesian coordinate systems.

The transformation equations for covariant and contravariant components are summarized below.

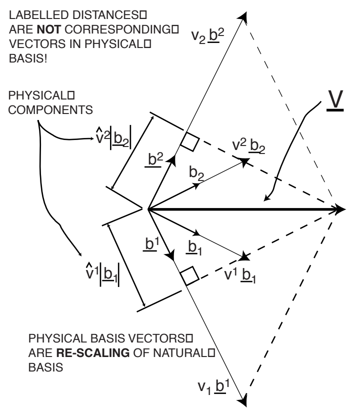
Contravariant: $\bar{v}^i = \dfrac{\partial \bar{x}^i}{\partial x^j} v^j$ and $v^i = \dfrac{\partial x^i}{\partial \bar{x}^j} \bar{v}^j$

Covariant: $\bar{v}_i = \dfrac{\partial x^j}{\partial \bar{x}^i} v_j$ and $v_i = \dfrac{\partial \bar{x}^j}{\partial x^i} \bar{v}_j$

II.3.6. Physical components

It may have caught your attention that some of the basis vectors that have been put forth are not dimensionally consistent; that is, the units vary from base vector to base vector, or even component to component. This stems from the fact that units of the coordinate systems themselves vary: consider polar coordinates, where the units are length and angle.

The basis vectors and components of vectors in general coordinates can be normalized to remove this lack of consistency - the components are then called *physical components*, which correspond to a new basis - the *physical basis*. The physical components and basis are found using these relationships:



$$\underline{v} = \hat{v}^k \hat{\underline{e}}_k$$

where the physical base vectors are:

$$\hat{\underline{e}}_k = \frac{\underline{e}_k}{\sqrt{g_{kk}}} \quad \text{(no sum)}$$

And the physical components corresponding to this basis are:

$$\hat{v}^k = v^k \sqrt{g_{kk}} \quad \text{(no sum)}$$

II.3.4 The metric tensor

The metric tensor is a second order tensor which describes fundamental geometric characteristics of a space "arithmetized" by $\underline{\mathbf{e}}_i$ and (x_1,x_2, etc.) (see Borisenko and Tarapov, pg. 32.) The dot-product of vectors in a general space are defined via the metric tensor (or vice versa):

$$g_{ij} = \underline{e}_i \cdot \underline{e}_j$$
$$g^{ij} = \underline{e}^i \cdot \underline{e}^j$$

Perhaps the most important role of the metric tensor is the fact that it defines "length" in an arbitrary, non-orthogonal coordinate system. Consider the metric tensor for natural, covariant basis vectors described by:

$$\underline{e}_i = \frac{\partial \underline{x}}{\partial x_i}$$

Then the metric tensor is:

$$g_{ij} = \frac{\partial \underline{x}}{\partial x_i} \cdot \frac{\partial \underline{x}}{\partial x_j}$$

Consider an the definition of an arc-length:

$$ds^2 = d\underline{x} \cdot d\underline{x}$$

Then...

$$ds^2 = \frac{\partial \underline{x}}{\partial x_i} dx^i \cdot \frac{\partial \underline{x}}{\partial x_j} dx^j = g_{ij} dx^i dx^j$$

Therefore, the "length" along a given curve in space is given by:

$$L = \int_a^b \sqrt{g_{ij} \frac{dx_i}{dt} \frac{dx_j}{dt}}\, dt$$

These concepts are essential when working with surfaces, where coordinate systems and vectors are defined in terms of the surface topology: SHELLS! These relationships will be explored further in studying the geometry of space curves.

II.4 Example

Consider the ordered n-tuple (1,0) in a rectangular Cartesian coordinate system. Use the ordered n-tuples (1,0) and (0,1) as a basis in the original coordinate system. Consider the transformation of this vector in a polar coordinate systems. The equations relating the coordinate systems are:

$$x_1 = \bar{x}_1 \cos(\bar{x}_2)$$
$$x_2 = \bar{x}_1 \sin(\bar{x}_2) \qquad 3.4.1$$

Using vector notation, the given vector is expanded about the chosen basis as:

$$\underline{v} = v^1\underline{e}_1 + v^2\underline{e}_2 = 1\underline{e}_1$$

The position vector of any point is given by:

$$\underline{x} = x_1\underline{e}_1 + x_2\underline{e}_2$$

Or, in terms of the barred coordinates and the **chosen** basis vectors,

$$\underline{x} = \overline{x}_1\cos(\overline{x}_2)\underline{e}_1 + \overline{x}_1\sin(\overline{x}_2)\underline{e}_2$$

The *natural basis vectors* (or covariant basis vectors) in the barred system are determined via the relationship

$$\underline{\overline{e}}_k = \frac{\partial \underline{x}}{\partial \overline{x}_k}$$

... and are found to be:

$$\underline{\overline{e}}_1 = (\cos\overline{x}_2, \sin\overline{x}_2)$$
$$\underline{\overline{e}}_2 = (-\overline{x}_1\sin\overline{x}_2, \overline{x}_1\cos\overline{x}_2)$$

Or, in terms of the base vectors in the original system:

$$\underline{\overline{e}}_1 = \cos\overline{x}_2\underline{e}_1 + \sin\overline{x}_2\underline{e}_2$$
$$\underline{\overline{e}}_2 = -\overline{x}_1\sin\overline{x}_2\underline{e}_1 + \overline{x}_1\cos\overline{x}_2\underline{e}_2$$

In the process of determining the natural basis vectors, we differentiated a vector: **caution must be urged** - for the simple case considered here, the base vectors in the given system are constant and thus their derivatives are zero. This generally need not be true, leading to complications; differentiation of general base vectors will be explored in subsequent lectures...

These are not the typical base vectors used in polar coordinates, as they are **not** orthonormal, though they **are** orthogonal. We will see at the conclusion of this example that the base vectors and components we typically use in polar coordinates are the *physical basis and components*.

So, once the natural basis is found, what are the contravariant components to the given vector, which correspond to the basis just derived? Since the components in the original system are **both** covariant **and** contravariant, no distinction need be made for the unbarred components. However, the contravariant components in the barred system are found via:

$$\overline{v}^j = \frac{\partial \overline{x}_j}{\partial x_k} v^k$$

Of course, the Jacobian $\frac{\partial \overline{x}_j}{\partial x_k}$ could be computed via eqn. 3.4.1; but this would require inverting the expressions, which leads to more complicated differentiation. Rather, we'll compute the inverse of the Jacobian, $\frac{\partial x_k}{\partial \overline{x}_j}$, then invert the expression, which is simple for a 2x2 matrix:

$$\frac{\partial \overline{x}_i}{\partial x_j} = \begin{bmatrix} \cos\overline{x}_2 & -\overline{x}_1\sin\overline{x}_2 \\ \sin\overline{x}_2 & \overline{x}_1\cos\overline{x}_2 \end{bmatrix}^{-1}$$

$$\frac{\partial \overline{x}_i}{\partial x_j} = \begin{bmatrix} \cos\overline{x}_2 & \sin\overline{x}_2 \\ \frac{\sin\overline{x}_2}{\overline{x}_1} & \frac{\cos\overline{x}_2}{\overline{x}_1} \end{bmatrix}$$

Using these relationships, we can directly compute both *the covariant and contravariant components of the vector* $\underline{v}$ *in the barred system*:

Contravariant: $\overline{v}^k = \left(\cos\overline{x}_2, \frac{-\sin\overline{x}_2}{\overline{x}_1}\right)$

Covariant: $\overline{v}_k = (\cos\overline{x}_2, -\overline{x}_1\sin\overline{x}_2)$

The invariance of the dot-product can be immediately verified; compare the product of $\underline{v}$ dotted with itself, in both coordinate systems:

$$\underline{v}\cdot\underline{v} = (1,0)\cdot(1,0) = 1$$
$$\underline{\overline{v}}\cdot\underline{\overline{v}} = \overline{v}^k\overline{v}_k = \cos^2(\overline{x}_2) + \frac{-\sin(\overline{x}_2)}{\overline{x}_1}\left(-\overline{x}_1\sin(\overline{x}_2)\right) = 1$$

So far, the natural basis vectors have been derived, as well as the contravariant components that correspond to this basis; the question remains - what are the reciprocal basis vectors to which the covariant components correspond?

$$\underline{\overline{e}}_i\cdot\underline{\overline{e}}^j = \delta_{ij}$$

$$\underline{\overline{e}}_1\cdot\underline{\overline{e}}^1 = 1$$
$$\underline{\overline{e}}_2\cdot\underline{\overline{e}}^1 = 0$$

Solving this linear system of equations for the contravariant basis vectors:

$$\underline{\overline{e}}^1 = (\cos\overline{x}_2, \sin\overline{x}_2)$$
$$\underline{\overline{e}}^2 = \left(\frac{-\sin\overline{x}_2}{\overline{x}_1}, \frac{\cos\overline{x}_2}{\overline{x}_1}\right)$$

Lastly, the physical components are found quite simply. Recall that the metric tensor (in the barred coordinate system) is defined by:

$$\bar{g}_{ij} = \underline{\bar{e}}_i \cdot \underline{\bar{e}}_j$$

where the bar denote it is the metric tensor for the polar coordinate space. Given the basis vectors in the barred system, the physical components are quite easily found:

$$\bar{g}_{11} = 1 \quad \text{and} \quad \bar{g}_{22} = \bar{x}_1^{\,2}$$

$$\hat{\underline{e}}_1 = \frac{\underline{e}_1}{\sqrt{g_{11}}} = \left(\cos\bar{x}_2, \sin\bar{x}_2\right)$$

$$\hat{\underline{e}}_2 = \frac{\underline{e}_2}{\sqrt{g_{22}}} = \left(-\sin\bar{x}_2, \cos\bar{x}_2\right)$$

And the components are given by...

$$\hat{\bar{v}}^1 = \bar{v}^1\sqrt{g_{11}} = \cos\bar{x}_2$$
$$\hat{\bar{v}}^2 = \bar{v}^2\sqrt{g_{22}} = -\sin\bar{x}_2$$

It is clear then, that the physical basis vectors correspond to $\underline{e}_r, \underline{e}_\theta$ - the base vectors we traditionally use in polar coordinates.

III. DIFFERENTIATION

In this section, differentiation of scalars and vectors is discussed, primarily in the familiar framework of constant basis vectors; derivatives of vectors expressed in terms of non-constant basis vectors are introduced to a small degree to set the stage for further development. Theorems related to vector fields are presented, with simple (yet meaningful) problems in vector mechanics to illustrate their utility.

III.1 Derivatives of scalars and vectors in rectangular Cartesian coordinate systems

A scalar field is a mathematical entity which has a scalar quantity associated with each point in space. If $(x_1,x_2,...x_n)$ are the coordinates describing the space, then a scalar field is defined as:

$$\Phi = f(x_1,x_2,...,x_n) = \Phi(x_1,x_2,...,x_n)$$

The **total derivative** of a scalar field with respect to some variable *t* is straightforward, and given by:

$$\frac{d\Phi}{dt} = \frac{\partial\Phi}{\partial x_i}\frac{\partial x_i}{\partial t}$$

where $\frac{\partial\Phi}{\partial x_i}$ and $\frac{\partial\Phi}{\partial t}$ are partial derivatives defined in the usual manner.

Suppose u^j are three functions defined in space, as in:

$$\begin{aligned} u^1 &= f_1(x_1,x_2,...,x_n) \\ u^2 &= f_2(x_1,x_2,...,x_n) \\ &\vdots \\ u^n &= f_n(x_1,x_2,...,x_n) \end{aligned}$$

If, at any (every) point in space, the functions u^j transform as components of $\underline{\mathbf{u}}$, then $\underline{\mathbf{u}}$ is *a vector field*, i.e. $\underline{\mathbf{u}} = f(x_1,x_2,...x_n)$. Note that in order to test these function to determine if they represent a vector field, one must have representations in two "arithmetized" coordinate systems, in order to check the transformation. **For all practical intents and purpose, one knows a priori whether or not it is a vector field from the application at hand.** (I.e. you *know* $\underline{\mathbf{u}}$ is vector field because it is representing displacement, velocity, etc.)

The derivative a vector field is defined as:

$$\frac{d\underline{u}}{dt} = \lim_{\Delta t\to 0}\frac{\underline{u}(t+\Delta t)-\underline{u}(t)}{\Delta t}$$

This definition implies that the derivative of a vector field is itself a vector field; subtracting two vectors results in a vector, while multiplying by $(1/\Delta t)$ merely scales the length of the vector. *For a constant basis that does not depend on t, and, if t is a scalar*, the derivative of a vector is simply:

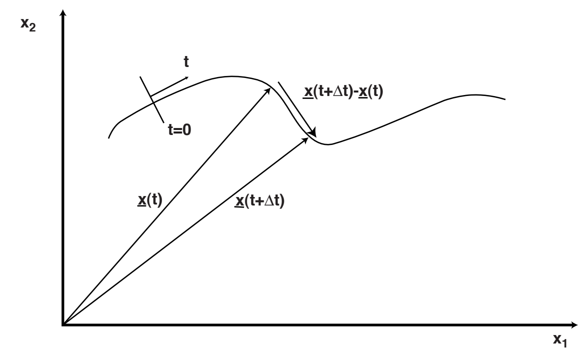
$$\frac{d\underline{u}}{dt} = \frac{du^j}{dt}\underline{e}^j$$

Thus, the derivative of a vector field which is a function of scalars, with respect to a scalar, is a vector as well.

Is a coordinate a scalar? That is, if I differentiate a vector with respect to a coordinate, do I get a vector? Not quite. This will be discussed in detail in Section 3.

III.2 The geometry of space curves

To illustrate derivatives with respect to a scalar (but not necessarily a coordinate), consider the derivative a vector field defined along a curve, parameterized by t:



The derivative of the vector field with respect to the parameter *t* is *the tangent vector field*:

$$\underline{t} = \frac{d\underline{x}}{dt}$$

Here $\underline{\mathbf{x}}$ is the a vector field which describes the position vector as a function of distance along a curve. That is, for all points on the curve:

$$x_i = x_i(t)$$

over some range, say $t_o \le t \le t_f$. Then, the length of a given section of curve is defined as:

$$L = \int_{t_o}^{t_f} \sqrt{g_{ij} \frac{dx_i}{dt} \frac{dx_j}{dt}} dt$$

This is the general expression for any "arithmetized" space: Since we're dealing with rectangular Cartesian space, the metric tensor is equal to the Kronecker delta, and hence:

$$L = \int_{t_o}^{t_f} \sqrt{\delta_{ij} \frac{dx_i}{dt} \frac{dx_j}{dt}} dt = \int_{t_o}^{t_f} \sqrt{\frac{dx_i}{dt} \frac{dx_i}{dt}} dt$$

It is important to note that t is not necessarily the arc length along the curve. The arc length, s, is defined as:

$$s = \int_0^t \sqrt{\frac{dx_i}{dt} \frac{dx_i}{dt}} dt$$

The arc length s is used to define the three fundamental vector fields which characterize the curve:

The **unit tangent vector field**: $\underline{t} = \frac{d\underline{x}}{ds}$,

the **unit principal normal vector**: $\underline{n} = \frac{1}{\kappa} \frac{d\underline{t}}{ds}$ where the **curvature** is defined as:

$$\kappa = \sqrt{\frac{d\underline{t}}{ds} \frac{d\underline{t}}{ds}},$$

and the **bi-normal vector field**: $\underline{b} = \underline{t} \times \underline{n}$.

These quantities combine to form what are known as **the Frenet-Serrat formulae**:

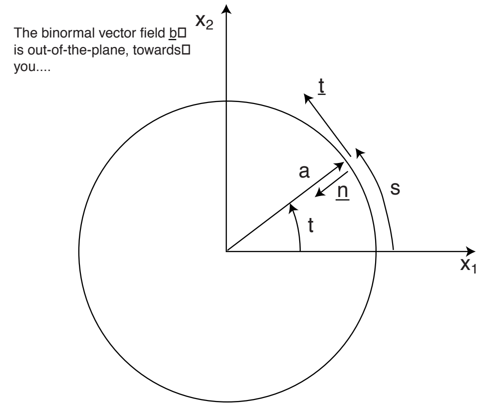
$$\frac{d\underline{t}}{ds} = \kappa \, \underline{n}$$
$$\frac{d\underline{n}}{ds} = -\kappa \, \underline{t} + \tau \, \underline{b}$$
$$\frac{d\underline{b}}{ds} = -\tau \, \underline{n}$$

where τ is defined as the **torsion** of the curve and is defined by:

$$\tau = \tau(s) = - \, \underline{n} \cdot \frac{d\underline{b}}{ds}$$

Derivation of these formulae follows naturally from the above definitions.

A quick and easy example is a circle of radius a:



The correct parameterization of this curve is:

$$x_1(t) = a \cos t$$
$$x_2(t) = a \sin t$$

Or, in terms of arc length (makes life a bit easier later):

$$s = \int_0^t \sqrt{\frac{dx_i}{dt} \frac{dx_i}{dt}} dt = \int_0^t \sqrt{(a \cos(t))^2 + (a \sin(t))^2} dt$$

$$s = a\,t$$

The position vector is thus:

$$\underline{x}(s) = a \cos\left(\frac{s}{a}\right)\underline{e}_1 + a \sin\left(\frac{s}{a}\right)\underline{e}_2$$

where the unit basis vectors are defined in the usual manner. The **unit tangent vector** of the curve is:

$$\underline{t}(s) = \frac{d\underline{x}}{ds} = -\sin\left(\frac{s}{a}\right)\underline{e}_1 + \cos\left(\frac{s}{a}\right)\underline{e}_2$$

which is consistent with what we would expect (see the figure above). The **unit normal** is:

$$\kappa = \sqrt{\frac{d\underline{t}}{ds} \frac{d\underline{t}}{ds}} = \frac{1}{a^2}$$

$$\underline{n}(s) = \frac{1}{\kappa}\frac{dt}{ds} = \cos\left(\frac{s}{a}\right)\underline{e}_1 + \sin\left(\frac{s}{a}\right)\underline{e}_2$$

III.3 The del operator ∇

Consider the equilibrium equations we all know:

$$\sigma_{ij,j} = 0$$

Does this imply

$$\frac{\partial \sigma_{rr}}{\partial r} + \frac{\partial \sigma_{r\theta}}{\partial \theta} = 0\,?$$

III.3.1 Defining the Del Operator

Consider the scalar *f* which is a function of coordinates: $f=f(x_1,x_2,x_3)$. The total derivative of f with respect to a given coordinate is:

$$\frac{df}{dx_k} = \frac{\partial f}{\partial x_i}\frac{\partial x_i}{\partial x_k}$$

If we denote

$$v_i \equiv \frac{\partial f}{\partial x_i},$$

Then v_i must be the covariant components of a vector. Why? Because uder a change of coordinates from $x_j \rightarrow \overline{x}_j$, the components of v_i in the barred system are:

$$\overline{v}_k = \frac{\partial f}{\partial \overline{x}_k} = \frac{\partial x_i}{\partial \overline{x}_k} v_i = \frac{\partial x_i}{\partial \overline{x}_k}\frac{\partial f}{\partial x_j}\frac{\partial x_j}{\partial x_i}.$$

That is, the components of v_i transform like covariant components of a vector, therefore they must be a vector. If we expand these covariant components about unit (contravariant) base vectors, we define the gradient of *f*:

$$\underline{v} \equiv \text{grad } f = \frac{\partial f}{\partial x_1}\underline{e}^1 + \frac{\partial f}{\partial x_2}\underline{e}^2 + \frac{\partial f}{\partial x_3}\underline{e}^3$$

The del operator, ∇, is defined as the operator which acts on a scalar function f to produce the gradient of f, denoted grad f:

$$\text{grad } f = \nabla f = \frac{\partial f}{\partial x_1}\underline{e}^1 + \frac{\partial f}{\partial x_2}\underline{e}^2 + \frac{\partial f}{\partial x_3}\underline{e}^3$$

Here, ∇f is defined as an open product, and should not be confused with the dot product, $\nabla \cdot \underline{u}$, which means something entirely different. Since we didn't define a coordinate system, the above must be a general result.

Thus, the del operator generally can be written as:

$$\nabla = \bar{\vec{\nabla}} = \underline{e}^n \frac{\partial}{\partial \xi_n}$$

or...

$$\nabla = \bar{\overleftarrow{\nabla}} = \frac{\partial}{\partial \xi_n}\underline{e}^n$$

where $\underline{\mathbf{e}}^n$ are the "contravariant" basis vector which are reciprocal to the natural basis!

It is important to note that ∇ is an operator - not really a vector, and not really a scalar. The operator can change the order of a tensor - i.e. a scalar to a vector. As such, the order of the operation can matter - i.e. whether you pre-multiply or post-multiply. When the operator is applied "from the left", as in the second expression, the differential operator is applied, with the basis vector post-multiplying.

III.3.2 Other derivations/definitions

The actual form of the ∇ operator can be derived many ways; some authors define the operator in terms of the dot product form of the total derivative of a scalar:

$$\frac{d\Phi}{dt} = \nabla\Phi \cdot \frac{d\underline{r}}{dt}$$

where $\underline{r}$ is the position vector of a point in space. For a position vector defined rectangular Cartesian coordinates:

$$\frac{d\Phi}{dt} = \nabla\Phi \cdot \left(\frac{\partial x_1}{\partial t}\underline{e}_1 + \frac{\partial x_2}{\partial t}\underline{e}_2 + \frac{\partial x_3}{\partial t}\underline{e}_3\right) = \frac{\partial \Phi}{\partial x_1}\frac{\partial x_1}{\partial t} + \frac{\partial \Phi}{\partial x_2}\frac{\partial x_2}{\partial t} + \frac{\partial \Phi}{\partial x_3}\frac{\partial x_3}{\partial t}$$

$$(\underline{a}+\underline{b}+\underline{c})\Phi \cdot \left(\frac{\partial x_1}{\partial t}\underline{e}_1 + \frac{\partial x_2}{\partial t}\underline{e}_2 + \frac{\partial x_3}{\partial t}\underline{e}_3\right) == \frac{\partial \Phi}{\partial x_1}\frac{\partial x_1}{\partial t} + \frac{\partial \Phi}{\partial x_2}\frac{\partial x_2}{\partial t} + \frac{\partial \Phi}{\partial x_3}\frac{\partial x_3}{\partial t}$$

So that

$$\nabla = (\underline{a}+\underline{b}+\underline{c}) = \frac{\partial}{\partial x_1}\underline{e}_1 + \frac{\partial}{\partial x_2}\underline{e}_2 + \frac{\partial}{\partial x_3}\underline{e}_3$$

This is the form of the operator for **rectangular Cartesian coordinate system only**. Note that the derivation derives what appears to be a slight difference - the post-multiplication of the basis vectors rather than pre-multiplication. Does it matter? Not for *this* coordinate system.

For the two coordinate systems we've considered in detail, the relevant basis are:

Rectangular Cartesian system: $\underline{e}^i = \underline{e}_i \neq f(x_1, x_2, x_3)$

Cylindrical polar coordinate system:

$$\underline{e}^1 = \underline{e}_R$$
$$\underline{e}^2 = \frac{1}{R}\underline{e}_\theta$$
$$\underline{e}^3 = \underline{e}_z$$

Thus, the ∇ operator for the **cylindrical polar coordinate system** is:

$$\boxed{\nabla = \underline{e}_r \frac{\partial}{\partial r} + \frac{1}{r}\frac{\partial}{\partial \theta}\underline{e}_\theta + \frac{\partial}{\partial z}\underline{e}_z}$$

For the more general case of **orthogonal, curvilinear coordinate systems**, the ∇ operator is given by:

$$\nabla = \frac{1}{\overline{g}_1}\underline{e}_1 \frac{\partial}{\partial \xi_1} + \frac{1}{\overline{g}_2}\underline{e}_2 \frac{\partial}{\partial \xi_2} + \frac{1}{\overline{g}_3}\underline{e}_3 \frac{\partial}{\partial \xi_3}$$

where: $\overline{g}_k = \sqrt{g_{kk}}$ (no sum). This form of the del operator is derived in Budiansky's chapter on general tensors and will be discussed in more detail in subsequent sections.

The **general expression** givn by:

$$\boxed{\nabla = g^{st} \frac{\partial \underline{r}}{\partial x_s} \frac{\partial}{\partial x_t}}$$

III.3.3 Operations with ∇

As mentioned earlier, the direction that ∇ "acts" makes a big difference. When the operator *precedes* a variable, it acts to create a new variable; this is referred to as the **open product**.

Open product with a scalar: $\nabla f = \underline{e}^n \frac{\partial f}{\partial \xi_n} = \frac{\partial f}{\partial \xi_n}\underline{e}^n = \underline{v}$

A VECTOR (1st order tensor)

Open product with a vector: $\nabla \underline{v} = \underline{e}^n \frac{\partial \underline{v}}{\partial \xi_n} = \nabla_n v^m \underline{e}_m \underline{e}^n = V_n^m \underline{e}_m \underline{e}^n$

A TENSOR (2nd order tensor)

If the operator is *after* the variable, it creates a new **operator**, not variable. For example:

$$f\nabla = f\overrightarrow{\nabla} = f\underline{e}^n \frac{\partial}{\partial \xi_n} = f\frac{\partial}{\partial \xi_1}\underline{e}^1 + f\frac{\partial}{\partial \xi_2}\underline{e}^2 + f\frac{\partial}{\partial \xi_3}\underline{e}^3$$

If there is an arrow to point out the direction which the operator acts, then the derivative is applied to the terms and the basis vector trails, as in:

$$f\overleftarrow{\nabla} = \frac{\partial f}{\partial \xi_n}\underline{e}^n = \frac{\partial f}{\partial \xi_1}\underline{e}^1 + \frac{\partial f}{\partial \xi_2}\underline{e}^2 + \frac{\partial f}{\partial \xi_3}\underline{e}^3$$

So, does $\overrightarrow{\nabla}\underline{v} = \underline{v}\overleftarrow{\nabla}$? **NO!** The open product of the del operator with a vector depends on the order:

$$\nabla \underline{v} = V_{mn}\underline{e}^m \underline{e}^n$$
$$\underline{v}\overleftarrow{\nabla} = V_{mn}\underline{e}^n \underline{e}^m$$

The distinction between these terms will become clearer when we discuss 2nd order tensors and dyads - that is, the meaning of $\underline{e}_i \underline{e}_j$. By way of example however, consider the small strain tensor defined as: $\underline{E} = \frac{1}{2}\left(\overrightarrow{\nabla}\underline{u} + \underline{u}\overleftarrow{\nabla}\right)$

∇ can be used with the **dot product** as well; this is referred to as a contraction if it serves to lower the order of a tensor:

Dot product with a scalar: **not defined**

Dot product with a vector: $\nabla \cdot \underline{v} = \underline{e}^n \frac{\partial}{\partial \xi_n} \cdot \underline{v} = \underline{e}^n \cdot \frac{\partial}{\partial \xi_n}\left(v^m \underline{e}_m\right) =$

$$\nabla \cdot \underline{v} = \underline{e}^n \cdot \frac{\partial v^m}{\partial \xi_n}\underline{e}_m + \underline{e}^n \cdot \frac{\partial \underline{e}_m}{\partial \xi_n} v^m$$

But what the hell is $\frac{\partial \underline{e}_m}{\partial \xi_n}$, the derivative of a vector? This is where we get into a jam, unless: *a*) the basis vectors are constant in space, or *b*) we know what covariant derivatives are (Section V.1 and V.2).

Well, we know the coordinate the derivative of basis vectors in the **rectangular Cartesian coordinate system** are zero (these basis vectors are constant), so for that case:

$$\nabla \cdot \underline{v} = \underline{e}^n \cdot \frac{\partial v^m}{\partial \xi_n}\underline{e}_m = \frac{\partial v^m}{\partial \xi_n}\underline{e}^n \cdot \underline{e}_m$$

which is can be reduced to the familiar version by noting that (for a RCCS!):

$$\underline{e}^n \cdot \underline{e}_m = \delta^n_m$$

which leads to:

$$\nabla \cdot \underline{v} = \frac{\partial v^m}{\partial x_m} = \frac{\partial v^1}{\partial x_1} + \frac{\partial v^2}{\partial x_2} + \frac{\partial v^3}{\partial x_3}$$

If vectors are not constant in space (i.e. the derivatives are non-zero), things are a bit more difficult - we'll do it later. The cylindrical polar coordinate system is a special case, and is presented in section III.6.

III.4 Scalar and vector fields and related theorems

III.4.1 Gradient, divergence, curl and Laplacian

This section explores the use of the ∇ operator in mathematics: naturally, this is a precursor to actually using the operator in physical problems, which will be done in the following section.

The **gradient** of a scalar field is defined by the open product of the del operator and the scalar field, i.e.:

$$\text{grad } \Phi \equiv \nabla\Phi$$

Note that this is a vector.

The **divergence** of a vector field (or tensor field) is defined by the dot product of the del operator and the field:

$$\text{div } \underline{u} \equiv \nabla \cdot \underline{u}$$

Note that this will result in a tensor field of lower order; i.e. the divergence of a vector field is a scalar field, the divergence of a 2nd order tensor is a vector field (1st order tensor), etc. For example, the equilibrium equation illustrates that the divergence of the stress tensor must be the body force vector: $\nabla \cdot \underline{\sigma} = \underline{b}$

The **curl** of a vector field is defined by the cross product of the del operator and the field:

$$\text{curl } \underline{u} \equiv \nabla \times \underline{u}$$

Note that the curl of a vector field is itself a vector field. A vector field is called *conservative* if the curl of the vector field is the zero vector.

The **Laplacian** is an operator defined by taking the dot product of the del operator with itself:

$$\Delta \equiv \nabla \cdot \nabla$$

This operator will come in handy in the second part of the course when we consider partial differential equations. What is the Laplacian for spherical coordinates?

III.4.2 Integral theorems

There are three closely-related and fundamental theorems in vector (and tensor) calculus which make use of these definitions. They are applicable to an enormous variety of problems in mechanics (solid, fluid, electromagnetism - pretty much everything. The reason for their widespread use stems from the fact that they allow one to change an integral over an area, surface or volume into an integral over a planar curve, space curve or area, respectively. Thus, they very often reduce the order of complexity of the integral.

Green's Theorem changes the integral of the curl of a vector field over an area into an integral over a ***closed*** planar curve:

$$\int_{\text{Area}} \left(\underline{n} \cdot \nabla \times \underline{p}\right) dA = \oint_{\substack{\text{Planar}\\ \text{Curve}}} \underline{p} \cdot d\underline{r}$$

where $\underline{\mathbf{p}}$ is any vector field, dA is the element area, $\underline{\mathbf{n}}$ is the vector normal to the plane and $\underline{\mathbf{r}}$ is the position vector. This is obviously appropriate to 2-D cases, e.g. plane stress or strain. For a two dimensional problem in rectangular Cartesian coordinates, the component form is:

$$\int_{\text{Area}} \left(\frac{\partial p^1}{\partial x_2} - \frac{\partial p^2}{\partial x_1}\right) dx_1 dx_2 = \oint_{\substack{\text{Planar}\\ \text{Curve}}} p^1 dx_1 + p^2 dx_2$$

Stoke's Theorem changes the integral of the curl of a vector field over a surface into an integral over a ***closed*** space curve:

$$\int_{\text{Surface Area}} \left(\underline{n} \cdot \nabla \times \underline{p}\right) dA = \oint_{\substack{\text{Space}\\ \text{Curve}}} \underline{p} \cdot d\underline{r}$$

where $\underline{\mathbf{n}}$ is the normal to the surface. The vector form of the theorem illustrates that Green's Theorem is closely related to Stoke's theorem; Stoke's theorem applies to a three-dimensional surface area, whereas Green's theorem is for an area contained in a plane. Stoke's theorem would be the appropriate theorem to change the integral over a surface "patch" on an egg to a curve enclosing the patch on the surface of the egg.

The **divergence theorem** (or **Gauss's Theorem**) changes the integral of the divergence of a vector field over a volume into the integral over a surface area:

$$\int_{\text{Volume}} \left(\nabla \cdot \underline{p}\right) dV = \oint_{\text{Surface Area}} \underline{p} \cdot \underline{n}\, dA$$

III.5 Problems in vector mechanics

Space curves:	Parallel force and straight trajectory Acceleration in terms of trajectory trihedral	
Laplacian operator:		Position vector and biharmonic
Curl and Conservative Fields:		Maxwell's Equation Electromagnetic disturbance Planetary motion
Green's Theorem:		Path independence of J-Integral Area of an ellipse
Divergence Theorem:		Equilibrium Symmetry of stress tensor Archimede's Law

III.6 Differentiation revisited: non-constant basis vectors

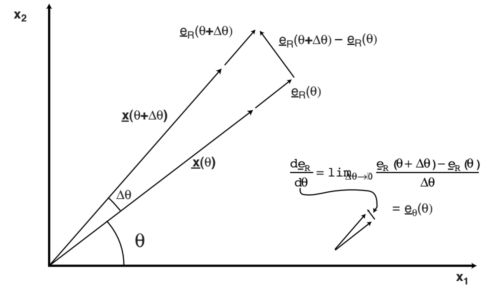
Earlier, we considered the dot product of ∇ and a vector in a space characterized by constant base vectors, but punted on the case with base vectors which varied with position - how do you take the derivative of a vector?

By the letter of the law, we might try this:

$$\frac{\partial \underline{v}}{\partial \xi_k} = \frac{\partial}{\partial \xi_k} v^i \underline{e}_i = \frac{\partial v^i}{\partial \xi_k} \underline{e}_i + v^i \frac{\partial \underline{e}_i}{\partial \xi_k}$$

We still have to find the derivative of a vector - this time the derivative of the basis vectors. The cylindrical polar coordinate system is a useful and popular one, can we get find the derivatives of basis vectors in this space?

This system provides an example which can be done graphically and illustrates "non-constant" basis vectors. A word of warning: using this technique to derive the derivatives of basis vectors in other spaces, even "simple" ones like spherical coordinates can be difficult, if not impossible. True vector differentiation of general vectors requires ***covariant derivatives*****,** which are left to the section on tensors.



Clearly, following the same "graphical" procedure as done in the figure above, we also can find:

$$\frac{\partial \underline{e}_R}{\partial R} = 0$$

and

$$\frac{\partial \underline{e}_\theta}{\partial \theta} = -\underline{e}_R$$

III.7 Summary

- The form of the gradient of a scalar is coordinate system independent (essentially because we do not have to take derivatives of vectors)
- The gradient and curl of a vector are defined in all coordinate systems, but the component form varies greatly; nonetheless, vector equations written in terms of del operator are valid in all coordinate systems.
- Malvern provides an excellent summary of gradient, divergence and curl in terms of *physical components* and basis vectors for typical coordinate systems. E.g. in polar coordinates, the curl is:

$$\nabla \times \underline{v} = \left(\frac{1}{r}\frac{\partial \hat{v}_z}{\partial \theta} - \frac{\partial \hat{v}_z}{\partial \theta}\right)\underline{\hat{e}}_r + \left(\frac{\partial \hat{v}_r}{\partial z} - \frac{\partial \hat{v}_z}{\partial r}\right)\underline{\hat{e}}_\theta + \left(\frac{1}{r}\frac{\partial (r\,\hat{v}_\theta)}{\partial \theta} - \frac{1}{r}\frac{\partial \hat{v}_r}{\partial \theta}\right)\underline{\hat{e}}_z$$

IV. INTRODUCTION TO TENSORS

IV.1 The tensor concept

A tensor is a mathematical object (used to represent a physical concept, e.g. strain) which obeys certain mathematical relationships; the mathematical relationships are easiest illustrated as transformation rules. That is, a tensor represented in one basis will transform to another basis according to standard set of rules.

- Zeroth order tensor = scalar
- First order tensor = vector

$$v^i = \frac{\partial x^i}{\partial \bar{x}^j}\bar{v}^j$$

- 2nd and higher order tensors = tensors

$$T_{kj} = \frac{\partial x_k}{\partial \bar{x}_r}\frac{\partial \bar{x}_s}{\partial x_j}\bar{T}_{rs}$$

Note that a second order tensor has **9 components**. Tensors of order two are generally referred to simply as tensors and the components always have two subscripts. Note that an object with two subscript may or may not be a tensor. There are certain tests which may be done to determine if an object is a tensor - these are referred to as **quotient laws**, which will be touched upon in examples given later in this chapter.

We have already introduced a second order tensor, the **metric tensor**:

$$\underline{G} = g_{ij} = \underline{e}_i \cdot \underline{e}_j$$

In so doing, we saw that the metric tensor "maps" one set of basis functions into another:

$$\underline{G} \cdot \underline{e}^j = \underline{e}_j$$

Thus a second order tensor can be thought of as a linear function which associates a given vector with another vector:

$$\underline{T} \cdot \underline{v} = \underline{u}$$

The stress tensor, considered in detail later, can be shown to "map" the normal of a surface into the traction on that surface:

$$\underline{\sigma} \cdot \underline{n} = \underline{t}$$

This relationship is known as **Cauchy's relation**, and is a staple of applied mechanics. If the second order tensor is thought of a mapping function, and there are multiple representations of a vector, there must be multiple representations of a tensor.

The second order tensor can map:

- Contravariant components of $\underline{v}$ to covariant components of $\underline{u}$:

  $$u_i = T_{ij}v^j$$

- Covariant components of $\underline{v}$ to contravariant components of $\underline{u}$:

  $$u^i = T^{ij}v_j$$

- Covariant (contravariant) components of $\underline{v}$ to covariant (contravariant) components of $\underline{u}$:

  $$u^i = T^i_j v^j \quad u_i = T_i^j v_j$$

Sometimes dots are used to remove ambiguity about which subscript comes first. Hence there are multiple representations of second order tensors:

COVARIANT: T_{ij}
CONTRAVARIANT: T^{ij}
MIXED: $T_i^{\ j}$ and $T^i_{\ j}$

Since tensors map vectors of a given dimension into vectors of the same dimension, they can be represented as square matrices:

$$T_{ij} = \underline{T} = [T]$$

Every second order tensor can be written as a matrix, but is every (3x3) matrix a tensor? NO. The elements of a matrix must transform like elements of tensor.

Consider matrix notation for a tensor - what operation is matrix multiplication? Multiplying tensors in matrix form is the dot product:

$$T^{ij}v_j = u^i$$

$$[T][v] = [u] \rightarrow \underline{T} \cdot \underline{v} = \underline{u}$$

Thus, the matrix forms of tensor equations are valid for ***all*** frames of reference, and imply various relationships between the various types of components of both the tensor and vector.

IV.2 Dyads, dyadics and 2nd order tensors[1]

A **dyad** is a mathematical object denoted by:

$$\underline{u}\ \underline{v}$$

where **u** and **v** are vectors, or 1st order tensors. The meaning of a dyad is simply that:

$$(\underline{u}\ \underline{v})\cdot\underline{y}=\underline{u}(\underline{v}\cdot\underline{y})$$

where **y** is any vector. Thus, the dyad is a second order tensor, which maps **y** into $\underline{u}(\underline{v}\cdot\underline{y})$, which is a vector multiplied by a scalar, i.e. another vector.

A **dyadic** is a sum of dyads, of the form:

$$\underline{T}=\underline{a}\ \underline{b}+\underline{c}\ \underline{d}+\underline{e}\ \underline{f}+\ldots$$

which simply means that:

$$\underline{T}\cdot\underline{v}=\underline{a}\ (\underline{b}\cdot\underline{v})+\underline{c}\ (\underline{d}\cdot\underline{v})+\underline{e}\ (\underline{f}\cdot\underline{v})+\ldots$$

If the vectors **a**, **b** and **c** are expanded about a set of general basis vectors, a second order tensor can be written as the sum of dyads:

$$\underline{T}=\underline{g}\ \underline{h}=\left(g^i\underline{e}_i\right)\left(h^j\underline{e}_j\right)=g^ih^j\ \underline{e}_i\ \underline{e}_j$$

Hence, a tensor can always be written as:

$$\underline{T}=T^{ij}\underline{e}_i\underline{e}_j$$

This should shed some light on the difference between $\bar{\nabla}\underline{u}$ and $\underline{u}\bar{\nabla}$, since:

$$\underline{T}=\bar{\nabla}\underline{u}=T_{ij}\underline{e}^i\underline{e}^j$$

whereas

$$\underline{S}=\underline{u}\bar{\nabla}=S_{ij}\underline{e}^j\underline{e}^i$$

Even though S_{ij} may be the same as T_{ij}, the tensor is not the same, as the coefficients in front of a given dyad may be different.

Obviously, you've dealt with tensors before and the base vectors (or dyads) are implied, as in the stress tensor:

$$\underline{\sigma}=\sigma^{ij}=\sigma^{ij}\underline{e}_i\underline{e}_j$$

---

1 This section is a mixture of Malvern's perspective and Budiansky's perspective.

The basic meaning of **T** is emphasized with the following expressions:

$$\underline{T}\cdot\underline{v}\equiv T^{ij}\underline{e}_i\underline{e}_j\cdot\underline{v}$$

$$\underline{T}\cdot\underline{v}=T^{ij}\underline{e}_i\left(\underline{e}_j\cdot\underline{v}\right)$$

Noting that

$$\underline{e}_j\cdot\underline{v}=\underline{e}_j\cdot v_i\underline{e}^i=v_i\underline{e}_j\cdot\underline{e}^i=v_i\delta_{ij}=v_j$$

we obtain

$$\underline{T}\cdot\underline{v}=\left(T^{ij}v_j\right)\underline{e}_i$$

Thus, if v_j is the jth covariant component of a vector, then $T^{ij}v_j$ is the ith contravariant component of another vector. This operation makes it clear that the dot product operation may not be commutative for the general case - i.e. **the order can matter**:

$$\underline{v}\cdot\underline{T}=\left(T^{ij}v_i\right)\underline{e}_j$$

This is similar, but distinctly different to the opposite order. Matrix notation implies this directly:

$$\underline{T}\cdot\underline{v}=[T][v]\neq[v][T]=\underline{v}\cdot\underline{T}$$

IV.3 Transformations

It has been shown that different basis vectors can be used to create different representations of the same tensor: T_{ij}, T^{ij}, $T^{i}{}_{.j}$, $T_i{}^j$, etc. So, given one representation of a tensor, how do we get another? Not surprisingly, the metric tensor holds the key:

Recall that:

$$\underline{e}_i=g_{ij}\underline{e}^j \text{ and } \underline{e}^i=g^{ij}\underline{e}_j$$

Using the dyadic form of a tensor **T**,

$$\underline{T}=T^{ij}\underline{e}_i\underline{e}_j=T_{pq}\underline{e}^p\underline{e}^q$$

Then

$$T^{ij}\underline{e}_i\underline{e}_j=T^{ij}g_{ip}\underline{e}^p\underline{e}_j$$

$$T^{ij}\underline{e}_i\underline{e}_j=T^{ij}g_{ip}\underline{e}^pg_{jq}\underline{e}^q$$

which implies along with the above:

$$T_{pq} = g_{ip} g_{jq} T^{ij}$$

Obviously, for rectangular Cartesian coordinates, the metric tensor components are given by the Kronecker delta, and there is no difference between contravariant and covariant components of a second order tensor.

Such an operation is often referred to as **lowering (or raising) the indices**. Similar operations result in mixed-component relations:

$$T^{i}_{\cdot j} = g_{jq} T^{iq}$$

$$T_{j}^{\cdot i} = g_{ip} T^{pj}$$

For these operations, the **dummy index is** the one that is **lowered**.

Typically, however, when one speaks of transformations, they do not refer to going from one representation to another (e.g. contravariant components to covariant components); rather, they refer to **transformation from one coordinate system to another**.

These transformations can be defined in terms of transformation coefficients:

$$\bar{T}_{ik} = a_{i}^{\cdot l} a^{k}_{\cdot m} T_{lm}$$

where a_{il} are the transformation coefficients from coordinate i in the barred system to coordinate l in the given system. Here, the a's are the coefficients of the **direct transformation**. The **inverse transformation** and its coefficients are:

$$T_{ik} = b_{i}^{\cdot l} b^{k}_{\cdot m} \bar{T}_{lm}$$

- For rectangular CARTESIAN systems, the transformation coefficients are the directional cosines, defined by

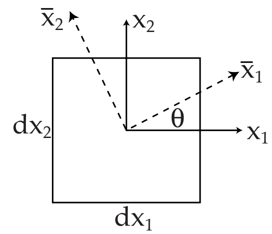
  $$a_{il} = \cos\theta_{i \rightarrow l} = \underline{e}_i \cdot \underline{\bar{e}}_l$$

- For general coordinate systems

$$T_{ij} = \frac{\partial \bar{x}_s}{\partial x_i}\frac{\partial \bar{x}_t}{\partial x_j}\bar{T}_{st} \text{ or } T^{ij} = \frac{\partial x^i}{\partial \bar{x}_r}\frac{\partial x^j}{\partial \bar{x}_t}\bar{T}^{rt}$$

$$\bar{T}^{ij} = \frac{\partial \bar{x}_i}{\partial x_r}\frac{\partial \bar{x}_j}{\partial x_t}T^{rt} \text{ or } \bar{T}_{ij} = \frac{\partial x^i}{\partial \bar{x}_r}\frac{\partial x^j}{\partial \bar{x}_t}T_{rt}$$

Note that a second order tensor transformation has **two** coefficients, rather than just one, as was seen for vector transformations. The stress tensor provides an example which may help to illustrate the transformation:



$$\underline{\sigma} \cdot \underline{n} = \underline{t}$$

$$\sigma_{ij} n_j = t_i$$

$$\bar{\sigma}_{ij}\bar{n}^j = \bar{t}_i = \frac{\partial x_r}{\partial \bar{x}_i} t_r = \frac{\partial x_r}{\partial \bar{x}_i}\sigma_{rp} n^p$$

Of course, without loss of generality, we can assume:

$$n^p = \frac{\partial x_p}{\partial \bar{x}_j}\bar{n}^j$$

So that:

$$\bar{\sigma}_{ij}\bar{n}^j = \frac{\partial x_r}{\partial \bar{x}_i}\frac{\partial x_p}{\partial \bar{x}_j}\sigma_{rp}\bar{n}^j$$

Because $\bar{n}^j$ post-multiplies either side in **exactly** the same manner, it can be dropped and we are left with the transformation rule from above.

Consider that the stress tensor is mapping a normal vector of a surface into the traction vector acting on that surface; then, to change this mapping device from one coordinate system to another requires that both vectors involved must be transformed to the new coordinate system. To put it another way, the concept of a stress has two essential elements: a force and an area. The force which produces the stress must be transformed to the new coordinate system, and the way that area is "measured" must be transformed to the new coordinate system as well. This interpretation implies **two** transformation coefficients in the transformation rule.

Base vectors can also be used to form the transformation law for tensors; the derivation below is taken from Budiansky's clear treatment of the matter (I think):

$$\underline{e}_i = \left(\underline{e}_i \cdot \underline{\bar{e}}^p\right)\underline{\bar{e}}_p$$

where the quantity in parenthesis represents the directional cosines for a Cartesian system. (It may also be interpreted as such for curvilinear but orthogonal systems?) Using this relationship:

$$\underline{T} = T^{ij}\underline{e}_i\underline{e}_j = \bar{T}^{pq}\underline{\bar{e}}_p\underline{\bar{e}}_q$$

$$\underline{T} = T^{ij}\left(\underline{e}_i \cdot \underline{\bar{e}}^p\right)\underline{\bar{e}}_p\left(\underline{e}_j \cdot \underline{\bar{e}}^q\right)\underline{\bar{e}}_q$$

$$\underline{T} = T^{ij}\left(\underline{e}_i \cdot \underline{\bar{e}}^p\right)\left(\underline{e}_j \cdot \underline{\bar{e}}^q\right)\underline{\bar{e}}_p\underline{\bar{e}}_q$$

Therefore: $$\bar{T}^{pq} = T^{ij}\left(\underline{e}_i \cdot \underline{\bar{e}}^p\right)\left(\underline{e}_j \cdot \underline{\bar{e}}^q\right)$$

The following equivalent relations (as well as many others) can be derived in a similar manner:

$$\bar{T}_{pq} = T^{ij}\left(\underline{e}_i \cdot \bar{\underline{e}}_p\right)\left(\underline{e}_j \cdot \bar{\underline{e}}_q\right)$$

$$\bar{T}^{p}_{\cdot q} = T_{ij}\left(\underline{e}^i \cdot \bar{\underline{e}}^p\right)\left(\underline{e}^j \cdot \bar{\underline{e}}_q\right)$$

IV.4 2nd Order tensors: eigenvalues and invariants

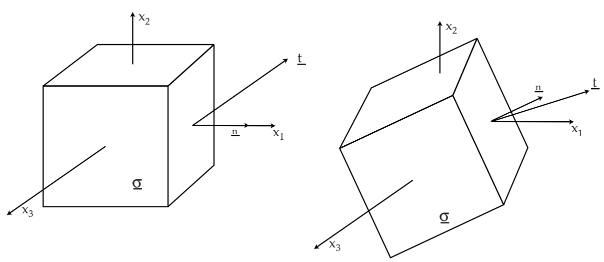
My notes say "Borisenko and Tarapov, pg. 109" at the top, which means I've probably followed their work closely, if not verbatim.

IV.4.1 Tensors and principal axes

Consider the stress tensor for a point in space, represented by a differential element. The relationship between the stress in the element and the tractions acting on a surface with normal **n** is given by Cauchy's relation:

$$\underline{\sigma} \cdot \underline{n} = \underline{t}$$

Thus, the orientation of the surface (or face) governs the tractions acting on that surface - the stress tensor "maps" the surface normal into the surface traction. Obviously, as the stress element is rotated, the direction of the traction vector changes, as the normal of the surface changes.

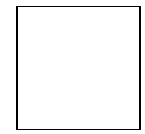


A meaningful (and popular) question is the following: is there an orientation where the traction vector points in the direction of the normal vector? If so, we can redefine a coordinate system which points in this direction, and life is made much simpler. WHY?

The question is then, what is **n** such that:

$$\underline{\sigma} \cdot \underline{n} = \lambda \underline{n}\,?$$

where λ is a scalar multiple. The problem should look very familiar in matrix notation:



The orientation where **t** has this property, i.e. it is aligned with the normal **n**. In this orientation, it is aligned with what are called the **principal axes**, and the tensor components expressed in this orientation are called the **principal values** of the tensor.

Why "principal" values/axes?

$$\sigma_{ij} n_j = \lambda n_i$$

which, for a 2-D problem becomes:

$$\sigma_{11} n_1 + \sigma_{12} n_2 = \lambda n_1$$

Clearly, $\sigma_{12} = 0$! In the coordinate system defined by the principal axes, there are no off-diagonal terms, and the stress tensor (as a mapping device) is considerably simpler.

The details of the 3-D case are presented below.

$$T_{ik} y_k = \lambda y_i$$

$$\left[T_{ik} - \lambda \delta_{ik}\right] y_k = 0$$

For a non-trivial solution, the determinant of the square matrix must be non-zero:

$$\begin{vmatrix} T_{11}-\lambda & T_{12} & T_{13} \\ T_{21} & T_{22}-\lambda & T_{23} \\ T_{31} & T_{32} & T_{33}-\lambda \end{vmatrix} = 0$$

A cubic **characteristic equation** results to solve for λ, which is used in the original question to solve for **y**, the **principal axes**. Note that three eigenvalues λ imply three principal axes, $\underline{y}^j$.

For general coordinates, care must be taken to ensure that the proper components (covariant or contravariant) are accounted for:

$$T_{ik} y^k = \lambda y_i\,.$$

We can't plug and chug in the usual way with this set of equations because it involves two types of components. (You know you have to have covariant components on the right, as the contravariant components on the left lower the index.) Before proceeding, we'll need to recast this expression in terms of only one set of components. Since $y_i = g_{ik} y^k$,

$$\left[T_{ik} - \lambda\delta_{ik}\right]y^k = g^{il}\left[T_{lk} - \lambda g_{ik}\right]y^k = 0$$

$$\left[T^{i}_{\cdot k} - \lambda g^{1}_{\cdot k}\right]y^k = 0.$$

Noting that: $g^{il}g_{ik} = \delta_{lk} = g^{l}_{\cdot k}$, the off-diagonal mutliples of l are zero, and we have the same eigenvalue problem, only with **mixed** components...

$$\begin{vmatrix} T^{1}_{\cdot 1} - \lambda & T^{1}_{\cdot 2} & T^{1}_{\cdot 3} \\ T^{2}_{\cdot 1} & T^{2}_{\cdot 2} - \lambda & T^{2}_{\cdot 3} \\ T^{3}_{\cdot 1} & T^{3}_{\cdot 2} & T^{3}_{\cdot 3} - \lambda \end{vmatrix} = 0$$

- **For general coordinates, we use the *mixed* components to make life easier**
- **For Cartesian systems, no distinction needs to made since:** $T^{ik} = T_{ik} = T^{i}_{k}$
- **If the tensor is symmetric, all λ's are real**
- **Recall that the principal directions (eigenvectors) are $\perp$**

*IV.4.2 The tensor ellipsoid for **rectangular Cartesian coordinate systems***

Consider the mapping:

$$T_{ik}v_k = w_i$$

In a coordinate system aligned with the principal axes:

$$\overline{T}_{ik}\overline{v}_k = \overline{w}_i$$

and off-diagonal terms are zero, so that:

$$\overline{T}_{ik}(i \neq k) = 0$$

$$\overline{w}_1 = \lambda_1\overline{v}_1$$
$$\overline{w}_2 = \lambda_2\overline{v}_2$$
$$\overline{w}_3 = \lambda_3\overline{v}_3$$

- If $\lambda_1 \neq \lambda_2 \neq \lambda_3$, then there are **three unique** principal directions (and hence there is only one orientation (or set of axes) which aligns the "mapped" vector with the original).

- If $\lambda_1 = \lambda_2 \neq \lambda_3$, then there are **two unique** principal directions, and the $\overline{x}_1 \leftrightarrow \overline{x}_2$ plane is a characteristic plane; applying the tensor to any vector $\underline{\mathbf{v}}$ in the $\overline{x}_1 \leftrightarrow \overline{x}_2$ plane changes the length of $\underline{\mathbf{v}}$, but does not rotate it. In other words, any two directions in this plane are principal directions.

- If $\lambda_1 = \lambda_2 = \lambda_3$, then applying $\underline{\mathbf{T}}$ to the vector $\underline{\mathbf{v}}$ maps a change in length but does not change its direction. This type of tensor is referred to as an **isotropic tensor**.

Consider a quadric surface, where the coordinates of any point on the surface satisfy:

$$T_{ik}x_i x_k = 1$$

For a planar ellipse: $T_{11}(x_1)^2 + T_{22}(x_2)^2 = 1$, that is there are no off-diagonal terms. Clearly, expressing any tensor $\underline{\mathbf{S}}$ in the coordinate system aligned with the principal axes, there are no off-diagonal terms - so the tensor in the coordinate system aligned with the principal axes ($\underline{\overline{S}}$ in our notation here) can be illustrated as an ellipsoid![2]

$$\overline{S}_{11}(\overline{x}_1)^2 + \overline{S}_{22}(\overline{x}_2)^2 + \overline{S}_{33}(\overline{x}_3)^2 = 1$$

$$\lambda_1(\overline{x}_1)^2 + \lambda_2(\overline{x}_2)^2 + \lambda_3(\overline{x}_3)^2 = 1$$

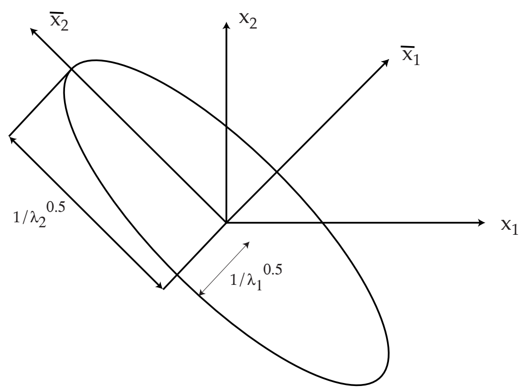
Or,

$$\left(\frac{\overline{x}_1}{1/\sqrt{\lambda_1}}\right)^2 + \left(\frac{\overline{x}_2}{1/\sqrt{\lambda_2}}\right)^2 + \left(\frac{\overline{x}_3}{1/\sqrt{\lambda_3}}\right)^2 = 1$$

For the 2-D case, we have:

$$\underline{\overline{S}} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \text{ where } \lambda_1 \neq \lambda_2$$

[2] Provided all λ's are greater than zero.



Obviously, when $\lambda_1 = \lambda_2$, the tensor is drawn as a circle: changing the orientation has no effect on the components of the tensor.

Borisenko and Tarapov illustrate the 3-D ellipsoid.

It should be emphasized that these transformations do nothing to the tensor itself, merely how that tensor is represented, i.e. the components of the tensor. As an example, finding the principal directions of the stress tensor, where there are no shear stresses, does **not** mean that the material is not experiencing shear - it is not experiencing shear defined in that orientation.

However, if the tensor is isotropic, it does not matter how you orient the coordinate system, there simply is no shear (off-diagonal terms). An example is a sphere loaded with a hydrostatic stress field - a sphere submersed in fluid. No matter how you orient the coordinate system, there is no shear. This is a critical aspect of plasticity - only shear causes plasticity, so the goal is to find the orientation where shear stress is a **maximum**. This is essentially the same problem discussed here.

IV.5 Examples of tensors

IV.5. 1 The moment of inertia tensor

The angular momentum of a group of n particles is given by;

$$\underline{\ell} = \sum_{j=1}^{n} m_j \left(\underline{r}_j \times \underline{v}_j \right)$$

Note that the use of the summation symbol suspends indicial notation for that variable. If the particles do not move with respect to each other or the distance to the origin does not change with time:

$$\underline{v}_j = \underline{\omega} \times \underline{r}_j$$

$$\underline{\ell} = \sum_{j=1}^{n} m_j \left(\underline{r}_j \times \left(\underline{\omega} \times \underline{v}_j \right) \right)$$

Using the identity of the triple cross-product:

$$\underline{\ell} = \sum_{j=1}^{n} m_j \left(\underline{\omega} \left(\underline{r}_j \cdot \underline{r}_j \right) - \underline{r}_j \left(\underline{\omega} \cdot \underline{r}_j \right) \right)$$

Or,

$$\ell_i = \sum_{j=1}^{n} m_j \left(\omega_i x_l^{(j)} x_l^{(j)} - \omega_k x_i^{(j)} x_k^{(j)} \right)$$

This form can be simplified to reveal the tensor character of the moment of inertia, $\underline{\mathbf{I}}$:

$$\ell_i = \omega_k I_{ik}$$

where: $$I_{ik} = \sum_{j=1}^{n} m_j \left(\delta_{ik} x_l^{(j)} x_l^{(j)} - x_i^{(j)} x_k^{(j)} \right)$$

Thus, the moment of inertia tensor maps the angular velocity into the angular momentum. Note that the tensor is symmetric. The diagonal terms represent the moments of inertia about the coordinate axes, where as the off-diagonal terms imply inertial coupling in dynamics (via the "products of inertia").

But how do we know $\underline{\mathbf{I}}$ is a tensor? If it transforms like a tensor, then we know it is a tensor. So, consider a new coordinate system with the same origin (i.e. no translations enters the transformation):

$$\bar{I}_{ik} = \sum_{j=1}^{n} m_j \left(\delta_{ik} \bar{x}_l^{(j)} \bar{x}_l^{(j)} - \bar{x}_i^{(j)} \bar{x}_k^{(j)} \right)$$

From the invariance of the dot product we know $\underline{x} \cdot \underline{x}$ doesn't depend on the coordinate system chosen, i.e.:

$$\bar{x}_l^{(j)} \bar{x}_l^{(j)} = x_l^{(j)} x_l^{(j)}$$

Secondly, we can transform the second term according to the standard transformation rules:

$$\bar{x}_i^{(j)} \bar{x}_k^{(j)} = \left(\alpha_{\bar{i}r} x_r^{(j)} \right) \left(\alpha_{\bar{k}s} x_s^{(j)} \right)$$

Note that (for numerous reasons):

$$\delta_{ik} = \alpha_{\bar{i}r} \alpha_{\bar{k}s} \delta_{rs}$$

Using these two expressions:

$$\bar{I}_{ik} = \sum_{j=1}^{n} m_j \left(\alpha_{\bar{i}r}\alpha_{\bar{k}s}\delta_{rs}\bar{x}_l^{(j)}\bar{x}_l^{(j)} - \alpha_{\bar{i}r}\alpha_{\bar{k}s}\bar{x}_r^{(j)}\bar{x}_s^{(j)}\right)$$

$$\bar{I}_{ik} = \alpha_{\bar{i}r}\alpha_{\bar{k}s} I_{rs}$$

Since $\underline{\mathbf{I}}$ obeys the proper transformation for a second order tensor, it must be one. QED[3].

IV.5.2. The deformation tensor

One of the most widely used tensors in engineering is the deformation tensor, which describes how a infinitessimal line changes length and orientation during an imposed deformation. The deformation tensor is a fundamental quantity which contains a wealth of information, including (*but not limited to*) the state of strain in the material.

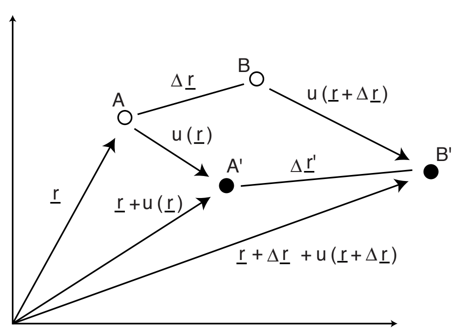
In this example, the generality of the deformation tensor is simplified to the strain tensor, which describes how much an infinitessimal line segment is "stretched." For simplicity's sake, we'll just consider a 2-D rectangular system; in the end, we'll attempt to define the strain tensor in general notation, which would enable easy calculation of the tensor components in a general system. In the following, the prime (e.g. A') denotes the deformed configuration.

$\Delta\underline{r} \equiv$ position of B *relative* to A BEFORE deformation

$\Delta\underline{r}' \equiv$ position of B *relative* to A AFTER deformation

Thus, we are interested in the change in length of $\Delta\underline{r}$ during deformation. Note that the components of $\Delta\underline{r}$ are:

$$\Delta\underline{r} = \Delta x_i \underline{e}_i$$



[3] QED = "quo erat demonstrandum" - "That which has been proven (demonstrated)".

The components of $\Delta\underline{r}'$ are:

$$\Delta\underline{r} = \Delta x'_i \underline{e}_i = \left[\Delta x_i + \frac{\partial u_i}{\partial x_k}\Delta x_k + O\left((\Delta x_k)^2\right)\right]\underline{e}_i$$

In the following, we neglect the higher order term; this does not imply we limit ourselves to small deformation or strain! It merely means we're looking at an infinitessimal line segment. The change in length of $\Delta\underline{r}$ during deformation is thus:

$$|\Delta\underline{r}'|^2 - |\Delta\underline{r}|^2 = \Delta x'_i \Delta x'_i - \Delta x_i \Delta x_i$$

$$|\Delta\underline{r}'|^2 - |\Delta\underline{r}|^2 = \left(\Delta x_i + \frac{\partial u_i}{\partial x_k}\Delta x_k\right)\left(\Delta x_i + \frac{\partial u_i}{\partial x_l}\Delta x_l\right) - \Delta x_i \Delta x_i$$

$$|\Delta\underline{r}'|^2 - |\Delta\underline{r}|^2 = \Delta x_i \Delta x_i + 2\frac{\partial u_i}{\partial x_k}\Delta x_i \Delta x_k + \frac{\partial u_i}{\partial x_k}\frac{\partial u_i}{\partial x_l}\Delta x_k \Delta x_l - \Delta x_i \Delta x_i$$

$$|\Delta\underline{r}'|^2 - |\Delta\underline{r}|^2 = \left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i} + \frac{\partial u_l}{\partial x_i}\frac{\partial u_l}{\partial x_k}\right)\Delta x_i \Delta x_k$$

$$|\Delta\underline{r}'|^2 - |\Delta\underline{r}|^2 = 2\varepsilon_{ik}\Delta x_i \Delta x_k$$

where

$$\varepsilon_{ik} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i} + \frac{\partial u_l}{\partial x_i}\frac{\partial u_l}{\partial x_k}\right)$$

Thus, the strain tensor "maps" the line segment with components Δx_i into the deformed system. For **small strains**, the third term is neglected. The test closely follows that of the previous example. (NOTE TO AUTHOR: you have this in your notes.)

IV.5.3. The metric tensor

The metric tensor has been discussed in previous sections. This section illustrates an alternative test for tensor character. The test states that if

$$g_{ij}x^i y^j = \text{scalar}$$

is true for any two arbitrary vectors, then $\underline{\mathbf{G}} = g^{ij}\underline{\mathbf{e}}_i\underline{\mathbf{e}}_j$ is a tensor.

$$g_{ij} = \underline{e}_i \cdot \underline{e}_j$$

$$\underline{x} = x^i \underline{e}_i = x_j \underline{e}^j$$

$$x^i \underline{e}_i \cdot \underline{e}_j = x_i \underline{e}^i \cdot \underline{e}_j$$

$$x^i g_{ij} = x_i \delta_{ij} \rightarrow x_j = g_{ij} x^i$$

So... $$\underline{x} \cdot \underline{y} = \text{scalar} = x^i y_i = g_{ij} x^i y^j$$

Another way to tell **G** is a tensor - it has the same form for any choice of basis vectors:

$$\underline{G} \equiv \underline{e}^i \underline{e}_i$$

That is, **G** is the collection of dyads formed by the contravariant and covariant basis vectors. If we choose a new set of basis vectors, denoted with bar:

$$\underline{e}_i = \left(\underline{e}_i \cdot \underline{\bar{e}}^p\right) \underline{\bar{e}}_p$$

And substitute, we get:

$$\underline{G} = \underline{e}^i \left(\underline{e}_i \cdot \underline{\bar{e}}^p\right) \underline{\bar{e}}_p = \underline{\bar{e}}^p \underline{\bar{e}}_p$$

which is exactly the same form we started with. Naturally, the components differ from case to case, as we only want to **use only one set** of either the covariant basis vectors or the contravariant basis vectors:

$$\underline{G} = g_{ij} \underline{e}^i \underline{e}^j$$

V. TENSOR CALCULUS

V.1 Differentiation in general coordinate systems: covariant derivatives

V.1.1 Covariant derivatives of vectors (first order tensors)

If you've been paying attention, you may recall how we considered derivatives of vectors previously, but limited ourselves to cases where the derivatives were either zero (RCCS) or we could figure them out without resorting to ugly, yet rigorous mathematics (polar coordinate system). We now consider taking derivatives in general coordinate systems, and this involves **covariant derivatives**. While this topic may seem esoteric (and may be labelled as such by some of your engineering colleagues), it provides the esssential framework for deriving governing equations in useful coordinate systems, **including those in finite element formulations for shells**. (You can run, but you can't hide.)

In the following, we introduce Christoffel symbols; it's easy to get lost in the mathematics and lose the physical meaning of what is going on. In a nutshell, these symbols allow one to calculate the derivatives of base vectors in general coordinate systems. This is critical (and unavoidable) if we want to use the del operator in these coordinate systems; you would need to do this to derive the equilibrium equation for a shell if the coordinate system were defined on the shell surface.

These notes are taken (with extreme gratitute that somebody explains it clearly) from Budiansky's chapter on tensors.

When we introduced derivatives of vectors earlier, we got about this far:

$$\frac{\partial \underline{v}}{\partial \xi_j} = \frac{\partial}{\partial \xi_j}\left(v^i \underline{e}_i\right)$$

$$\frac{\partial \underline{v}}{\partial \xi_j} = \frac{\partial v^i}{\partial \xi_j}\underline{e}_i + v^i \frac{\partial \underline{e}_i}{\partial \xi_j}$$

Now, DEFINE: $\frac{\partial \underline{e}_i}{\partial \xi_j} = \Gamma_{ij}^k \underline{e}_k$

Γ_{ij}^k is the **CHRISTOFFEL SYMBOL OF THE SECOND KIND**. By definition, the Christoffel symbol of the second kind is the kth contravariant component of the derivative with respect to ξ^j of the ith covariant base vector. Using the definition of the natural (or covariant) base vector:

$$\frac{\partial \underline{e}_i}{\partial \xi_j} = \frac{\partial}{\partial \xi_j}\left(\frac{\partial \underline{x}}{\partial \xi_i}\right) = \frac{\partial^2 \underline{x}}{\partial \xi_j \partial \xi_i}$$

So that the Christoffel symbol of the second kind is symmetric with respect to the lower indices (due to the fact that the order of the denominator does matter in the above):

$$\Gamma_{ij}^k = \Gamma_{ji}^k$$

This is particularly useful considering that Γ_{ij}^k represents 27 terms. (Symmetry eliminates several of them.)

We let $v^i{}_{,j}$ represent the contravariant derivative of v^i; that is, it will be the ith covariant component of the the vector which results from taking the derivative of **v**:

$$\frac{\partial \underline{v}}{\partial \xi_j} = v^i{}_{,j}\,\underline{e}_i$$

$$v^i{}_{,j} = \frac{\partial v^i}{\partial \xi_j} + v^k \Gamma_{kj}^i$$

where the second term represents the additional term needed for the covariant derivative of the **contravariant component**.

This now enables us to define the total derivative of an arbitrary vector:

$$d\underline{v} = \frac{\partial \underline{v}}{\partial \xi_j} d\xi_j = \left(v^i{}_{,j}\, d\xi_j\right)\underline{e}_i$$

$$d\underline{v} = \left(\frac{\partial v^i}{\partial \xi_j} + v^k \Gamma_{kj}^i\right) d\xi_j \underline{e}_i$$

$$d\underline{v} = \left(\frac{\partial v^i}{\partial \xi_j} d\xi_j + v^k \Gamma_{kj}^i d\xi_j\right)\underline{e}_i$$

It is important to note that $v^i{}_{,j}$ is a tensor, but $\frac{\partial v^i}{\partial \xi_j}$ is not necessarily a tensor. This is an excellent opportunity to test your understanding of tensors: hint - it has to do with how these quantities transform.

If you were playing close attention, you will note that we just took the covariant derivative of the **contravariant** component. What then, is the covariant derivative of the covariant component, i.e. $v_{i,j}$**?** Well, we can lower the indice using the metric tensor, as in:

$$v_{i,j} = g_{ki} v^k{}_{,j}$$

Or, we can directly calculate $v_{i,j}$ using the same procedure as above:

$$\frac{\partial \underline{v}}{\partial \xi_j} = v_{i,j}\, \underline{e}^i = \frac{\partial v_i}{\partial \xi_j} \underline{e}^i + v_i \frac{\partial \underline{e}^i}{\partial \xi_j}$$

$$\underline{e}_i \cdot \frac{\partial \underline{e}^k}{\partial \xi_j} = -\underline{e}^k \cdot \frac{\partial \underline{e}_i}{\partial \xi_j} = -\underline{e}^k \cdot \left(\Gamma^l_{ij} \underline{e}_l \right) = -\Gamma^k_{ij}$$

So that:

$$v_{i,j} = \frac{\partial v_i}{\partial \xi_j} - v_k \Gamma^k_{ij}$$

where the second term represents the additional term needed for the covariant derivative of the **covariant component.**

So, how in the world do you compute the Christoffel symbols? We start by noting that although Γ^k_{ij} **is** not a 3rd order tensor (since it doesn't transform as one), you can lower the subscript; when you do this, you get **the Christoffel symbol of the first kind**:

$$[ij,p] = g_{kp} \Gamma^k_{ij}$$

Using this, we can write the Christoffel symbols in terms of the metric tensor:

$$[ij,p] = \underline{e}_k \cdot \frac{\partial \underline{e}_i}{\partial \xi_j} = \underline{e}_k \cdot \frac{\partial \underline{x}}{\partial \xi_k} \cdot \frac{\partial^2 \underline{x}}{\partial \xi_i \partial \xi_j}$$

$$\frac{\partial g_{ij}}{\partial \xi_k} = \frac{\partial}{\partial \xi_k}\left(\underline{e}_i \cdot \underline{e}_j \right) = \underline{e}_i \cdot \frac{\partial \underline{e}_j}{\partial \xi_k} + \underline{e}_j \cdot \frac{\partial \underline{e}_i}{\partial \xi_k}$$

$$\frac{\partial g_{ij}}{\partial \xi_k} = [jk,i] + [ik,j]$$

These expressions can be combined to produce:

$$g^{pk} \Gamma^p_{ij} = [ij,k] = \frac{1}{2}\left[\frac{\partial g_{ik}}{\partial \xi_i} + \frac{\partial g_{jk}}{\partial \xi_j} - \frac{\partial g_{ij}}{\partial \xi_k} \right]$$

Thus, the Christoffel symbols really only depend on the metric tensor and its derivatives.

Malvern describes an additional approach to exploring the meaning of this mess. The big picture is that Malvern uses the fact that basis vectors in Cartesian coordinates do not depend on location, i.e. their spatial derivatives are zero. In the following , let x_i be Cartesian coordinates, and let $\underline{i}_j$ be the usual (position-independent) basis vectors in

this coordinate system. Also, let ξ_i and $\underline{e}_i$ be general curvilinear coordinates and (covariant) basis vectors.

$$\underline{x} = x_p \underline{i}_p$$

$$\underline{e}_i = \frac{\partial \underline{x}}{\partial \xi_j} = \frac{\partial x_p}{\partial \xi_j} \underline{i}_p$$

By definition:

$$\underline{i}_p = \frac{\partial \xi_m}{\partial x_p} \underline{e}_m$$

$$\frac{\partial \underline{e}_m}{\partial \xi_n} = \frac{\partial^2 x_p}{\partial \xi_m \partial \xi_n} \underline{i}_p + \frac{\partial x_p}{\partial \xi_j} \frac{\partial \underline{i}_p}{\partial \xi_j}$$

But the basis vectors in the rectangular system do not vary in space, so:

$$\frac{\partial \underline{i}_p}{\partial \xi_j} = 0$$

The result is:

$$\frac{\partial \underline{e}_m}{\partial \xi_n} = \frac{\partial^2 x_p}{\partial \xi_m \partial \xi_n} \underline{i}_p = \frac{\partial^2 x_p}{\partial \xi_m \partial \xi_n} \frac{\partial \xi_s}{\partial x_p} \underline{e}_s$$

or,

$$\frac{\partial \underline{e}_m}{\partial \xi_n} = \Gamma^s_{mn} \underline{e}_s = \begin{Bmatrix} s \\ m \quad n \end{Bmatrix} \underline{e}_s$$

Here then, is an alternative symbol and definition of the Christoffel symbol of the second kind:

$$\Gamma^s_{mn} = \begin{Bmatrix} s \\ m \quad n \end{Bmatrix} = \frac{\partial^2 x_p}{\partial \xi_m \partial \xi_n} \frac{\partial \xi_s}{\partial x_p}$$

This is a useful (perhaps the most useful) form, considering that we often know the connection between coordinte systems: e.g. $x_1 = r \cos \theta$ and $x_2 = r \sin \theta$. The Christoffel symbols of the first kind found via this method are:

$$[mn,r] = g_{rs} \Gamma^s_{mn} = \underline{e}_r \cdot \underline{e}_s \Gamma^s_{mn}$$

$$[mn,r] = \frac{\partial x_q}{\partial \xi_r}\frac{\partial x_k}{\partial \xi_s}\delta_{qk}\Gamma^s_{mn}$$

$$[mn,r] = \frac{\partial x_q}{\partial \xi_r}\frac{\partial x_k}{\partial \xi_s}\delta_{qk}\frac{\partial^2 x_p}{\partial \xi_m \partial \xi_m}\frac{\partial \xi_s}{\partial x_p}$$

$$[mn,r] = \frac{\partial x_q}{\partial \xi_r}\delta_{qk}\frac{\partial^2 x_p}{\partial \xi_m \partial \xi_n}\delta_{kp}$$

And finally,

$$[mn,r] = \frac{\partial^2 x_p}{\partial \xi_m \partial \xi_n}\frac{\partial x_q}{\partial \xi_r}\delta_{pq}$$

So does this imply that the Christoffel symbols depend on a Cartesian coordinate system? NO! The final expression above can be reworked to give the same expression derived earlier, which expresses the Christoffel symbols in terms of the metric tensor and its derivatives.

V.1.2 Covariant derivative of 2nd order tensors

While we are at it, we might as well define the covariant derivative of a second order tensor. Starting with the dyadic form makes things quite similar to our previous derivation:

$$\underline{T} = T^{ij}\underline{e}_i\underline{e}_j$$

$$\frac{\partial \underline{T}}{\partial \xi_k} = \frac{\partial T^{ij}}{\partial \xi_k}\underline{e}_i\underline{e}_j + T^{ij}\frac{\partial \underline{e}_i}{\partial \xi_k}\underline{e}_j + T^{ij}\underline{e}_i\frac{\partial \underline{e}_j}{\partial \xi_k}$$

$$\frac{\partial \underline{T}}{\partial \xi_k} = \frac{\partial T^{ij}}{\partial \xi_k}\underline{e}_i\underline{e}_j + T^{ij}\Gamma^p_{ik}\underline{e}_p\underline{e}_j + T^{ij}\Gamma^q_{jk}\underline{e}_i\underline{e}_q$$

Recall that:

$$\frac{\partial \underline{v}}{\partial \xi_k} = v^i{}_{,k}\,\underline{e}_i$$

Similarly, we denote:

$$\frac{\partial \underline{T}}{\partial \xi_k} = T^{ij}{}_{,k}\,\underline{e}_i\underline{e}_j$$

where:

$$\boxed{T^{ij}{}_{,p} = \frac{\partial T^{ij}}{\partial \xi_p} + T^{rj}\Gamma^i_{rp} + T^{ir}\Gamma^j_{rp}}$$

V.2 Divergence, curl and Laplacian in general coordinate systems

Now that we have defined covariant derivatives, we can explore the general use of the del operator for general coordinate systems. The divergence, curl and Laplacian for general bases vectors and coordinates are derived below. Not much needs to be said, as we are merely applying covariant derivatives and the definition of the del operator.

Divergence:

$$\nabla \cdot \underline{v} \equiv \text{div } \underline{v} = \underline{e}^j\frac{\partial}{\partial \xi_j}\cdot \underline{v}$$

$$= \underline{e}^j\frac{\partial v^p}{\partial \xi_j}\cdot \underline{e}_p + \underline{e}^j v^p \cdot \frac{\partial \underline{e}_p}{\partial \xi_j}$$

$$= \frac{\partial v^p}{\partial \xi_j}\underline{e}^j \cdot \underline{e}_p + v^p\underline{e}^i \cdot \Gamma^k_{pi}\underline{e}_k$$

$$= \frac{\partial v^p}{\partial \xi_j}\delta_{jp} + v^p\Gamma^k_{ip}\underline{e}^i \cdot \underline{e}_k$$

$$\nabla \cdot \underline{v} = \frac{\partial v^i}{\partial \xi_i} + v^p\Gamma^p_{ip}$$

Or, in terms of the metric tensor and its derivatives:

$$\nabla \cdot \underline{v} = \frac{\partial v^i}{\partial \xi_i} + v^p\left(\frac{1}{2}g^{is}\left(\frac{\partial g_{is}}{\partial \xi_p} + \frac{\partial g_{ps}}{\partial \xi_i} - \frac{\partial g_{ip}}{\partial \xi_s}\right)\right)$$

$$\nabla \cdot \underline{v} = \frac{\partial v^i}{\partial \xi_i} + v^p\left(\frac{1}{2}g^{is}\left(\frac{\partial g_{is}}{\partial \xi_p}\right)\right)$$

By determinant theory,

$$g \equiv \det|g_{ij}|$$

$$g^{is}\frac{\partial g^{is}}{\partial \xi_p} = \frac{1}{g}\frac{\partial g}{\partial \xi_j}$$

So that:

$$\nabla \cdot \underline{v} = \frac{\partial v^i}{\partial \xi_i} + v^p\left(\frac{1}{2g}\left(\frac{\partial g}{\partial \xi_p}\right)\right) = \frac{\partial v^i}{\partial \xi_i}\frac{\sqrt{g}}{\sqrt{g}} + \frac{v^p}{\sqrt{g}}\left(\frac{1}{2\sqrt{g}}\left(\frac{\partial g}{\partial \xi_p}\right)\right)$$

$$\nabla \cdot \underline{v} = \frac{1}{\sqrt{g}}\frac{\partial v^j}{\partial \xi_j}\frac{\partial}{\partial \xi_i}\left(v^i\sqrt{g}\right)$$

Curl:

$$\text{curl } \underline{v} = \nabla \times \underline{v} = \underline{e}^j \frac{\partial}{\partial \xi_j} \times \underline{v}$$

$$= \underline{e}^j \times \frac{\partial \underline{v}}{\partial \xi_j}$$

$$= \underline{e}^j \times v^i{}_{,j}\,\underline{e}_i$$

$$= \underline{e}^j \times v_{k,j}\,\underline{e}^k$$

$$= \left(\underline{e}^j \times \underline{e}^k\right) v_{k,j}$$

Note that:

$$\underline{e}^i \times \underline{e}^k = \frac{\underline{e}_j}{\sqrt{g}} \text{ for } i,j,k \text{ being cyclic permutation of } 1,2,3$$

$$\underline{e}^i \times \underline{e}^k = -\frac{\underline{e}_j}{\sqrt{g}} \text{ for } i,j,k \text{ being cyclic permutation of } 2,1,3$$

$= 0$ otherwise

$$\nabla \times \underline{v} = \sum_{i,j,k} \frac{\underline{e}_i}{\sqrt{g}}\left(v_{k,j} - v_{j,k}\right)$$

where :

$$v_{k,j} = \frac{\partial v_k}{\partial \xi_j} - \Gamma^l_{kj} v_l$$

Since $\Gamma^i_{kj} = \Gamma^i_{jk}$, then:

$$v_{i,j} - v_{j,i} = \frac{\partial v_i}{\partial \xi_j} - \frac{\partial v_j}{\partial \xi_i}$$

So the curl of a vector is:

$$\nabla \times \underline{v} = \sum_{i,j,k} \frac{\underline{e}_i}{\sqrt{g}}\left(\frac{\partial v_k}{\partial \xi_j} - \frac{\partial v_j}{\partial \xi_k}\right)$$

Laplacian:

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

Recall that ∇f is a vector, with the covariant components:

$$u_i = \frac{\partial f}{\partial \xi_i}$$

Then the contravariant components are found by raising the index:

$$u^i = g^{ij} u_j = g^{ij}\frac{\partial f}{\partial \xi_j}$$

So the divergence of grad f is:

$$\nabla^2 f = \nabla \cdot \nabla f = \frac{1}{\sqrt{g}}\frac{\partial}{\partial \xi_i}\left(\sqrt{g}\, g^{ij}\frac{\partial f}{\partial \xi_j}\right)$$

V.3 Budiansky's Classical Interlude: orthogonal curvilinear coordinates

Budiansky has a very clear and straightforward presentation which deals with covariant derivatives in curvilinear orthogonal coordinates. The derivation below is taken from his notes:

- ORTHOGONALITY OF COORDINATE SYSTEM IMPLIES THAT:

$$ds^2 = \alpha^2 d\xi^2 + \beta^2 d\eta^2 + \gamma^2 d\varphi^2$$

where the coordinate system is defined by (ξ, η, φ)

- FOR SUCH A SYSTEM, THE METRIC TENSOR IS DIAGONAL: THE COVARIANT COMPONENTS ARE:

$$g_{ij} = \begin{bmatrix} \alpha^2 & 0 & 0 \\ 0 & \beta^2 & 0 \\ 0 & 0 & \gamma^2 \end{bmatrix}$$

The contravariant components are: $g^{11} = \dfrac{1}{\alpha^2}, g^{22} = \dfrac{1}{\beta^2}, g^{33} = \dfrac{1}{\gamma^2}$

- CONSIDER THE DERIVATIVES OF THE PHYSICAL BASE VECTORS:

Base vectors: $\hat{e}_\xi = \dfrac{e_1}{\sqrt{g_{11}}} = \dfrac{e_1}{\alpha};\ \hat{e}_\eta = \dfrac{e_2}{\sqrt{g_{22}}} = \dfrac{e_2}{\beta};\ \hat{e}_\varphi = \dfrac{e_3}{\sqrt{g_{33}}} = \dfrac{e_3}{\gamma}$

$$\frac{\partial \hat{e}_\xi}{\partial \xi} = -\frac{1}{\alpha^2}\frac{\partial \alpha}{\partial \xi} e_1 + \frac{1}{\alpha}\frac{\partial e_1}{\partial \xi}$$

$$\frac{\partial \hat{e}_\xi}{\partial \xi} = -\frac{\partial \alpha}{\partial \xi}\frac{1}{\alpha^2} e_1 + \frac{1}{\alpha}\Gamma_{11}^k e_k$$

$$\frac{\partial \hat{e}_\xi}{\partial \xi} = -\frac{\partial \alpha}{\partial \xi}\frac{1}{\alpha^2} e_1 + \frac{1}{\alpha}\left(\Gamma_{11}^1 e_1 + \Gamma_{11}^2 e_2 + \Gamma_{11}^3 e_3\right)$$

$g_{kp}\Gamma_{ij}^p = [ij, k]$, but since off-diagonal terms in the metric tensor are zero,

$$\Gamma_{11}^k = \frac{[11, k]}{g_{kk}} \text{ (no sum)}$$

$$[11, k] = \frac{1}{2}\left\{\frac{\partial g_{1k}}{\partial \xi_1} + \frac{\partial g_{1k}}{\partial \xi_1} - \frac{\partial g_{11}}{\partial \xi_k}\right\}$$

$$[11,1] = \frac{1}{2}\frac{\partial}{\partial \xi}\left(\alpha^2\right) = \alpha\frac{\partial \alpha}{\partial \xi}$$

$$[11,3] = -\frac{1}{2}\frac{\partial}{\partial \varphi}\left(\alpha^2\right) = -\alpha\frac{\partial \alpha}{\partial \varphi}$$

So:

$$\frac{\partial \hat{e}_\xi}{\partial \xi} = -\frac{\partial \alpha}{\partial \xi}\frac{1}{\alpha^2} e_1 + \frac{1}{\alpha}\left(\frac{1}{\alpha}\frac{\partial \alpha}{\partial \xi} e_1 - \frac{\alpha}{\beta^2}\frac{\partial \alpha}{\partial \eta} e_2 - \frac{\alpha}{\gamma^2}\frac{\partial \alpha}{\partial \varphi} e_3\right)$$

$$= \frac{1}{\beta}\frac{\partial \alpha}{\partial \eta}\hat{e}_\eta - \frac{1}{\gamma}\frac{\partial \alpha}{\partial \varphi}\hat{e}_\varphi$$

And similarly,

$$\frac{\partial \hat{e}_\xi}{\partial \eta} = \frac{1}{\alpha}\frac{\partial \beta}{\partial \xi}\hat{e}_\eta$$

For polar coordinates: $(\xi,\eta,\varphi) \rightarrow (r,\theta,z)$

$$\alpha = 1$$
$$\beta = r \left(g_{22} = r^2\right)$$
$$\gamma = 1$$

$$\frac{\partial \hat{\underline{e}}_r}{\partial \theta} = (1)(1)\hat{\underline{e}}_\theta = \hat{\underline{e}}_\theta!$$

The same procedure yields:

$$\frac{\partial \hat{\underline{e}}_\eta}{\partial \eta} = -\frac{1}{\alpha}\frac{\partial \beta}{\partial \xi}\hat{\underline{e}}_\xi - \frac{1}{\gamma}\frac{\partial \beta}{\partial \varphi}\hat{\underline{e}}_\varphi$$

which implies (of course):

$$\frac{\partial \hat{\underline{e}}_\theta}{\partial \theta} = -\left(\frac{1}{1}\right)(1)\hat{\underline{e}}_r = -\hat{\underline{e}}_r$$

NOTE TO AUTHOR: the complete derivation (with C.S.) of this is in the notes.