UNM SUPPLEMENTAL BOOK DRAFT
June 2004

Curvilinear Analysis in a Euclidean Space

Presented in a framework and notation customized for students and professionals who are already familiar with Cartesian analysis in ordinary 3D physical engineering space.

Rebecca M. Brannon

Written by Rebecca Moss Brannon of Albuquerque NM, USA, in connection with adjunct teaching at the University of New Mexico. This document is the intellectual property of Rebecca Brannon.

June 2004

Table of contents

PREFACE

This document started out as a small set of notes assigned as supplemental reading and exercises for the graduate students taking my continuum mechanics course at the University of New Mexico (Fall of 1999). After the class, I posted the early versions of the manuscript on my UNM web page to solicit interactions from anyone on the internet who cared to comment. Since then, I have regularly fielded questions and added enhancements in response to the encouragement (and friendly hounding) that I received from numerous grad students and professors from all over the world.

Perhaps the most important acknowledgement I could give for this work goes to Prof. Howard "Buck" Schreyer, who introduced me to curvilinear coordinates when I was a student in his Continuum Mechanics class back in 1987. Buck's daughter, Lynn Betthenum (sp?) has also contributed to this work by encouraging her own grad students to review it and send me suggestions and corrections.

Although the vast majority of this work was performed in the context of my University appointment, I must acknowledge the professional colleagues who supported the continued refinement of the work while I have been a staff member and manager at Sandia National Laboratories in New Mexico.

Rebecca Brannon
rmbrann@me.unm.edu
June 2004

Curvilinear Analysis in a Euclidean Space

Rebecca M.Brannon, University of New Mexico, First draft: Fall 1998, Present draft: 6/17/04.

1. Introduction

This manuscript is a student's introduction on the mathematics of curvilinear coordinates, but can also serve as an information resource for practicing scientists. Being an introduction, we have made every effort to keep the analysis well connected to concepts that should be familiar to anyone who has completed a first course in regular-Cartesian-coordinate[1] (RCC) vector and tensor analysis. Our principal goal is to introduce *engineering* specialists to the *mathematician's* language of general curvilinear vector and tensor analysis. Readers who are already well-versed in functional analysis will probably find more rigorous manuscripts (such as [14]) more suitable. If you are completely new to the subject of general curvilinear coordinates or if you seek guidance on the basic machinery associated with non-orthonormal base vectors, then you will probably find the approach taken in this report to be unique and (comparatively) accessible. Many engineering students presume that they can get along in their careers just fine without ever learning any of this stuff. Quite often, that's true. Nonetheless, there will undoubtedly crop up times when a system operates in a skewed or curved coordinate system, and a basic knowledge of curvilinear coordinates makes life a lot easier. Another reason to learn curvilinear coordinates — even if you never explicitly apply the knowledge to any practical problems — is that you will develop a far deeper understanding of Cartesian tensor analysis.

Learning the basics of curvilinear analysis is an essential first step to reading much of the older materials modeling literature, and the theory is still needed today for non-Euclidean surface and quantum mechanics problems. We added the proviso "older" to the materials modeling literature reference because more modern analyses are typically presented using *structured notation* (also known as Gibbs, symbolic, or direct notation) in which the highest-level fundamental meaning of various operations are called out by using a notation that does *not* explicitly suggest the procedure for actually performing the operation. For example, $\underset{\sim}{\mathbf{a}} \bullet \underset{\sim}{\mathbf{b}}$ would be the structure notation for the vector dot product whereas $a_1 b_1 + a_2 b_2 + a_3 b_3$ would be the procedural notation that clearly shows how to compute the dot product but has the disad-

1. Here, "regular" means that the basis is right, rectangular, and normalized. "Right" means the basis forms a right-handed system (i.e., crossing the first base vector into the second results in a third vector that has a positive dot product with the third base vectors). "Rectangular" means that the base vectors are mutually perpendicular. "Normalized" means that the base vectors are dimensionless and of unit length. "Cartesian" means that all three coordinates have the same physical units [12, p90]. The last "C" in the RCC abbreviation stands for "coordinate" and its presence implies that the basis is itself defined in a manner that is coupled to the coordinates. Specifically, the basis is always tangent to the coordinate grid. A goal of this paper is to explore the implications of removing the constraints of RCC systems. What happens when the basis is not rectangular? What happens when coordinates of different dimensions are used? What happens when the basis is selected independently from the coordinates?

vantage of being applicable only for RCC systems. The *same* operation would be computed differently in a non-RCC system — the fundamental operation itself doesn't change; instead the method for computing it changes depending on the system you adopt. Operations such as the dot and cross products are known to be invariant when expressed using combined component+basis notation. Anyone who *chooses* to perform such operations using Cartesian components will obtain the same result as anyone else who opts to use general curvilinear components — provided that both researchers understand the connections between the two approaches! Every now and then, the geometry of a problem clearly calls for the use of non-orthonormal or spatially varying base vectors. Knowing the basics of curvilinear coordinates permits analysts to choose the approach that most simplifies their calculations. This manuscript should be regarded as providing two services: (1) enabling students of Cartesian analysis to solidify their knowledge by taking a foray into curvilinear analysis and (2) enabling engineering professionals to read older literature wherein it was (at the time) considered more "rigorous" or stylish to present all analyses in terms of general curvilinear analysis.

In the field of materials modeling, the stress tensor is regarded as a function of the strain tensor and other material state variables. In such analyses, the material often contains certain "preferred directions" such as the direction of fibers in a composite matrix, and curvilinear analysis becomes useful if those directions are not orthogonal. For plasticity modeling, the machinery of non-orthonormal base vectors can be useful to understand six-dimensional stress space, and it is especially useful when analyzing the response of a material when the stress resides at a so-called yield surface "vertex". Such a vertex is defined by the convergence of two or more surfaces having different and generally non-orthogonal orientation normals, and determination of whether or not a trial elastic stress rate is progressing into the "cone of limiting normals" becomes quite straightforward using the formal mathematics of non-orthonormal bases.

This manuscript is broken into three key parts: Syntax, Algebra, and Calculus. Chapter 2 introduces the most common coordinate systems and iterates the distinction between irregular bases and curvilinear coordinates; that chapter introduces the several fundamental quantities (such as metrics) which appear with irresistible frequency throughout the literature of generalized tensor analysis. Chapter 3 shows how Cartesian formulas for basic vector and tensor operations must be altered for non-Cartesian systems. Chapter 4 covers basis and coordinate transformations, and it provides a gentle introduction to the fact that base vectors can vary with position.

The fact that the underlying base vectors might be non-normalized, non-orthogonal, and/or non-right-handed is the essential focus of Chapter 4. By contrast, Chapter 5 focuses on how extra terms must appear in gradient expressions (in addition to the familiar terms resulting from spatial variation of scalar and vector components); these extra terms account for the fact that the coordinate base vectors vary in space. The fact that different base vectors can be used at different points in space is an essential feature of curvilinear coordinates analysis.

The vast majority of engineering applications use one of the coordinate systems illustrated in Fig. 1.1. Of these, the rectangular Cartesian coordinate system is the most popular choice. For all three systems in Fig. 1.1, the base vectors are unit vectors. The base vectors are also mutually perpendicular, and the ordering is "right-handed" (*i.e.,* the third base vector is obtained by crossing the first into the second). Each base vector points in the direction that the position vector moves if one coordinate is increased, holding the other two coordinates constant; thus, base vectors for spherical and cylindrical coordinates vary with position. This is a crucial concept: although the coordinate system has only one origin, there can be an infinite number base vectors because the base vector orientations can depend on position.

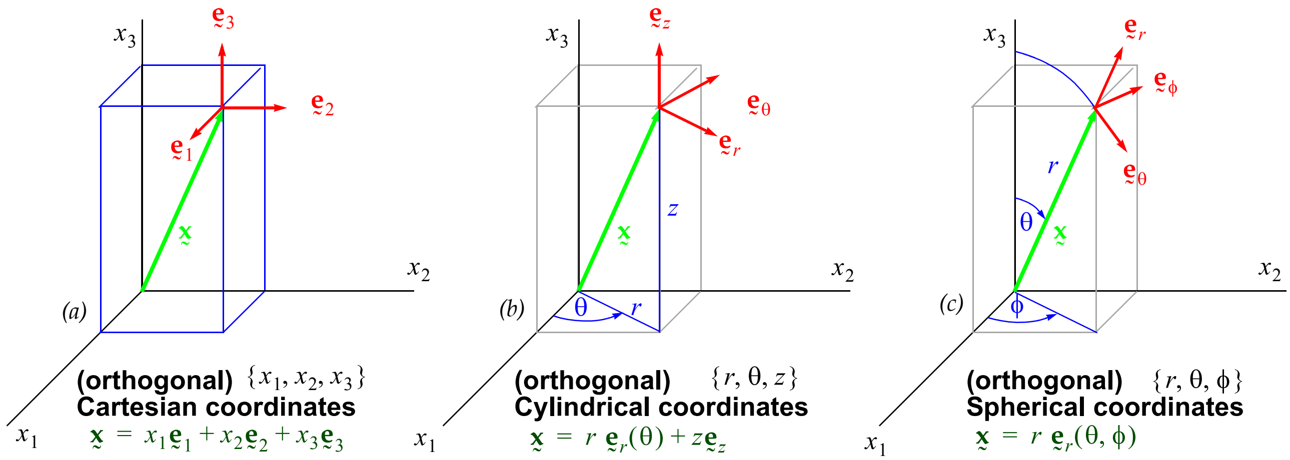


FIGURE 1.1 The most common engineering coordinate systems. Note that all three systems are orthogonal because the associated base vectors are mutually perpendicular. The cylindrical and spherical coordinate systems are inhomogeneous because the base vectors vary with position. As indicated, $\underset{\sim}{e}_r$ depends on θ for cylindrical coordinates and $\underset{\sim}{e}_r$ depends on both θ and ψ for spherical coordinates.

Most practicing engineers can get along just fine without ever having to learn the theory behind general curvilinear coordinates. Naturally, every engineer must, at some point, deal with cylindrical and spherical coordinates, but they can look up whatever formulas they need in handbook tables. So why bother learning about generalized curvilinear coordinates? Different people have different motivations for studying general curvilinear analysis. Those dealing with general relativity, for example, must be able to perform tensor analysis on four dimensional curvilinear manifolds. Likewise, engineers who analyze shells and membranes in 3D space greatly benefit from general tensor analysis. Reading the literature of continuum mechanics — especially the older work — demands an understanding of the notation. Finally, the topic is just plain interesting in its own right. James Simmonds [7] begins his book on tensor analysis with the following wonderful quote:

> *The magic of this theory will hardly fail to impose itself on anybody who has truly understood it; it represents a genuine triumph of the method of absolute differential calculus, founded by Gauss, Riemann, Ricci, and Levi-Civita.*
>
> —Albert Einstein[1]

An important message articulated in this quote is the suggestion that, once you have mastered tensor analysis, you will begin to recognize its basic concepts in many other seemingly unrelated fields of study. Your knowledge of tensors will therefore help you master a broader range of subjects.

1. From: “Contribution to the Theory of General Relativity,” 1915; as quoted and translated by C. Lanczos in The Einstein Decade, p213.

This document is a teaching and learning tool. To assist with this goal, you will note that the text is color-coded as follows:

BLUE $\Rightarrow$ definition
RED $\Rightarrow$ important concept

Please direct comments to rmbrann@me.unm.edu

1.1 Vector and Tensor Notation

The tensorial order of quantities will be indicated by the number of underlines. For example, s is a scalar, $\underset{\sim}{\mathbf{v}}$ is a vector, $\underset{\approx}{\mathbf{T}}$ is a second-order tensor, $\underset{\approx\!\!\sim}{\xi}$ is a third order tensor, *etc.* We follow Einstein's summation convention where repeated indices are to be summed (this rule will be later clarified for curvilinear coordinates).

You, the reader, are presumed familiar with basic operations in Cartesian coordinates (dot product, cross-product, determinant, etc.). Therefore, we may define our *structured* terminology and notational conventions by telling you their meanings in terms ordinary Cartesian coordinates. A principal purpose of this document is to show how these *same* structured operations must be computed using *different procedures* when using non-RCC systems. In this section, where we are merely explaining the meanings of the non-indicial notation structures, we will use standard RCC conventions that components of vectors and tensors are identified by subscripts that take on the values 1, 2, and 3. Furthermore, when exactly two indices are repeated in a single term, they are understood to be summed from 1 to 3. Later on, for non-RCC systems, the conventions for subscripts will be generalized.

The term "array" is often used for any matrix having one dimension equal to 1. This document focuses exclusively on ordinary 3D physical space. Thus, unless otherwise indicated, the word "array" denotes either a 3×1 or a 1×3 matrix. Any array of three numbers may be expanded as the sum of the array components times the corresponding primitive *basis* arrays:

$$\begin{Bmatrix} v_1 \\ v_2 \\ v_3 \end{Bmatrix} = v_1 \begin{Bmatrix} 1 \\ 0 \\ 0 \end{Bmatrix} + v_2 \begin{Bmatrix} 0 \\ 1 \\ 0 \end{Bmatrix} + v_3 \begin{Bmatrix} 0 \\ 0 \\ 1 \end{Bmatrix} \tag{1.1}$$

Everyday engineering problems typically characterize vectors using only the regular Cartesian (orthonormal right-handed) laboratory basis, $\{\underset{\sim}{\mathbf{e}}_1, \underset{\sim}{\mathbf{e}}_2, \underset{\sim}{\mathbf{e}}_3\}$. Being orthonormal, the base vectors have the property that $\underset{\sim}{\mathbf{e}}_i \bullet \underset{\sim}{\mathbf{e}}_j = \delta_{ij}$, where δ_{ij} is the Kronecker delta and the indices (i and j) take values from 1 to 3. Equation (1.1) is the matrix-notation equivalent of the usual expansion of a vector as a sum of components times base vectors:

$$\underset{\sim}{\mathbf{v}} = v_1 \underset{\sim}{\mathbf{e}}_1 + v_2 \underset{\sim}{\mathbf{e}}_2 + v_3 \underset{\sim}{\mathbf{e}}_3 \tag{1.2}$$

More compactly,

$$\underset{\sim}{\mathbf{v}} = v_k \underset{\sim}{\mathbf{e}}_k \tag{1.3}$$

Many engineering problems (e.g., those with spherical or cylindrical symmetry) become extremely complicated when described using the orthonormal laboratory basis, but they simplify superbly when phrased in terms of some other basis, $\{\underset{\sim}{\mathbf{E}}_1, \underset{\sim}{\mathbf{E}}_2, \underset{\sim}{\mathbf{E}}_3\}$. Most of the time, this "other" basis is also a "regular" basis, meaning that it is orthonormal ($\underset{\sim}{\mathbf{E}}_i \bullet \underset{\sim}{\mathbf{E}}_j = \delta_{ij}$) and right handed ($\underset{\sim}{\mathbf{E}}_3 = \underset{\sim}{\mathbf{E}}_1 \times \underset{\sim}{\mathbf{E}}_2$) — the only difference is that the new basis is oriented differently than the laboratory basis. The best choice for this other, more convenient, basis might vary in space. Note, for example, that the bases for spherical and cylindrical coordinates illustrated in Fig. 1.1 are orthonormal and right-handed (and therefore "regular") even though a different set of base vectors is used at each point in space. This harks back to our earlier comment that the properties of being orthonormal and curvilinear are distinct — one does not imply or exclude the other.

Generalized curvilinear coordinates show up when studying quantum mechanics or shell theory (or even when interpreting a material deformation from the perspective of a person who translates, rotates, and stretches along with the material). For most of these advanced physics problems, the governing equations are greatly simplified when expressed in terms of an "irregular" basis (i.e., one that is not orthogonal, not normalized, and/or not right-handed). To effectively study curvilinear coordinates and irregular bases, the reader must practice constant vigilance to keep track of what *particular basis* a set of components is referenced to. When working with irregular bases, it is customary to construct a complementary or "dual" basis that is intimately related to the original irregular basis. Additionally, even though it is might not be convenient for the application at hand, the regular laboratory basis still exists. Sometimes a quantity is most easily interpreted using yet other bases. For example, a tensor is usually described in terms of the laboratory basis or some "applications" basis, but we all know that the tensor is particularly simplified if it is expressed in terms of its *principal* basis. Thus, any engineering problem might involve the simultaneous use of many different bases. If the basis is changed, then the components of vectors and tensors must change too. To emphasize the inextricable interdependence of components and bases, vectors are routinely expanded in the form of components times base vectors $\underset{\sim}{\mathbf{v}} = v_1 \underset{\sim}{\mathbf{e}}_1 + v_2 \underset{\sim}{\mathbf{e}}_2 + v_3 \underset{\sim}{\mathbf{e}}_3$.

What is it that distinguishes vectors from simple 3×1 arrays of numbers? The answer is that the component array for a vector is determined by the underlying basis and this component array must change is a very particular manner when the basis is changed. Vectors have (by definition) an invariant quality with respect to a change of basis. Even though the components themselves change when a basis changes, they must change in a very specific way — it they don't change that way, then the thing you are dealing with (whatever it may be) is not a vector. Even though components change when the basis changes, the *sum* of the components times the base vectors remains the same. Suppose, for example, that $\{\underset{\sim}{\mathbf{e}}_1, \underset{\sim}{\mathbf{e}}_2, \underset{\sim}{\mathbf{e}}_3\}$ is the regular laboratory basis and $\{\underset{\sim}{\mathbf{E}}_1, \underset{\sim}{\mathbf{E}}_2, \underset{\sim}{\mathbf{E}}_3\}$ is some alternative orthonormal right-handed basis. Let the components of a vector $\underset{\sim}{\mathbf{v}}$ with respect to the lab basis be denoted by v_k and let $\tilde{v}_k$ denote the components with respect to the second basis. The invariance of vectors requires that the basis expansion of the vector must give the same result regardless of which basis is used. Namely,

$$v_k \underset{\sim}{\mathbf{e}}_k = \tilde{v}_k \underset{\sim}{\mathbf{E}}_k \tag{1.4}$$

The relationship between the $\tilde{v}_k$ components and the v_k components can be easily characterized as follows:

$$\text{Dotting both sides of (1.4) by } \underset{\sim}{\mathbf{E}}_m \text{ gives: } v_k(\underset{\sim}{\mathbf{e}}_k \bullet \underset{\sim}{\mathbf{E}}_m) = \tilde{v}_k(\underset{\sim}{\mathbf{E}}_k \bullet \underset{\sim}{\mathbf{E}}_m) \tag{1.5a}$$

$$\text{Dotting both sides of (1.4) by } \underset{\sim}{\mathbf{e}}_m \text{ gives: } v_k(\underset{\sim}{\mathbf{e}}_k \bullet \underset{\sim}{\mathbf{e}}_m) = \tilde{v}_k(\underset{\sim}{\mathbf{E}}_k \bullet \underset{\sim}{\mathbf{e}}_m) \tag{1.5b}$$

Note that $\tilde{v}_k(\underset{\sim}{\mathbf{E}}_k \bullet \underset{\sim}{\mathbf{E}}_m) = \tilde{v}_k \delta_{km} = \tilde{v}_m$. Similarly, $v_k(\underset{\sim}{\mathbf{e}}_k \bullet \underset{\sim}{\mathbf{e}}_m) = v_k \delta_{km} = v_m$. We can define a set of nine numbers (known as direction cosines) $L_{ij} \equiv \underset{\sim}{\mathbf{e}}_i \bullet \underset{\sim}{\mathbf{E}}_j$. Therefore, $\underset{\sim}{\mathbf{e}}_k \bullet \underset{\sim}{\mathbf{E}}_m = L_{km}$ and $\underset{\sim}{\mathbf{E}}_k \bullet \underset{\sim}{\mathbf{e}}_m = L_{mk}$, where we have used the fact that the dot product is commutative ($\underset{\sim}{\mathbf{a}} \bullet \underset{\sim}{\mathbf{b}} = \underset{\sim}{\mathbf{b}} \bullet \underset{\sim}{\mathbf{a}}$ for any vectors $\underset{\sim}{\mathbf{a}}$ and $\underset{\sim}{\mathbf{b}}$). With these observations and definitions, Eq. (1.5) becomes

$$v_k L_{km} = \tilde{v}_m \tag{1.6a}$$

$$v_m = \tilde{v}_k L_{mk} \tag{1.6b}$$

These relationships show how the $\{v_k\}$ components are related to the $\tilde{v}_m$ components. Satisfying these relationships is often the identifying characteristic used to identify whether or not something really is a vector (as opposed to a simple collection of three numbers). This discussion was limited to changing from one regular (i.e., orthonormal right-handed) to another. Later on, the concepts will be revisited to derive the change of component formulas that apply to irregular bases. The key point (which is exploited throughout the remainder of this document) is that, although the vector components themselves change with the basis, the *sum* of components times base vectors in invariant.

The statements made above about vectors also have generalizations to tensors. For example, the analog of Eq. (1.1) is the expansion of a 3×3 matrix into a sum of individual components times base tensors:

$$\begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} = A_{11}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + A_{12}\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \ldots + A_{33}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1.7}$$

Looking at a tensor in this way helps clarify why tensors are often treated as nine-dimensional vectors: there are nine components and nine associated "base tensors." Just as the intimate relationship between a vector and its components is emphasized by writing the vector in the form of Eq. (1.4), the relationship between a tensor's components and the underlying basis is emphasized by writing tensors as the sum of components times "basis dyads". Specifically in direct correspondence to Eq. (1.7) we write

$$\underset{\approx}{\mathbf{A}} = A_{11}\underset{\sim}{\mathbf{e}}_1\underset{\sim}{\mathbf{e}}_1 + A_{12}\underset{\sim}{\mathbf{e}}_1\underset{\sim}{\mathbf{e}}_2 + \ldots + A_{33}\underset{\sim}{\mathbf{e}}_3\underset{\sim}{\mathbf{e}}_3 . \tag{1.8}$$

Two vectors written side-by-side are to be multiplied **dyadically**, for example, $\underset{\sim}{\mathbf{a}}\underset{\sim}{\mathbf{b}}$ is a second-order tensor **dyad** with Cartesian *ij* components $a_i b_j$. Any tensor can be expressed as a linear combination of the nine possible **basis dyads**. Specifically the dyad $\underset{\sim}{\mathbf{e}}_i\underset{\sim}{\mathbf{e}}_j$ corresponds to an RCC component matrix that has zeros everywhere except "1" in the ij position, which was what enabled us to write Eq. (1.7) in the more compact form of Eq. (1.8). Even a dyad $\underset{\sim}{\mathbf{a}}\underset{\sim}{\mathbf{b}}$ itself can be expanded in terms of basis dyads as $\underset{\sim}{\mathbf{a}}\underset{\sim}{\mathbf{b}} = a_i b_i \underset{\sim}{\mathbf{e}}_i\underset{\sim}{\mathbf{e}}_j$. Dyadic multiplication is often alternatively denoted with the symbol ⊗. For example, $\underset{\sim}{\mathbf{a}} \otimes \underset{\sim}{\mathbf{b}}$ means the same thing as $\underset{\sim}{\mathbf{a}}\underset{\sim}{\mathbf{b}}$. We prefer using no "⊗" symbol for dyadic multiplication because it allows more appealing identities such as $(\underset{\sim}{\mathbf{a}}\underset{\sim}{\mathbf{b}}) \bullet \underset{\sim}{\mathbf{c}} = \underset{\sim}{\mathbf{a}}(\underset{\sim}{\mathbf{b}} \bullet \underset{\sim}{\mathbf{c}})$.

Working with third-order tensors requires introduction of **triads**, which are denoted structurally by three vectors written side-by-side. Specifically, $\underset{\sim}{\mathbf{u}}\underset{\sim}{\mathbf{v}}\underset{\sim}{\mathbf{w}}$ is a third-order tensor with RCC *ijk* components $u_i v_j w_k$. Any third-order tensor can always be expressed as a linear combination of the fundamental basis triads. The concept of dyadic multiplication extends similarly to fourth and higher-order tensors.

Using our summation notation that repeated indices are to be summed, the standard component-basis expression for a tensor (Eq. 1.8) can be written

$$\underset{\approx}{\mathbf{A}} = A_{ij}\underset{\sim}{\mathbf{e}}_i\underset{\sim}{\mathbf{e}}_j . \tag{1.9}$$

The components of a tensor change when the basis changes, but the sum of components times basis dyads remains invariant. Even though a tensor comprises many components and basis triads, it is this *sum* of individual parts that's unique and physically meaningful.

A **raised single dot** is the **first-order inner product**. For example, in terms of a Cartesian basis, $\underset{\sim}{\mathbf{u}} \bullet \underset{\sim}{\mathbf{v}} = u_k v_k$. When applied between tensors of higher or mixed orders, the single dot continues to denote the first order inner product; that is, adjacent vectors in the basis dyads are dotted together so that $\underset{\approx}{\mathbf{A}} \bullet \underset{\approx}{\mathbf{B}} = (A_{ij}\underset{\sim}{\mathbf{e}}_i\underset{\sim}{\mathbf{e}}_j) \bullet (B_{pq}\underset{\sim}{\mathbf{e}}_p\underset{\sim}{\mathbf{e}}_q) = (A_{ij}B_{pq}\delta_{jp}\underset{\sim}{\mathbf{e}}_i\underset{\sim}{\mathbf{e}}_q) = A_{ij}B_{jq}\underset{\sim}{\mathbf{e}}_i\underset{\sim}{\mathbf{e}}_q$. Here δ_{ij} is the Kronecker delta, defined to equal 1 if $i=j$ and 0 otherwise. The common operation, $\underset{\approx}{\mathbf{A}} \bullet \underset{\sim}{\mathbf{u}}$ denotes a first order vector whose i^{th} Cartesian component is $A_{ij}u_j$. If, for example, $\underset{\sim}{\mathbf{w}} = \underset{\approx}{\mathbf{A}} \bullet \underset{\sim}{\mathbf{u}}$, then the RCC components of $\underset{\sim}{\mathbf{w}}$ may be found by the matrix multiplication:

$$\underset{\sim}{\mathbf{w}} = \underset{\approx}{\mathbf{A}} \bullet \underset{\sim}{\mathbf{u}} \text{ implies (for RCC) } w_i = A_{ij}u_j \text{, or } \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \tag{1.10}$$

Similarly, $\underset{\sim}{\mathbf{v}} \bullet \underset{\approx}{\mathbf{B}} = v_k B_{km}\underset{\sim}{\mathbf{e}}_m = \underset{\approx}{\mathbf{B}}^{\mathrm{T}} \bullet \underset{\sim}{\mathbf{v}}$, where the superscript "T" denotes the tensor **transpose** (*i.e.*, $B_{ij}^T = B_{ji}$). Note that the effect of the raised single dot is to sum adjacent indices. Applying similar heuristic notational interpretation, the reader can verify that $\underset{\sim}{\mathbf{u}} \bullet \underset{\approx}{\mathbf{C}} \bullet \underset{\sim}{\mathbf{v}}$ must be a scalar computed in RCC by $u_i C_{ij} v_j$.

A first course in tensor analysis, for example, teaches that the **cross-product** between two vectors $\underset{\sim}{\mathbf{v}} \times \underset{\sim}{\mathbf{w}}$ is a new vector obtained in RCC by

$$\underset{\sim}{\mathbf{v}} \times \underset{\sim}{\mathbf{w}} = (v_2 w_3 - v_3 w_2)\underset{\sim}{\mathbf{e}}_1 + (v_3 w_1 - v_1 w_3)\underset{\sim}{\mathbf{e}}_2 + (v_1 w_2 - v_2 w_1)\underset{\sim}{\mathbf{e}}_3 \tag{1.11}$$

or, more compactly,

$$\underset{\sim}{\mathbf{v}} \times \underset{\sim}{\mathbf{w}} = \varepsilon_{ijk} v_j w_k \underset{\sim}{\mathbf{e}}_i, \tag{1.12}$$

where ε_{ijk} is the permutation symbol

$\varepsilon_{ijk} = 1$ if $ijk = 123, 231$, or 312
$\varepsilon_{ijk} = -1$ if $ijk = 321, 132$, or 213
$\varepsilon_{ijk} = 0$ if any of the indices i, j, or k are equal. (1.13)

Importantly,

$$\underset{\sim}{\mathbf{e}}_i \times \underset{\sim}{\mathbf{e}}_j = \varepsilon_{ijk} \underset{\sim}{\mathbf{e}}_k \tag{1.14}$$

This permits us to alternatively write Eq. (1.12) as

$$\underset{\sim}{\mathbf{v}} \times \underset{\sim}{\mathbf{w}} = v_j w_k (\underset{\sim}{\mathbf{e}}_j \times \underset{\sim}{\mathbf{e}}_k) \tag{1.15}$$

We will employ a self-defining notational structure for *all* conventional vector operations. For example, the expression $\underset{\approx}{\mathbf{C}} \times \underset{\sim}{\mathbf{v}}$ can be immediately inferred to mean

$$\underset{\approx}{\mathbf{C}} \times \underset{\sim}{\mathbf{v}} = C_{ij} \underset{\sim}{\mathbf{e}}_i \underset{\sim}{\mathbf{e}}_i \times v_k \underset{\sim}{\mathbf{e}}_k = C_{ij} v_k \underset{\sim}{\mathbf{e}}_i \underset{\sim}{\mathbf{e}}_i \times \underset{\sim}{\mathbf{e}}_k = C_{ij} v_k \varepsilon_{mik} \underset{\sim}{\mathbf{e}}_m \tag{1.16}$$

The "**triple scalar-valued product**" is denoted with square brackets around a list of three vectors and is defined $[\underset{\sim}{\mathbf{u}}, \underset{\sim}{\mathbf{v}}, \underset{\sim}{\mathbf{w}}] \equiv \underset{\sim}{\mathbf{u}} \bullet (\underset{\sim}{\mathbf{v}} \times \underset{\sim}{\mathbf{w}})$. Note that

$$\varepsilon_{ijk} = [\underset{\sim}{\mathbf{e}}_i, \underset{\sim}{\mathbf{e}}_j, \underset{\sim}{\mathbf{e}}_k] \tag{1.17}$$

We denote the **second-order inner product** by a **"double dot" colon**. For rectangular Cartesian components, the second-order inner product sums adjacent *pairs* of components. For example, $\underset{\approx}{\mathbf{A}} : \underset{\approx}{\mathbf{B}} = A_{ij} B_{ij}$, $\underset{\approx}{\xi} : \underset{\approx}{\mathbf{C}} = \xi_{ijk} C_{jk} \underset{\sim}{\mathbf{e}}_i$, and $\underset{\approx}{\xi} : \underset{\sim}{\mathbf{a}} \underset{\sim}{\mathbf{b}} = \xi_{ijk} a_j b_k \underset{\sim}{\mathbf{e}}_i$. Caution: many authors insidiously use the term "inner product" for the similar looking scalar-valued operation $A_{ij} B_{ji}$, but this operation is not an inner product because it fails the positivity axiom required for any inner product.

1.2 Homogeneous coordinates

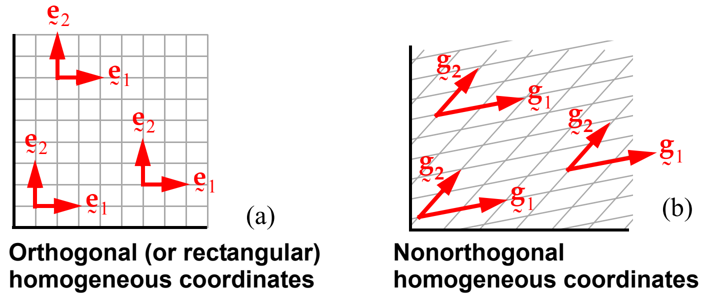


FIGURE 1.2 Homogeneous coordinates. The base vectors are the same at all points in space. This condition is possible only if the coordinate grid is formed by straight lines.

A coordinate system is called "**homogeneous**" if the associated base vectors are the same throughout space. A basis is "**orthogonal**" (or "**rectangular**") if the base vectors are everywhere mutually perpendicular. Most authors use the term "Cartesian coordinates" to refer to the conventional *orthonormal* homogeneous right-handed system of Fig. 1.1a. As seen in Fig. 1.2b, a homogeneous system is not required to be orthogonal. Furthermore, no coordinate system is required to have unit base vectors. The opposite of homogeneous is "curvilinear," and Fig. 1.3 below shows that a coordinate system can be both curvilinear and orthogonal. In short, the properties of being "orthogonal" or "homogeneous" are independent (one does not imply or exclude the other).

1.3 Curvilinear coordinates

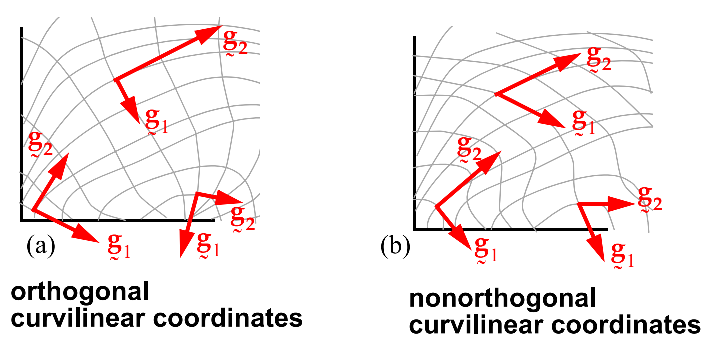


FIGURE 1.3 Curvilinear coordinates. The base vectors are still tangent to coordinate lines. The left system is curvilinear and orthogonal (the coordinate lines always meet at right angles).

The **coordinate grid** is the family of lines along which only one coordinate varies. If the grid has at least some curved lines, the coordinate system is called "**curvilinear**," and, as shown in Fig. 1.3, the associated base vectors (tangent to the grid lines) necessarily change with position, so curvilinear systems are always inhomogeneous. The system in Fig. 1.3a has base vectors that are everywhere orthogonal, so it is simultaneously curvilinear and orthogonal. Note from Fig. 1.1 that conventional cylindrical and spherical coordinates are both orthogonal and curvilinear. Incidentally, no matter what type of coordinate system is used, base vectors need not be of unit length; they only need to point in the direction that the position vector would move when changing the associated coordinate, holding others constant.[1] We will call a basis "**regular**" if it consists of a right-handed orthonormal triad. The systems in Fig. 1.3 have irregular associated base vectors. The system in Fig 1.3a can be "regularized" by normalizing the base vectors. Cylindrical and spherical systems are examples of regularized curvilinear systems.

In Section 2, we introduce mathematical tools for both irregular homogeneous and irregular curvilinear coordinates first deals with the possibility that the base vectors might be non-orthogonal, non-normalized, and/or non-right-handed. Section 3 shows that the component formulas for many operations such as the dot product take on forms that are different from the regular (right-handed orthonormal) formulas. The distinction between homogeneous and curvilinear coordinates becomes apparent in Section 5, where the derivative of a vector or higher order tensor requires additional terms to account for the variation of curvilinear base vectors with position. By contrast, homogeneous base vectors do not vary with position, so the tensor calculus formulas look very much like their Cartesian counterparts, even if the associated basis is irregular.

1.4 Difference between Affine (non-metric) and Metric spaces

As discussed by Papastavridis [12], there are situations where the axes used to define a space don't have the same physical dimensions, and there is no possibility of comparing the units of one axis against the units of another axis. Such spaces are called "affine" or "non-metric." The apropos example cited by Papastavridis is "thermodynamic state space" in which the pressure, volume, and temperature of a fluid are plotted against one another. In such a space, the concept of lengths (and therefore angles) between two points becomes meaningless. In affine geometries, we are only interested in properties that remain invariant under arbitrary scale and angle changes of the axes.

The remainder of this document is dedicated to *metric* spaces such as the ordinary physical 3D space that we all (hopefully) live in.

2. Dual bases for irregular bases

Suppose there are compelling physical reasons to use an irregular basis $\{\underset{\sim}{\mathbf{g}}_1, \underset{\sim}{\mathbf{g}}_2, \underset{\sim}{\mathbf{g}}_3\}$. Here, "irregular" means the basis might be nonorthogonal, non-normalized, and/or non-right-handed. In this section we develop tools needed to derive modified component formulas for tensor operations such as the dot product. For tensor *algebra*, it is irrelevant whether the basis is homogeneous or curvilinear; all that matters is the possibility that the base vectors

1. Strictly speaking, it is not necessary to require that the base vectors have any relationship whatsoever with the coordinate lines. If desired, for example, we *could* use arbitrary curvilinear coordinates while taking the basis to be everywhere aligned with the laboratory basis. In this document, however, the basis is *always* assumed tangent to coordinate lines. Such a basis is called the "associated" basis.

might not be orthogonal and/or might not be of unit length and/or might not form a right-handed system. Again, keep in mind that we will be deriving new procedures for computing the operations, but the ultimate result and meanings for the operations will be unchanged. If, for example, you had two vectors expressed in terms of an irregular basis, then you could always transform those vectors into conventional RCC expansions in order to compute the dot product. The point of this section is to deduce *faster* methods that permit you to obtain the *same result* directly from the irregular vector components without having to transform to RCC.

To simplify the discussion, we will assume that the underlying space is our ordinary 3D physical Euclidean space.[1] Whenever needed, we may therefore assume there exists a *right-handed* orthonormal **laboratory basis**, $\{\underset{\sim}{\mathbf{e}}_1, \underset{\sim}{\mathbf{e}}_2, \underset{\sim}{\mathbf{e}}_3\}$ where $\underset{\sim}{\mathbf{e}}_i \bullet \underset{\sim}{\mathbf{e}}_j = \delta_{ij}$. This is particularly convenient because we can then claim that there exists a transformation tensor $\underset{\approx}{\mathbf{F}}$ such that

$$\underset{\sim}{\mathbf{g}}_i = \underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{e}}_i . \tag{2.1}$$

If this transformation tensor is written in component form with respect to the laboratory basis, then the i^{th} column of the matrix [F] contains the components of the $\underset{\sim}{\mathbf{g}}_i$ base vector with respect to the laboratory basis. In terms of the *lab* components of [F], Eq. (2.1) can be written

$$\underset{\sim}{\mathbf{g}}_i = F_{ji} \underset{\sim}{\mathbf{e}}_j \tag{2.2}$$

Comparing Eq. (2.1) with our formula for how to dot a tensor into a vector [Eq. (1.10)], you might wonder why Eq. (2.2) involves F_{ji} instead of F_{ij}. After all, Eq. (1.10) appears to be telling us that adjacent indices should be summed, but Eq. (2.2) shows the summation index j being summed with the farther (first) index on the tensor. There's a subtle and important phenomenon here that needs careful attention whenever you deal with equations like (2.1) that really represent *three separate equations* for each value of i from 1 to 3. To unravel the mystery, let's start by changing the symbol used for the free index in Eq. (2.1) by writing it equivalently by $\underset{\sim}{\mathbf{g}}_k = \underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{e}}_k$. Now, applying Eq. (1.10) gives $(\underset{\sim}{\mathbf{g}}_k)_i = F_{ij}(\underset{\sim}{\mathbf{e}}_k)_j$. Any vector, $\underset{\sim}{\mathbf{v}}$, can be expanded as $\underset{\sim}{\mathbf{v}} = v_i \underset{\sim}{\mathbf{e}}_i$. Applying this identity with $\underset{\sim}{\mathbf{v}}$ replaced by $\underset{\sim}{\mathbf{g}}_k$ gives $\underset{\sim}{\mathbf{g}}_k = (\underset{\sim}{\mathbf{g}}_k)_i \underset{\sim}{\mathbf{e}}_i$, or, $\underset{\sim}{\mathbf{g}}_k = F_{ij}(\underset{\sim}{\mathbf{e}}_k)_j \underset{\sim}{\mathbf{e}}_i$. The expression $(\underset{\sim}{\mathbf{e}}_k)_j$ represents the j^{th} lab component of $\underset{\sim}{\mathbf{e}}_k$, so it must equal δ_{kj}. Consequently, $\underset{\sim}{\mathbf{g}}_k = F_{ij}\delta_{kj}\underset{\sim}{\mathbf{e}}_i = F_{ik}\underset{\sim}{\mathbf{e}}_i$, which is equivalent to Eq. (2.2).

Incidentally, the transformation tensor $\underset{\approx}{\mathbf{F}}$ may be written in a purely dyadic form as

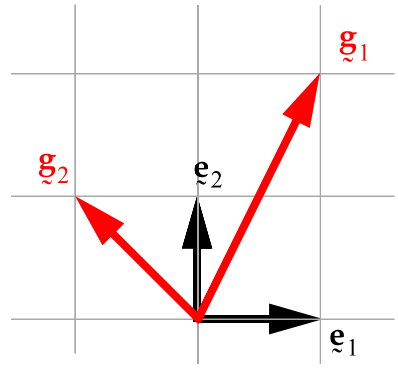
$$\underset{\approx}{\mathbf{F}} = \underset{\sim}{\mathbf{g}}_k \underset{\sim}{\mathbf{e}}_k \tag{2.3}$$

---

1. To quote from Ref. [7], "Three-dimensional Euclidean space, E_3, may be characterized by a set of axioms that expresses relationships among primitive, undefined quantities called points, lines, etc. These relationships so closely correspond to the results of ordinary measurements of distance in the physical world that, until the appearance of general relativity, it was thought that Euclidean geometry was *the* kinematic model of the universe."

Study Question 2.1 Consider the following irregular base vectors expressed in terms of the laboratory basis:

$$\underset{\sim}{\mathbf{g}}_1 = \underset{\sim}{\mathbf{e}}_1 + 2\underset{\sim}{\mathbf{e}}_2$$
$$\underset{\sim}{\mathbf{g}}_2 = -\underset{\sim}{\mathbf{e}}_1 + \underset{\sim}{\mathbf{e}}_2$$
$$\underset{\sim}{\mathbf{g}}_3 = \underset{\sim}{\mathbf{e}}_3 \ .$$



(a) Explain why this basis is irregular.

(b) Find the 3×3 matrix of components of the transformation tensor $\underset{\approx}{F}$ with respect to the laboratory basis.

Partial Answer: *(a) The term "regular" is defined on page 10.*
(b) In terms of the lab basis, the component array of the first base vector is $\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$*, so this must be the first column of the [F] matrix.*

This transformation tensor $\underset{\approx}{\mathbf{F}}$ is defined for the *specific* irregular basis of interest as it relates to the laboratory basis. The transformation tensor for a different pair of bases will be different. This does not imply that $\underset{\approx}{\mathbf{F}}$ is not a tensor. Readers who are familiar with continuum mechanics may be wondering whether our basis transformation tensor $\underset{\approx}{\mathbf{F}}$ has anything to do with the deformation gradient tensor $\underset{\approx}{\mathbf{F}}$ used to describe continuum motion. The answer is "no." In general, the tensor $\underset{\approx}{\mathbf{F}}$ in this document merely represents the relationship between the laboratory basis and the irregular basis. Even though our $\underset{\approx}{\mathbf{F}}$ tensor is generally unrelated to the deformation gradient $\underset{\approx}{\mathbf{F}}$ tensor from continuum mechanics, it's still interesting to consider the special case in which these two tensors *are* the same. If a deforming material is conceptually "painted" with an orthogonal grid in its reference state, then this grid will deform with the material, thereby providing a natural "embedded" curvilinear coordinate system with an associated "natural" basis that is everywhere tangent to the painted grid lines. When this "natural" embedded basis is used, our transformation tensor $\underset{\approx}{\mathbf{F}}$ will be identical to the deformation gradient tensor $\underset{\approx}{\mathbf{F}}$. The component forms of many constitutive material models become intoxicatingly simple in structure when expressed using an embedded basis (it remains a point of argument, however, whether or not simple structure implies intuitiveness). The embedded basis *co-varies* with the grid lines — in other words, these vectors stay always tangent to the grid lines and they stretch in proportion with the stretching of the grid lines. For this reason, the embedded basis is called the **covariant** basis. Later on, we will introduce a companion triad of vectors, called the **contravariant** basis, that does not move with the grid lines; instead we will find that the contravariant basis moves in a way that it remains always perpendicular to material planes that do co-vary with the deformation. When a plane of particles moves with the material, its normal does not generally move with the material!

Our tensor $\underset{\approx}{\mathbf{F}}$ can be seen as characterizing a transformation operation that will take you from the orthonormal laboratory base vectors to the irregular base vectors. The three irregular base vectors, $\{\underset{\sim}{\mathbf{g}}_1, \underset{\sim}{\mathbf{g}}_2, \underset{\sim}{\mathbf{g}}_3\}$ form a triad, which in turn defines a parallelepiped. The volume of the parallelepiped is given by the **Jacobian** of the transformation tensor $\underset{\approx}{\mathbf{F}}$, defined by

$$\boxed{J \equiv \det[\underset{\approx}{\mathbf{F}}]}\,. \tag{2.4}$$

Geometrically, the Jacobian J in Eq. (2.4) equals the volume of the parallelepiped formed by the covariant base vectors $\{\underset{\sim}{\mathbf{g}}_1, \underset{\sim}{\mathbf{g}}_2, \underset{\sim}{\mathbf{g}}_3\}$. To see why this triple scalar product is identically equal to the determinant of the transformation tensor $\underset{\approx}{\mathbf{F}}$, we now introduce the **direct notation definition of a determinant**:

> The determinant, $\det[\underset{\approx}{\mathbf{F}}]$ (also called the **Jacobian**), of a tensor $\underset{\approx}{\mathbf{F}}$ is the unique scalar satisfying
>
> $$[\underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{u}}, \underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{v}}, \underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{w}}] = \det[\underset{\approx}{\mathbf{F}}]\,[\underset{\sim}{\mathbf{u}}, \underset{\sim}{\mathbf{v}}, \underset{\sim}{\mathbf{w}}] \qquad \text{for all vectors } \{\underset{\sim}{\mathbf{u}}, \underset{\sim}{\mathbf{v}}, \underset{\sim}{\mathbf{w}}\}\,. \tag{2.5}$$

Geometrically this strange-looking definition of the determinate states that if a parallelepiped is formed by three vectors, $\{\underset{\sim}{\mathbf{u}}, \underset{\sim}{\mathbf{v}}, \underset{\sim}{\mathbf{w}}\}$, and a transformed parallelepiped is formed by the three transformed vectors $\{\underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{u}}, \underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{v}}, \underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{w}}\}$, then the ratio of the transformed volume to the original volume will have a unique value, regardless what three vectors are chosen to form original the parallelepiped! This volume ratio is the determinant of the transformation tensor.

Since Eq. (2.5) must hold for all vectors, *it must hold for any particular choices of those vectors*. Suppose we choose to identify $\{\underset{\sim}{\mathbf{u}}, \underset{\sim}{\mathbf{v}}, \underset{\sim}{\mathbf{w}}\}$ with the underlying orthonormal basis $\{\underset{\sim}{\mathbf{e}}_1, \underset{\sim}{\mathbf{e}}_2, \underset{\sim}{\mathbf{e}}_3\}$. Then, recalling from Eq. (2.4) that $\det[\underset{\approx}{\mathbf{F}}]$ is denoted by the Jacobian J, Eq. (2.5) becomes $[\underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{e}}_1, \underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{e}}_2, \underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{e}}_3]=J[\underset{\sim}{\mathbf{e}}_1, \underset{\sim}{\mathbf{e}}_2, \underset{\sim}{\mathbf{e}}_3]$. The underlying Cartesian basis $\{\underset{\sim}{\mathbf{e}}_1, \underset{\sim}{\mathbf{e}}_2, \underset{\sim}{\mathbf{e}}_3\}$ is orthonormal and right-handed, so $[\underset{\sim}{\mathbf{e}}_1, \underset{\sim}{\mathbf{e}}_2, \underset{\sim}{\mathbf{e}}_3]=1$. Recalling from Eq. (2.1) that the covariant basis is obtained by the transformation $\underset{\sim}{\mathbf{g}}_i = \underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{e}}_i$, we get

$$[\underset{\sim}{\mathbf{g}}_1, \underset{\sim}{\mathbf{g}}_2, \underset{\sim}{\mathbf{g}}_3] = J\,, \tag{2.6}$$

which completes the proof that the Jacobian J can be computed by taking the determinant of the Cartesian transformation tensor or by simply taking the triple scalar product of the covariant base vectors, whichever method is more convenient:

$$\boxed{J = \det[\underset{\approx}{\mathbf{F}}] = \underset{\sim}{\mathbf{g}}_1 \bullet (\underset{\sim}{\mathbf{g}}_2 \times \underset{\sim}{\mathbf{g}}_3) \equiv [\underset{\sim}{\mathbf{g}}_1, \underset{\sim}{\mathbf{g}}_2, \underset{\sim}{\mathbf{g}}_3]}\,. \tag{2.7}$$

The set of vectors $\{\underset{\sim}{\mathbf{g}}_1, \underset{\sim}{\mathbf{g}}_2, \underset{\sim}{\mathbf{g}}_3\}$ forms a basis if and only if $\underset{\approx}{\mathbf{F}}$ is invertible — i.e., the Jacobian must be nonzero. By choice, the laboratory basis $\{\underset{\sim}{\mathbf{e}}_1, \underset{\sim}{\mathbf{e}}_2, \underset{\sim}{\mathbf{e}}_3\}$ is regular and therefore right-handed. Hence, the irregular basis $\{\underset{\sim}{\mathbf{g}}_1, \underset{\sim}{\mathbf{g}}_2, \underset{\sim}{\mathbf{g}}_3\}$ is

right-handed if $J > 0$
left-handed if $J < 0$. (2.8)

2.1 Modified summation convention

Given that $\{\underset{\sim}{\mathbf{g}}_1, \underset{\sim}{\mathbf{g}}_2, \underset{\sim}{\mathbf{g}}_3\}$ is a basis, we know there exist unique coefficients $\{a^1, a^2, a^3\}$ such that any vector $\underset{\sim}{\mathbf{a}}$ can be written $\underset{\sim}{\mathbf{a}} = a^1\underset{\sim}{\mathbf{g}}_1 + a^2\underset{\sim}{\mathbf{g}}_2 + a^3\underset{\sim}{\mathbf{g}}_3$. Using Einstein's summation notation, you may write this expansion as

$$\underset{\sim}{\mathbf{a}} = a^i\underset{\sim}{\mathbf{g}}_i \tag{2.9}$$

By convention, components with respect to an irregular basis $\{\underset{\sim}{\mathbf{g}}_1, \underset{\sim}{\mathbf{g}}_2, \underset{\sim}{\mathbf{g}}_3\}$ are identified with *superscripts*,[1] rather than subscripts. Summations always occur on different levels — a superscript is always paired with a subscript in these implied summations. The summation convention rules for an irregular basis are:

1. An index that appears exactly once in any term is called a "**free index**," and it must appear exactly once in *every* term in the expression.
2. Each particular free index must appear at the same level in every term. Distinct free indices may permissibly appear at different levels.
3. Any index that appears exactly twice in a given term is called a **dummy sum index** and implies summation from 1 to 3. No index may appear more than twice in a single term.
4. Given a dummy sum pair, one index must appear at the **upper "contravariant" level**, and one must appear at the **lower "covariant" level**.
5. Exceptions to the above rules must be clearly indicated whenever the need arises.
   - Exceptions of rule #1 are extremely rare in tensor analysis because rule #1 can never be violated in any well-formed tensor expression. However, exceptions to rule #1 do regularly appear in non-tensor (matrix) equations. For example, one might define a matrix $[A]$ with components given by $A_{ij} = v_i + v_j$. Here, both i and j are free indices, and the right-hand-side of this equation violates rule #1 because the index i occurs exactly once in the first term but not in the second term. This definition of the A_{ij} numbers is certainly well-defined in a matrix sense[2], but the equation is a violation of tensor index rule #1. Consequently if you *really do* wish to use the equation $A_{ij} = v_i + v_j$ to define some matrix $[A]$, then you should include a parenthetical comment that the tensor index conventions are not to be applied — otherwise your readers will think you made a typo.
   - Exceptions of rules #2 and #4 can occur when working with a regular (right-handed orthonormal) basis because it turns out that there is no distinction between covariant and contravariant components when the basis is regular. For example, v_i is identically equal to v^i when the basis is regular. That's why indicial expressions in most engineering publications show *all* components using only subscripts.
   - Exceptions of rule #3 sometimes occur when the indices are actually referenced to a *particular* basis and are not intended to apply to *any* basis. Consider, for example, how you would need to handle an exception to rule #3 when defining the i^{th} principle direction $\underset{\sim}{\mathbf{p}}_i$ and eigenvalue λ_i associated with some tensor $\underset{\approx}{\mathbf{T}}$. You would have to write something like "$\underset{\approx}{\mathbf{T}} \bullet \underset{\sim}{\mathbf{p}}_i = \lambda_i \underset{\sim}{\mathbf{p}}_i$ (no sum over index i)" in order to call attention to the fact that the index i is supposed to be a *free* index, not summed. Another exception to rule #3 occurs when an index appears only once, but you *really do* wish for a summation over that index. In that case you must explicitly show the summation sign in front of the equation. Similarly, if you *really do* wish for an index to appear more than twice, then you must explicitly indicate whether that index is free or summed.

1. The superscripts are only indexes, not exponents. For example, a^2 is the *second* contravariant component of a vector $\underset{\sim}{\mathbf{a}}$ — it is not the square of some quantity a. If your work *does* involve some scalar quantity "a", then you should typeset its square as $(a)^2$ whenever there is any chance for confusion.
2. This equation is not well defined as an indicial *tensor* equation because it will not transform properly under a basis change. The concept of what constitutes a well-formed tensor operation will be discussed in more detail later.

We arbitrarily elected to place the index on the lower level of our basis $\{\underset{\sim}{\mathbf{g}}_1, \underset{\sim}{\mathbf{g}}_2, \underset{\sim}{\mathbf{g}}_3\}$, so (recalling rule #4) we call it the **"covariant" basis**. The coefficients $\{a^1, a^2, a^3\}$ have the index on the upper level and are therefore called the **contravariant components** of the vector $\underset{\sim}{\mathbf{a}}$. Later on, we will define ***co*variant components** $\{a_1, a_2, a_3\}$ with respect to a carefully defined complementary **"*contra*variant" basis** $\{\underset{\sim}{\mathbf{g}}^1, \underset{\sim}{\mathbf{g}}^2, \underset{\sim}{\mathbf{g}}^3\}$. We will then have two ways to write the vector: $\underset{\sim}{\mathbf{a}} = a^i \underset{\sim}{\mathbf{g}}_i = a_i \underset{\sim}{\mathbf{g}}^i$. Keep in mind: we have not yet indicated how these contra- and co-variant components are computed, or what they mean physically, or why they are useful. For now, we are just introducing the standard "high-low" notation used in the study of irregular bases. *(You may find the phrase "co-go-below" helpful to remember the difference between co- and contra-variant.)*

We will eventually show there are *four* ways to write a second-order tensor. We will introduce contravariant components T^{ij} and covariant components T_{ij}, such that $\underset{\approx}{\mathbf{T}} = T^{ij} \underset{\sim}{\mathbf{g}}_i \underset{\sim}{\mathbf{g}}_j = T_{ij} \underset{\sim}{\mathbf{g}}^i \underset{\sim}{\mathbf{g}}^j$. We will also introduce "mixed" components $T^i_{\bullet j}$ and $T_i^{\bullet j}$ such that $\underset{\approx}{\mathbf{T}} = T^i_{\bullet j} \underset{\sim}{\mathbf{g}}_i \underset{\sim}{\mathbf{g}}^j = T_i^{\bullet j} \underset{\sim}{\mathbf{g}}^i \underset{\sim}{\mathbf{g}}_j$. Note the use of a "dot" to serve as a place holder to indicate the order of the indices (the order of the indices is dictated by the order of the dyadic basis pair). As shown in Section 3.4, use of a "dot" placeholder is necessary only for *nonsymmetric* tensors. (namely, we will find that symmetric tensor components satisfy the property that $T^i_{\bullet j} = T_j^{\bullet i}$, *so the placement of the "dot" is inconsequential for symmetric tensors.)*. In professionally typeset manuscripts, the dot placeholder might not be necessary because, for example, $T^i{}_j$ can be typeset in a manner that is clearly distinguishable from $T_j{}^i$. The dot placeholders are more frequently used in handwritten work, where individuals have unreliable precision or clarity of penmanship. Finally, the number of dot placeholders used in an expression is typically kept to the minimum necessary to clearly demark the order of the indices. For example, $T^{i\bullet}_{\bullet j}$ means the same thing as $T^i_{\bullet j}$. Either of these expressions clearly show that the indices are supposed to be ordered as "i followed by j," not vice versa. Thus, only one dot is enough to serve the purpose of indicating order. Similarly, $T^{ij}_{\bullet\bullet}$ means the same thing as T^{ij}, but the dots serve no clarifying purpose for this case when all indices are on the same level (thus, they are omitted). The importance of clearly indicating the *order* of the indices is inadequately emphasized in some texts.[1]

1. For example, Ref. [4] fails to clearly indicate index ordering. They use neither well-spaced typesetting nor dot placeholders, which can be confusing.

IMPORTANT: As proved later (Study question 2.6), there is no difference between covariant and contravariant components whenever the basis is orthonormal. Hence, for example, $\{\underset{\sim}{\mathbf{e}}^1, \underset{\sim}{\mathbf{e}}^2, \underset{\sim}{\mathbf{e}}^3\}$ is the same as $\{\underset{\sim}{\mathbf{e}}_1, \underset{\sim}{\mathbf{e}}_2, \underset{\sim}{\mathbf{e}}_3\}$. Nevertheless, in order to always satisfy rules #2 and #4 of the sum conventions, we rewrite all familiar orthonormal formulas so that the summed subscripts are on different levels. Furthermore, throughout this document, the following are all equivalent symbols for the **Kronecker delta**:

$$\delta_{ij},\ \delta_i^j,\ \delta^{ij} \quad = \begin{cases} 1 & \text{if } i=j \\ 0 & \text{if } i \neq j \end{cases}. \tag{2.10}$$

BEWARE: as discussed in Section 3.5, the set of δ^{ij} values should be regarded as indexed symbols as defined above, not as components of any particular tensor. Yes, it's true that δ^{ij} are components of the identity tensor $\underset{\approx}{\mathbf{I}}$ with respect to the underlying rectangular Cartesian basis $\{\underset{\sim}{\mathbf{e}}_1, \underset{\sim}{\mathbf{e}}_2, \underset{\sim}{\mathbf{e}}_3\}$, but they are *not* the contravariant components of the identity tensor with respect to an irregular $\{\underset{\sim}{\mathbf{g}}_1, \underset{\sim}{\mathbf{g}}_2, \underset{\sim}{\mathbf{g}}_3\}$ basis. Likewise, δ_{ij} are *not* the covariant components of the identity tensor with respect to the irregular basis. Interestingly, the mixed-level Kronecker delta components, δ_i^j, do turn out to be the mixed components of the identity tensor with respect to either basis! Most of the time, we will be concerned only with the components of tensors with respect to the irregular basis. The Kronecker delta is important in its own right. This is one reason why we denote the identity tensor $\underset{\approx}{\mathbf{I}}$ by a symbol different from its components. Later on, we will note the importance of the permutation symbol ε_{ijk} (which equals +1 if *ijk* ={123, 231, or 312}, -1 if *ijk*={321, 213, or 132}, and zero otherwise). The permutation symbol represents the components of the alternating tensor with respect to the *any* regular (*i.e.,* right-handed orthonormal basis), but not with respect to an irregular basis. Consequently, we will represent the alternating *tensor* by a different symbol $\underset{\approx}{\xi}$ so that we can continue to use the permutation *symbol* ε_{ijk} as an independent indexed quantity. Tracking the basis to which components are referenced is one the most difficult challenges of curvilinear coordinates.

Important notation glitch Square brackets [ ] will be used to indicate a 3×3 matrix, and braces { } will indicate a 3×1 matrix containing vector components. For example, $\{v^i\}$ denotes the 3×1 matrix that contains the contravariant components of a vector $\underset{\sim}{\mathbf{v}}$. Similarly, $[T_{ij}]$ is the matrix that contains the covariant components of a second-order tensor $\underset{\approx}{\mathbf{T}}$, and $[T^{ij}]$ will be used to denote the *matrix* containing the contravariant components of $\underset{\approx}{\mathbf{T}}$. Any indices appearing inside a matrix merely indicate the co/contravariant nature of the matrix — they are not interpreted in the same way as indices in an indicial expression. The indices *ij* merely to indicate the (high/low/mixed) *level* of the matrix components. The rules on page 15 apply only to proper indicial equations, not to equations involving matrices. We will later prove, for example, that the determinant $\det\underset{\approx}{\mathbf{T}}$ can be computed by the determinant of the mixed components of $\underset{\approx}{\mathbf{T}}$. Thus, we might write $\det\underset{\approx}{\mathbf{T}} = \det[T^i_{\bullet j}]$. The brackets around $T^i_{\bullet j}$ indicate that this equation involves the *matrix* of mixed components, so the rules on page 15 do not apply to the ij indices. It's okay that i and j don't appear on the left-hand side.

2.2 The metric coefficients and the dual contravariant basis

For reasons that will soon become apparent, we introduce a symmetric set of numbers g_{ij}, called the "**metric coefficients**," defined

$$\boxed{g_{ij} \equiv \underset{\sim}{\mathbf{g}}_i \bullet \underset{\sim}{\mathbf{g}}_j} \,. \tag{2.11}$$

When the space is Euclidean, the $\{\underset{\sim}{\mathbf{g}}_1, \underset{\sim}{\mathbf{g}}_2, \underset{\sim}{\mathbf{g}}_3\}$ base vectors can be expressed as linear combinations of the underlying orthonormal laboratory basis and the above set of dot products can be computed using the ordinary orthonormal basis formulas.[1]

We also introduce a **dual "contravariant" basis** $\{\underset{\sim}{\mathbf{g}}^1, \underset{\sim}{\mathbf{g}}^2, \underset{\sim}{\mathbf{g}}^3\}$ defined such that

$$\boxed{\underset{\sim}{\mathbf{g}}^i \bullet \underset{\sim}{\mathbf{g}}_j = \delta^i_j} \,. \tag{2.12}$$

Geometrically, Eq. (2.12) requires that the first contravariant base vector $\underset{\sim}{\mathbf{g}}^1$ must be perpendicular to both $\underset{\sim}{\mathbf{g}}_2$ and $\underset{\sim}{\mathbf{g}}_3$, so it must be of the form $\underset{\sim}{\mathbf{g}}^1 = \alpha(\underset{\sim}{\mathbf{g}}_2 \times \underset{\sim}{\mathbf{g}}_3)$. The *as-yet* undetermined scalar α is determined by requiring that $\underset{\sim}{\mathbf{g}}^1 \bullet \underset{\sim}{\mathbf{g}}_1$ equal unity.

$$\underset{\sim}{\mathbf{g}}^1 \bullet \underset{\sim}{\mathbf{g}}_1 = [\alpha(\underset{\sim}{\mathbf{g}}_2 \times \underset{\sim}{\mathbf{g}}_3)] \bullet \underset{\sim}{\mathbf{g}}_1 = \alpha \underset{\sim}{\mathbf{g}}_1 \bullet (\underset{\sim}{\mathbf{g}}_2 \times \underset{\sim}{\mathbf{g}}_3) = \alpha J \overset{\text{"set"}}{=} 1 \tag{2.13}$$

In the second-to-last step, we recognized the triple scalar product, $\underset{\sim}{\mathbf{g}}_1 \bullet (\underset{\sim}{\mathbf{g}}_2 \times \underset{\sim}{\mathbf{g}}_3)$, to be the Jacobian J defined in Eq. (2.7). In the last step we asserted that the result must equal unity. Consequently, the scalar α is merely the reciprocal of the Jacobian:

$$\alpha = \frac{1}{J} \tag{2.14}$$

All three contravariant base vectors can be determined similarly to eventually give the final result:

$$\underset{\sim}{\mathbf{g}}^1 = \frac{1}{J}(\underset{\sim}{\mathbf{g}}_2 \times \underset{\sim}{\mathbf{g}}_3), \qquad \underset{\sim}{\mathbf{g}}^2 = \frac{1}{J}(\underset{\sim}{\mathbf{g}}_3 \times \underset{\sim}{\mathbf{g}}_1), \qquad \underset{\sim}{\mathbf{g}}^3 = \frac{1}{J}(\underset{\sim}{\mathbf{g}}_1 \times \underset{\sim}{\mathbf{g}}_2) \,. \tag{2.15}$$

where

$$J = \underset{\sim}{\mathbf{g}}_1 \bullet (\underset{\sim}{\mathbf{g}}_2 \times \underset{\sim}{\mathbf{g}}_3) \equiv [\underset{\sim}{\mathbf{g}}_1, \underset{\sim}{\mathbf{g}}_2, \underset{\sim}{\mathbf{g}}_3] \,. \tag{2.16}$$

---

1. Note: If the space is *not* Euclidean, then an orthonormal basis does *not* exist, and the metric coefficients g_{ij} must be specified *a priori*. Such a space is called Riemannian. Shell and membrane theory deals with 2D curved Riemannian manifolds embedded in 3D space. The geometry of general relativity is that of a *four-dimensional* Riemannian manifold. For further examples of Riemannian spaces, see, e.g., Refs. [5, 7].

An alternative way to obtain the dual contravariant basis is to assert that $\{\underset{\sim}{\mathbf{g}}^1, \underset{\sim}{\mathbf{g}}^2, \underset{\sim}{\mathbf{g}}^3\}$ is in fact a basis; we may therefore demand that coefficients L_{ik} must exist such that each covariant base vector $\underset{\sim}{\mathbf{g}}_i$ can be written as a linear combination of the contravariant basis: $\underset{\sim}{\mathbf{g}}_i = L_{ik}\underset{\sim}{\mathbf{g}}^k$. Dotting both sides with $\underset{\sim}{\mathbf{g}}_k$ and imposing Eqs. (2.11) and (2.12) shows that the transformation coefficients must be identical to the covariant metric coefficients: $L_{ik} = g_{ik}$. Thus

$$\underset{\sim}{\mathbf{g}}_i = g_{ik}\underset{\sim}{\mathbf{g}}^k \quad . \tag{2.17}$$

This equation may be solved for the contravariant basis. Namely,

$$\boxed{\underset{\sim}{\mathbf{g}}^i = g^{ik}\underset{\sim}{\mathbf{g}}_k} \, , \tag{2.18}$$

where the matrix of ***contravariant* metric components** g^{ij} is obtained by inverting the covariant metric matrix $[g_{ij}]$. Dotting both sides of Eq. (2.18) by $\underset{\sim}{\mathbf{g}}^j$ we note that

$$\boxed{g^{ij} = \underset{\sim}{\mathbf{g}}^i \bullet \underset{\sim}{\mathbf{g}}^j} \quad , \tag{2.19}$$

which is similar in form to Eq. (2.11).

In later analyses, keep in mind that g^{ij} is the inverse of the g_{ij} matrix. Furthermore, both metric matrices are symmetric. Thus, whenever these are multiplied together with a contracted index, the result is the Kronecker delta:

$$\boxed{g^{ik}g_{kj} = g^{ki}g_{kj} = g^{ik}g_{jk} = g^{ki}g_{jk} = \delta^i_j} \, . \tag{2.20}$$

Another quantity that will appear frequently in later analyses is the determinant g_o of the covariant g_{ij} metric matrix and the determinant g^o of the contravariant g^{ij} metric matrix:

$$g_o \equiv \det\begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{bmatrix} \qquad \text{and} \qquad g^o \equiv \det\begin{bmatrix} g^{11} & g^{12} & g^{13} \\ g^{21} & g^{22} & g^{23} \\ g^{31} & g^{32} & g^{33} \end{bmatrix} \, . \tag{2.21}$$

Recalling that the $[g^{ij}]$ matrix is the inverse of the $[g_{ij}]$ matrix, we note that

$$g_o = \frac{1}{g^o} \, . \tag{2.22}$$

Furthermore, as shown in Study Question 2.5, g_o is related to the Jacobian J from Eqs. (2.4) and (2.6) by

$$g_o = J^2 \, , \tag{2.23}$$

Thus

$$g^o = \frac{1}{J^2} \, . \tag{2.24}$$

Non-trivial lemma[1] Note that Eq. (2.21) shows that g_o may be regarded as a function of the nine components of the g_{ij} matrix. Taking the partial derivative of g_o with respect to a particular g_{ij} component gives a result that is identical to the signed subminor (also called the cofactor) associated with that component. The "subminor" is a number associated with each matrix position that is equal to the determinant of the 2×2 submatrix obtained by striking out the i^{th} row and j^{th} column of the original matrix. To obtain the cofactor (i.e., the *signed subminor*) associated with the *ij* position, the subminor is multiplied by $(-1)^{i+j}$. Denoting this cofactor by g_{ij}^C, we have

$$\frac{\partial g_o}{\partial g_{ij}} = g_{ij}^C \tag{2.25}$$

Any book on matrix analysis will include a proof that the *inverse* of a matrix $[A]$ can be obtained by taking the transpose of the cofactor matrix and dividing by the determinant:

$$[A]^{-1} = \frac{[A]^{CT}}{\det[A]} \tag{2.26}$$

From which it follows that $[A]^C = (\det[A])[A]^{-T}$. Applying this result to the case that $[A]$ is the *symmetric* matrix $[g_{ij}]$, and recalling that $\det[g_{ij}]$ is denoted g_o, and also recalling that $[g_{ij}]^{-1}$ is given by $[g^{ij}]$, Eq. (2.25) can be written

$$\frac{\partial g_o}{\partial g_{ij}} = g_o g^{ij} \tag{2.27}$$

Similarly,

$$\frac{\partial g^o}{\partial g^{ij}} = g^o g_{ij} \tag{2.28}$$

With these equations, we are introduced for the first time to a new index notation rule: subscript indices that appear in the "denominator" of a derivative should be regarded as superscript indices in the expression as a whole. Similarly, superscripts in the denominator should be regarded as subscripts in the derivative as a whole. With this convention, there is no violation of the index rule that requires free indices to be on the same level in all terms.

Recalling Eq. (2.23), we can apply the chain rule to note that

$$\frac{\partial g_o}{\partial g_{ij}} = 2J\frac{\partial J}{\partial g_{ij}} \tag{2.29}$$

Equation (2.27) permits us to express the left-hand-side of this equation in terms of the *contra-*

1. This side-bar can be skipped without impacting your ability to read subsequent material.

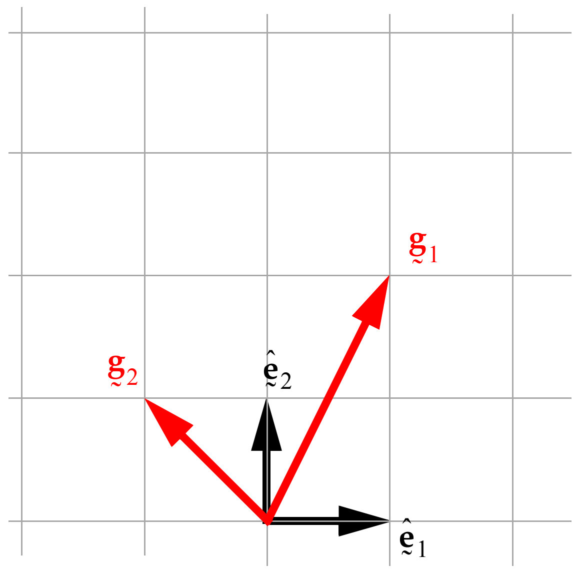
variant metric. Thus, we may solve for the derivative on the right-hand-side to obtain

$$\frac{\partial J}{\partial g_{ij}} = \frac{1}{2} J g^{ij} \tag{2.30}$$

where we have again recalled that $g_o = J^2$.

Study Question 2.2 Let $\{\underset{\sim}{\mathbf{e}}_1, \underset{\sim}{\mathbf{e}}_2, \underset{\sim}{\mathbf{e}}_3\}$ represent the ordinary orthonormal laboratory basis. Consider the following irregular base vectors:

$$\underset{\sim}{\mathbf{g}}_1 = \underset{\sim}{\mathbf{e}}_1 + 2\underset{\sim}{\mathbf{e}}_2$$
$$\underset{\sim}{\mathbf{g}}_2 = -\underset{\sim}{\mathbf{e}}_1 + \underset{\sim}{\mathbf{e}}_2$$
$$\underset{\sim}{\mathbf{g}}_3 = \underset{\sim}{\mathbf{e}}_3 \,.$$



(a) Construct the metric coefficients g_{ij}.

(b) Construct $[g^{ij}]$ by inverting $[g_{ij}]$.

(c) Construct the contravariant (dual) basis by directly using the formula of Eq. (2.15). Sketch the dual basis in the picture at right and visually verify that it satisfies the condition of Eq. (2.12).

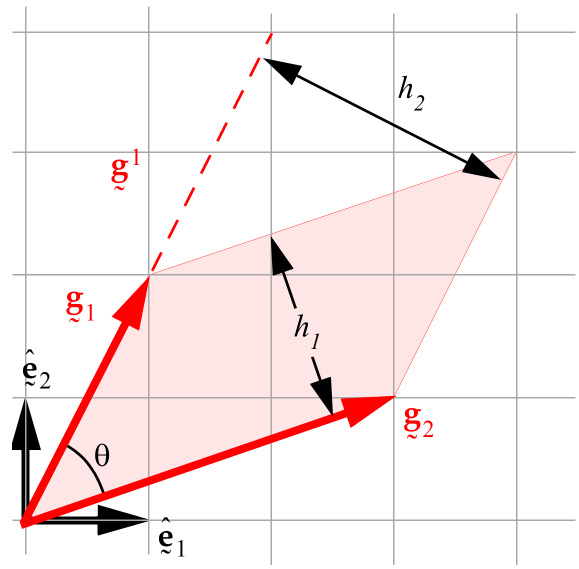
(d) Confirm that the formula of Eq. (2.18) gives the same result as derived in part (c).

(e) Redo parts (a) through (d) if $\underset{\sim}{\mathbf{g}}_3$ is now replaced by $\underset{\sim}{\mathbf{g}}_3 = 5\underset{\sim}{\mathbf{e}}_3$.

Partial Answer: *(a)* $g_{11}=5$, $g_{13}=0$, $g_{22}=2$ *(b)* $g^{11}=2/9$, $g^{33}=1$, $g^{21}=-1/9$ *(c)* $J=3$, $\underset{\sim}{\mathbf{g}}^2=\frac{1}{3}(\underset{\sim}{\mathbf{g}}_3 \times \underset{\sim}{\mathbf{g}}_1)=-\frac{2}{3}\underset{\sim}{\mathbf{e}}_1+\frac{1}{3}\underset{\sim}{\mathbf{e}}_2$. *(d)* $\underset{\sim}{\mathbf{g}}^2 = g^{21}\underset{\sim}{\mathbf{g}}_1 + g^{22}\underset{\sim}{\mathbf{g}}_2 + g^{23}\underset{\sim}{\mathbf{g}}_3 = -\frac{1}{9}\underset{\sim}{\mathbf{g}}_1 + \frac{5}{9}\underset{\sim}{\mathbf{g}}_2 = -\frac{2}{3}\hat{\underset{\sim}{\mathbf{e}}}_1 + \frac{1}{3}\hat{\underset{\sim}{\mathbf{e}}}_2$.

Study Question 2.3 Let $\{\underset{\sim}{\mathbf{e}}_1, \underset{\sim}{\mathbf{e}}_2, \underset{\sim}{\mathbf{e}}_3\}$ represent the ordinary orthonormal laboratory basis. Consider the following irregular base vectors:

$$\underset{\sim}{\mathbf{g}}_1 = \underset{\sim}{\mathbf{e}}_1 + 2\underset{\sim}{\mathbf{e}}_2$$
$$\underset{\sim}{\mathbf{g}}_2 = 3\underset{\sim}{\mathbf{e}}_1 + \underset{\sim}{\mathbf{e}}_2$$
$$\underset{\sim}{\mathbf{g}}_3 = -7\underset{\sim}{\mathbf{e}}_3 \ .$$



(a) is this basis right handed?

(b) Compute $[g^{ij}]$ and $[g_{ij}]$.

(c) Construct the contravariant (dual) basis.

(d) Prove that $\underset{\sim}{\mathbf{g}}^1$ points in the direction of the outward normal to the upper surface of the shaded parallelogram.

(d) Prove that h_k (see drawing label) is the reciprocal of the length of $\underset{\sim}{\mathbf{g}}^k$.

(e) Prove that the area of the face of the parallelepiped whose normal is parallel to $\underset{\sim}{\mathbf{g}}^k$ is given by the Jacobian J times the magnitude of $\underset{\sim}{\mathbf{g}}^k$.

Partial Answer: *(a) yes -- note the direction of the third base vector (b)*

A faster way to get the metric coefficients: Suppose you already have the $[F_{ij}]$ with respect to the regular laboratory basis. We now prove that the $[g_{ij}]$ matrix can by obtained by $[F]^T[F]$. Recall that $\underset{\sim}{\mathbf{g}}_i = \underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{e}}_i$. Therefore

$$\underset{\sim}{\mathbf{g}}_i \bullet \underset{\sim}{\mathbf{g}}_j = (\underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{e}}_i) \bullet (\underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{e}}_j) = (\underset{\sim}{\mathbf{e}}_i \bullet \underset{\approx}{\mathbf{F}}^T) \bullet (\underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{e}}_j) = \underset{\sim}{\mathbf{e}}_i \bullet (\underset{\approx}{\mathbf{F}}^T \bullet \underset{\approx}{\mathbf{F}}) \bullet \underset{\sim}{\mathbf{e}}_j \ , \tag{2.31}$$

The left hand side is g_{ij}, and the right hand side represents the *ij* components of $[F]^T[F]$ with respect to the regular laboratory basis, which completes the proof that the covariant metric coefficients g_{ij} are equal to the *ij* *laboratory* components of the tensor $\underset{\approx}{\mathbf{F}}^T \bullet \underset{\approx}{\mathbf{F}}$

Study Question 2.4 Using the [F] matrix from Study Question 2.1, verify that $[F]^T[F]$ gives the same matrix for g_{ij} as computed in Question 2.2.

Study Question 2.5 Recall that the covariant basis may be regarded as a transformation of the right-handed orthonormal laboratory basis. Namely, $\underset{\sim}{\mathbf{g}}_i = \underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{e}}_i$.

(a) Prove that the *contravariant* basis is obtained by the inverse transpose transformation:

$$\underset{\sim}{\mathbf{g}}^i = \underset{\approx}{\mathbf{F}}^{-T} \bullet \underset{\sim}{\mathbf{e}}^i, \tag{2.32}$$

where $\underset{\sim}{\mathbf{e}}^i$ is the same as $\underset{\sim}{\mathbf{e}}_i$.

(b) Prove that the contravariant metric coefficients g^{ij} are equal to the *ij* *laboratory* components of the tensor $\underset{\approx}{\mathbf{F}}^{-1} \bullet \underset{\approx}{\mathbf{F}}^{-T}$.

(c) Prove that, for any invertible tensor $\underset{\approx}{\mathbf{S}}$, the tensor $\underset{\approx}{\mathbf{S}}^T \bullet \underset{\approx}{\mathbf{S}}$ is positive definite. Explain why this implies that the $[g_{ij}]$ and g^{ij} matrices are positive definite.

(d) Use the result from part (b) to prove that the determinant g_o of the covariant metric matrix g_{ij} equals the square of the Jacobian J^2.

Partial Answer: *(a) Substitute Eqs (2.1) and (2.32) into Eq. (2.12) and use the fact that* $\underset{\sim}{\mathbf{e}}^i \bullet \underset{\sim}{\mathbf{e}}_j = \delta^i_j$. *(b) Follow logic similar to Eq. (2.31). (c) A correct proof must use the fact that* $\underset{\approx}{\mathbf{S}}$ *is invertible. Reminder: a tensor* $\underset{\approx}{\mathbf{A}}$ *is positive definite iff* $\underset{\sim}{\mathbf{u}} \bullet \underset{\approx}{\mathbf{A}} \bullet \underset{\sim}{\mathbf{u}} > 0 \quad \forall \underset{\sim}{\mathbf{u}} \neq \underset{\sim}{\mathbf{0}}$. *The dot product positivity property states that* $\underset{\sim}{\mathbf{v}} \bullet \underset{\sim}{\mathbf{v}} \begin{cases} > 0 \text{ if } \underset{\sim}{\mathbf{v}} \neq \underset{\sim}{\mathbf{0}} \\ = 0 \text{ iff } \underset{\sim}{\mathbf{v}} = \underset{\sim}{\mathbf{0}} \end{cases}$. *(d) Easy: the determinant of a product is the product of the determinants.*

Study Question 2.6 SPECIAL CASE (orthonormal systems)

Suppose that the metric coefficients happen to equal the Kronecker delta:

$$g_{ij} = \delta_{ij} \tag{2.33}$$

(a) Explain why this condition implies that the covariant base vectors are orthonormal.
(b) Does the above equation tell us anything about the handedness of the systems?
(c) Explain why orthonormality implies that contravariant base vectors are identically equal to the covariant base vectors.

(a) By definition, $g_{ij} \equiv \underset{\sim}{\mathbf{g}}_i \bullet \underset{\sim}{\mathbf{g}}_j$, *so the fact that these dot products equal the Kronecker delta says, for example, that* $|\underset{\sim}{\mathbf{g}}_1| = 1$ *(i.e., it is a unit vector) and* $\underset{\sim}{\mathbf{g}}_1$ *is perpendicular to* $\underset{\sim}{\mathbf{g}}_2$, *etc. (b) There is no info about the handedness of the system. (c) By definition, the first contravariant base vector must be*

perpendicular to the second and third covariant vectors, and it must satisfy $\underset{\sim}{\mathbf{g}}^1 \bullet \underset{\sim}{\mathbf{g}}_1 = 1$*, which (with some thought) leads to the conclusion that* $\underset{\sim}{\mathbf{g}}^1 = \underset{\sim}{\mathbf{g}}_1$.

Super-fast way to get the dual (contravariant) basis and metrics Recall [Eq. (2.1)] that the covariant basis can be connected to the lab basis through the equation

$$\underset{\sim}{\mathbf{g}}_i = \underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{e}}_i \tag{2.34}$$

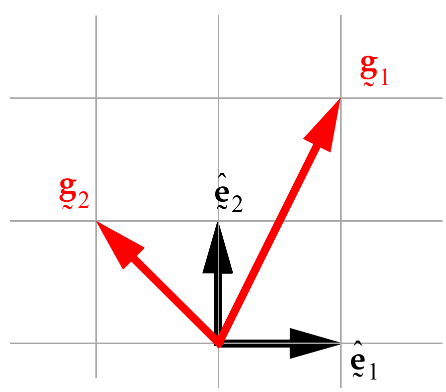
When we introduced this equation, we explained that the i^{th} column of the lab component matrix [F] would contain the lab components of $\underset{\sim}{\mathbf{g}}_i$. Later, in Study Question 2.5, Eq. (2.34), we asserted that

$$\underset{\sim}{\mathbf{g}}^i = \underset{\approx}{\mathbf{F}}^{-T} \bullet \underset{\sim}{\mathbf{e}}^i , \tag{2.35}$$

Consequently, we may conclude that the i^{th} column of $[F]^{-T}$ must contain the lab components of $\underset{\sim}{\mathbf{g}}^i$. This means that the i^{th} *row* of $[F]^{-1}$ must contain the i^{th} contravariant base vector.

Study Question 2.7 Let $\{\underset{\sim}{\mathbf{e}}_1, \underset{\sim}{\mathbf{e}}_2, \underset{\sim}{\mathbf{e}}_3\}$ represent the ordinary orthonormal laboratory basis. Consider the following irregular base vectors:

$$\underset{\sim}{\mathbf{g}}_1 = \underset{\sim}{\mathbf{e}}_1 + 2\underset{\sim}{\mathbf{e}}_2$$
$$\underset{\sim}{\mathbf{g}}_2 = -\underset{\sim}{\mathbf{e}}_1 + \underset{\sim}{\mathbf{e}}_2$$
$$\underset{\sim}{\mathbf{g}}_3 = \underset{\sim}{\mathbf{e}}_3 .$$



(a) Construct the $[F]$ matrix by putting the lab components of $\underset{\sim}{\mathbf{g}}_i$ into the i^{th} column.

(b) Find $[F]^{-1}$

(c) Find $\underset{\sim}{\mathbf{g}}^i$ from the i^{th} row of $[F]^{-1}$.

(d) Directly compute g^{ij} from the result of part (c).

(e) Compare the results with those found in earlier study questions and comment on which method was fastest.

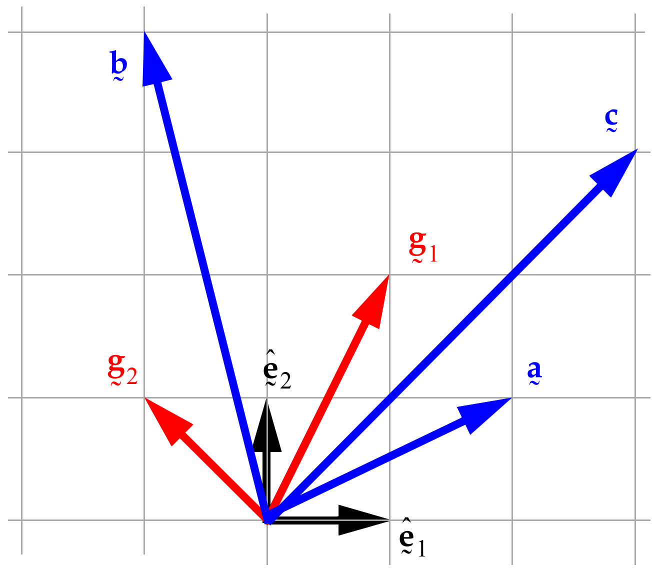
Partial Answer: *(a)* $[F] = \begin{bmatrix} 1 & -1 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ *(b)* $[F]^{-1} = \begin{bmatrix} 1/3 & 1/3 & 0 \\ -2/3 & 1/3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(c), $\underset{\sim}{\mathbf{g}}^2 = -\frac{2}{3}\underset{\sim}{\mathbf{e}}_1 + \frac{1}{3}\underset{\sim}{\mathbf{e}}_2$. *(d)* $g^{12} = \underset{\sim}{\mathbf{g}}^1 \bullet \underset{\sim}{\mathbf{g}}^2 =$ dot product of 1st and 2nd rows $= -\frac{1}{9}$

(e) All results the same. This way was computationally faster!

Study Question 2.8 Consider the same irregular base vectors in Study Question 2.2. Namely, $\underset{\sim}{\mathbf{g}}_1 = \underset{\sim}{\mathbf{e}}_1 + 2\underset{\sim}{\mathbf{e}}_2$, $\underset{\sim}{\mathbf{g}}_2 = -\underset{\sim}{\mathbf{e}}_1 + \underset{\sim}{\mathbf{e}}_2$, $\underset{\sim}{\mathbf{g}}_3 = \underset{\sim}{\mathbf{e}}_3$.

Now consider three vectors, $\underset{\sim}{\mathbf{a}}$, $\underset{\sim}{\mathbf{b}}$, and $\underset{\sim}{\mathbf{c}}$, in the 1-2 plane as shown.



(a) Referring to the sketch, graphically demonstrate that $\underset{\sim}{\mathbf{g}}_1 - \underset{\sim}{\mathbf{g}}_2$ gives a vector identically equal to $\underset{\sim}{\mathbf{a}}$. Thereby explain why the contravariant components of $\underset{\sim}{\mathbf{a}}$ are: $a^1=1$, $a^2=-1$, and $a^3=0$.

(b) Using similar geometrical arguments, find the contravariant components of $\underset{\sim}{\mathbf{b}}$.

(c) Find the contravariant components of $\underset{\sim}{\mathbf{c}}$.

(d) The path from the tail to the tip of a vector can be decomposed into parts that are parallel to the base vectors. For example, the vector $\underset{\sim}{\mathbf{b}}$ can be viewed as a segment equal to $\underset{\sim}{\mathbf{g}}_1$ plus a segment equal to $2\underset{\sim}{\mathbf{g}}_2$. What are the lengths of each of these individual segments?

(e) In general, when any given vector $\underset{\sim}{\mathbf{u}}$ is broken into segments parallel to the covariant base vectors $\{\underset{\sim}{\mathbf{g}}_1, \underset{\sim}{\mathbf{g}}_2, \underset{\sim}{\mathbf{g}}_3\}$, how are the lengths of these segments related (if at all) to the contravariant components $\{u^1, u^2, u^3\}$ of the vector?

Partial Answer: *(a) We know that $a^1=1$ because it's the coefficient of $\underset{\sim}{\mathbf{g}}_1$ in the expression $\underset{\sim}{\mathbf{a}} = \underset{\sim}{\mathbf{g}}_1 - \underset{\sim}{\mathbf{g}}_2$. (b) Draw a path from the origin to the tip of the vector $\underset{\sim}{\mathbf{b}}$ such that the path consists of straight line segments that are always parallel to one of the base vectors $\underset{\sim}{\mathbf{g}}_1$ or $\underset{\sim}{\mathbf{g}}_2$. You can thereby demonstrate graphically that $\underset{\sim}{\mathbf{b}} = \underset{\sim}{\mathbf{g}}_1 + 2\underset{\sim}{\mathbf{g}}_2$. Hence $b^1=1$, $b^2=$ etc (c) You're on your own. (d) The length of the second segment equals the magnitude of $2\underset{\sim}{\mathbf{g}}_2$, or $2\sqrt{5}$. (e) In general, when a vector $\underset{\sim}{\mathbf{u}}$ is broken into segments parallel to the covariant basis, then the length of the segment parallel to $\underset{\sim}{\mathbf{g}}_i$ must equal $|u^i|\sqrt{g_{ii}}$, with no implied sum on i. Lesson here? Because the base vectors are not necessarily of unit length, the meaning of the contravariant component u^i must never be confused with the length of the associated line segment! The contravariant component u^i is merely the coefficient of $\underset{\sim}{\mathbf{g}}_i$ in the linear expansion $\underset{\sim}{\mathbf{u}} = u^i\underset{\sim}{\mathbf{g}}_i$. The component's magnitude equals the length of the segment divided by the length of the base vector!*

2.3 Transforming components by raising and lower indices.

At this point, we have not yet identified a procedural means of determining the contravariant components $\{a^1, a^2, a^3\}$ or covariant components $\{a_1, a_2, a_3\}$ — we merely assert that they *exist*. The sets $\{\underset{\sim}{\mathbf{g}}_1, \underset{\sim}{\mathbf{g}}_2, \underset{\sim}{\mathbf{g}}_3\}$ and $\{\underset{\sim}{\mathbf{g}}^1, \underset{\sim}{\mathbf{g}}^2, \underset{\sim}{\mathbf{g}}^3\}$ are each bases, so we know that we may express any vector $\underset{\sim}{\mathbf{a}}$ as

$$\underset{\sim}{\mathbf{a}} = a^k \underset{\sim}{\mathbf{g}}_k = a_k \underset{\sim}{\mathbf{g}}^k \; . \tag{2.36}$$

Dotting both sides of this equation by $\underset{\sim}{\mathbf{g}}^i$ gives $\underset{\sim}{\mathbf{a}} \bullet \underset{\sim}{\mathbf{g}}^i = a^k \underset{\sim}{\mathbf{g}}_k \bullet \underset{\sim}{\mathbf{g}}^i = a_k \underset{\sim}{\mathbf{g}}^k \bullet \underset{\sim}{\mathbf{g}}^i$. Equation (2.12) allows us to simplify the second term as $a^k \underset{\sim}{\mathbf{g}}_k \bullet \underset{\sim}{\mathbf{g}}^i = a^k \delta_k^i = a^i$. Using Eq. (2.19), the third term simplifies as $a_k \underset{\sim}{\mathbf{g}}^k \bullet \underset{\sim}{\mathbf{g}}^i = a_k g^{ki} = g^{ik} a_k$, where the last step utilized the fact that the g^{ik} matrix is symmetric. Thus, we conclude

$$\boxed{a^i = \underset{\sim}{\mathbf{a}} \bullet \underset{\sim}{\mathbf{g}}^i} \tag{2.37}$$

$$\boxed{a^i = g^{ik} a_k} \; . \tag{2.38}$$

Similarly, by dotting both sides of Eq. (2.36) by $\underset{\sim}{\mathbf{g}}_i$ you can show that

$$\boxed{a_i = \underset{\sim}{\mathbf{a}} \bullet \underset{\sim}{\mathbf{g}}_i} \tag{2.39}$$

$$\boxed{a_i = g_{ik} a^k} \; . \tag{2.40}$$

These very important equations provide a means of transforming between types of components. In Eq. (2.40), note how the g_{ik} matrix effectively "lowers" the index on a^k, changing it to an "i". Similarly, in Eq. (2.38), the g^{ik} matrix "raises" the index on a_k, changing it to an "i". Thus the metric coefficients serve a role that is very similar to that of the Kronecker delta in orthonormal theory. Incidentally, note that Eqs. (2.17) and (2.18) are also examples of raising and lowering indices.

Study Question 2.9 Simplify the following expressions so that there are no metric coefficients:

$$a^m g_{ml}, \quad g^{pk} u_p, \quad f_n g^{ni}, \quad r_i g^{ij} s_j, \quad g^{ij} b^k g_{kj}, \quad g_{ij} g^{jk} \tag{2.41}$$

Partial Answer: *For the first one, the index "m" is repeated. The g_{ml} allows you to lower the superscript "m" on a^m so that it becomes an "l" when the metric coefficient is removed. The final simplified result for the first expression is a_l. Be certain that your final simplified results still have the same free indices on the same level as they were in the original expression. The very last expression has a twist: recognizing the repeated index j, we can lower the index j on g^{jk} and turn it into an "i" when we remove g_{ij} so that a simplified expression becomes $g_{ij}g^{jk} = g_i{}^k$. Alternatively, if we prefer to remove the g^{jk}, we could raise the index j on g_{ij}, making it into a "k" so that an alternative simplification of the very last expression becomes $g_{ij}g^{jk} = g_i{}^k$. Either method gives the same result. The final simplification comes from recalling that the mixed metric components are identically equal to the Kronecker delta. Thus, $g_{ij}g^{jk} = \delta_i^k$, which is merely one of the expressions given in Eq. (2.20).*

Finding contravariant vector components — classical method In Study Question 2.8, we used geometrical methods to find the contravariant components of several vectors. Eqs (2.37) through (2.40) provide us with an *algebraic* means of determining the contravariant and covariant components of vectors. Namely, given a basis $\{\underset{\sim}{\mathbf{g}}_1, \underset{\sim}{\mathbf{g}}_2, \underset{\sim}{\mathbf{g}}_3\}$ and a vector $\underset{\sim}{\mathbf{a}}$, the contra- and co-variant components of the vector are determined as follows:

STEP 1. Compute the covariant metric coefficients $g_{ij} = \underset{\sim}{\mathbf{g}}_i \bullet \underset{\sim}{\mathbf{g}}_j$.

STEP 2. Compute the contravariant metric coefficients g^{ij} by inverting $[g_{ij}]$.

STEP 3. Compute the contravariant basis $\underset{\sim}{\mathbf{g}}^i = g^{ij}\underset{\sim}{\mathbf{g}}_j$.

STEP 4. Compute the covariant components $a_i = \underset{\sim}{\mathbf{a}} \bullet \underset{\sim}{\mathbf{g}}_i$.

STEP 5. Compute the contravariant components $a^i = \underset{\sim}{\mathbf{a}} \bullet \underset{\sim}{\mathbf{g}}^i$. Alternatively, $a^i = g^{ij}a_j$.

Finding contravariant vector components — accelerated method If the lab components of the vector $\underset{\sim}{\mathbf{a}}$ are available, then you can quickly compute the covariant and contravariant components by noting that $a_i = \underset{\sim}{\mathbf{a}} \bullet \underset{\sim}{\mathbf{g}}_i = \underset{\sim}{\mathbf{a}} \bullet \underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{e}}_i = (\underset{\approx}{\mathbf{F}}^T \bullet \underset{\sim}{\mathbf{a}}) \bullet \underset{\sim}{\mathbf{e}}_i$ and, similarly, $a^i = (\underset{\approx}{\mathbf{F}}^{-1} \bullet \underset{\sim}{\mathbf{a}}) \bullet \underset{\sim}{\mathbf{e}}_i$. The steps are as follows:

STEP 1. Construct the $[F]$ matrix by putting the lab components of $\underset{\sim}{\mathbf{g}}_i$ into the i^{th} column.

STEP 2. The covariant components are found from $[F]^T\{a\}$, using lab components.

STEP 3. The contravariant components are found from $[F]^{-1}\{a\}$, using lab components.

Study Question 2.10 Using the basis and vectors from Study Question 2.8, apply the classical step-by-step algorithm to compute the covariant and contravariant components of the vectors, $\underset{\sim}{\mathbf{a}}$ and $\underset{\sim}{\mathbf{b}}$. Check whether the accelerated method gives the same answer. Be sure to verify that your contravariant components agree with those determined geometrically in Question 2.8.

Partial Answer: *Classical method: First get the contra- and co-variant base vectors, then apply the formulas:*

$a_2 = \underset{\sim}{\mathbf{a}} \bullet \underset{\sim}{\mathbf{g}}_2 = (2\underset{\sim}{\mathbf{e}}_1 + \underset{\sim}{\mathbf{e}}_2) \bullet (-\underset{\sim}{\mathbf{e}}_1 + \underset{\sim}{\mathbf{e}}_2) = -1$, *and so on giving* $\{a_1, a_2, a_3\} = \{4, -1, 0\}$.

$a^2 = \underset{\sim}{\mathbf{a}} \bullet \underset{\sim}{\mathbf{g}}^2 = (2\underset{\sim}{\mathbf{e}}_1 + \underset{\sim}{\mathbf{e}}_2) \bullet (-\tfrac{2}{3}\underset{\sim}{\mathbf{e}}_1 + \tfrac{1}{3}\underset{\sim}{\mathbf{e}}_2) = -1$, *and so on giving* $\{a^1, a^2, a^3\} = \{1, -1, 0\}$.

ACCELERATED METHOD:

covariant components: $[F]^T\{a\}_{\text{lab}} = \begin{bmatrix} 1 & 2 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ -1 \\ 0 \end{bmatrix}$ . . . *agrees!*

contravariant components: $[F]^{-1}\{a\}_{\text{lab}} = \begin{bmatrix} 1/3 & 1/3 & 0 \\ -2/3 & 1/3 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ . . . *agrees!*

By applying techniques similar to those used to derive Eqs. (2.37) through (2.40), the following analogous results can be shown to hold for tensors:

$$\underset{\approx}{\mathbf{T}} = T^{ij}\underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}_j = T_{ij}\underset{\sim}{\mathbf{g}}^i\underset{\sim}{\mathbf{g}}^j = T^i_{\bullet j}\underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}^j = T_i^{\bullet j}\underset{\sim}{\mathbf{g}}^i\underset{\sim}{\mathbf{g}}_j \tag{2.42}$$

$$T^{ij} = \underset{\sim}{\mathbf{g}}^i \bullet \underset{\approx}{\mathbf{T}} \bullet \underset{\sim}{\mathbf{g}}^j = T_{mn}g^{im}g^{jn} = T^i_{\bullet k}g^{kj} = T_k^{\bullet j}g^{ki} \tag{2.43a}$$

$$T_{ij} = \underset{\sim}{\mathbf{g}}_i \bullet \underset{\approx}{\mathbf{T}} \bullet \underset{\sim}{\mathbf{g}}_j = T^{mn}g_{mi}g_{nj} = T^k_{\bullet j}g_{ki} = T_i^{\bullet k}g_{kj} \tag{2.43b}$$

$$T^i_{\bullet j} = \underset{\sim}{\mathbf{g}}^i \bullet \underset{\approx}{\mathbf{T}} \bullet \underset{\sim}{\mathbf{g}}_j = T^{ik}g_{kj} = T_{kj}g^{ki} = T_m^{\bullet n}g^{mi}g_{nj} \tag{2.43c}$$

$$T_i^{\bullet j} = \underset{\sim}{\mathbf{g}}_i \bullet \underset{\approx}{\mathbf{T}} \bullet \underset{\sim}{\mathbf{g}}^j = T^{kj}g_{ki} = T_{ik}g^{kj} = T^m_{\bullet n}g_{mi}g^{nj} . \tag{2.43d}$$

Again, observe how the metric coefficients play a role similar to that of the Kronecker delta for orthonormal bases. For example, in Eq. (2.43a), the expression $T_{mn}g^{im}g^{jn}$ "becomes" T^{ij} as follows: upon seeing g^{im}, you look for a subscript i or m on T. Finding the subscript m, you replace it with an i and change its level from a subscript to a superscript. The metric coefficient g^{jn} similarly raises the subscript n on T_{mn} to a become a superscript j on T^{ij}.

Incidentally, if the $[F]$ matrix and lab components $[T]_{\text{lab}}$ are available for $\underset{\approx}{\mathbf{T}}$, then

$$[T^{ij}] = [F]^{-1}[T]_{\text{lab}}[F]^{-T},\ [T_{ij}] = [F]^T[T]_{\text{lab}}[F],\ [T^i_{\bullet j}] = [F]^{-1}[T]_{\text{lab}}[F],\text{ and } [T_i^{\bullet j}] = [F]^T[T]_{\text{lab}}[F]^{-T} \tag{2.44}$$

These are nifty quick-answer formulas, but for general discussions, Eqs. (2.43) are really more meaningful and those equations are applicable even when you *don't* have lab components.

Our final index changing properties involve the *mixed* Kronecker delta itself. The mixed Kronecker delta δ_i^j changes the index symbol *without* changing its level. For example,

$$v^i\delta_i^j = v^j \qquad v_i\delta_j^i = v_j \tag{2.45}$$

$$T^{ij}\delta_i^k = T^{kj} \qquad \delta_j^n T_i^{\bullet j}\delta_m^i = T_m^{\bullet n} \qquad \text{etc.} \tag{2.46}$$

Contrast this with the metric coefficients g_{ij} and g^{ij}, which change both the symbol and the level of the sub/superscripts.

Study Question 2.11 Let $\underset{\approx}{\mathbf{H}}$ denote a fourth-order tensor. Use the method of raising and lowering indices to simplify the following expressions so that there are no metric coefficients or Kronecker deltas (g^{ij}, g_{ij}, or δ_i^j) present.

(a) $H_{ijkl}g^{im}g^{jn}$ (b) $\delta_m^q H_{ipqn}g^{is}g^{nj}$ (c) $H^{i\bullet k\bullet}_{\bullet j\bullet l}g_{im}g^{jn}g^{mp}$

Multiply the following covariant, or mixed components by appropriate combinations of the metric coefficients to convert them all to pure contravariant components (*i.e.,* all superscripts).

(d) $H^{ij\bullet\bullet}_{\bullet\bullet kp}$ (e) H_{ijkl} (f) $H^{i\bullet k\bullet}_{\bullet j\bullet l}$

(g) Are we having fun yet?

Partial Answer: *(a) Seeing the g^{im}, you can get rid of it if you can find a subscript "i" or "m." Finding "i" as a subscript on H, you raise its level and change it to an "m," leaving a "dot" placeholder where the old subscript lived. After similarly eliminating the g^{jn}, the final simplified result is $H^{mn\bullet\bullet}_{\bullet\bullet kl}$. (b) $H^{s\bullet\bullet j}_{\bullet pm\bullet}$ (c) For the first and last "g" factor, you can lower the "i" to make it an "m" and then raise the "m" to make it a "p." Alternatively, you can recognize the opportunity to apply Eq. (2.20) so that $g_{im}g^{mp} = \delta_i^p$, and then Eq. (2.46) allows you to change the original superscript "i" on the H tensor to a "p" leaving its level unchanged. (d) The i and j are already on the upper level, so you leave them alone. To raise the "k" subscript, you multiply by g^{km}. The "p" is raised similarly so that the final result is $H^{ijmn} = H^{ij\bullet\bullet}_{\bullet\bullet kp}g^{km}g^{pn}$. (g) Yes, of course!*

3. Tensor algebra for general bases

A general basis is one that is not restricted to be orthogonal, normalized, or right-handed. Throughout this section, it is important to keep in mind that structured (direct) notation formulas remain unchanged. Here we show how the *component* formulas take on different forms for general bases that are permissibly nonorthogonal, non-normalized, and/or non-right-handed.

3.1 The vector inner ("dot") product for general bases.

Presuming that the contravariant components of two vectors $\underset{\sim}{\mathbf{a}}$ and $\underset{\sim}{\mathbf{b}}$ are known, we may expand the direct notation for the dot product as

$$\underset{\sim}{\mathbf{a}} \bullet \underset{\sim}{\mathbf{b}} = (a^i \underset{\sim}{\mathbf{g}}_i) \bullet (b^j \underset{\sim}{\mathbf{g}}_j) = a^i b^j (\underset{\sim}{\mathbf{g}}_i \bullet \underset{\sim}{\mathbf{g}}_j) \ . \tag{3.1}$$

Substituting the definition (2.11) into Eq. (3.1), the formula for the dot product becomes

$$\boxed{\underset{\sim}{\mathbf{a}} \bullet \underset{\sim}{\mathbf{b}} = a^i b^j g_{ij}} = a^1 b^1 g_{11} + a^1 b^2 g_{12} + a^1 b^3 g_{13} + a^2 b^1 g_{21} + \ldots + a^3 b^3 g_{33} \ . \tag{3.2}$$

This result can be written in other forms by raising and lowering indices. Noting, for example, that $b^j g_{ij} = b_i$, we obtain a much simpler formula for the dot product

$$\boxed{\underset{\sim}{\mathbf{a}} \bullet \underset{\sim}{\mathbf{b}} = a^i b_i} = a^1 b_1 + a^2 b_2 + a^3 b_3 \ . \tag{3.3}$$

The simplicity of this formula and its close similarity to the familiar formula from orthonormal theory is one of the principal reasons for introducing the dual bases. We can lower the index on a^i by writing it as $g^{ki} a_k$ so we obtain another formula for the dot product:

$$\boxed{\underset{\sim}{\mathbf{a}} \bullet \underset{\sim}{\mathbf{b}} = a_k b_i g^{ki}} = a_1 b_1 g^{11} + a_1 b_2 g^{12} + a_1 b_3 g^{13} + a_2 b_1 g^{21} + \ldots + a_3 b_3 g^{33} \ . \tag{3.4}$$

Finally, we recognize the index raising combination $b_i g^{ki} = b^k$, to obtain yet another formula:

$$\boxed{\underset{\sim}{\mathbf{a}} \bullet \underset{\sim}{\mathbf{b}} = a_i b^i} = a_1 b^1 + a_2 b^2 + a_3 b^3 \ . \tag{3.5}$$

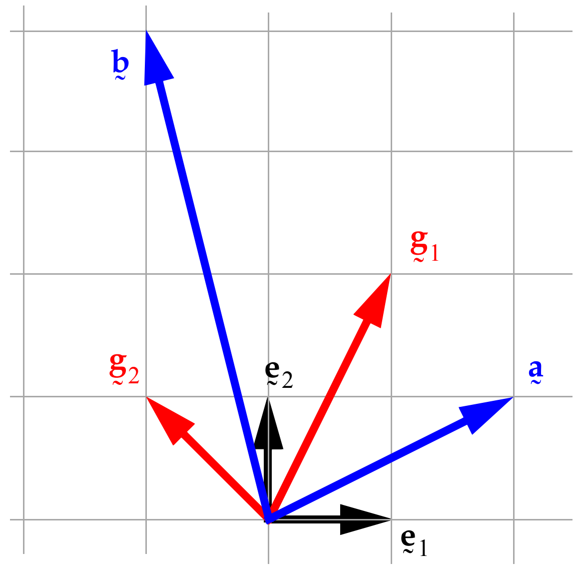
Whenever new results are derived for a general basis, it is always advisable to ensure that they all reduce to a familiar form whenever the basis is orthonormal. For the special case that the basis $\{\underset{\sim}{\mathbf{g}}_1, \underset{\sim}{\mathbf{g}}_2, \underset{\sim}{\mathbf{g}}_3\}$ happens to be orthonormal, we note that $g_{ij} = \delta_{ij}$. Furthermore, for an orthonormal basis, there is no difference between contravariant and covariant components. Thus Eqs. 3.2, 3.3, 3.4, and 3.5 do indeed all reduce to the usual orthonormal formula.

Study Question 3.1 Consider the following nonorthonormal base vectors:

$$\underset{\sim}{\mathbf{g}}_1 = \underset{\sim}{\mathbf{e}}_1 + 2\underset{\sim}{\mathbf{e}}_2$$

$$\underset{\sim}{\mathbf{g}}_2 = -\underset{\sim}{\mathbf{e}}_1 + \underset{\sim}{\mathbf{e}}_2 \qquad \text{and} \qquad \underset{\sim}{\mathbf{g}}_3 = \underset{\sim}{\mathbf{e}}_3$$

Consider two vectors, $\underset{\sim}{\mathbf{a}}$ and $\underset{\sim}{\mathbf{b}}$, lying in the 1-2 plane as drawn to scale in the figure.



(a) Express $\underset{\sim}{\mathbf{a}}$ and $\underset{\sim}{\mathbf{b}}$ in terms of the $\underset{\sim}{\mathbf{e}}$ lab basis.

(b) Compute $\underset{\sim}{\mathbf{a}} \bullet \underset{\sim}{\mathbf{b}}$ in the ordinary familiar manner by using lab components.

(c) Use Eqs. (3.2) and (3.3) to compute $\underset{\sim}{\mathbf{a}} \bullet \underset{\sim}{\mathbf{b}}$. Does the result agree with part (b)?

Partial Answer: *(a)* $\underset{\sim}{\mathbf{a}} = 2\underset{\sim}{\mathbf{e}}_1 + \underset{\sim}{\mathbf{e}}_2$ *(b)* $\underset{\sim}{\mathbf{a}} \bullet \underset{\sim}{\mathbf{b}} = 2$ *(c) yes, of course it agrees!*

Index contraction The vector dot product is a special case of a more general operation, which we will call index contraction. Given two free indices in an expression, we say that we "contract" them when they are turned into dummy summation indices according to the following rules:

1. if the indices started at the same level, multiply them by a metric coefficient

2. if the indices are at different levels, multiply by a Kronecker delta (and simplify).

In both cases, the indices on the metric coefficient or Kronecker delta must be the same symbol as the indices being contracted, but at different levels so that they become dummy summation symbols upon application of the summation conventions. The symbol C_1^2 denotes contraction of the first and second indices; C_2^8 would denote contraction of the second and eighth indices (assuming, of course, that the expression has eight free indices to start with!). Index contraction is *actually* basis contraction and it is the ordering of the basis that dictates the contraction.

Consider, for example, the expression $v_i W^{jk} Z_{lmn}$. This expression has six free indices, and we will suppose that it therefore corresponds to a sixth-order tensor with the associated base vectors being ordered the same as the free indices $\underset{\sim}{\mathbf{g}}^i \underset{\sim}{\mathbf{g}}_j \underset{\sim}{\mathbf{g}}_k \underset{\sim}{\mathbf{g}}^l \underset{\sim}{\mathbf{g}}^m \underset{\sim}{\mathbf{g}}^n$. To contract the first and second indices (i and j), which lie on different levels, we must dot the first and second base vectors into themselves, resulting in the Kronecker delta δ_j^i. Thus, contracting the first and second indices of $v_i W^{jk} Z_{lmn}$ gives $\delta^i_j v_i W^{jk} Z_{lmn}$, which simplifies to $v_i W^{ik} Z_{lmn}$. Note that the contraction operation has reduced the order of the result from a sixth-order tensor down to a fourth-order tensor. To compute the contraction C^6_4, we must dot the fourth base vector $\underset{\sim}{\mathbf{g}}^l$ into the sixth base vector $\underset{\sim}{\mathbf{g}}^n$, which results in g^{ln}. Thus, contracting the fourth and sixth indices of $v_i W^{jk} Z_{lmn}$ gives $g^{ln} v_i W^{jk} Z_{lmn}$. To summarize,

Index contraction is equivalent to dotting together the base vectors associated with the identified indices. C^β_α denotes contraction of the α^{th} and β^{th} base vectors.

No matter what basis is used to expand a tensor, the result of a contraction operation will be the same. In other words, the contraction operation is invariant under a change in basis. By this, we mean that you can apply the contract and then change the basis or vice versa — the result will be the same.

Incidentally, note that the vector dot product $\underset{\sim}{\mathbf{u}} \bullet \underset{\sim}{\mathbf{v}}$ can be expressed in terms of a contraction as C^2_1 operating on the *dyad* $\underset{\sim}{\mathbf{u}}\underset{\sim}{\mathbf{v}}$. The formal C^β_α notation for index contraction is rarely used, but the *phrase* "index contraction" is very common.

3.2 Other dot product operations

Using methods similar to those in Section 3.1, indicial forms of the operation $\underset{\approx}{\mathbf{T}} \bullet \underset{\sim}{\mathbf{v}}$ are found to be

$$\underset{\approx}{\mathbf{T}} \bullet \underset{\sim}{\mathbf{v}} = T^{ij} v_j \underset{\sim}{\mathbf{g}}_i = T_{ij} v^j \underset{\sim}{\mathbf{g}}^i = T^i_{\bullet j} v^j \underset{\sim}{\mathbf{g}}_i \quad \text{etc.}$$

In all these formulas, the dot product results in a summation between adjacent indices on opposite levels. If you know only T_{ij} and v_k, you must first raise an index to apply the dot product. For example,

$$\underset{\approx}{\mathbf{T}} \bullet \underset{\sim}{\mathbf{v}} = T_{ij} v_k g^{kj} \underset{\sim}{\mathbf{g}}^i \ . \tag{3.6}$$

In general, the correct indicial expression for any operation involving dot products can be derived by starting with the direct notation and expanding each argument in whatever form happens to be available. This approach typically leads to the opportunity to apply one or more of Eqs. (2.11), (2.12), or (2.19). For example, suppose you know two tensors in the forms $\underset{\approx}{\mathbf{A}} = A^{ij} \underset{\sim}{\mathbf{g}}_i \underset{\sim}{\mathbf{g}}_j$ and $\underset{\approx}{\mathbf{B}} = B^i_{\bullet j} \underset{\sim}{\mathbf{g}}_i \underset{\sim}{\mathbf{g}}^j$. Then

$$\underset{\approx}{\mathbf{A}} \bullet \underset{\approx}{\mathbf{B}} = (A^{ij} \underset{\sim}{\mathbf{g}}_i \underset{\sim}{\mathbf{g}}_j) \bullet (B^m_{\bullet n} \underset{\sim}{\mathbf{g}}_m \underset{\sim}{\mathbf{g}}^n) = A^{ij} B^m_{\bullet n} \underset{\sim}{\mathbf{g}}_i (\underset{\sim}{\mathbf{g}}_j \bullet \underset{\sim}{\mathbf{g}}_m) \underset{\sim}{\mathbf{g}}^n = A^{ij} B^m_{\bullet n} g_{jm} \underset{\sim}{\mathbf{g}}_i \underset{\sim}{\mathbf{g}}^n \ . \tag{3.7}$$

Note the change in dummy sum indices from *ij* to *mn* required to avoid violation of the summation conventions. For the final step, we merely applied Eq. (2.11). If desired, the above result may be simplified by using the g_{jm} to lower the "*j*" superscript on A^{ij} (changing it to an

"m") so that

$$\underset{\approx}{\mathbf{A}} \bullet \underset{\approx}{\mathbf{B}} = A^{i}_{\bullet m} B^{m}_{\bullet n} \underset{\sim}{\mathbf{g}}_i \underset{\sim}{\mathbf{g}}^n \ . \tag{3.8}$$

Alternatively, we could have used the g_{jm} to lower the "m" superscript on B (changing it to a "j") so that

$$\underset{\approx}{\mathbf{A}} \bullet \underset{\approx}{\mathbf{B}} = A^{ij} B_{jn} \underset{\sim}{\mathbf{g}}_i \underset{\sim}{\mathbf{g}}^n \ . \tag{3.9}$$

These are all valid expressions for the composition of two tensors. Note that the final result in the last three equations involved components times the basis dyad $\underset{\sim}{\mathbf{g}}_i \underset{\sim}{\mathbf{g}}^n$. Hence those components represent the mixed "${}^{i}_{\bullet n}$" components of $\underset{\approx}{\mathbf{A}} \bullet \underset{\approx}{\mathbf{B}}$.

Study Question 3.2 Complete the following table, which shows the indicial components for all of the sixteen possible ways to express the operation $\underset{\approx}{\mathbf{A}} \bullet \underset{\approx}{\mathbf{B}}$, depending on what type of components are available for $\underset{\approx}{\mathbf{A}}$ and $\underset{\approx}{\mathbf{B}}$. The shaded cells have the simplest form because they do not involve metric coefficients.

	$B_{nj} \underset{\sim}{\mathbf{g}}^n \underset{\sim}{\mathbf{g}}^j$	$B^{nj} \underset{\sim}{\mathbf{g}}_n \underset{\sim}{\mathbf{g}}_j$	$B^{n}_{\bullet j} \underset{\sim}{\mathbf{g}}_n \underset{\sim}{\mathbf{g}}^j$	$B_{n}^{\bullet j} \underset{\sim}{\mathbf{g}}^n \underset{\sim}{\mathbf{g}}_j$
$A_{im} \underset{\sim}{\mathbf{g}}^i \underset{\sim}{\mathbf{g}}^m$	$A_{im} B_{nj} g^{mn} \underset{\sim}{\mathbf{g}}^i \underset{\sim}{\mathbf{g}}^j$	$A_{im} B^{mj} \underset{\sim}{\mathbf{g}}^i \underset{\sim}{\mathbf{g}}_j$		$A_{im} B_{n}^{\bullet j} g^{mn} \underset{\sim}{\mathbf{g}}^i \underset{\sim}{\mathbf{g}}_j$
$A^{im} \underset{\sim}{\mathbf{g}}_i \underset{\sim}{\mathbf{g}}_m$	$A^{im} B_{mj} \underset{\sim}{\mathbf{g}}_i \underset{\sim}{\mathbf{g}}^j$		$A^{im} B^{n}_{\bullet j} g_{mn} \underset{\sim}{\mathbf{g}}_i \underset{\sim}{\mathbf{g}}^j$	
$A^{i}_{\bullet m} \underset{\sim}{\mathbf{g}}_i \underset{\sim}{\mathbf{g}}^m$			$A^{i}_{\bullet m} B^{m}_{\bullet j} \underset{\sim}{\mathbf{g}}_i \underset{\sim}{\mathbf{g}}^j$	$A^{i}_{\bullet m} B_{n}^{\bullet j} g^{mn} \underset{\sim}{\mathbf{g}}_i \underset{\sim}{\mathbf{g}}_j$
$A_{i}^{\bullet m} \underset{\sim}{\mathbf{g}}^i \underset{\sim}{\mathbf{g}}_m$	$A_{i}^{\bullet m} B_{mj} \underset{\sim}{\mathbf{g}}^i \underset{\sim}{\mathbf{g}}^j$	$A_{i}^{\bullet m} B^{nj} g_{mn} \underset{\sim}{\mathbf{g}}^i \underset{\sim}{\mathbf{g}}_j$		

3.3 The transpose operation

The transpose $\underset{\approx}{\mathbf{B}}^T$ of a tensor $\underset{\approx}{\mathbf{B}}$ is defined in direct notation such that

$$\underset{\sim}{\mathbf{u}} \bullet \underset{\approx}{\mathbf{B}}^T \bullet \underset{\sim}{\mathbf{v}} = \underset{\sim}{\mathbf{v}} \bullet \underset{\approx}{\mathbf{B}} \bullet \underset{\sim}{\mathbf{u}} \text{ for all vectors } \underset{\sim}{\mathbf{u}} \text{ and } \underset{\sim}{\mathbf{v}} \ . \tag{3.10}$$

Since this must hold for all vectors, it must hold for any particular choice. Taking $\underset{\sim}{\mathbf{u}}=\underset{\sim}{\mathbf{g}}_i$ and $\underset{\sim}{\mathbf{v}}=\underset{\sim}{\mathbf{g}}_j$, we see that

$$(\underset{\approx}{\mathbf{B}}^T)_{ij} = B_{ji} \ . \tag{3.11}$$

Similarly, you can show that

$$(\underset{\approx}{\mathbf{B}}^T)^{ij} = B^{ji} \ . \tag{3.12}$$

The mixed components of the transpose are more different. Namely, taking $\underset{\sim}{\mathbf{u}}=\underset{\sim}{\mathbf{g}}_i$ and $\underset{\sim}{\mathbf{v}}=\underset{\sim}{\mathbf{g}}^j$,

$$(\underset{\approx}{\mathbf{B}}^T)_i^{\bullet j} = B^j_{\bullet i} \,. \tag{3.13}$$

Note that the high/low level of the indices remains unchanged. Only their ordering is switched. Similarly,

$$(\underset{\approx}{\mathbf{B}}^T)^i_{\bullet j} = B_j^{\bullet i} \,. \tag{3.14}$$

All of the above relations could have been derived in the following alternative manner: The transpose $(\underset{\sim}{\mathbf{u}}\underset{\sim}{\mathbf{v}})^T$ of any dyad is simply $\underset{\sim}{\mathbf{v}}\underset{\sim}{\mathbf{u}}$. Therefore, knowing that the transpose of a sum is the sum of the transposes and that any tensor can be written as a sum of dyads, we can write:

$$\underset{\approx}{\mathbf{B}}^T = (B_{ij}\underset{\sim}{\mathbf{g}}^i\underset{\sim}{\mathbf{g}}^j)^T = B_{ij}(\underset{\sim}{\mathbf{g}}^i\underset{\sim}{\mathbf{g}}^j)^T = B_{ij}\underset{\sim}{\mathbf{g}}^j\underset{\sim}{\mathbf{g}}^i \,. \tag{3.15}$$

The coefficient of $\underset{\sim}{\mathbf{g}}^i\underset{\sim}{\mathbf{g}}^j$ is B_{ij}. Thus, $(\underset{\approx}{\mathbf{B}}^T)_{ji} = B_{ij}$, which is the same as Eq. (3.11).

3.4 Symmetric Tensors

A tensor $\underset{\approx}{\mathbf{A}}$ is symmetric only if it equals its own transpose. Therefore, referring to Eqs. (3.11) through (3.15) the components of a symmetric tensor satisfy

$$A_{ij} = A_{ji}, \quad A^{ij} = A^{ji}, \quad A^i_{\bullet j} = A_j^{\bullet i} \,. \tag{3.16}$$

The last relationship shows that the "dot placeholder" is unnecessary for symmetric tensors, and we may write simply A_i^j without ambiguity.

3.5 The identity tensor for a general basis.

The identity tensor $\underset{\approx}{\mathbf{I}}$ is the unique symmetric tensor for which

$$\underset{\sim}{\mathbf{u}} \bullet \underset{\approx}{\mathbf{I}} \bullet \underset{\sim}{\mathbf{v}} = \underset{\sim}{\mathbf{u}} \bullet \underset{\sim}{\mathbf{v}} \text{ for all vectors } \underset{\sim}{\mathbf{u}} \text{ and } \underset{\sim}{\mathbf{v}} \,. \tag{3.17}$$

Since this must hold for all vectors, it must hold for any particular choice. Taking $\underset{\sim}{\mathbf{u}}=\underset{\sim}{\mathbf{g}}_i$ and $\underset{\sim}{\mathbf{v}}=\underset{\sim}{\mathbf{g}}_j$, we find that $\underset{\sim}{\mathbf{g}}_i \bullet \underset{\approx}{\mathbf{I}} \bullet \underset{\sim}{\mathbf{g}}_j = \underset{\sim}{\mathbf{g}}_i \bullet \underset{\sim}{\mathbf{g}}_j$. The left-hand side represents the *ij* covariant components of the identity, and the right hand side is g_{ij}. By taking $\underset{\sim}{\mathbf{u}}=\underset{\sim}{\mathbf{g}}^i$ and $\underset{\sim}{\mathbf{u}}=\underset{\sim}{\mathbf{g}}^j$, it can be similarly shown that the contravariant components of the identity are g^{ij}. By taking $\underset{\sim}{\mathbf{u}}=\underset{\sim}{\mathbf{g}}^i$ and $\underset{\sim}{\mathbf{u}}=\underset{\sim}{\mathbf{g}}_j$, the mixed components of the identity are found to equal δ_i^j. Thus,

$$\boxed{\underset{\approx}{\mathbf{I}} = g^{ij}\underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}_j = g_{ij}\underset{\sim}{\mathbf{g}}^i\underset{\sim}{\mathbf{g}}^j = \delta_i^j\underset{\sim}{\mathbf{g}}^i\underset{\sim}{\mathbf{g}}_j = \delta_j^i\underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}^j} \tag{3.18}$$

This result further validates raising and lowering indices by multiplying by appropriate metric coefficients — such an operation merely represents dotting by the identity tensor.

3.6 Eigenproblems and similarity transformations

The **eigenproblem** for a tensor $\underset{\approx}{\mathbf{T}}$ requires the determination of all eigenvectors $\underset{\sim}{\mathbf{u}}$ and corresponding eigenvalues λ such that

$$\underset{\approx}{\mathbf{T}} \bullet \underset{\sim}{\mathbf{u}} = \lambda \underset{\sim}{\mathbf{u}} \ . \tag{3.19}$$

In this form, the eigenvector $\underset{\sim}{\mathbf{u}}$ is called the **"right" eigenvector**. We can also define a **"left" eigenvector** $\underset{\sim}{\mathbf{v}}$ such that

$$\underset{\sim}{\mathbf{v}} \bullet \underset{\approx}{\mathbf{T}} = \lambda \underset{\sim}{\mathbf{v}} \ . \tag{3.20}$$

In other words, the left eigenvectors are the right eigenvectors of $\underset{\approx}{\mathbf{T}}^T$. The characteristic equation for $\underset{\approx}{\mathbf{T}}^T$ is the same as that for $\underset{\approx}{\mathbf{T}}$, so the eigenvalues are the same for both the right and left eigenproblems. Consider a particular eigenpair $(\lambda_k, \underset{\sim}{\mathbf{u}}_k)$:

$$\underset{\approx}{\mathbf{T}} \bullet \underset{\sim}{\mathbf{u}}_k = \lambda_k \underset{\sim}{\mathbf{u}}_k \ . \qquad \text{(no sum on } k\text{)} \tag{3.21}$$

Dot from the left by a left eigenvector $\underset{\sim}{\mathbf{v}}_m$, noting that $\underset{\sim}{\mathbf{v}}_m \bullet \underset{\approx}{\mathbf{T}} = \lambda_m \underset{\sim}{\mathbf{v}}_m$ (no sum on m). Then

$$\lambda_m \underset{\sim}{\mathbf{v}}_m \bullet \underset{\sim}{\mathbf{u}}_k = \lambda_k \underset{\sim}{\mathbf{v}}_m \bullet \underset{\sim}{\mathbf{u}}_k \ . \qquad \text{(no sum on } k\text{)} \tag{3.22}$$

Rearranging,

$$(\lambda_m - \lambda_k)(\underset{\sim}{\mathbf{v}}_m \bullet \underset{\sim}{\mathbf{u}}_k) = 0 \ . \qquad \text{(no sums)} \tag{3.23}$$

From this result we conclude that the left and right eigenvectors corresponding distinct $(\lambda_m \neq \lambda_k)$ eigenvalues are orthogonal. This motivates **renaming the eigenvectors using dual basis notation**:

$$\text{Rename } \underset{\sim}{\mathbf{u}}_k = \underset{\sim}{\mathbf{p}}_k .$$
$$\text{Rename } \underset{\sim}{\mathbf{v}}_m = \underset{\sim}{\mathbf{p}}^m . \tag{3.24}$$

The magnitudes of eigenvectors are arbitrary, so we can select normalization such that the renamed vectors are truly dual bases:

$$\underset{\sim}{\mathbf{p}}^m \bullet \underset{\sim}{\mathbf{p}}_k = \delta_k^m \ . \tag{3.25}$$

Nonsymmetric tensors might not possess a complete set of eigenvectors (i.e., the geometric multiplicity of eigenvalues might be less than their algebraic multiplicity). If, however, the tensor $\underset{\approx}{\mathbf{T}}$ happens to possess a complete (or "**spanning**") set of eigenvectors, then those eigenvectors form an acceptable basis. The mixed components of $\underset{\approx}{\mathbf{T}}$ with respect to this principal basis are then

$$T^i_{\bullet j} = \underset{\sim}{\mathbf{p}}^i \bullet \underset{\approx}{\mathbf{T}} \bullet \underset{\sim}{\mathbf{p}}_j = \underset{\sim}{\mathbf{p}}^i \bullet (\lambda_j \underset{\sim}{\mathbf{p}}_j) = \lambda_j \delta^i_j \ . \qquad \text{(no sum on } j\text{)} \tag{3.26}$$

Ah! Whenever a tensor $\underset{\approx}{\mathbf{T}}$ possesses a complete set of eigenvectors, it is diagonal in its *mixed* principal basis! Stated differently,

$$[T^i_{\bullet j}] = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}_{\underset{\sim}{\mathbf{p}}_i \underset{\sim}{\mathbf{p}}^j} . \tag{3.27}$$

or

$$\underset{\approx}{\mathbf{T}} = \sum_{k=1}^{3} \lambda_k \underset{\sim}{\mathbf{p}}_k \underset{\sim}{\mathbf{p}}^k . \tag{3.28}$$

In matrix analysis books, this result is usually presented as a similarity transformation. As was done in Eqs. (2.2) and (2.32), we can invoke the existence of a basis transformation tensor $\underset{\approx}{\mathbf{F}}$, such that

$$\underset{\sim}{\mathbf{p}}_k = \underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{e}}_k \text{ and } \underset{\sim}{\mathbf{p}}^k = \underset{\approx}{\mathbf{F}}^{-T} \bullet \underset{\sim}{\mathbf{e}}^k . \tag{3.29}$$

The columns of the matrix of $\underset{\approx}{\mathbf{F}}$ with respect to the laboratory $\{\underset{\sim}{\mathbf{e}}_1, \underset{\sim}{\mathbf{e}}_2, \underset{\sim}{\mathbf{e}}_3\}$ basis are simply the right eigenvectors expressed in the lab basis:

$$[\underset{\approx}{\mathbf{F}}]_{\underset{\sim}{\mathbf{e}}\underset{\sim}{\mathbf{e}}} = [\{\underset{\sim}{\mathbf{p}}_1\}_{\underset{\sim}{\mathbf{e}}}\{\underset{\sim}{\mathbf{p}}_2\}_{\underset{\sim}{\mathbf{e}}}\{\underset{\sim}{\mathbf{p}}_3\}_{\underset{\sim}{\mathbf{e}}}] . \tag{3.30}$$

With Eq. (3.29), the diagonalization result (3.28) can be written as similarity transformation. Namely,

$$\underset{\approx}{\mathbf{T}} = \sum_{k=1}^{3} \lambda_k (\underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{e}}_k)(\underset{\approx}{\mathbf{F}}^{-T} \bullet \underset{\sim}{\mathbf{e}}^k) = \underset{\approx}{\mathbf{F}} \bullet \underset{\approx}{\Lambda} \bullet \underset{\approx}{\mathbf{F}}^{-1} , \tag{3.31}$$

where

$$\underset{\approx}{\Lambda} \equiv \sum_{k=1}^{3} \lambda_k \underset{\sim}{\mathbf{e}}_k \underset{\sim}{\mathbf{e}}^k \rightarrow \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}_{\underset{\sim}{\mathbf{e}}_k \underset{\sim}{\mathbf{e}}^k} . \tag{3.32}$$

Study Question 3.3 Consider a tensor having components with respect to the orthonormal laboratory basis given by

$$[\underset{\approx}{\mathbf{T}}] = \begin{bmatrix} 2 & 1 & -2 \\ 1 & 2 & -2 \\ 0 & 0 & 2 \end{bmatrix}_{\underset{\sim}{\mathbf{e}}_k \underset{\sim}{\mathbf{e}}^k} . \tag{3.33}$$

Prove that this tensor has a spanning set of eigenvectors $\{\underset{\sim}{\mathbf{p}}_1, \underset{\sim}{\mathbf{p}}_2, \underset{\sim}{\mathbf{p}}_3\}$. Find the laboratory components of the basis transformation tensor $\underset{\approx}{\mathbf{F}}$, such that $\underset{\sim}{\mathbf{p}}_k = \underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{e}}_k$. Also verify Eq. (3.31) that $\underset{\approx}{\mathbf{T}}$ is similar to a tensor that is diagonal in the laboratory basis, with the diagonal components being equal to the eigenvalues of $\underset{\approx}{\mathbf{T}}$.

Partial Answer: $[\underset{\approx}{\mathbf{F}}] = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 2 & 1 \\ 0 & 1 & 0 \end{bmatrix}$.

The above discussion focused on eigenproblems. We now mention the more general relationship between similar components and similarity transformations. When two tensors $\underset{\approx}{\mathbf{T}}$ and $\underset{\approx}{\mathbf{S}}$ possess the same components but with respect to different *mixed* bases, i.e., when

$$\underset{\approx}{\mathbf{T}} = \alpha^{i}_{\bullet j} \underset{\sim}{\mathbf{g}}_i \underset{\sim}{\mathbf{g}}^j \quad \text{and} \quad \underset{\approx}{\mathbf{S}} = \alpha^{i}_{\bullet j} \underset{\sim}{\mathbf{G}}_i \underset{\sim}{\mathbf{G}}^j , \quad \text{(identical } \textit{mixed} \text{ components)} \tag{3.34}$$

then the tensors are similar. In other words, there exists a transformation tensor $\underset{\approx}{\mathbf{F}}$ such that

$$\underset{\sim}{\mathbf{g}}_k = \underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{G}}_k \quad \text{and therefore} \quad \underset{\approx}{\mathbf{T}} = \underset{\approx}{\mathbf{F}} \bullet \underset{\approx}{\mathbf{S}} \bullet \underset{\approx}{\mathbf{F}}^{-1} , \tag{3.35}$$

In the context of continuum mechanics, the transformation tensor $\underset{\approx}{\mathbf{F}}$ is typically the deformation gradient tensor. It is important to have at least a vague notion of this concept in order to communicate effectively with researchers who prefer to do all analyses in general curvilinear coordinates. To them, the discovery that two tensors have identical mixed components with respect to different bases has great significance, whereas you might find it more meaningful to recognize this situation as merely a similarity transformation.

Study Question 3.4 Suppose two tensors $\underset{\approx}{\mathbf{T}}$ and $\underset{\approx}{\mathbf{S}}$ possess the *same* contravariant components with respect to *different* bases. In other words,

$$\underset{\approx}{\mathbf{T}} = \beta^{ij}\underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}_j \quad \text{and} \quad \underset{\approx}{\mathbf{S}} = \beta^{ij}\underset{\sim}{\mathbf{G}}_i\underset{\sim}{\mathbf{G}}_j \quad \text{(same contravariant components)} \tag{3.36}$$

Demonstrate by direct substitution that

$$\underset{\approx}{\mathbf{T}} = \underset{\approx}{\mathbf{F}} \bullet \underset{\approx}{\mathbf{S}} \bullet \underset{\approx}{\mathbf{F}}^T , \tag{3.37}$$

where $\underset{\approx}{\mathbf{F}}$ is a basis transformation tensor defined such that

$$\underset{\sim}{\mathbf{g}}_k = \underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{G}}_k . \tag{3.38}$$

3.7 The alternating tensor

When using a nonorthonormal or non-right-handed basis, it is a good idea to use a different symbol $\underset{\approx\!\approx}{\xi}$ for the alternating tensor so that the permutation symbol ε_{ijk} may retain its usual meaning. As we did for the Kronecker delta, we will assign the same meaning to the **permutation symbol** regardless of the contra/covariant level of the indices. Namely,

$$\varepsilon_{ijk},\ \varepsilon^{ijk},\ \varepsilon_{ij\bullet}^{\bullet\bullet k},\ \text{etc.} = \begin{cases} +1 \ \text{if}\ ijk = 123, 231, 312 \\ -1 \ \text{if}\ ijk = 321, 213, 132 \\ 0 \quad \text{otherwise} \end{cases} . \tag{3.39}$$

Important: These quantities are all defined the same regardless of the contra- or co- level at which the indices are placed. The above defined permutation symbols are the components of the alternating tensor with respect to any regular *right-handed* orthonormal laboratory basis, so they do not transform co/contravariant level via the metric tensors. It is not allowed to raise or lower indices on the ordinary permutation symbol with the metric tensors. A similar situation was encountered in connection with the Kronecker delta of Eq. (2.10).

Through the use of the permutation symbol, the three formulas written explicitly in Eq. (2.15) can be expressed compactly as a *single* indicial equation:

$$\varepsilon_{ijm}\underset{\sim}{\mathbf{g}}^m = \frac{1}{J}(\underset{\sim}{\mathbf{g}}_i \times \underset{\sim}{\mathbf{g}}_j) . \tag{3.40}$$

In terms of an orthonormal right-handed basis, the alternating tensor is

$$\underset{\approx\!\approx}{\xi} = \varepsilon_{ijk}\underset{\sim}{\mathbf{e}}^i\underset{\sim}{\mathbf{e}}^j\underset{\sim}{\mathbf{e}}^k . \tag{3.41}$$

In terms of a general basis, the basis form for the alternating tensor is

$$\underset{\approx}{\xi} = \xi_{ijk}\underset{\sim}{\mathbf{g}}^i\underset{\sim}{\mathbf{g}}^j\underset{\sim}{\mathbf{g}}^k = \xi^{ijk}\underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}_j\underset{\sim}{\mathbf{g}}_k = \xi_{ij\bullet}^{\bullet\bullet k}\underset{\sim}{\mathbf{g}}^i\underset{\sim}{\mathbf{g}}^j\underset{\sim}{\mathbf{g}}_k = \text{etc.}, \tag{3.42}$$

where

$$\xi_{ijk} = [\underset{\sim}{\mathbf{g}}_i, \underset{\sim}{\mathbf{g}}_j, \underset{\sim}{\mathbf{g}}_k], \quad \xi^{ijk} = [\underset{\sim}{\mathbf{g}}^i, \underset{\sim}{\mathbf{g}}^j, \underset{\sim}{\mathbf{g}}^k], \quad \xi_{ij\bullet}^{\bullet\bullet k} = [\underset{\sim}{\mathbf{g}}_i, \underset{\sim}{\mathbf{g}}_j, \underset{\sim}{\mathbf{g}}^k], \quad \text{etc.} \tag{3.43}$$

Using Eq. (3.40),

$$\xi_{ijk} = (\underset{\sim}{\mathbf{g}}_i \times \underset{\sim}{\mathbf{g}}_j) \bullet \underset{\sim}{\mathbf{g}}_k = (J\varepsilon_{ijm}\underset{\sim}{\mathbf{g}}^m) \bullet \underset{\sim}{\mathbf{g}}_k = J\varepsilon_{ijm}\delta_k^m . \tag{3.44}$$

Thus, simplifying the last expression,

$$\boxed{\xi_{ijk} = J\varepsilon_{ijk}} . \tag{3.45}$$

Hence, the covariant alternating tensor components simply equal to the permutation symbol times the Jacobian. This result could have been derived in the following alternative way: Substituting Eq. (2.1) into Eq. (3.43) and using the direct notation definition of determinant gives

$$\xi_{ijk} = [\underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{e}}_i, \underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{e}}_j, \underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{e}}_k] = J[\underset{\sim}{\mathbf{e}}_i, \underset{\sim}{\mathbf{e}}_j, \underset{\sim}{\mathbf{e}}_k] = J\varepsilon_{ijk} . \tag{3.46}$$

Study Question 3.5 Use Eq. (2.32), in Eq. (3.43) to prove that

$$\xi^{ijk} = \frac{1}{J}\varepsilon^{ijk} \tag{3.47}$$

Study Question 3.6 Consider a basis that is orthonormal but left-handed (see Eq. 2.8). Prove that

$$\xi^{ijk} = -\varepsilon^{ijk} . \tag{3.48}$$

This is why some textbooks claim to "define" the permutation symbol to be its negative when the basis is left-handed. We do *not* adopt this practice. We keep the permutation *symbol* unchanged. For a left-handed basis, the permutation symbol stays the same, but the components of the alternating *tensor* change sign.

3.8 Vector valued operations

CROSS PRODUCT In direct notation, the cross product is defined

$$\underset{\sim}{\mathbf{u}} \times \underset{\sim}{\mathbf{v}} = \underset{\approx}{\xi} : \underset{\sim}{\mathbf{u}} \underset{\sim}{\mathbf{v}} \; . \tag{3.49}$$

Hence,

$$\underset{\sim}{\mathbf{u}} \times \underset{\sim}{\mathbf{v}} = \xi_{ijk} u^j v^k \underset{\sim}{\mathbf{g}}^i \; . \tag{3.50}$$

Alternatively,

$$\underset{\sim}{\mathbf{u}} \times \underset{\sim}{\mathbf{v}} = \xi^{ijk} u_j v_k \underset{\sim}{\mathbf{g}}_i \; . \tag{3.51}$$

Axial vectors Some textbooks choose to identify certain vectors such as angular velocity as "different" because, according to these books, they take on different form in a left-handed basis. This viewpoint is objectionable since it intimates that physical quantities are affected by the choice of bases. Those textbooks handle these supposedly different "axial" vectors by redefining the left-handed cross-product to be negative of the right-hand definition. Malvern suggests alternatively redefining the permutation symbol to be its negative for left-hand basis. Malvern's suggested sign change is handled automatically by defining the alternating tensor as we have above. Specifically Eqs. (3.44) and (2.8) show that the alternating tensor components ξ_{ijk} will automatically change sign for left-handed systems. Hence, with this approach, there is no need for special formulas for left-hand bases.

3.9 Scalar valued operations

TENSOR INNER (double dot) PRODUCT The inner product between two dyads, $\underset{\sim}{\mathbf{a}}\underset{\sim}{\mathbf{b}}$ and $\underset{\sim}{\mathbf{r}}\underset{\sim}{\mathbf{s}}$ is a scalar defined

$$(\underset{\sim}{\mathbf{a}}\underset{\sim}{\mathbf{b}}):(\underset{\sim}{\mathbf{r}}\underset{\sim}{\mathbf{s}}) \equiv (\underset{\sim}{\mathbf{a}} \bullet \underset{\sim}{\mathbf{r}})(\underset{\sim}{\mathbf{b}} \bullet \underset{\sim}{\mathbf{s}}) \tag{3.52}$$

When applied to base vectors, this result tells us that

$$\begin{aligned}
&(\underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}_j):(\underset{\sim}{\mathbf{g}}_m\underset{\sim}{\mathbf{g}}_n) = g_{im}g_{jn} && (\underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}_j):(\underset{\sim}{\mathbf{g}}^m\underset{\sim}{\mathbf{g}}^n) = \delta_i^m\delta_j^n \\
&(\underset{\sim}{\mathbf{g}}^i\underset{\sim}{\mathbf{g}}^j):(\underset{\sim}{\mathbf{g}}_m\underset{\sim}{\mathbf{g}}_n) = \delta_m^i\delta_n^j && (\underset{\sim}{\mathbf{g}}^i\underset{\sim}{\mathbf{g}}^j):(\underset{\sim}{\mathbf{g}}^m\underset{\sim}{\mathbf{g}}^n) = g^{im}g^{jn} \\
&(\underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}^j):(\underset{\sim}{\mathbf{g}}_m\underset{\sim}{\mathbf{g}}_n) = g_{im}\delta_n^j && (\underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}^j):(\underset{\sim}{\mathbf{g}}^m\underset{\sim}{\mathbf{g}}^n) = \delta_i^m g^{jn} \\
&\text{etc.}
\end{aligned} \tag{3.53}$$

The "etc" stands for the many other ways we could possibly mix up the level of the indices; the basic trend should be clear.

The inner product between two tensors, $\underset{\approx}{\mathbf{A}}$ and $\underset{\approx}{\mathbf{B}}$, is defined to be distributive over addition. In other words, each tensor can be expanded in terms of known components times base vectors and then Eq. (3.53) can be applied to the base vectors. If, for example, the mixed $A^i_{\bullet j}$

components of $\underset{\approx}{\mathbf{A}}$ and the contravariant B^{mn} components of $\underset{\approx}{\mathbf{B}}$ are known, then the inner product between $\underset{\approx}{\mathbf{A}}$ and $\underset{\approx}{\mathbf{B}}$ is

$$\underset{\approx}{\mathbf{A}}:\underset{\approx}{\mathbf{B}} = (A^{i}_{\bullet j}\underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}^j):(B^{mn}\underset{\sim}{\mathbf{g}}_m\underset{\sim}{\mathbf{g}}_n) = A^{i}_{\bullet j}B^{mn}(\underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}^j):(\underset{\sim}{\mathbf{g}}_m\underset{\sim}{\mathbf{g}}_n) = A^{i}_{\bullet j}B^{mn}g_{im}\delta^j_n = A^{i}_{\bullet j}B^{mj}g_{im} \tag{3.54}$$

Stated differently, to find the inner product between two tensors, $\underset{\approx}{\mathbf{A}}$ and $\underset{\approx}{\mathbf{B}}$, you should contract[1] the adjacent indices *pairwise* so that the first index in $\underset{\approx}{\mathbf{A}}$ is contracted with the first index in $\underset{\approx}{\mathbf{B}}$ and the second index in $\underset{\approx}{\mathbf{A}}$ is contracted with the second index in $\underset{\approx}{\mathbf{B}}$. To clarify, here are some expressions for the tensor inner product for various possibilities of which components of each tensor are available:

$$\underset{\approx}{\mathbf{A}}:\underset{\approx}{\mathbf{B}} = A_{ij}B^{ij} = A^{ij}B_{ij} = A^{i}_{\bullet j}B^{mj}g_{im} = A^{i}_{\bullet j}B_{i}^{\bullet j} = A_{ij}B_{mn}g^{im}g^{jn} = \dots \tag{3.55}$$

Note that

$$\underset{\approx}{\mathbf{A}}:\underset{\approx}{\mathbf{B}} = \mathrm{tr}(\underset{\approx}{\mathbf{A}}^T \bullet \underset{\approx}{\mathbf{B}}) = \mathrm{tr}(\underset{\approx}{\mathbf{A}} \bullet \underset{\approx}{\mathbf{B}}^T) \tag{3.56}$$

where "tr" is the trace operation.

TENSOR MAGNITUDE The magnitude of a tensor $\underset{\approx}{\mathbf{A}}$ is defined

$$\|\underset{\approx}{\mathbf{A}}\| = \sqrt{\underset{\approx}{\mathbf{A}}:\underset{\approx}{\mathbf{A}}} \tag{3.57}$$

Based on the result of Eq. (3.55), note that the magnitude of a tensor is *not* found by simply summing the squares of the tensor components (and rooting the result). Instead, the tensor magnitude is computed by rooting the summation running over each component of $\underset{\approx}{\mathbf{A}}$ multiplied by its *dual* component. If, for example, the components $A^{i}_{\bullet j}$ are known, then the dual components $A_{i}^{\bullet j}$ must be computed by $A_{i}^{\bullet j} = g_{im}A^{m}_{\bullet n}g^{jn}$ so that

$$\underset{\approx}{\mathbf{A}}:\underset{\approx}{\mathbf{A}} = A^{i}_{\bullet j}A_{i}^{\bullet j} = (A^{i}_{\bullet j}A^{m}_{\bullet n})g_{im}g^{jn} \tag{3.58}$$

TRACE The trace an operation in which the two base vectors of a second order tensor are contracted[2] together, resulting in a scalar. In direct notation, the trace of a tensor $\underset{\approx}{\mathbf{B}}$ can be defined $\mathrm{tr}\underset{\approx}{\mathbf{B}} = \underset{\approx}{\mathbf{I}}:\underset{\approx}{\mathbf{B}}$. There are four ways to write the tensor $\underset{\approx}{\mathbf{B}}$:

$$\underset{\approx}{\mathbf{B}} = B_{ij}\underset{\sim}{\mathbf{g}}^i\underset{\sim}{\mathbf{g}}^j = B^{ij}\underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}_j = B^{i}_{\bullet j}\underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}^j = B_{i}^{\bullet j}\underset{\sim}{\mathbf{g}}^i\underset{\sim}{\mathbf{g}}_j . \tag{3.59}$$

The double dot operation is distributive over addition. Furthermore, for any vectors $\underset{\sim}{\mathbf{u}}$ and $\underset{\sim}{\mathbf{v}}$, $\underset{\approx}{\mathbf{I}}:(\underset{\sim}{\mathbf{u}}\underset{\sim}{\mathbf{v}}) = \underset{\sim}{\mathbf{u}} \bullet \underset{\sim}{\mathbf{v}}$. Therefore, the trace is found by contracting the base vectors to give

$$\mathrm{tr}\underset{\approx}{\mathbf{B}} = B_{ij}\underset{\sim}{\mathbf{g}}^i \bullet \underset{\sim}{\mathbf{g}}^j = B^{ij}\underset{\sim}{\mathbf{g}}_i \bullet \underset{\sim}{\mathbf{g}}_j = B^{i}_{\bullet j}\underset{\sim}{\mathbf{g}}_i \bullet \underset{\sim}{\mathbf{g}}^j = B_{i}^{\bullet j}\underset{\sim}{\mathbf{g}}^i \bullet \underset{\sim}{\mathbf{g}}_j , \tag{3.60}$$

1. The definition of index "contraction" is given on page 31.
2. The definition of index "contraction" is given on page 31.

or

$$\mathrm{tr}\,\underset{\approx}{\mathbf{B}} = B_{ij}g^{ij} = B^{ij}g_{ij} = B^{k}_{\bullet k} = B_{k}^{\bullet k} \ . \tag{3.61}$$

Note that the last two expressions are the closest in form to the familiar formula for the trace in orthonormal bases.

TRIPLE SCALAR-VALUED VECTOR PRODUCT In direct notation,

$$[\underset{\sim}{\mathbf{u}}, \underset{\sim}{\mathbf{v}}, \underset{\sim}{\mathbf{w}}] \equiv \underset{\sim}{\mathbf{u}} \bullet (\underset{\sim}{\mathbf{v}} \times \underset{\sim}{\mathbf{w}}) \ . \tag{3.62}$$

Thus, applying Eq. (3.50),

$$[\underset{\sim}{\mathbf{u}}, \underset{\sim}{\mathbf{v}}, \underset{\sim}{\mathbf{w}}] = \xi_{ijk}u^i v^j w^k \ . \tag{3.63}$$

DETERMINANT The determinant of a tensor $\underset{\approx}{\mathbf{T}}$ is defined in direct notation as

$$[\underset{\approx}{\mathbf{T}} \bullet \underset{\sim}{\mathbf{u}}, \underset{\approx}{\mathbf{T}} \bullet \underset{\sim}{\mathbf{v}}, \underset{\approx}{\mathbf{T}} \bullet \underset{\sim}{\mathbf{w}}] = \det\underset{\approx}{\mathbf{T}}[\underset{\sim}{\mathbf{u}}, \underset{\sim}{\mathbf{v}}, \underset{\sim}{\mathbf{w}}] \quad \text{for all vectors} \quad \{\underset{\sim}{\mathbf{u}}, \underset{\sim}{\mathbf{v}}, \underset{\sim}{\mathbf{w}}\} \ . \tag{3.64}$$

Recalling that the triple scalar product is defined $[\underset{\sim}{\mathbf{u}}, \underset{\sim}{\mathbf{v}}, \underset{\sim}{\mathbf{w}}] = \underset{\sim}{\mathbf{u}} \bullet (\underset{\sim}{\mathbf{v}} \times \underset{\sim}{\mathbf{w}})$ we may apply previous formulas for the dot and cross product to conclude that

$$\xi_{ijk}T^{ip}T^{jq}T^{kr} = (\det\underset{\approx}{\mathbf{T}})\xi^{pqr} \ , \tag{3.65}$$

Recalling Eqs. (3.44) and (3.47), we may introduce the ordinary permutation symbol to write this result as

$$\varepsilon_{ijk}T^{ip}T^{jq}T^{kr} = \frac{\det\underset{\approx}{\mathbf{T}}}{J^2}\varepsilon^{pqr} \ . \tag{3.66}$$

Now that we are using the ordinary permutation symbol, we may interpret this result as a matrix equation. Specifically, the left hand side represents the determinant of the contravariant component matrix T^{ij} times the permutation symbol ε^{pqr}. Therefore, the above equation implies that $\det[T^{ij}] = (\det\underset{\approx}{\mathbf{T}})/J^2$. Recalling Eq. (2.23),

$$\det\underset{\approx}{\mathbf{T}} = g_o\det[T^{ij}] \ . \tag{3.67}$$

Similarly, it can be shown that

$$\det\underset{\approx}{\mathbf{T}} = g^o\det[T_{ij}] \ . \tag{3.68}$$

In other words, if you have the matrix of covariant components, you compute the determinant of the tensor $\underset{\approx}{\mathbf{T}}$ by finding the determinant of the $[T_{ij}]$ matrix and multiplying the result by g^o.

Suppose the components of $\underset{\approx}{\mathbf{T}}$ are known in *mixed* form. Then Eq. (3.64) gives

$$\xi_{ijk}T^{i}_{\bullet p}T^{j}_{\bullet q}T^{k}_{\bullet r} = \det\underset{\approx}{\mathbf{T}}\,\xi_{pqr} \, . \tag{3.69}$$

Recalling Eqs. (3.45), we may introduce the ordinary permutation symbol. Noting that the J's on each side cancel, the above result as

$$\varepsilon_{ijk} T^i_{\bullet p} T^j_{\bullet q} T^k_{\bullet r} = \det \underset{\approx}{\mathbf{T}} \varepsilon_{pqr} . \tag{3.70}$$

Therefore, the determinant of the tensor $\underset{\approx}{\mathbf{T}}$ is equal to the determinant of the matrix of *mixed* components:

$$\det \underset{\approx}{\mathbf{T}} = \det[T^i_{\bullet j}] . \tag{3.71}$$

Likewise, it can be shown that

$$\det \underset{\approx}{\mathbf{T}} = \det[T_i^{\bullet j}] . \tag{3.72}$$

3.10 Tensor-valued operations

TRANSPOSE We have already discussed one tensor-valued operation, the transpose. Specifically, consider a tensor expressed in one of its four possible ways:

$$\underset{\approx}{\mathbf{T}} = T^{ij} \underset{\sim}{\mathbf{g}}_i \underset{\sim}{\mathbf{g}}_j = T_{ij} \underset{\sim}{\mathbf{g}}^i \underset{\sim}{\mathbf{g}}^j = T^i_{\bullet j} \underset{\sim}{\mathbf{g}}_i \underset{\sim}{\mathbf{g}}^j = T_i^{\bullet j} \underset{\sim}{\mathbf{g}}^i \underset{\sim}{\mathbf{g}}_j . \tag{3.73}$$

In Section 3.3, we showed that the transpose may be obtained by transposing the matrix base dyads.

$$\underset{\approx}{\mathbf{T}}^T = T^{ij} \underset{\sim}{\mathbf{g}}_j \underset{\sim}{\mathbf{g}}_i = T_{ij} \underset{\sim}{\mathbf{g}}^j \underset{\sim}{\mathbf{g}}^i = T^i_{\bullet j} \underset{\sim}{\mathbf{g}}^j \underset{\sim}{\mathbf{g}}_i = T_i^{\bullet j} \underset{\sim}{\mathbf{g}}_j \underset{\sim}{\mathbf{g}}^i . \tag{3.74}$$

Most people are more accustomed to thinking of transposing components rather than base vectors. Referring to the above result, we note that

$$(\underset{\approx}{\mathbf{T}}^T)^{ji} = T^{ij}, \qquad \text{or} \quad \underset{\sim}{\mathbf{g}}^j \bullet \underset{\approx}{\mathbf{T}}^T \bullet \underset{\sim}{\mathbf{g}}^i = \underset{\sim}{\mathbf{g}}^i \bullet \underset{\approx}{\mathbf{T}} \bullet \underset{\sim}{\mathbf{g}}^j \tag{3.75a}$$

$$(\underset{\approx}{\mathbf{T}}^T)_{ji} = T_{ij}, \qquad \text{or} \quad \underset{\sim}{\mathbf{g}}_j \bullet \underset{\approx}{\mathbf{T}}^T \bullet \underset{\sim}{\mathbf{g}}_i = \underset{\sim}{\mathbf{g}}_i \bullet \underset{\approx}{\mathbf{T}} \bullet \underset{\sim}{\mathbf{g}}_j \tag{3.75b}$$

$$(\underset{\approx}{\mathbf{T}}^T)_j^{\bullet i} = T^i_{\bullet j}, \qquad \text{or} \quad \underset{\sim}{\mathbf{g}}_j \bullet \underset{\approx}{\mathbf{T}}^T \bullet \underset{\sim}{\mathbf{g}}^i = \underset{\sim}{\mathbf{g}}^i \bullet \underset{\approx}{\mathbf{T}} \bullet \underset{\sim}{\mathbf{g}}_j \tag{3.75c}$$

$$(\underset{\approx}{\mathbf{T}}^T)^j_{\bullet i} = T_i^{\bullet j}, \qquad \text{or} \quad \underset{\sim}{\mathbf{g}}^j \bullet \underset{\approx}{\mathbf{T}}^T \bullet \underset{\sim}{\mathbf{g}}_i = \underset{\sim}{\mathbf{g}}_i \bullet \underset{\approx}{\mathbf{T}} \bullet \underset{\sim}{\mathbf{g}}^j \tag{3.75d}$$

Using our *matrix* notational conventions, these equations would be written

$$[(\underset{\approx}{\mathbf{T}}^T)^{mn}] = [T^{ij}]^T . \tag{3.76a}$$

$$[(\underset{\approx}{\mathbf{T}}^T)_{mn}] = [T_{ij}]^T . \tag{3.76b}$$

$$[(\underset{\approx}{\mathbf{T}}^T)_m^{\bullet n}] = [T^i_{\bullet j}]^T . \tag{3.76c}$$

$$[(\underset{\approx}{\mathbf{T}}^T)^m_{\bullet n}] = [T_i^{\bullet j}]^T . \tag{3.76d}$$

Here, we have intentionally used different indices (m and n) on the left-hand-side to remind the reader that matrix equations are not subject to the same index rules. The indices are present within the matrix brackets only to indicate to the reader which components (covariant, contravariant, or mixed) are contained in the matrix.

Equation (3.76a) says that the *matrix* of contravariant components of $[\underset{\approx}{\mathbf{T}}^T]$ are obtained by taking the transpose of the matrix of contravariant components of the original tensor $[\underset{\approx}{\mathbf{T}}]$. Similarly, Eq. (3.76b) says that the covariant component matrix for $\underset{\approx}{\mathbf{T}}^T$ is just the transpose of the covariant matrix for $\underset{\approx}{\mathbf{T}}$.

The *mixed* component formulas are more tricky. Equation (3.76c) states that the *low-high* mixed components of $\underset{\approx}{\mathbf{T}}^T$ are obtained by taking the transpose of the *high-low* components of $\underset{\approx}{\mathbf{T}}$. Likewise, Eq. (3.76d) states that the *high-low* mixed components of $\underset{\approx}{\mathbf{T}}^T$ are obtained by taking the transpose of the *low-high* components of $\underset{\approx}{\mathbf{T}}$.

INVERSE Consider a tensor $\underset{\approx}{\mathbf{T}}$. Let $\underset{\approx}{\mathbf{U}} = \underset{\approx}{\mathbf{T}}^{-1}$. Then

$$\underset{\approx}{\mathbf{T}} \bullet \underset{\approx}{\mathbf{U}} = \underset{\approx}{\mathbf{I}}. \tag{3.77}$$

Hence

$$T_{ik}U^{kj} = \delta_i^j. \tag{3.78}$$

Hence, the *contra*variant components of $\underset{\approx}{\mathbf{T}}^{-1}$ are obtained by inverting the T_{ij} matrix of *covariant* components.

Alternatively note a different component form of Eq. (3.77) is

$$T^i_{\bullet k}U^k_{\bullet j} = \delta^i_j. \tag{3.79}$$

This result states that the *high-low* mixed ${}^{i}_{\bullet j}$ components of $\underset{\approx}{\mathbf{T}}^{-1}$ are obtained by inverting the *high-low* mixed $[T^i_{\bullet j}]$ matrix.

The formulas for inverses may be written in our matrix notation as

$$[(\underset{\approx}{\mathbf{T}}^{-1})^{mn}] = [T_{ij}]^{-1}. \tag{3.80a}$$

$$[(\underset{\approx}{\mathbf{T}}^{-1})_{mn}] = [T^{ij}]^{-1}. \tag{3.80b}$$

$$[(\underset{\approx}{\mathbf{T}}^{-1})_m^{\bullet n}] = [(\underset{\approx}{\mathbf{T}}^{-1})_i^{\bullet j}]^{-1}. \tag{3.80c}$$

$$[(\underset{\approx}{\mathbf{T}}^{-1})^m_{\bullet n}] = [(\underset{\approx}{\mathbf{T}}^{-1})^i_{\bullet j}]^{-1}. \tag{3.80d}$$

COFACTOR Without proof, we claim that similar methods can be used to show that the matrix of *high-low* mixed ${}^{i}_{\bullet j}$ components of $\underset{\approx}{\mathbf{T}}^C$ are the cofactor of the *low-high* $[T_i^{\bullet j}]$ matrix. Note that the ${}^{i}_{\bullet j}$ components come from the ${}_{i}^{\bullet j}$ matrix. The contravariant matrix for $\underset{\approx}{\mathbf{T}}^C$ will equal the cofactor of the *covariant* matrix for $\underset{\approx}{\mathbf{T}}$ *times the metric scale factor* g^o. The full set of formulas is

$$[(\underset{\approx}{\mathbf{T}}^C)^{mn}] = g^o[T_{ij}]^C. \tag{3.81a}$$

$$[(\underset{\approx}{\mathbf{T}}^C)_{mn}] = g_o[T^{ij}]^C. \tag{3.81b}$$

$$[(\underset{\approx}{\mathbf{T}}^C)_m^{\bullet n}] = [T^i_{\bullet j}]^C. \tag{3.81c}$$

$$[(\underset{\approx}{\mathbf{T}}^C)^m_{\bullet n}] = [T_i^{\bullet j}]^C. \tag{3.81d}$$

DYAD By direct expansion, $\underset{\sim}{\mathbf{u}}\underset{\sim}{\mathbf{v}} = (u^i\underset{\sim}{\mathbf{g}}_i)(v^j\underset{\sim}{\mathbf{g}}_j) = u^i v^j \underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}_j$. Thus,

$$(\underset{\sim}{\mathbf{u}}\underset{\sim}{\mathbf{v}})^{ij} = u^i v^j. \tag{3.82}$$

Similarly,

$$(\underset{\sim}{\mathbf{u}}\underset{\sim}{\mathbf{v}})_{ij} = u_i v_j. \tag{3.83}$$

$$(\underset{\sim}{\mathbf{u}}\underset{\sim}{\mathbf{v}})^i_{\bullet j} = u^i v_j. \tag{3.84}$$

$$(\underset{\sim}{\mathbf{u}}\underset{\sim}{\mathbf{v}})_i^{\bullet j} = u_i v^j. \tag{3.85}$$

4. Basis and Coordinate transformations

This chapter discusses how vector and tensor components transform when the basis changes. This chapter may be skipped without any negative impact on the understandability of later chapters.

Chapter 3 had focused exclusively on the implications of using an irregular basis. For vector and tensor algebra (Chapter 3), all that matters is the possibility that the basis might be non-normalized and/or non-right-handed and/or non-orthogonal. For vector and tensor algebra, it doesn't matter whether or not the basis changes from one point in space to another — that's because algebraic operations are always applied *at a single location*. By contrast, differential operations such as the spatial gradient (discussed later in Chapter 5) are dramatically affected by whether or not the basis is fixed or spatially varying. This present chapter provides a gentle transition between the topics of Chapter 3 to the topics of Chapter 4 by outlining the transformation rules that govern how components of a vector with respect to one irregular basis will change if a different basis is used *at that same point in space*.

Throughout this document, we have asserted that any vector $\underset{\sim}{\mathbf{v}}$ can (and should) be regarded as invariant in the sense that the vector itself is unchanged upon a change in basis. The vector can be expanded as a sum of components v^i times corresponding base vectors $\underset{\sim}{\mathbf{g}}_i$:

$$\underset{\sim}{\mathbf{v}} = v^i \underset{\sim}{\mathbf{g}}_i \tag{4.1}$$

It's true that the individual components will change if the basis is changed, but the *sum* of components times base vectors will be invariant. Consequently, a vector's components must change in a very specific way if a different basis is used.

Basis transformation discussions are complicated and confusing because you have to consider two different systems at the same time. Since each system has both a covariant and a contravariant basis, talking about two different systems entails keeping track of *four* different basis triads. To help with this book-keeping nightmare, we will now refer to the first system as the "A" system and the other system as the "B" system. Each contravariant index (which is a superscript) will be accompanied by a subscript (either A or B) to indicate which system the index refers to. Likewise, each covariant index (a subscript) will now be accompanied by a superscript (A or B) to indicate the associated system. You should regard the system indicator (A or B) to serve the same sort of role as the "dot placeholder" discussed on page 16 — they are not indices. With this convention, we may now say

$$\{\underset{\sim}{\mathbf{g}}_1^A, \underset{\sim}{\mathbf{g}}_2^A, \underset{\sim}{\mathbf{g}}_3^A\} \text{ are the covariant base vectors for system-A} \tag{4.2}$$

$$\{\underset{\sim}{\mathbf{g}}_A^1, \underset{\sim}{\mathbf{g}}_A^2, \underset{\sim}{\mathbf{g}}_A^3\} \text{ are the contravariant base vectors for system-A} \tag{4.3}$$

$$\{\underset{\sim}{\mathbf{g}}_1^B, \underset{\sim}{\mathbf{g}}_2^B, \underset{\sim}{\mathbf{g}}_3^B\} \text{ are the covariant base vectors for system-B} \tag{4.4}$$

$$\{\underset{\sim}{\mathbf{g}}_B^1, \underset{\sim}{\mathbf{g}}_B^2, \underset{\sim}{\mathbf{g}}_B^3\} \text{ are the contravariant base vectors for system-B} \tag{4.5}$$

The system flag, A or B, acts as a "dot-placeholder" and moves along with its associated index in operations. Now that we are dealing with up to four basis triads, the components of vectors must also be flagged to indicate the associated basis. Hence, depending on which of the above four bases are used, the basis expansion of a vector can be any of the following:

$$\underset{\sim}{\mathbf{v}} = v_A^i \underset{\sim}{\mathbf{g}}_i^A \tag{4.6}$$

$$\underset{\sim}{\mathbf{v}} = v_i^A \underset{\sim}{\mathbf{g}}_A^i \tag{4.7}$$

$$\underset{\sim}{\mathbf{v}} = v_B^i \underset{\sim}{\mathbf{g}}_i^B \tag{4.8}$$

$$\underset{\sim}{\mathbf{v}} = v_B^i \underset{\sim}{\mathbf{g}}_i^B \tag{4.9}$$

Keep in mind that the system label (A or B) is to be regarded as a "dot-placeholder," not an implicitly summed index. The metric coefficients for any given system are defined in the usual way and the Kronecker delta relationship between contravariant and covariant base vectors *within a single system* still holds. Specifically,

$$g_{ij}^{AA} = \underset{\sim}{\mathbf{g}}_i^A \bullet \underset{\sim}{\mathbf{g}}_j^A \qquad g_{i\,A}^{A\,j} = \underset{\sim}{\mathbf{g}}_i^A \bullet \underset{\sim}{\mathbf{g}}_A^j = \delta_i^j \tag{4.10}$$

$$g_{A\,j}^{i\,A} = \underset{\sim}{\mathbf{g}}_A^i \bullet \underset{\sim}{\mathbf{g}}_j^A = \delta_j^i \qquad g_{AA}^{ij} = \underset{\sim}{\mathbf{g}}_A^i \bullet \underset{\sim}{\mathbf{g}}_A^j \tag{4.11}$$

$$g_{ij}^{BB} = \underset{\sim}{\mathbf{g}}_i^B \bullet \underset{\sim}{\mathbf{g}}_j^B \qquad g_{i\,B}^{B\,j} = \underset{\sim}{\mathbf{g}}_i^B \bullet \underset{\sim}{\mathbf{g}}_B^j = \delta_i^j \tag{4.12}$$

$$g_{B\,j}^{i\,B} = \underset{\sim}{\mathbf{g}}_B^i \bullet \underset{\sim}{\mathbf{g}}_j^B = \delta_j^i \qquad g_{BB}^{ij} = \underset{\sim}{\mathbf{g}}_B^i \bullet \underset{\sim}{\mathbf{g}}_B^j \tag{4.13}$$

In previous chapters (which dealt with only a single system), we showed that the matrix containing the g^{ij} components could be obtained by inverting the $[g_{ij}]$ matrix. This relationship still holds within each *individual* system (A or B) listed above. Namely,

$$g_{ik}^{AA} g_{AA}^{kj} = g_{i\,A}^{A\,j} = \delta_i^j \tag{4.14}$$

$$g_{ik}^{BB} g_{BB}^{kj} = g_{i\,B}^{B\,j} = \delta_i^j \tag{4.15}$$

Now that we are considering two systems simultaneously, we can further define new matrices that inter-relate the two systems:

$$g_{ij}^{AB} = \underset{\sim}{\mathbf{g}}_i^A \bullet \underset{\sim}{\mathbf{g}}_j^B \qquad g_{AB}^{ij} = \underset{\sim}{\mathbf{g}}_A^i \bullet \underset{\sim}{\mathbf{g}}_B^j \tag{4.16a}$$

$$g_{ij}^{BA} = \underset{\sim}{\mathbf{g}}_i^B \bullet \underset{\sim}{\mathbf{g}}_j^A \qquad g_{BA}^{ij} = \underset{\sim}{\mathbf{g}}_B^i \bullet \underset{\sim}{\mathbf{g}}_A^j \tag{4.16b}$$

$$g_{A\,j}^{i\,B} = \underset{\sim}{\mathbf{g}}_A^i \bullet \underset{\sim}{\mathbf{g}}_j^B \qquad g_{i\,B}^{A\,j} = \underset{\sim}{\mathbf{g}}_i^A \bullet \underset{\sim}{\mathbf{g}}_B^j \tag{4.16c}$$

$$g_{B\,j}^{i\,A} = \underset{\sim}{\mathbf{g}}_B^i \bullet \underset{\sim}{\mathbf{g}}_j^A \qquad g_{i\,A}^{B\,j} = \underset{\sim}{\mathbf{g}}_i^B \bullet \underset{\sim}{\mathbf{g}}_A^j \tag{4.16d}$$

Study Question 4.1 Consider the following irregular base vectors (the A-system):

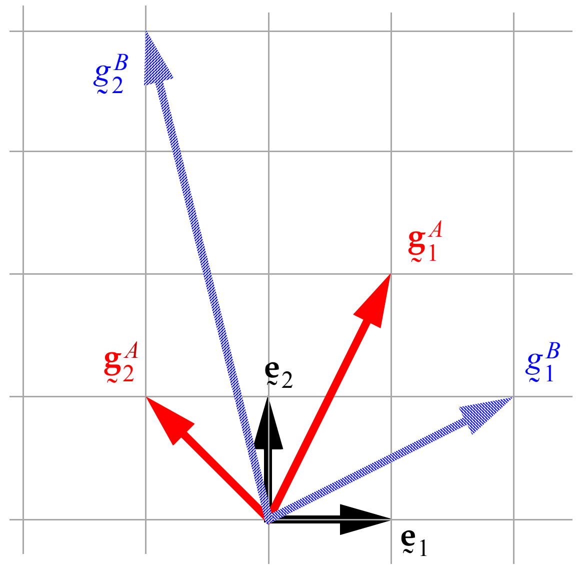
$$\underset{\sim}{\mathbf{g}}_1^A = \underset{\sim}{\mathbf{e}}_1 + 2\underset{\sim}{\mathbf{e}}_2$$

$$\underset{\sim}{\mathbf{g}}_2^A = -\underset{\sim}{\mathbf{e}}_1 + \underset{\sim}{\mathbf{e}}_2 \quad \text{and} \quad \underset{\sim}{\mathbf{g}}_3^A = \underset{\sim}{\mathbf{e}}_3$$

Additionally consider a second B-system:

$$\underset{\sim}{\mathbf{g}}_1^B = 2\underset{\sim}{\mathbf{e}}_1 + \underset{\sim}{\mathbf{e}}_2$$

$$\underset{\sim}{\mathbf{g}}_2^B = -\underset{\sim}{\mathbf{e}}_1 + 4\underset{\sim}{\mathbf{e}}_2 \quad \text{and} \quad \underset{\sim}{\mathbf{g}}_3^B = \underset{\sim}{\mathbf{e}}_3$$



(a) Find the contravariant and covariant metrics for each individual system.

(b) Find the contravariant basis for each individual system.

(c) Directly apply Eq. (4.16) to find the coupling matrices. Then verify that $[g_{ij}^{AB}] = [g_{ij}^{BA}]^T$.

Partial Answer: *(a)* $[g_{ij}^{AA}] = \begin{bmatrix} 5 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $[g^{ij}_{AA}] = \begin{bmatrix} 2/9 & -1/9 & 0 \\ -1/9 & 5/9 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $g_{ij}^{BB} = \begin{bmatrix} 5 & 2 & 0 \\ 2 & 17 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $g_{ij}^{AA} = \begin{bmatrix} 17/81 & -1/81 & 0 \\ -1/81 & 5/81 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

(b) $\underset{\sim}{\mathbf{g}}_A^1 = \frac{1}{3}\underset{\sim}{\mathbf{e}}_1 + \frac{1}{3}\underset{\sim}{\mathbf{e}}_2$, $\underset{\sim}{\mathbf{g}}_A^2 = \frac{-2}{3}\underset{\sim}{\mathbf{e}}_1 + \frac{1}{3}\underset{\sim}{\mathbf{e}}_2$, $\underset{\sim}{\mathbf{g}}_B^1 = \frac{4}{9}\underset{\sim}{\mathbf{e}}_1 + \frac{1}{9}\underset{\sim}{\mathbf{e}}_2$, $\underset{\sim}{\mathbf{g}}_B^2 = \frac{-1}{9}\underset{\sim}{\mathbf{e}}_1 + \frac{2}{9}\underset{\sim}{\mathbf{e}}_2$

(c) $[g_{ij}^{AB}] = \begin{bmatrix} 4 & 7 & 0 \\ -1 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $[g_{ij}^{BA}] = \begin{bmatrix} 4 & -1 & 0 \\ 7 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $[g^{iA}_{Bj}] = \begin{bmatrix} 2/3 & -1/3 & 0 \\ 1/3 & 1/3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $[g^{Aj}_{iB}] = \begin{bmatrix} 2/3 & 1/3 & 0 \\ -1/3 & 1/3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$,

$[g^{ij}_{AB}] = \begin{bmatrix} 5/27 & 1/27 & 0 \\ -7/27 & 4/27 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $[g^{ij}_{BA}] = \begin{bmatrix} 5/27 & -7/27 & 0 \\ 1/27 & 4/27 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $[g^{iB}_{Aj}] = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $[g^{Bj}_{iA}] = \begin{bmatrix} 1 & -1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

From the commutativity of the dot product, Eq. (4.16) implies that

$$g_{ij}^{AB} = g_{ji}^{BA}, \qquad g^{ij}_{AB} = g^{ij}_{BA}, \qquad g^{iB}_{Aj} = g^{Bi}_{jA}, \quad \text{and} \qquad g_{i\ B}^{A\ j} = g^{jA}_{Bi}. \tag{4.17}$$

System metrics (i.e., "g" matrices that involve AA or BB) are components of a tensor (the identity tensor). Coupling matrix components (i.e., ones that involve A and B) are not components of a tensor — instead, a coupling matrix characterizes interrelationship between two bases.

The first relationship in Eq. (4.17) does not say that the matrix containing g_{ij}^{AB} is symmetric. Instead this equation is saying that the matrix containing g_{ij}^{AB} may be obtained by the transpose of the *different* matrix that contains g_{ij}^{BA}. Given a "g" *matrix* of any type, you can transform it immediately to other types of g-matrices by using the following rules:

- To exchange system labels (A or B) left-right, apply a transpose.
- To exchange system labels (A or B) up-down, apply an inverse.
- To exchange system labels (A or B) along a diagonal, apply an inverse transpose.

For example, if g_{ij}^{AB} is available, then

$$\overset{\textbf{left-right}}{[g_{**}^{BA}] = [g_{**}^{AB}]^T}, \qquad \overset{\textbf{up-down}}{g_{AB}^{**} = [g_{**}^{AB}]^{-1}} \qquad \overset{\textbf{diagonal}}{g_{BA}^{**} = [g_{**}^{AB}]^{-T}} \tag{4.18}$$

Similarly, if g_{iB}^{Aj} is available, then

$$\overset{\textbf{left-right}}{[g_{B*}^{*A}] = [g_{*B}^{A*}]^T}, \qquad \overset{\textbf{up-down}}{g_{A*}^{*B} = [g_{*B}^{A*}]^{-1}} \qquad \overset{\textbf{diagonal}}{g_{*A}^{B*} = [g_{*B}^{A*}]^{-T}} \tag{4.19}$$

Here, and in all upcoming matrix equations involving "g" components, a star (*) is inserted where indices normally reside in indicial equations. These equations show how to move two system labels simultaneously.

To move only one system label to a new location, you need a matrix product. In matrix products, abutting system labels must be on opposite levels (the system labels are *not* summed — index rules apply only to indicial expressions, not to matrix expressions. For example, what operation would you need to move only the "B" label in $[g_{B*}^{*A}]$ from the bottom to the top? The answer is that you need to use the B-system metric as follows:

$$[g_{**}^{BA}] = [g_{**}^{BB}][g_{B*}^{*A}] \qquad \text{(this is the matrix form of } g_{ij}^{BA} = g_{ik}^{BB} g_{Bj}^{kA}\text{)} \tag{4.20}$$

This behavior also applies to transforming bases. For example,

$$\underset{\sim}{\mathbf{g}}_i^A = g_{iB}^{Ak} \underset{\sim}{\mathbf{g}}_k^B \tag{4.21}$$

All of the "g" matrices may be obtained only from knowledge of one system metric (same system labels) and one coupling matrix (different system labels).

Study Question 4.2 Suppose you are studying a problem involving two bases, A and B, and you know one system metric g_{ij}^{BB} and one coupling matrix g_{iA}^{Bj} as follows:

$$[g_{**}^{BB}] = \begin{bmatrix} 4 & 2 & 3 \\ 2 & 4 & -1 \\ 3 & -1 & 5 \end{bmatrix} \quad \text{and} \quad [g_{*A}^{B*}] = \begin{bmatrix} 1 & 2 & -4 \\ 0 & -1 & 2 \\ 3 & 0 & -2 \end{bmatrix} \tag{4.22}$$

(a) Find $[g_{**}^{AA}]$ and $[g_{AA}^{**}]$

(b) Express the $\underset{\sim}{\mathbf{g}}_i^A$ base vectors in terms of the $\underset{\sim}{\mathbf{g}}_k^B$ base vectors.

(b) Express the $\underset{\sim}{\mathbf{g}}_A^i$ base vectors in terms of the $\underset{\sim}{\mathbf{g}}_k^B$ base vectors.

Partial Answer: *(a) We are given a regular metric and a mixed coupling matrix. Start by expressing* $[g_{**}^{AA}]$ *as the product of any AB matrix times a BB metric times a BA matrix:* $[g_{**}^{AA}] = [g_{*B}^{A*}][g_{**}^{BB}][g_{B*}^{*A}]$. *We now need to get* $[g_{*B}^{A*}]$ *expressed in terms of the* <u>*given*</u> *coupling matrix* $[g_{*A}^{B*}]$. *In other words, we need to swap A and B along the diagonal, which requires an inverse-transpose:* $[g_{*B}^{A*}] = [g_{*A}^{B*}]^{-T}$. *Next, we need to express the coupling matrix* $[g_{B*}^{*A}]$ *in terms of the given coupling matrix* $[g_{*A}^{B*}]$. *This requires only an up-down motion of the system labels, so that's just an inverse:* $[g_{B*}^{*A}] = [g_{*A}^{B*}]^{-1}$. *Putting it all back in the first equation gives* $[g_{**}^{AA}] = [g_{*A}^{B*}]^{-T}[g_{**}^{BB}][g_{*A}^{B*}]^{-1} = \begin{bmatrix} 1 & 2 & -4 \\ 0 & -1 & 2 \\ 3 & 0 & -2 \end{bmatrix}^{-T} \begin{bmatrix} 4 & 2 & 3 \\ 2 & 4 & -1 \\ 3 & -1 & 5 \end{bmatrix} \begin{bmatrix} 1 & 2 & -4 \\ 0 & -1 & 2 \\ 3 & 0 & -2 \end{bmatrix} = \begin{bmatrix} \frac{19}{2} & \frac{5}{2} & \frac{57}{2} \\ 5 & 11 & 49 \\ -\frac{9}{2} & \frac{1}{2} & -\frac{15}{2} \end{bmatrix}$

(b) $\underset{\sim}{\mathbf{g}}_i^A = g_{iB}^{Ak}\underset{\sim}{\mathbf{g}}_k^B$. *The matrix in this relation is* $[g_{*B}^{A*}] = [g_{*A}^{B*}]^{-T} = \begin{bmatrix} 1 & 3 & \frac{3}{2} \\ 2 & 5 & 3 \\ 0 & -1 & -\frac{1}{2} \end{bmatrix}$. *Therefore*

$\underset{\sim}{\mathbf{g}}_1^A = \underset{\sim}{\mathbf{g}}_1^B + 3\underset{\sim}{\mathbf{g}}_2^B + \frac{3}{2}\underset{\sim}{\mathbf{g}}_3^B$, $\underset{\sim}{\mathbf{g}}_2^A = 2\underset{\sim}{\mathbf{g}}_1^B + 5\underset{\sim}{\mathbf{g}}_2^B + 3\underset{\sim}{\mathbf{g}}_3^B$, *and* $\underset{\sim}{\mathbf{g}}_3^A = -\underset{\sim}{\mathbf{g}}_2^B - \frac{1}{2}\underset{\sim}{\mathbf{g}}_3^B$.

(c) is similar.

We showed in previous sections how the covariant components of a vector are related to the contravariant components *from the same system*. For example, we showed that $v_i = g_{ij}v^j$ when considering only one system. When considering more than one basis at a time, this result needs to explicitly show the system (A or B) flags:

$$v_i^A = g_{ij}^{AA} v_A^j \tag{4.23}$$

$$v_i^B = g_{ij}^{BB} v_B^j \tag{4.24}$$

These equations show how to transform components *within a single system.* To transform components from one system to another, the formulas are

$$v_i^A = g_{iB}^{Ak} v_k^B \qquad v_i^A = g_{ik}^{AB} v_B^k \tag{4.25}$$

$$v_A^i = g_{AB}^{ik} v_k^B \qquad v_A^i = g_{Ak}^{iB} v_B^k \tag{4.26}$$

Equivalently,

$$v_i^A = v_k^B g_{Bi}^{kA} \qquad v_i^A = v_B^k g_{ki}^{BA} \tag{4.27}$$

$$v_A^i = v_k^B g_{BA}^{ki} \qquad v_A^i = v_B^k g_{kA}^{Bi} \tag{4.28}$$

Tensor transformations are similar. For example

$$T_{ij}^{AA} = T_{pq}^{BB} g_{Bi}^{pA} g_{Bj}^{qA} \qquad \text{and} \qquad T_{Bj}^{iB} = T_{AA}^{pq} g_{pB}^{Ai} g_{qj}^{AB} \tag{4.29}$$

These formulas are constructed one index position at a time. In the last formula, for example, the first free index position on the T is the pair $\begin{smallmatrix} i \\ B \end{smallmatrix}$. On the other side of the equation, the first index pair on the T is $\begin{smallmatrix} p \\ A \end{smallmatrix}$. Because p is a dummy summed index, the "g" must have $\begin{smallmatrix} A \\ p \end{smallmatrix}$. Thus, the first "$g$" is a combination of the free index pair on the left side and the "flipped" index pair on the right g_{pB}^{Ai}. The "T" in the above equations may be moved to the middle if you like it better that way:

$$T_{ij}^{AA} = g_{Bi}^{pA} T_{pq}^{BB} g_{Bj}^{qA} \qquad \text{and} \qquad T_{Bj}^{iB} = g_{pB}^{Ai} T_{AA}^{pq} g_{qj}^{AB} \tag{4.30}$$

If you wanted to write these as matrix equations, you can write them in a way that dummy summed indices are abutting, keeping in mind that the g components are unchanged when indices *along with their system label* are moved left-to-right. The g components require an inverse on their matrix forms to move system labels up-down. Thus, the matrix forms for these transformation formulas are

$$[T_{**}^{AA}] = [g_{B*}^{*A}]^T [T_{**}^{BB}][g_{B*}^{*A}] \qquad \text{and} \qquad T_{Bj}^{iB} = [g_{B*}^{*A}]^T [T_{AA}^{**}][g_{B*}^{*A}]^{-T}[g_{**}^{AA}] \tag{4.31}$$

Numerous matrix versions of component transformation formulas may be written. In this case, we wrote the matrix expressions on the assumption that the known coupling and metric matrices were $[g_{B*}^{*A}]$ and $[g_{**}^{AA}]$.

Study Question 4.3 Consider the same irregular base vectors in Study Question 4.1. Namely,

$$\mathbf{g}_1^A = \mathbf{e}_1 + 2\mathbf{e}_2$$

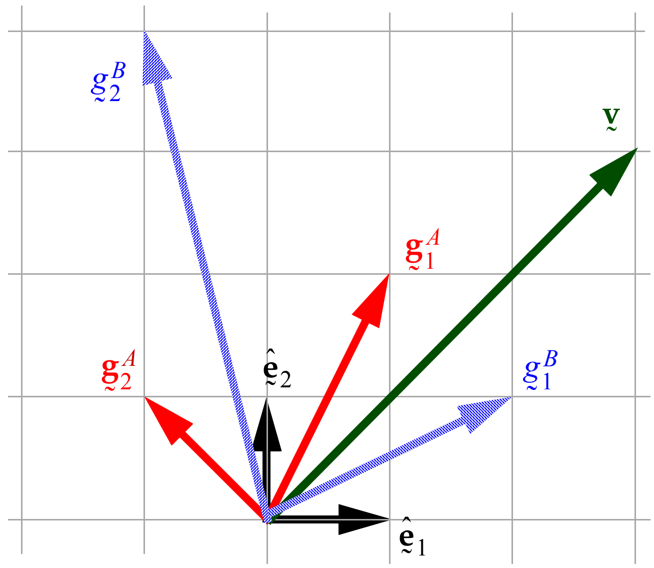
$$\mathbf{g}_2^A = -\mathbf{e}_1 + \mathbf{e}_2 \quad \text{and} \quad \mathbf{g}_3^A = \mathbf{e}_3$$

and

$$\mathbf{g}_1^B = 2\mathbf{e}_1 + \mathbf{e}_2$$

$$\mathbf{g}_2^B = -\mathbf{e}_1 + 4\mathbf{e}_2 \quad \text{and} \quad \mathbf{g}_3^B = \mathbf{e}_3$$

Let $\mathbf{v} = 3\mathbf{e}_1 + 3\mathbf{e}_2$.



(a) Determine the covariant and contravariant components of $\mathbf{v}$ with respect to systems A and B. (i.e., find v_i^A, v_A^i, v_i^B, v_B^i).

(b) Demonstrate graphically that your answers to part (a) make sense.

(c) Verify the following:

$$v_i^A = g_{ik}^{AB} v_B^k \qquad v_i^A = g_{iB}^{Ak} v_k^B$$

$$v_A^i = g_{Ak}^{iB} v_B^k \qquad v_A^i = g_{AB}^{ik} v_k^B$$

(d) Fill in the question marks to make the following equation true:

$$v_i^B = g_{??}^{??} v_?^A$$

Partial Answer: *(a)* $v_1^A = 9$, $v_2^A = 0$, $v_A^1 = 2$, $v_A^2 = -1$, $v_1^B = 9$, $v_2^B = 9$, $v_B^1 = \frac{5}{3}$, $v_B^2 = \frac{1}{3}$.

(b) The first third and fourth results of part (a) imply that $\mathbf{v} = v_A^1 \mathbf{g}_1^A + v_A^2 \mathbf{g}_2^A = 2\mathbf{g}_1^A - \mathbf{g}_2^A$. *Drawing grid lines parallel to* $\mathbf{g}_1^A$ *and* $\mathbf{g}_2^A$*, the path from the origin to the tip of vector* $\mathbf{v}$ *may be reached by first traversing along a line segment equal to* $2\mathbf{g}_1^A$*, and then moving anti-parallel to* $\mathbf{g}_2^A$ *to reach the tip.*

(c) Using results from study question 4.1, $v_i^A = g_{ik}^{AB} v_B^k$ *becomes in matrix form*
$$\begin{bmatrix} 9 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 & 7 & 0 \\ -1 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5/3 \\ 1/3 \\ 0 \end{bmatrix}.$$

Multiplying this out verifies that it is true.

(d) The answer is $v_i^B = g_{iA}^{Bk} v_k^A$.

Coordinate transformations A principal goal of this chapter is to show how the v_B^i components must be related to the v_A^i components. We are seeking *component* transformations resulting from a change of *basis*. Strictly speaking, coordinates may always be selected independently from the basis. This, however, is rarely done. Usually the basis is taken to be defined to be the one that is naturally defined by the coordinate system grid lines. The natural covariant basis co-varies with the coordinates — i.e., each covariant base vector always points tangent to the coordinate a grid line and has a length that varies in proportion to the grid spacing density. The natural contravariant basis contra-varies with the grid — i.e., each contravariant base vector points normal to a surface of constant coordinate value. When people talk about component transformations resulting from a change in *coordinates*, they implicitly assuming that the basis is coupled to the choice of coordinates. There is an unfortunate implicit assumption in many publications that vector and tensor components *must* be coupled to the spatial coordinates. In fact, it's fine to use a basis that is selected entirely independently from the coordinates. For example, to study loads on a ferris wheel, you might decide to use a fixed Cartesian basis that is aligned with gravity, while also using cylindrical (r, θ, z) *coordinates* to identify points in space.

Spatial **coordinates** are any three numbers that uniquely define a location in space. For example, Cartesian coordinates are frequently denoted $\{x, y, z\}$, cylindrical coordinates are $\{r, \theta, z\}$, and so on. Any quantity that is known to vary in space can be written as a function of the coordinates. Coordinates uniquely define the position vector $\underset{\sim}{\mathbf{x}}$, but *coordinates* are not the same thing as *components* of the position vector. For example, the position vector in cylindrical coordinates is $\underset{\sim}{\mathbf{x}} = r\underset{\sim}{\mathbf{e}}_r + z\underset{\sim}{\mathbf{e}}_z$. Notice that position vector is *not* $r\underset{\sim}{\mathbf{e}}_r + \theta\underset{\sim}{\mathbf{e}}_\theta + z\underset{\sim}{\mathbf{e}}_z$. The components of the position vector are $\{r, \textbf{zero}, z\}$, not $\{r, \theta, z\}$. The position vector has only *two* nonzero components. It has no $\theta\underset{\sim}{\mathbf{e}}_\theta$ term. Does this mean that the position vector depends only on r and z? Nope. The position vector's dependence on θ is buried *implicitly* in the dependence of $\underset{\sim}{\mathbf{e}}_r$ on θ.

Whether the basis is chosen to be coupled to the coordinates is entirely up to you. If convenient to do so (as in the ferris wheel example), you might choose to use cylindrical coordinates but a *Cartesian* basis so that the position vector would be written $\underset{\sim}{\mathbf{x}} = (r\cos\theta)\underset{\sim}{\mathbf{e}}_x + (r\sin\theta)\underset{\sim}{\mathbf{e}}_y + z\underset{\sim}{\mathbf{e}}_z$. In this case, the chosen basis is the Cartesian laboratory basis $\{\underset{\sim}{\mathbf{e}}_x, \underset{\sim}{\mathbf{e}}_y, \underset{\sim}{\mathbf{e}}_z\}$, which is entirely uncoupled from the coordinates $\{r, \theta, z\}$.

To speak in generality, we will denote the three spatial coordinates by $\{\eta^1, \eta^2, \eta^3\}$. If, for example, cylindrical coordinates are used, then $\eta^1 = r$, $\eta^2 = \theta$, and $\eta^3 = z$. If spherical coordinates are used, then $\eta^1 = r$, $\eta^2 = \theta$, and $\eta^3 = \psi$. If Cartesian coordinates are used, then $\eta^1 = x$, $\eta^2 = y$, and $\eta^3 = z$.

The position vector is truly a function of all three $\{\eta^1, \eta^2, \eta^3\}$ coordinates. The basis can be selected independently from the choice of coordinates, However, we can always define the ***natural* basis** [12] that *is* associated with the choice of coordinates by

$$\underset{\sim}{\mathbf{g}}_i \equiv \frac{\partial \underset{\sim}{\mathbf{x}}}{\partial \eta^i} \tag{4.32}$$

The i^{th} natural covariant base vector $\underset{\sim}{\mathbf{g}}_i$ equals the derivative of the position vector $\underset{\sim}{\mathbf{x}}$ with respect to the i^{th} coordinate. Hence, this base vector points in the direction of increasing η^i and it is tangent to the grid line along which the other two coordinates remain constant. Examples of the natural basis for various coordinate systems are provided in Section 5.3. The natural basis defined in Eq. (4.32) is clearly coupled to the coordinates themselves. Consequently, a change in coordinates will result in a change of the natural basis. Thus, the components of any vector expressed using the *natural* basis will change upon a change in coordinates. Having the basis be coupled to the coordinates is a *choice*. One can alternatively choose the basis to be uncoupled from the coordinates (see the discussion of non-holonomicity in Ref. [12]).

Each particular value of the coordinates $\{\eta^1, \eta^2, \eta^3\}$ defines a unique location $\underset{\sim}{\mathbf{x}}$ in space. Conversely, each location in space corresponds to a unique set of $\{\eta^1, \eta^2, \eta^3\}$ coordinates. That means the coordinates themselves can be regarded as functions of the position vector, and we can therefore take the spatial gradients of the coordinates. As further explained in Section 5.3, the ***contravariant* natural base vector** is defined to equal these coordinate gradients:

$$\underset{\sim}{\mathbf{g}}^j \equiv \frac{\partial \eta^j}{\partial \underset{\sim}{\mathbf{x}}} \tag{4.33}$$

Being the gradient of η^j, the contravariant base vector $\underset{\sim}{\mathbf{g}}^j$ must point normal to surfaces of constant η^j. Proving that these are indeed the contravariant base vectors associated with the covariant natural basis defined in Eq. (4.32) is simply a matter of applying the chain rule to demonstrate that $\underset{\sim}{\mathbf{g}}_i \bullet \underset{\sim}{\mathbf{g}}^j$ comes out to equal the Kronecker delta:

$$\underset{\sim}{\mathbf{g}}_i \bullet \underset{\sim}{\mathbf{g}}^j = \frac{\partial \underset{\sim}{\mathbf{x}}}{\partial \eta^i} \bullet \frac{\partial \eta^j}{\partial \underset{\sim}{\mathbf{x}}} = \frac{\partial \eta^j}{\partial \eta^i} = \delta_i^j \tag{4.34}$$

The metric coefficients corresponding to Eq. (4.32) are

$$g_{ij} = \underset{\sim}{\mathbf{g}}_i \bullet \underset{\sim}{\mathbf{g}}_j = \frac{\partial \underset{\sim}{\mathbf{x}}}{\partial \eta^i} \bullet \frac{\partial \underset{\sim}{\mathbf{x}}}{\partial \eta^j} \tag{4.35}$$

Many textbooks present this result in pure component form by writing the vector $\underset{\sim}{\mathbf{x}}$ in terms of its cartesian components as $\underset{\sim}{\mathbf{x}} = x_k \underset{\sim}{\mathbf{e}}_k$ so that the above expression becomes

$$g_{ij} = \sum_{k=1}^{3} \frac{\partial x_k}{\partial \eta^i} \frac{\partial x_k}{\partial \eta^j} \qquad \text{for the natural basis} \tag{4.36}$$

When written this way, it's important to recognize that x_k are the *Cartesian* components of the position vector (not the covariant components with respect to an irregular basis). That's why we introduced the summation in Eq. (4.36) -- otherwise, we would be violating the high-low rule of the summation conventions.

Similarly, the contravariant metric coefficients are

$$g^{ij} = \sum_{k=1}^{3} \frac{\partial\eta^i}{\partial x_k}\frac{\partial\eta^j}{\partial x_k} \qquad \text{for the natural basis} \tag{4.37}$$

These are the expressions for the metric coefficients when the natural basis is used. Let us reiterate that there is no law that says the basis used to described vectors and tensors must necessarily be coupled to the choice of spatial coordinates in any way. We cite these relationships to assist those readers who wish to make connections between our own present formalism and other published discourses on curvilinear coordinates.

In the upcoming discussion of coordinate transformations, we will *not* assume that the metric coefficients are given by the above expressions. Because our transformation equations will be expressed in terms of metrics only, they will apply even when the chosen basis is *not* the natural basis. When we derive the component transformation rules, we will include both the general basis transformation expression and its specialization to the case that the basis is chosen to be the *natural* basis.

4.1 What is a vector? What is a tensor?

The bane of many rookie graduate students is the common qualifier question: "what is a vector?" The frustration stems, in part, from the fact that the "correct" answer depends on who is doing the asking. Different people have different answers to this question. More than likely, the professor follows up with the even harder question "what is a tensor?"

Definition #0 (used for undergraduates): A college freshman is typically told that a vector is something with length and direction, and nothing more is said (certainly no mention is made of tensors!). The idea of a tensor is tragically withheld from most undergraduates.

Definition #1 (classical, but our least favorite): Some professors merely want their student to say that there are two kinds of vectors: A contravariant vector is a set of three numbers $\{v^1, v^2, v^3\}$ that transform according to $v_B^k = v_A^i g_{iB}^{Ak}$ when changing from an "A-basis" to a "B-basis", whereas a covariant vector is a set of three numbers $\{v_1, v_2, v_3\}$ that transform such that $v_k^B = v_i^A g_{Ak}^{iB}$. These professors typically wish for their students to define *four* different kinds of tensors, the contravariant tensor being a set of nine numbers that transform according to $T_{BB}^{ij} = T_{AA}^{mn} g_{Bm}^{iA} g_{Bn}^{jA}$, the mixed tensors being ones whose components transform according to $T_{iB}^{Bj} = T_{mA}^{An} g_{im}^{BA} g_{BA}^{jn}$ or $T_{Bj}^{iB} = T_{An}^{mA} g_{BA}^{im} g_{jn}^{BA}$, etc.

Definition #2 (our preference for ordinary engineering applications): We regard a vector to be an entity that exists independent of our act of observing it — independent of our act of representing it quantitatively. Of course, to decide if something is an engineering vector, we demand that it must nevertheless lend itself to a quantitative representation as a sum of three components $\{v^1, v^2, v^3\}$ times[1] three base vectors $\{\underset{\sim}{\mathbf{g}}_1, \underset{\sim}{\mathbf{g}}_2, \underset{\sim}{\mathbf{g}}_3\}$, and we *test* whether or not we have a vector by verifying that $v_B^k = v_A^i g_{iB}^{Ak}$ holds upon a change of basis. The base vectors themselves are regarded as abstractions defined in terms of geometrical fundamental primitives. Namely, the base vectors are regarded as directed line segments between points in space, which are presumed (by axiom) to exist. The key distinction between definition #2 and #1 is that definition #2 makes no distinction between covariant vectors and contravariant vectors. Definition #2 instead uses the terms "covariant components" and "contravariant components" — these are components of the *same* vector.

Recall that definition of a *second-order tensor* under our viewpoint required introduction of new objects, called dyads, that could be constructed from vectors. From there, we asserted that any collection of nine numbers could be assembled as a sum of these numbers times basis dyads $\underset{\sim}{\mathbf{g}}_i \underset{\sim}{\mathbf{g}}_j$. A tensor would then be defined as any linear combination of tensor dyads for which the coefficients of these dyads (i.e., the tensor components) follow the transformation rules outlined in the preceding section.

Definition #3 (for mathematicians or for advanced engineering) Mathematicians define vectors as being "members of a vector space." A vector space must comprise certain basic components:

A1. A field R must exist. *(For engineering applications, the field is the set of reals.)*

A2. There must be a discerning definition of membership in a set V.

A3. There must be a rule for multiplying a scalar α times a member $\underset{\sim}{\mathbf{v}}$ of V. Furthermore, this multiplication rule must be *proved* closed in V:
If $\alpha \in R$ and $\underset{\sim}{\mathbf{v}} \in V$ then $\alpha \underset{\sim}{\mathbf{v}} \in V$

A4. There must be a rule for adding two members of V.
Furthermore, this vector addition rule must be *proved* closed in V:
If $\underset{\sim}{\mathbf{v}} \in V$ and $\underset{\sim}{\mathbf{w}} \in V$ then $\underset{\sim}{\mathbf{v}} + \underset{\sim}{\mathbf{w}} \in V$

A5. There must be a well defined process for determining whether two members of V are equal.

A6. The multiplication and addition rules must satisfy the following rules:

- $\underset{\sim}{\mathbf{v}} + \underset{\sim}{\mathbf{w}} = \underset{\sim}{\mathbf{w}} + \underset{\sim}{\mathbf{v}}$ and $\alpha \underset{\sim}{\mathbf{v}} = \underset{\sim}{\mathbf{v}} \alpha$
- $\underset{\sim}{\mathbf{u}} + (\underset{\sim}{\mathbf{v}} + \underset{\sim}{\mathbf{w}}) = (\underset{\sim}{\mathbf{u}} + \underset{\sim}{\mathbf{v}}) + \underset{\sim}{\mathbf{w}}$
- There must exist a zero vector $\underset{\sim}{\mathbf{0}} \in V$ such that $\underset{\sim}{\mathbf{v}} + \underset{\sim}{\mathbf{0}} = \underset{\sim}{\mathbf{v}}$.

1. The act of multiplying a scalar times a vector is presumed to be well defined and to satisfy the rules outlined in definition #3.

Unfortunately, textbooks seem to fixate on item A6, completely neglecting the far more subtle and difficult items A2, A3, A4, and A5. Engineering vectors are more than something with length and direction. Likewise, engineering vectors are more than simply an array of three numbers. When people define vectors according to the way their components change upon a change of basis (definitions #1 and #2), they are implicitly addressing axiom A2. Our "definition #2" is a special case of this more general definition. In general, axiom A2 is the most difficult axiom to satisfy when discussing *specific* vector spaces.

Whenever possible, vector spaces are typically supplemented with the definition of an inner product (here denoted $(\underset{\sim}{\mathbf{a}}, \underset{\sim}{\mathbf{b}})$), which is a scalar-valued binary[1] operation between two vectors, $\underset{\sim}{\mathbf{a}}$ and $\underset{\sim}{\mathbf{b}}$, that must satisfy the following rules:

A7. $(\underset{\sim}{\mathbf{a}}, \underset{\sim}{\mathbf{b}}) = (\underset{\sim}{\mathbf{b}}, \underset{\sim}{\mathbf{a}})$

A8. $(\underset{\sim}{\mathbf{a}}, \underset{\sim}{\mathbf{a}}) > 0$ if $\underset{\sim}{\mathbf{a}} \neq \underset{\sim}{\mathbf{0}}$ and $(\underset{\sim}{\mathbf{a}}, \underset{\sim}{\mathbf{a}}) = \underset{\sim}{\mathbf{0}}$ only if $\underset{\sim}{\mathbf{a}} = \underset{\sim}{\mathbf{0}}$.

An inner product space is just a vector space that has an inner product.

Study Question 4.4

(a) Apply definition #3 to prove that the set of all real-valued continuous functions of one variable (on a domain being reals) is a vector space.

(b) Suppose that f and g are members of this vector space. Prove that the scalar-valued functional

$$\int_{-\infty}^{\infty} f(x)g(x)dx \tag{4.38}$$

is an acceptable definition of the inner product, (f, g).

(c) Explain what a Taylor-series expansion $f(x)$ has in common with a basis expansion ($\underset{\sim}{\mathbf{f}} = f_i \underset{\sim}{\mathbf{g}}_i$) of a vector. A Fourier-series expansion is related to the Taylor series in a way that is analogous to what operation from ordinary vector analysis in 3D space?

(d) Explain how the binary function $f(x)g(y)$ is analogous to the dyad $\underset{\sim}{\mathbf{f}}\underset{\sim}{\mathbf{g}}$ from conventional vector/tensor analysis.

(e) Explain why the dummy variables x and y play roles similar to the indices i and j

(f) Explain why a general binary function $a(x, y)$ is analogous to a conventional tensor components A_{ij}.

(g) Assuming that the goal of this document is to give you greater insight into vector and tensor analysis in 3D space, do you think this problem contributes toward that goal? Explain?

Partial Answer: *(a) You may assume that the set of reals constitutes an adequate field, but you may obtain a careful definition of the term "field" in Reference [6]. In order to deal with axiom A2, you will need to locate a carefully crafted definition of what it means to be a "real valued continuous func-*

1.The term "binary" is just an fancy way of saying that the function has *two* arguments.

tion of one (real) variable" [see References 8, 9, or 10 if you are stumped]. (b) Demonstrate that axioms A7 and A8 hold true. (c) The Taylor series expansion writes a function as a linear combination of more simple functions, namely, the powers x^n, for n ranging from 0 to ∞. The Fourier series expansion writes a function as a linear combination of trigonometric functions. (d) The dyad $\underset{\sim}{\mathbf{f}}\underset{\sim}{\mathbf{g}}$ has components $f_i g_j$, where the indices i and j range from 1 to 3. The function f(x)g(y) has values defined for values of the independent variables x and y ranging from $-\infty$ to ∞. A dyad $\underset{\sim}{\mathbf{f}}\underset{\sim}{\mathbf{g}}$ is the most primitive of tensors, and it is defined such that it has no meaning (other than book-keeping) unless it operates on an arbitrary vector so that $\underset{\sim}{\mathbf{f}}\underset{\sim}{\mathbf{g}} \bullet \underset{\sim}{\mathbf{v}}$ means $\underset{\sim}{\mathbf{f}}(\underset{\sim}{\mathbf{g}} \bullet \underset{\sim}{\mathbf{v}})$. Likewise, we can define the most primitive function of two variables, f(x)g(y) so that it takes on meaning when appearing in an operation on a general function $v(y)$ so that f(x)g(y) is defined so that it means $\int_{-\infty}^{\infty} f(x)g(y)v(y)dy = f(x)\int_{-\infty}^{\infty} g(y)v(y)dy$. (e) see previous answer (f).

When addressing the question of whether or not something is indeed a tensor, you must commit yourself to which of the definitions discussed on page 55 you wish to use. When we cover the topic of curvilinear calculus, we will encounter the Christoffel symbols Γ_{ij}^k and Γ_{kij}. These three-index quantities characterize how the *natural* curvilinear basis varies in space. Their definition is based upon your choice of basis, $\{\underset{\sim}{\mathbf{g}}_1, \underset{\sim}{\mathbf{g}}_2, \underset{\sim}{\mathbf{g}}_3\}$. Naturally it stands to reason that choosing some other basis will still permit you to construct Christoffel symbols for that system. Any review of the literature will include a statement that the Christoffel symbols "are not third-order tensors." This statement merely means that

$$\Gamma_{ijB}^{BBk} \neq \Gamma_{pqA}^{AAr} g_{ip}^{BA} g_{jq}^{BA} g_{BA}^{kr} \tag{4.39}$$

Note that it is perfectly acceptable for you to construct a *distinct* tensors defined

$$\underset{\approx}{\Gamma}^A = \Gamma_{ijA}^{AAk} \underset{\sim}{\mathbf{g}}_A^m \underset{\sim}{\mathbf{g}}_A^n \underset{\sim}{\mathbf{g}}_p^A \quad \text{and} \quad \underset{\approx}{\Gamma}^B = \Gamma_{ijB}^{BBk} \underset{\sim}{\mathbf{g}}_B^m \underset{\sim}{\mathbf{g}}_B^n \underset{\sim}{\mathbf{g}}_p^B \tag{4.40}$$

Equation (4.39) tells us that these two tensors are not equal. That is,

$$\underset{\approx}{\Gamma}^A \neq \underset{\approx}{\Gamma}^B \tag{4.41}$$

Stated differently, for each basis, there exists a *basis-dependent* Christoffel tensor. This should not be disturbing. After all, the base vectors *themselves* are, by definition, basis-dependent, but that doesn't mean they aren't vectors. Changing the basis will change the associated Christoffel tensor. A particular choice for the basis is itself (of course) basis dependent -- change the basis, and you will (obviously) change the base vectors themselves. You can always construct other vectors and tensors from the basis, but you would naturally expect these tensors to change upon a change of basis if the new definition of the tensor is defined the same as the old definition except that the new base vectors are used. What's truly remarkable is that there exist certain combinations of base vectors and basis-referenced definitions of components that turn out to be *invariant* under a change of basis.

Consider two tensors constructed from the metrics g_{ij}^{AA} and g_{ij}^{BB}:

$$\underset{\approx}{\mathbf{G}}^{AA} = g_{ij}^{AA}\underset{\sim}{\mathbf{g}}_A^i\underset{\sim}{\mathbf{g}}_A^j \qquad \text{and} \qquad \underset{\approx}{\mathbf{G}}^{BB} = g_{ij}^{BB}\underset{\sim}{\mathbf{g}}_B^i\underset{\sim}{\mathbf{g}}_B^j \tag{4.42}$$

These two tensors, it turns out, are equal even though they were constructed using *different* bases and *different* components! Even though $\underset{\sim}{\mathbf{g}}_A^i \neq \underset{\sim}{\mathbf{g}}_B^i$ and even though $g_{ij}^{AA} \neq g_{ij}^{BB}$, it turns out that the differences cancel each other out in the combinations defined in Eq. (4.42) so that both tensors are equal. In fact, as was shown in Eq. (3.18), these tensors are simply the identity tensor!

$$\underset{\approx}{\mathbf{G}}^{AA} = \underset{\approx}{\mathbf{G}}^{BB} = \underset{\approx}{\mathbf{I}} \tag{4.43}$$

This is an exceptional situation. More often than not, when you construct a tensor by multiplying an ordered array of numbers by corresponding base vectors in different systems, you will end up with two different tensors. To clarify, suppose that f is a function (scalar, vector, or tensor valued) which can be constructed from the basis. Presume that the *same* function can be applied to any basis, and furthermore presume that the results will be different, so we must denote the results by different symbols.

$$\underset{\sim}{F}^A = f(\underset{\sim}{\mathbf{g}}_1^A, \underset{\sim}{\mathbf{g}}_2^A, \underset{\sim}{\mathbf{g}}_3^A) \text{ and if } \underset{\sim}{F}^B = f(\underset{\sim}{\mathbf{g}}_1^B, \underset{\sim}{\mathbf{g}}_2^B, \underset{\sim}{\mathbf{g}}_3^B) \tag{4.44}$$

In the older literature, the components of $\underset{\sim}{F}^A$ were called tensor if and only if $\underset{\sim}{F}^A = \underset{\sim}{F}^B$. In more modern treatments, the *function* f is called basis invariant if $\underset{\sim}{F}^A = \underset{\sim}{F}^B$, whereas the function f is basis-*dependent* if $\underset{\sim}{F}^A \neq \underset{\sim}{F}^B$. These two tensors are equal only if the components of $\underset{\sim}{F}^A$ with respect to one basis happen to equal the components of $\underset{\sim}{F}^B$ with respect to the *same* basis. It's not uncommon for two distinct tensors to have the same components with respect to *different* bases — such a result would *not* imply equality of the two tensors.

4.2 Coordinates are not the same thing as components

All base vectors in this section are to be regarded as the natural basis so that they are related to the coordinates η^k by $\underset{\sim}{\mathbf{g}}_k = \partial\underset{\sim}{\mathbf{x}}/\partial\eta^k$. Recall that we have typeset the three coordinates as η^k. Despite this typesetting, these numbers are *not* contravariant components of any vector $\underset{\sim}{\eta}$. By this we mean that if you compare a different set of components $\bar{\eta}^k$ from a second system, they are *not* generally related to the original set of components via a vector transformation rule. We may only presume that transformation rules exist such that each $\bar{\eta}^k$ can be expressed as a function of the η^m coordinates. If coordinates were also vector components, then they would have to be related by a transformation formula of the form

$$\bar{\eta}^k = \frac{\partial\bar{\eta}^k}{\partial\eta^m}\eta^m \qquad \leftarrow \quad \textit{valid for homogeneous, but not curvilinear coordinates} \tag{4.45}$$

A coordinate transformation of this form is a *linear*, which holds only for homogenous (i.e., *non*-curvilinear) coordinate systems. Even though the η^k coordinates are not generally components of a vector, their *increments* $d\eta^k$ do transform like vectors. That is,

$$d\bar{\eta}^k = \frac{\partial\bar{\eta}^k}{\partial\eta^m}d\eta^m \quad \leftarrow \quad \text{this } \underline{\underline{\text{is}}} \text{ true for both homogenous and curvilinear coordinates!} \tag{4.46}$$

For curvilinear systems, the transformation formulas that express coordinates $\bar{\eta}^k$ in terms of η^m are *nonlinear* functions, but their *increments* are linear for the same reason that the increment of a nonlinear function $y = f(x)$ is an expression that is linear with respect to increments: $dy = f'(x)dx$. Geometrically, at any point on a curvy line, we can construct a *straight* tangent to that line.

4.3 Do there exist a "complementary" or "dual" coordinates?

Recall that the coordinates are typeset as η^k. We've explained that the coordinate *increments* $d\eta^k$ transform as vectors even though the coordinates themselves are not components of vectors. The $d\eta^k$ increments are the components of the position increment vector $d\underset{\sim}{\mathbf{x}}$. In other words,

$$(d\underset{\sim}{\mathbf{x}})^k = d\eta^k \tag{4.47}$$

To better emphasize the order of operation, we should write this as

$$(d\underset{\sim}{\mathbf{x}})^k = d(\eta^k) \tag{4.48}$$

On the left side, we have the k^{th} component of the vector differential, whereas on the right side we have the differential of the k^{th} coordinate.

The covariant component of the position increment is well-defined:

$$(d\underset{\sim}{\mathbf{x}})_i = g_{ik}(d\underset{\sim}{\mathbf{x}})^k \tag{4.49}$$

Looking back at Eq. (4.48), natural question would be: is it possible to define complementary or dual coordinates η_i that are related to the baseline coordinates η^i such that

$$(d\underset{\sim}{\mathbf{x}})_i = d(\eta_i) \quad \leftarrow \quad \text{this is NOT possible in general} \tag{4.50}$$

We will now prove (by contradiction) that such dual coordinates do *not* generally exist. If such coordinates did exist, then it would have to be true that

$$d(\eta_i) = g_{ij}d(\eta^j) \tag{4.51}$$

For the η_i to exist, this equation would need to be integrable. In other words, the expression $g_{ij}d(\eta^j)$ would have to be an exact differential, which means that the metrics would have to be given by

$$g_{ij} = \frac{\partial\eta_i}{\partial\eta^j} \tag{4.52}$$

and therefore, the second partials would need to be interchangeable:

$$\frac{\partial g_{ij}}{\partial \eta^k} = \frac{\partial^2 \eta_i}{\partial \eta^j \partial \eta^k} = \frac{\partial^2 \eta_i}{\partial \eta^k \partial \eta^j} = \frac{\partial g_{ik}}{\partial \eta^j} \tag{4.53}$$

This constraint is not satisfied by curvilinear systems. Consider, for example, cylindrical coordinates, $(\eta^1 = r,\ \eta^2 = \theta,\ \eta^3 = z)$, where the metric matrix is

$$[g_{ij}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.54}$$

Note that

$$\frac{\partial g_{22}}{\partial \eta^1} = \frac{\partial g_{22}}{\partial r} = 2r, \quad \text{but} \quad \frac{\partial g_{21}}{\partial \eta^2} = \frac{\partial g_{21}}{\partial \theta} = 0 \tag{4.55}$$

Hence, since these are not equal, the integrability condition of Eq. (4.53) is violated and therefore there do *not* exist dual coordinates.

5. Curvilinear calculus

Chapter 3 focused on algebra, where the only important issue was the possibility that the basis might be irregular (i.e., nonorthogonal, nonnormalized, and/or non-right-handed). In that chapter, it made no difference whether the coordinates are homogeneous or curvilinear. For homogeneous coordinates (where the base vectors are the same at all points in space), tensor calculus formulas for the gradient are similar to the formulas for ordinary rectangular Cartesian coordinates. By contrast, when the coordinates are curvilinear, there are new additional terms in gradient formulas to account for the variation of the base vectors with position.

5.1 A introductory example

Before developing the general curvilinear theory, it is useful to first show how to develop the formulas *without* using the full power of curvilinear calculus. Suppose an engineering problem is most naturally described using cylindrical coordinates, which are related to the laboratory Cartesian coordinates by

$$x_1 = r\cos\theta, \quad x_2 = r\sin\theta, \quad x_3 = z. \tag{5.1}$$

The base vectors for cylindrical coordinates are related to the lab base vectors by

$$\begin{aligned} \underset{\sim}{\mathbf{e}}_r &= \cos\theta\,\underset{\sim}{\mathbf{e}}_1 + \sin\theta\,\underset{\sim}{\mathbf{e}}_2 \\ \underset{\sim}{\mathbf{e}}_\theta &= -\sin\theta\,\underset{\sim}{\mathbf{e}}_1 + \cos\theta\,\underset{\sim}{\mathbf{e}}_2 \\ \underset{\sim}{\mathbf{e}}_z &= \underset{\sim}{\mathbf{e}}_3\,. \end{aligned} \tag{5.2}$$

Suppose you need the gradient of a scalar, $ds/d\underset{\sim}{\mathbf{x}}$. If s is written as a function of x_i, then the familiar Cartesian formula applies:

$$\frac{ds}{d\underset{\sim}{\mathbf{x}}} = \frac{\partial s}{\partial x_1}\underset{\sim}{\mathbf{e}}_1 + \frac{\partial s}{\partial x_2}\underset{\sim}{\mathbf{e}}_2 + \frac{\partial s}{\partial x_3}\underset{\sim}{\mathbf{e}}_3\,. \tag{5.3}$$

Because curvilinear coordinates were selected, the function $s(r, \theta, z)$ is probably simple in form. If you were a sucker for punishment, you could substitute the inverse relationships,

$$r = \sqrt{x_1^2 + x_1^2}, \quad \theta = \tan^{-1}\left(\frac{x_2}{x_1}\right), \quad z = x_3\,. \tag{5.4}$$

Then you'd have the function $s(x_1, x_2, x_3)$, with which you could directly apply Eq. (5.3). This approach is unsatisfactory for several reasons. Strictly speaking, Eq. (5.4) is incorrect because the arctangent has *two* solutions in the range from 0 to 2π; the correct formula would have to be the two-argument arctangent. Furthermore, actually computing the derivative would be tedious, and the final result would be expressed in terms of the lab Cartesian basis instead of the cylindrical basis.

The better approach is to look at the problem from a more academic slant by applying the chain rule. Given that $s = s(r, \theta, z)$ then

$$\frac{ds}{d\underset{\sim}{\mathbf{x}}} = \left(\frac{\partial s}{\partial r}\right)_{\theta, z} \frac{dr}{d\underset{\sim}{\mathbf{x}}} + \left(\frac{\partial s}{\partial \theta}\right)_{r, z} \frac{d\theta}{d\underset{\sim}{\mathbf{x}}} + \left(\frac{\partial s}{\partial z}\right)_{r, \theta} \frac{dz}{d\underset{\sim}{\mathbf{x}}}, \tag{5.5}$$

where the subscripts indicate which variables are held constant in the derivatives. The derivatives of the cylindrical coordinates with respect to $\underset{\sim}{\mathbf{x}}$ can be computed *a priori*. For example, the quantity $dr/d\underset{\sim}{\mathbf{x}}$ is the gradient of r. Physically, we know that the gradient of a quantity is perpendicular to surfaces of constant values of that quantity. Surfaces of constant r are cylinders of radius r, so we know that $dr/d\underset{\sim}{\mathbf{x}}$ must be perpendicular to the cylinder. In other words, we know that $dr/d\underset{\sim}{\mathbf{x}}$ must be parallel to $\underset{\sim}{\mathbf{e}}_r$. The gradients of the cylindrical coordinates are obtained by applying the ordinary Cartesian formula:

$$\begin{aligned}
\frac{dr}{d\underset{\sim}{\mathbf{x}}} &= \frac{\partial r}{\partial x_1}\underset{\sim}{\mathbf{e}}_1 + \frac{\partial r}{\partial x_2}\underset{\sim}{\mathbf{e}}_2 + \frac{\partial r}{\partial x_3}\underset{\sim}{\mathbf{e}}_3 \\
\frac{d\theta}{d\underset{\sim}{\mathbf{x}}} &= \frac{\partial \theta}{\partial x_1}\underset{\sim}{\mathbf{e}}_1 + \frac{\partial \theta}{\partial x_2}\underset{\sim}{\mathbf{e}}_2 + \frac{\partial \theta}{\partial x_3}\underset{\sim}{\mathbf{e}}_3 \\
\frac{dz}{d\underset{\sim}{\mathbf{x}}} &= \frac{\partial z}{\partial x_1}\underset{\sim}{\mathbf{e}}_1 + \frac{\partial z}{\partial x_2}\underset{\sim}{\mathbf{e}}_2 + \frac{\partial z}{\partial x_3}\underset{\sim}{\mathbf{e}}_3 .
\end{aligned} \tag{5.6}$$

We must determine these derivatives *implicitly* by using Eqs. (5.1) to first compute the derivatives of the Cartesian coordinates with respect to the cylindrical coordinates. We can arrange the nine possible derivatives of Eqs. (5.1) in a matrix:

$$\begin{bmatrix} \frac{\partial x_1}{\partial r} & \frac{\partial x_1}{\partial \theta} & \frac{\partial x_1}{\partial z} \\ \frac{\partial x_2}{\partial r} & \frac{\partial x_2}{\partial \theta} & \frac{\partial x_2}{\partial z} \\ \frac{\partial x_3}{\partial r} & \frac{\partial x_3}{\partial \theta} & \frac{\partial x_3}{\partial z} \end{bmatrix} = \begin{bmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{5.7}$$

The inverse derivatives are obtained inverting this matrix to give

$$\begin{bmatrix} \frac{\partial r}{\partial x_1} & \frac{\partial r}{\partial x_2} & \frac{\partial r}{\partial x_3} \\ \frac{\partial \theta}{\partial x_1} & \frac{\partial \theta}{\partial x_2} & \frac{\partial \theta}{\partial x_3} \\ \frac{\partial z}{\partial x_1} & \frac{\partial z}{\partial x_2} & \frac{\partial z}{\partial x_3} \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\frac{\sin\theta}{r} & \frac{\cos\theta}{r} & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{5.8}$$

Substituting these derivatives into Eq. (5.6) gives

$$\frac{dr}{d\underset{\sim}{\mathbf{x}}} = \cos\theta\,\underset{\sim}{\mathbf{e}}_1 + \sin\theta\,\underset{\sim}{\mathbf{e}}_2$$
$$\frac{d\theta}{d\underset{\sim}{\mathbf{x}}} = -\frac{\sin\theta}{r}\underset{\sim}{\mathbf{e}}_1 + \frac{\cos\theta}{r}\underset{\sim}{\mathbf{e}}_2$$
$$\frac{dz}{d\underset{\sim}{\mathbf{x}}} = \underset{\sim}{\mathbf{e}}_3\,. \tag{5.9}$$

Referring to Eq. (5.2) we conclude that

$$\boxed{\begin{aligned}\frac{dr}{d\underset{\sim}{\mathbf{x}}} &= \underset{\sim}{\mathbf{e}}_r\\ \frac{d\theta}{d\underset{\sim}{\mathbf{x}}} &= \frac{\underset{\sim}{\mathbf{e}}_\theta}{r}\\ \frac{dz}{d\underset{\sim}{\mathbf{x}}} &= \underset{\sim}{\mathbf{e}}_3\end{aligned}} \tag{5.10}$$

This result was rather tedious to derive, but it was a one-time-only task. For any coordinate system, it is always a good idea to derive the coordinate gradients and save them for later use.

Now that we have the coordinate gradients, Eq. (5.5) becomes

$$\boxed{\frac{ds}{d\underset{\sim}{\mathbf{x}}} = s_{,r}\underset{\sim}{\mathbf{e}}_r + \frac{s_{,\theta}}{r}\underset{\sim}{\mathbf{e}}_\theta + s_{,z}\underset{\sim}{\mathbf{e}}_3}\,, \tag{5.11}$$

where the commas are a shorthand notation for derivatives. This is the formula for the gradient of a scalar that you would find in a handbook.

The above analysis showed that high-powered curvilinear theory is not necessary to derive the formula for the gradient of a scalar function of the cylindrical coordinates. All that's needed is knowledge of tensor analysis in Cartesian coordinates. However, for more complicated coordinate systems, a fully developed curvilinear theory is indispensable.

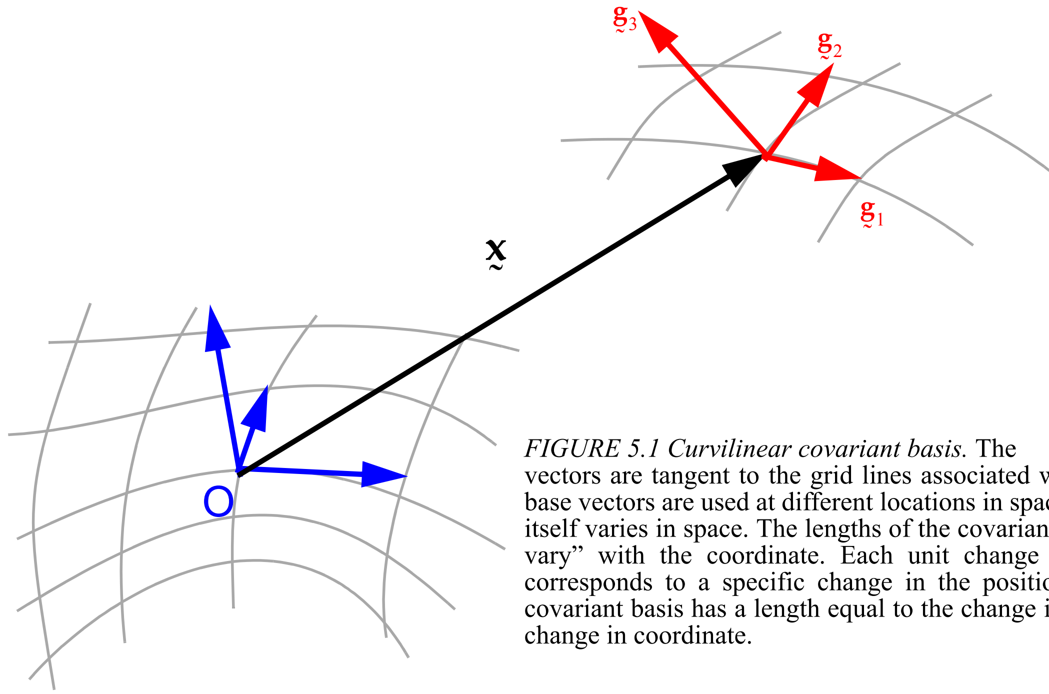
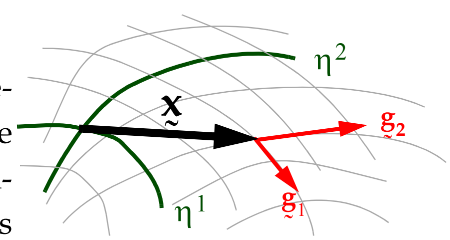


FIGURE 5.1 Curvilinear covariant basis. The covariant base vectors are tangent to the grid lines associated with η^k. Different base vectors are used at different locations in space because the grid itself varies in space. The lengths of the covariant base vectors "co-vary" with the coordinate. Each unit change in the coordinate corresponds to a specific change in the position vector, and the covariant basis has a length equal to the change in position per unit change in coordinate.

5.2 Curvilinear coordinates



In what follows, three coordinates $\{\eta^1, \eta^2, \eta^3\}$ are presumed to identify the position of a point $\mathbf{x}$ in space. These coordinates are identified with superscripts merely by convention. As sketched in Figure 5.1, the associated base vectors at any point in space are tangent to the grid lines at that point. The base vector $\mathbf{g}_k$ points in the direction of increasing η^k. Importantly, the position vector $\mathbf{x}$ is *not* generally equal to $\eta^k \mathbf{g}_k$. For example, the position vector for spherical coordinates is simply $\mathbf{x} = r\mathbf{e}_r$ (*not* $\mathbf{x} = r\mathbf{e}_r + \theta\mathbf{e}_\theta + \phi\mathbf{e}_\psi$); for spherical coordinates, the dependence of the position vector on the coordinates θ and ϕ is hidden in dependence of $\mathbf{e}_r$ on θ and ϕ.

5.3 The "associated" curvilinear covariant basis

Given three coordinates $\{\eta^1, \eta^2, \eta^3\}$ that identify the location in space, the associated covariant base vectors are defined to be tangent to the grid lines along which only one coordinate varies (the others being held constant). By definition, the three coordinates $\{\eta^1, \eta^2, \eta^3\}$ identify the position in space:

$$\mathbf{x} = \mathbf{x}(\eta^1, \eta^2, \eta^3); \quad \textit{i.e.,}\ \mathbf{x} \text{ is a function of } \eta^1, \eta^2, \text{ and } \eta^3. \tag{5.12}$$

The i^{th} covariant base vector is defined to point in the direction that $\mathbf{x}$ moves when η^k is increased, holding the other coordinates constant. Hence, the natural definition of the i^{th} covariant basis is the partial derivative of $\underset{\sim}{\mathbf{x}}$ with respect to η^i:

$$\boxed{\underset{\sim}{\mathbf{g}}_i \equiv \frac{\partial \underset{\sim}{\mathbf{x}}}{\partial \eta^i}}. \tag{5.13}$$

Note the following new summation convention: the superscript on η^i is interpreted as a *subscript* in the final result because η^i appears in the "denominator" of the derivative.

Note that the natural covariant base vectors are not necessarily unit vectors. Furthermore, because the coordinates $\{\eta^1, \eta^2, \eta^3\}$ do not necessarily have the same physical units, the natural base vectors themselves will have different physical units. For example, cylindrical coordinates $\{r, \theta, z\}$ have dimensions of length, radians, and length, respectively. Consequently, two of the base vectors ($\underset{\sim}{\mathbf{g}}_1$ and $\underset{\sim}{\mathbf{g}}_3$) are dimensionless, but $\underset{\sim}{\mathbf{g}}_2$ has dimensions of length. Such a situation is not unusual and should not be alarming. We will find that the *components* of a vector will have physical dimensions appropriate to ensure that the each term in the sum of components times base vectors will have the same physical dimensions as the vector itself. Again this point harks back to the fact that neither components nor base vectors are invariant, but the *sum* of components times base vectors is invariant.

Equation (5.12) states that the coordinates uniquely identify the position in space. Conversely, any position in space corresponds to a unique set of coordinates. That is, each coordinate may be regarded as a single-valued function of position vector:

$$\eta^j = \eta^j(\underset{\sim}{\mathbf{x}}). \tag{5.14}$$

By the chain rule, note that

$$\frac{d\eta^j}{d\underset{\sim}{\mathbf{x}}} \bullet \frac{\partial \underset{\sim}{\mathbf{x}}}{\partial \eta^i} = \frac{\partial \eta^j}{\partial \eta^i} = \delta^j_i. \tag{5.15}$$

Therefore, recalling Eq. (5.13), the contravariant dual basis must be given by

$$\boxed{\underset{\sim}{\mathbf{g}}^j \equiv \frac{\partial \eta^j}{\partial \underset{\sim}{\mathbf{x}}}}. \tag{5.16}$$

This derivative is the spatial gradient of η^j. Hence, as sketched in Fig. 5.2, each contravariant base vector $\underset{\sim}{\mathbf{g}}^j$ is normal to surfaces of constant η^j, and it points in the direction of increasing η^j.

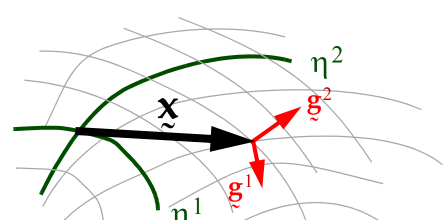


FIGURE 5.2 Curvilinear contravariant basis. The contravariant base vectors are normal to surfaces of constant η^k.

Starting with Eq. (5.12), the increment in the position vector is given by

$$d\underset{\sim}{\mathbf{x}} = \frac{\partial \underset{\sim}{\mathbf{x}}}{\partial \eta^k} d\eta^k, \tag{5.17}$$

or, recalling Eq. (5.13),

$$d\underset{\sim}{\mathbf{x}} = d\eta^k \, \underset{\sim}{\mathbf{g}}_k \,. \tag{5.18}$$

Now we note a key difference between homogeneous and curvilinear coordinates:

ix. For homogeneous coordinates, each of the $\underset{\sim}{\mathbf{g}}_k$ base vectors is same throughout space — they are independent of the coordinates $\{\eta^1, \eta^2, \eta^3\}$. Hence, for homogeneous coordinates, Eq. (5.18) can be integrated to show that $\underset{\sim}{\mathbf{x}} = \eta^k \, \underset{\sim}{\mathbf{g}}_k$.

x. For curvilinear coordinates, each $\underset{\sim}{\mathbf{g}}_k$ varies in space — they depend on the coordinates $\{\eta^1, \eta^2, \eta^3\}$. Hence, for curvilinear coordinates, Eq. (5.18) does *not* imply that $\underset{\sim}{\mathbf{x}} = \eta^k \, \underset{\sim}{\mathbf{g}}_k$.

EXAMPLE Consider cylindrical coordinates: $\eta^1 = r$, $\eta^2 = \theta$, $\eta^3 = z$.

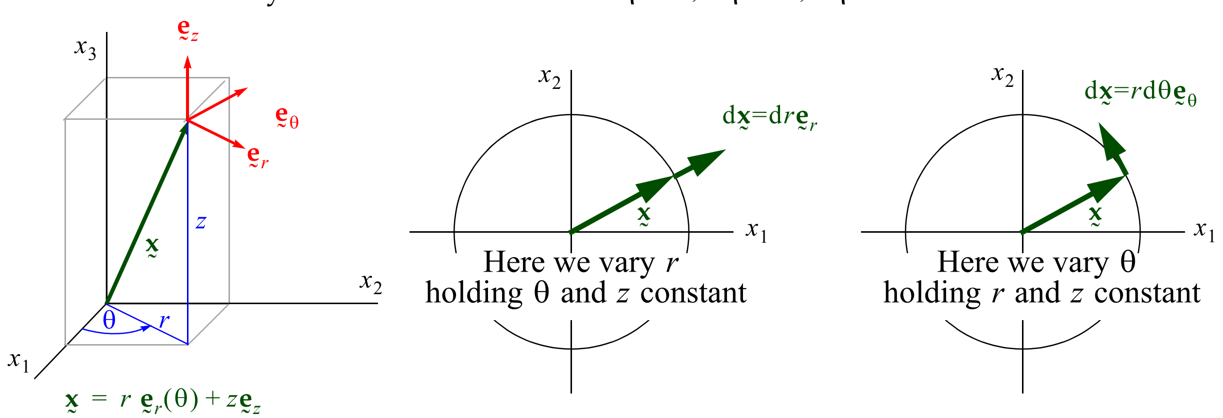


FIGURE 5.3 Covariant basis for cylindrical coordinates. The views down the z-axis show how the covariant base vectors can be determined graphically by varying one coordinate, holding the others constant.

Figure 5.3 illustrates how, for simple enough coordinate systems, you can determine the covariant base vectors graphically. Recall that $\underset{\sim}{\mathbf{g}}_1 = (\partial\underset{\sim}{\mathbf{x}}/\partial\eta^1)_{\eta^2,\,\eta^3}$ and therefore that $\underset{\sim}{\mathbf{g}}_1$ is the coefficient of $d\eta^1$ when $\underset{\sim}{\mathbf{x}}$ is varied by differentially changing η^1 holding η^2 and η^3 constant. For cylindrical coordinates, when the radial coordinate, $\eta^1 = r$, is varied holding the others constant, the position vector $\underset{\sim}{\mathbf{x}}$ moves such $d\underset{\sim}{\mathbf{x}} = dr\underset{\sim}{\mathbf{e}}_r$, and therefore $\underset{\sim}{\mathbf{g}}_1$ must equal $\underset{\sim}{\mathbf{e}}_r$ because it is the coefficient of dr. Similarly, $\underset{\sim}{\mathbf{g}}_2$ must be equal to $r\underset{\sim}{\mathbf{e}}_\theta$ since that is the coefficient of $d\theta$ in $d\underset{\sim}{\mathbf{x}}$ when the second coordinate, $\eta^2 = \theta$, is varied holding the others constant. Summarizing,

$$\underset{\sim}{\mathbf{g}}_1 = \underset{\sim}{\mathbf{e}}_r, \qquad \underset{\sim}{\mathbf{g}}_2 = r\underset{\sim}{\mathbf{e}}_\theta, \qquad \text{and} \qquad \underset{\sim}{\mathbf{g}}_3 = \underset{\sim}{\mathbf{e}}_z \,. \tag{5.19}$$

To derive these results analytically (rather than geometrically), we utilize the underlying rectangular Cartesian basis $\{\underset{\sim}{\mathbf{e}}_1, \underset{\sim}{\mathbf{e}}_2, \underset{\sim}{\mathbf{e}}_3\}$ to write the position vector as

$$\underset{\sim}{\mathbf{x}} = x_1\underset{\sim}{\mathbf{e}}_1 + x_2\underset{\sim}{\mathbf{e}}_2 + x_3\underset{\sim}{\mathbf{e}}_3 \,, \tag{5.20}$$

where

$$x_1 = r\cos\theta \tag{5.21a}$$

$$x_2 = r\sin\theta \tag{5.21b}$$

$$x_3 = z\,. \tag{5.21c}$$

Then

$$\underset{\sim}{\mathbf{g}}_1 = \left(\frac{\partial\underset{\sim}{\mathbf{x}}}{\partial r}\right)_{\theta,\,z} = \left(\frac{\partial x_1}{\partial r}\right)_{\theta,\,z}\underset{\sim}{\mathbf{e}}_1 + \left(\frac{\partial x_2}{\partial r}\right)_{\theta,\,z}\underset{\sim}{\mathbf{e}}_2 + \left(\frac{\partial x_3}{\partial r}\right)_{\theta,\,z}\underset{\sim}{\mathbf{e}}_3 = \cos\theta\,\underset{\sim}{\mathbf{e}}_1 + \sin\theta\,\underset{\sim}{\mathbf{e}}_2 \tag{5.22a}$$

$$\underset{\sim}{\mathbf{g}}_2 = \left(\frac{\partial\underset{\sim}{\mathbf{x}}}{\partial \theta}\right)_{r,\,z} = \left(\frac{\partial x_1}{\partial \theta}\right)_{r,\,z}\underset{\sim}{\mathbf{e}}_1 + \left(\frac{\partial x_2}{\partial \theta}\right)_{r,\,z}\underset{\sim}{\mathbf{e}}_2 + \left(\frac{\partial x_3}{\partial \theta}\right)_{r,\,z}\underset{\sim}{\mathbf{e}}_3 = -r\sin\theta\,\underset{\sim}{\mathbf{e}}_1 + r\cos\theta\,\underset{\sim}{\mathbf{e}}_2 \tag{5.22b}$$

$$\underset{\sim}{\mathbf{g}}_3 = \left(\frac{\partial\underset{\sim}{\mathbf{x}}}{\partial z}\right)_{r,\,\theta} = \left(\frac{\partial x_1}{\partial z}\right)_{r,\,\theta}\underset{\sim}{\mathbf{e}}_1 + \left(\frac{\partial x_2}{\partial z}\right)_{r,\,\theta}\underset{\sim}{\mathbf{e}}_2 + \left(\frac{\partial x_3}{\partial z}\right)_{r,\,\theta}\underset{\sim}{\mathbf{e}}_3 = \underset{\sim}{\mathbf{e}}_3\,. \tag{5.22c}$$

which we note is equivalent to the graphically derived Eqs. (5.19). Also note that $\underset{\sim}{\mathbf{g}}_1$ and $\underset{\sim}{\mathbf{g}}_3$ are dimensionless, whereas $\underset{\sim}{\mathbf{g}}_2$ has physical dimensions of length.

The metric coefficients g_{ij} for cylindrical coordinates are derived in the usual way:

$$g_{11} = \underset{\sim}{\mathbf{g}}_1 \bullet \underset{\sim}{\mathbf{g}}_1 = 1 \qquad g_{12} = \underset{\sim}{\mathbf{g}}_1 \bullet \underset{\sim}{\mathbf{g}}_2 = 0 \qquad g_{13} = \underset{\sim}{\mathbf{g}}_1 \bullet \underset{\sim}{\mathbf{g}}_3 = 0 \tag{5.23a}$$

$$g_{12} = \underset{\sim}{\mathbf{g}}_2 \bullet \underset{\sim}{\mathbf{g}}_1 = 0 \qquad g_{22} = \underset{\sim}{\mathbf{g}}_2 \bullet \underset{\sim}{\mathbf{g}}_2 = r^2 \qquad g_{23} = \underset{\sim}{\mathbf{g}}_2 \bullet \underset{\sim}{\mathbf{g}}_3 = 0 \tag{5.23b}$$

$$g_{31} = \underset{\sim}{\mathbf{g}}_3 \bullet \underset{\sim}{\mathbf{g}}_1 = 0 \qquad g_{32} = \underset{\sim}{\mathbf{g}}_3 \bullet \underset{\sim}{\mathbf{g}}_2 = 0 \qquad g_{33} = \underset{\sim}{\mathbf{g}}_3 \bullet \underset{\sim}{\mathbf{g}}_3 = 1\,, \tag{5.23c}$$

or, in matrix form,

$$[g_{ij}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \qquad \leftarrow \text{ cylindrical coordinates} \tag{5.24}$$

Whenever the metric tensor comes out to be diagonal as it has here, the coordinate system is orthogonal and the base vectors at each point in space are mutually perpendicular. Inverting the covariant $[g_{ij}]$ matrix gives the contravariant metric coefficients:

$$[g^{ij}] = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/r^2 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \qquad \leftarrow \text{ cylindrical coordinates} \tag{5.25}$$

The contravariant dual basis can be derived in the usual way by applying the formula of Eq (2.18):

$$\underset{\sim}{\mathbf{g}}^1 = g^{11}\underset{\sim}{\mathbf{g}}_1 + g^{12}\underset{\sim}{\mathbf{g}}_2 + g^{13}\underset{\sim}{\mathbf{g}}_3 = \underset{\sim}{\mathbf{g}}_1 = \underset{\sim}{\mathbf{e}}_r \tag{5.26a}$$

$$\underset{\sim}{\mathbf{g}}^2 = g^{21}\underset{\sim}{\mathbf{g}}_1 + g^{22}\underset{\sim}{\mathbf{g}}_2 + g^{23}\underset{\sim}{\mathbf{g}}_3 = \frac{\underset{\sim}{\mathbf{g}}_2}{r^2} = \frac{\underset{\sim}{\mathbf{e}}_\theta}{r} \tag{5.26b}$$

$$\underset{\sim}{\mathbf{g}}^3 = g^{31}\underset{\sim}{\mathbf{g}}_1 + g^{32}\underset{\sim}{\mathbf{g}}_2 + g^{33}\underset{\sim}{\mathbf{g}}_3 = \underset{\sim}{\mathbf{g}}_3 = \underset{\sim}{\mathbf{e}}_z. \tag{5.26c}$$

With considerably more effort, we can alternatively derive these results by directly applying Eq. (5.16). To do this, we must express the coordinates $\{r, \theta, z\}$ in terms of $\{x_1, x_2, x_3\}$:

$$r = \sqrt{x_1^2 + x_2^2} \tag{5.27a}$$

$$\theta = \tan^{-1}\left(\frac{x_2}{x_1}\right) \tag{5.27b}$$

$$z = x_3. \tag{5.27c}$$

Then applying Eq. (5.16) gives

$$\underset{\sim}{\mathbf{g}}^1 = \frac{dr}{d\underset{\sim}{\mathbf{x}}} = \frac{\partial r}{\partial x_1}\underset{\sim}{\mathbf{e}}_1 + \frac{\partial r}{\partial x_2}\underset{\sim}{\mathbf{e}}_2 + \frac{\partial r}{\partial x_3}\underset{\sim}{\mathbf{e}}_3 = \frac{x_1}{r}\underset{\sim}{\mathbf{e}}_1 + \frac{x_2}{r}\underset{\sim}{\mathbf{e}}_2 = \cos\theta\,\underset{\sim}{\mathbf{e}}_1 + \sin\theta\,\underset{\sim}{\mathbf{e}}_2 = \underset{\sim}{\mathbf{e}}_r \tag{5.28a}$$

$$\underset{\sim}{\mathbf{g}}^2 = \frac{d\theta}{d\underset{\sim}{\mathbf{x}}} = \frac{\partial \theta}{\partial x_1}\underset{\sim}{\mathbf{e}}_1 + \frac{\partial \theta}{\partial x_2}\underset{\sim}{\mathbf{e}}_2 + \frac{\partial \theta}{\partial x_3}\underset{\sim}{\mathbf{e}}_3 = \left(\frac{-x_2}{x_1^2}\right)\cos^2\theta\,\underset{\sim}{\mathbf{e}}_1 + \left(\frac{1}{x_1}\right)\cos^2\theta\,\underset{\sim}{\mathbf{e}}_2 = \frac{1}{r}\underset{\sim}{\mathbf{e}}_\theta \tag{5.28b}$$

$$\underset{\sim}{\mathbf{g}}^3 = \frac{dz}{d\underset{\sim}{\mathbf{x}}} = \frac{\partial z}{\partial x_1}\underset{\sim}{\mathbf{e}}_1 + \frac{\partial z}{\partial x_2}\underset{\sim}{\mathbf{e}}_2 + \frac{\partial z}{\partial x_3}\underset{\sim}{\mathbf{e}}_3 = \underset{\sim}{\mathbf{e}}_3 = \underset{\sim}{\mathbf{e}}_z. \tag{5.28c}$$

which agrees with the previous result in Eq. (5.26).

Because the curvilinear basis $\{\underset{\sim}{\mathbf{g}}_1, \underset{\sim}{\mathbf{g}}_2, \underset{\sim}{\mathbf{g}}_3\}$ associated with cylindrical coordinates is orthogonal, the dual vectors $\{\underset{\sim}{\mathbf{g}}^1, \underset{\sim}{\mathbf{g}}^2, \underset{\sim}{\mathbf{g}}^3\}$ are in exactly the same directions but have different magnitudes. It is common practice to use the associated ortho*normal* basis $\{\underset{\sim}{\mathbf{e}}_r, \underset{\sim}{\mathbf{e}}_\theta, \underset{\sim}{\mathbf{e}}_z\}$ as we have in the above equations. This practice also eliminates the sometimes confusing complication of dimensional base vectors. When a coordinate system is orthogonal, the shared orthonormal basis is called the "**physical basis**." Notationally, vector components $\{v_1, v_2, v_3\}$ or $\{v^1, v^2, v^3\}$ are defined $v_k \equiv \underset{\sim}{\mathbf{g}}_k \bullet \underset{\sim}{\mathbf{v}}$ and $v^k \equiv \underset{\sim}{\mathbf{g}}^k \bullet \underset{\sim}{\mathbf{v}}$, respectively. Because cylindrical coordinates are orthogonal, it is conventional to also define so-called "physical components" with respect to the orthonormalized cylindrical basis.

For cylindrical coordinates, the physical components are denoted $\{v_r, v_\theta, v_z\}$. Whenever you derive a formula in terms of the general covariant and contravariant vector components, it's a good idea to convert the final result to physical coordinates and the physical basis. For cylindrical coordinates, these conversion formulas are:

$$v_1 = \underset{\sim}{\mathbf{g}}_1 \bullet \underset{\sim}{\mathbf{v}} = \underset{\sim}{\mathbf{e}}_r \bullet \underset{\sim}{\mathbf{v}} = v_r \qquad v^1 = \underset{\sim}{\mathbf{g}}^1 \bullet \underset{\sim}{\mathbf{v}} = \underset{\sim}{\mathbf{e}}_r \bullet \underset{\sim}{\mathbf{v}} = v_r \tag{5.29a}$$

$$v_2 = \underset{\sim}{\mathbf{g}}_2 \bullet \underset{\sim}{\mathbf{v}} = (r\underset{\sim}{\mathbf{e}}_\theta) \bullet \underset{\sim}{\mathbf{v}} = rv_\theta \qquad v^2 = \underset{\sim}{\mathbf{g}}^2 \bullet \underset{\sim}{\mathbf{v}} = \left(\frac{1}{r}\underset{\sim}{\mathbf{e}}_\theta\right) \bullet \underset{\sim}{\mathbf{v}} = \frac{v_\theta}{r} \tag{5.29b}$$

$$v_3 = \underset{\sim}{\mathbf{g}}_3 \bullet \underset{\sim}{\mathbf{v}} = \underset{\sim}{\mathbf{e}}_z \bullet \underset{\sim}{\mathbf{v}} = v_z \qquad v^3 = \underset{\sim}{\mathbf{g}}^3 \bullet \underset{\sim}{\mathbf{v}} = \underset{\sim}{\mathbf{e}}_z \bullet \underset{\sim}{\mathbf{v}} = v_z. \tag{5.29c}$$

Study Question 5.1 For spherical coordinates, $\{\eta^1 = r,\ \eta^2 = \theta,\ \eta^3 = \phi\}$, the underlying rectangular Cartesian coordinates are

$$x_1 = r\sin\theta\cos\phi \tag{5.30a}$$

$$x_2 = r\sin\theta\sin\phi \tag{5.30b}$$

$$x_3 = r\cos\theta. \tag{5.30c}$$

(a) Follow the above example to prove that the covariant basis for spherical coordinates is

$$\underset{\sim}{\mathbf{g}}_1 = \underset{\sim}{\mathbf{e}}_r, \quad \text{where} \quad \underset{\sim}{\mathbf{e}}_r = \sin\theta\cos\phi\,\underset{\sim}{\mathbf{e}}_1 + \sin\theta\sin\phi\,\underset{\sim}{\mathbf{e}}_2 + \cos\theta\,\underset{\sim}{\mathbf{e}}_3 \tag{5.31a}$$

$$\underset{\sim}{\mathbf{g}}_2 = r\underset{\sim}{\mathbf{e}}_\theta, \quad \text{where} \quad \underset{\sim}{\mathbf{e}}_\theta = \cos\theta\cos\phi\,\underset{\sim}{\mathbf{e}}_1 + \cos\theta\sin\phi\,\underset{\sim}{\mathbf{e}}_2 - \sin\theta\,\underset{\sim}{\mathbf{e}}_3 \tag{5.31b}$$

$$\underset{\sim}{\mathbf{g}}_3 = r\sin\theta\,\underset{\sim}{\mathbf{e}}_\phi \quad \text{where} \quad \underset{\sim}{\mathbf{e}}_\phi = -\sin\phi\,\underset{\sim}{\mathbf{e}}_1 + \cos\phi\,\underset{\sim}{\mathbf{e}}_2. \tag{5.31c}$$

(b) Prove that the dual contravariant basis for spherical coordinates is

$$\underset{\sim}{\mathbf{g}}^1 = \underset{\sim}{\mathbf{e}}_r, \tag{5.32a}$$

$$\underset{\sim}{\mathbf{g}}^2 = \frac{1}{r}\underset{\sim}{\mathbf{e}}_\theta, \tag{5.32b}$$

$$\underset{\sim}{\mathbf{g}}^3 = \frac{1}{r\sin\theta}\underset{\sim}{\mathbf{e}}_\phi. \tag{5.32c}$$

(c) As was done for cylindrical coordinates in Eq. (5.29) show that the spherical covariant and contravariant vector components are related to the spherical "physical" components by

$$v_1 = v_r \qquad v^1 = v_r \tag{5.33a}$$

$$v_2 = rv_\theta \qquad v^2 = \frac{v_\theta}{r} \tag{5.33b}$$

$$v_3 = (r\sin\theta)v_\phi \qquad v^3 = \frac{v_\phi}{r\sin\theta}. \tag{5.33c}$$

5.4 The gradient of a scalar

Section 5.1 provided a simplified example for finding the formula for the gradient of a scalar function of cylindrical coordinates. Now we outline the procedure for more complicated general curvilinear coordinates. Recall Eq. (5.18):

$$d\underset{\sim}{\mathbf{x}} = d\eta^k \, \underset{\sim}{\mathbf{g}}_k \, , \tag{5.34}$$

where

$$\underset{\sim}{\mathbf{g}}_k = \frac{\partial \underset{\sim}{\mathbf{x}}}{\partial \eta^k} . \tag{5.35}$$

Dotting both sides of Eq. (5.34) with $\underset{\sim}{\mathbf{g}}^i$ shows that

$$\boxed{d\eta^i = \underset{\sim}{\mathbf{g}}^i \bullet d\underset{\sim}{\mathbf{x}}} . \tag{5.36}$$

This important equation is crucial in determining expressions for gradient operations. Consider, for example, a scalar-valued field

$$s = s(\eta^1, \eta^2, \eta^3) . \tag{5.37}$$

The increment in this function is given by

$$\mathrm{d}s = \frac{\partial s}{\partial \eta^k} \mathrm{d}\eta^k \, , \tag{5.38}$$

or, using Eq. (5.36),

$$\mathrm{d}s = \frac{\partial s}{\partial \eta^k} \underset{\sim}{\mathbf{g}}^k \bullet \mathrm{d}\underset{\sim}{\mathbf{x}} \, , \tag{5.39}$$

which holds for all $\mathrm{d}\underset{\sim}{\mathbf{x}}$. The direct notation definition of the gradient $\mathrm{d}s / \mathrm{d}\underset{\sim}{\mathbf{x}}$ of a scalar field is

$$\mathrm{d}s = \frac{\mathrm{d}s}{\mathrm{d}\underset{\sim}{\mathbf{x}}} \bullet \mathrm{d}\underset{\sim}{\mathbf{x}} \qquad \forall \mathrm{d}\underset{\sim}{\mathbf{x}} . \tag{5.40}$$

Comparing the above two equations gives the formula for the gradient of a scalar in curvilinear coordinates:

$$\boxed{\frac{\mathrm{d}s}{\mathrm{d}\underset{\sim}{\mathbf{x}}} = \frac{\partial s}{\partial \eta^k} \underset{\sim}{\mathbf{g}}^k} . \tag{5.41}$$

Notice that this formula is very similar in form to the familiar formula for rectangular Cartesian coordinates. Gradient formulas won't look significantly different until we compute vector gradients in the next section.

Example: cylindrical coordinates Consider a scalar field

$$s = s(r, \theta, z) . \tag{5.42}$$

Applying Eq. (5.41) with the contravariant basis of Eq. (5.26) gives

$$\frac{\mathrm{d}s}{\mathrm{d}\underset{\sim}{\mathbf{x}}} = \frac{\partial s}{\partial r}\underset{\sim}{\mathbf{e}}_r + \frac{\partial s}{\partial \theta}\frac{\underset{\sim}{\mathbf{e}}_\theta}{r} + \frac{\partial s}{\partial z}\underset{\sim}{\mathbf{e}}_z \,, \tag{5.43}$$

which is the gradient formula typically found in math handbooks.

Study Question 5.2 Follow the above example [using Eq. (5.31) in Eq. (5.41)] to prove that the formula for the gradient of a scalar s in spherical coordinates is

$$\frac{\mathrm{d}s}{\mathrm{d}\underset{\sim}{\mathbf{x}}} = \frac{\partial s}{\partial r}\underset{\sim}{\mathbf{e}}_r + \frac{\partial s}{\partial \theta}\frac{\underset{\sim}{\mathbf{e}}_\theta}{r} + \frac{\partial s}{\partial \phi}\frac{\underset{\sim}{\mathbf{e}}_\phi}{r\sin\theta} \,. \tag{5.44}$$

5.5 Gradient of a vector -- "simplified" example

Now let's look at the gradient of a vector. If the vector is expressed in Cartesian coordinates, the formula for the gradient is

$$\frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}} = \frac{\partial v_i}{\partial x_j}\underset{\sim}{\mathbf{e}}_i\underset{\sim}{\mathbf{e}}_j \,. \tag{5.45}$$

Suppose a vector $\underset{\sim}{\mathbf{v}}$ is expressed in the cylindrical coordinates:

$$\underset{\sim}{\mathbf{v}} = v_r\underset{\sim}{\mathbf{e}}_r + v_\theta\underset{\sim}{\mathbf{e}}_\theta + v_z\underset{\sim}{\mathbf{e}}_z \,. \tag{5.46}$$

The components v_r, v_θ, and v_z are presumably known as functions of the $\{r, \theta, z\}$ coordinates. Importantly, *the base vectors also vary with the coordinates!*

First, recall the product rule for the gradient of a scalar times a vector:

$$\frac{d(s\underset{\sim}{\mathbf{u}})}{d\underset{\sim}{\mathbf{x}}} = \underset{\sim}{\mathbf{u}}\frac{ds}{d\underset{\sim}{\mathbf{x}}} + s\frac{d\underset{\sim}{\mathbf{u}}}{d\underset{\sim}{\mathbf{x}}} \,. \tag{5.47}$$

Applying the product rule to each of the terms in Eq. (5.46), the gradient of $d\underset{\sim}{\mathbf{v}}/d\underset{\sim}{\mathbf{x}}$ gives

$$\begin{aligned}\frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}} &= \underset{\sim}{\mathbf{e}}_r\frac{dv_r}{d\underset{\sim}{\mathbf{x}}} + \underset{\sim}{\mathbf{e}}_\theta\frac{dv_\theta}{d\underset{\sim}{\mathbf{x}}} + \underset{\sim}{\mathbf{e}}_z\frac{dv_z}{d\underset{\sim}{\mathbf{x}}} \\ &\quad + v_r\frac{d\underset{\sim}{\mathbf{e}}_r}{d\underset{\sim}{\mathbf{x}}} + v_\theta\frac{d\underset{\sim}{\mathbf{e}}_\theta}{d\underset{\sim}{\mathbf{x}}} + v_z\frac{d\underset{\sim}{\mathbf{e}}_z}{d\underset{\sim}{\mathbf{x}}} \,.\end{aligned} \tag{5.48}$$

Eq. (5.11) applies to the gradient of any scalar. Hence, the first three terms of Eq. (5.48) can be written

$$\frac{dv_r}{d\underset{\sim}{\mathbf{x}}} = v_{r,r}\underset{\sim}{\mathbf{e}}_r + \frac{v_{r,\theta}}{r}\underset{\sim}{\mathbf{e}}_\theta + v_{r,z}\underset{\sim}{\mathbf{e}}_3$$

$$\frac{dv_\theta}{d\underset{\sim}{\mathbf{x}}} = v_{\theta,r}\underset{\sim}{\mathbf{e}}_r + \frac{v_{\theta,\theta}}{r}\underset{\sim}{\mathbf{e}}_\theta + v_{\theta,z}\underset{\sim}{\mathbf{e}}_3$$

$$\frac{dv_z}{d\underset{\sim}{\mathbf{x}}} = v_{z,r}\underset{\sim}{\mathbf{e}}_r + \frac{v_{z,\theta}}{r}\underset{\sim}{\mathbf{e}}_\theta + v_{z,z}\underset{\sim}{\mathbf{e}}_3 . \tag{5.49}$$

Now we need formulas for the gradients of the base vectors. Applying the product rule to Eq. (5.2), noting that the gradients of the Cartesian basis are all zero, gives

$$\frac{d\underset{\sim}{\mathbf{e}}_r}{d\underset{\sim}{\mathbf{x}}} = (-\sin\theta\underset{\sim}{\mathbf{e}}_1 + \cos\theta\underset{\sim}{\mathbf{e}}_2)\frac{d\theta}{d\underset{\sim}{\mathbf{x}}} = \underset{\sim}{\mathbf{e}}_\theta\frac{d\theta}{d\underset{\sim}{\mathbf{x}}}$$

$$\frac{d\underset{\sim}{\mathbf{e}}_\theta}{d\underset{\sim}{\mathbf{x}}} = (-\cos\theta\underset{\sim}{\mathbf{e}}_1 - \sin\theta\underset{\sim}{\mathbf{e}}_2)\frac{d\theta}{d\underset{\sim}{\mathbf{x}}} = -\underset{\sim}{\mathbf{e}}_r\frac{d\theta}{d\underset{\sim}{\mathbf{x}}}$$

$$\frac{d\underset{\sim}{\mathbf{e}}_z}{d\underset{\sim}{\mathbf{x}}} = \underset{\approx}{\mathbf{0}} . \tag{5.50}$$

Applying (5.10),

$$\boxed{\begin{aligned} \frac{d\underset{\sim}{\mathbf{e}}_r}{d\underset{\sim}{\mathbf{x}}} &= \frac{1}{r}\underset{\sim}{\mathbf{e}}_\theta\underset{\sim}{\mathbf{e}}_\theta \\ \frac{d\underset{\sim}{\mathbf{e}}_\theta}{d\underset{\sim}{\mathbf{x}}} &= -\frac{1}{r}\underset{\sim}{\mathbf{e}}_r\underset{\sim}{\mathbf{e}}_\theta \\ \frac{d\underset{\sim}{\mathbf{e}}_z}{d\underset{\sim}{\mathbf{x}}} &= \underset{\approx}{\mathbf{0}} . \end{aligned}} \tag{5.51}$$

Substituting Eqs. (5.49) and (5.51) into (5.48) gives

$$\begin{aligned} \frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}} = {} & \underset{\sim}{\mathbf{e}}_r\left(v_{r,r}\underset{\sim}{\mathbf{e}}_r + \frac{v_{r,\theta}}{r}\underset{\sim}{\mathbf{e}}_\theta + v_{r,z}\underset{\sim}{\mathbf{e}}_3\right) \\ & + \underset{\sim}{\mathbf{e}}_\theta\left(v_{\theta,r}\underset{\sim}{\mathbf{e}}_r + \frac{v_{\theta,\theta}}{r}\underset{\sim}{\mathbf{e}}_\theta + v_{\theta,z}\underset{\sim}{\mathbf{e}}_3\right) \\ & \underset{\sim}{\mathbf{e}}_z\left(v_{z,r}\underset{\sim}{\mathbf{e}}_r + \frac{v_{z,\theta}}{r}\underset{\sim}{\mathbf{e}}_\theta + v_{z,z}\underset{\sim}{\mathbf{e}}_3\right) \\ & + v_r\left(\frac{1}{r}\underset{\sim}{\mathbf{e}}_\theta\underset{\sim}{\mathbf{e}}_\theta\right) + v_\theta\left(-\frac{1}{r}\underset{\sim}{\mathbf{e}}_r\underset{\sim}{\mathbf{e}}_\theta\right) . \end{aligned} \tag{5.52}$$

Collecting terms gives

$$\begin{aligned}\frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}} = &(v_{r,r})\underset{\sim}{\mathbf{e}}_r\underset{\sim}{\mathbf{e}}_r + \left(\frac{v_{r,\theta}}{r} - \frac{v_\theta}{r}\right)\underset{\sim}{\mathbf{e}}_r\underset{\sim}{\mathbf{e}}_\theta + (v_{r,z})\underset{\sim}{\mathbf{e}}_r\underset{\sim}{\mathbf{e}}_z \\ &+ (v_{\theta,r})\underset{\sim}{\mathbf{e}}_\theta\underset{\sim}{\mathbf{e}}_r + \left(\frac{v_{\theta,\theta}}{r} + \frac{v_r}{r}\right)\underset{\sim}{\mathbf{e}}_\theta\underset{\sim}{\mathbf{e}}_\theta + (v_{\theta,z})\underset{\sim}{\mathbf{e}}_\theta\underset{\sim}{\mathbf{e}}_z \\ &+ (v_{z,r})\underset{\sim}{\mathbf{e}}_z\underset{\sim}{\mathbf{e}}_r + \left(\frac{v_{z,\theta}}{r}\right)\underset{\sim}{\mathbf{e}}_z\underset{\sim}{\mathbf{e}}_\theta + v_{z,z}\underset{\sim}{\mathbf{e}}_z\underset{\sim}{\mathbf{e}}_z\,. \end{aligned} \tag{5.53}$$

This result is usually given in textbooks in matrix form with respect to the cylindrical basis:

$$\boxed{\left[\frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}}\right] = \begin{bmatrix} v_{r,r} & \dfrac{v_{r,\theta}-v_\theta}{r} & v_{r,z} \\ v_{\theta,r} & \dfrac{v_{\theta,\theta}+v_r}{r} & v_{\theta,z} \\ v_{z,r} & \dfrac{v_{z,\theta}}{r} & v_{z,z} \end{bmatrix}}\,. \tag{5.54}$$

There's no doubt that this result required a considerable amount of effort to derive. Typically, these kinds of formulas are compiled in the appendices of most tensor analysis reference books. The appendix of R.B. Bird's book on macromolecular hydrodynamics is particularly well-organized and error-free. If, however, you use Bird's appendix, you will notice that the components given for the gradient of a vector seem to be the transpose of what we have presented above; that's because Bird (and some others) define the gradient of a tensor to be the transpose of our definition. Before using anyone's gradient table, you should always ascertain which definition the author uses.

Now we are going to perform the same sort of analysis to show how the gradient is determined for general curvilinear coordinates.

5.6 Gradient of a vector in curvilinear coordinates

The formula for the gradient of a scalar in curvilinear coordinates was not particularly tough to derive and comprehend — it didn't look profoundly different from the formula for rectangular Cartesian coordinates. Taking gradients of vectors, however, begins a new nightmare. Consider a vector field,

$$\underset{\sim}{\mathbf{v}} = \underset{\sim}{\mathbf{v}}(\eta^1, \eta^2, \eta^3)\,. \tag{5.55}$$

Each component of the vector is of course a function of the coordinates, but for general curvilinear coordinates, so are the base vectors! Written out,

$$\underset{\sim}{\mathbf{v}} = v^i(\eta^1, \eta^2, \eta^3) \ \underset{\sim}{\mathbf{g}}_i(\eta^1, \eta^2, \eta^3). \tag{5.56}$$

Therefore the increment $d\underset{\sim}{\mathbf{v}}$ involves both increments dv^i of the components and increments $d\underset{\sim}{\mathbf{g}}_i$ of the base vectors:

$$d\underset{\sim}{\mathbf{v}} = dv^i \underset{\sim}{\mathbf{g}}_i + v^i d\underset{\sim}{\mathbf{g}}_i. \tag{5.57}$$

Applying the chain rule and using Eq. (5.36), the component increments can be written

$$dv^i = \frac{\partial v^i}{\partial \eta^j} d\eta^j = \frac{\partial v^i}{\partial \eta^j} \underset{\sim}{\mathbf{g}}^j \bullet d\underset{\sim}{\mathbf{x}}. \tag{5.58}$$

Similarly, the base vector increments are

$$d\underset{\sim}{\mathbf{g}}_i = \frac{\partial \underset{\sim}{\mathbf{g}}_i}{\partial \eta^j} d\eta^j = \frac{\partial \underset{\sim}{\mathbf{g}}_i}{\partial \eta^j} \underset{\sim}{\mathbf{g}}^j \bullet d\underset{\sim}{\mathbf{x}}. \tag{5.59}$$

Substituting these results into Eq. (5.57) and rearranging gives

$$d\underset{\sim}{\mathbf{v}} = \frac{\partial v^i}{\partial \eta^j} \underset{\sim}{\mathbf{g}}_i \underset{\sim}{\mathbf{g}}^j \bullet d\underset{\sim}{\mathbf{x}} + v^i \frac{\partial \underset{\sim}{\mathbf{g}}_i}{\partial \eta^j} \underset{\sim}{\mathbf{g}}^j \bullet d\underset{\sim}{\mathbf{x}}, \tag{5.60}$$

which holds for all $d\underset{\sim}{\mathbf{x}}$. Recall the gradient $d\underset{\sim}{\mathbf{v}} / d\underset{\sim}{\mathbf{x}}$ of a vector is defined in direct notation such that

$$d\underset{\sim}{\mathbf{v}} = \left[\frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}}\right] \bullet d\underset{\sim}{\mathbf{x}} \qquad \forall d\underset{\sim}{\mathbf{x}}. \tag{5.61}$$

Comparing the above two equations gives us a formula for the gradient:

$$\frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}} = \frac{\partial v^i}{\partial \eta^j} \underset{\sim}{\mathbf{g}}_i \underset{\sim}{\mathbf{g}}^j + v^i \frac{\partial \underset{\sim}{\mathbf{g}}_i}{\partial \eta^j} \underset{\sim}{\mathbf{g}}^j. \tag{5.62}$$

Incidentally, these equations serve as further examples of how a superscript in the "denominator" is regarded as a subscript in the summation convention.

Christoffel Symbols Note that the nine $\partial \underset{\sim}{\mathbf{g}}_i / \partial \eta^j$ vectors [*i.e.*, the coefficients of v_i in the last term of Eq. (5.62)] are strictly properties of the coordinate system and they may be computed and tabulated *a priori*. This family of system-dependent vectors is denoted

$$\boxed{\underset{\sim}{\Gamma}_{ij} \equiv \frac{\partial \underset{\sim}{\mathbf{g}}_i}{\partial \eta^j}}. \tag{5.63}$$

Recalling Eq. (5.35), note that

$$\underset{\sim}{\Gamma}_{ij} = \frac{\partial^2 \underset{\sim}{\mathbf{x}}}{\partial \eta^j \partial \eta^i} = \frac{\partial^2 \underset{\sim}{\mathbf{x}}}{\partial \eta^i \partial \eta^j} = \underset{\sim}{\Gamma}_{ji}. \tag{5.64}$$

Thus, only six of the nine $\underset{\sim}{\Gamma}_{ij}$ vectors are independent due to the symmetry in ij. The k^{th} contravariant component of $\underset{\sim}{\Gamma}_{ij}$ is obtained by dotting $\underset{\sim}{\mathbf{g}}^k$ into $\underset{\sim}{\Gamma}_{ij}$:

$$\boxed{\Gamma_{ij}^k \equiv \underset{\sim}{\mathbf{g}}^k \bullet \underset{\sim}{\Gamma}_{ij}}. \tag{5.65}$$

The quantity Γ_{ij}^k is called the **Christoffel symbol of the second kind**.

Important comment about notation: Even though the Christoffel symbols Γ_{ij}^k have three indices, they are *not* components of a third-order tensor. By this, we mean that a second basis $\underset{\sim}{\bar{\mathbf{g}}}^k$ will have a different set of Christoffel symbols $\bar{\Gamma}_{ij}^k$ and, as discussed below, they are that are not obtainable via a simple tensor transformation of the Christoffel symbols of the original system [i.e., $\bar{\Gamma}_{ij}^k \neq \Gamma_{mn}^p(\underset{\sim}{\mathbf{g}}^m \bullet \underset{\sim}{\bar{\mathbf{g}}}_m)(\underset{\sim}{\mathbf{g}}^n \bullet \underset{\sim}{\bar{\mathbf{g}}}_n)(\underset{\sim}{\mathbf{g}}_p \bullet \underset{\sim}{\bar{\mathbf{g}}}^k)$]. Instead of the notation Γ_{ij}^k, Christoffel symbols are therefore frequently denoted in the literature with the odd-looking notation $\{{k \atop ij}\}$. The fact that Christoffel symbols are not components of a tensor is, of course, a strong justification for avoiding typesetting Christoffel symbols in the same way as tensors. However, to be 100% consistent, proponents of this notational ideology would — by the same arguments — be compelled to *not* typeset *coordinates* using the notation η^k which erroneously makes it look as though the η^k are contravariant components of some vector $\underset{\sim}{\eta}$, even though they aren't. Being comfortable with the typesetting η^k, we are also comfortable with the typesetting Γ_{ij}^k. The key is for the analyst to recognize that neither of these symbols connote tensors. Instead, they are "basis-intrinsic" quantities (i.e., indexed quantities whose meaning is defined for a particular basis and whose connection with counterparts from a different basis are *not* obtained via a tensor transformation rule). Of course, the base vectors themselves are basis-intrinsic objects. Any new object that is constructed from basis-intrinsic quantities should be *itself* regarded as basis-intrinsic until proven otherwise. For example, the metric $g_{ij} \equiv \underset{\sim}{\mathbf{g}}_i \bullet \underset{\sim}{\mathbf{g}}_j$ were initially regarded as basis-intrinsic because they were constructed from basis-intrinsic objects (the base vectors), but it was proved that they turned out to also satisfy the tensor transformation rule. Consequently, even though the metric matrix is constructed from basis-intrinsic quantities, it turns out to *not* be basis intrinsic itself (the metric components are components of the identity tensor).

On page 86, we define Christoffel symbols of the "first" kind, which are useful in Riemannian spaces where there is no underlying rectangular Cartesian basis. In general, if the term "Christoffel symbol" is used by itself, it should be taken to mean the Christoffel symbol of the *second* kind defined above. Christoffel symbols may appear rather arcane, but keep in mind that these quantities simply characterize how the base vectors vary in space. Christoffel are also sometimes called the "**affinities**" [12].

By virtue of Eq. (5.64), note that

$$\boxed{\Gamma_{ij}^{k} = \Gamma_{ji}^{k}} . \tag{5.66}$$

Like the components F_{ij} of the tensor $\underset{\sim}{\mathbf{F}}$, the Christoffel symbols Γ_{ij}^{k} are properties of the *particular* coordinate system and its basis. Consequently, the Christoffel symbols are not the components of a third-order tensor. Specifically, if we were to consider some *different* curvilinear coordinate system and compute the Christoffel symbols for that system, the result would not be the same as would be obtained by a tensor transformation of the Christoffel symbols of the first coordinate system to the second system. This is in contrast to the metric coefficients g_{ij} which *do* happen to be components of a tensor (namely, the identity tensor).

Recalling that Γ_{ij}^{k} is the k^{th} contravariant component of $\underset{\sim}{\Gamma}_{ij}$ and by Eq. (5.63) $\underset{\sim}{\Gamma}_{ij} = \partial\underset{\sim}{\mathbf{g}}_i / \partial\eta^j$, we conclude that the variation of the base vectors in space is completely characterized by the Christoffel symbols. Namely,

$$\boxed{\frac{\partial\underset{\sim}{\mathbf{g}}_i}{\partial\eta^j} = \Gamma_{ij}^{k}\underset{\sim}{\mathbf{g}}_k} . \tag{5.67}$$

Increments in the base vectors By the chain rule, the increment in the covariant base vector can always be written

$$d\underset{\sim}{\mathbf{g}}_i = \frac{\partial\underset{\sim}{\mathbf{g}}_i}{\partial\eta^j} d\eta^j \tag{5.68}$$

or, using the notation introduced in Eq. (5.67)

$$\boxed{d\underset{\sim}{\mathbf{g}}_i = \Gamma_{ij}^{k}\underset{\sim}{\mathbf{g}}_k d\eta^j} \tag{5.69}$$

Manifold torsion Recall that the Christoffel symbols are not components of basis-independent tensors. Consider, however [12], the anti-symmetric part of the Christoffel symbols:

$$2\Gamma_{[ij]}^{k} \equiv \Gamma_{ij}^{k} - \Gamma_{ji}^{k} \tag{5.70}$$

As long as Eq. (5.66) holds, then the manifold torsion will be zero. For a non-holomomic system, it's possible that the manifold torsion will be nonzero, but it will turn out to be a basis independent (i.e., "free" vector). Henceforth, we assume that Eq. (5.66) holds true, so no further mention will be made of the manifold torsion.

EXAMPLE: Christoffel symbols for cylindrical coordinates In terms of the underlying rectangular Cartesian basis, the covariant base vectors from Eq. (5.19) can be written

$$\underset{\sim}{\mathbf{g}}_1 = \cos\theta\underset{\sim}{\mathbf{e}}_1 + \sin\theta\underset{\sim}{\mathbf{e}}_2 \tag{5.71a}$$

$$\underset{\sim}{\mathbf{g}}_2 = -r\sin\theta\underset{\sim}{\mathbf{e}}_1 + r\cos\theta\underset{\sim}{\mathbf{e}}_2 \tag{5.71b}$$

$$\underset{\sim}{\mathbf{g}}_3 = \underset{\sim}{\mathbf{e}}_3 . \tag{5.71c}$$

Therefore, applying Eq. (5.63), using $\{\eta^1=r,\ \eta^2=\theta,\ \eta^3=z\}$

$$\underset{\sim}{\Gamma}_{11} = \frac{\partial\underset{\sim}{\mathbf{g}}_1}{\partial r} = \underset{\sim}{0} \quad \underset{\sim}{\Gamma}_{12} = \frac{\partial\underset{\sim}{\mathbf{g}}_1}{\partial\theta} = -\sin\theta\underset{\sim}{\mathbf{e}}_1 + \cos\theta\underset{\sim}{\mathbf{e}}_2 = \frac{1}{r}\underset{\sim}{\mathbf{g}}_2 \quad \underset{\sim}{\Gamma}_{13} = \frac{\partial\underset{\sim}{\mathbf{g}}_1}{\partial z} = \underset{\sim}{0} \tag{5.72a}$$

$$\underset{\sim}{\Gamma}_{21} = \underset{\sim}{\Gamma}_{12} \quad \underset{\sim}{\Gamma}_{22} = \frac{\partial\underset{\sim}{\mathbf{g}}_2}{\partial\theta} = -r\cos\theta\underset{\sim}{\mathbf{e}}_1 - r\sin\theta\underset{\sim}{\mathbf{e}}_2 = -r\underset{\sim}{\mathbf{g}}_1 \quad \underset{\sim}{\Gamma}_{23} = \frac{\partial\underset{\sim}{\mathbf{g}}_2}{\partial z} = \underset{\sim}{0} \tag{5.72b}$$

$$\underset{\sim}{\Gamma}_{31} = \underset{\sim}{\Gamma}_{13} \quad \underset{\sim}{\Gamma}_{32} = \underset{\sim}{\Gamma}_{23} \quad \underset{\sim}{\Gamma}_{33} = \frac{\partial\underset{\sim}{\mathbf{g}}_3}{\partial z} = \underset{\sim}{0} . \tag{5.72c}$$

Noting that Γ_{ij}^k is the coefficient of $\underset{\sim}{\mathbf{g}}_k$ in the expression for $\underset{\sim}{\Gamma}_{ij}$, we find that only three of the 27 Christoffel symbols are nonzero. Namely, noting that Γ_{12}^2 is the coefficient of $\underset{\sim}{\mathbf{g}}_2$ in $\underset{\sim}{\Gamma}_{12}$ and noting that Γ_{22}^1 the coefficient of $\underset{\sim}{\mathbf{g}}_1$ in $\underset{\sim}{\Gamma}_{22}$,

$$\Gamma_{12}^2 = \Gamma_{21}^2 = \frac{1}{r} \quad \text{and} \quad \Gamma_{22}^1 = -r \quad (\text{all other } \Gamma_{ij}^k = 0). \tag{5.73}$$

If you look up cylindrical Christoffel symbols in a handbook, you will probably find the subscripts (1,2,3) replaced with the coordinate symbols (r,θ,z) for clarity so they are listed as

$$\Gamma_{r\theta}^{\theta} = \Gamma_{\theta r}^{\theta} = \frac{1}{r} \quad \text{and} \quad \Gamma_{\theta\theta}^{r} = -r \quad (\text{all other } \Gamma_{ij}^k = 0). \tag{5.74}$$

Study Question 5.3 Using the spherical covariant base vectors in Eq. (5.31), prove that

$$\underset{\sim}{\Gamma}_{11} = \underset{\sim}{0} \qquad \underset{\sim}{\Gamma}_{12} = \frac{1}{r}\underset{\sim}{\mathbf{g}}_2 \qquad \underset{\sim}{\Gamma}_{13} = \frac{1}{r}\underset{\sim}{\mathbf{g}}_3 \tag{5.75a}$$

$$\underset{\sim}{\Gamma}_{21} = \underset{\sim}{\Gamma}_{12} \qquad \underset{\sim}{\Gamma}_{22} = -r\underset{\sim}{\mathbf{g}}_1 \qquad \underset{\sim}{\Gamma}_{23} = \cot\phi\,\underset{\sim}{\mathbf{g}}_3 \tag{5.75b}$$

$$\underset{\sim}{\Gamma}_{31} = \underset{\sim}{\Gamma}_{13} \qquad \underset{\sim}{\Gamma}_{32} = \underset{\sim}{\Gamma}_{23} \qquad \underset{\sim}{\Gamma}_{33} = -r\sin^2\theta\,\underset{\sim}{\mathbf{g}}_1 - \sin\theta\cos\theta\,\underset{\sim}{\mathbf{g}}_2\,. \tag{5.75c}$$

Therefore show that the Christoffel symbols for spherical $\{r, \theta, \phi\}$ coordinates are

$$\Gamma^{\theta}_{r\theta} = \Gamma^{\theta}_{\theta r} = \frac{1}{r}, \qquad \Gamma^{\phi}_{r\phi} = \Gamma^{\phi}_{\phi r} = \frac{1}{r}, \tag{5.76a}$$

$$\Gamma^{r}_{\theta\theta} = -r, \qquad \Gamma^{\phi}_{\theta\phi} = \Gamma^{\phi}_{\phi\theta} = \cot\phi \tag{5.76b}$$

$$\Gamma^{r}_{\phi\phi} = -r\sin^2\theta, \qquad \Gamma^{\theta}_{\phi\phi} = -\sin\theta\cos\theta \qquad (\text{all other } \Gamma^k_{ij} = 0). \tag{5.76c}$$

Partial Answer: *To find the vector $\underset{\sim}{\Gamma}_{12}$, first differentiate $\underset{\sim}{\mathbf{g}}_1$ from Eq. (5.31a) with respect to η_2 (i.e., with respect to θ). You then recognize from (5.31b) that the result is $\underset{\sim}{\Gamma}_{12} = \underset{\sim}{\mathbf{e}}_\theta$. Then $t\Gamma^2_{12}$ is found by dotting $\underset{\sim}{\Gamma}_{12}$ by $\underset{\sim}{\mathbf{g}}^2$ from Eq. (5.32b). Hence $\Gamma^2_{12} = \Gamma^{\theta}_{r\theta} = \underset{\sim}{\Gamma}_{12} \bullet \underset{\sim}{\mathbf{g}}^2 = \underset{\sim}{\mathbf{e}}_\theta \bullet (\underset{\sim}{\mathbf{e}}_\theta / r) = 1/r$.*

Covariant differentiation of contravariant components Substituting Eq. (5.67) into Eq. (5.62) gives

$$\frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}} = \frac{\partial v^i}{\partial \eta^j}\underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}^j + v^i\Gamma^k_{ij}\underset{\sim}{\mathbf{g}}_k\underset{\sim}{\mathbf{g}}^j\,, \tag{5.77}$$

or, changing the dummy summation indices so that the basis dyads will have the same subscripts in both terms,

$$\frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}} = \frac{\partial v^i}{\partial \eta^j}\underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}^j + v^k\Gamma^i_{kj}\underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}^j\,. \tag{5.78}$$

We now introduce a compact notation called **covariant vector differentiation**:

$$\boxed{v^i{}_{/j} \equiv \frac{\partial v^i}{\partial \eta^j} + v^k\Gamma^i_{kj}}\,. \tag{5.79}$$

The notation for covariant differentiation varies widely: the slash in $v^i{}_{/j}$ is also often denoted with a comma $v^i{}_{,j}$ although many writers use a comma to denote ordinary partial differentiation. Keep in mind: the Christoffel terms in Eq. (5.79) account for the variation of the base vectors in space. Using covariant differentiation, the gradient of a vector is then written compactly as

$$\boxed{\frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}} = v^i_{/j}\underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}^j}\,. \tag{5.80}$$

Example: Gradient of a vector in cylindrical coordinates Recalling Eq. (5.72), we can apply Eq. (5.79) to obtain

$$v^1{}_{/1} = \frac{\partial v^1}{\partial \eta^1} + v^1\Gamma^1_{11} + v^2\Gamma^1_{21} + v^3\Gamma^1_{31} = \frac{\partial v^1}{\partial r} \tag{5.81rr}$$

$$v^1{}_{/2} = \frac{\partial v^1}{\partial \eta^2} + v^1\Gamma^1_{12} + v^2\Gamma^1_{22} + v^3\Gamma^1_{32} = \frac{\partial v^1}{\partial \theta} - v^2 r \tag{5.81rt}$$

$$v^1{}_{/3} = \frac{\partial v^1}{\partial \eta^3} + v^1\Gamma^1_{13} + v^2\Gamma^1_{23} + v^3\Gamma^1_{33} = \frac{\partial v^1}{\partial z} \tag{5.81rz}$$

$$v^2{}_{/1} = \frac{\partial v^2}{\partial \eta^1} + v^1\Gamma^2_{11} + v^2\Gamma^2_{21} + v^3\Gamma^2_{31} = \frac{\partial v^2}{\partial r} \tag{5.81tr}$$

$$v^2{}_{/2} = \frac{\partial v^2}{\partial \eta^2} + v^1\Gamma^2_{12} + v^2\Gamma^2_{22} + v^3\Gamma^2_{32} = \frac{\partial v^2}{\partial \theta} + \frac{v^1}{r} \tag{5.81tt}$$

$$v^2{}_{/3} = \frac{\partial v^2}{\partial \eta^3} + v^1\Gamma^2_{13} + v^2\Gamma^2_{23} + v^3\Gamma^2_{33} = \frac{\partial v^2}{\partial z} \tag{5.81tz}$$

$$v^3{}_{/1} = \frac{\partial v^3}{\partial \eta^1} + v^1\Gamma^3_{11} + v^2\Gamma^3_{21} + v^3\Gamma^3_{31} = \frac{\partial v^3}{\partial r} \tag{5.81zr}$$

$$v^3{}_{/2} = \frac{\partial v^3}{\partial \eta^2} + v^1\Gamma^3_{12} + v^2\Gamma^3_{22} + v^3\Gamma^3_{32} = \frac{\partial v^3}{\partial \theta} \tag{5.81zt}$$

$$v^3{}_{/3} = \frac{\partial v^3}{\partial \eta^3} + v^1\Gamma^3_{13} + v^2\Gamma^3_{23} + v^3\Gamma^3_{33} = \frac{\partial v^3}{\partial z} \tag{5.81zz}$$

Hence, the gradient of a vector is given by

$$\begin{aligned}\frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}} = & \left(\frac{\partial v^1}{\partial r}\right)\underset{\sim}{\mathbf{g}}_1\underset{\sim}{\mathbf{g}}^1 + \left(\frac{\partial v^1}{\partial \theta} - v^2 r\right)\underset{\sim}{\mathbf{g}}_1\underset{\sim}{\mathbf{g}}^2 + \left(\frac{\partial v^1}{\partial z}\right)\underset{\sim}{\mathbf{g}}_1\underset{\sim}{\mathbf{g}}^3 \\ & + \left(\frac{\partial v^2}{\partial r}\right)\underset{\sim}{\mathbf{g}}_2\underset{\sim}{\mathbf{g}}^1 + \left(\frac{\partial v^2}{\partial \theta} + \frac{v^1}{r}\right)\underset{\sim}{\mathbf{g}}_2\underset{\sim}{\mathbf{g}}^2 + \left(\frac{\partial v^2}{\partial z}\right)\underset{\sim}{\mathbf{g}}_2\underset{\sim}{\mathbf{g}}^3 \\ & + \left(\frac{\partial v^3}{\partial r}\right)\underset{\sim}{\mathbf{g}}_3\underset{\sim}{\mathbf{g}}^1 + \left(\frac{\partial v^3}{\partial \theta}\right)\underset{\sim}{\mathbf{g}}_3\underset{\sim}{\mathbf{g}}^2 + \left(\frac{\partial v^3}{\partial z}\right)\underset{\sim}{\mathbf{g}}_3\underset{\sim}{\mathbf{g}}^3 .\end{aligned} \tag{5.82}$$

Substituting Eqs. (5.19), (5.26), and (5.29) into the above formula gives

$$\begin{aligned}\frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}} = & \left(\frac{\partial v_r}{\partial r}\right)\underset{\sim}{\mathbf{e}}_r\underset{\sim}{\mathbf{e}}_r + \left(\frac{\partial v_r}{\partial \theta} - v_\theta\right)\underset{\sim}{\mathbf{e}}_r\left(\frac{\underset{\sim}{\mathbf{e}}_\theta}{r}\right) + \left(\frac{\partial v_r}{\partial z}\right)\underset{\sim}{\mathbf{e}}_r\underset{\sim}{\mathbf{e}}_z \\ & + \left(\frac{1}{r}\frac{\partial v_\theta}{\partial r}\right)(r\underset{\sim}{\mathbf{e}}_\theta)\underset{\sim}{\mathbf{e}}_r + \left(\frac{1}{r}\frac{\partial v_\theta}{\partial \theta} + \frac{v_r}{r}\right)(r\underset{\sim}{\mathbf{e}}_\theta)\left(\frac{\underset{\sim}{\mathbf{e}}_\theta}{r}\right) + \left(\frac{1}{r}\frac{\partial v_\theta}{\partial z}\right)(r\underset{\sim}{\mathbf{e}}_\theta)\underset{\sim}{\mathbf{e}}_z \\ & + \left(\frac{\partial v_z}{\partial r}\right)\underset{\sim}{\mathbf{e}}_z\underset{\sim}{\mathbf{e}}_r + \left(\frac{\partial v_z}{\partial \theta}\right)\underset{\sim}{\mathbf{e}}_z\left(\frac{\underset{\sim}{\mathbf{e}}_\theta}{r}\right) + \left(\frac{\partial v_z}{\partial z}\right)\underset{\sim}{\mathbf{e}}_z\underset{\sim}{\mathbf{e}}_z .\end{aligned} \tag{5.83}$$

Upon simplification, the matrix of $d\underset{\sim}{\mathbf{v}}/d\underset{\sim}{\mathbf{x}}$ with respect to the usual ortho*normalized* basis $\{\underset{\sim}{\mathbf{e}}_r, \underset{\sim}{\mathbf{e}}_\theta, \underset{\sim}{\mathbf{e}}_z\}$ is

$$\left[\frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}}\right] = \begin{bmatrix} \dfrac{\partial v_r}{\partial r} & \dfrac{1}{r}\left(\dfrac{\partial v_r}{\partial\theta} - v_\theta\right) & \dfrac{\partial v_r}{\partial z} \\ \dfrac{\partial v_\theta}{\partial r} & \left(\dfrac{1}{r}\dfrac{\partial v_\theta}{\partial\theta} + \dfrac{v_r}{r}\right) & \dfrac{\partial v_\theta}{\partial z} \\ \dfrac{\partial v_z}{\partial r} & \dfrac{1}{r}\dfrac{\partial v_z}{\partial\theta} & \dfrac{\partial v_z}{\partial z} \end{bmatrix} \text{ w.r.t } \{\underset{\sim}{\mathbf{e}}_r, \underset{\sim}{\mathbf{e}}_\theta, \underset{\sim}{\mathbf{e}}_z\} \text{ unit basis.} \tag{5.84}$$

This is the formula usually cited in handbooks.

Study Question 5.4 Again we will duplicate the methods of the preceding example for spherical coordinates. However, rather than duplicating the entire hideous analyses, we here compute only the $r\phi$ component of $d\underset{\sim}{\mathbf{v}}/d\underset{\sim}{\mathbf{x}}$. In particular, we seek

$$\underset{\sim}{\mathbf{e}}_r \bullet \frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}} \bullet \underset{\sim}{\mathbf{e}}_\phi. \tag{5.85}$$

(a) Noting from Eqs. (5.31) and (5.32) how $\underset{\sim}{\mathbf{e}}_r$ and $\underset{\sim}{\mathbf{e}}_\phi$ are related to $\{\underset{\sim}{\mathbf{g}}_1, \underset{\sim}{\mathbf{g}}_2, \underset{\sim}{\mathbf{g}}_3\}$, show that

$$\underset{\sim}{\mathbf{e}}_r \bullet \frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}} \bullet \underset{\sim}{\mathbf{e}}_\phi = \frac{1}{r\sin\theta}\underset{\sim}{\mathbf{g}}^1 \bullet \frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}} \bullet \underset{\sim}{\mathbf{g}}_3. \tag{5.86}$$

(b) Explain why Eq. (5.80) therefore implies that

$$\underset{\sim}{\mathbf{e}}_r \bullet \frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}} \bullet \underset{\sim}{\mathbf{e}}_\phi = \frac{v^1_{;3}}{r\sin\theta}. \tag{5.87}$$

(c) With respect to the orthonormal basis $\{\underset{\sim}{\mathbf{e}}_r, \underset{\sim}{\mathbf{e}}_\theta, \underset{\sim}{\mathbf{e}}_\phi\}$ recall from Eq. (5.33) that $v^1 = v_r$ and $v^3 = v_\phi / r\sin\theta$. Use the Christoffel symbols of Eq. (5.76) in the formula (5.79) to show that

$$v^1{}_{/3} = \frac{\partial v_r}{\partial\phi} + v_\phi(-\sin\theta). \tag{5.88}$$

(d) The final step is to substitute this result into Eq. (5.87) to deduce that

$$\underset{\sim}{\mathbf{e}}_r \bullet \frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}} \bullet \underset{\sim}{\mathbf{e}}_\phi = \frac{1}{r\sin\theta}\frac{\partial v_r}{\partial\phi} - \frac{v_\phi}{r}. \tag{5.89}$$

Cite a textbook (or other reference) that tabulates formulas for gradients in spherical coordinates. Does your Eq. (5.89) agree with the formula for the $r\phi$ component of $d\underset{\sim}{\mathbf{v}}/d\underset{\sim}{\mathbf{x}}$ provided in the textbook?

Covariant differentiation of covariant components Recalling, Eq. (5.63), we now consider a similarly-defined *basis-dependent* vector:

$$\underset{\sim}{\mathrm{P}}{}_i^k \equiv \frac{\partial \underset{\sim}{\mathbf{g}}^k}{\partial \eta^i}, \tag{5.90}$$

and analogously to Eq. (5.65) we define

$$\boxed{\mathrm{P}_{ij}^k \equiv \underset{\sim}{\mathrm{P}}{}_i^k \bullet \underset{\sim}{\mathbf{g}}_j}. \tag{5.91}$$

Given a vector $\underset{\sim}{\mathbf{v}}$ for which the covariant v_k components are known, an analysis similar to that of the previous sections eventually reveals that the gradient of the vector can be written:

$$\frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}} = v_{i/j}\underset{\sim}{\mathbf{g}}^i\underset{\sim}{\mathbf{g}}^j, \tag{5.92}$$

where

$$v_{i/j} \equiv \frac{\partial v_i}{\partial \eta^j} + v_k \mathrm{P}_{ij}^k. \tag{5.93}$$

The question naturally arises: what connection, if any, does P_{ij}^k have with Γ_{ij}^k? To answer this question, differentiate both sides of Eq. (2.12) with respect to η^k:

$$\frac{\partial \underset{\sim}{\mathbf{g}}^i}{\partial \eta^k} \bullet \underset{\sim}{\mathbf{g}}_j + \underset{\sim}{\mathbf{g}}^i \bullet \frac{\partial \underset{\sim}{\mathbf{g}}_j}{\partial \eta^k} = 0, \tag{5.94}$$

or

$$\underset{\sim}{\mathrm{P}}{}_k^i \bullet \underset{\sim}{\mathbf{g}}_j + \underset{\sim}{\mathbf{g}}^i \bullet \underset{\sim}{\Gamma}_{jk} = 0 \text{ and therefore } \boxed{\mathrm{P}_{kj}^i + \Gamma_{kj}^i = 0}. \tag{5.95}$$

In other words, P_{ij}^k is just the negative of Γ_{ij}^k. Consequently, equation (5.92) becomes

$$\frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}} = v_{i/j}\underset{\sim}{\mathbf{g}}^i\underset{\sim}{\mathbf{g}}^j \text{ where } \qquad v_{i/j} \equiv \frac{\partial v_i}{\partial \eta^j} - v_k \Gamma_{ij}^k. \tag{5.96}$$

We now have an equation for increments of the *contravariant* base vectors

$$\boxed{d\underset{\sim}{\mathbf{g}}^k = (-\Gamma_{ij}^k \underset{\sim}{\mathbf{g}}^j) d\eta^i} \tag{5.97}$$

which should be compared with Eq. (5.69)

Product rules for covariant differentiation Most of the usual rules of differential calculus apply. For example,

$$(v_i w_j)_{/k} = v_i w_{j/k} + v_{i/k} w_j \qquad \text{and} \qquad (v^i w_j)_{/k} = v^i w_{j/k} + v^i{}_{/k} w_j, \text{ etc.} \tag{5.98}$$

5.7 Backward and forward gradient operators

The gradient $d\underset{\sim}{v}/d\underset{\sim}{x}$ that we defined in the preceding section is really a *backward* operating gradient. By this we mean that the ij component follows an index ordering that is analogous to the index ordering on the dyad $\underset{\sim}{\mathbf{v}}\underset{\sim}{\mathbf{d}}$. This dyad has components $v_i d_j$, which is comparable to the index ordering in the *Cartesian ij* component formula $\partial v_i/\partial x_j$.

In Cartesian coordinates, the gradient $d\underset{\sim}{v}/d\underset{\sim}{x}$ is a second order tensor whose *Cartesian ij* components are given by $\partial v_i/\partial x_j$. But couldn't we just have well have defined a second order tensor whose *Cartesian ij* components would be $\partial v_j/\partial x_i$. The only difference between these two choices is that the index placement is swapped. Thus, the transpose of one choice gives the other choice.

Following definitions and notation used by Malvern [11], we will denote the backward operating gradient by $\underset{\sim}{\mathbf{v}}\overleftarrow{\nabla}$ and the forward operating gradient by $\overrightarrow{\nabla}\underset{\sim}{\mathbf{v}}$. The "$\overrightarrow{\nabla}$" operates on arguments to its right, while the $\overleftarrow{\nabla}$ operates on arguments to its left. As mentioned earlier, our previously defined vector gradient is actually a backward del definition:

$$\underset{\sim}{\mathbf{v}}\overleftarrow{\nabla} \text{ means the same thing as } \frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}}. \qquad \text{Thus, } \underset{\sim}{\mathbf{v}}\overleftarrow{\nabla} = v^i_{/j}\underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}^j \tag{5.99}$$

$$\overrightarrow{\nabla}\underset{\sim}{\mathbf{v}} \text{ means the same thing as } \left(\frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}}\right)^T \qquad \text{Thus, } \overrightarrow{\nabla}\underset{\sim}{\mathbf{v}} = v^j_{/i}\underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}^j \tag{5.100}$$

The component ordering for the forward gradient operator is defined in a manner that is analogous to the dyad $\underset{\sim}{\mathbf{d}}\underset{\sim}{\mathbf{v}} = (d^i v_j)\underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}^j$, whereas the component ordering for the backward gradient is analogous to that on the dyad $\underset{\sim}{\mathbf{v}}\underset{\sim}{\mathbf{d}} = (v^j d_i)\underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}^j$. This leads naturally to the question of whether or not it is possible to define right and left gradient operators in a manner that permits some "heuristic assistance."

Recall that

$$\frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}} = \frac{\partial v^i}{\partial \eta^j}\underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}^j + v^k\Gamma^i_{kj}\underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}^j \tag{5.101}$$

Suppose that we wish to identify a left-acting operator $\overleftarrow{\nabla}$ such that

$$\frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}} = (v^i\underset{\sim}{\mathbf{g}}_i)\overleftarrow{\nabla} \tag{5.102}$$

Let's suppose we desire to define this operator such that it follows a product rule so that

$$(v^i\underset{\sim}{\mathbf{g}}_i)\overleftarrow{\nabla} = \underset{\sim}{\mathbf{g}}_i[(v^i)\overleftarrow{\nabla}] + v^i[(\underset{\sim}{\mathbf{g}}_i)\overleftarrow{\nabla}] \tag{5.103}$$

This suggests that we should define

$$\boxed{(v^i)\overleftarrow{\nabla} = \frac{\partial v^i}{\partial \eta^j}\underset{\sim}{\mathbf{g}}^j} \quad \text{and} \quad \boxed{(\underset{\sim}{\mathbf{g}}_k)\overleftarrow{\nabla} = \Gamma^i_{kj}\underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}^j} \tag{5.104}$$

Similarly, for the forward operation gradient, we define

$$\vec{\nabla}(v^i \underset{\sim}{\mathbf{g}}_i) = \left[\vec{\nabla}(v^i)\right]\underset{\sim}{\mathbf{g}}_i + v^i\left[\vec{\nabla}\underset{\sim}{\mathbf{g}}_i\right] \tag{5.105}$$

from which it follows

$$\vec{\nabla}(v^i) = \frac{\partial v^i}{\partial \eta^j}\underset{\sim}{\mathbf{g}}^j \quad \text{and} \quad \vec{\nabla}(\underset{\sim}{\mathbf{g}}_k) = \Gamma^i_{kj}\underset{\sim}{\mathbf{g}}^j\underset{\sim}{\mathbf{g}}_i \tag{5.106}$$

These definitions are cute, but we caution against using them too cavalierly. A careful application of the original definition of the gradient is probably safer.

5.8 Divergence of a vector

The divergence of a vector $\underset{\sim}{\mathbf{v}}$ is denoted $\nabla \bullet \underset{\sim}{\mathbf{v}}$ and it is defined by

$$\nabla \bullet \underset{\sim}{\mathbf{v}} = \mathrm{tr}\left(\frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}}\right) \tag{5.107}$$

The simplicity of the component expression for the trace depends on the expression chosen for the gradient operation. For example, taking the trace of Eq. (5.96) gives a valid but somewhat ugly expression:

$$\nabla \bullet \underset{\sim}{\mathbf{v}} = v_{i/j}(\underset{\sim}{\mathbf{g}}^i \bullet \underset{\sim}{\mathbf{g}}^j) = v_{i/j}(g^{ij}) = \left[\frac{\partial v_i}{\partial \eta^j} - v_k\Gamma^k_{ij}\right]g^{ij} \tag{5.108}$$

A much simpler expression for the divergence can be obtained by taking the trace of Eq. (5.80) to give

$$\nabla \bullet \underset{\sim}{\mathbf{v}} = v^i_{/j}\underset{\sim}{\mathbf{g}}_i \bullet \underset{\sim}{\mathbf{g}}^j = v^i_{/j}\delta^j_i = v^i_{/i} \tag{5.109}$$

Written out explicitly, this result is

$$\boxed{\nabla \bullet \underset{\sim}{v} = \frac{\partial v^i}{\partial \eta^i} + v^k\Gamma^i_{ki}} \tag{5.110}$$

It will later be shown [in Eq. (5.126)] that

$$\Gamma^i_{ki} = \frac{1}{J}\frac{\partial J}{\partial \eta^k} \tag{5.111}$$

Therefore, Eq. (5.110) gives the very useful formula [7]:

$$\boxed{\nabla \bullet \underset{\sim}{v} = \frac{1}{J}\frac{\partial (Jv^k)}{\partial \eta^k}} \tag{5.112}$$

5.9 Curl of a vector

The curl of a vector $\underset{\sim}{\mathbf{v}}$ is denoted $\vec{\nabla} \times \underset{\sim}{\mathbf{v}}$ and it is defined to equal the axial vector[1] associated with $\vec{\nabla}\underset{\sim}{\mathbf{v}}$. Thus, recalling Eq. (5.100),

$$\vec{\nabla} \times \underset{\sim}{\mathbf{v}} = -\frac{1}{2}\underset{\approx}{\xi}\mathbf{:}\left(\frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}}\right)^T = \frac{1}{2}\underset{\approx}{\xi}\mathbf{:}\left(\frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}}\right) \tag{5.113}$$

In the last step, we have used the skew-symmetry of the permutation tensor. Substituting Eq. (5.77) gives

$$\vec{\nabla} \times \underset{\sim}{\mathbf{v}} = \frac{1}{2}\underset{\approx}{\xi}\mathbf{:}\left[\left(\frac{\partial v^i}{\partial \eta^j} + v^k\Gamma^i_{kj}\right)\underset{\sim}{\mathbf{g}}_i\underset{\sim}{\mathbf{g}}^j\right] \tag{5.114}$$

This result can be written as

$$\vec{\nabla} \times \underset{\sim}{\mathbf{v}} = \frac{1}{2}\left(\frac{\partial v^i}{\partial \eta^j} + v^k\Gamma^i_{kj}\right)\underset{\sim}{\mathbf{g}}_i \times \underset{\sim}{\mathbf{g}}^j \tag{5.115}$$

We can alternatively start with the covariant expression for the gradient given in Eq. (5.96). Namely,

$$\frac{d\underset{\sim}{\mathbf{v}}}{d\underset{\sim}{\mathbf{x}}} = v_{i/j}\underset{\sim}{\mathbf{g}}^i\underset{\sim}{\mathbf{g}}^j \text{ where } \qquad v_{i/j} \equiv \frac{\partial v_i}{\partial \eta^j} - v_k\Gamma^k_{ij}\,. \tag{5.116}$$

which gives

$$\vec{\nabla} \times \underset{\sim}{\mathbf{v}} = \frac{1}{2}\underset{\approx}{\xi}\mathbf{:}[v_{i/j}\underset{\sim}{\mathbf{g}}^i\underset{\sim}{\mathbf{g}}^j] = \frac{1}{2}\xi^{kij}v_{i/j}\underset{\sim}{\mathbf{g}}_k \tag{5.117}$$

Recalling from Eq. (3.47) that $\xi^{kij} = \frac{1}{J}\varepsilon^{ijk}$, the expanded component form of this result is

$$\vec{\nabla} \times \underset{\sim}{\mathbf{v}} = \frac{1}{2J}[(v_{2/3} - v_{3/2})\underset{\sim}{\mathbf{g}}_1 + (v_{3/1} - v_{1/3})\underset{\sim}{\mathbf{g}}_2 + (v_{1/2} - v_{2/1})\underset{\sim}{\mathbf{g}}_1] \tag{5.118}$$

where (recall)

$$J = (\underset{\sim}{\mathbf{g}}_1 \times \underset{\sim}{\mathbf{g}}_2) \bullet \underset{\sim}{\mathbf{g}}_3 \tag{5.119}$$

1. The axial vector associated with any tensor $\underset{\approx}{\mathbf{T}}$ is defined by $-\frac{1}{2}\underset{\approx}{\xi}\mathbf{:}\underset{\approx}{\mathbf{T}}$, where $\underset{\approx}{\xi}$ is the permutation tensor.

5.10 Gradient of a tensor

Using methods similar to the preceding sections, one can apply the direct notation definition of the gradient $d\mathbf{T}/d\mathbf{x}$ to eventually prove that

$$\frac{d\mathbf{T}}{d\mathbf{x}} = T^{ij}{}_{/k}\mathbf{g}_i\mathbf{g}_j\mathbf{g}^k \text{, where } T^{ij}{}_{/k} \equiv \frac{\partial T^{ij}}{\partial \eta^k} + T^{mj}\Gamma^i_{mk} + T^{im}\Gamma^j_{mk} \,, \tag{5.120}$$

or, if the tensor is known in its covariant components,

$$\frac{d\mathbf{T}}{d\mathbf{x}} = T_{ij/k}\mathbf{g}_i\mathbf{g}_j\mathbf{g}^k \text{, where } T_{ij/k} \equiv \frac{\partial T_{ij}}{\partial \eta^k} - T_{mj}\Gamma^m_{ik} - T_{im}\Gamma^m_{jk} \,. \tag{5.121}$$

For mixed tensor components, we have

$$\frac{d\mathbf{T}}{d\mathbf{x}} = T^i_{\bullet j/k}\mathbf{g}_i\mathbf{g}^j\mathbf{g}^k \text{, where} T^i_{\bullet j/k} \equiv \frac{\partial T^i_{\bullet j}}{\partial \eta^k} + T^m_{\bullet j}\Gamma^i_{mk} - T^i_{\bullet m}\Gamma^m_{jk} \,. \tag{5.122}$$

Note that there is a Christoffel term for each subscript on the tensor. The term is negative if the summed index is a subscript on T and positive if the summed index is a superscript on T.

Ricci's lemma Recall from Eq. (3.18) that the metric coefficients are components of the identity tensor. Knowing that the gradient of the identity tensor must be zero, it follows that

$$g_{ij/k} = 0 \quad \text{and} \quad g^{ij}{}_{/k} = 0 \,. \tag{5.123}$$

Corollary Recall that the Jacobian J equals the volume of the parallelepiped formed by the three base vectors $[\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3]$. Since the base vectors vary with position, it follows that J varies with position. In other words, the Jacobian J may be regarded as a function of the coordinates. Taking the derivative of the Jacobian with respect to the coordinate η^k and applying the chain rule gives

$$\frac{\partial J}{\partial \eta^k} = \frac{\partial J}{\partial g_{ij}}\frac{\partial g_{ij}}{\partial \eta_k} = \frac{1}{2}Jg^{ij}\frac{\partial g_{ij}}{\partial \eta_k} \text{, where we have used Eq. (2.30).} \tag{5.124}$$

Recalling from Eq. (5.123) that the metric components have vanishing *covariant derivatives*, Eq. (5.121) tells us that

$$\frac{\partial g_{ij}}{\partial \eta^k} = g_{mj}\Gamma^m_{ik} + g_{mi}\Gamma^m_{jk} \tag{5.125}$$

Thus, Eq. (5.124) becomes

$$\frac{\partial J}{\partial \eta^k} = \frac{1}{2}Jg^{ij}[g_{mj}\Gamma^m_{ik} + g_{mi}\Gamma^m_{jk}]$$

$$= \frac{1}{2}J[\delta^i_m\Gamma^m_{ik} + \delta^j_m\Gamma^m_{jk}]$$

$$= \frac{1}{2}J[\Gamma^m_{mk} + \Gamma^m_{mk}]$$
$$= J\Gamma^m_{mk} \tag{5.126}$$

or

$$\Gamma^m_{mk} = \frac{1}{J}\frac{\partial J}{\partial \eta^k} \tag{5.127}$$

5.11 Christoffel symbols of the first kind

Section 5.6 showed how to compute the Christoffel symbols of the second kind by directly applying the formula:

$$\Gamma^k_{ij} = \underset{\sim}{\mathbf{g}}^k \bullet \frac{\partial \underset{\sim}{\mathbf{g}}_i}{\partial \eta^j}. \tag{5.128}$$

This method was tractable because the space was Euclidean and we could therefore utilize the existence of an underlying rectangular Cartesian basis to determine the change in the base vectors with respect to the coordinates. However, for non-Euclidean spaces, you have only the metric coefficients g_{ij}, and the simplest way to compute the Christoffel symbols of the second kind then begins by first computing the **Christoffel symbols of the first kind** Γ_{ijk}, which are related to the Christoffel symbols of the second kind by

$$\Gamma_{ijk} = \Gamma^m_{ij} g_{mk}. \tag{5.129}$$

These Γ_{ijk} are frequently denoted in the literature as $[ij, k]$, presumably to emphasize that they are not components of any third-order-tensor, in the sense discussed on page __. Substituting Eq. (5.128) into Eq. (5.129) gives

$$\Gamma_{ijk} = \left(\underset{\sim}{\mathbf{g}}^m \bullet \frac{\partial \underset{\sim}{\mathbf{g}}_i}{\partial \eta^j}\right) g_{mk} = \left(\frac{\partial \underset{\sim}{\mathbf{g}}_i}{\partial \eta^j}\right) \bullet \underset{\sim}{\mathbf{g}}^m g_{mk} = \frac{\partial \underset{\sim}{\mathbf{g}}_i}{\partial \eta^j} \bullet \underset{\sim}{\mathbf{g}}_k, \tag{5.130}$$

where the final term results from lowering the index on $\underset{\sim}{\mathbf{g}}^m$. Thus, using Eq. (5.63),

$$\Gamma_{ijk} = \underset{\sim}{\Gamma}_{ij} \bullet \underset{\sim}{\mathbf{g}}_k. \tag{5.131}$$

Recalling that $\underset{\sim}{\Gamma}_{ij} = \underset{\sim}{\Gamma}_{ji}$, this result also reveals that Γ_{ijk} is symmetric in its first two indices:

$$\Gamma_{ijk} = \Gamma_{jik}. \tag{5.132}$$

Now note that

$$\frac{\partial(g_{ik})}{\partial \eta^j} = \frac{\partial(\underset{\sim}{\mathbf{g}}_i \bullet \underset{\sim}{\mathbf{g}}_k)}{\partial \eta^j} = \frac{\partial \underset{\sim}{\mathbf{g}}_i}{\partial \eta^j} \bullet \underset{\sim}{\mathbf{g}}_k + \underset{\sim}{\mathbf{g}}_i \bullet \frac{\partial \underset{\sim}{\mathbf{g}}_k}{\partial \eta^j} = \underset{\sim}{\Gamma}_{ij} \bullet \underset{\sim}{\mathbf{g}}_k + \underset{\sim}{\mathbf{g}}_i \bullet \underset{\sim}{\Gamma}_{kj}. \tag{5.133}$$

Using Eq. (5.131) gives[1]

$$\frac{\partial g_{ik}}{\partial \eta^j} = \Gamma_{ijk} + \Gamma_{kji} . \tag{5.134}$$

This relationship can be easily remembered by noting the structure of the indices. Note, for example, that the middle subscript on both Γ's is the same as that on the coordinate η^j.

Direct substitution of Eq. (5.134) into (5.135) validates the following expression that can be used to directly obtain the Christoffel symbols of the first kind given only the metric coefficients:

$$\boxed{\Gamma_{ijk} = \frac{1}{2}\left[\frac{\partial g_{jk}}{\partial \eta^i} + \frac{\partial g_{ki}}{\partial \eta^j} - \frac{\partial g_{ij}}{\partial \eta^k}\right]} . \tag{5.135}$$

This formula (which can be verified by substituting Eq. (5.134) into (5.135) represents the easiest way to obtain the Christoffel symbols of the first kind when only the metric coefficients are known. This formula has easily remembered index structure: for Γ_{ijk}, the index symbols on each of the η coordinates are ordered i, j, k. Once the Γ_{ijk} are known, the Christoffel symbols of the second kind are obtained by solving Eq. (5.129) to give

$$\Gamma_{ij}^n = \Gamma_{ijk} g^{kn} . \tag{5.136}$$

5.12 The fourth-order Riemann-Christoffel curvature tensor

Recall that this document has limited its scope to curvilinear systems embedded in a *Euclidean space* of the same dimension. When the Euclidean space is three-dimensional, then this document has focused on alternative coordinate systems that still define points in this space and therefore require three coordinates. An example of a two-dimensional curvilinear space embedded in Euclidean space is the surface of a sphere. Only two coordinates are required to specify a point on a sphere and this two dimensional space is "*Riemannian*" (i.e., non-Euclidean) because it is not possible for us to construct a rectangular coordinate grid on a sphere. In ordinary engineering mechanics applications, the mathematics of reduced dimension spaces within ordinary physical space is needed to study plate and beam theory.

We have focused on the three-dimensional Euclidean space that we *think* we live it. In this modern (post-Einstein) era, this notion of a Euclidean physical world is now recognized to be only an approximation (referred to as Newtonian space). Einstein and colleagues threw a wrench in our thinking by introducing the notion that space and time are *curved*.

We have now mentioned two situations (prosaic shell/beam theory and exciting relativity theory) where understanding some basics of non-Euclidean spaces is needed. We have men-

1. An alternative way to obtain this result is to simply apply Eq. (5.129) to Eq. (5.125).

tioned that a Riemannian space is one that does not permit construction of a rectangular coordinate grid. How is this statement cast in mathematical terms? In other words, what process is needed to decide if a space is Euclidean or Riemannian in the first place. The answer is tied to a special fourth-order tensor called the Riemann-Christoffel tensor soon to be defined. If this tensor turns out to be zero, then your space is Euclidean. Otherwise, it is Riemannian.

The Riemann-Christoffel tensor is defined

$$R_{ijkl} = \frac{\partial \Gamma_{jli}}{\partial \eta^k} - \frac{\partial \Gamma_{jki}}{\partial \eta^l} + \Gamma_{ilp}\Gamma^p_{jk} - \Gamma_{ikp}\Gamma^p_{jl} \tag{5.137}$$

or

$$R_{ijkl} = \frac{1}{2}\left(\frac{\partial^2 g_{il}}{\partial \eta^j \partial \eta^k} - \frac{\partial^2 g_{jl}}{\partial \eta^i \partial \eta^k} - \frac{\partial^2 g_{ik}}{\partial \eta^j \partial \eta^l} + \frac{\partial^2 g_{jk}}{\partial \eta^i \partial \eta^l}\right) + g^{mn}(\Gamma_{jkn}\Gamma_{ilm} - \Gamma_{jlm}\Gamma_{ikn}) \tag{5.138}$$

This tensor is skew-symmetric in the first two indices and in the last two indices. It is major symmetric:

$$R_{ijkl} = -R_{jikl} = -R_{ijlk} = R_{klij} \tag{5.139}$$

In a two-dimensional space, only the R_{1212} component is independent, the others being either zero or related to this component by $R_{1212} = -R_{2112} = -R_{1221} = R_{2121}$.

Note from Eq. (5.137) that R_{ijkl} depends on the Christoffel symbols. For a Cartesian system, all Christoffel symbols are zero. Hence, in a Cartesian system, $R_{ijkl} = 0$. Thus, if a space is *capable* of supporting a Cartesian system (i.e., if the space is Euclidean), then the Riemann-Christoffel tensor must be zero. This would be true even if you aren't actually using a Cartesian system. For example, ordinary 3D Newtonian space is Euclidean and therefore its Riemann-Christoffel tensor must be zero even if you happen to employ a different set of three curvilinear coordinates such as spherical coordinates (r, θ, ϕ). This would follow because the transformation relations are linear. Hence, if there exists a Cartesian system (in which the Riemann-Christoffel tensor is zero), a linear transformation to a different, possibly curvilinear, system would result again in the zero tensor. For the Riemann-Christoffel tensor to *not* be zero, you would have to be working in a reduced dimension space [such as the *surface* of a sphere where the coordinates are (θ, ϕ)]. The Riemann-Christoffel tensor is, therefore, a measure of curvature of a Riemannian space. Because the Riemann-Christoffel tensor transforms like a tensor, it is *not* a basis-intrinsic quantity despite the fact that it has been here defined in terms of basis-intrinsic quantities. A similar situation was encountered with the metric coefficients g_{ij}, which were defined $g_{ij} = \underset{\sim}{\mathbf{g}}_i \bullet \underset{\sim}{\mathbf{g}}_j$. The metrics g_{ij} are components of the identity tensor. Therefore the product of these components times base vectors $g_{ij}\underset{\sim}{\mathbf{g}}^i\underset{\sim}{\mathbf{g}}^j$ is basis-independent (it equals the identity tensor). Similarly, the product of the Riemann-Christoffel components times basis vectors is basis-independent:

$$\underset{\approx}{\mathbf{R}} = R_{ijkl}\underset{\sim}{\mathbf{g}}^i\underset{\sim}{\mathbf{g}}^j\underset{\sim}{\mathbf{g}}^k\underset{\sim}{\mathbf{g}}^l \tag{5.140}$$

or

$$\mathbf{R} = \left[\frac{\partial\Gamma_{jli}}{\partial\eta^k} - \frac{\partial\Gamma_{jki}}{\partial\eta^l} + \Gamma_{ilp}\Gamma^p_{jk} - \Gamma_{ikp}\Gamma^p_{jl}\right]\mathbf{g}^i\mathbf{g}^j\mathbf{g}^k\mathbf{g}^l \tag{5.141}$$

or

$$\mathbf{R} = \left[\frac{\partial(\boldsymbol{\Gamma}_{jl}\bullet\mathbf{g}_i)}{\partial\eta^k} - \frac{\partial(\boldsymbol{\Gamma}_{jk}\bullet\mathbf{g}_i)}{\partial\eta^l} + (\boldsymbol{\Gamma}_{il}\bullet\mathbf{g}_p)(\boldsymbol{\Gamma}_{jk}\bullet\mathbf{g}^p) - (\boldsymbol{\Gamma}_{ik}\bullet\mathbf{g}_p)(\boldsymbol{\Gamma}_{jl}\bullet\mathbf{g}^p)\right]\mathbf{g}^i\mathbf{g}^j\mathbf{g}^k\mathbf{g} \tag{5.142}$$

or

$$\mathbf{R} = [\boldsymbol{\Gamma}_{jlk}\bullet\mathbf{g}_i + \boldsymbol{\Gamma}_{jl}\bullet\boldsymbol{\Gamma}_{ik} - \boldsymbol{\Gamma}_{jkl}\bullet\mathbf{g}_i - \boldsymbol{\Gamma}_{jk}\bullet\boldsymbol{\Gamma}_{il} + (\boldsymbol{\Gamma}_{il}\bullet\mathbf{g}_p)(\boldsymbol{\Gamma}_{jk}\bullet\mathbf{g}^p) - (\boldsymbol{\Gamma}_{ik}\bullet\mathbf{g}_p)(\boldsymbol{\Gamma}_{jl}\bullet\mathbf{g}^p)]\mathbf{g}^i\mathbf{g}^j\mathbf{g}^k\mathbf{g}^l \tag{5.143}$$

where

$$\boldsymbol{\Gamma}_{ijk} = \frac{\partial\boldsymbol{\Gamma}_{ij}}{\partial\eta^k} = \frac{\partial^2\mathbf{g}_i}{\partial\eta^j\partial\eta^k} = \frac{\partial^3\mathbf{x}}{\partial\eta^i\partial\eta^j\partial\eta^k} \tag{5.144}$$

Noting that the indices on this tensor may be permuted in any manner, note that the first and third terms in Eq. (5.143) may be canceled, giving

$$\mathbf{R} = [\boldsymbol{\Gamma}_{jl}\bullet\boldsymbol{\Gamma}_{ik} - \boldsymbol{\Gamma}_{jk}\bullet\boldsymbol{\Gamma}_{il} + \boldsymbol{\Gamma}_{il}\bullet\boldsymbol{\Gamma}_{jk} - \boldsymbol{\Gamma}_{ik}\bullet\boldsymbol{\Gamma}_{jl}]\mathbf{g}^i\mathbf{g}^j\mathbf{g}^k\mathbf{g}^l \tag{5.145}$$

or, rearranging slightly,

$$\mathbf{R} = -[\boldsymbol{\Gamma}_{ik}\bullet\boldsymbol{\Gamma}_{jl} - \boldsymbol{\Gamma}_{il}\bullet\boldsymbol{\Gamma}_{jk} + \boldsymbol{\Gamma}_{jk}\bullet\boldsymbol{\Gamma}_{il} - \boldsymbol{\Gamma}_{jl}\bullet\boldsymbol{\Gamma}_{ik}]\mathbf{g}^i\mathbf{g}^j\mathbf{g}^k\mathbf{g}^l \tag{5.146}$$

or

$$\mathbf{R} = -[g_{ikjl} - g_{iljk} + g_{jkil} - g_{jlik}]\mathbf{g}^i\mathbf{g}^j\mathbf{g}^k\mathbf{g}^l \tag{5.147}$$

where

$$g_{ijkl} \equiv \boldsymbol{\Gamma}_{ij}\bullet\boldsymbol{\Gamma}_{kl} = \frac{\partial^2\mathbf{x}}{\partial\eta^i\partial\eta^j}\bullet\frac{\partial^2\mathbf{x}}{\partial\eta^k\partial\eta^l} \tag{5.148}$$

6. Embedded bases and objective rates

We introduced the mapping tensor $\mathbf{F}$ in Eq. (2.1) to serve as a mere "helper" tensor. Namely, if the basis $\{\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3\}$ exists in a Euclidean space, then we defined

$$\mathbf{g}_i = \mathbf{F}\bullet\mathbf{E}_i \tag{6.1}$$

Here, the basis, $\{\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3\}$ is the same as the "laboratory" basis that we had originally used in Section 2. In continuum mechanics (especially the older literature), a special "convected" coordinate system is often adopted in which the coordinates η^k of a point currently located a a position $\underset{\sim}{\mathbf{x}}$ at time t are identified to equal the *initial* Cartesian coordinates X^k of the same material particle at time 0. For this special case, the coordinate grid is identical to the deformation of the initial coordinate grid when it is convected along with the material particles. For this special case, the mapping tensor $\underset{\approx}{\mathbf{F}}$ in Eq. (6.1) is identical to the deformation gradient tensor from traditional continuum mechanics. In this physical context, the term "covariant" is especially apropos because the covariant base vector $\underset{\sim}{\mathbf{g}}_i$ varies coincidently with the underlying material particles. That is, the unit cube is defined by three unit vectors $\{\underset{\sim}{\mathbf{E}}_1, \underset{\sim}{\mathbf{E}}_2, \underset{\sim}{\mathbf{E}}_3\}$ will, upon deformation, deform to a parallelepiped whose sides move with the material and the sides of this parallelepiped are given by the three convected base vectors $\{\underset{\sim}{\mathbf{g}}_1, \underset{\sim}{\mathbf{g}}_2, \underset{\sim}{\mathbf{g}}_3\}$.

By contrast, consider the contravariant base vectors, which (recall) are related to the mapping tensor by

$$\underset{\sim}{\mathbf{g}}^k = \underset{\approx}{\mathbf{F}}^{-1} \bullet \underset{\sim}{\mathbf{E}}^k \tag{6.2}$$

These base vectors do *not* move coincidently with the material particles. Instead, the contravariant base vectors move somewhat contrary to the material motion. In particular, if $\{\underset{\sim}{\mathbf{E}}^1, \underset{\sim}{\mathbf{E}}^2, \underset{\sim}{\mathbf{E}}^3\}$ are the outward unit normals to the unit cube, then material deformation will generally move the cube to a parallelepiped whose faces now have outward unit normals parallel to $\{\underset{\sim}{\mathbf{g}}^1, \underset{\sim}{\mathbf{g}}^2, \underset{\sim}{\mathbf{g}}^3\}$. In general, the motion of the normal to a plane moves somewhat contrary to the motion of material fibers that were originally parallel to the plane's normal.

Of course, it is not really necessary for the initial coordinate system to be Cartesian itself. When the initial coordinate system is permitted to be curvilinear, then we will denote its associated set of base vectors by $\{\underset{\sim}{\mathbf{G}}_1, \underset{\sim}{\mathbf{G}}_2, \underset{\sim}{\mathbf{G}}_3\}$. As before, these *covariant* base vectors deform to new orientations given by

$$\underset{\sim}{\mathbf{g}}_i = \underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{G}}_i \tag{6.3}$$

The associated dual basis is given by

$$\underset{\sim}{\mathbf{g}}^k = \underset{\approx}{\mathbf{F}}^{-T} \bullet \underset{\sim}{\mathbf{G}}^k \tag{6.4}$$

Here $\underset{\approx}{\mathbf{F}}$ is here retaining its meaning as the physical deformation gradient tensor, which necessitates introducing new symbols to denote the mapping tensors for the individual curvilinear bases. Namely, we will presume the existence of tensors $\underset{\approx}{\mathbf{h}}$ and $\underset{\approx}{\mathbf{H}}$ such that

$$\underset{\sim}{\mathbf{g}}_i = \underset{\approx}{\mathbf{h}} \bullet \underset{\sim}{\mathbf{E}}_i \text{ and } \underset{\sim}{\mathbf{g}}^k = \underset{\approx}{\mathbf{h}}^{-T} \bullet \underset{\sim}{\mathbf{E}}^k \tag{6.5}$$

$$\underset{\sim}{\mathbf{G}}_i = \underset{\approx}{\mathbf{H}} \bullet \underset{\sim}{\mathbf{E}}_i \text{ and } \underset{\sim}{\mathbf{G}}^k = \underset{\approx}{\mathbf{H}}^{-T} \bullet \underset{\sim}{\mathbf{E}}^k \tag{6.6}$$

from which it follows that

$$g_{ij} = \underset{\sim}{\mathbf{E}}_i \bullet \underset{\approx}{y} \bullet \underset{\sim}{\mathbf{E}}_j \text{ and } g^{ij} = \underset{\sim}{\mathbf{E}}^i \bullet \underset{\approx}{y}^{-1} \bullet \underset{\sim}{\mathbf{E}}^j, \text{ where } \underset{\approx}{y} \equiv \underset{\approx}{\mathbf{h}}^T \bullet \underset{\approx}{\mathbf{h}} \tag{6.7}$$

$$G_{ij} = \underset{\sim}{\mathbf{E}}_i \bullet \underset{\approx}{\mathbf{Y}} \bullet \underset{\sim}{\mathbf{E}}_j \text{ and } G^{ij} = \underset{\sim}{\mathbf{E}}^i \bullet \underset{\approx}{\mathbf{Y}}^{-1} \bullet \underset{\sim}{\mathbf{E}}^j, \text{ where } \underset{\approx}{\mathbf{Y}} \equiv \underset{\approx}{\mathbf{H}}^T \bullet \underset{\approx}{\mathbf{H}} \tag{6.8}$$

Furthermore, substituting Eqs. (6.5) and (6.6) into (6.3) implies that

$$\underset{\approx}{\mathbf{F}} = \underset{\approx}{\mathbf{h}} \bullet \underset{\approx}{\mathbf{H}}^{-1} \tag{6.9}$$

In continuum mechanics, the Seth-Hill generalized strain [13] measure is defined in direct notation as

$$\underset{\approx}{e} = \frac{1}{k}[(\underset{\approx}{\mathbf{F}}^T \bullet \underset{\approx}{\mathbf{F}})^{k/2} - \underset{\approx}{\mathbf{I}}] \tag{6.10}$$

Here, the Seth-Hill parameter k is selected according to whichever strain measure the analyst prefers. Namely,

$k = -1$	-->	True/Swainger strain
$k=1$	-->	engineering/Cauchy strain
$k=2$	-->	Green/Lagrange strain
$k=-2$	-->	Almansi
$k=0$	-->	logarithmic/Hencky strain

Of course, the case of $k=0$ must be applied in the limit.

The Lagrangian strain measure corresponds to choosing $k=2$ to give

$$\underset{\approx}{e}^{\text{Lagr}} = \frac{1}{2}[(\underset{\approx}{\mathbf{F}}^T \bullet \underset{\approx}{\mathbf{F}}) - \underset{\approx}{\mathbf{I}}] \tag{6.11}$$

The "Euler" strain measure corresponds to choosing $k=-2$ to give

$$\underset{\approx}{e}^{\text{Euler}} = \frac{1}{2}[\underset{\approx}{\mathbf{I}} - (\underset{\approx}{\mathbf{F}}^T \bullet \underset{\approx}{\mathbf{F}})^{-1}] \tag{6.12}$$

Strain measures in older literature look drastically different from this because they are typically expressed in terms of initial and deformed metric tensors. What is the connection? First note from Eq. (6.3) that

$$g_{ij} = \underset{\sim}{\mathbf{G}}_i \bullet \underset{\approx}{\mathbf{F}}^T \bullet \underset{\approx}{\mathbf{F}} \bullet \underset{\sim}{\mathbf{G}}_j \tag{6.13}$$

Furthermore, by definition,

$$G_{ij} \equiv \underset{\sim}{\mathbf{G}}_i \bullet \underset{\sim}{\mathbf{G}}_j = \underset{\sim}{\mathbf{G}}_i \bullet \underset{\approx}{\mathbf{I}} \bullet \underset{\sim}{\mathbf{G}}_j \tag{6.14}$$

Thus, dotting Eq. (6.11) from the left by $\underset{\sim}{\mathbf{G}}_i$ and from the right by $\underset{\sim}{\mathbf{G}}_j$ gives

$$\underset{\sim}{\mathbf{G}}_i \bullet \underset{\approx}{e}^{\text{Lagr}} \bullet \underset{\sim}{\mathbf{G}}_j = \frac{1}{2}[g_{ij} - G_{ij}] \tag{6.15}$$

This result shows that the difference between deformed and initial covariant metrics (which appears frequently in older literature) is identically equal to the covariant components of the Lagrangian strain tensor with respect to the initial basis.

Similarly, you can show that

$$\underset{\sim}{\mathbf{G}}^i \bullet \underset{\approx}{e}^{\text{Euler}} \bullet \underset{\sim}{\mathbf{G}}^j = \frac{1}{2}[G^{ij} - g^{ij}] \tag{6.16}$$

In general, if you wish to convert a component equation from the older literature into modern invariant direct notation form, you can use Eqs. (6.5), (6.6) and (6.9) as long as you can deduce which basis applies to the component formulas. Converting an index equation to direct symbolic form is harder than the reverse (which is a key argument used in favor of direct symbolic formulations of governing equations). Consider, for example, the definition of the second Piola Kirchhoff tensor:

$$\underset{\approx}{\mathbf{s}} \equiv \underset{\approx}{\mathbf{F}}^{-1} \bullet \underset{\approx}{\tau} \bullet \underset{\approx}{\mathbf{F}}^{-T}, \text{ where } \underset{\approx}{\tau} = J\underset{\approx}{\sigma}. \tag{6.17}$$

Here, $J = \det(\underset{\approx}{\mathbf{F}})$ and $\underset{\approx}{\sigma}$ is the Cauchy stress. The tensor $\underset{\approx}{\tau}$ is called the “Kirchhoff” stress, and it is identically equal to the Cauchy stress for incompressible materials. Dotting both sides of the above equation by the initial contravariant base vectors gives

$$\underset{\sim}{\mathbf{G}}^i \bullet \underset{\approx}{\mathbf{s}} \bullet \underset{\sim}{\mathbf{G}}^j = \underset{\sim}{\mathbf{G}}^i \bullet \underset{\approx}{\mathbf{F}}^{-1} \bullet \underset{\approx}{\tau} \bullet \underset{\approx}{\mathbf{F}}^{-T} \bullet \underset{\sim}{\mathbf{G}}^j \tag{6.18}$$

or, using Eq. (6.4),

$$\underset{\sim}{\mathbf{G}}^i \bullet \underset{\approx}{\mathbf{s}} \bullet \underset{\sim}{\mathbf{G}}^j = \underset{\sim}{\mathbf{g}}^i \bullet \underset{\approx}{\tau} \bullet \underset{\sim}{\mathbf{g}}^j \tag{6.19}$$

This shows that the contravariant components of the second Piola Kirchhoff tensor with respect to the *initial* basis are equal to the contravariant components of the Kirchhoff tensor with respect to the *current* basis. Results like this are worth noting and recording in your personal file so that you can quickly convert older curvilinear constitutive model equations into modern direct notation form.

7. Concluding remarks.

R.B. Bird et al. [1] astutely remark "...some authors feel it is stylish to use general tensor notation even when it is not needed, and hence it is necessary to be familiar with the use of base vectors and covariant and contravariant components in order to study the literature..."

Modern researchers realize that operations such as the dot product, cross product, and gradient are proper tensor operations. Such operations commute with basis and/or coordinate transformations as long as the computational procedure for evaluating them is defined in a manner appropriate for the underlying coordinates or basis. In other words, one can apply the applicable formula for such an operation in *any convenient system* and transform the result to a second system, and the result will be the same as if the operation had been applied directly in the second system at the outset.

Once an operation is known to be a proper tensor operation, it is endowed with a structured (direct) notation symbolic form. Derivations of new identities usually begin by casting a direct notation formula in a conveniently simple system such as a principal basis or maybe just ordinary rectangular Cartesian coordinates. From there, proper tensor operations are performed within that system. In the final step, the result is re-cast in structured (direct) notation. A structured result can then be justifiably expressed into *any other system* because all the operations used in the derivation had been proper operations. Consequently, one should always perform derivations using only proper tensor operations using whatever coordinate system makes the analysis accessible to the largest audience of readers. The last step is to cast the final result back in direct notation, thereby allowing it to be recast in any other desired basis or coordinate system.

8. REFERENCES

All of the following references are *books*, not journal articles. This choice has been made to emphasize that the subject of curvilinear analysis has been around for centuries -- there exists no new literature for us to quote that has not already been quoted in one or more of the following texts. So why, you might ask, is this new manuscript called for? The answer is that the present manuscript is designed specifically for new students and researchers. It contains many heuristically derived results. Any self-respecting student should follow up this manuscript's elementary introduction with careful reading of the following references.

Greenberg's applied math book should be part of any engineer's personal library. Readers should make an effort to read all footnotes and student exercises in Greenberg's book — they're both enlightening and entertaining! Simmond's book is a good choice for continued pursuit of curvilinear coordinates. Tables of gradient formulas for particular coordinate systems can be found in the appendix of Bird, Armstrong, and Hassager's book.

[1] Bird, R.B., Armstrong, R.C., Hassager, O., **Dynamics of Polymeric Liquids**. Wiley. (1987).

[2] Schreyer, H.L., **Introduction to Continuum Mechanics**, University of New Mexico internal report. (1976)

[3] Arfken, G.B and Weber, H.J., **Mathematical Methods for Physicists**. Academic Press (1995).

[4] McConnel, A.J., **Applications of Tensor Analysis**, Dover, NY, (1957)

[5] Lovelock, David and Rund, Hanno. **Tensors, differential forms, and variational principles.**, Wiley, NY (1975).

[6] Ramkrishna, D. and Amundson, NR. **Linear Operator Methods in Chemical Engineering**, Prentice-Hall, New Jersey, (1985).

[7] Simmonds, James G. **A Brief on Tensor Analysis**, Springer-Verlag, NY (1994).

[8] Buck, R.C., **Advanced Calculus**, McGraw-Hill, NY (1978)

[9] Rudin, W., **Real and Complex Analysis**, 3rd Ed., McGraw-Hill, NY (1987).

[10] Greenberg, M.D., **Foundations of applied mathematics**, Prentice Hall, New Jersey (1978).

[11] Malvern, L.E., **Introduction to the mechanics of a continuous medium.** Prentice-Hall, New Jersey (1969).

[12] Papastravridis, John G., **Tensor Calculus and Analytical Dynamics**, CRC Press, (1998).

[13] Narasimhan, Mysore N., **Principles of continuum mechanics.**, Wiley-Interscience, NY (1993.

[14]Abraham, R., Marsden, J.E., and Ratiu, T. **Manifolds, Tensor Analysis, and Applications, 2nd Ed.** (Vol. 75 in series entitled: **Applied Mathematical Sciences**). Springer, N.Y. (1988).