

Contents

Part A. Mathematical Preliminaries

Chapter A1. Some Basic Concepts

A1.1. Sets and Lists

A1.2. Functions

A1.3. Groups

Chapter A2. Linear Spaces

A2.1. Definition of a Linear Space

A2.2. Finite-Dimensional Linear Spaces

A2.3. Inner Product Spaces. Normed Linear Spaces

A2.4. Linear Transformations

A2.5. Linear Functionals. Dual Bases. Covariant and Contravariant Components

A2.6. Transformations of Bases

A2.7. Volume Orientation Functionals. Oriented Spaces

A2.8. The Vector Product

Chapter A3. Second-Order Tensors

A3.1. Definition of a Second-Order Tensor. The Tensor Product

- A3.2. Second-Order Tensors in Finite-Dimensional Linear Spaces
- A3.3. Transformations of Bases
- A3.4. Boundedness, Continuity, Operator Norm
- A3.5. Multiplication of Second-Order Tensors
- A3.6. Transposition, Symmetric Tensors
- A3.7. The Trace of a Second-Order Tensor. The Trace Inner Product
- A3.8. The Determinant of a Second-Order Tensor
- A3.9. The Inverse of a Second-Order Tensor
- A3.10. Orthogonal Second-Order Tensors
- A3.11. Eigenvalues, Eigenvectors, and Eigenspaces of a Second-Order Tensor
- A3.12. Direct Sum Decompositions, Invariant Linear Subspaces
- A3.13. The Spectral Theorem for Symmetric Tensors in Finite-Dimensional Spaces
- A3.14. The Hamilton-Cayley Theorem
- A3.15. Positive Definite Tensors. Tensor Square Roots. The Polar Decomposition Theorem

Chapter A4. Higher-Order Tensors

Chapter A5. Differentiation on Finite-Dimensional Linear Spaces

- A5.1. Metric Spaces. Open Sets
- A5.2. The Derivative
- A5.3. Functions of a Real Variable. The Ordinary Derivative
- A5.4. Real-Valued Functions. The Gradient
- A5.5. Elementary Rules of Differentiation
- A5.6. Derivatives of Products
- A5.7. Derivatives of Composite Functions. The Chain Rule
- A5.8. Some Additional Derivatives. The Invariants and the Inverse of a Tensor
- A5.9. The Mean-Value Theorem for Derivatives
- A5.10. Higher Derivatives. The CP Spaces
- A5.11. The Inverse Function Theorem. The Tensor Square Root
- A5.12. Functions of Several Variables. Partial Derivatives
- A5.13. The Codomain a Product Space
- A5.14. Identification with the Component Approach

Chapter A6. Point Spaces

- A6.1. Definition of a Point Space
- A6.2. Finite-Dimensional Point Spaces. Cartesian Coordinate Systems
- A6.3. Euclidean Point Spaces. Point Set Topology
- A6.4. Fields
- A6.5. Scalar Fields
- A6.6. Vector Fields
- A6.7. Second-Order Tensor Fields

- A6.8. Point Fields
- A6.9. Some Identities
- A6.10. Some Integral Theorems
- A6.11. General Coordinate Systems
- A6.12. Some Particular Coordinate Systems
- A6.13. Coordinate Transformations
- A6.14. Differentiation of Scalar Fields
- A6.15. Differentiation of Vector Fields. CHRISTOFFEL Symbols
- A6.16. Differentiation of Second-Order Tensor Fields
- A6.17. Differentiation of Point Fields

Chapter A7. The Tensor Exponential Function

v

Chapter AB. Isotropic Tensors and Isotropic Tensor Functions

AB.1. Isotropic Tensors

AB.2. Invariant Sets and Functions

AB.3. Isotropic Scalar-Valued Functions of a Vector

AB.4. Isotropic Scalar-Valued Functions of a Tensor

AB.5. Isotropic Tensor-Valued Functions of a Tensor

Part B. General Continuum Mechanics

Chapter B1. Kinematics

B1.1. Deformations

B1.2. Rigid Deformations. The Extension Theorem

B1.3. Proper Rotations

B1.4. Homogeneous Deformations

B1.5. The Geometry of General Deformations

B1.6. Isochoric Deformations. The Localization Theorem

B1.7. Bodies and Configurations

B1.8. Motions. Material and Spatial Descriptions. Functions of Position and Time

B1.9. Path Lines, Stream Lines, and Streak Lines

B1.10. Rigid Motions

- B1.11. The Geometry of Motions
- B1.12. Isochoric Motions. Parts of a Body
- B1.13. Reference Configurations. The Transport Theorem
- B1.14. The Referential Description of Rigid and Isochoric Motions
- B1.15. Coordinate Analysis in the Referential Description
- B1.16. The Present Configuration as the Reference Configuration.
The Relative Description
- B1.17. The Relative Description of Isochoric and Rigid Motions
- B1.18. Coordinate Analysis in the Relative Description

Chapter B2. Mass. Momentum. Kinetic Energy

- B2.1. Mass Density. Conservation of Mass
- B2.2. Mass Conservation in Isochoric and Rigid Motions
- B2.3. Center of Mass
- B2.4. Momentum, Kinetic Energy
- B2.5. Momentum and Kinetic Energy for Rigid Motions

Chapter B3. Force. Stress. Momentum Balance

- B3.1. The Stress Principle and Balance of Momentum

- B3.2. Momentum Balance for Rigid Bodies
- B3.3. The Theorem of Virtual Power
- B3.4. The Stress Tensor and the Stress Equations of Motion
- B3.5. Additional Properties of the Stress Tensor
- B3.6. Engineering Stress

Chapter B4. A Theory of Constitutive Equations

- B4.1. Simple Materials. The Response Function
- B4.2. The Principle of Material Indifference
- B4.3. Observer Transformations. Another Approach to Material Indifference
- B4.4. Material Indifference in the Relative Description
- B4.5. Material Symmetry
- B4.6. NOLL's Classification. Isotropic, Fluid, and Solid Materials
- B4.7. Material Symmetry in the Relative Description
- B4.8. Internal Constraints
- B4.9. Incompressible, Inextensible, and Rigid Bodies
- B4.10. Material Symmetry of Constrained Bodies

Part C. Examples of Simple Materials

Chapter C1. Some Fluids

- C1.1. REINER - RIVLIN Fluids
- C1.2. NEWTONIAN Fluids
- C1.3. RIVLIN - ERICKSEN Fluids
- C1.4. Viscometric Flows of Incompressible Simple Fluids

Chapter C2. Elasticity

- C2.1. The Elastic Response Function
- C2.2. Material Symmetry
- C2.3. Some Equilibrium Solutions for Compressible, Isotropic, Homogeneous, Elastic Solids
- C2.4. Some Equilibrium Solutions for Incompressible, Isotropic, Homogeneous, Elastic Solids

A6.4, Fields

Functions whose domains are sets of points are a common occurrence in the physical sciences, and it is traditional to refer to such "functions of position" as fields.¹ e.g., one speaks of the temperature field, the velocity field, etc. Here, not yet having gotten to any physics, we shall speak of scalar fields, vector fields, etc.; and, of course, by this terminology mean scalar-valued functions of position, vector-valued functions of position, etc.

To be precise let $f: D \rightarrow W$, where D is an open subset of E_n , an n -dimensional Euclidean point space with translation space V_n , and W is a finite-dimensional linear space.² Of course, W can be the special linear spaces $\mathbb{R}$ and $L(V_n)$. We have chosen notation that is as close as possible to that employed in the previous chapter because we are particularly interested in the differentiation of fields. Here the domain space E_n is not a linear space, but it is a metric space³, and that is all we need. Guided by Definition

¹ This terminology should not be confused with the algebraic concept of scalar field, examples of which are the real number field and the complex number field.

² This notation is understood throughout this section.

³ Cf. Theorem A6.3.1.

A5.2.2, we make

Definition A6.4.1. The field $f: D \rightarrow W$ is differentiable at the point $x \in D$ if $\exists$ a linear transformation

$$Df(x): V_n \rightarrow W \quad (V_n \text{ is the translation space of } E_n.)$$

$\exists$ for sufficiently small $|u|$ ($u \in V_n$)

$$f(x+u) - f(x) = Df(x)[u] + o(u) \text{ as } u \rightarrow 0.$$

The linear transformation $Df(x)$ is called the derivative' of f at x .

Of course, in view of Definition A6.1.1, for $x \in D \subset E_n$ and $u \in V_n$, $x+u$ is a point in E_n ; in fact, in D for $|u|$ sufficiently small.

Since Definition A6.4.1 is essentially the same as Definition A5.2.2, all of the results of Chapter 5 carry over to fields provided their values reside in finite-dimensional linear spaces.

In older work in mechanics it is customary to deal with fields in terms of the components of the function values and the

'More often than not in continuum mechanics, derivatives of fields are called gradients and denoted by $\nabla f(x)$ instead of by $Df(x)$. We shall reserve the gradient terminology and the ∇ notation for scalar-valued functions (cf. §A5.4).

coordinates of the arguments. Usually, this obscures the general theory, but it is useful in the solution of particular problems. We take a limited look at this approach next.

Let $\{O; \underline{e}_{\langle i \rangle}\}$ be a reference frame¹ for E_n with the basis $\{\underline{e}_{\langle i \rangle}\}$ for the translation space $\mathcal{V}_n$ orthonormal. The position vector of a typical point $\underline{x} \in E_n$ relative to the origin O can be expressed as

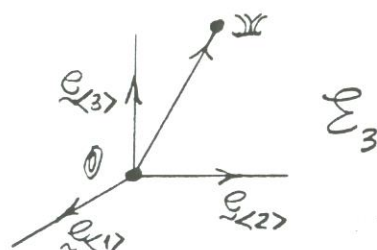
$$\overrightarrow{O\underline{x}} = \underline{x} = x^i \underline{e}_{\langle i \rangle} . \quad)^2$$

Hence,

$$\underline{x} = O + x^i \underline{e}_{\langle i \rangle}$$

and

$$x^i = \underline{e}_{\langle i \rangle} \cdot \overrightarrow{O\underline{x}} .$$



This one-to-one correspondence between points $\underline{x}$ in E_n and lists $\underline{\hat{x}} = (x^1, x^2, \dots, x^n)$ in $\mathbb{R}^n$ is called a rectangular Cartesian coordinate system for E_n . It differs from the Cartesian coordinate system of Definition A6.2.4 only in that the basis for $\mathcal{V}_n$ is orthonormal.

Now let us return to our general setting of $\underline{f}: D \rightarrow W$. Taking $\{\underline{e}_{\langle 1 \rangle}, \underline{e}_{\langle 2 \rangle}, \dots, \underline{e}_{\langle m \rangle}\}$ to be an orthonormal

¹See Definition A6.2.3.

²It would be appropriate to write $x^{\langle i \rangle}$ instead of x^i , but this is usually not done for coordinates even in the rectangular case.

basis for the range space $\mathcal{W}$, we can write

$$f(x) = f(0 + x^i \underline{e}_{\langle i \rangle}) = \sum_{i=1}^m \hat{f}^{\langle i \rangle}(x^1, x^2, \dots, x^n) \underline{e}_{\langle i \rangle}$$

just as we did in § A5.14. Now we are in a position to deal with the differentiation of fields in terms of the ordinary partial derivatives of component fields, and all of the results developed in § A5.14 are available to us. Note, however, that the dimensions m and n are reversed from the usage there.

Next we consider individually the kinds of fields that will be particularly important to us in continuum mechanics. Accordingly, for the remainder of this chapter, $\mathcal{D}$ is an open set in a three-dimensional Euclidean point space $\mathcal{E}_3$, the translation space associated with $\mathcal{E}_3$ is the three-dimensional vector space $\mathcal{V}_3$, $\{\underline{e}_{\langle i \rangle}\}$ is an orthonormal basis for $\mathcal{V}_3$, $\mathcal{O}$ is a point in $\mathcal{E}_3$ so that $\{\mathcal{O}; \underline{e}_{\langle i \rangle}\}$ is a reference frame for $\mathcal{E}_3$, and x^i are the coordinates of $x \in \mathcal{E}_3$ w.r.t. $\{\mathcal{O}; \underline{e}_{\langle i \rangle}\}$. Often, we shall denote the list $(x^1, x^2, x^3) \in \mathbb{R}^3$ as $\hat{x}$.

Supplementary Reading
~~nothing more~~

Same as for § A6.3

A6.5. Scalar Fields

Refer to the agreements set out at the end of §A6.4.
 Consider a scalar field $f: D \rightarrow \mathbb{R}$. By Definition A6.4.1,
 f is differentiable at the point $\underline{x} \in D$ if $\exists$ a linear transformation $Df(\underline{x}) \in \text{Lin}(\mathcal{V}_3, \mathbb{R})$ $\exists$ for sufficiently small $|\underline{u}|$ ($\underline{u} \in \mathcal{V}_3$)

$$f(\underline{x} + \underline{u}) - f(\underline{x}) = Df(\underline{x})[\underline{u}] + o(\underline{u}) \text{ as } \underline{u} \rightarrow \underline{0}.$$

Since here the derivative $Df(\underline{x})$ is a linear functional on $\mathcal{V}_3$,
 $\exists$ a unique vector $\nabla f(\underline{x}) \in \mathcal{V}_3$, called the gradient of f at $\underline{x}$, $\exists$

$$Df(\underline{x})[\underline{u}] = \nabla f(\underline{x}) \cdot \underline{u} .)^1$$

In terms of coordinates, for $\underline{x} \in D$ we write

$$f(\underline{x}) = \hat{f}(\hat{\underline{x}}) = \hat{f}(x^1, x^2, x^3);$$

of course, $\hat{\underline{x}} = (x^1, x^2, x^3) \in \hat{D} \subset \mathbb{R}^3$, where $\hat{D}$ is the open subset of $\mathbb{R}^3$ that the rectangular Cartesian coordinate system associates with $D \subset \mathcal{E}_3$. By Theorem A5.14.4, if f is differentiable at $\underline{x}$, then $\hat{f}$ has ordinary first partial derivatives at $\hat{\underline{x}}$; in fact,

$$\hat{f}_{,i}(\hat{\underline{x}}) = \underline{e}_{\langle i \rangle} \cdot \nabla f(\underline{x})$$

and

$$\nabla f(\underline{x}) = \sum_{i=1}^3 \hat{f}_{,i}(\hat{\underline{x}}) \underline{e}_{\langle i \rangle} .$$

¹ cf. Theorem A5.4.1.

By Theorem A5.14.2, $f \in CP(D, \mathbb{R})$ iff $\hat{f}$ is
continuous and has continuous ordinary partial derivatives
on $\hat{D}$ of all orders up to and including p .

It is worth noting that in no way have we taken advantage of the fact that the dimension of the underlying point space is 3. If the Cartesian coordinate system employed were not rectangular, some minor adjustments would have to be made.

~~Supplementary Reading~~

Same as for §A6.3

A6.6. Vector Fields

Refer to the agreements set out at the end of § A6.4.
 Consider a vector field $\underline{v}: D \rightarrow \mathcal{V}_3$. By Definition A6.4.1,
 $\underline{v}$ is differentiable at the point $\underline{x} \in D$ if $\exists$ a linear transformation $D\underline{v}(\underline{x}) \in \text{Lin } \mathcal{V}_3$ $\exists$ for sufficiently small $|\underline{u}|$ ($\underline{u} \in \mathcal{V}_3$)

$$\underline{v}(\underline{x} + \underline{u}) - \underline{v}(\underline{x}) = D\underline{v}(\underline{x})[\underline{u}] + o(\underline{u}) \text{ as } \underline{u} \rightarrow \underline{0}.$$

Of course, $D\underline{v}(\underline{x})$ is a second-order tensor. In the continuum mechanics literature, one often sees $D\underline{v}(\underline{x})$ written as $\nabla \underline{v}(\underline{x})$ and referred to as the gradient of the vector field $\underline{v}$ at the point $\underline{x}$.

In terms of coordinates and components, for $\underline{x} \in D$ we write

$$\underline{v}(\underline{x}) = v^{<i>}(\underline{x}) \underline{e}_{<i>} = \hat{v}^{<i>}(\hat{\underline{x}}) \underline{e}_{<i>} = \hat{v}^{<i>}(x^1, x^2, x^3) \underline{e}_{<i>} ;)^1$$

with $\hat{\underline{x}} = (x^1, x^2, x^3) \in \hat{D} \subset \mathbb{R}^3$, where $\hat{D}$ is the open subset of $\mathbb{R}^3$ that the rectangular Cartesian coordinate system associates with $D \subset \mathcal{E}_3$. By Theorem A5.14.1, if $\underline{v}$ is differentiable at $\underline{x}$, then the component functions $\hat{v}^{<i>}$ have ordinary first partial derivatives at $\hat{\underline{x}}$; in fact,

$$\hat{v}^{<i>}_{,j}(\hat{\underline{x}}) = \underline{e}_{<i>} \cdot D\underline{v}(\underline{x})[\underline{e}_{<j>}].$$

¹Here, the summation convention is in effect.

and

$$D_{\underline{v}}(\underline{x}) = \sum_{i=1}^3 \sum_{j=1}^3 \hat{v}_{\langle i \rangle, j}(\underline{x}) \underline{e}_{\langle i \rangle} \otimes \underline{e}_{\langle j \rangle}$$

By Theorem A5.14.2, $\underline{v} \in C^p(D, V_3)$ iff each component function $\hat{v}_{\langle i \rangle}$ is continuous and has continuous ordinary partial derivatives on $\hat{D}$ of all orders up to and including p .

In the foregoing, the fact that the dimension of the underlying space is 3 has not been utilized; however, if the Cartesian coordinate system were not rectangular, there would have to be some minor adjustments.

The following easy result will be used repeatedly.

Theorem A6.6.1. Let the vector field $\underline{v}: D \rightarrow V_3$ be differentiable on D and have uniform derivative on D , say $\overset{\circ}{\underline{v}}$. Then if D is polygonally connected,

$$\underline{v}(\underline{x}) = \underline{v}(\underline{y}) + \overset{\circ}{\underline{v}}(\underline{x} - \underline{y}) \quad \forall \underline{x}, \underline{y} \in D.$$

Proof. Consider the function $\hat{\underline{v}}: D \rightarrow V_3$ defined by

$$\hat{\underline{v}}(\underline{x}) = \underline{v}(\underline{x}) - \overset{\circ}{\underline{v}}(\underline{x} - \underline{y})$$

for $\underline{y} \in D$ arbitrarily fixed in D . By Theorems A5.5.1 and A5.5.4(i), $\hat{\underline{v}}$ is differentiable on D and

$$D\hat{V} = 0 \text{ on } D.$$

Then by Theorem A5.9.3, $\hat{V}$ is uniform on D ; in particular,

$$\hat{V}(x) = \hat{V}(y)$$

or

$$V(x) - \overset{\circ}{V}(x-y) = V(y). \quad \square$$

Two important auxiliary fields associated with a differentiable vector field are given in the following definition.

Definition A6.6.1. Let the vector field $\underline{v}: D \rightarrow \mathcal{V}_3$ be
differentiable at $\underline{x}$. Then the divergence of $\underline{v}$ at $\underline{x}$ is

$$\operatorname{div} \underline{v}(\underline{x}) = \operatorname{tr} D\underline{v}(\underline{x}) ; \quad)^{0,1}$$

and when $\mathcal{V}_3$ is oriented,² the curl of $\underline{v}$ at $\underline{x}$ is

$$\operatorname{curl} \underline{v}(\underline{x}) = 2 \operatorname{dual} \operatorname{skw} D\underline{v}(\underline{x}). \quad)^{3,4}$$

Common alternative notations are

$$\operatorname{div} \underline{v} = \nabla \cdot \underline{v} \quad \text{and} \quad \operatorname{curl} \underline{v} = \nabla \times \underline{v}.$$

Of course, the divergence and curl can be worked out in terms of coordinates and components.

Theorem A6.6.2. Let the vector field $\underline{v}: D \rightarrow \mathcal{V}_3$ be
differentiable at $\underline{x}$. Then

$$\operatorname{div} \underline{v}(\underline{x}) = \hat{\nabla}^{(i)},_i(\hat{\underline{v}}) \quad (\text{summation implied}) ;$$

and when $\mathcal{V}_3$ is oriented and the basis $\{\underline{e}^{(i)}\}$ is positive,⁵

⁰ See Definition A3.7.1. ¹ The divergence is defined the same way

² See Definition A2.7.4. when $\underline{v}: D \subset \mathcal{E}_n \rightarrow \mathcal{V}_n$.

³ See Theorem A3.6.9 and Definition A3.6.3.

⁵ See Definition A2.7.6. ⁴ The curl is a strictly 3-dimensional notion.

A6.6.5

$$\text{curl } \underline{v}(\underline{x}) = \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \epsilon_{ijk} \hat{v}^{(k)}_{,j}(\underline{x}) \underline{e}_{(i)}$$

Proof. Exercise A6.6.1. $\square$

Supplementary Reading

Same as for § A6.3

A6.7. Second-Order Tensor Fields

Refer to the agreements set out at the end of §A6.4. Consider a second-order tensor field $\underline{T}: \mathcal{D} \rightarrow \text{Lin } \mathcal{V}_3$. By Definition A6.4.1, $\underline{T}$ is differentiable at the point $\underline{x} \in \mathcal{D}$ if $\exists$ a linear transformation $D\underline{T}(\underline{x}) \in \text{Lin}(\mathcal{V}_3, \text{Lin } \mathcal{V}_3) \ni$ for sufficiently small $|\underline{u}|$ ($\underline{u} \in \mathcal{V}_3$)

$$\underline{T}(\underline{x} + \underline{u}) - \underline{T}(\underline{x}) = D\underline{T}(\underline{x})[\underline{u}] + o(\underline{u}) \text{ as } \underline{u} \rightarrow \underline{0}.$$

Of course, $D\underline{T}(\underline{x})$ is a third-order tensor. In the continuum mechanics literature, one often sees $D\underline{T}(\underline{x})$ written as $\nabla \underline{T}(\underline{x})$ and referred to as the gradient of the tensor field $\underline{T}$ at the point $\underline{x}$.

In terms of coordinates and components, for $\underline{x} \in \mathcal{D}$ we write

$$\underline{T}(\underline{x}) = T^{<ij>}(\underline{x}) \underline{e}_{<i>} \otimes \underline{e}_{<j>} = \hat{T}^{<ij>}(\hat{\underline{x}}) \underline{e}_{<i>} \otimes \underline{e}_{<j>} = \hat{T}^{<ij>}(x^1, x^2, x^3) \underline{e}_{<i>} \otimes \underline{e}_{<j>},$$

with $\hat{\underline{x}} = (x^1, x^2, x^3) \in \hat{\mathcal{D}} \subset \mathbb{R}^3$, where $\hat{\mathcal{D}}$ is the open subset of $\mathbb{R}^3$ that the rectangular Cartesian coordinate system associates with $\mathcal{D} \subset \mathcal{E}_3$. By Theorem A5.14.1, if $\underline{T}$ is differentiable at $\underline{x}$, then the component functions $\hat{T}^{<ij>}$ have ordinary first partial derivatives at $\hat{\underline{x}}$; in fact,

$$\hat{T}^{<ij>}_{,k}(\hat{\underline{x}}) = (\underline{e}_{<i>} \otimes \underline{e}_{<j>}) \cdot D\underline{T}(\underline{x})[\underline{e}_{<k>}]$$

$$= \underline{e}_{<i>} \cdot (D\underline{T}(\underline{x})[\underline{e}_{<k>}] \underline{e}_{<j>}) \quad (\text{Theorem A3.7.4(iii)})$$

and

$$D\hat{T}(x) = \sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 \hat{T}^{<i,j>}_{,k}(\hat{x}) (e_{<i>} \otimes e_{<j>}) \otimes e_{<k>}$$

By Theorem A5.14.2, $\hat{T} \in CP(\hat{D}, \text{Lin } V_3)$ iff each component function $\hat{T}^{<i,j>}$ is continuous and has continuous ordinary partial derivatives on $\hat{D}$ of all orders up to and including p .

The fact that the dimension of the underlying space is 3 has not played an essential role in the above summary; however, if the Cartesian coordinate system were not rectangular, some minor adjustments would have to be made.

The only auxiliary field that we shall need to associate with a differentiable tensor field is its divergence. This notion is developed in the next theorem. The reason for this peculiar approach will become evident later when we extend the familiar divergence theorem for vector fields to tensor fields.

(second-order)

Theorem A6.7.1. Let the tensor field $\underline{T}: D \rightarrow \text{Lin } \mathcal{V}_3$ be
differentiable at $\underline{x}$. Then for arbitrary uniform $\underline{a} \in \mathcal{V}_3$,
define the vector field $\underline{v}_a: D \rightarrow \mathcal{V}_3$ by

$$\underline{v}_a(\underline{x}) = (\underline{T}(\underline{x}))^T \underline{a}.$$

(independent of $\underline{a}$)

Then $\underline{v}_a$ is differentiable at $\underline{x}$ and $\exists$ a unique vector called
the divergence of $\underline{T}$ at $\underline{x}$ and denoted¹ by $\text{div } \underline{T}(\underline{x}) \ni$

$$\text{div } \underline{v}_a(\underline{x}) = (\text{div } \underline{T}(\underline{x})) \cdot \underline{a}. \quad)^2$$

Proof. Since $\underline{T}$ is differentiable at $\underline{x}$,

$$\begin{aligned} \underline{v}_a(\underline{x} + \underline{u}) - \underline{v}_a(\underline{x}) &= (\underline{T}(\underline{x} + \underline{u}))^T \underline{a} - (\underline{T}(\underline{x}))^T \underline{a} \quad)^3 \\ &= (\underline{T}(\underline{x} + \underline{u}) - \underline{T}(\underline{x}))^T \underline{a} \\ &= (D\underline{T}(\underline{x})[\underline{u}] + o(\underline{u}))^T \underline{a} \\ &= (D\underline{T}(\underline{x})[\underline{u}])^T \underline{a} + o(\underline{u}) \quad \text{as } \underline{u} \rightarrow \underline{0}. \end{aligned}$$

¹ Often $\text{div } \underline{T}$ is denoted by $\nabla \cdot \underline{T}$.

² The definition of $\text{div } \underline{T}$ is unchanged when $\underline{T}: D \subset \mathbb{E}_n \rightarrow \text{Lin } \mathcal{V}_n$.

³ Of course, one can use the theorems of Chapter 5 to work out $D\underline{v}_a(\underline{x})$. Do this as Exercise A6.7.1.

Since $DI_{\sim}(x) \in \text{Lin}(V_3, \text{Lin } V_3)$, the mapping

$$\underline{u} \mapsto (DI_{\sim}(x)[\underline{u}])^T \underline{a}$$

is linear. Thus, γ_a is differentiable at x and

$$DV_{\sim a}(x)[\underline{u}] = (DI_{\sim}(x)[\underline{u}])^T \underline{a}.$$

Now, as was observed above, in terms of coordinates and components,

$$DI_{\sim}(x)[\underline{u}] = \sum_i \sum_j \sum_k \hat{T}^{<ij>}_{jk}(\hat{x}) \underbrace{[(\underline{e}_{\sim<i>} \otimes \underline{e}_{\sim<j>}) \otimes \underline{e}_{\sim<k>}]}_{(e_{\sim<k>} \cdot \underline{u})(\underline{e}_{\sim<i>} \otimes \underline{e}_{\sim<j>})} \underline{u}$$

$\therefore$ by Theorem A3.6.3, (Exercise A5.10.1) $(e_{\sim<k>} \cdot \underline{u})(\underline{e}_{\sim<i>} \otimes \underline{e}_{\sim<j>})$,

$$(DI_{\sim}(x)[\underline{u}])^T = \sum_i \sum_j \sum_k \hat{T}^{<ij>}_{jk}(\hat{x}) (e_{\sim<k>} \cdot \underline{u})(\underline{e}_{\sim<j>} \otimes \underline{e}_{\sim<i>});$$

so that

$$\begin{aligned} (DI_{\sim}(x)[\underline{u}])^T \underline{a} &= \sum_i \sum_j \sum_k \hat{T}^{<ij>}_{jk}(\hat{x}) (e_{\sim<k>} \cdot \underline{u}) \underbrace{(\underline{e}_{\sim<j>} \otimes \underline{e}_{\sim<i>}) \underline{a}}_{(\underline{e}_{\sim<i>} \cdot \underline{a}) \underline{e}_{\sim<j>}} \\ &\quad \underbrace{(\underline{e}_{\sim<i>} \cdot \underline{a})(\underline{e}_{\sim<k>} \cdot \underline{u})}_{(\underline{e}_{\sim<j>} \otimes \underline{e}_{\sim<k>})} \underline{e}_{\sim<j>} \\ &\quad (\underline{e}_{\sim<j>} \otimes \underline{e}_{\sim<k>}) \underline{u}. \end{aligned}$$

Hence,

$$DV_{\sim a}(x)[\underline{u}] = \sum_i \sum_j \sum_k \hat{T}^{<ij>}_{jk}(\hat{x}) (\underline{e}_{\sim<i>} \cdot \underline{a})(\underline{e}_{\sim<j>} \otimes \underline{e}_{\sim<k>}) \underline{u}$$

or

$$DV_{\sim a}(x) = \sum_i \sum_j \sum_k \hat{T}^{<ij>}_{jk}(\hat{x}) (\underline{e}_{\sim<i>} \cdot \underline{a}) \underline{e}_{\sim<j>} \otimes \underline{e}_{\sim<k>}.$$

Then by Definitions A6.6.1 and A3.7.1,

$$\begin{aligned} \operatorname{div}_{\underline{V}_a}(\underline{T}) &= \sum_i \sum_j \sum_k \hat{T}^{<ij>}{}_{,k}(\underline{x}) (e_{<i>} \cdot \underline{a}) \underbrace{tu(e_{<j>} \otimes e_{<k>})}_{S_{jk}} \\ &= \left(\sum_i \sum_j \sum_k \hat{T}^{<ij>}{}_{,j}(\underline{x}) e_{<i>} \right) \cdot \underline{a}. \end{aligned}$$

Thus, the vector

$$\operatorname{div} \underline{T}(\underline{x}) := \sum_i \sum_j \sum_k \hat{T}^{<ij>}{}_{,j}(\underline{x}) e_{<i>}$$

has the property that for arbitrary $\underline{a}$

$$\operatorname{div}_{\underline{V}_a}(\underline{T}) = (\operatorname{div} \underline{T}(\underline{x})) \cdot \underline{a}.$$

There can be no other such vector, because if $\underline{t}$ were one, we would have

$$(\operatorname{div} \underline{T}(\underline{x})) \cdot \underline{a} = \underline{t} \cdot \underline{a} \quad \forall \underline{a} \in \underline{V}_3$$

which implies¹ $\operatorname{div} \underline{T}(\underline{x}) = \underline{t}$. $\square$

The above proof also supplies

Theorem A6.7.2. Let the (second-order) tensor field $\underline{T}: D \rightarrow \operatorname{Lin} \underline{V}_3$ be differentiable at $\underline{x}$. Then

$$\operatorname{div} \underline{T}(\underline{x}) = \hat{T}^{<ij>}{}_{,j}(\underline{x}) e_{<i>} \text{ (summation implied).}$$

¹ Cf. Theorem A2.3.4.

AG.7.6

Supplementary Reading

Same as for §A6.3

A6.8. Point Fields

Refer to the agreements set out at the end of § A6.4.
Consider a point field $p: D \rightarrow E_3$. Since E_3 is not a linear space, strictly speaking, Definition A6.4.1 does not apply to this situation. On the other hand, since the translation space is a linear space, Definition A6.4.1 does make sense in the present context. Formally, we make

Definition A6.8.1. The point field $p: D \rightarrow E_3$ is differentiable at the point $x \in D$ if $\exists$ a linear transformation

$$Dp(x): \mathcal{V}_3 \rightarrow \mathcal{V}_3 \quad (\mathcal{V}_3 \text{ is the translation space of } \mathcal{V}_3)$$

$\exists$ for sufficiently small $|u|$ ($u \in \mathcal{V}_3$)

$$p(x+u) - p(x) = Dp(x)[u] + o(u) \text{ as } u \rightarrow 0.$$

The linear transformation $Dp(x)$ is called the derivative¹ of p at x .²

This definition makes sense because the point difference

¹ In continuum mechanics, often called the gradient of p at x and written as $\nabla p(x)$.

² With obvious minor adjustments, we use the same definition when the domain space is E_m and the range space is E_n .

$p(x+u) - p(x)$ is a vector in the translation space V_x . Thus, the values, $Dp(x)[u]$, of the linear approximating function $Dp(x)$ should, indeed, be vectors; and the statement

$$p(x+u) - p(x) = Dp(x)[u] + o(u) \text{ as } u \rightarrow 0$$

has the usual interpretation:

$$(p(x+u) - p(x)) - Dp(x)[u] = o(u) \text{ as } u \rightarrow 0.$$

Of course, $Dp(x)$ is a second-order tensor.

Once Definition A6.8.1 is in place, the theory of differentiation of point fields is identical to the theory of differentiation of vector fields. We need not repeat it.

Supplementary Reading

Same as for § A6.3

Supplementary Reading
~~information~~

same as for $\S$ A6.3

A6.9. Some Identities

It should be clear that the various differentiation rules derived in Chapter 5 apply equally well to fields. In this section we gather a few needed rules on the differentiation of fields that are not given there. We continue with the agreements set out at the end of §A6.4 and at the beginning of §A5.2.

For the most part, our first two theorems here can be read off from Theorem A5.5.1. However, because of the possible confusion that is attendant to the fact that points are not elements of a linear space while point differences are, it seems prudent to separate them out.

Theorem A6.9.1. Let the point fields $p: D \rightarrow \mathcal{E}_3$ and $q: D \rightarrow \mathcal{E}_3$ be differentiable at point $x \in D$. Then the vector field $v: D \rightarrow \mathcal{V}_3$ defined by

$$v(x) = p(x) - q(x)$$

also is differentiable at x , and

$$Dv(x) = Dp(x) - Dq(x).$$

Proof. Exercise A6.9.1. $\square$

Theorem A6.9.2. Let the point field $p: D \rightarrow E_3$ and
the vector field $v: D \rightarrow E_3$ be differentiable at the point
 $x \in D$. Then the point field $q: D \rightarrow E_3$ defined by

$$q(x) = p(x) + v(x)$$

also is differentiable at x , and

$$Dq(x) = Dp(x) + Dv(x),$$

Proof. Exercise A6.9.2. $\square$

Theorem A 6.9.3. Let $f \in C^1(D, \mathbb{R})$. Then the tensor field
 $\underline{T}: D \rightarrow \text{Lin } V_3$ defined by

$$\underline{T}(x) = f(x) \underline{I}$$

has a continuous divergence on D which is given by

$$\text{div } \underline{T}(x) = \nabla f(x).$$

Proof. In the spirit of the remarks on pp. A 5.14.18-19, we shall give two distinct but equally rigorous demonstrations.

Direct Approach. By Theorem A 5.5.2, the uniform tensor $\underline{I}$ is differentiable on D and has zero derivative. Then by Theorem A 5.6.4,¹ the product $\underline{T} = f \underline{I}$ is differentiable on D , and

$$D\underline{T}(x)[u] = Df(x)[u] \underline{I} = (\nabla f(x) \cdot u) \underline{I}.$$

Next, in accordance with Theorem A 6.7.1, consider

$$\underline{V}_a(x) = (\underline{T}(x))^T \underline{a} = \underline{T}(x) \underline{a}$$

since $\underline{T}(x) = f(x) \underline{I}$. By Theorems A 5.5.2 and A 5.6.6, the

¹See especially the remarks following the theorem.

product $\underline{v}_a$ is differentiable on D , and

$$\begin{aligned} D\underline{v}_a(\underline{x})[\underline{u}] &= D\underline{T}(\underline{x})[\underline{u}]\underline{a} = (\nabla f(\underline{x}) \cdot \underline{u})\underline{a} \\ &= (a \otimes \nabla f(\underline{x}))\underline{u}. \end{aligned}$$

Thus,

$$D\underline{v}_a(\underline{x}) = a \otimes \nabla f(\underline{x});$$

and

$$\operatorname{div} \underline{v}_a(\underline{x}) = \operatorname{tr} D\underline{v}_a(\underline{x}) = \underline{a} \cdot \nabla f(\underline{x}) = (\operatorname{div} \underline{T}(\underline{x})) \cdot \underline{a}.$$

$\therefore$

$$\operatorname{div} \underline{T}(\underline{x}) = \nabla f(\underline{x}).$$

The continuity of $\operatorname{div} \underline{T}$ on D follows from the above formula and the hypothesis that $f \in C^1(D, \mathbb{R})$.

Components - Coordinates Approach. Since $\underline{T}(\underline{x}) = f(\underline{x})\underline{I}$,

$$\hat{T}^{<ij>}(\hat{\underline{x}}) = \hat{f}(\hat{\underline{x}}) \delta_{ij}.$$

Since $f \in C^1(D, \mathbb{R})$, $\hat{f}$ is continuous on $\hat{D}$ and has continuous ordinary first partial derivatives on $\hat{D}$. $\therefore$ the component functions $\hat{T}^{<ij>}$ are continuous on $\hat{D}$ and have continuous ordinary first partial derivatives on $\hat{D}$. Consequently, $\underline{T} \in C^1(D, \operatorname{Lin} \mathbb{V}_3)$. In particular, $\underline{T}$ is differentiable on D , and by Theorem A6.7.2,

$$\operatorname{div} \underline{T}(\underline{x}) = \hat{T}^{<ij>},_{j}(\hat{\underline{x}}) \underline{e}_{<i>} = \sum_i \sum_j \hat{f}_{,j}(\hat{\underline{x}}) \delta_{ij} \underline{e}_{<i>}$$

$$= \sum_i \hat{f}_{,i}(\hat{x}) \underline{e}_{\langle i \rangle} = \nabla f(x). \quad)^1$$

Again, the continuity of $\operatorname{div} \underline{T}$ follows from the continuity of ∇f . $\square$

A generalization of Theorem A6.9.3, which we shall need is

Theorem A6.9.4. Let $f \in C^1(D, \mathbb{R})$ and $\underline{S} \in C^1(D, \operatorname{Lin} V_3)$.
Then the tensor field $\underline{T}: D \rightarrow \operatorname{Lin} V_3$ defined by

$$\underline{T}(x) = f(x) \underline{S}(x)$$

has a continuous divergence on D which is given by

$$\operatorname{div} \underline{T}(x) = f(x) \operatorname{div} \underline{S}(x) + \underline{S}(x) \nabla f(x).$$

Proof. Exercise A6.9.3. $\square$

Theorem A6.9.5. Let $\underline{u}, \underline{v} \in C^1(D, V_3)$. Then the tensor
field $\underline{T}: D \rightarrow \operatorname{Lin} V_3$ defined by

$$\underline{T}(x) = \underline{u}(x) \otimes \underline{v}(x)$$

has a continuous divergence on D which is given by

$$\operatorname{div} (\underline{u} \otimes \underline{v})(x) = D\underline{u}(x) [\underline{v}(x)] + (\operatorname{div} \underline{v}(x)) \underline{u}(x).$$

Proof. Exercise A6.9.4. $\square$

¹Here we use Theorem A5.14.4. cf. the remarks on p. A6.5.1,

Theorem A6.9.6. Let $\underline{T} \in C^1(\mathcal{D}, \text{Lin } \mathcal{V}_3)$ and $\underline{u} \in C^1(\mathcal{D}, \mathcal{V}_3)$.
Then the vector field $\underline{v}: \mathcal{D} \rightarrow \mathcal{V}_3$ defined by

$$\underline{v}(\underline{x}) = \underline{T}(\underline{x}) \underline{u}(\underline{x})$$

has a continuous divergence on $\mathcal{D}$ which is given by

$$\text{div } \underline{v}(\underline{x}) = [\text{div } \underline{T}^T(\underline{x})] \underline{u}(\underline{x}) + \underline{T}^T(\underline{x}) \cdot \underline{D}\underline{u}(\underline{x}).$$

Proof. Exercise A6.9.5. $\square$

An easier result of the same sort is

Theorem A6.9.7. Let $f \in C^1(\mathcal{D}, \mathbb{R})$ and $\underline{u} \in C^1(\mathcal{D}, \mathcal{V}_3)$.
Then the vector field $\underline{v}: \mathcal{D} \rightarrow \mathcal{V}_3$ defined by

$$\underline{v}(\underline{x}) = f(\underline{x}) \underline{u}(\underline{x})$$

has a continuous divergence on $\mathcal{D}$ which is given by

$$\text{div } \underline{v}(\underline{x}) = [\nabla f(\underline{x})] \cdot \underline{u}(\underline{x}) + f(\underline{x}) \text{div } \underline{u}(\underline{x}).$$

Proof. Exercise A6.9.6. $\square$

A6.10. Some Integral Theorems

While our treatment of differentiation was not comprehensive, it did set the foundation of the subject. For the more difficult topic of integration, we can not take the time to do this; and we assume that the student is already familiar with the integration of scalar and vector fields over arcs, surfaces, and regions in E_3 . Here we mainly collect together some useful results in our notation.

Definition A6.10.1 A set $R \subset E_3$ is a bounded regular region iff it is a bounded region¹ whose boundary ∂R is the union of a finite number of nonintersecting closed regular surfaces, this latter term being used in the sense of Kellogg.²

It is important to realize that the boundary of a regular region may have edges and corners where surface pieces fit together. Regular regions are general enough to cover the needs of continuum mechanics, but special enough that the following result obtains.

¹We use region as a synonym for pathwise connected point set.

²See KELLOGG's Foundations of Potential Theory, p. 112 or GURTIN's The Linear Theory of Elasticity, p. 13.

Theorem A6.10.1. (Divergence Theorem) Let the vector and tensor fields $\underline{u}$ and $\underline{T}$ be smooth on $\overline{Q}$,¹ where Q is a bounded regular region. Then if $\underline{n}$ denotes the outward unit normal to ∂Q ,

$$\int_Q \operatorname{div} \underline{u} \, dV = \int_{\partial Q} \underline{u} \cdot \underline{n} \, dA$$

and

$$\int_Q \operatorname{div} \underline{T} \, dV = \int_{\partial Q} \underline{T} \underline{n} \, dA.$$

Proof. The first assertion is the classical divergence theorem, whose proof is available in KELLOGG's book.

To establish the second assertion, we let $\underline{a}$ be an arbitrary uniform vector and use the definition² of $\operatorname{div} \underline{T}$ to get

$$\underline{a} \cdot \int_Q \operatorname{div} \underline{T} \, dV = \int_Q (\operatorname{div} \underline{T}) \cdot \underline{a} \, dV = \int_Q \operatorname{div} (\underline{T}^T \underline{a}) \, dV.$$

Now we can apply the classical divergence theorem to the smooth

(Continuous on $\overline{Q}$ and

¹Smooth on the closure $\overline{Q}$ means smooth on the interior $\overset{\circ}{Q}$ (i.e., differentiable with continuous derivative) with the existence of a continuous extension of the derivative from $\overset{\circ}{Q}$ to $\overline{Q}$.

²Theorem A6.7.1.

vector field $\underline{T}\underline{a}$ to get

$$\underline{a} \cdot \int_Q \operatorname{div} \underline{T} \, dV = \int_{\partial Q} (\underline{T}\underline{a}) \cdot \underline{n} \, dA = \int_{\partial Q} \underline{a} \cdot \underline{T}\underline{n} \, dA = \underline{a} \cdot \int_{\partial Q} \underline{T}\underline{n} \, dA.$$

Since $\underline{a}$ is arbitrary, this implies¹ the desired result. $\square$

Theorem A6.10.2. (Change of Variable Theorem for Volume Integrals) Let $\underline{f}: D \rightarrow E_3$ and $g: \underline{f}(D) \rightarrow \mathbb{R}$, where D is an open region in E_3 . Suppose that $\underline{f}$ is one-to-one on D , that $\underline{f}$ is smooth on D , and that g is continuous on $\underline{f}(D)$. Then for every closed bounded regular region $Q \subset D$

$$\begin{aligned} \int_{\underline{f}(Q)} g \, dV &= \int_Q g \circ \underline{f} \, |\det D\underline{f}| \, dV \\ &= \int_Q g(\underline{f}(p)) |\det D\underline{f}(p)| \, dV_p. \end{aligned}$$

Proof. See APOSTOL's Mathematical Analysis, pp. 271-274. $\square$

Theorem A6.10.3. (Mean Value Theorem for Volume Integrals) Let $\underline{f}: \bar{Q} \rightarrow \mathbb{R}$, where Q is a bounded regular region in E_3 . Suppose that $\underline{f}$ is continuous on $\bar{Q}$. Then $\exists \xi \in Q \ni$

¹Cf. Theorem A2.3.4.

$$\int_R f dV = f(\bar{x}) \int_R dV.$$

Proof. See APOSTOL's Mathematical Analysis, p. 269. $\square$

A similar mean value theorem holds for surface integrals.

Supplementary Reading

Same as for §A6.3 plus

APOSTOL, Mathematical Analysis

GURTIN, The Linear Theory of Elasticity in Bd. VIa/2 of Handbuch der Physik.

KELLOGG, Foundations of Potential Theory.

A6.11. General Coordinate Systems

Recall from Definition A6.2.3 that a reference frame for an n -dimensional Euclidean point space E_n is a point $O \in E_n$ together with a basis $\{\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n\}$ for the associated translation space V_n . Then the position vector of any point $P \in E_n$ w.r.t. the origin O admits the representation

$$\vec{OP} = p^i \underline{e}_i.$$

This sets up one-to-one correspondence between points in E_n and lists in $\mathbb{R}^n$:

$$P \mapsto (p^1, p^2, \dots, p^n) = (\vec{OP} \cdot \underline{e}^1, \vec{OP} \cdot \underline{e}^2, \dots, \vec{OP} \cdot \underline{e}^n)$$

and

$$(p^1, p^2, \dots, p^n) \mapsto P = O + p^i \underline{e}_i,$$

which we called a Cartesian coordinate system for E_n (Definition A6.2.4). Here, we generalize this idea by assigning in a natural way a different basis to each point of E_n . As the first step we make

Definition A6.11.1. A coordinate system¹ for an n -dimensional Euclidean point space E_n is a one-to-one correspondence between points $P \in E_n$ and lists $(x^1, x^2, \dots, x^n)$ in $\mathbb{R}^n$ which is two times continuously differentiable in each direction.

¹Often, general coordinate system or curvilinear coordinate system.

More precisely, a coordinate system for E_n consists of a point-valued function on an n -dimensional interval $\mathcal{Q} = \mathcal{Q}_1 \times \mathcal{Q}_2 \times \dots \times \mathcal{Q}_n$

$$\hat{p} \in C^2(\mathcal{Q}, E_n) \quad [p = \hat{p}(x^1, x^2, \dots, x^n)]$$

together with its inverse

$$\hat{x}^i \in C^2(E_n, \mathcal{Q}_i) \quad [x^i = \hat{x}^i(p)], \quad i=1, 2, \dots, n,$$

which is a system of n scalar fields on E_n . The numbers $x^1, x^2, \dots, x^n$ are the coordinates of the point p relative to the coordinate system.

In practice, we permit the one-to-one correspondence to break down on certain subsets of E_n — the set of singular points of the coordinate system.

For the remainder of this section it is understood that a particular coordinate system has been given, i.e., the functions $\hat{p}$ and $\hat{x}^i$ ($i=1, 2, \dots, n$) have been prescribed.

Theorem A6.11.1. At a nonsingular point $p \in E_n$, the vectors

$$e_i(p) = \hat{p}_{,i}(\hat{x}^1(p), \hat{x}^2(p), \dots, \hat{x}^n(p)), \quad i=1, 2, \dots, n$$

form a basis for $\mathcal{V}_n$ — the natural basis at point p of the coordinate system. The reference frame

$\{\underline{p}; \underline{e}_1(\underline{p}), \underline{e}_2(\underline{p}), \dots, \underline{e}_n(\underline{p})\}$ is called the natural frame at $\underline{p}$ of the coordinate system.

Proof. In view of Theorem A2.2.8, it suffices to show that the set $\{\underline{e}_1(\underline{p}), \underline{e}_2(\underline{p}), \dots, \underline{e}_n(\underline{p})\}$ is linearly independent. Accordingly, consider the statement

$$c^j \underline{e}_j(\underline{p}) = \underline{0}.$$

This $\Rightarrow$

$$\begin{aligned} 0 &= \nabla \hat{x}^i(\underline{p}) \cdot \underline{e}_j(\underline{p}) c^j \quad (\underline{0} \cdot \underline{a} = 0) \\ &= \nabla \hat{x}^i(\underline{p}) \cdot \hat{p}_{,j}(\hat{x}^1(\underline{p}), \hat{x}^2(\underline{p}), \dots, \hat{x}^n(\underline{p})) c^j \quad (\text{substitution}) \\ &= \nabla \hat{x}^i(\hat{p}(x^1, x^2, \dots, x^n)) \cdot \hat{p}_{,j}(x^1, x^2, \dots, x^n) c^j \quad (\text{substitution}). \end{aligned}$$

Now, the composition

$$x^i = \hat{x}^i(\hat{p}(x^1, x^2, \dots, x^n)), \quad i=1, 2, \dots, n$$

defines the identity map on $\mathbb{R}^n$, and $\therefore$ by the chain rule¹,

$$\delta_j^i = \nabla \hat{x}^i(\hat{p}(x^1, x^2, \dots, x^n)) \cdot \hat{p}_{,j}(x^1, x^2, \dots, x^n) \quad (*)$$

Hence,

$$c^j \underline{e}_j(\underline{p}) = \underline{0}$$

¹ The details of the calculation are left as Exercise A6.11.1.

⇒

$$0 = \delta_j^i c_j = c^i;$$

and we conclude from Theorem A2.2.2 that the set $\{\underline{e}_1(p), \underline{e}_2(p), \dots, \underline{e}_n(p)\}$ is linearly independent. $\square$

Equation (*) in the above proof states that

$$\nabla \hat{x}^i(p) \cdot \underline{e}_j(p) = \delta_j^i, \quad i, j = 1, 2, \dots, n,$$

which by Theorem A2.5.5 characterizes the set $\{\nabla \hat{x}^1(p), \nabla \hat{x}^2(p), \dots, \nabla \hat{x}^n(p)\}$ as the dual basis of $\{\underline{e}_1(p), \underline{e}_2(p), \dots, \underline{e}_n(p)\}$. Thus, we have

Theorem A 6.11.2. At a nonsingular point $p \in E_n$, the vectors

$$\underline{e}^i(p) = \nabla \hat{x}^i(p), \quad i = 1, 2, \dots, n$$

comprise the dual basis of the natural basis at p .

Now that we have the natural basis $\{\underline{e}_1(p), \underline{e}_2(p), \dots, \underline{e}_n(p)\}$ and its dual $\{\underline{e}^1(p), \underline{e}^2(p), \dots, \underline{e}^n(p)\}$ at point p for a given coordinate system for E_n , essentially everything that we did before regarding component representations of vectors and tensors carries over to this more general scheme. Of course, now we must be cognizant of the fact that these bases depend on position. E.g., from Definition 2.5.2 and Theorems

A2.5.8 and A3.2.2, we have

Theorem A6.11.3. Let $\underline{u}$ and $\underline{T}$ be a vector field and a second-order tensor field on E_n . Then $\underline{u}$ admits the representations

$$\underline{u}(p) = u^i(p) \underline{e}_i(p) = u_i(p) \underline{e}^i(p),$$

where

$$u^i(p) = \underline{e}^i(p) \cdot \underline{u}(p) \text{ and } u_i(p) = \underline{e}_i(p) \cdot \underline{u}(p)$$

are the contravariant and covariant components of the vector field $\underline{u}$ at the point p relative to the given coordinate system; and $\underline{T}$ admits the representations

$$\underline{T}(p) = T^{ij}(p) \underline{e}_i(p) \otimes \underline{e}_j(p) = T_{ij}(p) \underline{e}^i(p) \otimes \underline{e}^j(p)$$

$$= T^{i \cdot j}(p) \underline{e}^i(p) \otimes \underline{e}_j(p) = T^i_{\cdot j}(p) \underline{e}_i(p) \otimes \underline{e}^j(p),$$

where

$$T^{ij}(p) = \underline{e}^i(p) \cdot \underline{T}(p) \underline{e}^j(p), \quad T_{ij}(p) = \underline{e}_i(p) \cdot \underline{T}(p) \underline{e}_j(p),$$

and

$$T^{i \cdot j}(p) = \underline{e}_i(p) \cdot \underline{T}(p) \underline{e}^j(p), \quad T^i_{\cdot j}(p) = \underline{e}^i(p) \cdot \underline{T}(p) \underline{e}_j(p)$$

are the contravariant, covariant, and mixed components of the tensor field $\underline{T}$ at the point p relative to the given coordinate system.

The g symbols or components of the identity tensor ¹ play the same role in this more general scheme as they did before ². In the present context, the traditional terminology is given by

Definition A6.11.2. The components of the metric tensor of the given coordinate system are

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

$$g_i^{\cdot j}(p) = \underline{e}_i(p) \cdot \underline{e}^j(p) \quad , \quad g^{\cdot i}_j(p) = \underline{e}^i(p) \cdot \underline{e}_j(p)$$

at point p .

Of special importance is

Definition A6.11.3. A coordinate system is orthogonal if at each nonsingular point p the vectors of the natural basis are orthogonal; i.e., if

$$g_{ij}(p) = 0 \quad (i \neq j) .$$

Most of the coordinate systems used in applications are orthogonal, but generally the vectors of the natural basis are

¹ See Definition A2.5.3 and Theorem A3.2.7.

² Cf. Theorems A2.5.10, 11, 12 and A3.2.8.

not all of unit length so that the natural basis is not orthonormal)¹.
This leads us to

Definition A6.11.4. Let $\{\underline{e}_1(p), \underline{e}_2(p), \dots, \underline{e}_n(p)\}$ be the natural basis at a nonsingular point p of an orthogonal coordinate system. Then the associated orthonormal basis at p is defined by

$$\underline{e}_{\langle i \rangle}(p) = \frac{\underline{e}_i(p)}{|\underline{e}_i(p)|} = \frac{\underline{e}_i(p)}{\sqrt{g_{ii}(p)}} \quad (\text{no sum}), \quad i=1, 2, \dots, n.$$

Components of tensor fields relative to the associated orthonormal basis are called physical components and are distinguished by the marks $\langle \rangle$ around the indices)².

It is an easy matter to calculate physical components in terms of regular tensor components.

Theorem A6.11.4. Let the coordinate system be orthogonal. Then at a nonsingular point the physical components of a vector field $\underline{v}$ and a second-order tensor field $\underline{T}$ are given by

$$v^{\langle i \rangle} = \sqrt{g_{ii}} v^i = \frac{v_i}{\sqrt{g_{ii}}}$$

¹ Cf. Definition A2.3.8.

² Cf. Definition A2.3.9.

and

$$T^{<\underline{i}\underline{j}>} = \sqrt{g_{\underline{i}\underline{i}} g_{\underline{j}\underline{j}}} T^{\underline{i}\underline{j}} = \frac{T_{\underline{i}\underline{j}}}{\sqrt{g_{\underline{i}\underline{i}} g_{\underline{j}\underline{j}}}},$$

where underlining an index suspends the summation convention on that index.

Proof. Exercise A6.11.2. $\square$

There is considerable confusion in the literature concerning physical components. When one sees statements such as "The cylindrical components of the velocity $\underline{v}$ are (v_r, v_θ, v_z) ," it often is not clear whether these are ordinary covariant components or physical components. We shall have no difficulty here because of our consistent use of the marks $< >$ to denote orthonormal bases and physical components,

AG.11.9

Supplementary Reading

BOWEN and WANG, Introduction to Vectors and Tensors,
Vol. 2, Vector and Tensor Analysis

COLEMAN, MARKOVITZ, and NOLL, Viscometric Flows of
Non-Newtonian Fluids (Appendix on Mathematical Concepts)

LICHNEROWICZ, Tensor Calculus

NOLL, Finite-Dimensional Spaces

THOMAS, Concepts from Tensor Analysis and Differential Geometry

TRUESDELL and NOLL, "The non-linear field theories of mechanics",
Handbuch der Physik, vol III/3 (edit. FLÜGGE) 1965.

A6.12. Some Particular Coordinate Systems

In this section, we consider some of the more commonly used coordinate systems.

Cartesian Coordinates¹

Let $\{O; \underline{e}_1, \underline{e}_2, \dots, \underline{e}_n\}$ be a fixed reference frame for E_n . Then we can express the position vector of any point $P \in E_n$ relative to the origin O by

$$\overrightarrow{OP} = x^i \underline{e}_i,$$

where

$$x^i = \underline{e}^i \cdot \overrightarrow{OP} =: \hat{x}^i(P) \in \mathbb{R}, \quad i=1, 2, \dots, n,$$

with $\{\underline{e}'_1, \underline{e}'_2, \dots, \underline{e}'_n\}$ being the dual basis to $\{\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n\}$.

To get the inverse relationship, we note that

$$P = O + \overrightarrow{OP} = O + x^i \underline{e}_i =: \hat{P}(x^1, x^2, \dots, x^n).$$

This one-to-one correspondence between points $P \in E_n$ and n -tuples $(x^1, x^2, \dots, x^n)$ in $\mathbb{R}^n$ defines a Cartesian coordinate system for E_n . Since the one-to-one correspondence never breaks down, a Cartesian coordinate system has no singular points.

¹G. Definition A6.2.4 and the remarks introducing §A6.11.

The vectors of the natural basis are given by $\rangle'$

$$\underline{e}_i(p) = \hat{p}_{,i}(\hat{x}^1(p), \hat{x}^2(p), \dots, \hat{x}^n(p)) = \underline{e}_i.$$

Thus, for a Cartesian coordinate system, each point has the same natural basis,

A rectangular Cartesian coordinate system is simply a Cartesian coordinate system where the defining basis $\{\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n\}$ is orthonormal. Often the adjective rectangular is suppressed, so confusion is possible.

Circular Cylindrical Coordinates

Let $\{O; \underline{e}_{\langle x \rangle}, \underline{e}_{\langle y \rangle}, \underline{e}_{\langle z \rangle}\}$ be a fixed reference frame for E_3 with the basis $\{\underline{e}_{\langle x \rangle}, \underline{e}_{\langle y \rangle}, \underline{e}_{\langle z \rangle}\}$ for $\forall_3$ orthonormal. Then we can write the position vector of any point $p \in E_3$ relative to the origin O as

$$\overrightarrow{OP} = x \underline{e}_{\langle x \rangle} + y \underline{e}_{\langle y \rangle} + z \underline{e}_{\langle z \rangle},$$

where

$$x = \underline{e}_{\langle x \rangle} \cdot \overrightarrow{OP} =: \hat{x}(p) \in \mathbb{R},$$

$$y = \underline{e}_{\langle y \rangle} \cdot \overrightarrow{OP} =: \hat{y}(p) \in \mathbb{R},$$

¹ Cf. Theorem A6.11.1.

and

$$z = \underline{e}_{\langle z \rangle} \cdot \vec{0p} =: \hat{z}(p) \in \mathbb{R}.$$

Define

$$r = \sqrt{x^2 + y^2} = \sqrt{[\hat{x}(p)]^2 + [\hat{y}(p)]^2} =: \hat{r}(p) \in [0, \infty)$$

and

$$\theta = \tan^{-1} \frac{y}{x} = \tan^{-1} \frac{\hat{y}(p)}{\hat{x}(p)} =: \hat{\theta}(p) \in [0, 2\pi)$$

so that

$$x = r \cos \theta, \quad y = r \sin \theta.$$

Then

$$p = 0 + \vec{0p} = 0 + r \cos \theta \underline{e}_{\langle x \rangle} + r \sin \theta \underline{e}_{\langle y \rangle} + z \underline{e}_{\langle z \rangle} =: \hat{p}(r, \theta, z).$$

one-to-one

This correspondence between points $p \in \mathbb{E}_3$ and 3-tuples $(r, \theta, z) \in [0, \infty) \times [0, 2\pi) \times (-\infty, \infty)$ is called a circular cylindrical coordinate system.

Exercise A6.12.1. Determine the set of singular points for a given circular cylindrical coordinate system.

The vectors of the natural basis are

$$\underline{e}_r(p) = \frac{\partial \hat{p}}{\partial r}(\hat{r}(p), \hat{\theta}(p), \hat{z}(p)) = \cos \theta \underline{e}_{\langle x \rangle} + \sin \theta \underline{e}_{\langle y \rangle},$$

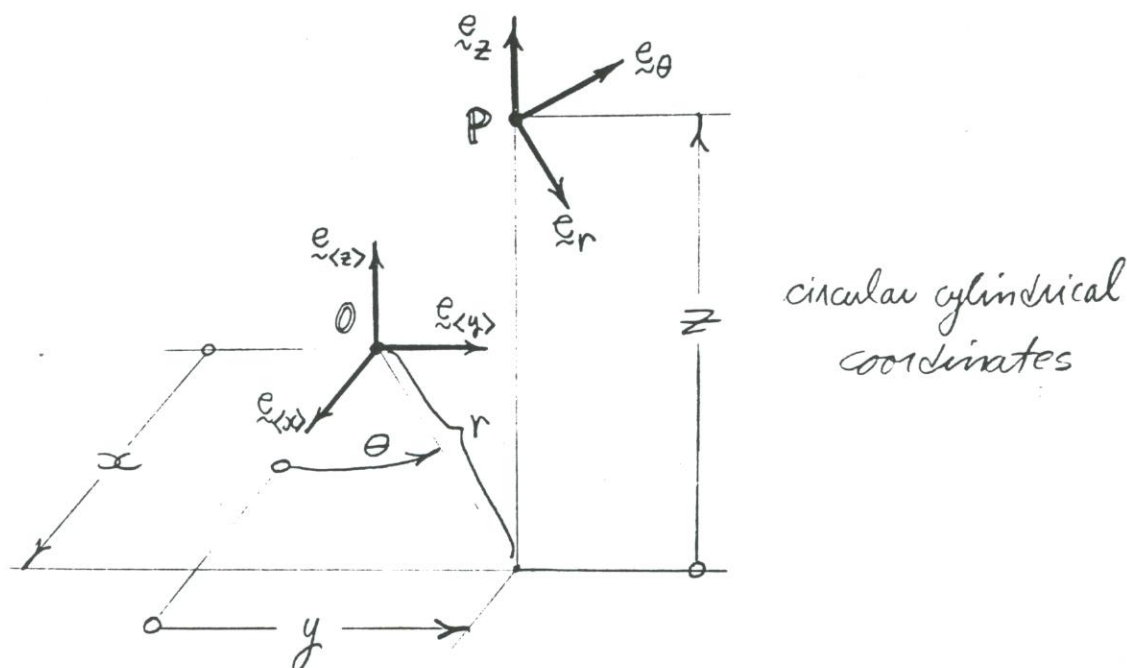
$$\underline{e}_\theta(p) = \frac{\partial \hat{p}}{\partial \theta}(\hat{r}(p), \hat{\theta}(p), \hat{z}(p)) = -r \sin \theta \underline{e}_{\langle x \rangle} + r \cos \theta \underline{e}_{\langle y \rangle},$$

$$\underline{e}_z(p) = \frac{\partial \hat{p}}{\partial z}(\hat{r}(p), \hat{\theta}(p), \hat{z}(p)) = \underline{e}_{\langle z \rangle}.$$

Quite obviously the natural basis is orthogonal so that a circular cylindrical coordinate system is orthogonal; however,

the natural basis is not orthonormal since

$$|\tilde{e}_\theta(p)| = r.$$



The above figure makes it clear that the typical point P is defined by the intersection of three surfaces: a circular cylinder of radius r whose axis is the z -axis; a plane parallel to the xy -plane and located at height z ; a plane perpendicular to the xy -plane which contains the z -axis and makes an angle θ with the xz -plane. Older treatments of curvilinear coordinate systems were often based on such coordinate surfaces.

Spherical Coordinates

Let $\{O; \underline{e}_{\langle x \rangle}, \underline{e}_{\langle y \rangle}, \underline{e}_{\langle z \rangle}\}$ be a fixed reference frame for E_3 with the basis $\{\underline{e}_{\langle x \rangle}, \underline{e}_{\langle y \rangle}, \underline{e}_{\langle z \rangle}\}$ orthonormal. Then we can write the position vector of any point $P \in E_3$ relative to the origin O as

$$\vec{OP} = x \underline{e}_{\langle x \rangle} + y \underline{e}_{\langle y \rangle} + z \underline{e}_{\langle z \rangle},$$

where

$$x = \underline{e}_{\langle x \rangle} \cdot \vec{OP} =: \hat{x}(P) \in \mathbb{R}, \text{ etc.}$$

Define

$$r = \sqrt{x^2 + y^2 + z^2} = \sqrt{[\hat{x}(P)]^2 + [\hat{y}(P)]^2 + [\hat{z}(P)]^2} =: \hat{r}(P) \in [0, \infty),$$

$$\phi = \tan^{-1} \frac{y}{x} = \tan^{-1} \frac{\hat{y}(P)}{\hat{x}(P)} =: \hat{\phi}(P) \in [0, 2\pi),$$

and

$$\theta = \cos^{-1} \frac{z}{\sqrt{x^2 + y^2 + z^2}} = \cos^{-1} \frac{\hat{z}(P)}{\sqrt{[\hat{x}(P)]^2 + [\hat{y}(P)]^2 + [\hat{z}(P)]^2}} =: \hat{\theta}(P) \in [0, \pi]$$

so that

$$x = r \sin \theta \cos \phi,$$

$$y = r \sin \theta \sin \phi,$$

$$z = r \cos \theta.$$

Then

$$P = O + \vec{OP} = O + r \sin \theta \cos \phi \underline{e}_{\langle x \rangle} + r \sin \theta \sin \phi \underline{e}_{\langle y \rangle} + r \cos \theta \underline{e}_{\langle z \rangle} =: \hat{P}(r, \theta, \phi).$$

one-to-one

This correspondence between points $P \in E_3$ and 3-tuples $(r, \theta, \phi) \in$
is called a spherical coordinate system.

$$[0, \infty) \times [0, 2\pi] \times [0, \pi]$$

Exercise A6.12.2. Determine the set of singular points for a given spherical coordinate system.

The vectors of the natural basis are

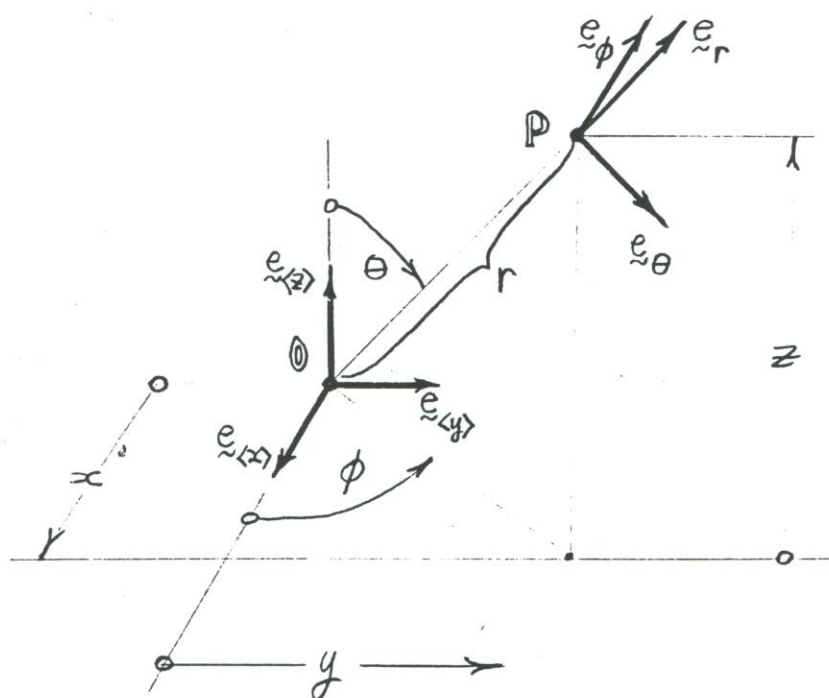
$$\tilde{e}_r(p) = \frac{\partial \hat{P}}{\partial r}(\hat{r}(p), \hat{\theta}(p), \hat{\phi}(p)) = \sin\theta \cos\phi \tilde{e}_{\langle x \rangle} + \sin\theta \sin\phi \tilde{e}_{\langle y \rangle} + \cos\theta \tilde{e}_{\langle z \rangle},$$

$$\tilde{e}_\theta(p) = \frac{\partial \hat{P}}{\partial \theta}(\hat{r}(p), \hat{\theta}(p), \hat{\phi}(p)) = r \cos\theta \cos\phi \tilde{e}_{\langle x \rangle} + r \cos\theta \sin\phi \tilde{e}_{\langle y \rangle} - r \sin\theta \tilde{e}_{\langle z \rangle},$$

$$\tilde{e}_\phi(p) = \frac{\partial \hat{P}}{\partial \phi}(\hat{r}(p), \hat{\theta}(p), \hat{\phi}(p)) = -r \sin\theta \sin\phi \tilde{e}_{\langle x \rangle} + r \sin\theta \cos\phi \tilde{e}_{\langle y \rangle}.$$

Again the natural basis is orthogonal, so that a spherical coordinate system is orthogonal, but it is not orthonormal since

$$|\tilde{e}_r(p)| = r \quad \text{and} \quad |\tilde{e}_\phi(p)| = r \sin\theta.$$



spherical
coordinates

Again the typical point p is at the intersection of three surfaces: a sphere of radius r centered at the origin O ; a cone of half angle Θ with vertex at O whose axis is the z -axis; a plane perpendicular to the xy -plane which contains the z -axis and makes an angle ϕ with the xz -plane.

~~Supplementary Reading~~

Same as for $\S$ A6.11

A6.13. Coordinate Transformations

Here we investigate transformations or changes of coordinate systems. As in § A6.11, we let the one-to-one correspondence

$$p = \hat{p}(x^1, x^2, \dots, x^n),$$

$$x^i = \hat{x}^i(p), \quad i=1, 2, \dots, n$$

between points $p \in E_n$ and n -tuples $(x^1, x^2, \dots, x^n) \in \mathbb{R}^n$ denote a typical coordinate system for E_n . A new coordinate system is defined by a new correspondence, say

$$p = \bar{\hat{p}}(\bar{x}^1, \bar{x}^2, \dots, \bar{x}^n),$$

$$\bar{x}^i = \bar{\hat{x}}^i(p), \quad i=1, 2, \dots, n.$$

In accordance with Definition A6.11.1, the domain of $\bar{\hat{p}}$, which is also the range of $\bar{\hat{x}}^i$ ($i=1, 2, \dots, n$), is an n -dimensional

interval $\bar{I} = \bar{I}_1 \times \bar{I}_2 \times \dots \times \bar{I}_n$. Substitution of $\mathbf{p} = \hat{\mathbf{p}}(x^1, x^2, \dots, x^n)$ into $\bar{x}^i = \hat{x}^i(\mathbf{p})$ yields

$$\bar{x}^i = \hat{x}^i(\hat{\mathbf{p}}(x^1, x^2, \dots, x^n)),$$

which is more conventionally and less precisely written as

$$\bar{x}^i = \hat{x}^i(\hat{\mathbf{p}}(x^1, x^2, \dots, x^n)) =: \bar{x}^i(x^1, x^2, \dots, x^n), \quad i=1, 2, \dots, n. \quad (\text{CT})$$

Similarly, by starting with $x^i = \hat{x}^i(\mathbf{p})$ we are lead to the inverse relations

$$x^i = \hat{x}^i(\hat{\mathbf{p}}(\bar{x}^1, \bar{x}^2, \dots, \bar{x}^n)) =: x^i(\bar{x}^1, \bar{x}^2, \dots, \bar{x}^n), \quad i=1, 2, \dots, n. \quad (\text{ICT})$$

Definition A6.13.1. The one-to-one correspondence between
coordinates x^i and coordinates $\bar{x}^i$ of the same point in E_n
defined by (CT) and (ICT) is a coordinate transformation.

Of course, the development above breaks down at any point which is a singular point for either of the two coordinate systems under consideration. Accordingly, a coordinate transformation breaks down where either set of coordinates corresponds to a singular point of the respective coordinate system. In what follows, this possibility is understood; it is too cumbersome to keep repeating the disclaimer. Note that our introductions of cylindrical and spherical coordinates were in reality coordinate transformations from rectangular Cartesian coordinates.

It is not difficult to see that the set of all coordinate transformations (for a given E_n) is a group with respect to composition. We leave the formulation and proof of this for the reader as Exercise A6.13.1.

By the chain rule, the smoothness built into coordinate systems via Definition A6.11.1 carries over to coordinate transformations. Formally, we have

Theorem A6.13.1. The transformation $(x^1, x^2, \dots, x^n) \mapsto \bar{x}^i$ of $\bar{Q}$ onto $\bar{Q}$ defined by (CT) has continuous first and second partial derivatives ($i=1, 2, \dots, n$). Similarly, the inverse transformation $(\bar{x}^1, \bar{x}^2, \dots, \bar{x}^n) \mapsto x^i$ of $\bar{Q}$ onto Q defined by (ICT) has continuous first and second partial derivatives.

Proof. Exercise A6.13.2. $\square$

The vectors of the natural basis of the new (or barred) coordinate system are, of course, given by $\bar{e}_i$ ¹

$$\bar{e}_i(p) = \bar{\hat{P}}_{,i}(\hat{x}^1(p), \hat{x}^2(p), \dots, \hat{x}^n(p)).$$

We can relate them to the vectors of the natural basis of the original (or unbarred) coordinate system as follows. Write

$$\begin{aligned} \bar{\hat{P}}(\bar{x}^1, \bar{x}^2, \dots, \bar{x}^n) &= p = \hat{P}(x^1, x^2, \dots, x^n) \\ &= \hat{P}(x^1(\bar{x}^1, \bar{x}^2, \dots, \bar{x}^n), \dots, x^n(\bar{x}^1, \bar{x}^2, \dots, \bar{x}^n)). \end{aligned}$$

¹cf. Theorem A6.11.1.

Then by the appropriate variant of the chain rule,

$$\bar{p}_{,i}(\bar{x}', \dots, \bar{x}^n) = \hat{p}_{,j}(\hat{x}'(\bar{x}', \dots, \bar{x}^n), \dots, \hat{x}^n(\bar{x}', \dots, \bar{x}^n)) x^j_{,i}(\bar{x}', \dots, \bar{x}^n).$$

Writing the x 's and the $\bar{x}$'s in terms of p , we get

$$\bar{e}_{,i}(p) = x^j_{,i}(\hat{x}'(p), \dots, \hat{x}^n(p)) \underline{e}_j(p).$$

It follows from Theorem A6.11.3 and the uniqueness of components w.r.t. a given base)² that

$$x^j_{,i}(\hat{x}'(p), \dots, \hat{x}^n(p)) = \underline{e}^j(p) \cdot \bar{e}_{,i}(p). \quad (*)$$

Similarly,

$$\bar{e}^i(p) = \nabla \hat{x}^i(p).$$

But

$$\hat{x}^i(p) = \bar{x}^i = \bar{x}^i(x', \dots, x^n) = \bar{x}^i(\hat{x}'(p), \dots, \hat{x}^n(p));$$

and again by the chain rule,

$$\nabla \hat{x}^i(p) = \bar{x}^i_{,j}(\hat{x}'(p), \dots, \hat{x}^n(p)) \nabla \hat{x}^j(p), \quad)^3$$

¹ Work this out as Exercise A6.13.3.

² Cf. Theorem A2.2.11.

³ Work this out as Exercise A6.13.4.

or

$$\bar{e}^i(p) = \bar{x}_{,j}^i(\hat{x}^1(p), \dots, \hat{x}^n(p)) \underline{e}^j(p).$$

Again it follows from Theorem A6.11.3 that

$$\bar{x}_{,j}^i(\hat{x}^1(p), \dots, \hat{x}^n(p)) = \underline{e}_j(p) \cdot \bar{e}^i(p). \quad (**)$$

We see from (*) and (**), that these partial derivatives may be identified with the α 's and β 's of Theorem A2.6.1; specifically

$$\left. \begin{aligned} \bar{x}_{,j}^i(\hat{x}^1(p), \dots, \hat{x}^n(p)) &= \alpha_j^i, \\ x_{,j}^i(\bar{x}^1(p), \dots, \bar{x}^n(p)) &= \beta_j^i. \end{aligned} \right\} \quad (-7-)$$

Of course, here the α 's and β 's and the various bases vary from point to point, but that does not affect the validity of Theorem A2.6.1; and hence we can read off

Theorem A6.13.2. The vectors of the natural basis and its dual transform according to the following rules under a coordinate transformation: ¹

$$\bar{e}_i(p) = \frac{\partial x^j}{\partial \bar{x}^i}(\bar{x}^1(p), \dots, \bar{x}^n(p)) \underline{e}_j(p),$$

$$\bar{e}^i(p) = \frac{\partial \bar{x}^i}{\partial x^j}(\hat{x}^1(p), \dots, \hat{x}^n(p)) \underline{e}^j(p),$$

¹Here, for obvious mnemonic reasons, we revert to the LEIBNITZ notation for the partial derivatives.

$$\tilde{e}_i = \frac{\partial \bar{x}^j}{\partial x^i} \tilde{e}_j ,$$

$$\tilde{e}^i = \frac{\partial x^i}{\partial \bar{x}^j} \tilde{e}^j .$$

The identification (-7-) can also be combined with Theorems A2.6.2 and A3.3.1 to get the transformation rules for vector and tensor components. Thus, we have

Theorem A6.13.3. The components of a vector field u and a second order tensor field T transform according to the following rules under a coordinate transformation:

$$\bar{u}_i(p) = \frac{\partial x^j}{\partial \bar{x}^i} (\hat{x}^1(p), \dots, \hat{x}^n(p)) u_j(p) ,$$

$$\bar{u}^i(p) = \frac{\partial \bar{x}^i}{\partial x^j} (\hat{x}^1(p), \dots, \hat{x}^n(p)) u^j(p) ,$$

$$u_i = \frac{\partial \bar{x}^j}{\partial x^i} \bar{u}_j ,$$

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

$$\bar{T}_{ij} = \frac{\partial x^k}{\partial \bar{x}^i} \frac{\partial x^l}{\partial \bar{x}^j} T_{kl} ,$$

$$\bar{T}^{ij} = \frac{\partial \bar{x}^i}{\partial x^k} \frac{\partial \bar{x}^j}{\partial x^l} T^{kl} ,$$

$$\bar{T}_{i \cdot j} = \frac{\partial x^k}{\partial \bar{x}^i} \frac{\partial \bar{x}^j}{\partial x^l} T_{k \cdot l} ,$$

$$\bar{T}^{\cdot i j} = \frac{\partial \bar{x}^i}{\partial x^k} \frac{\partial \bar{x}^j}{\partial x^l} T^{\cdot k l} ,$$

$$T_{ij} = \frac{\partial \bar{x}^k}{\partial x^i} \frac{\partial \bar{x}^l}{\partial x^j} \bar{T}_{kl},$$

$$T^{ij} = \frac{\partial x^i}{\partial \bar{x}^k} \frac{\partial x^j}{\partial \bar{x}^l} \bar{T}^{kl},$$

$$T_{i \cdot j} = \frac{\partial \bar{x}^k}{\partial x^i} \frac{\partial \bar{x}^l}{\partial x^j} \bar{T}_{k \cdot l},$$

$$T^{\cdot i}_{\cdot j} = \frac{\partial x^i}{\partial \bar{x}^k} \frac{\partial x^j}{\partial \bar{x}^l} \bar{T}^{\cdot k}_{\cdot l}.$$

Such transformation rules are the starting point for general¹ tensor analysis via the component-transformation rule approach (see the discussion in Chapter A4).

Supplementary Reading

Same as for § A6.11

¹ General here refers to the fact that the bases are those associated with a general or curvilinear (as opposed to Cartesian) coordinate system.

A6.14. Differentiation of Scalar Fields

Let $f: D \rightarrow \mathbb{R}$ be a differentiable scalar field on an open set $D \subset E_n$.¹ We adopt the customary but loose convention of writing

$$f(p) = f(\hat{p}(x^1, x^2, \dots, x^n)) =: f(x^1, x^2, \dots, x^n),$$

where

$$p = \hat{p}(x^1, x^2, \dots, x^n), \quad x^i = \hat{x}^i(p), \quad i = 1, 2, \dots, n$$

defines a given coordinate system for E_n .

By the chain rule (Theorem A5.7.1), the "function" $f(x^1, x^2, \dots, x^n)$ ² is also differentiable. Then by Theorem A5.12.5 it has first partial derivatives, which by the chain rule are given by

$$D_i f(x^1, x^2, \dots, x^n) [\alpha_i] = Df(\hat{p}(x^1, x^2, \dots, x^n)) [D_i \hat{p}(x^1, x^2, \dots, x^n) [\alpha_i]].$$

By the linearity of derivatives,

$$\delta_i D_i f(x^1, \dots, x^n) [1] = \delta_i Df(\hat{p}(x^1, x^2, \dots, x^n)) [D_i \hat{p}(x^1, x^2, \dots, x^n) [1]].$$

¹ Cf. Definition A6.4.1 and the introduction to § A6.5.

² Here, our looseness exacts its toll.

which by Theorem A5.3.1 can be expressed as

$$\begin{aligned}
 f_{,i}(x^1, x^2, \dots, x^n) &= Df(\hat{p}(x^1, x^2, \dots, x^n)) [\hat{p}_{,i}(x^1, x^2, \dots, x^n)]^1 \\
 &= \nabla f(\hat{p}(x^1, x^2, \dots, x^n)) \cdot \hat{p}_{,i}(x^1, x^2, \dots, x^n) \quad (\text{Thm. A5.4.1}) \\
 &= \nabla f(p) \cdot \hat{p}_{,i}(\hat{x}^1(p), \hat{x}^2(p), \dots, \hat{x}^n(p)) \quad (\text{substitution}) \\
 &= \nabla f(p) \cdot e_i(p) \quad (\text{Theorem A6.11.1}) \\
 &= (\nabla f(p))_i \quad (\text{Theorem A6.11.3}).
 \end{aligned}$$

We can get the contravariant components of $\nabla f(p)$ by using the g symbols to raise the index)²:

$$(\nabla f(p))^i = g^{ij}(p) (\nabla f(p))_j.$$

Hence, we have established

Theorem A6.14.1. Let the scalar field f be differentiable at the nonsingular point p . Then the components of its gradient are given by

$$(\nabla f(p))_i = f_{,i}(\hat{x}^1(p), \hat{x}^2(p), \dots, \hat{x}^n(p))$$

¹ Cf. Theorem A5.12.8.

² Cf. Theorem A2.5.11.

and

$$(\nabla f(p))^i = g^{ij}(p) f_{,j}(\hat{x}^1(p), \hat{x}^2(p), \dots, \hat{x}^n(p)).$$

Thus, even for general coordinates ¹ the covariant components of the gradient of a scalar field are just the ordinary first partial derivatives w.r.t. the corresponding coordinates. In the following sections, we shall see that the situation is not so simple for higher-order tensor fields.

Supplementary Reading

Same as for § A6.11

¹ See p. § A6.5 for the rectangular Cartesian case.

A6.15. Differentiation of Vector Fields. CHRISTOFFEL Symbols

Let $\underline{v}: D \rightarrow \mathcal{V}_n$ be a differentiable vector field on an open set $D \subset E_n$,¹ and as in the last section write

$$\underline{v}(\underline{p}) = \underline{v}(\hat{\underline{p}}(x^1, x^2, \dots, x^n) =: \underline{v}(x^1, x^2, \dots, x^n),$$

where

$$\underline{p} = \hat{\underline{p}}(x^1, x^2, \dots, x^n), \quad x^i = \hat{x}^i(\underline{p}), \quad i = 1, 2, \dots, n$$

defines a given coordinate system for E_n .

By the chain rule, the "function" $\underline{v}(x^1, x^2, \dots, x^n)$ has ordinary first partial derivatives

$$\begin{aligned} \underline{v}_{,i}(x^1, x^2, \dots, x^n) &= D\underline{v}(\hat{\underline{p}}(x^1, x^2, \dots, x^n) [\hat{\underline{p}}_{,i}(x^1, x^2, \dots, x^n)])^2 \\ &= D\underline{v}(\underline{p}) [\hat{\underline{p}}_{,i}(\hat{x}^1(\underline{p}), \hat{x}^2(\underline{p}), \dots, \hat{x}^n(\underline{p}))] \text{ (substitution)} \\ &= D\underline{v}(\underline{p}) [\underline{e}_{,i}(\underline{p})] \text{ (Theorem A6.11.1)} \\ &= D\underline{v}(\underline{p}) \underline{e}_{,i}(\underline{p}) \text{ (tensor notation).} \end{aligned}$$

But by Theorem A6.11.3 we can also write

¹ cf. Definition A6.4.1 and the introduction to § A6.6.

² Work this out as Exercise A6.15.1.

$$\underline{v}(p) = v^k(p) \underline{e}_k(p) ,$$

or

$$\underline{v}(x^1, x^2, \dots, x^n) = v^k(x^1, x^2, \dots, x^n) \underline{e}_k(x^1, x^2, \dots, x^n),$$

where, of course,

$$v^k(x^1, x^2, \dots, x^n) := v^k(\hat{p}(x^1, x^2, \dots, x^n))$$

and

$$\underline{e}_{\sim k}(x^1, x^2, \dots, x^n) := \underline{e}_{\sim k}(\hat{p}(x^1, x^2, \dots, x^n)).$$

Differentiation of this expression yields

$$\begin{aligned} \underline{v}_{,i}(x^1, x^2, \dots, x^n) &= v^k_{,i}(x^1, x^2, \dots, x^n) \underline{e}_k(x^1, x^2, \dots, x^n) \\ &\quad + v^k(x^1, x^2, \dots, x^n) \underline{e}_{k,i}(x^1, x^2, \dots, x^n). \end{aligned}$$

The appearance here of $\underline{e}_{k,i}(x^1, x^2, \dots, x^n)$ is why we assumed in Definition A6.11.1 that $\hat{p}$ is two-times continuously differentiable. Equating our two expressions for $\underline{v}_{,i}$, we get (with the arguments of the functions suppressed)

$$(D\underline{v}) \underline{e}_{\sim i} = v^k_{,i} \underline{e}_{\sim k} + v^k \underline{e}_{k,i};$$

and $\therefore$

$$\underline{e}^j \cdot [(D\underline{v}) \underline{e}_{\sim i}] = v^k_{,i} \underline{e}^j \cdot \underline{e}_{\sim k} + v^k \underline{e}^j \cdot \underline{e}_{k,i}.$$

Now, by Theorem A6.11.3,

$$\underline{e}^j \cdot [(D\underline{v}) \underline{e}_i] = (D\underline{v})^j_i ;$$

and by Definition A6.11.2 and Theorem A2.5.10 (ii),

$$\underline{e}^j \cdot \underline{e}_k = g^j_k = S^j_k .$$

Hence, we have proven

Theorem A6.15.1. Let the vector field $\underline{v}$ be differentiable at the nonsingular point $\underline{p}$. Then the mixed components of its derivative are given by

$$(D\underline{v}(\underline{p}))^i_j = v^i_{,j}(\hat{x}^1(\underline{p}), \hat{x}^2(\underline{p}), \dots, \hat{x}^n(\underline{p})) + \{^i_{kj}\} v^k(\hat{x}^1(\underline{p}), \hat{x}^2(\underline{p}), \dots, \hat{x}^n(\underline{p}))$$

where

$$\{^i_{kj}\} = \underline{e}^i(\hat{x}^1(\underline{p}), \hat{x}^2(\underline{p}), \dots, \hat{x}^n(\underline{p})) \cdot \underline{e}_{k,j}(\hat{x}^1(\underline{p}), \hat{x}^2(\underline{p}), \dots, \hat{x}^n(\underline{p}))$$

is the CHRISTOFFEL symbol of the second kind.

Of course, we can get any of the other components of $D\underline{v}(\underline{p})$ by using the g symbols to raise and lower indices.)¹

In the tensor analysis literature based on the component-transformation rule approach, one commonly sees the notations

$$v^i_{,j} + \{^i_{kj}\} v^k = v^i_{;j} = v^i_{|j} = \nabla_j v^i ,$$

¹ Cf. Theorem A3.2.8.

and there this quantity is called the j^{th} covariant derivative of the i^{th} contravariant component of the vector field $\underline{v}$.

For the purposes of calculation, it is convenient to introduce a new symbol by formally lowering the upper index on $\{k_j^i\}$.

Definition A6.15.1. The CHRISTOFFEL symbol of the first kind is defined by

$$[k_j, i] = g_{ir} \{k_j^r\}.$$

Often the CHRISTOFFEL symbols are denoted by

$$\{k_j^i\} = \Gamma_{kj}^i \quad \text{and} \quad [k_j, i] = \Gamma_{kji};$$

however, this notation has the unfortunate feature of suggesting that they are tensor components.

Exercise A6.15.2. Derive the transformation rule for $\{k_j^i\}$ under a coordinate transformation and conclude thereby that the CHRISTOFFEL symbols are not tensor components.

Theorem A6.15.2. The CHRISTOFFEL symbols have the following properties:

$$(i) \quad \left\{ \begin{smallmatrix} i \\ k_j \end{smallmatrix} \right\} = \left\{ \begin{smallmatrix} i \\ j_k \end{smallmatrix} \right\} ;$$

$$(ii) \quad [k_j, i] = [j_k, i] ;$$

$$(iii) \quad \left\{ \begin{smallmatrix} i \\ k_j \end{smallmatrix} \right\} = g^{ir} [k_j, r] ;$$

$$(iv) \quad [k_j, i] = \underline{e}_i \cdot \underline{e}_{k,j} ;$$

$$(v) \quad [k_j, i] = \frac{1}{2} (g_{ij,k} + g_{ik,j} - g_{kji}) .$$

Proof. The proof of (i) - (iv) constitute Exercises A6.15.3
- 6. To prove (v), we note from Definition A6.11.2 that

$$g_{ij,k} = (\underline{e}_i \cdot \underline{e}_j)_{,k} = \underline{e}_{i,k} \cdot \underline{e}_j + \underline{e}_i \cdot \underline{e}_{j,k} .$$

By Theorem A6.11.1,

$$\begin{aligned} \underline{e}_{i,j} &= \hat{P}_{,ij} \\ &= \hat{P}_{,ji} \quad (p \in C^2) \\ &= \underline{e}_{j,i} \quad (\text{Theorem A6.11.1}) . \end{aligned}$$

These last two results $\Rightarrow$

$$\begin{aligned} g_{ij,k} + g_{ik,j} - g_{kji} &= 2 \underline{e}_i \cdot \underline{e}_{k,j} \\ &= 2 [k_j, i] \quad \text{by (iv). } \square \end{aligned}$$

Parts (v) and (iii) of Theorem A6.15.2 provide an easy way to calculate the CHRISTOFFEL symbols. In the component-transformation rule approach to tensors, the coordinate system would be characterized by specifying the components of the metric tensor (the g 's), and (v) and (iii) would be the definitions of the CHRISTOFFEL symbols.

Turning to the divergence of a differentiable vector field $\underline{v}$, we have by Definition A6.6.1 that

$$\begin{aligned}\operatorname{div} \underline{v} &= \operatorname{tr} D\underline{v} \\ &= (D\underline{v})^i_{\cdot i} \quad (\text{Theorem A3.7.1}).\end{aligned}$$

Then Theorem A6.15.1 yields

Theorem A6.15.3. Let $\underline{v}$ be a differentiable vector field.

Then

$$\operatorname{div} \underline{v} = v^i_{,i} + \{^i_{ki}\} v^k.$$

The curl of a differentiable vector field $\underline{v}$ is a bit more involved. Here, we must necessarily restrict our consideration to E_3 and suppose that the associated translation space $\mathcal{V}_3$ has been oriented by the assignment of a volume orientation function Δ . Then by Definition A6.6.1,

$$\operatorname{curl} \underline{v} = 2 \text{ axial skew } D\underline{v}.$$

By Theorem A3.6.9,

$$(\text{curl } \underline{v})^i = \frac{-1}{\Delta(\underline{e}_1, \underline{e}_2, \underline{e}_3)} \varepsilon^{ijk} (D\underline{v})_{[jk]}^i,$$

where

$$(D\underline{v})_{[jk]}^i = \frac{1}{2} [(D\underline{v})_{jk}^i - (D\underline{v})_{kj}^i]$$

by Theorem A3.6.6. By the skew symmetry of the alternating symbol, we can write

$$(\text{curl } \underline{v})^i = \frac{-1}{\Delta(\underline{e}_1, \underline{e}_2, \underline{e}_3)} \varepsilon^{ijk} (D\underline{v})_{jk}^i.$$

By Theorem A3.2.8,

$$(D\underline{v})_{jk}^i = g_{jl} (D\underline{v})_{\cdot k}^l;$$

and

$$(D\underline{v})_{\cdot k}^l = v_{\cdot k}^l + \{m_k\}^l v^m$$

by Theorem A6.15.1. Thus,

$$\begin{aligned} (\text{curl } \underline{v})^i &= \frac{-1}{\Delta(\underline{e}_1, \underline{e}_2, \underline{e}_3)} \varepsilon^{ijk} g_{jl} (D\underline{v})_{\cdot k}^l \\ &= \frac{1}{\Delta(\underline{e}_1, \underline{e}_2, \underline{e}_3)} \varepsilon^{ijk} g_{kl} (D\underline{v})_{\cdot j}^l, \end{aligned}$$

where $(Dv)_0^l \cdot j$ is known.

By Theorem A2.7.6,

$$[\Delta(\underline{e}_1, \underline{e}_2, \underline{e}_3)]^2 = \det[g_{ij}];$$

whence

$$\Delta(\underline{e}_1, \underline{e}_2, \underline{e}_3) = \pm \sqrt{\det[g_{ij}]}.$$

We must choose the + sign if the natural basis $\{\underline{e}_1, \underline{e}_2, \underline{e}_3\}$ is positive and the - sign if it is negative.¹ This, of course, depends on the coordinate system and on the choice of orientation of the translation space. However, given a coordinate system and a volume orientation function, if it happens that $\Delta(\underline{e}_1, \underline{e}_2, \underline{e}_3) < 0$, then

$$\Delta(\underline{e}_2, \underline{e}_1, \underline{e}_3) = -\Delta(\underline{e}_1, \underline{e}_2, \underline{e}_3) > 0. \quad)^2$$

Thus, we can always arrange for the natural basis to be positive by our labeling of the coordinates.

In summary, we have

Theorem A6.15.4. Let v be a differentiable vector field on an open subset of E_3 . Let the associated translation space be

¹ See Definition A2.7.6.

² Cf. Axiom (O2) of Definition A2.7.3.

oriented and label the coordinates in such a way that the natural basis is positive at the point of interest. Then at this point,

where

$$(\text{curl } \underline{v})^i = \frac{1}{\sqrt{g}} \varepsilon^{ijk} g_{ke} (D\underline{v})^e_{\cdot j} ,$$

and

$$g = \det [g_{mn}]$$

$$(D\underline{v})^e_{\cdot j} = v^e_{,j} + \left\{ \begin{matrix} e \\ mj \end{matrix} \right\} v^m .$$

We also can get at the components of the curl via the component - transformation rule approach. Here, the idea is to take the expressions for the components of the curl w.r.t. a rectangular Cartesian coordinate system and transform them to components w.r.t. a general coordinate system. It will be instructive to work through this, because it will force us to touch base with several of the key ideas in the component - transformation rule approach to tensors.

First of all, we know from Theorem A6.6.2 that the components of the curl w.r.t. a rectangular Cartesian coordinate system whose basis $\{\underline{e}_{(1)}, \underline{e}_{(2)}, \underline{e}_{(3)}\}$ is positive are given by

$$(\text{curl } \underline{v})^{(i)} = \sum_{j,k=1}^3 \varepsilon^{ijk} v^{(k)}_{,j}$$

Let us denote the rectangular Cartesian coordinate system used here by $\{\mathbb{P} \leftrightarrow x\}$.

From Theorem A6.15.1, we recognize $v^{(k)}_{,j}$ as the components of $D\underline{v}$ w.r.t. $\{\mathbb{P} \leftrightarrow x\}$; i.e.,

$$(D\underline{v})^k_{\cdot j} = v^{(k)}_{,j}.$$

This is because $\{^i_{kj}\} = 0$ for $\{\mathbb{P} \leftrightarrow x\}$. In fact, since covariant and contravariant components coalesce for orthonormal bases, we could just as well write

$$(D\underline{v})_{<kj>} = v^{(k)}_{,j}.$$

At this stage ε^{ijk} is just a symbol. However, by the considerations of Chpt. A4, we can view ε^{ijk} as the components w.r.t. $\{\mathbb{P} \leftrightarrow x\}$ of the third-order tensor

$$\varepsilon^{ijk} \underline{e}_{\langle i \rangle} \otimes \underline{e}_{\langle j \rangle} \otimes \underline{e}_{\langle k \rangle}.$$

To get the components w.r.t. a general coordinate system $\{p \leftrightarrow \bar{x}\}$, we just use the appropriate transformation rule:

$$\begin{aligned} \bar{\varepsilon}^{ijk} &= \frac{\partial \bar{x}^i}{\partial x^l} \frac{\partial \bar{x}^j}{\partial x^m} \frac{\partial \bar{x}^k}{\partial x^n} \varepsilon^{lmn} \quad)^1 \\ &= \det \left[\frac{\partial \bar{x}^p}{\partial x^q} \right] \varepsilon^{ijk} \end{aligned}$$

by a standard result from the theory of determinants (of matrices). We can get a handle on $\det \left[\frac{\partial \bar{x}^p}{\partial x^q} \right]$ as follows. The contravariant components of $\underline{I}$ w.r.t. the rectangular coordinate system $\{p \leftrightarrow x\}$ are

$$g^{ij} = \underline{e}^{\langle i \rangle} \cdot \underline{e}^{\langle j \rangle} = \delta^{ij},$$

and w.r.t. the general coordinate system $\{p \leftrightarrow \bar{x}\}$ they are $\bar{g}^{ij}$. $\therefore$ by the appropriate transformation rule,

$$\bar{g}^{ij} = \frac{\partial \bar{x}^i}{\partial x^m} \frac{\partial \bar{x}^j}{\partial x^n} g^{mn} = \sum_{m=1}^3 \frac{\partial \bar{x}^i}{\partial x^m} \frac{\partial \bar{x}^j}{\partial x^m}.$$

Next, we view these two-index symbols as elements of matrices and take the determinant to get

$$\det [\bar{g}^{ij}] = \left(\det \left[\frac{\partial \bar{x}^p}{\partial x^q} \right] \right)^2,$$

which by Theorem A2.5.10 becomes

¹Again see Chpt. A4; also of Theorem A6.13.3.

$$\left(\det \left[\frac{\partial \bar{x}^p}{\partial x^q} \right] \right)^2 = \left(\det [\bar{g}_{ij}] \right)^{-1}.$$

Before taking the square root, we need to know whether $\det \left[\frac{\partial \bar{x}^p}{\partial x^q} \right]$ should be positive or negative. Consider the transformation rules for the vectors of the natural basis :)

$$\underline{\bar{e}}_{\langle i \rangle} = \frac{\partial \bar{x}^j}{\partial x^i} \underline{\bar{e}}_j.$$

$$\begin{aligned} \Delta(\underline{\bar{e}}_{\langle 1 \rangle}, \underline{\bar{e}}_{\langle 2 \rangle}, \underline{\bar{e}}_{\langle 3 \rangle}) &= \Delta\left(\frac{\partial \bar{x}^j}{\partial x^1} \underline{\bar{e}}_j, \frac{\partial \bar{x}^k}{\partial x^2} \underline{\bar{e}}_k, \frac{\partial \bar{x}^l}{\partial x^3} \underline{\bar{e}}_l\right) \\ &= \frac{\partial \bar{x}^j}{\partial x^1} \frac{\partial \bar{x}^k}{\partial x^2} \frac{\partial \bar{x}^l}{\partial x^3} \Delta(\underline{\bar{e}}_j, \underline{\bar{e}}_k, \underline{\bar{e}}_l) \\ &= \frac{\partial \bar{x}^j}{\partial x^1} \frac{\partial \bar{x}^k}{\partial x^2} \frac{\partial \bar{x}^l}{\partial x^3} \epsilon_{jkl} \Delta(\underline{\bar{e}}_1, \underline{\bar{e}}_2, \underline{\bar{e}}_3) \\ &= \det \left[\frac{\partial \bar{x}^p}{\partial x^q} \right] \Delta(\underline{\bar{e}}_1, \underline{\bar{e}}_2, \underline{\bar{e}}_3), \end{aligned}$$

where we have used the linearity and skew-symmetry of Δ and the definition of the determinant of a matrix. Now, we started out with $\Delta(\underline{e}_{\langle 1 \rangle}, \underline{e}_{\langle 2 \rangle}, \underline{e}_{\langle 3 \rangle}) = 1$, and if we require that the natural basis of the general coordinate system $\{p \leftrightarrow \bar{x}\}$ also be positive, then

$$\det \left[\frac{\partial \bar{x}^p}{\partial x^q} \right] > 0.$$

This condition is generally assumed ($\forall$ coordinate systems

$\{p \leftrightarrow x\}$ and $\{p \leftrightarrow \bar{x}\}$ in the component-transformation rule approach. It guarantees that all natural bases are like-handed. We shall adopt it here. Then from the result at the top of the previous page,

$$\det \left[\frac{\partial \bar{x}^p}{\partial x^q} \right] = \frac{1}{\sqrt{\bar{g}}} ,$$

where

$$\bar{g} = \det [\bar{g}_{ij}] .$$

Returning now to the top of p. A6.15.11, we see that

$$\bar{\epsilon}^{ijk} = \frac{1}{\sqrt{\bar{g}}} \epsilon^{ijk} .$$

In other words, the 3rd-order tensor whose contravariant components w.r.t. a general coordinate system $\{p \leftrightarrow \bar{x}\}$ (whose natural basis is positive) are

$$\bar{\epsilon}^{ijk} = \frac{1}{\sqrt{\bar{g}}} \epsilon^{ijk}$$

has components ϵ^{ijk} relative to a rectangular Cartesian coordinate system (with positive basis). We have introduced the ϵ notation so that ϵ^{ijk} is always just the alternating symbol. This is a standard device, but some authors interchange the roles of ϵ and $\bar{\epsilon}$.

Returning to our starting point on p. A6.15.10, we can

now say that w.r.t. the rectangular coordinate system $\{p \leftrightarrow x\}$

$$(\text{curl } \underline{v})^{(i)} = e^{ijk} (D \underline{v})_{kj} ,$$

and all of the ingredients are tensor components (w.r.t. $\{p \leftrightarrow x\}$). This is enough to allow us to conclude that an equation of the same form holds amongst the components of these tensors w.r.t. any other coordinate system $\{p \leftrightarrow \bar{x}\}$; i.e.,

$$\overline{(\text{curl } \underline{v})^i} = \bar{e}^{ijk} \overline{(D \underline{v})_{kj}} .$$

Specifically, what we need is

Theorem A6.15.5. Let a^i, B_{ij}, C^{ijk} and $\bar{a}^i, \bar{b}_{ij}, \bar{C}^{ijk}$ be components w.r.t. general coordinate systems $\{p \leftrightarrow x\}$ and $\{p \leftrightarrow \bar{x}\}$ of a first-order tensor $\underline{a}$, a second-order tensor $\underline{B}$, and a third-order tensor $\underline{C}$. Then

$$a^i = C^{ijk} B_{kj} \Rightarrow \bar{a}^i = \bar{C}^{ijk} \bar{B}_{kj} .$$

Proof. We just use appropriate transformation rules.

$$\bar{a}^i = \frac{\partial \bar{x}^i}{\partial x^l} a^l = \frac{\partial \bar{x}^i}{\partial x^l} C^{ljk} B_{kj} .$$

But

$$C^{ljk} = \frac{\partial x^l}{\partial \bar{x}^m} \frac{\partial x^j}{\partial \bar{x}^n} \frac{\partial x^k}{\partial \bar{x}^p} \bar{C}^{mnp} , \quad B_{kj} = \frac{\partial \bar{x}^q}{\partial x^k} \frac{\partial \bar{x}^r}{\partial x^j} \bar{B}_{qr} ;$$

so that

$$\begin{aligned}\bar{a}^i &= \frac{\partial \bar{x}^i}{\partial x^l} \left(\frac{\partial x^l}{\partial \bar{x}^m} \frac{\partial x^j}{\partial \bar{x}^n} \frac{\partial x^k}{\partial \bar{x}^p} \bar{c}^{mnp} \right) \left(\frac{\partial \bar{x}^q}{\partial x^k} \frac{\partial \bar{x}^r}{\partial x^j} \bar{b}_{qr} \right) \\ &= \left(\frac{\partial \bar{x}^i}{\partial x^l} \frac{\partial x^l}{\partial \bar{x}^m} \right) \left(\frac{\partial \bar{x}^r}{\partial x^j} \frac{\partial x^j}{\partial \bar{x}^n} \right) \left(\frac{\partial \bar{x}^q}{\partial x^k} \frac{\partial x^k}{\partial \bar{x}^p} \right) \bar{c}^{mnp} \bar{b}_{qr} .\end{aligned}$$

By (†) on p.A6.13.5 and Theorem A2.6.1,

$$\frac{\partial \bar{x}^i}{\partial x^l} \frac{\partial x^l}{\partial \bar{x}^m} = \alpha_l^i \beta_m^l = \delta_m^i . \quad)^1$$

$$\therefore \bar{a}^i = \delta_m^i \delta_n^r \delta_p^q \bar{c}^{mnp} \bar{b}_{qr} = \bar{c}^{irg} \bar{b}_{gr} . \quad \square$$

So now we can say that for any coordinate system $\{p \leftrightarrow x\}$ with positive natural basis,

$$(\text{curl } \underline{v})^i = e^{ijk} (D \underline{v})_{kj} . \quad)^2$$

But

$$e^{ijk} = \frac{1}{\sqrt{g}} \varepsilon^{ijk} ,$$

and

$$(D \underline{v})_{kj} = g_{ke} (D \underline{v})^e_j ;$$

so that

¹ This also follows immediately from the chain rule; i.e.,

$$\frac{\partial \bar{x}^i}{\partial x^l} \frac{\partial x^l}{\partial \bar{x}^m} = \frac{\partial \bar{x}^i}{\partial \bar{x}^m} = \delta_m^i .$$

² Here, we call the general coordinate system $\{p \leftrightarrow x\}$ rather than $\{p \leftrightarrow \bar{x}\}$ so that our final formula is not cluttered with overbars.

$$(\text{curl } \underline{v})^i = \frac{1}{\sqrt{g}} \varepsilon^{ijk} g_{kl} (Dv)^l_j ,$$

which is in agreement with Theorem A6.15.4.

Supplementary Reading

Same as for § A6.11

A6.16. Differentiation of Second-Order Tensor Fields

Let $\underline{T}: \mathcal{D} \rightarrow \text{Lin } \mathcal{V}_n$ be a differentiable second-order tensor field on an open set $\mathcal{D} \subset E_n$,¹ and as in the previous two sections write

$$\underline{T}(\mathcal{P}) = \underline{T}(\hat{\mathcal{P}}(x^1, x^2, \dots, x^n)) =: \underline{T}(x^1, x^2, \dots, x^n),$$

where

$$\mathcal{P} = \hat{\mathcal{P}}(x^1, x^2, \dots, x^n), \quad x^i = \hat{x}^i(\mathcal{P}), \quad i = 1, 2, \dots, n$$

defines a given coordinate system for E_n .

All that we need is the divergence, $\text{div } \underline{T}$. We got at this indirectly in Theorem A6.7.1 by introducing the vector field

$$\underline{v}_a(\mathcal{P}) = (\underline{T}(\mathcal{P}))^T \underline{a}$$

for arbitrary uniform $\underline{a} \in \mathcal{V}_n$. Then $\text{div } \underline{T}(\mathcal{P})$ is the unique vector with the property that

$$\text{div } \underline{v}_a(\mathcal{P}) = (\text{div } \underline{T}(\mathcal{P})) \cdot \underline{a} \quad \forall \text{ uniform } \underline{a} \in \mathcal{V}_n.$$

In terms of components,

¹ Cf. Definition A6.4.1 and the introduction to §A6.7.

$$V_a^i = T^{ji} a_j ;$$

and $\therefore$ by Theorem A6.15.3,

$$\begin{aligned} \operatorname{div} V_a &= V_{a,i}^i + \{^i_{ki}\} V_a^k \\ &= T^{ji}_{,i} a_j + T^{ji} a_{j,i} + \{^i_{ki}\} T^{jk} a_j . \end{aligned}$$

Now, $a_j = \underline{e}_j \cdot \underline{a}$; and since $\underline{a}$ does not depend on position,

$$a_{j,i} = \underline{e}_{j,i} \cdot \underline{a} = (\underline{e}_{j,i})^k a_k = \underline{e}^k \cdot (\underline{e}_{j,i}) a_k = \{^k_{ji}\} a_k .$$

Note that relative to a general coordinate system a uniform vector field does not have uniform components. Thus,

$$\begin{aligned} \operatorname{div} V_a &= \left(T^{ji}_{,i} + \{^j_{ki}\} T^{ki} + \{^i_{ki}\} T^{jk} \right) a_j \\ &= (\operatorname{div} T) \cdot \underline{a} = (\operatorname{div} T)^j a_j . \end{aligned}$$

Since this must hold $\forall \underline{a}$, we have

Theorem A6.16.1. Let T be a differentiable second-order tensor field. Then the contravariant components of its divergence are given by

$$(\operatorname{div} T)^i = T^{ij}_{,j} + \{^i_{kj}\} T^{kj} + \{^j_{kj}\} T^{ik} .$$

You may find the following formulas from Truesdell and Noll's "Nonlinear Field Theories of Mechanics" to be useful.

6A. Appendix. Cylindrical and spherical co-ordinates. The following formulae are frequently used in this article and are likely to be useful also in future work on special solutions.

α) Cylindrical co-ordinates r, θ, z :

Components of the unit tensor:

$$g_{rr} = g^{\theta\theta} = g_{zz} = g^{zz} = 1, \quad g_{\theta\theta} = r^2, \quad g^{\theta\theta} = \frac{1}{r^2}. \quad (6A.1)$$

Non-zero Christoffel symbols:

$$\left\{ \begin{matrix} \theta \\ r \theta \end{matrix} \right\} = \left\{ \begin{matrix} \theta \\ \theta r \end{matrix} \right\} = \frac{1}{r}, \quad \left\{ \begin{matrix} r \\ \theta \theta \end{matrix} \right\} = -r. \quad (6A.2)$$

Contravariant components of the acceleration $\ddot{x}$ of a velocity field $\dot{x} = \dot{x}(x, t)$:

$$\ddot{x}^r = \partial_t \dot{x}^r + \dot{x}^k \partial_k \dot{x}^r - r(\dot{x}^\theta)^2, \quad \ddot{x}^\theta = \partial_t \dot{x}^\theta + \dot{x}^k \partial_k \dot{x}^\theta + \frac{2}{r} \dot{x}^r \dot{x}^\theta, \quad \ddot{x}^z = \partial_t \dot{x}^z + \dot{x}^k \partial_k \dot{x}^z. \quad (6A.3)$$

Physical components of the divergence of a symmetric second-order tensor field T :

$$\left. \begin{aligned} (\operatorname{div} T)\langle r \rangle &= \partial_r T\langle rr \rangle + \frac{1}{r} \partial_\theta T\langle r\theta \rangle + \partial_z T\langle rz \rangle + \frac{T\langle rr \rangle - T\langle \theta\theta \rangle}{r}, \\ (\operatorname{div} T)\langle \theta \rangle &= \partial_r T\langle r\theta \rangle + \frac{1}{r} \partial_\theta T\langle \theta\theta \rangle + \partial_z T\langle \theta z \rangle + \frac{2}{r} T\langle r\theta \rangle, \\ (\operatorname{div} T)\langle z \rangle &= \partial_r T\langle rz \rangle + \frac{1}{r} \partial_\theta T\langle \theta z \rangle + \partial_z T\langle zz \rangle + \frac{1}{r} T\langle rz \rangle. \end{aligned} \right\} \quad (6A.4)$$

β) Spherical co-ordinates r, θ, φ :

Components of the unit tensor:

$$g_{rr} = g^{\theta\theta} = 1, \quad g_{\theta\theta} = r^2, \quad g^{\theta\theta} = \frac{1}{r^2}, \quad g_{\varphi\varphi} = (r \sin \theta)^2, \quad g^{\varphi\varphi} = \frac{1}{(r \sin \theta)^2}. \quad (6A.5)$$

Non-zero Christoffel symbols:

$$\left. \begin{aligned} \left\{ \begin{matrix} \theta \\ \theta r \end{matrix} \right\} &= \left\{ \begin{matrix} \theta \\ r \theta \end{matrix} \right\} = \left\{ \begin{matrix} \varphi \\ \varphi r \end{matrix} \right\} = \left\{ \begin{matrix} \varphi \\ r \varphi \end{matrix} \right\} = \frac{1}{r}, \quad \left\{ \begin{matrix} r \\ \theta \theta \end{matrix} \right\} = -r, \quad \left\{ \begin{matrix} r \\ \varphi \varphi \end{matrix} \right\} = -r \sin^2 \theta, \\ \left\{ \begin{matrix} \varphi \\ \varphi \theta \end{matrix} \right\} &= \left\{ \begin{matrix} \varphi \\ \theta \varphi \end{matrix} \right\} = \cot \theta, \quad \left\{ \begin{matrix} \theta \\ \varphi \varphi \end{matrix} \right\} = -\sin \theta \cos \theta. \end{aligned} \right\} \quad (6A.6)$$

Contravariant components of the acceleration $\ddot{x}$ of a velocity field $\dot{x} = \dot{x}(x, t)$:

$$\left. \begin{aligned} \ddot{x}^r &= \partial_t \dot{x}^r + \dot{x}^k \partial_k \dot{x}^r - r[(\dot{x}^\theta)^2 + \sin^2 \theta (\dot{x}^\varphi)^2], \\ \ddot{x}^\theta &= \partial_t \dot{x}^\theta + \dot{x}^k \partial_k \dot{x}^\theta + \frac{2}{r} \dot{x}^r \dot{x}^\theta - \sin \theta \cos \theta (\dot{x}^\varphi)^2, \\ \ddot{x}^\varphi &= \partial_t \dot{x}^\varphi + \dot{x}^k \partial_k \dot{x}^\varphi + \frac{2}{r} \dot{x}^r \dot{x}^\varphi + 2 \cot \theta \dot{x}^\theta \dot{x}^\varphi. \end{aligned} \right\} \quad (6A.7)$$

Physical components of the divergence of a symmetric second-order tensor field T :

$$\left. \begin{aligned} (\operatorname{div} T)\langle r \rangle &= \partial_r T\langle rr \rangle + \frac{1}{r} \partial_\theta T\langle r\theta \rangle + \frac{1}{r \sin \theta} \partial_\varphi T\langle r\varphi \rangle + \\ &\quad + \frac{1}{r} [2 T\langle rr \rangle - T\langle \theta\theta \rangle - T\langle \varphi\varphi \rangle + \cot \theta T\langle r\theta \rangle], \\ (\operatorname{div} T)\langle \theta \rangle &= \partial_r T\langle r\theta \rangle + \frac{1}{r} \partial_\theta T\langle \theta\theta \rangle + \frac{1}{r \sin \theta} \partial_\varphi T\langle \theta\varphi \rangle + \\ &\quad + \frac{1}{r} [3 T\langle r\theta \rangle + \cot \theta (T\langle \theta\theta \rangle - T\langle \varphi\varphi \rangle)], \\ (\operatorname{div} T)\langle \varphi \rangle &= \partial_r T\langle r\varphi \rangle + \frac{1}{r} \partial_\theta T\langle \theta\varphi \rangle + \frac{1}{r \sin \theta} \partial_\varphi T\langle \varphi\varphi \rangle + \\ &\quad + \frac{1}{r} [3 T\langle r\varphi \rangle + 2 \cot \theta T\langle \theta\varphi \rangle]. \end{aligned} \right\} \quad (6A.8)$$

A6.16.4

Supplementary Reading

Same as for § A6.11

Chapter A8

Isotropic Tensors and Isotropic Tensor Functions

Various isotropic tensors and isotropic tensor-valued functions of tensors appear repeatedly in continuum mechanics, especially in the theories of particular classes of materials. It is the purpose of this chapter to lay down precise definitions and to establish useful representation theorems.

A8.1. Isotropic Tensors

In the older literature an isotropic (m^{th} -order) tensor was generally defined to be one whose components were the same with respect to all orthonormal bases, i.e., the components were invariant under changes of orthonormal bases. We shall use this approach to motivate an alternative definition which agrees with the manner in which the notion of isotropy arises naturally in the newer literature. We proceed by considering tensors of the first few lower orders individually.

Zero-order tensors or scalars may be thought of as having one component that is the same for all bases.¹ Thus, all scalars are isotropic; alternatively, one may take the point of view that the concept of isotropy does not arise in the context of scalars.

Next, consider a first-order tensor or vector $\underline{u}$ in the n -dimensional inner product space V_n . Let $\{\underline{e}_{\langle 1 \rangle}, \underline{e}_{\langle 2 \rangle}, \dots, \underline{e}_{\langle n \rangle}\}$ and $\{\bar{\underline{e}}_{\langle 1 \rangle}, \bar{\underline{e}}_{\langle 2 \rangle}, \dots, \bar{\underline{e}}_{\langle n \rangle}\}$ be two orthonormal bases for V_n , so that $\underline{u}$ admits the representations

$$\underline{u} = u^{\langle i \rangle} \underline{e}_{\langle i \rangle} = \bar{u}^{\langle i \rangle} \bar{\underline{e}}_{\langle i \rangle} \quad . \quad)^2$$

¹ See the discussion in §A4.

² Cf. Theorem A2.5.3.

By Theorem A3.10.8, $\exists$ a unique orthogonal tensor $\underline{Q} \ni$ the orthonormal bases are related by

$$\underline{\bar{e}}_{\langle i \rangle} = \underline{Q} \underline{e}_{\langle i \rangle} ; \quad)^1$$

and then by Theorem A2.6.2, the components of $\underline{u}$ are related by

$$\bar{u}^{\langle i \rangle} = \sum_{j=1}^n Q_{\langle j \rangle \langle i \rangle} u^{\langle j \rangle}.$$

In direct notation, this is equivalent (the details are left as Exercise A8.1.1) to

$$\underline{v} = \underline{Q}^T \underline{u},$$

where $\underline{v}$ is the vector

$$\underline{v} = \bar{u}^{\langle i \rangle} \underline{\bar{e}}_{\langle i \rangle}.$$

If $\underline{u}$ is isotropic in the sense that $\bar{u}^{\langle i \rangle} = u^{\langle i \rangle}$, then

$$\underline{v} = u^{\langle i \rangle} \underline{\bar{e}}_{\langle i \rangle} = \underline{u}. \text{ Thus,}$$

$$\underline{u} = \underline{Q}^T \underline{u};$$

conversely, if this last equation holds, then $\bar{u}^{\langle i \rangle} = u^{\langle i \rangle}$. Furthermore, we note that since $\underline{Q}^T = \underline{Q}^{-1}$ (Theorem A3.10.2), the statement $\underline{u} = \underline{Q}^T \underline{u}$ is equivalent to

$$\underline{Q} \underline{u} = \underline{u}.$$

These remarks lead us to

Moreover, $\{\underline{e}_{\langle 1 \rangle}, \underline{e}_{\langle 2 \rangle}, \dots, \underline{e}_{\langle n \rangle}\}$ and $\{\underline{\bar{e}}_{\langle 1 \rangle}, \underline{\bar{e}}_{\langle 2 \rangle}, \dots, \underline{\bar{e}}_{\langle n \rangle}\}$ are like-handed iff $\underline{Q} \in \text{Orth}^+ \mathcal{V}_n$.

Definition A8.1.1. A vector $\underline{u} \in V_n$ is isotropic if

$$\underline{Q}\underline{u} = \underline{u} \quad \forall \underline{Q} \in \text{Orth } V_n;$$

$\underline{u}$ is hemitropic¹ if

$$\underline{Q}\underline{u} = \underline{u} \quad \forall \underline{Q} \in \text{Orth}^+ V_n.$$

Since orthogonal tensors preserve magnitudes (Theorem A3.10.1), $\underline{Q}\underline{u}$ has the same length as $\underline{u}$; and geometrically, the operation $\underline{u} \mapsto \underline{Q}\underline{u}$ can be thought of as "rotating" the vector $\underline{u}$ into the vector $\underline{Q}\underline{u}$. Thus, isotropic vectors have the property that rotations change neither their lengths nor their directions.)² From this point of view the following result₁ is not surprising.

Theorem A8.1.1. If a vector $\underline{u} \in V_n$ is either hemitropic or isotropic, then $\underline{u} = \underline{0}$. ((n ≥ 2))

Conversely, $\underline{0}$ is isotropic and hence hemitropic.

Proof. Assume that $\underline{u}$ is hemitropic so that

¹ Sometimes isotropic and then an isotropic is referred to as isotropic with a center of symmetry. Needless to say, care must be taken in drawing on results from the literature in this area.

² Cf. Theorem A2.3.13.

$$\underline{Q}\underline{u} = \underline{u} \quad \forall \underline{Q} \in \text{Orth}^+ V_n.$$

Suppose that $\underline{u} \neq \underline{0}$. Then

$$\underline{e}_{\langle 1 \rangle} := \frac{1}{|\underline{u}|} \underline{u}$$

has the property that $|\underline{e}_{\langle 1 \rangle}| = 1$. Since $\underline{e}_{\langle 1 \rangle} \neq \underline{0}$, the singleton $\{\underline{e}_{\langle 1 \rangle}\}$ is linearly independent (Exercise A2.2.1).

Consequently, by Theorem A2.2.9, $\{\underline{e}_{\langle 1 \rangle}\}$ can be extended to a basis for V_n , which can be used to construct via the proof of Theorem A2.3.15 an orthonormal basis

$\{\underline{e}_{\langle 1 \rangle}, \underline{e}_{\langle 2 \rangle}, \dots, \underline{e}_{\langle n \rangle}\}$ whose first element is $\underline{e}_{\langle 1 \rangle}$.

Obviously, the sets $\{\underline{e}_{\langle 1 \rangle}, \underline{e}_{\langle 2 \rangle}, \dots, \underline{e}_{\langle n \rangle}\}$ and $\{\tilde{\underline{e}}_{\langle 1 \rangle}, \tilde{\underline{e}}_{\langle 2 \rangle}, \dots, \tilde{\underline{e}}_{\langle n \rangle}\}$ defined by

$$\underline{\bar{e}}_{\langle 1 \rangle} = \tilde{\underline{e}}_{\langle 1 \rangle} = \underline{e}_{\langle 2 \rangle},$$

$$\underline{\bar{e}}_{\langle 2 \rangle} = \underline{e}_{\langle 1 \rangle}, \quad \tilde{\underline{e}}_{\langle 2 \rangle} = -\underline{e}_{\langle 1 \rangle},$$

$$\underline{\bar{e}}_{\langle i \rangle} = \tilde{\underline{e}}_{\langle i \rangle} = \underline{e}_{\langle i \rangle} \quad \text{for } i = 3, 4, \dots, n$$

are both also orthonormal bases for V_n . Let Δ denote a choice of volume-orientation functional for V_n .¹ Then by the multilinearity², we have

¹ Cf. Theorem A2.7.2.

² See Definition A2.7.3.

$$\Delta(\bar{e}_{\langle 1 \rangle}, \bar{e}_{\langle 2 \rangle}, \dots, \bar{e}_{\langle n \rangle}) = -\Delta(\tilde{e}_{\langle 1 \rangle}, \tilde{e}_{\langle 2 \rangle}, \dots, \tilde{e}_{\langle n \rangle}).$$

Consequently, the basis $\{\underline{e}_{\langle 1 \rangle}, \underline{e}_{\langle 2 \rangle}, \dots, \underline{e}_{\langle n \rangle}\}$ is like-handed¹ with either the basis $\{\bar{e}_{\langle 1 \rangle}, \bar{e}_{\langle 2 \rangle}, \dots, \bar{e}_{\langle n \rangle}\}$ or the basis $\{\tilde{e}_{\langle 1 \rangle}, \tilde{e}_{\langle 2 \rangle}, \dots, \tilde{e}_{\langle n \rangle}\}$. For definiteness, suppose that $\{\underline{e}_{\langle 1 \rangle}, \underline{e}_{\langle 2 \rangle}, \dots, \underline{e}_{\langle n \rangle}\}$ and $\{\bar{e}_{\langle 1 \rangle}, \bar{e}_{\langle 2 \rangle}, \dots, \bar{e}_{\langle n \rangle}\}$ are like-handed. Then by Theorem A3.10.8 $\exists$ a unique $Q \in \text{Orth}^+ V_n \ni$

$$Q \underline{e}_{\langle i \rangle} = \bar{e}_{\langle i \rangle}, \quad i \in \mathbb{I}_n.$$

In particular,

$$Q \underline{e}_{\langle 1 \rangle} = \bar{e}_{\langle 1 \rangle} = \underline{e}_{\langle 2 \rangle}.$$

But

$$Q \underline{e}_{\langle 1 \rangle} = Q\left(\frac{1}{|\underline{u}|} \underline{u}\right) \quad (\text{definition of } \underline{e}_{\langle 1 \rangle})$$

$$= \frac{1}{|\underline{u}|} Q \underline{u} \quad (\text{linearity of } Q)$$

$$= \frac{1}{|\underline{u}|} \underline{u} \quad (\text{hemitropy of } \underline{u})$$

$$= \underline{e}_{\langle 1 \rangle} \quad (\text{definition of } \underline{e}_{\langle 1 \rangle}).$$

Thus, we have the contradiction that $\underline{e}_{\langle 1 \rangle} = \underline{e}_{\langle 2 \rangle}$; and $\therefore \underline{u} = \underline{0}$.

If $\underline{u}$ is isotropic, it is necessarily hemitropic; and from the above argument, $\underline{u} = \underline{0}$.

Since all tensors take $\underline{Q}$ into $\underline{Q}$, the converse is obvious. $\square$

Exercise A8.1.2. The above demonstration is clearly restricted to the case of $n \geq 2$. Consider the $n=1$ case.

¹ See Definition A2.7.6.

For second-order tensors, the considerations are similar. Let $\{\underline{e}_{\langle 1 \rangle}, \underline{e}_{\langle 2 \rangle}, \dots, \underline{e}_{\langle n \rangle}\}$ and $\{\bar{\underline{e}}_{\langle 1 \rangle}, \bar{\underline{e}}_{\langle 2 \rangle}, \dots, \bar{\underline{e}}_{\langle n \rangle}\}$ be two orthonormal bases for V_n so that $\underline{T} \in \text{Lin } V_n$ admits the representations

$$\underline{T} = T^{\langle ij \rangle} \underline{e}_{\langle i \rangle} \otimes \underline{e}_{\langle j \rangle} = \bar{T}^{\langle ij \rangle} \bar{\underline{e}}_{\langle i \rangle} \otimes \bar{\underline{e}}_{\langle j \rangle} \quad . \quad)'$$

By Theorems A3.3.1 and A3.10.8,

$$\bar{T}^{\langle ij \rangle} = \sum_{k=1}^n \sum_{l=1}^n Q_{\langle ki \rangle} Q_{\langle lj \rangle} T^{\langle kl \rangle} ,$$

where $\underline{Q}$ is the orthogonal tensor connecting the bases ($\bar{\underline{e}}_{\langle i \rangle} = \underline{Q} \underline{e}_{\langle i \rangle}$). In direct notation, this is equivalent to

$$\underline{\bar{S}} = \underline{Q}^T \underline{T} \underline{Q} ,$$

where $\underline{\bar{S}}$ is the tensor

$$\underline{\bar{S}} = \bar{T}^{\langle ij \rangle} \bar{\underline{e}}_{\langle i \rangle} \otimes \bar{\underline{e}}_{\langle j \rangle} .$$

For $\underline{T}$ isotropic in the sense that $\bar{T}^{\langle ij \rangle} = T^{\langle ij \rangle}$, we get

$$\underline{T} = \underline{Q}^T \underline{T} \underline{Q} ,$$

which is more conveniently expressed as

$$\underline{Q} \underline{T} \underline{Q}^T = \underline{T} .$$

Thus, we are lead to

¹ See Theorem A3.2.2.

Definition A8.1.2. A second-order tensor $\underline{I} \in \text{Lin } \mathcal{V}_n$ is isotropic if

$$\underline{Q} \underline{I} \underline{Q}^T = \underline{I} \quad \forall \underline{Q} \in \text{Orth } \mathcal{V}_n;$$

$\underline{I}$ is hemitropic¹ if

$$\underline{Q} \underline{I} \underline{Q}^T = \underline{I} \quad \forall \underline{Q} \in \text{Orth}^+ \mathcal{V}_n.$$

The next two theorems show that the notion of isotropy has some nontrivial content in the context of second-order tensors.

Theorem A8.1.2. If a symmetric second-order tensor $\underline{S} \in \text{Sym } \mathcal{V}_n$ is either hemitropic or isotropic, then it has the form

$$\underline{S} = \sigma \underline{I}, \quad \sigma \in \mathbb{R}. \quad)^2$$

Conversely, a second-order tensor of this form is isotropic and hence hemitropic.

Proof.³ Assume that $\underline{S}$ is hemitropic so that

¹ Sometimes isotropic and then our isotropic is called isotropic with a center of symmetry. Again, caution is advised when drawing on results from the literature.

² Often tensors of this form are said to be spherical or hydrostatic.

³ This proof was shown to the author by Dr. LAYACHI HADJI in 1986.

$$\underline{\underline{Q}} \underline{\underline{S}} \underline{\underline{Q}}^T = \underline{\underline{S}} \quad \forall \quad \underline{\underline{Q}} \in \text{Orth}^+ V_n;$$

equivalently,

$$\underline{\underline{Q}} \underline{\underline{S}} = \underline{\underline{S}} \underline{\underline{Q}} \quad \forall \quad \underline{\underline{Q}} \in \text{Orth}^+ V_n;$$

i.e., $\underline{\underline{S}}$ commutes with every proper orthogonal tensor. Since $\underline{\underline{S}}$ is symmetric, it has an eigenpair $(\sigma, \underline{\underline{e}})$ by the Spectral Theorem (Theorem A1.13.1):

$$\underline{\underline{S}} \underline{\underline{e}} = \sigma \underline{\underline{e}}, \quad |\underline{\underline{e}}| = 1. \quad)^1$$

Then

$$(\underline{\underline{S}} \underline{\underline{Q}}) \underline{\underline{e}} = (\underline{\underline{Q}} \underline{\underline{S}}) \underline{\underline{e}} \quad (\text{hemitropy and tensor equality})$$

$$= \underline{\underline{Q}} (\underline{\underline{S}} \underline{\underline{e}}) \quad (\text{tensor multiplication})$$

$$= \underline{\underline{Q}} (\sigma \underline{\underline{e}}) \quad ((\sigma, \underline{\underline{e}}) \text{ an eigenpair of } \underline{\underline{S}})$$

$$= \sigma (\underline{\underline{Q}} \underline{\underline{e}}) \quad (\text{linearity of } \underline{\underline{Q}})$$

$$= \sigma \underline{\underline{I}} (\underline{\underline{Q}} \underline{\underline{e}}) \quad (\text{definition of identity tensor})$$

$$= (\sigma \underline{\underline{I}}) (\underline{\underline{Q}} \underline{\underline{e}}) \quad (\text{definition of scalar multiple of a tensor});$$

and by application of the definition of tensor multiplication to the l.h.s., we have

$$\underline{\underline{S}} (\underline{\underline{Q}} \underline{\underline{e}}) = (\sigma \underline{\underline{I}}) (\underline{\underline{Q}} \underline{\underline{e}}) \quad \forall \quad \underline{\underline{Q}} \in \text{Orth}^+ V_n.$$

Now the vector $\underline{\underline{e}}$ is fixed and of unit length, but by

¹Here, for later convenience, we have normalized $\underline{\underline{e}}$; cf. Theorem A3.11.2 and the subsequent discussion.

the arbitrariness of $\underline{Q} \in \text{Orth}^+ V_n$, $\underline{Q}\underline{e}$ may be viewed as an arbitrary unit vector in V_n . (Establish this fact as Exercise A8.1.3.) Then by Theorem A3.1.2, we conclude that

$$\underline{S} = \sigma \underline{I}.$$

If $\underline{S}$ is isotropic, then it is necessarily hemitropic; and from the argument above, $\underline{S} = \sigma \underline{I}$.

The truth of the converse assertion follows from a trivial calculation. $\square$

The restriction to symmetric tensors in the previous theorem is easily removed if the dimension of the underlying linear space is odd.

Theorem A8.1.3. If a second-order tensor $\underline{T} \in \text{Lin } V_n$ (n odd) is either hemitropic or isotropic, then it has the form

$$\underline{T} = \tau \underline{I}, \quad \tau \in \mathbb{R}.$$

Conversely, a second-order tensor of this form is isotropic and hence hemitropic.

Proof. Consider the proof of the previous theorem. The symmetry of $\underline{S}$ was used only to guarantee that $\underline{S}$ possess an eigenpair. This is also true of $\underline{T} \in \text{Lin } V_n$ when n is odd. To see this, we

note that by Theorem A3.11.7 a necessary and sufficient condition for τ to be an eigenvalue of $\underline{T}$ is that

$$\det(\underline{T} - \tau \underline{I}) = 0;$$

i.e., τ must be a real root of the characteristic polynomial, which by Theorem A3.11.8 is a polynomial of degree n . By the Fundamental Theorem of Algebra, an n^{th} degree polynomial has exactly n roots (generally complex). But here the coefficients in the polynomial are real, so any complex roots occur in conjugate pairs; i.e., if p is a root, then so is $\bar{p}$. Then since n is odd, there must be at least one real root. Thus, there is an eigenpair, and the previous proof applies. $\square$

It is interesting to note that in the above theorem the hemitropic or isotropic tensor $\underline{T} \in \text{Lin } V_n$ (n odd) turns out to have the form $\underline{T} = \tau \underline{I}$. Thus, $\underline{T}$ is, in fact, symmetric! We can see what is going on here by means of

Theorem A8.1.4. If a second-order tensor $\underline{T} \in \text{Lin } V_n$ is isotropic (hemitropic), then its symmetric and skew parts are each isotropic (hemitropic).

Proof. Exercise A8.1.4. (Hint: Take the transpose of the defining statement $\underline{Q} \underline{T} \underline{Q}^T = \underline{T}$ and add and subtract the result to $\underline{Q} \underline{T} \underline{Q}^T = \underline{T}$.) $\square$

In view of Theorems A8.1.3 and 4, the next result is not surprising.

Theorem A8.1.5. If a skew tensor $\underline{W} \in \text{Skw } \mathcal{V}_n$ (n odd) is either hemitropic or isotropic, then $\underline{W} = \underline{0}$.

Conversely, $\underline{0}$ is isotropic and hence hemitropic.

Proof. Suppose that $\underline{W}$ is either hemitropic or isotropic. By Theorem A8.1.3, $\underline{W}$ has the form

$$\underline{W} = \omega \underline{I}, \quad \omega \in \mathbb{R}.$$

But then

$$\underline{W}^T = \omega \underline{I},$$

which is consistent with $\underline{W}^T = -\underline{W}$ iff $\omega = 0$, i.e. $\underline{W} = \underline{0}$.

The converse is trivial. $\square$

For an alternative proof, we could have proceeded as in the proof of Theorem A8.1.3 by starting with an eigenpair for $\underline{W}$, the existence of which is guaranteed by Theorem A3.11.11 — with the eigenvalue being 0.

Supplementary Reading

JEFFRIES, Cartesian Tensors

PRAGER, Introduction to Mechanics of Continua

AB.2. Invariant Sets and Functions

As preparation for the concept of isotropic functions, we need

Definition AB.2.1. Let $\mathcal{G} \subset \text{Orth } \mathcal{V}_n$. Then a set $\mathcal{U} \subset \mathcal{V}_n$ is invariant under $\mathcal{G}$ if

$$\underline{u} \in \mathcal{U}, \underline{Q} \in \mathcal{G} \Rightarrow \underline{Q}\underline{u} \in \mathcal{U},$$

and a set $\mathcal{Q} \subset \text{Lin } \mathcal{V}_n$ is invariant under $\mathcal{G}$ if

$$\underline{A} \in \mathcal{Q}, \underline{Q} \in \mathcal{G} \Rightarrow \underline{Q}\underline{A}\underline{Q}^T \in \mathcal{Q}.$$

Occasionally, the following trivial observation will be useful.

Theorem AB.2.1. If a set $\mathcal{U} \subset \mathcal{V}_n$ ($\mathcal{Q} \subset \text{Lin } \mathcal{V}_n$) is invariant under $\mathcal{G} \subset \text{Orth } \mathcal{V}_n$, then it is invariant under any $\mathcal{G}' \subset \mathcal{G}$.

Many of the sets of vectors and tensors that arise naturally in continuum physics are invariant under the full orthogonal group.

Theorem AB.2.2. The sets $\mathcal{V}_n$, $\{\underline{x} \in \mathcal{V}_n : \underline{x} \neq \underline{0}\} \subset \mathcal{V}_n$ and $\text{Lin } \mathcal{V}_n$, $\text{Inv } \mathcal{V}_n$, $\text{Inv}^+ \mathcal{V}_n$, $\text{Unim } \mathcal{V}_n$, $\text{Unim}^+ \mathcal{V}_n$, $\text{Orth } \mathcal{V}_n$, $\text{Sym } \mathcal{V}_n$, $\text{Skw } \mathcal{V}_n$, $\text{Psym } \mathcal{V}_n \subset \text{Lin } \mathcal{V}_n$ are invariant under $\text{Orth } \mathcal{V}_n$.

Proof. For the most part, the theorem follows from straight forward calculations. The most difficult cases are $\{ \underline{v} \in \underline{V}_n : \underline{v} \neq \underline{0} \}$ and $\text{Psym} \underline{V}_n$; they constitute Exercise AB.2.1. $\square$

Turning now to invariant functions of a vector, we have

Definition AB.2.2. Let $\mathcal{U} \subset \underline{V}_n$ be invariant under $\mathcal{G} \subset \text{Orth } \underline{V}_n$. Then a scalar-valued function $f: \mathcal{U} \rightarrow \mathbb{R}$ is invariant under $\mathcal{G}$ if

$$f(\underline{Q}\underline{u}) = f(\underline{u}) \quad \forall \underline{u} \in \mathcal{U} \text{ and } \forall \underline{Q} \in \mathcal{G};$$

a vector-valued function $\underline{f}: \mathcal{U} \rightarrow \underline{V}_n$ is invariant under $\mathcal{G}$ if

$$\underline{f}(\underline{Q}\underline{u}) = \underline{Q}\underline{f}(\underline{u}) \quad \forall \underline{u} \in \mathcal{U} \text{ and } \forall \underline{Q} \in \mathcal{G};$$

and a second-order tensor-valued function $\underline{F}: \mathcal{U} \rightarrow \text{Lin } \underline{V}_n$ is invariant under $\mathcal{G}$ if

$$\underline{F}(\underline{Q}\underline{u}) = \underline{Q}\underline{F}(\underline{u})\underline{Q}^T \quad \forall \underline{u} \in \mathcal{U} \text{ and } \forall \underline{Q} \in \mathcal{G}.$$

Note that these definitions would be nonsensical if the domains of the various functions were not invariant.

Similarly, for invariant functions of a tensor, we have

Definition AB.2.3. Let $\mathcal{Q} \subset \text{Lin } V_n$ be invariant under
 $\mathcal{H} \subset \text{Orth } V_n$. Then a scalar-valued function $g: \mathcal{Q} \rightarrow \mathbb{R}$ is
invariant under $\mathcal{H}$ if

$$g(\underline{\underline{Q}} \underline{\underline{A}} \underline{\underline{Q}}^T) = g(\underline{\underline{A}}) \quad \forall \underline{\underline{A}} \in \mathcal{Q} \text{ and } \forall \underline{\underline{Q}} \in \mathcal{H};$$

a vector-valued function $g: \mathcal{Q} \rightarrow V_n$ is invariant under $\mathcal{H}$ if

$$\underline{\underline{g}}(\underline{\underline{Q}} \underline{\underline{A}} \underline{\underline{Q}}^T) = \underline{\underline{Q}} \underline{\underline{g}}(\underline{\underline{A}}) \quad \forall \underline{\underline{A}} \in \mathcal{Q} \text{ and } \forall \underline{\underline{Q}} \in \mathcal{H};$$

and a second-order tensor-valued function $\underline{\underline{G}}: \mathcal{Q} \rightarrow \text{Lin } V_n$ is
invariant under $\mathcal{H}$ if

$$\underline{\underline{G}}(\underline{\underline{Q}} \underline{\underline{A}} \underline{\underline{Q}}^T) = \underline{\underline{Q}} \underline{\underline{G}}(\underline{\underline{A}}) \underline{\underline{Q}}^T \quad \forall \underline{\underline{A}} \in \mathcal{Q} \text{ and } \forall \underline{\underline{Q}} \in \mathcal{H}.$$

Again, analogous to Theorem AB.2.1, we trivially have

Theorem AB.2.3. Let the set $\mathcal{U} \subset V_n$ ($\mathcal{Q} \subset \text{Lin } V_n$) be
invariant under $\mathcal{H} \subset \text{Orth } V_n$. If a scalar-, vector-, or second-
order tensor-valued function on $\mathcal{U}$ ($\mathcal{Q}$) is invariant under
 $\mathcal{H}$, then it is invariant under any $\mathcal{H}' \subset \mathcal{H}$.

Clearly, the co-domains of the functions in Definitions AB.2.2 and 3 can be restricted without any adjustments to the definitions. For example, if $\mathcal{Q} \subset \text{Lin } V_n$ is invariant under $\mathcal{H} \subset \text{Orth } V_n$, then the symmetric second-order tensor-valued function $\underline{\underline{G}}': \mathcal{Q} \rightarrow \text{Sym } V_n$ is invariant under $\mathcal{H}$ if

$$\underline{\underline{G}}'(\underline{\underline{Q}} \underline{\underline{A}} \underline{\underline{Q}}^T) = \underline{\underline{Q}} \underline{\underline{G}}'(\underline{\underline{A}}) \underline{\underline{Q}}^T \quad \forall \underline{\underline{A}} \in \mathcal{Q} \text{ and } \forall \underline{\underline{Q}} \in \mathcal{H}.$$

When the various branches of continuum physics are combined, as for example in continuum thermo-mechanics, one often needs certain generalizations and hybrids of Definitions A.B.2.2 and 3. For example, in thermoelasticity, we shall need

Definition A.B.2.4. Let $\mathcal{U} \subset \mathcal{V}_n$ and $\mathcal{Q} \subset \text{Lin } \mathcal{V}_n$ be invariant under $\mathcal{G} \subset \text{Orth } \mathcal{V}_n$. Then a scalar-valued function $h: \mathcal{U} \times \mathcal{Q} \rightarrow \mathbb{R}$ is invariant under $\mathcal{G}$ if

$$h(\underline{Q}\underline{u}, \underline{Q}\underline{A}\underline{Q}^T) = h(\underline{u}, \underline{A}) \quad \forall \underline{u} \in \mathcal{U}, \forall \underline{A} \in \mathcal{Q}, \text{ and } \forall \underline{Q} \in \mathcal{G};$$

a vector-valued function $\underline{h}: \mathcal{U} \times \mathcal{Q} \rightarrow \mathcal{V}_n$ is invariant under $\mathcal{G}$ if

$$\underline{h}(\underline{Q}\underline{u}, \underline{Q}\underline{A}\underline{Q}^T) = \underline{Q}\underline{h}(\underline{u}, \underline{A}) \quad \forall \underline{u} \in \mathcal{U}, \forall \underline{A} \in \mathcal{Q}, \text{ and } \forall \underline{Q} \in \mathcal{G};$$

and a second-order tensor-valued function $\underline{H}: \mathcal{U} \times \mathcal{Q} \rightarrow \text{Lin } \mathcal{V}_n$ is invariant under $\mathcal{G}$ if

$$\underline{H}(\underline{Q}\underline{u}, \underline{Q}\underline{A}\underline{Q}^T) = \underline{Q}\underline{H}(\underline{u}, \underline{A})\underline{Q}^T \quad \forall \underline{u} \in \mathcal{U}, \forall \underline{A} \in \mathcal{Q}, \text{ and } \forall \underline{Q} \in \mathcal{G}.$$

When the set $\mathcal{G} \subset \text{Orth } V_n$ in Definitions A8.2.2 - 4 is the full orthogonal group $\text{Orth } V_n$, the functions are said to be isotropic. The remaining sections of this chapter are devoted to establishing representation theorems for various isotropic functions.

Supplementary Reading

GURTIN, An Introduction to Continuum Mechanics

TRUESDELL and NOLL, "The non-linear field theories of mechanics",
Handbuch der Physik, vol III/3 (edit. S. Flügge) 1965

A8.3. Isotropic Scalar-Valued Functions of a Vector

Formalizing the remark which ended the previous section, we have

Definition A8.3.1. Let $\mathcal{U} \subset V_n$ be invariant under $\text{Orth } V_n$ ($\text{Orth}^+ V_n$). Then a scalar-valued function of a vector, $f: \mathcal{U} \rightarrow \mathbb{R}$, is isotropic (hemitropic¹) if it is invariant under $\text{Orth } V_n$ ($\text{Orth}^+ V_n$); i.e., if

$$f(\underline{Qu}) = f(\underline{u}) \quad \forall \underline{u} \in \mathcal{U} \text{ and } \forall Q \in \text{Orth } V_n \text{ (Orth}^+ V_n \text{)}.$$

As was observed in the paragraph preceding Theorem A8.1.1, orthogonal tensors preserve lengths, and thus the operation $\underline{u} \mapsto \underline{Qu}$ may be viewed as "rotating" $\underline{u}$ into $\underline{Qu}$. Then since an isotropic function $f: \mathcal{U} \rightarrow \mathbb{R}$ has the property that $f(\underline{Qu}) = f(\underline{u})$, it is insensitive to the direction of its argument; and we are lead to the following representation theorem.

Theorem A8.3.1. Let $\mathcal{U} \subset V_n$ ($n \geq 2$) be invariant under $\text{Orth } V_n$. If a scalar-valued function of a vector, $f: \mathcal{U} \rightarrow \mathbb{R}$, is either hemitropic or isotropic, then $\exists$ a unique function $\psi: \mathcal{U} \rightarrow \mathbb{R}$, where $\mathcal{U} := \{|\underline{u}| : \underline{u} \in \mathcal{U}\}$, $\ni$

¹Sometimes isotropic and then our isotropic is referred to as isotropic with a center of symmetry. As was the case throughout §B.1, care must be taken in drawing on results from the literature in this area.

$$f(\underline{u}) = \vartheta(|\underline{u}|), \quad \underline{u} \in \mathcal{U}. \quad)^1$$

Conversely, a function of this form is isotropic and hence² hemitropic.

Proof. Assume that f is hemitropic so that

$$f(Q\underline{u}) = f(\underline{u}) \quad \forall \underline{u} \in \mathcal{U} \text{ and } \forall Q \in \text{Orth}^+ \mathcal{V}_n.$$

To establish the representation, it suffices to prove that

$$|\underline{\bar{u}}| = |\underline{u}| \Rightarrow f(\underline{\bar{u}}) = f(\underline{u});$$

here, of course, $\underline{u}, \underline{\bar{u}} \in \mathcal{U} \subset \mathcal{V}_n$. Suppose then that $|\underline{\bar{u}}| = |\underline{u}|$.

If $|\underline{\bar{u}}| = |\underline{u}| = 0$, then $\underline{\bar{u}} = \underline{u} = \underline{0}$; ³ and trivially, $f(\underline{\bar{u}}) = f(\underline{u})$.

If $|\underline{\bar{u}}| = |\underline{u}| \neq 0$, then

$$\underline{e}_{\langle 1 \rangle} := \frac{\underline{u}}{|\underline{u}|} \quad \text{and} \quad \underline{\bar{e}}_{\langle 1 \rangle} := \frac{\underline{\bar{u}}}{|\underline{\bar{u}}|}$$

are unit vectors; and as in the proof of Theorem 8.1.1, the

¹We have taken $\mathcal{U}$ to be invariant under the full orthogonal group $\text{Orth } \mathcal{V}_n$ so as to be able to treat isotropy and hemitropy together; obviously, by separating the treatments, we could get slightly sharper results.

²Gf. Theorem A8.2.3.

³See Theorem A2.3.6 (ii).

singletons $\{e_{\langle 1 \rangle}\}$ and $\{\bar{e}_{\langle 1 \rangle}\}$ can be extended to like-handed orthonormal bases $\{e_{\langle 1 \rangle}, e_{\langle 2 \rangle}, \dots, e_{\langle n \rangle}\}$ and $\{\bar{e}_{\langle 1 \rangle}, \bar{e}_{\langle 2 \rangle}, \dots, \bar{e}_{\langle n \rangle}\}$ for V_n . Moreover, $\exists Q \in \text{Orth}^+ V_n \ni$

$$\bar{e}_{\langle i \rangle} = Q e_{\langle i \rangle}, \quad i \in \Pi_n.$$

In particular,

$$\bar{e}_{\langle 1 \rangle} = Q e_{\langle 1 \rangle},$$

or

$$\frac{\bar{u}}{|\bar{u}|} = Q \left(\frac{u}{|u|} \right)$$

$$= \frac{1}{|u|} Q u \quad (\text{linearity of } Q)$$

$\Rightarrow$:

$$\bar{u} = Q u \quad (|\bar{u}| = |u|).$$

Then by the assumed hermitropy,

$$f(\bar{u}) = f(Q u) = f(u),$$

and the validity of the representation is proven.

Suppose that $\hat{\vartheta}: \mathcal{U} \rightarrow \mathbb{R}$ is another function with the same property. Then

$$f(u) = \vartheta(|u|) = \hat{\vartheta}(|u|), \quad u \in \mathcal{U}.$$

$\therefore \hat{\vartheta}$ and ϑ agree on $\mathcal{U}$; i.e., $\hat{\vartheta} = \vartheta$.

The truth of the converse assertion follows from a
an easy calculation, the details of which are left as
Exercise A8.3.1. $\square$

Supplementary Reading

same as for $\S$ A8.2 plus

NOLL, "Representations of certain isotropic tensor functions", Archiv
der Mathematik 21 (1970) 87-90.

WANG, "On representations for isotropic functions. Part I. Isotropic
functions of symmetric tensors and vectors", Arch. Rational
Mech. Anal. 33 (1969) 249-267.

——, "On representations for isotropic functions. Part II. Isotropic
functions of skew-symmetric tensors, symmetric tensors, and
vectors", Arch. Rational Mech. Anal. 33 (1969) 268-287.

SMITH, "On a fundamental error in two papers of C.-C. Wang 'On
representations for isotropic functions, Parts I and II'",
Arch. Rational Mech. Anal. 36 (1970) 161-165.

WANG, "A new representation theorem for isotropic functions: An answer
to G.F. Smith's criticism of my papers on representations

for isotropic functions. Part 1. Scalar-valued isotropic functions", Arch. Rational Mech. Anal. 36 (1970) 166-197.

—, "A new representation theorem for isotropic functions: An answer to G. F. Smith's criticism of my papers on representations for isotropic functions. Part 2. Vector-valued isotropic functions, symmetric tensor-valued isotropic functions, and skew-symmetric tensor-valued isotropic functions", Arch. Rational Mech. Anal. 36 (1970) 198-223.

A8.4. Isotropic Scalar-Valued Functions of a Tensor

Again, formalizing the remark which ended § A8.2, we have

Definition A8.4.1 Let $\mathcal{Q} \subset \text{Lin } V_n$ be invariant under
 $\text{Orth } V_n$ ($\text{Orth}^+ V_n$). Then a scalar-valued function of a second-
order tensor, $g: \mathcal{Q} \rightarrow \mathbb{R}$, is isotropic (hemitropic)¹ if it
is invariant under $\text{Orth } V_n$ ($\text{Orth}^+ V_n$); i.e., if

$$g(\underline{Q} \underline{A} \underline{Q}^T) = g(\underline{A}) \quad \forall \underline{A} \in \mathcal{Q} \text{ and } \forall \underline{Q} \in \text{Orth } V_n \text{ (Orth}^+ V_n \text{)}.$$

As has been emphasized repeatedly, in view of the alternative labeling in this area, one must be very careful in drawing on the literature. However, for as the underlying linear space will generally be 3-dimensional, and in this case notions of isotropy and hemitropy for scalar-valued functions of a tensor coincide. Specifically, we have

Theorem A8.4.1. Let $\mathcal{Q} \subset \text{Lin } V_n$, where the dimension n is odd.
Then a scalar-valued function of a tensor, $g: \mathcal{Q} \rightarrow \mathbb{R}$, is
isotropic iff it is hemitropic.

Proof. Suppose that g is isotropic. Then by the above definition, g is invariant under $\text{Orth } V_n$. By Theorem A8.2.3, g is necessarily invariant under the subset $\text{Orth}^+ V_n \subset \text{Orth } V_n$;

¹ Sometimes isotropic and then our isotropic is referred to as isotropic with a center of symmetry.

i.e., g is hemitropic.

Conversely, suppose that g is hemitropic. Then by the above definition g is invariant under $\text{Orth}^+ V_n$; i.e.,

$$\underline{A} \in \mathcal{Q}, \underline{Q} \in \text{Orth}^+ V_n \Rightarrow \underline{Q} \underline{A} \underline{Q}^T \in \mathcal{Q}, g(\underline{Q} \underline{A} \underline{Q}^T) = g(\underline{A}). \quad (*)$$

Now consider any $\underline{A} \in \mathcal{Q}$ and any $\underline{Q} \in \text{Orth} V_n$. For $\underline{Q}$ there are exactly two possibilities: either $\underline{Q} \in \text{Orth}^+ V_n$ or $\underline{Q} \notin \text{Orth}^+ V_n$. If $\underline{Q} \in \text{Orth}^+ V_n$, then by (*)

$$\underline{Q} \underline{A} \underline{Q}^T \in \mathcal{Q} \text{ and } g(\underline{Q} \underline{A} \underline{Q}^T) = g(\underline{A}).$$

If $\underline{Q} \notin \text{Orth}^+ V_n$, then let $\underline{Q} = -I$. In this case, it follows from Theorem A3.8.2 (iv) that

$$\det(-\underline{Q}) = (-1)^n \det \underline{Q} = (-1)^n (-1) = 1.$$

$\therefore -\underline{Q} \in \text{Orth}^+ V_n$ and by (*),

$$(-\underline{Q}) \underline{A} (-\underline{Q})^T \in \mathcal{Q} \text{ and } g((-\underline{Q}) \underline{A} (-\underline{Q})^T) = g(\underline{A}).$$

But

$$(-\underline{Q}) \underline{A} (-\underline{Q})^T = \underline{Q} \underline{A} \underline{Q}^T; \quad)^1$$

so in this case as well,

¹ Cf. Theorems A3.6.3 (iii) and A3.5.1 (iii).

$$\underline{\underline{Q}} \underline{\underline{A}} \underline{\underline{Q}}^T \in \mathcal{Q} \text{ and } g(\underline{\underline{Q}} \underline{\underline{A}} \underline{\underline{Q}}^T) = g(\underline{\underline{A}}).$$

Thus, g is isotropic. $\square$

Important examples of isotropic scalar-valued functions of a tensor are given in the following two theorems.

Theorem A8.4.2. The functions tr and $\det$ on $\text{Lin } \mathcal{V}_n$ are each isotropic.

Proof. Let $\underline{\underline{A}} \in \text{Lin } \mathcal{V}_n$ and $\underline{\underline{Q}} \in \text{Orth } \mathcal{V}_n$. Then

$$\begin{aligned} \text{tr}(\underline{\underline{Q}} \underline{\underline{A}} \underline{\underline{Q}}^T) &= \text{tr}(\underline{\underline{Q}}^T \underline{\underline{Q}} \underline{\underline{A}}) & (\text{tr}(\underline{\underline{S}} \underline{\underline{T}}) &= \text{tr}(\underline{\underline{T}} \underline{\underline{S}})) \\ &= \text{tr} \underline{\underline{A}} & (\underline{\underline{Q}}^T \underline{\underline{Q}} &= \underline{\underline{I}}, \underline{\underline{I}} \underline{\underline{A}} = \underline{\underline{A}}). \end{aligned}$$

$\therefore \text{tr}$ is isotropic.

Again for any $\underline{\underline{A}} \in \text{Lin } \mathcal{V}_n$ and $\underline{\underline{Q}} \in \text{Orth } \mathcal{V}_n$,

$$\begin{aligned} \det(\underline{\underline{Q}} \underline{\underline{A}} \underline{\underline{Q}}^T) &= (\det \underline{\underline{Q}})(\det \underline{\underline{A}})(\det \underline{\underline{Q}}^T) & (\det(\underline{\underline{S}} \underline{\underline{T}}) &= (\det \underline{\underline{S}})(\det \underline{\underline{T}})) \\ &= (\det \underline{\underline{Q}})^2 \det \underline{\underline{A}} & (\det \underline{\underline{I}}^T &= \det \underline{\underline{I}}) \\ &= \det \underline{\underline{A}} & (\det \underline{\underline{Q}} &= \pm 1). \end{aligned}$$

$\therefore \det$ is isotropic. $\square$

Theorem AB.4.3. The fundamental invariants¹ defined
on $\text{Lin } \mathcal{V}_3$ by

$$\zeta_1(\underline{A}) = \text{tr } \underline{A},$$

$$\zeta_2(\underline{A}) = \frac{1}{2} [(\text{tr } \underline{A})^2 - \text{tr } \underline{A}^2],$$

and

$$\zeta_3(\underline{A}) = \det \underline{A}$$

are each isotropic.

Proof. The isotropy of the first and third invariants follows from the previous theorem. The isotropy of the second invariant is left as Exercise AB.4.1. $\square$

The following representation theorem shows that any isotropic scalar-valued function of a symmetric tensor on $\mathcal{V}_3$ is a function of the fundamental invariants of the tensor. This, in fact, is the basis for the terminology "fundamental invariants"; for in older language the theorem says that any invariant of a symmetric tensor (meaning any scalar-valued function of the tensor invariant under $\text{Orth } \mathcal{V}_3$) can be expressed in terms of the fundamental invariants.

Theorem AB.4.4. Let $\mathcal{Q} \subset \text{Sym } \mathcal{V}_3$ be invariant under $\text{Orth } \mathcal{V}_3$. If a scalar-valued function of a symmetric tensor, $g: \mathcal{Q} \rightarrow \mathbb{R}$, is isotropic, then $\exists$ a unique function $\varphi: \mathcal{I}_A \rightarrow \mathbb{R}$, where

¹ Cf. Theorem A3.11.8.

$$\mathcal{I}_A = \{(\zeta_1(\underline{A}), \zeta_2(\underline{A}), \zeta_3(\underline{A})) : \underline{A} \in \mathcal{Q}\},$$

$\Rightarrow$

$$g(\underline{A}) = \varphi(\zeta_1(\underline{A}), \zeta_2(\underline{A}), \zeta_3(\underline{A})).$$

Conversely, a function of this form is isotropic.¹

Proof. Assume that g is isotropic. To establish the validity of the representation, it suffices to show that

$$g(\bar{\underline{A}}) = g(\underline{A})$$

whenever

$$\zeta_1(\bar{\underline{A}}) = \zeta_1(\underline{A}), \zeta_2(\bar{\underline{A}}) = \zeta_2(\underline{A}), \zeta_3(\bar{\underline{A}}) = \zeta_3(\underline{A}); \quad (*)$$

here, of course, $\underline{A}, \bar{\underline{A}} \in \mathcal{Q} \subset \text{Sym } \mathcal{V}_3$. Suppose then that $(*)$ holds. Since $\bar{\underline{A}}$ and $\underline{A}$ have the same fundamental invariants, they have the same characteristic polynomials.² Consequently, they have the same spectra³, say $\{a_1, a_2, a_3\}$. Then $\bar{\underline{A}}$ and $\underline{A}$ have the spectral representations⁴

$$\bar{\underline{A}} = \sum_{i=1}^3 a_i \bar{\underline{e}}_{\langle i \rangle} \otimes \bar{\underline{e}}_{\langle i \rangle} \quad \text{and} \quad \underline{A} = \sum_{i=1}^3 a_i \underline{e}_{\langle i \rangle} \otimes \underline{e}_{\langle i \rangle},$$

where $\{\bar{\underline{e}}_{\langle 1 \rangle}, \bar{\underline{e}}_{\langle 2 \rangle}, \bar{\underline{e}}_{\langle 3 \rangle}\}$ and $\{\underline{e}_{\langle 1 \rangle}, \underline{e}_{\langle 2 \rangle}, \underline{e}_{\langle 3 \rangle}\}$ are orthonormal bases

¹Recall from Theorem AB.4.1 that here ($n=3$) isotropy and hemitropy coincide.

²Cf. Theorem A3.11.8.

³Cf. Definition A3.13.1.

⁴See Theorem A3.13.1.

of eigenvectors of $\bar{A}$ and A , respectively. By Theorem A3.10.8, $\exists Q \in \text{Orth } \mathbb{R}_3 \ni$

$$Q \bar{e}_{\langle i \rangle} = e_{\langle i \rangle}, \quad i \in \mathbb{I}_3.$$

Then

$$\begin{aligned} Q \bar{A} Q^T &= Q \left(\sum_{i=1}^3 a_i \bar{e}_{\langle i \rangle} \otimes \bar{e}_{\langle i \rangle} \right) Q^T \quad (\text{substitution}) \\ &= \sum_{i=1}^3 a_i Q (\bar{e}_{\langle i \rangle} \otimes \bar{e}_{\langle i \rangle}) Q^T \quad (\text{homogeneity and distributivity of tensor multiplication}) \\ &= \sum_{i=1}^3 a_i (Q \bar{e}_{\langle i \rangle}) \otimes (Q \bar{e}_{\langle i \rangle}) \quad (\text{Exercise A3.6.6}) \\ &= \sum_{i=1}^3 a_i e_{\langle i \rangle} \otimes e_{\langle i \rangle} = A \quad (\text{substitution}). \end{aligned}$$

Then

$$\begin{aligned} g(\bar{A}) &= g(Q \bar{A} Q^T) \quad (\text{substitution}) \\ &= g(A) \quad (\text{isotropy of } g), \end{aligned}$$

and the validity of the representation is proven.

Suppose that $\hat{\varphi}: \mathcal{D}_A \rightarrow \mathbb{R}$ is another function with the same property. Then

$$g(\bar{A}) = \varphi(\ell_1(\bar{A}), \ell_2(\bar{A}), \ell_3(\bar{A})) = \hat{\varphi}(\ell_1(\bar{A}), \ell_2(\bar{A}), \ell_3(\bar{A})), \quad \bar{A} \in \mathcal{Q}.$$

$\therefore \varphi$ and $\hat{\varphi}$ agree on $\mathcal{D}_A$; i.e., $\hat{\varphi} = \varphi$.

The proof of the converse assertion is left as Exercise AB.4.2. $\square$

Supplementary Reading

Same as for EAB.3

A8.5. Isotropic Tensor-Valued Functions of a Tensor

Here the formalization of the remark which ended §A8.2 is

Definition A8.5.1. Let $\mathcal{Q} \subset \text{Lin } V_n$ be invariant under $\text{Orth } V_n$ ($\text{Orth}^+ V_n$). Then a second-order tensor-valued function of a second-order tensor, $\underline{G}: \mathcal{Q} \rightarrow \text{Lin } V_n$, is isotropic (hemitropic¹) if it is invariant under $\text{Orth } V_n$ ($\text{Orth}^+ V_n$), i.e., if

$$\underline{G}(\underline{Q}\underline{A}\underline{Q}^T) = \underline{Q}\underline{G}(\underline{A})\underline{Q}^T \quad \forall \underline{A} \in \mathcal{Q} \text{ and } \forall \underline{Q} \in \text{Orth } V_n \text{ (Orth}^+ V_n \text{)}.$$

Again,² since the elements of both the domain and co-domain are even-order tensors, the notions of isotropy and hemitropy coincide when the dimension of the underlying linear space is odd. Formally, we have

Theorem A8.5.1. Let $\mathcal{Q} \subset \text{Lin } V_n$, where the dimension n is odd. Then a tensor-valued function of a tensor, $\underline{G}: \mathcal{Q} \rightarrow \text{Lin } V_n$, is isotropic iff it is hemitropic.

Proof. A proof similar to the proof of Theorem A8.4.1 is easily constructed; the details are left as Exercise A8.5.1. $\square$

¹ Sometimes isotropic and then our isotropic is referred to as isotropic with a center of symmetry.

² Cf. Theorem A8.4.1.

Integral powers provide important examples of isotropic tensor-valued functions of a tensor.

Theorem AB.5.2. The functions $\underline{G}: \text{Lin } \mathcal{V}_n \rightarrow \text{Lin } \mathcal{V}_n$ and
 $\underline{H}: \text{Inv } \mathcal{V}_n \rightarrow \text{Inv } \mathcal{V}_n$ defined by

$$\underline{G}(\underline{A}) = \underline{A}^m \quad \text{and} \quad \underline{H}(\underline{A}) = \underline{A}^{-m} \quad (m \geq 0 \text{ an integer})$$

are both isotropic.

Proof. Consider $\underline{G}$ in the case of $m=0$. Then $\underline{G}(\underline{A}) = \underline{A}^0 = \underline{I}$,
 and for any $\underline{Q} \in \text{Orth } \mathcal{V}_n$

$$\underline{G}(\underline{Q}\underline{A}\underline{Q}^T) = \underline{I}$$

$$= \underline{Q}\underline{Q}^T \quad (\underline{Q} \in \text{Orth } \mathcal{V}_n)$$

$$= \underline{Q}\underline{I}\underline{Q}^T \quad (\text{property of } \underline{I})$$

$$= \underline{Q}\underline{G}(\underline{A})\underline{Q}^T \quad (\text{substitution}).$$

$\therefore \underline{G}$ is isotropic when $m=0$.

When $m > 0$,

$$\underline{G}(\underline{Q}\underline{A}\underline{Q}^T) = (\underline{Q}\underline{A}\underline{Q}^T)^m$$

$$\begin{aligned}
& \overbrace{= (QAQ^T)(QAQ^T) \cdots (QAQ^T)}^{m \text{ times}} \quad (\text{defn. of exponentiation}) \\
&= \underline{Q} \underline{A} (\underline{Q}^T \underline{Q}) \underline{A} (\underline{Q}^T \underline{Q}) \cdots (\underline{Q}^T \underline{Q}) \underline{A} \underline{Q}^T \quad (\text{associativity}) \\
&= \underline{Q} \underline{A} \underline{I} \underline{A} \underline{I} \cdots \underline{I} \underline{A} \underline{Q}^T \quad (\underline{Q} \in \text{Orth } V_n) \\
&= \underline{Q} \underline{A} \underline{A} \cdots \underline{A} \underline{Q}^T \quad (\text{property of } \underline{I}) \\
&= \underline{Q} \underline{A}^m \underline{Q}^T \quad (\text{defn. of exponentiation}) \\
&= \underline{Q} \underline{G(A)} \underline{Q}^T \quad (\text{substitution}).
\end{aligned}$$

$\therefore \underline{G}$ is isotropic when $m > 0$.

The isotropy of $\underline{H}$ is established similarly; the details are left as Exercise AB.5.2. $\square$

The main result of this section will be a representation theorem for isotropic symmetric tensor-valued functions of a symmetric tensor. But first we need some preparation.

Theorem AB.5.3, (Commutation Theorem) Suppose that
 $\underline{S}, \underline{T} \in \text{Lin } V$ commute. Then each eigenspace¹ of $\underline{S}$ is invariant²
under $\underline{T}$, "Conversely", if each eigenspace of $\underline{S} \in \text{Sym } V_3$ is invariant

¹See Definition A3.11.2. At this level of generality, $\underline{S}$ may not have any eigenvalues.

²See Definition A3.12.2.

under $\underline{T} \in \text{Lin } \underline{V}_3$, then $\underline{S}$ and $\underline{T}$ commute.

Proof. Suppose that $\underline{S}$ and $\underline{T}$ commute. Let $\underline{u}$ belong to the eigenspace of $\underline{S}$ corresponding to the eigenvalue λ . Then $\underline{S}\underline{u} = \lambda\underline{u}$, and

$$\begin{aligned} \underline{S}(\underline{T}\underline{u}) &= (\underline{S}\underline{T})\underline{u} && \text{(defn. of tensor multiplication)} \\ &= (\underline{T}\underline{S})\underline{u} && (\underline{S} \text{ and } \underline{T} \text{ commute}) \\ &= \underline{T}(\underline{S}\underline{u}) && \text{(defn. of tensor multiplication)} \\ &= \underline{T}(\lambda\underline{u}) && (\underline{u} \text{ in eigenspace corresponding to } \lambda) \\ &= \lambda(\underline{T}\underline{u}) && \text{(linearity of } \underline{T}). \end{aligned}$$

Thus, $\underline{T}\underline{u}$ belongs to the eigenspace of $\underline{S}$ corresponding to λ ; i.e., this space is invariant under $\underline{T}$.

"Conversely", suppose that each eigenspace of $\underline{S} \in \text{Sym } \underline{V}_3$ is invariant under $\underline{T}$. The demonstration follows from Theorem A3.13.6. If the three eigenvalues of $\underline{S}$ are distinct, then $\underline{S}$ has three eigenspaces $\underline{E}_1$, $\underline{E}_2$, and $\underline{E}_3$; moreover,

$$\underline{V}_3 = \underline{E}_1 \oplus \underline{E}_2 \oplus \underline{E}_3,$$

Thus, each $\underline{v} \in \underline{V}_3$ admits the unique decomposition

$$\underline{v} = \sum_{i=1}^3 \underline{v}_i, \quad \underline{v}_i \in \mathcal{E}_i.$$

Then

$$\begin{aligned}
 (\underline{S}\underline{T})\underline{v} &= (\underline{S}\underline{T})\left(\sum_{i=1}^3 \underline{v}_i\right) && \text{(substitution)} \\
 &= \underline{S}\left(\underline{T}\sum_{i=1}^3 \underline{v}_i\right) && \text{(defn. of tensor multiplication)} \\
 &= \underline{S}\left(\sum_{i=1}^3 \underline{T}\underline{v}_i\right) && \text{(linearity of } \underline{T}) \\
 &= \sum_{i=1}^3 \underline{S}(\underline{T}\underline{v}_i) && \text{(linearity of } \underline{S}) \\
 &= \sum_{i=1}^3 \lambda_i \underline{T}\underline{v}_i && (\underline{T}\underline{v}_i \in \mathcal{E}_i) \\
 &= \sum_{i=1}^3 \underline{T} \lambda_i \underline{v}_i && \text{(linearity of } \underline{T}) \\
 &= \sum_{i=1}^3 \underline{T}(\underline{S}\underline{v}_i) && (\underline{v}_i \in \mathcal{E}_i) \\
 &= \sum_{i=1}^3 (\underline{T}\underline{S})\underline{v}_i && \text{(defn. of tensor multiplication)} \\
 &= (\underline{T}\underline{S})\left(\sum_{i=1}^3 \underline{v}_i\right) && \text{(linearity of } \underline{T}\underline{S}) \\
 &= (\underline{T}\underline{S})\underline{v} && \text{(substitution)}.
 \end{aligned}$$

By the arbitrariness of $\underline{v}$, $\underline{S}\underline{T} = \underline{T}\underline{S}$; i.e., $\underline{S}$ and $\underline{T}$ commute.

The proof is essentially the same when $\underline{S}$ has repeated eigenvalues. The demonstration is left as Exercise A8.5.3. $\square$

An important consequence of the Commutation Theorem is

Theorem A8.5.4. (Transfer Theorem) Let $\mathcal{Q} \subset \text{Sym } V_3$ be invariant under $\text{Orth } V_3$ and let $\underline{G}: \mathcal{Q} \rightarrow \text{Lin } V_3$ be an isotropic tensor-valued function of a symmetric tensor. Then any eigenvector of $\underline{A} \in \mathcal{Q}$ is an eigenvector of $\underline{G}(\underline{A})$.

Proof. Denote the spectrum of $\underline{A} \in \mathcal{Q}$ by $\{\lambda_1, \lambda_2, \lambda_3\}$ and let $\{\underline{\hat{e}}_{\langle 1 \rangle}, \underline{\hat{e}}_{\langle 2 \rangle}, \underline{\hat{e}}_{\langle 3 \rangle}\}$ be a corresponding eigenvector basis of $\underline{A}$.¹ As usual, there are three cases to be considered which are defined by the multiplicities of the eigenvalues,

(i) $(\lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \lambda_1)$. By Theorem A3.13.5, $\underline{A}$ has three eigenspaces

$$\mathcal{E}_i = \text{Lsp} \{ \underline{\hat{e}}_{\langle i \rangle}^* \}, \quad i \in \mathbb{I}_3,$$

Let $\underline{e}$ be any eigenvector of $\underline{A}$. Then $\underline{e} \neq \underline{0}$, and $\underline{e} \in \mathcal{E}_i$ for some choice of i . Suppose, for definiteness, that $\underline{e} \in \mathcal{E}_1$; then $\underline{e}$ has the form

$$\underline{e} = \alpha \underline{\hat{e}}_{\langle 1 \rangle}^*, \quad \alpha \neq 0. \quad (*)$$

Let $\underline{Q} \in \text{Orth } V_3$ be the reflection across the plane perpendicular to $\underline{\hat{e}}_{\langle 1 \rangle}^*$ ²:

¹ Cf. the Spectral Theorem (Theorem A3.13.1).

² See Definition A2.8.1 and Theorem A2.8.3

$$Q \tilde{e}_{\langle 1 \rangle}^* = -\tilde{e}_{\langle 1 \rangle}^*, \quad Q \tilde{e}_{\langle 2 \rangle}^* = \tilde{e}_{\langle 2 \rangle}^*, \quad Q \tilde{e}_{\langle 3 \rangle}^* = \tilde{e}_{\langle 3 \rangle}^*. \quad)^1 (**)$$

Exercise A8.5.4. Confirm analytically that Q so defined is orthogonal,

Each eigenspace of A is obviously invariant under this Q , and $\therefore$ by the Commutation Theorem, $QA = AQ$, which is equivalent to

$$QAQ^T = A.$$

Then since G is isotropic,

$$Q G(A) Q^T = G(QAQ^T) = G(A),$$

or

$$Q G(A) = G(A) Q; \quad (\#)$$

i.e., Q also commutes with $G(A)$. By $(\#)$, $(*)$, and $(**)$,

$$Q G(A) \tilde{e} = G(A) Q \tilde{e} = G(A) Q (\alpha \tilde{e}_{\langle 1 \rangle}^*) = G(A) (-\alpha \tilde{e}_{\langle 1 \rangle}^*) = -G(A) \tilde{e}. \quad (\#\#)$$

Thus, Q maps $G(A) \tilde{e}$ into its negative. To see what this means, we use the fact that $\{\tilde{e}_{\langle 1 \rangle}^*, \tilde{e}_{\langle 2 \rangle}^*, \tilde{e}_{\langle 3 \rangle}^*\}$ is a basis for V_3 to write

$$G(A) \tilde{e} = \alpha \tilde{e}_{\langle i \rangle}^*.$$

¹ Cf. Theorem A3.2.1.

Then by (**) and the linearity of Q ,

$$Q \underline{G(A)} \underline{e} = \alpha^{(1)} Q \underline{e}_{(1)}^* = -\alpha^{(1)} \underline{e}_{(1)}^* + \alpha^{(2)} \underline{e}_{(2)}^* + \alpha^{(3)} \underline{e}_{(3)}^*,$$

But by (##),

$$Q \underline{G(A)} \underline{e} = -\underline{G(A)} \underline{e} = -\alpha^{(1)} \underline{e}_{(1)}^* - \alpha^{(2)} \underline{e}_{(2)}^* - \alpha^{(3)} \underline{e}_{(3)}^*;$$

and $\therefore$ by the uniqueness of components relative to a given basis,

$$\alpha^{(2)} = -\alpha^{(2)} = 0 \text{ and } \alpha^{(3)} = -\alpha^{(3)} = 0,$$

Thus,

$$\underline{G(A)} \underline{e} = \alpha^{(1)} \underline{e}_{(1)}^* = \frac{\alpha^{(1)}}{\alpha} \underline{e};$$

i.e., $\underline{e}$ also is an eigenvector of $\underline{G(A)}$.

If $\underline{e} \in \mathcal{E}_2$ ($\mathcal{E}_3$), we can get the same result in the same way by choosing Q to be the reflection across the plane perpendicular to $\underline{e}_{(2)}^*$ ($\underline{e}_{(3)}^*$).

(ii) ($\lambda_1 \neq \lambda_2 = \lambda_3$). This case can be handled in the same fashion as case (i). The details are left as Exercise AB.5.5.

(iii). ($\lambda_1 = \lambda_2 = \lambda_3 = \lambda$). In this case, $\underline{A}$ has the spherical form.

$$\underline{A} = \lambda \underline{I}.)'$$

¹ See Theorem A3.13.5,

Obviously, every nonzero vector is an eigenvector of $\underline{A}$, and $\therefore$ we must show that every nonzero vector also is an eigenvector of $\underline{G}(\underline{A})$.

Since $\underline{A}$ has the above form,

$$\underline{Q} \underline{A} \underline{Q}^T = \underline{A} \quad \forall \underline{Q} \in \text{Orth } \mathcal{V}_3.$$

Then since $\underline{G}$ is isotropic,

$$\underline{Q} \underline{G}(\underline{A}) \underline{Q}^T = \underline{G}(\underline{Q} \underline{A} \underline{Q}^T) = \underline{G}(\underline{A})$$

or

$$\underline{Q} \underline{G}(\underline{A}) \underline{Q}^T = \underline{G}(\underline{A}) \quad \forall \underline{Q} \in \text{Orth } \mathcal{V}_3.$$

Thus, $\underline{G}(\underline{A})$ is an isotropic¹ tensor; and by Theorem A8.1.3, it admits the representation

$$\underline{G}(\underline{A}) = \beta \underline{I}, \quad \beta \in \mathbb{R}.$$

Consequently, every nonzero vector is an eigenvector of $\underline{G}(\underline{A})$.² $\square$

As an interesting corollary of the Transfer Theorem, we have

¹ See Definition A8.1.2.

² We note that the lengthier demonstration which established case (i) also can be used to prove case (iii).

Theorem A8.5.5.¹ Let $\mathcal{Q} \subset \text{Sym } V_3$ be invariant under
 $\text{Orth } V_3$ and let $G: \mathcal{Q} \rightarrow \text{Lin } V_3$ be an isotropic tensor-valued
function of a symmetric tensor. Then $G(\mathcal{Q}) \subset \text{Sym } V_3$; i.e.,
the values of G also are symmetric.

Proof. By the Spectral Theorem (Theorem A3.13.1), A
has an eigenvector basis, say $\{\underline{\underline{e}}_{\langle 1 \rangle}^*, \underline{\underline{e}}_{\langle 2 \rangle}^*, \underline{\underline{e}}_{\langle 3 \rangle}^*\}$. By the
Transfer Theorem, $\underline{\underline{e}}_{\langle i \rangle}^*$ also is an eigenvector of $G(\underline{\underline{A}})$;
call the associated eigenvalue μ_i . By Theorem A3.2.2,
 $G(\underline{\underline{A}})$ has the representation

$$G(\underline{\underline{A}}) = (G(\underline{\underline{A}}))^{\langle ij \rangle} \underline{\underline{e}}_{\langle i \rangle}^* \otimes \underline{\underline{e}}_{\langle j \rangle}^*,$$

where

$$(G(\underline{\underline{A}}))^{\langle ij \rangle} = \underline{\underline{e}}_{\langle i \rangle}^* \cdot G(\underline{\underline{A}}) \underline{\underline{e}}_{\langle j \rangle}^*$$

$$= \underline{\underline{e}}_{\langle i \rangle}^* \cdot (\mu_j \underline{\underline{e}}_{\langle j \rangle}^*) \quad ((\mu_j, \underline{\underline{e}}_{\langle j \rangle}^*) \text{ an eigenvector pair of } G(\underline{\underline{A}}))$$

$$= \mu_j \underline{\underline{e}}_{\langle i \rangle}^* \cdot \underline{\underline{e}}_{\langle j \rangle}^* \quad (\text{linearity of inner product})$$

$$= \mu_j \delta_{ij} \quad (\{\underline{\underline{e}}_{\langle i \rangle}^*\} \text{ orthonormal}).$$

$\therefore$

$$G(\underline{\underline{A}}) = \sum_{i=1}^3 \mu_i \underline{\underline{e}}_{\langle i \rangle}^* \otimes \underline{\underline{e}}_{\langle i \rangle}^*,$$

which is symmetric by Theorem A3.6.3 (ii), (iii), and (vi). $\square$

The final step in our preparation for the representation

¹This result was shown to me by Dr. LAYACHI HADJI in 1986.

theorem for a symmetric tensor-valued function of a symmetric tensor is

Theorem AB.5.6. (WANG'S Lemma) Let $\underline{A} \in \text{Sym } \mathcal{V}_3$ have the spectrum $\{\lambda_1, \lambda_2, \lambda_3\}$, and let $\{\underline{\tilde{e}}_{\langle 1 \rangle}, \underline{\tilde{e}}_{\langle 2 \rangle}, \underline{\tilde{e}}_{\langle 3 \rangle}\}$ be a corresponding orthonormal basis of eigenvectors.¹

(i) $(\lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \lambda_1)$. Then $\underline{A}$ has the spectral decomposition

$$\underline{A} = \sum_{i=1}^3 \lambda_i \underline{\tilde{e}}_{\langle i \rangle}^* \otimes \underline{\tilde{e}}_{\langle i \rangle}.$$

The set $\{\underline{I}, \underline{A}, \underline{A}^2\}$ is linearly independent, and

$$\text{Lsp}\{\underline{I}, \underline{A}, \underline{A}^2\} = \text{Lsp}\{\underline{\tilde{e}}_{\langle 1 \rangle}^* \otimes \underline{\tilde{e}}_{\langle 1 \rangle}, \underline{\tilde{e}}_{\langle 2 \rangle}^* \otimes \underline{\tilde{e}}_{\langle 2 \rangle}, \underline{\tilde{e}}_{\langle 3 \rangle}^* \otimes \underline{\tilde{e}}_{\langle 3 \rangle}\}.$$

(ii) $(\lambda_1 \neq \lambda_2 = \lambda_3)$. Then $\underline{A}$ has the spectral decomposition

$$\underline{A} = \lambda_1 \underline{\tilde{e}}_{\langle 1 \rangle}^* \otimes \underline{\tilde{e}}_{\langle 1 \rangle} + \lambda_2 (\underline{I} - \underline{\tilde{e}}_{\langle 1 \rangle}^* \otimes \underline{\tilde{e}}_{\langle 1 \rangle})$$

¹ Cf. the Spectral Theorem (Theorem A3.13.1).

² Note that in this case the normalized eigenvectors $\underline{\tilde{e}}_{\langle i \rangle}$ are unique to within sense (Theorem A3.13.5), and this ambiguity multiplies out of $\underline{\tilde{e}}_{\langle i \rangle}^* \otimes \underline{\tilde{e}}_{\langle i \rangle}$.

³ Note that in this case the normalized eigenvector $\underline{\tilde{e}}_{\langle 1 \rangle}$ is unique to within sense (Theorem A3.13.5), and this ambiguity multiplies out of $\underline{\tilde{e}}_{\langle 1 \rangle}^* \otimes \underline{\tilde{e}}_{\langle 1 \rangle}$.

The set $\{\underline{I}, \underline{A}\}$ is linearly independent, and

$$\text{Lsp}\{\underline{I}, \underline{A}\} = \text{Lsp}\{\underline{e}_{<i>}^* \otimes \underline{e}_{<i>}^*, \underline{I} - \underline{e}_{<i>}^* \otimes \underline{e}_{<i>}^*\}.$$

Proof. (i) $(\lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \lambda_1)$. The spectral decomposition

$$\underline{A} = \sum_{i=1}^3 \lambda_i \underline{e}_{<i>}^* \otimes \underline{e}_{<i>}^* \quad (*)$$

follows from the Spectral Theorem (Theorem A3.13.1). In order to establish the linear independence of $\{\underline{I}, \underline{A}, \underline{A}^2\}$, consider the linear combination

$$\alpha \underline{A}^2 + \beta \underline{A} + \gamma \underline{I} = \underline{0}. \quad (**)$$

Operating on $\underline{e}_{<i>}^*$ with $(**)$ in which $\underline{A}$ is represented by $(*)$, we get (Exercise A8.5.6)

$$(\alpha \lambda_i^2 + \beta \lambda_i + \gamma) \underline{e}_{<i>}^* = \underline{0}, \quad i \in \Pi_3.$$

Consequently,

$$\alpha \lambda_i^2 + \beta \lambda_i + \gamma = 0, \quad i \in \Pi_3;$$

or

$$\left. \begin{aligned} \lambda_1^2 \alpha + \lambda_1 \beta + \gamma &= 0, \\ \lambda_2^2 \alpha + \lambda_2 \beta + \gamma &= 0, \\ \lambda_3^2 \alpha + \lambda_3 \beta + \gamma &= 0. \end{aligned} \right\} \quad (\#)$$

View $(\#)$ as a system of homogeneous, linear, algebraic equations

for α, β , and δ . The determinant of the coefficient matrix is the VANDERMONDE determinant

$$\begin{vmatrix} \lambda_1^2 & \lambda_1 & 1 \\ \lambda_2^2 & \lambda_2 & 1 \\ \lambda_3^2 & \lambda_3 & 1 \end{vmatrix} = (\lambda_1 - \lambda_2)(\lambda_2 - \lambda_3)(\lambda_1 - \lambda_3),$$

which clearly does not vanish in the present case. $\therefore \alpha = \beta = \delta = 0$), and we conclude from Theorem A2.2.2 that the set $\{\underline{I}, \underline{A}, \underline{A}^2\}$ is linearly independent.

Next consider

$$L := \text{Lsp} \{ \underline{e}_{\langle 1 \rangle}^* \otimes \underline{e}_{\langle 1 \rangle}^*, \underline{e}_{\langle 2 \rangle}^* \otimes \underline{e}_{\langle 2 \rangle}^*, \underline{e}_{\langle 3 \rangle}^* \otimes \underline{e}_{\langle 3 \rangle}^* \}.$$

By Theorems A3.6.3(vi) and A2.1.12, L is a linear subspace of $\text{Sym } V_3$; and consequently by Theorem A2.1.11, L itself is a linear space. By Theorem A3.7.5, the set $\{ \underline{e}_{\langle 1 \rangle}^* \otimes \underline{e}_{\langle 1 \rangle}^*, \underline{e}_{\langle 2 \rangle}^* \otimes \underline{e}_{\langle 2 \rangle}^*, \underline{e}_{\langle 3 \rangle}^* \otimes \underline{e}_{\langle 3 \rangle}^* \}$ is orthonormal; and $\therefore$ by Theorem A2.3.14 it is linearly independent. Since it necessarily spans L , it follows from Definition A2.2.2 that it is a basis for L . We then conclude from Definition A2.2.3 that L is 3-dimensional.

Now by direct calculation from (*),

¹Cf. Theorem A3.8.7

$$\underline{A}^2 = \sum_{i=1}^3 \lambda_i^2 \underline{\underline{e}}_{\langle i \rangle}^* \otimes \underline{\underline{e}}_{\langle i \rangle}^* \in \mathcal{L};$$

by (*) itself, $\underline{A} \in \mathcal{L}$; and by the resolution of the identity¹,

$$\underline{I} = \sum_{i=1}^3 \underline{\underline{e}}_{\langle i \rangle}^* \otimes \underline{\underline{e}}_{\langle i \rangle}^* \in \mathcal{L}.$$

Thus, the linearly independent set of three elements $\{\underline{I}, \underline{A}, \underline{A}^2\}$ is contained in the 3-dimensional linear space $\mathcal{L}$. $\therefore$ by Theorem A2.2.B, it is a basis for $\mathcal{L}$; and then by Definition A2.2.2, it spans $\mathcal{L}$. Hence,

$$\text{Lsp}\{\underline{I}, \underline{A}, \underline{A}^2\} = \mathcal{L} = \text{Lsp}\{\underline{\underline{e}}_{\langle 1 \rangle}^* \otimes \underline{\underline{e}}_{\langle 1 \rangle}^*, \underline{\underline{e}}_{\langle 2 \rangle}^* \otimes \underline{\underline{e}}_{\langle 2 \rangle}^*, \underline{\underline{e}}_{\langle 3 \rangle}^* \otimes \underline{\underline{e}}_{\langle 3 \rangle}^*\}.$$

(ii) $(\lambda_1 \neq \lambda_2 = \lambda_3)$. This case can be handled by the same kind of reasoning. The details are left as Exercise AB.5.7. $\square$

We are now in position to establish our first representation theorem for an isotropic symmetric tensor-valued function of a symmetric tensor.

Theorem AB.5.7.² Let $\mathcal{Q} \subset \text{Sym } V_3$ be invariant

¹ See p. A3.13.13.

² RIVLIN and ERICKSEN, "Stress-deformation relations for isotropic materials," J. Rational Mech. Anal., 4 (1955) 323-425. When $\mathcal{Q}$ is a polynomial function, the theorem is much older.

under Orth V_3 . If a symmetric tensor-valued function of a symmetric tensor, $G: \mathcal{Q} \rightarrow \text{Sym } V_3$, is isotropic, then $\exists$ unique functions

$$\alpha_0, \alpha_1, \alpha_2, \beta_0, \beta_1, \gamma_0 : \mathcal{D}_A \rightarrow \mathbb{R},$$

where

$$\mathcal{D}_A = \{ (\ell_1(\underline{A}), \ell_2(\underline{A}), \ell_3(\underline{A})) : \underline{A} \in \mathcal{Q} \},$$

$\Rightarrow$

$$G(\underline{A}) = \begin{cases} \alpha_0(\ell_1(\underline{A}), \ell_2(\underline{A}), \ell_3(\underline{A})) \underline{I} + \alpha_1(\cdot) \underline{A} + \alpha_2(\cdot) \underline{A}^2 & \text{if } \lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \lambda_1, \\ \beta_0(\ell_1(\underline{A}), \ell_2(\underline{A}), \ell_3(\underline{A})) \underline{I} + \beta_1(\cdot) \underline{A} & \text{if } \lambda_1 \neq \lambda_2 = \lambda_3, \\ \gamma_0(\ell_1(\underline{A}), \ell_2(\underline{A}), \ell_3(\underline{A})) \underline{I} & \text{if } \lambda_1 = \lambda_2 = \lambda_3, \end{cases}$$

where $\ell_1(\underline{A})$, $\ell_2(\underline{A})$, and $\ell_3(\underline{A})$ are the fundamental invariants of $\underline{A}$, and $\{\lambda_1, \lambda_2, \lambda_3\}$ is the spectrum of $\underline{A}$.

Conversely, a function of this form is isotropic.¹

Proof. The converse assertion follows immediately from Theorems A8.4.4 and A8.5.2. The details are left as Exercise A8.5.8.

To establish the representation, assume that G is

¹Recall from Theorem A8.5.1 that here ($n=3$) isotropy and hemitropy coincide.

isotropic, and consider any $\underline{A} \in \mathcal{Q}$. Denote the spectrum of $\underline{A}$ by $\{\lambda_1, \lambda_2, \lambda_3\}$, and let $\{\underline{e}_{\langle 1 \rangle}^*, \underline{e}_{\langle 2 \rangle}^*, \underline{e}_{\langle 3 \rangle}^*\}$ be an associated eigenvector basis.¹ By the Transfer Theorem, each $\underline{e}_{\langle i \rangle}^*$ also is an eigenvector of $\underline{G}(\underline{A})$. Consequently, $\underline{G}(\underline{A})$ has the spectral resolution

$$\underline{G}(\underline{A}) = \sum_{i=1}^3 \mu_i \underline{e}_{\langle i \rangle}^* \otimes \underline{e}_{\langle i \rangle}^*. \quad (*)$$

The remainder of the analysis falls into the three standard cases determined by the multiplicities of the eigenvalues of $\underline{A}$.

(i) $(\lambda_1 \neq \lambda_2 \neq \lambda_3 \neq \lambda_1)$. By (*),

$$\underline{G}(\underline{A}) \in \text{Lsp} \{ \underline{e}_{\langle 1 \rangle}^* \otimes \underline{e}_{\langle 1 \rangle}^*, \underline{e}_{\langle 2 \rangle}^* \otimes \underline{e}_{\langle 2 \rangle}^*, \underline{e}_{\langle 3 \rangle}^* \otimes \underline{e}_{\langle 3 \rangle}^* \};$$

but in this case, we know from Wang's Lemma that

$$\text{Lsp} \{ \underline{e}_{\langle 1 \rangle}^* \otimes \underline{e}_{\langle 1 \rangle}^*, \underline{e}_{\langle 2 \rangle}^* \otimes \underline{e}_{\langle 2 \rangle}^*, \underline{e}_{\langle 3 \rangle}^* \otimes \underline{e}_{\langle 3 \rangle}^* \} = \text{Lsp} \{ \underline{I}, \underline{A}, \underline{A}^2 \}.$$

Moreover, $\{\underline{I}, \underline{A}, \underline{A}^2\}$ is a basis for $\text{Lsp} \{ \underline{e}_{\langle 1 \rangle}^* \otimes \underline{e}_{\langle 1 \rangle}^*, \underline{e}_{\langle 2 \rangle}^* \otimes \underline{e}_{\langle 2 \rangle}^*, \underline{e}_{\langle 3 \rangle}^* \otimes \underline{e}_{\langle 3 \rangle}^* \}$.²
 $\therefore \exists$ unique scalars $\overline{\alpha}_0(\underline{A}), \overline{\alpha}_1(\underline{A}), \overline{\alpha}_2(\underline{A}) \ni$

$$\underline{G}(\underline{A}) = \overline{\alpha}_0(\underline{A}) \underline{I} + \overline{\alpha}_1(\underline{A}) \underline{A} + \overline{\alpha}_2(\underline{A}) \underline{A}^2. \quad (**)$$

Next, we use the isotropy of $\underline{G}$ to show that the functions

¹ Cf. the Spectral Theorem (Theorem A3.13.1).

² See the proof of Wang's Lemma.

$\bar{\alpha}_0, \bar{\alpha}_1$, and $\bar{\alpha}_2$ (on $\mathcal{Q}$ to $\mathbb{R}$) are each isotropic. Since $\underline{G}$ is isotropic,

$$\underline{G}(\underline{A}) - \underline{Q}^T \underline{G}(\underline{Q} \underline{A} \underline{Q}^T) \underline{Q} = \underline{0} \quad \forall \underline{Q} \in \text{Orth } \mathcal{V}_3 \quad (*)$$

On using the representation $(**)$ in $(*)$, one is quickly¹ lead to

$$[\bar{\alpha}_0(\underline{A}) - \bar{\alpha}_0(\underline{Q} \underline{A} \underline{Q}^T)] \underline{I} + [\bar{\alpha}_1(\underline{A}) - \bar{\alpha}_1(\underline{Q} \underline{A} \underline{Q}^T)] \underline{A} + [\bar{\alpha}_2(\underline{A}) - \bar{\alpha}_2(\underline{Q} \underline{A} \underline{Q}^T)] \underline{A}^2 = \underline{0};$$

and by the linear independence of $\{\underline{I}, \underline{A}, \underline{A}^2\}$, the terms in brackets must vanish. Thus, for each $i \in \{0, 1, 2\}$

$$\bar{\alpha}_i(\underline{Q} \underline{A} \underline{Q}^T) = \bar{\alpha}_i(\underline{A}) \quad \forall \underline{A} \in \mathcal{Q} \text{ and } \forall \underline{Q} \in \text{Orth } \mathcal{V}_3;$$

i.e., each $\bar{\alpha}_i$ is an isotropic scalar-valued function of a symmetric tensor. Then by the representation Theorem A8.4.4 $\exists$ unique functions $\alpha_i : \mathcal{D}_A \rightarrow \mathbb{R} \ni$

$$\bar{\alpha}_i(\underline{A}) = \alpha_i(\ell_1(\underline{A}), \ell_2(\underline{A}), \ell_3(\underline{A})), \quad i=0, 1, 2.$$

This completes the proof for the case of distinct eigenvalues.

(ii) $(\lambda_1 \neq \lambda_2 = \lambda_3)$. In this case, it follows from Theorem A3.13.5 (ii) that the eigenspaces of $\underline{A}$ are $\mathcal{E}_1 = \text{Lsp}\{\underline{e}_{<1>}^*\}$ and $\mathcal{E}_2 = \mathcal{E}_1^\perp = \text{Lsp}\{\underline{e}_{<2>}^*, \underline{e}_{<3>}^*\}$. Consider $\underline{G}(\underline{A})$, and recall

$$\underline{G}(\underline{A}) = \sum_{i=1}^3 \mu_i \underline{e}_{< i >}^* \otimes \underline{e}_{< i >}^* \quad (*),$$

¹Use of Theorem A8.5.2 speeds up the process.

In terms of the multiplicities of the eigenvalues of $\underline{G}(\underline{A})$, there are the three standard cases.

(i)' ($\mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_1$). By Theorem A3.13.5 (i), the eigenspaces of $\underline{G}(\underline{A})$ are $\mathcal{E}_i' = \text{Lsp} \{ \tilde{\underline{E}}_{\langle i \rangle}^* \}$, $i \in \Pi_3$; here, the prime emphasizes that we are looking at the eigenspaces of $\underline{G}(\underline{A})$ rather than the eigenspaces of $\underline{A}$.

Now, $\tilde{\underline{E}}_{\langle 2 \rangle}^* + \tilde{\underline{E}}_{\langle 3 \rangle}^* \in \mathcal{E}_2$ is an eigenvector of $\underline{A}$; and $\therefore$ by the Transfer Theorem, it also is an eigenvector of $\underline{G}(\underline{A})$. Thus, $\tilde{\underline{E}}_{\langle 2 \rangle}^* + \tilde{\underline{E}}_{\langle 3 \rangle}^*$ must belong to $\mathcal{E}_1'$ or $\mathcal{E}_2'$ or $\mathcal{E}_3'$; but, clearly, it does not. Consequently, in case (ii), $\underline{G}(\underline{A})$ does not have three distinct eigenvalues, i.e., subcase (i)' is empty.

In considering the case of $\underline{G}(\underline{A})$ having two distinct eigenvalues, it is not sufficient to examine just one representative situation such as $\mu_1 \neq \mu_2 = \mu_3$ because our labeling here is tied in with the structure of $\underline{A}$ through (*).

(ii)' ($\mu_1 \neq \mu_2 = \mu_3$). By Theorem A3.13.5 (ii) and (*), $\underline{G}(\underline{A})$ has the two eigenspaces $\mathcal{E}_1' = \text{Lsp} \{ \tilde{\underline{E}}_{\langle 1 \rangle}^* \} = \mathcal{E}_1$ and $\mathcal{E}_2' = (\mathcal{E}_1')^\perp = \mathcal{E}_1^\perp$; and the kind of incompatibility with the structure of $\underline{A}$ which arose in subcase (i)' does not occur here. In this case, it also follows from Theorem A3.13.5 (ii) that the spectral resolution of $\underline{G}(\underline{A})$ can be written as

$$\underline{G}(\underline{A}) = \mu_1 \underline{E}_{\langle 1 \rangle}^* \otimes \underline{E}_{\langle 1 \rangle}^* + \mu_2 (\underline{I} - \underline{E}_{\langle 1 \rangle}^* \otimes \underline{E}_{\langle 1 \rangle}^*).$$

So $\underline{G}(\underline{A}) \in \text{Lsp} \{ \underline{E}_{\langle 1 \rangle}^* \otimes \underline{E}_{\langle 1 \rangle}^*, \underline{I} - \underline{E}_{\langle 1 \rangle}^* \otimes \underline{E}_{\langle 1 \rangle}^* \}$, and by Wang's lemma

$$\underline{G}(\underline{A}) \in \text{Lsp} \{ \underline{I}, \underline{A} \}.$$

(ii)' $(\mu_2 \neq \mu_1 = \mu_3)$. By Theorem A3.13.5 (ii) and (*), $\underline{G}(\underline{A})$ has the two eigenspaces $\mathcal{E}_1' = \text{Lsp} \{ \underline{E}_{\langle 2 \rangle}^* \}$ and $\mathcal{E}_2' = (\mathcal{E}_1')^\perp = \text{Lsp} \{ \underline{E}_{\langle 1 \rangle}^*, \underline{E}_{\langle 3 \rangle}^* \}$. As in subcase (i)',

$$\underline{E}_{\langle 2 \rangle}^* + \underline{E}_{\langle 3 \rangle}^* \in \mathcal{E}_2' \text{ but } \underline{E}_{\langle 2 \rangle}^* + \underline{E}_{\langle 3 \rangle}^* \notin \mathcal{E}_1' \text{ or } \mathcal{E}_2'$$

— in contradiction of the Transfer Theorem. $\therefore$ subcase (ii)'_b cannot obtain.

--- (ii)'_c $(\mu_3 \neq \mu_2 = \mu_1)$. This subcase is empty also. The details are left as Exercise AB.5.9.

(iii)' $(\mu_1 = \mu_2 = \mu_3 =: \mu)$. In this subcase, it follows from the Spectral Theorem (Theorem A3.13.1) that $\underline{G}(\underline{A}) = \mu \underline{I}$. Thus,

$$\underline{G}(\underline{A}) \in \text{Lsp} \{ \underline{I}, \underline{A} \}.$$

Hence, in all possibilities allowed in case (ii),

$$\underline{G}(\underline{A}) \in \text{Lsp} \{ \underline{I}, \underline{A} \}.$$

Since $\{\underline{I}, \underline{A}\}$ is a linearly independent set¹, it is a basis for $\text{Lsp}\{\underline{I}, \underline{A}\}$; and $\exists$ unique scalars $\overline{\beta}_0(\underline{A})$ and $\overline{\beta}_1(\underline{A}) \ni$

$$\underline{G}(\underline{A}) = \overline{\beta}_0(\underline{A})\underline{I} + \overline{\beta}_1(\underline{A})\underline{A}.$$

By the same kind of argument as used in case (i), it is easy to see that the functions $\overline{\beta}_0$ and $\overline{\beta}_1$ (on $\mathcal{Q}$ to $\mathbb{R}$) are isotropic (Work this out as Exercise AB.5.10.). Then by Theorem AB.4.4, $\exists$ unique functions $\beta_i: \mathcal{D}_A \rightarrow \mathbb{R} \ni$

$$\overline{\beta}_i(\underline{A}) = \beta_i(\underline{L}_1(\underline{A}), \underline{L}_2(\underline{A}), \underline{L}_3(\underline{A})), \quad i=0,1.$$

This completes the proof for the case of exactly two distinct eigenvalues.

(iii) ($\lambda_1 = \lambda_2 = \lambda_3 =: \lambda$). In this case, we see from Theorem A3.13.5 (iii) that $\underline{A}$ has one eigenspace, $\mathcal{E}_1 = \mathcal{V}_3$. Thus, by the Transfer Theorem, every nonzero vector is an eigenvector of $\underline{G}(\underline{A})$. The same kind of analysis as that used in case (ii) shows that $\mu_1 = \mu_2 = \mu_3 =: \mu$ (Work this out as Exercise AB.5.11.). Then (*) becomes $\underline{G}(\underline{A}) = \mu \underline{I}$, and on writing $\mu = \overline{\delta}(\underline{A})$, we have

$$\underline{G}(\underline{A}) = \overline{\delta}(\underline{A})\underline{I}.$$

By the isotropy of $\underline{G}$, for any $\underline{Q} \in \text{Orth } \mathcal{V}_3$ and $\underline{A} \in \mathcal{Q}$,

¹ See Wang's Lemma.

$$\underline{G}(\underline{Q}\underline{A}\underline{Q}^T) = \underline{Q} \underline{G}(\underline{A}) \underline{Q}^T$$

or

$$\bar{\gamma}(\underline{Q}\underline{A}\underline{Q}^T)\underline{I} = \underline{Q} [\bar{\gamma}(\underline{A})\underline{I}] \underline{Q}^T = \bar{\gamma}(\underline{A})\underline{I}.$$

Thus, $\bar{\gamma} : \mathcal{Q} \rightarrow \mathbb{R}$ is isotropic, and by Theorem A8.4.4, $\exists$ a unique function $\gamma_0 : \mathcal{D}_A \rightarrow \mathbb{R} \ni$

$$\bar{\gamma}(\underline{A}) = \gamma_0(\ell_1(\underline{A}), \ell_2(\underline{A}), \ell_3(\underline{A})).$$

This completes the proof for the case of exactly one eigenvalue. $\square$

In continuum mechanics, it is customary to observe that the representations in Theorem A8.5.7 appropriate to repeated eigenvalues are special cases of the representation that obtains when the eigenvalues are distinct. Thus, we have the following alternative representation theorem.

Theorem A8.5.8. Let $\mathcal{Q} \subset \text{Sym}^+ \mathbb{V}_3$ be invariant under $\text{Orth } \mathbb{V}_3$. If a symmetric tensor-valued function of a symmetric tensor, $\underline{G} : \mathcal{Q} \rightarrow \text{Sym } \mathbb{V}_3$, is isotropic, then $\exists$ (not necessarily unique) functions

$$\varphi_0, \varphi_1, \varphi_2 : \mathcal{D}_A \rightarrow \mathbb{R},$$

where

$$\mathcal{D}_A = \{(\ell_1(\underline{A}), \ell_2(\underline{A}), \ell_3(\underline{A})) : \underline{A} \in \mathcal{Q}\}$$

$\ni$

$$\underline{G}(\underline{A}) = \varphi_0(\ell_1(\underline{A}), \ell_2(\underline{A}), \ell_3(\underline{A}))\underline{I} + \varphi_1(\cdot)\underline{A} + \varphi_2(\cdot)\underline{A}^2,$$

where $\ell_1(A)$, $\ell_2(A)$, and $\ell_3(A)$ are the fundamental invariants of A .

Conversely, a function of this form is isotropic.¹

The single formula representation of the above theorem is so much simpler than the three formula representation of Theorem A8.5.7, that it seems silly even to bother to state Theorem A8.5.7 as a formal result. However, there are some subtleties that should not go unnoticed. Given an isotropic function $\underline{G}: \mathcal{Q} \rightarrow \text{Sym}^2 \mathbb{V}_3$, the functions φ_0, φ_1 , and φ_2 in the single formula representation are not in general uniquely determined; moreover, they generally do not inherit the smoothness properties of $\underline{G}$. In fact, it is possible for $\underline{G}$ to be continuous but for the φ 's to be discontinuous at $\underline{A}$'s with repeated eigenvalues. Some work has been done on this by SERRIN.²

When $\underline{A}$ is invertible, it follows from the Hamilton-Cayley Theorem (Theorem A3.14.3) that

$$\underline{A}^2 = \ell_1(\underline{A})\underline{A} - \ell_2(\underline{A})\underline{I} + \ell_3(\underline{A})\underline{A}^{-1};$$

consequently, in this case, the single formula representation and the first equation in the three formula representation admit the

¹Recall from Theorem A8.5.1 that here ($n=3$) isotropy and hemitropy coincide.

²SERRIN, "The derivation of the stress-deformation relations for a Stokesian fluid", J. Maths. Mech. **8** (1959) 459-469,

alternative expression

$$\underline{\mathcal{G}}(\underline{A}) = \mathcal{V}_0(\mathcal{L}_1(\underline{A}), \mathcal{L}_2(\underline{A}), \mathcal{L}_3(\underline{A})) \underline{I} + \mathcal{V}_1(\cdot) \underline{A} + \mathcal{V}_{-1}(\cdot) \underline{A}^{-1},$$

where $\mathcal{V}_0, \mathcal{V}_1$, and $\mathcal{V}_{-1}$ are functions on $\mathcal{Q}_A$ to $\mathbb{R}$, and now

$$\mathcal{Q}_A = \{\mathcal{L}_1(\underline{A}), \mathcal{L}_2(\underline{A}), \mathcal{L}_3(\underline{A}) : \underline{A} \in \mathcal{Q} \subset \text{Inv } \mathcal{V}_3\}.$$

Finally, we shall need a representation for an isotropic linear symmetric tensor-valued function of a symmetric tensor. We should be able to get this by specializing Theorem A8.5.7 (alternatively Theorem A8.5.8) to the linear case, but for a rigorous development it is easier to proceed directly.

Theorem A8.5.9. If a linear symmetric tensor-valued function of a symmetric tensor, $\underline{\mathcal{G}} : \text{Sym } \mathcal{V}_3 \rightarrow \text{Sym } \mathcal{V}_3$, is isotropic, then $\exists$ unique real constants μ and λ $\ni$

$$\underline{\mathcal{G}}(\underline{A}) = \lambda (\text{tr } \underline{A}) \underline{I} + 2\mu \underline{A}, \quad \underline{A} \in \text{Sym } \mathcal{V}_3.$$

Conversely, a linear function of this form is isotropic.¹

Proof. Suppose that $\underline{\mathcal{G}} : \text{Sym } \mathcal{V}_3 \rightarrow \text{Sym } \mathcal{V}_3$ is both linear and isotropic. Let $\mathcal{U} = \{\underline{u} \in \mathcal{V}_3 : |\underline{u}| = 1\}$ denote the set of all unit vectors, and consider $\underline{\mathcal{G}}(\underline{e} \otimes \underline{e})$ for $\underline{e} \in \mathcal{U}$. We leave it as Exercise A8.5.12 to show that $\underline{e} \otimes \underline{e}$ has spectrum $\{0, 0, 1\}$ and eigenspaces $\text{Lsp}\{\underline{e}\}$ and $\{\underline{e}\}^\perp$. Thus, $\underline{e} \otimes \underline{e}$ has exactly

¹Recall from Theorem A8.5.1 that here ($n=3$) isotropy and hemitropy coincide.

two distinct eigenvalues; and by case (ii) of the proof of Theorem A8.5.7,

$$\underline{G}(\underline{e} \otimes \underline{e}) \in \text{Lsp} \{ \underline{I}, \underline{e} \otimes \underline{e} \}$$

But by Wang's Lemma, the set $\{ \underline{I}, \underline{e} \otimes \underline{e} \}$ is linearly independent; $\therefore$ it is a basis for $\text{Lsp} \{ \underline{I}, \underline{e} \otimes \underline{e} \}$, and $\exists$ unique scalars $\lambda(\underline{e})$ and $\mu(\underline{e}) \ni$

$$\underline{G}(\underline{e} \otimes \underline{e}) = \lambda(\underline{e}) \underline{I} + 2\mu(\underline{e}) \underline{e} \otimes \underline{e}. \quad (*)$$

For any other $\underline{\bar{e}} \in \mathcal{U}$, we likewise have

$$\underline{G}(\underline{\bar{e}} \otimes \underline{\bar{e}}) = \lambda(\underline{\bar{e}}) \underline{I} + 2\mu(\underline{\bar{e}}) \underline{\bar{e}} \otimes \underline{\bar{e}}. \quad (**)$$

Arguing as in the proof of Theorem A8.1.1, we can always find an orthogonal tensor $\underline{Q} \ni \underline{\bar{e}} = \underline{Q} \underline{e}$. Then by Exercise A3.6.6,

$$\underline{\bar{e}} \otimes \underline{\bar{e}} = \underline{Q}(\underline{e} \otimes \underline{e})\underline{Q}^T.$$

Then

$$\underline{G}(\underline{\bar{e}} \otimes \underline{\bar{e}}) = \underline{G}(\underline{Q}(\underline{e} \otimes \underline{e})\underline{Q}^T) \quad (\text{substitution})$$

$$= \underline{Q} \underline{G}(\underline{e} \otimes \underline{e}) \underline{Q}^T \quad (\text{isotropy})$$

$$= \underline{Q} [\lambda(\underline{e}) \underline{I} + 2\mu(\underline{e}) \underline{e} \otimes \underline{e}] \underline{Q}^T \quad (\text{representation } *)$$

$$= \lambda(\underline{e}) \underline{I} + 2\mu(\underline{e}) \underline{\bar{e}} \otimes \underline{\bar{e}} \quad (\text{tensor manipulations})$$

which on comparison with the representation (**) yields

$$[\lambda(\bar{e}) - \lambda(\underline{e})] \underline{I} + [2\mu(\bar{e}) - 2\mu(\underline{e})] \bar{e} \otimes \bar{e} = \underline{0}.$$

Again by Wang's Lemma, the set $\{\underline{I}, \bar{e} \otimes \bar{e}\}$ is linearly independent; and $\therefore$ the terms in the brackets must vanish. Thus,

$$\lambda(\bar{e}) = \lambda(\underline{e}) \text{ and } \mu(\bar{e}) = \mu(\underline{e});$$

i.e., $\lambda(\underline{e})$ and $\mu(\underline{e})$ are constants, say λ and μ . Then (*) becomes

$$\underline{G}(\underline{e} \otimes \underline{e}) = \lambda \underline{I} + 2\mu \underline{e} \otimes \underline{e}, \quad \underline{e} \in \mathcal{U}, \quad (\#)$$

Now consider an arbitrary $\underline{A} \in \text{Sym } \mathcal{V}_3$. By the Spectral Theorem (Theorem A3.13.1),

$$\underline{A} = \sum_{i=1}^3 \lambda_i \underline{\hat{e}}_{\langle i \rangle}^* \otimes \underline{\hat{e}}_{\langle i \rangle}^*.$$

Then

$$\begin{aligned} \underline{G}(\underline{A}) &= \underline{G}\left(\sum_{i=1}^3 \lambda_i \underline{\hat{e}}_{\langle i \rangle}^* \otimes \underline{\hat{e}}_{\langle i \rangle}^*\right) && \text{(substitution)} \\ &= \sum_{i=1}^3 \lambda_i \underline{G}(\underline{\hat{e}}_{\langle i \rangle}^* \otimes \underline{\hat{e}}_{\langle i \rangle}^*) && \text{(linearity of } \underline{G} \text{)} \\ &= \sum_{i=1}^3 \lambda_i (\lambda \underline{I} + 2\mu \underline{\hat{e}}_{\langle i \rangle}^* \otimes \underline{\hat{e}}_{\langle i \rangle}^*) && \text{(representation \#)} \\ &= \lambda \left(\sum_{i=1}^3 \lambda_i\right) \underline{I} + 2\mu \left(\sum_{i=1}^3 \lambda_i \underline{\hat{e}}_{\langle i \rangle}^* \otimes \underline{\hat{e}}_{\langle i \rangle}^*\right) && \text{(tensor manipulations)} \\ &= \lambda(t_{\underline{A}}) \underline{I} + 2\mu \underline{A} && \text{(Theorems A3.11.8 and A3.13.4, and substitution).} \end{aligned}$$

The converse assertion is a trivial calculation. $\square$

Supplementary Reading

Same as for § AB.3 plus

SERRIN, "Mathematical principles of classical fluid mechanics",
Handbuch der Physik, vol. VIII/1 (1959) 125-263,
(edit. S. Flügge).

WANG, Mathematical Principles of Mechanics and Electromagnetism,
Part A: Analytical and Continuum Mechanics (1979)