NOTES ON CONTINUUM MECHANICS

ROUBEN ROSTAMIAN

Abstract. Here are brief notes on some of the central topics of continuum mechanics. These represent my take on the presentations in Gonzalez and Stuart [5], Gurtin [7], and Chadwick [2].

Contents

Date: September 10, 2023.

1. Notation

$\mathbb{E}_3$	= the three-dimensional Euclidean space
$\mathcal{V}$	= the linear space of vectors in $\mathbb{E}_3$
$\mathcal{L}$	= the linear space of second order tensors $\mathcal{V} \to \mathcal{V}$
$\mathcal{L}^+$	= the set of all tensors $\boldsymbol{A} \in \mathcal{L}$ with $\det \boldsymbol{A} > 0$
$\mathcal{L}_{\mathrm{sym}}$	= the linear subspace of symmetric tensors in $\mathcal{L}$
$\mathcal{L}_{\mathrm{skew}}$	= the linear subspace of skew-symmetric tensors in $\mathcal{L}$
$\mathcal{L}^+_{\mathrm{sym}}$	= the set of all symmetric and positive definite tensors
$\mathcal{L}_{\mathrm{orth}}$	= the linear space of orthogonal tensors in $\mathcal{L}$
$\mathcal{L}^+_{\mathrm{orth}}$	= the set of all tensors $\boldsymbol{Q} \in \mathcal{L}_{\mathrm{orth}}$ with $\det \boldsymbol{Q} > 0$

2. Vectors

Throughout these notes, $\mathbb{E}_3$ refers to the three-dimensional Euclidean space. The elements of $\mathbb{E}_3$ are *points*. The oriented line segment that extends from the point $\boldsymbol{x}$ to point $\boldsymbol{y}$ is called a *vector* and is written $\boldsymbol{v} = \boldsymbol{y} - \boldsymbol{x}$. We write $\mathcal{V}$ for the set of all vectors. Two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ which are parallel and have common orientation are regarded equal. Thus, $\mathcal{V}$ is the *equivalence class* of all vectors under this concept of equality.

We write $\|\boldsymbol{v}\|$ for the length of the vector $\boldsymbol{v}$. The *zero vector*, written $\boldsymbol{0}$, is the vector of zero length. The product $\alpha\boldsymbol{v}$ a vector $\boldsymbol{v} \in \mathcal{V}$ and a number $\alpha \in \mathbb{R}$, is obtained by scaling $\boldsymbol{v}$'s length by the factor α. A negative α both scales and reverses $\boldsymbol{v}$'s orientation. The sum $\boldsymbol{u}+\boldsymbol{v}$ of vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{V}$ is the vector that coincides with the diagonal of the parallelogram formed by $\boldsymbol{u}$ and $\boldsymbol{v}$.

The vectors $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$ are *linearly independent* if the equation $\alpha\boldsymbol{u} + \beta\boldsymbol{v} + \gamma\boldsymbol{w} = \boldsymbol{0}$, $\alpha, \beta, \gamma \in \mathbb{R}$, implies that $\alpha = \beta = \gamma = 0$.

The *scalar product* $\boldsymbol{u} \cdot \boldsymbol{v}$ of vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{V}$ is defined as $\boldsymbol{u} \cdot \boldsymbol{v} = \|\boldsymbol{u}\|\,\|\boldsymbol{v}\| \cos\theta$, where $\theta \in [0, \pi]$ is the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$. Let us note that if $\boldsymbol{u}$ and $\boldsymbol{v}$ are orthogonal, then $\boldsymbol{u} \cdot \boldsymbol{v} = 0$. Also note that $\boldsymbol{u} \cdot \boldsymbol{u} = \|\boldsymbol{u}\|^2$.

The *vector product* (also known as the *cross-product*) $\boldsymbol{u}\times\boldsymbol{v}$ of vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{V}$ is a vector $\boldsymbol{w}$ constructed as follows. If either of $\boldsymbol{u}$ or $\boldsymbol{v}$ is zero, or if $\boldsymbol{u}$ and $\boldsymbol{v}$ are collinear, then $\boldsymbol{w}$ is the zero vector. Otherwise, $\boldsymbol{w}$ is a vector of length $\|\boldsymbol{u}\|\,\|\boldsymbol{v}\| \sin\theta$ and is perpendicular to the plane formed by $\boldsymbol{u}$ and $\boldsymbol{v}$. Here $\theta \in [0, \pi]$ is the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$, as before. The conditions stipulated above determine $\boldsymbol{w}$ up to a multiplicative factor of ± 1. To disambiguate, we pick the one which conforms to the *right-hand rule*. It follows then $\boldsymbol{u}\times\boldsymbol{v} = -\boldsymbol{v}\times\boldsymbol{u}$. The

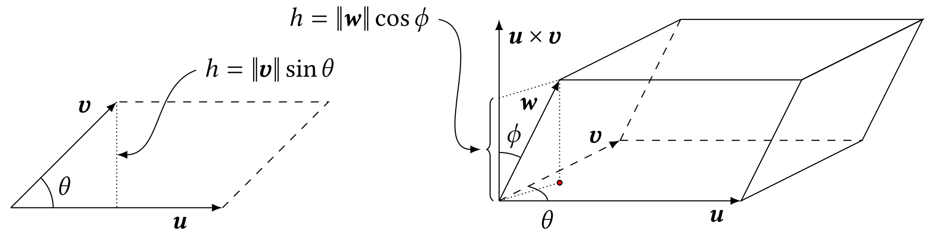


FIGURE 1. On the left: The area A of the parallelogram formed by the vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ is

$$A = \|\boldsymbol{u}\|\, h = \|\boldsymbol{u}\|\, \|\boldsymbol{v}\| \sin\theta = \|\boldsymbol{u} \times \boldsymbol{v}\|.$$

On the right: The volume V of the parallelepiped formed by the vectors $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ is the area of the base times the height, that is

$$V = \|\boldsymbol{u} \times \boldsymbol{v}\|\, h = \|\boldsymbol{u} \times \boldsymbol{v}\|\, \|\boldsymbol{w}\| \cos\phi = \big|(\boldsymbol{u} \times \boldsymbol{v}) \cdot \boldsymbol{w}\big| = \big|[\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}]\big|.$$

left diagram in Figure 1 shows that the area of the parallelogram formed by the vectors is $\boldsymbol{u}$ and $\boldsymbol{v}$ is $\|\boldsymbol{u} \times \boldsymbol{v}\|$.

Proposition 1.

$$\|\boldsymbol{u} \times \boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 \|\boldsymbol{v}\|^2 - (\boldsymbol{u} \cdot \boldsymbol{v})^2 \quad \textit{for all } \boldsymbol{u}, \boldsymbol{v} \in \mathcal{V}. \tag{1}$$

Proof. Let $\theta \in [0, \pi]$ be the angle between $\boldsymbol{u}$ and $\boldsymbol{v}$. We know that $\boldsymbol{u} \cdot \boldsymbol{v} = \|\boldsymbol{u}\|\, \|\boldsymbol{v}\| \cos\theta$, and $\|\boldsymbol{u} \times \boldsymbol{v}\| = \|\boldsymbol{u}\|\, \|\boldsymbol{v}\| \sin\theta$. Therefore

$$\begin{aligned} \|\boldsymbol{u} \times \boldsymbol{v}\|^2 &= \|\boldsymbol{u}\|^2 \|\boldsymbol{v}\|^2 \sin^2\theta = \|\boldsymbol{u}\|^2 \|\boldsymbol{v}\|^2 (1 - \cos^2\theta) \\ &= \|\boldsymbol{u}\|^2 \|\boldsymbol{v}\|^2 - \|\boldsymbol{u}\|^2 \|\boldsymbol{v}\|^2 \cos^2\theta = \|\boldsymbol{u}\|^2 \|\boldsymbol{v}\|^2 - (\boldsymbol{u} \cdot \boldsymbol{v})^2. \quad \square \end{aligned}$$

The *scalar triple product* $[\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}]$ of any three vectors $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$ is defined as $[\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}] = (\boldsymbol{u} \times \boldsymbol{v}) \cdot \boldsymbol{w}$. It follows from geometry that $[\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}]$ is the volume of the parallelepiped formed by the three vectors $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$. Consequently, $[\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}] = 0$ if and only if the vectors are linearly dependent.

The right diagram in Figure 1 shows that the volume of the parallelepiped formed by the vectors is $\boldsymbol{u}$, $\boldsymbol{v}$, and $\boldsymbol{w}$ is $\big|[\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}]\big|$. If the three vectors form a right-handed system, then the absolute value signs are redundant and the volume is simply $[\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}]$.

In the rest of these notes, we shorten the term *scalar triple product* to *triple product.*

Proposition 2. *For any $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$ we have*

$$[\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}] = [\boldsymbol{v}, \boldsymbol{w}, \boldsymbol{u}] = [\boldsymbol{w}, \boldsymbol{u}, \boldsymbol{v}] = -[\boldsymbol{u}, \boldsymbol{w}, \boldsymbol{v}] = -[\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{w}] = -[\boldsymbol{w}, \boldsymbol{v}, \boldsymbol{u}].$$

In words, the triple product $[\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}]$ is invariant under the cyclic permutation of its members, and it changes sign under a non-cyclic permutation.

Proof. Interchanging the first and second members of a triple product reverses its sign:

$$[\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}] = (\boldsymbol{u} \times \boldsymbol{v}) \cdot \boldsymbol{w} = -(\boldsymbol{v} \times \boldsymbol{u}) \cdot \boldsymbol{w} = -[\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{w}].$$

Furthermore, from the observation that the triple product is zero if its vectors are linearly dependent, it follows that

$$\begin{aligned} 0 &= [\boldsymbol{u}, \boldsymbol{v} + \boldsymbol{w}, \boldsymbol{v} + \boldsymbol{w}] \\ &= [\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{v} + \boldsymbol{w}] + [\boldsymbol{u}, \boldsymbol{w}, \boldsymbol{v} + \boldsymbol{w}] \\ &= [\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{v}] + [\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}] + [\boldsymbol{u}, \boldsymbol{w}, \boldsymbol{v}] + [\boldsymbol{u}, \boldsymbol{w}, \boldsymbol{w}] \\ &= [\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}] + [\boldsymbol{u}, \boldsymbol{w}, \boldsymbol{v}]. \end{aligned}$$

We conclude that $[\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}] = -[\boldsymbol{u}, \boldsymbol{w}, \boldsymbol{v}]$, that is, interchanging the second and third members of a triple product reverses its sign as well. The proposition's assertion follows from repeated interchanges of the triple product's members; see Exercise 1. $\square$

A *frame* is an orthonormal triplet of vectors. The orthonormality of the frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ is conveniently expressed in terms of the *Kronecker delta* δ_{ij}:

$$\boldsymbol{e}_i \cdot \boldsymbol{e}_j = \delta_{ij}, \quad i, j = 1, 2, 3, \tag{2}$$

where

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases} \tag{3}$$

We say that the frame is *right-handed* if

$$\boldsymbol{e}_1 \times \boldsymbol{e}_2 = \boldsymbol{e}_3, \quad \boldsymbol{e}_2 \times \boldsymbol{e}_3 = \boldsymbol{e}_1, \quad \boldsymbol{e}_3 \times \boldsymbol{e}_1 = \boldsymbol{e}_2. \tag{4}$$

Throughout these notes *all frames are right-handed.* The right-handedness is implicitly assumed even when it is not made explicit.

We will find it convenient to express the three equations in (4) as a single equation involving symbolic indices i, j, k, similar to that in (2). Toward that end, let us observe that for any i and j in the set $\{1, 2, 3\}$, we should be able to express the vector $\boldsymbol{e}_i \times \boldsymbol{e}_j$ as a linear combination of the frame's three vectors since the frame forms a basis for $\mathcal{V}$. Thus,

$$\boldsymbol{e}_i \times \boldsymbol{e}_j = \sum_{k=1}^{3} \epsilon_{ijk} \boldsymbol{e}_k, \tag{5}$$

The coefficients ϵ_{ijk} are determined as follows. If $i = j$, then $\boldsymbol{e}_i \times \boldsymbol{e}_j = \boldsymbol{0}$, while if $i \neq j$, then $\boldsymbol{e}_i \times \boldsymbol{e}_j = \pm \boldsymbol{e}_k$, where k is that element of the index set $\{1, 2, 3\}$ which is other than i and j. The plus or minus sign is determined according to whether the sequence of indices i, j, k is a cyclic or non-cyclic permutation of $\{1, 2, 3\}$. We conclude that the coefficients ϵ_{ijk} are given by

$$\epsilon_{ijk} = \begin{cases} 1 & \text{if } ijk \text{ is a cyclic permutation of } 123, \\ -1 & \text{if } ijk \text{ is a non-cyclic permutation of } 123, \\ 0 & \text{otherwise (that is, } ijk \text{ contains a repeated index).} \end{cases} \tag{6}$$

The ϵ_{ijk} is known as the *permutation symbol* or the *alternator.*

Any frame is a basis in $\mathcal{V}$, and therefore any vector $\boldsymbol{u} \in \mathcal{V}$ may be be expressed as a linear combination of the frame's members.

$$\boldsymbol{u} = u_1\, \boldsymbol{e}_1 + u_2\, \boldsymbol{e}_2 + u_3\, \boldsymbol{e}_3. \tag{7}$$

The coefficients u_1, u_2, and u_3 are called the *components of* $\boldsymbol{u}$ in the frame. Multiplying this by $\boldsymbol{e}_1$ we get $\boldsymbol{u}\cdot\boldsymbol{e}_1 = u_1\boldsymbol{e}_1\cdot\boldsymbol{e}_1 + u_2\boldsymbol{e}_2\cdot\boldsymbol{e}_1 + u_3\boldsymbol{e}_3\cdot\boldsymbol{e}_1 = u_1$ since $\boldsymbol{e}_1\cdot\boldsymbol{e}_1 = 1$ and $\boldsymbol{e}_1\cdot\boldsymbol{e}_2 = \boldsymbol{e}_1\cdot\boldsymbol{e}_3 = 0$. Similarly, $\boldsymbol{u}\cdot\boldsymbol{e}_2 = u_2$ and $\boldsymbol{u}\cdot\boldsymbol{e}_3 = u_3$. We conclude that

$$\boldsymbol{u} = (\boldsymbol{u}\cdot\boldsymbol{e}_1)\,\boldsymbol{e}_1 + (\boldsymbol{u}\cdot\boldsymbol{e}_2)\,\boldsymbol{e}_2 + (\boldsymbol{u}\cdot\boldsymbol{e}_3)\,\boldsymbol{e}_3 \quad \text{for all } \boldsymbol{u}\in\mathcal{V}. \tag{8}$$

The summation convention. Although expressions such as (7) and (8) may be compactly presented through the sigma notation $\boldsymbol{u} = \sum_{i=1}^{3} u_i\,\boldsymbol{e}_i$ and $\boldsymbol{u} = \sum_{i=1}^{3}(\boldsymbol{u}\cdot\boldsymbol{e}_i)\,\boldsymbol{e}_i$, it is quite common to adopt a more economical notation by dropping the sigma symbol altogether and writing them as $\boldsymbol{u} = u_i\,\boldsymbol{e}_i$ and $\boldsymbol{u} = (\boldsymbol{u}\cdot\boldsymbol{e}_i)\,\boldsymbol{e}_i$, under the implicit assumption that terms involving repeated indices are summed over as the repeated index ranges over 1, 2, 3. This is known as the *summation convention* or the *Einstein notation* as it was introduced by Albert Einstein in his paper on general relativity in 1916. By the same token, the vector multiplication in (5) may be expressed as $\boldsymbol{e}_i\times\boldsymbol{e}_j = \epsilon_{ijk}\,\boldsymbol{e}_k$. We will use the summation convention wherever possible throughout these notes.

In the expression $\boldsymbol{e}_i\times\boldsymbol{e}_j = \epsilon_{ijk}\,\boldsymbol{e}_k$, the symbols i and j are called *free indices* since they may take on any of the values of 1, 2, 3, but their values are unspecified at the moment. The symbol k is called a *dummy index* since it takes on the values 1, 2, 3 in the summation and then it disappears. The name of a dummy index is immaterial. For instance, $\epsilon_{ijk}\,\boldsymbol{e}_k = \epsilon_{ijq}\,\boldsymbol{e}_q$ because both sides expand to the same thing upon the evaluation of the implied summation.

To illustrate the economy of notation provided by the summation convention, consider the frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ and, following (8), express any two vectors $\boldsymbol{u}, \boldsymbol{v}\in V$ in terms of their components in the frame as $\boldsymbol{u} = u_i\,\boldsymbol{e}_i$ and $\boldsymbol{v} = v_j\,\boldsymbol{e}_j$, where the summation convention is in force. Then the dot product of the two vectors takes the form

$$\boldsymbol{u}\cdot\boldsymbol{v} = (u_i\boldsymbol{e}_i)\cdot(v_j\boldsymbol{e}_j) = u_iv_j\,\boldsymbol{e}_i\cdot\boldsymbol{e}_j = u_iv_j\,\delta_{ij} = u_iv_i. \tag{9}$$

In the last step of that calculation we have set $v_j\delta_{ij} = v_i$. To see why, let's revert to the sigma notation, $v_j\delta_{ij} = \sum_{j=1}^{3} v_j\delta_{ij}$ and observe that as j runs from 1 to 3, the value of δ_{ij} is nonzero only when j hits the value of i, and thus, the summation collapses to a single term, v_i.

In general, the expression $a_j\delta_{ij}$ always collapses to a_i, for any indexed variable a. This is known as *the transfer property of* δ_{ij}.

In the special case of $\boldsymbol{v} = \boldsymbol{u}$, equation (9) yields an expression for the length of $\boldsymbol{u}$ in terms of its components in the frame:

$$\|\boldsymbol{u}\|^2 = \boldsymbol{u}\cdot\boldsymbol{u} = u_iu_i$$

As another application, let $\boldsymbol{u} = u_i\,\boldsymbol{e}_i$ and $\boldsymbol{v} = v_j\,\boldsymbol{e}_j$, as before, and then calculate the vector product $\boldsymbol{u}\times\boldsymbol{v}$ in terms of the components of $\boldsymbol{u}$ and $\boldsymbol{v}$:

$$\boldsymbol{u}\times\boldsymbol{v} = (u_i\boldsymbol{e}_i)\times(v_j\boldsymbol{e}_j) = u_iv_j\,(\boldsymbol{e}_i\times\boldsymbol{e}_j) = u_iv_j\,(\epsilon_{ijk}\boldsymbol{e}_k) = \epsilon_{ijk}u_iv_j\,\boldsymbol{e}_k. \tag{10}$$

Building upon this calculation, we see that

$$(\boldsymbol{u}\times\boldsymbol{v})\cdot\boldsymbol{w} = (\epsilon_{ijk}u_iv_j\,\boldsymbol{e}_k)\cdot(w_p\boldsymbol{e}_p) = \epsilon_{ijk}u_iv_jw_p(\boldsymbol{e}_k\cdot\boldsymbol{e}_p) = \epsilon_{ijk}u_iv_jw_p\delta_{kp} = \epsilon_{ijk}u_iv_jw_k,$$

where in the last step we have applied the transfer property of the Kronecker delta. We conclude that

$$[\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}] = \epsilon_{ijk} u_i v_j w_k. \tag{11}$$

The special case of $[\boldsymbol{e}_i, \boldsymbol{e}_j, \boldsymbol{e}_k]$ is particularly interesting, and therefore let us make a record of it. From (5) we have $\boldsymbol{e}_i \times \boldsymbol{e}_j = \epsilon_{ijp}\boldsymbol{e}_p$. Therefore

$$[\boldsymbol{e}_i, \boldsymbol{e}_j, \boldsymbol{e}_k] = (\boldsymbol{e}_i \times \boldsymbol{e}_j) \cdot \boldsymbol{e}_k = \epsilon_{ijp}\boldsymbol{e}_p \cdot \boldsymbol{e}_k = \epsilon_{ijp}\delta_{pk} = \epsilon_{ijk},$$

that is

$$[\boldsymbol{e}_i, \boldsymbol{e}_j, \boldsymbol{e}_k] = \epsilon_{ijk}. \tag{12}$$

3. Linear functionals

We deviate from our focus on the previous section's three-dimensional vector space $\mathcal{V}$ and consider a general, abstract, vector space X, not necessarily three-dimensional, and not even equipped with a dot product. The generality is not essential to our work since eventually we will apply the results to $\mathcal{V}$, but the abstract nature of X helps to see the lines of argument more clearly.

A *functional* is linear function $f : X \to R$.[1] Let f be a functional on X and let N be its null space, that is

$$N = \left\{ \boldsymbol{x} \in X : f(\boldsymbol{x}) = 0 \right\}.$$

We leave the proofs of the following statements as exercises:

(1) N is a linear subspace of X, that is, if $\boldsymbol{u}$ and $\boldsymbol{v}$ are in N, and $\alpha, \beta \in \mathbb{R}$, then $\alpha\boldsymbol{u} + \beta\boldsymbol{v} \in N$.

(2) Suppose f is not identically zero (and therefore N is *not* the entire X). Then there exists a $\boldsymbol{q} \in X$ so that every $\boldsymbol{x} \in X$ admits the decomposition

$$\boldsymbol{x} = \boldsymbol{p} + \beta\boldsymbol{q},$$

where $\boldsymbol{p} \in N$ and $\beta \in \mathbb{R}$ depend on $\boldsymbol{x}$.

Hint: Let $\boldsymbol{q} \in X$ be such that $f(\boldsymbol{q}) \neq 0$ (why is there such a $\boldsymbol{q}$?) and $\boldsymbol{p} = \boldsymbol{x} - \dfrac{f(\boldsymbol{x})}{f(\boldsymbol{q})}\boldsymbol{q}$. Show that $\boldsymbol{p} \in N$.

Remark 1. Let M be the span of the vector $\boldsymbol{q}$. What we have shown above says that $X = N + M$ where M is one dimensional. In other words, the null space of a nontrivial function $f : X \to \mathbb{R}$ is of co-dimension 1.

(3) Suppose X is equipped with a dot product. Then for each functional f there exists a unique $\boldsymbol{y} \in X$ so that

$$f(\boldsymbol{x}) = \boldsymbol{x} \cdot \boldsymbol{y} \quad \text{for all } \boldsymbol{x} \in X.$$

Hint: First, address the easy special case when f is identically zero. Next, suppose f is not identically zero and let N be its null space. From the previous problem we know that the co-dimension of N is 1. Pick $\boldsymbol{n}$ to be a unit vector orthogonal

[1] The word "functional" is a short way of saying "a linear function whose range is $\mathbb{R}$". There is nothing particularly deep about it.

to N. Then any $\boldsymbol{x} \in X$ may be deocomposed as $\boldsymbol{x} = \boldsymbol{p} + \beta \boldsymbol{n}$ where $\boldsymbol{p} \in N$. Show that $f(\boldsymbol{x}) = \boldsymbol{x} \cdot \big(f(\boldsymbol{n})\,\boldsymbol{n}\big)$ and therefore $\boldsymbol{y} = f(\boldsymbol{n})\,\boldsymbol{n}$.

Remark 2. The vector $\boldsymbol{y}$ is called *the representation of* f.

Remark 3. With only minor changes, the statement above generalizes to the infinite-dimensional Hilbert space where it is called the *Riesz Representation Theorem.*

(4) Let X be a vector space equipped with a dot product. Consider the linear mapping $A : X \to X$. For a fixed $\boldsymbol{y} \in X$, define $f : X \to \mathbb{R}$ through $f(\boldsymbol{x}) = \big(A\boldsymbol{x}\big) \cdot \boldsymbol{y}$. Verify that f is linear. Let $\boldsymbol{y}^*$ be a representative of f (see the previous problem). Then $\big(A\boldsymbol{x}\big) \cdot \boldsymbol{y} = \boldsymbol{x} \cdot \boldsymbol{y}^*$ for all $\boldsymbol{x} \in X$. This construction associates with every $\boldsymbol{y} \in X$ a unique vector $\boldsymbol{y}^* \in X$. The mapping $A^* : \boldsymbol{y} \mapsto \boldsymbol{y}^*$ is called the *adjoint* of A. Verify that A^* is linear.

Remark 4. In view of $\boldsymbol{y}^* = A^*\boldsymbol{y}$, we have the frequently used identity

$$\big(A\boldsymbol{x}\big) \cdot \boldsymbol{y} = \boldsymbol{x} \cdot \big(A^*\boldsymbol{y}\big). \tag{13}$$

4. Second order tensors

A linear function $\boldsymbol{A} : \mathcal{V} \to \mathcal{V}$ is called a *second order tensor.* We write $\boldsymbol{\mathcal{L}}$ for the set of all second order tensors[2]. Thus, if $\boldsymbol{A} \in \boldsymbol{\mathcal{L}}$, $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{V}$, and $\alpha, \beta \in \mathbb{R}$, we have

$$\boldsymbol{A}(\alpha\boldsymbol{u} + \beta\boldsymbol{v}) = \alpha\boldsymbol{A}\boldsymbol{u} + \beta\boldsymbol{A}\boldsymbol{v}.$$

We will encounter *fourth order tensors* later in these notes, but when there is no chance of misunderstanding, we say *tensor* when we mean a second order tensor.

We equip the set of tensors with a linear space structure by defining the operations of addition and scalar multiplication on that set, as follows:

$$(\boldsymbol{A} + \boldsymbol{B})(\boldsymbol{u}) = \boldsymbol{A}\boldsymbol{u} + \boldsymbol{B}\boldsymbol{u} \quad \text{and} \quad (\alpha\boldsymbol{A})\boldsymbol{u} = \alpha(\boldsymbol{A}\boldsymbol{u}) \tag{14}$$

for all $\boldsymbol{A}, \boldsymbol{B} \in \boldsymbol{\mathcal{L}}$, $\alpha \in \mathbb{R}$, and $\boldsymbol{u} \in \mathcal{V}$.

The *zero tensor* $\boldsymbol{0}$ maps every vector to the zero vector, and the *identity tensor* $\boldsymbol{I}$ maps every vector to itself. Thus

$$\boldsymbol{0}\boldsymbol{u} = \boldsymbol{0}, \quad \boldsymbol{I}\boldsymbol{u} = \boldsymbol{u}, \qquad \text{for all } \boldsymbol{u} \in \mathcal{V}.$$

In continuum mechanics, the adjoint A^* of A (recall (13)) is traditionally called the *transpose* of $\boldsymbol{A}$ and is written $\boldsymbol{A}^T$. Thus

$$\boldsymbol{u} \cdot \big(\boldsymbol{A}^T\boldsymbol{v}\big) = \big(\boldsymbol{A}\boldsymbol{u}\big) \cdot \boldsymbol{v} \qquad \text{for all } \boldsymbol{u}, \boldsymbol{v} \in \mathcal{V}. \tag{15}$$

It follows that

$$\big(\boldsymbol{A}^T\big)^T = \boldsymbol{A}, \quad (\alpha\boldsymbol{A} + \beta\boldsymbol{B})^T = \alpha\boldsymbol{A}^T + \beta\boldsymbol{B}^T, \quad (\boldsymbol{A}\boldsymbol{B})^T = \boldsymbol{B}^T\boldsymbol{A}^T$$

for all $\boldsymbol{A}, \boldsymbol{B} \in \boldsymbol{\mathcal{L}}$ and all $\alpha, \beta \in \mathbb{R}$.

[2] Thus, a second order tensor is what is called a *linear operator* on the vector space $\mathcal{V}$ in operator theory and functional analysis. Following the well-established tradition, we write $\boldsymbol{A}\boldsymbol{u}$ rather than $\boldsymbol{A}(\boldsymbol{u})$ whenever unambiguous, to indicate the action of the tensor $\boldsymbol{A}$ on the vector $\boldsymbol{u}$.

A tensor $\boldsymbol{A}$ is called *symmetric*[3] if $\boldsymbol{A}^T = \boldsymbol{A}$, and *skew-symmetric* if if $\boldsymbol{A}^T = -\boldsymbol{A}$. The identity

$$\boldsymbol{A} = \frac{1}{2}\big(\boldsymbol{A} + \boldsymbol{A}^T\big) + \frac{1}{2}\big(\boldsymbol{A} - \boldsymbol{A}^T\big)$$

shows that any tensor may be expressed as the sum of a symmetric and skew-symmetric tensors.

Theorem 1. *Associated with any tensor $\boldsymbol{A} \in \mathcal{L}$ there are three numbers, called its principal invariants and written $\iota_1(\boldsymbol{A})$, $\iota_2(\boldsymbol{A})$, $\iota_3(\boldsymbol{A})$, such that for all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \mathcal{V}$ we have:*

$$[\boldsymbol{A}\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}] + [\boldsymbol{u}, \boldsymbol{A}\boldsymbol{v}, \boldsymbol{w}] + [\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{A}\boldsymbol{w}] = \iota_1(\boldsymbol{A})\,[\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}], \tag{16a}$$

$$[\boldsymbol{u}, \boldsymbol{A}\boldsymbol{v}, \boldsymbol{A}\boldsymbol{w}] + [\boldsymbol{A}\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{A}\boldsymbol{w}] + [\boldsymbol{A}\boldsymbol{u}, \boldsymbol{A}\boldsymbol{v}, \boldsymbol{w}] = \iota_2(\boldsymbol{A})\,[\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}], \tag{16b}$$

$$[\boldsymbol{A}\boldsymbol{u}, \boldsymbol{A}\boldsymbol{v}, \boldsymbol{A}\boldsymbol{w}] = \iota_3(\boldsymbol{A})\,[\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}]. \tag{16c}$$

Proof. We will prove (16a) and leave the proofs of (16b) and (16c) as exercises to the reader.

Pick any frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ and express $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ in terms of components, that is, $\boldsymbol{u} = u_i\boldsymbol{e}_i$, $\boldsymbol{v} = v_j\boldsymbol{e}_j$, $\boldsymbol{w} = w_k\boldsymbol{e}_k$. Then

$$\begin{aligned}
&[\boldsymbol{A}\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}] + [\boldsymbol{u}, \boldsymbol{A}\boldsymbol{v}, \boldsymbol{w}] + [\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{A}\boldsymbol{w}] \\
&\qquad = [\boldsymbol{A}(u_i\boldsymbol{e}_i), v_j\boldsymbol{e}_j, w_k\boldsymbol{e}_k] + [u_i\boldsymbol{e}_i, \boldsymbol{A}(v_j\boldsymbol{e}_j), w_k\boldsymbol{e}_k] + [u_i\boldsymbol{e}_i, v_j\boldsymbol{e}_j, \boldsymbol{A}(w_k\boldsymbol{e}_k)] \\
&\qquad = u_i v_j w_k\Big([\boldsymbol{A}\boldsymbol{e}_i, \boldsymbol{e}_j, \boldsymbol{e}_k] + [\boldsymbol{e}_i, \boldsymbol{A}\boldsymbol{e}_j, \boldsymbol{e}_k] + [\boldsymbol{e}_i, \boldsymbol{e}_j, \boldsymbol{A}\boldsymbol{e}_k]\Big) \\
&\qquad = u_i v_j w_k M_{ijk},
\end{aligned} \tag{17}$$

where we have set

$$M_{ijk} = [\boldsymbol{A}\boldsymbol{e}_i, \boldsymbol{e}_j, \boldsymbol{e}_k] + [\boldsymbol{e}_i, \boldsymbol{A}\boldsymbol{e}_j, \boldsymbol{e}_k] + [\boldsymbol{e}_i, \boldsymbol{e}_j, \boldsymbol{A}\boldsymbol{e}_k].$$

Let us verify that M_{ijk} is unchanged by a cyclic permutation of indices, that is, $\{i, j, k\} \to \{k, i, j\}$. We have

$$\begin{aligned}
M_{kij} &= [\boldsymbol{A}\boldsymbol{e}_k, \boldsymbol{e}_i, \boldsymbol{e}_j] + [\boldsymbol{e}_k, \boldsymbol{A}\boldsymbol{e}_i, \boldsymbol{e}_j] + [\boldsymbol{e}_k, \boldsymbol{e}_i, \boldsymbol{A}\boldsymbol{e}_j] \\
&= [\boldsymbol{e}_i, \boldsymbol{e}_j, \boldsymbol{A}\boldsymbol{e}_k] + [\boldsymbol{A}\boldsymbol{e}_i, \boldsymbol{e}_j, \boldsymbol{e}_k] + [\boldsymbol{e}_i, \boldsymbol{A}\boldsymbol{e}_j, \boldsymbol{e}_k] \\
&= M_{ijk}.
\end{aligned}$$

Let us verify that M_{ijk} is unchanged by a non-cyclic permutation of indices, e.g., $\{i, j, k\} \to \{j, i, k\}$. We have

$$\begin{aligned}
M_{jik} &= [\boldsymbol{A}\boldsymbol{e}_j, \boldsymbol{e}_i, \boldsymbol{e}_k] + [\boldsymbol{e}_j, \boldsymbol{A}\boldsymbol{e}_i, \boldsymbol{e}_k] + [\boldsymbol{e}_j, \boldsymbol{e}_i, \boldsymbol{A}\boldsymbol{e}_k] \\
&= -[\boldsymbol{e}_i, \boldsymbol{A}\boldsymbol{e}_j, \boldsymbol{e}_k] - [\boldsymbol{e}_i, \boldsymbol{e}_j, \boldsymbol{A}\boldsymbol{e}_k] - [\boldsymbol{e}_i, \boldsymbol{e}_j, \boldsymbol{A}\boldsymbol{e}_k] \\
&= -M_{ijk}.
\end{aligned}$$

Let us verify that $M_{ijk} = 0$ if any of its indices is repeated. Take, for instance, $i = j$. Then

$$\begin{aligned}
M_{iik} &= [\boldsymbol{A}\boldsymbol{e}_i, \boldsymbol{e}_i, \boldsymbol{e}_k] + [\boldsymbol{e}_i, \boldsymbol{A}\boldsymbol{e}_i, \boldsymbol{e}_k] + [\boldsymbol{e}_i, \boldsymbol{e}_i, \boldsymbol{A}\boldsymbol{e}_k] \\
&= [\boldsymbol{A}\boldsymbol{e}_i, \boldsymbol{e}_i, \boldsymbol{e}_k] - [\boldsymbol{A}\boldsymbol{e}_i, \boldsymbol{e}_i, \boldsymbol{e}_k] + [\boldsymbol{e}_i, \boldsymbol{e}_i, \boldsymbol{A}\boldsymbol{e}_k] \\
&= 0.
\end{aligned}$$

[3]The equivalent term in operator theory is *self-adjoint*.

We see that M_{ijk} is unchanged by a cyclic permutation of indices, its sign reverses by a non-cyclic permutation of indices, and is zero when any there are repeated indices. From the previous lemma it follows that $M_{ijk} = C\epsilon_{ijk}$ for some C. The factor C may be evaluated by taking taking any cyclic permutation of $\{1, 2, 3\}$ for $\{i, j, k\}$. Thus, $M_{ijk} = \epsilon_{ijk}M_{123}$. Consequently, equation (17) takes the form

$$[\boldsymbol{Au}, \boldsymbol{v}, \boldsymbol{w}] + [\boldsymbol{u}, \boldsymbol{Av}, \boldsymbol{w}] + [\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{Aw}] = u_i v_j w_k \epsilon_{ijk} M_{123},$$

which in view of (11) becomes

$$[\boldsymbol{Au}, \boldsymbol{v}, \boldsymbol{w}] + [\boldsymbol{u}, \boldsymbol{Av}, \boldsymbol{w}] + [\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{Aw}] = M_{123}[\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}], \tag{18}$$

Comparing this with (16a) we conclude that $\iota_1(\boldsymbol{A}) = M_{123}$, that is

$$\iota_1(\boldsymbol{A}) = [\boldsymbol{Ae}_1, \boldsymbol{e}_2, \boldsymbol{e}_3] + [\boldsymbol{e}_1, \boldsymbol{Ae}_2, \boldsymbol{e}_3] + [\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{Ae}_3].$$

To complete the proof, we observe that despite the appearances, M_{123} (and therefore $\iota_1(\boldsymbol{A})$) is independent of the choice of the frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$. Indeed, if it varied with the choice of the frame, then (18) couldn't hold since $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ are independent of the frame. □

Remark 5. The preceding proof establishes an explicit formula for the first principal invariant $\iota_1(\boldsymbol{A})$. For future reference, here we summarize this here along with the corresponding formulas for ι_2 and ι_3 which you will derive in the exercises:

$$\iota_1(\boldsymbol{A}) = [\boldsymbol{Ae}_1, \boldsymbol{e}_2, \boldsymbol{e}_3] + [\boldsymbol{e}_1, \boldsymbol{Ae}_2, \boldsymbol{e}_3] + [\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{Ae}_3], \tag{19a}$$

$$\iota_2(\boldsymbol{A}) = [\boldsymbol{e}_1, \boldsymbol{Ae}_2, \boldsymbol{Ae}_3] + [\boldsymbol{Ae}_1, \boldsymbol{e}_2, \boldsymbol{Ae}_3] + [\boldsymbol{Ae}_1, \boldsymbol{Ae}_2, \boldsymbol{e}_3], \tag{19b}$$

$$\iota_3(\boldsymbol{A}) = [\boldsymbol{Ae}_1, \boldsymbol{Ae}_2, \boldsymbol{Ae}_3]. \tag{19c}$$

Let's reiterate that the principal invariants $\iota_1(\boldsymbol{A})$, $\iota_2(\boldsymbol{A})$, and $\iota_3(\boldsymbol{A})$ are independent of the choice of the frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$.

5. The dyadic product

The *dyadic product* (also known as the *tensor product*) of vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathcal{V}$, is the second order tensor $\boldsymbol{u} \otimes \boldsymbol{v}$ defined through its action on vectors:

$$(\boldsymbol{u} \otimes \boldsymbol{v})\,\boldsymbol{x} = (\boldsymbol{v} \cdot \boldsymbol{x})\,\boldsymbol{u}, \quad \text{for all } \boldsymbol{x} \in \mathcal{V}. \tag{20}$$

Proposition 3. *For all* $\boldsymbol{A}, \boldsymbol{B} \in \mathcal{L}$, $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}, \boldsymbol{a}, \boldsymbol{b} \in \mathcal{V}$, *and* $\alpha, \beta \in \mathbb{R}$, *we have*

$$(\alpha\boldsymbol{u} + \beta\boldsymbol{v}) \otimes \boldsymbol{w} = \alpha\boldsymbol{u} \otimes \boldsymbol{w} + \beta\boldsymbol{v} \otimes \boldsymbol{w}, \tag{21a}$$

$$(\boldsymbol{u} \otimes \boldsymbol{v})^T = \boldsymbol{v} \otimes \boldsymbol{u}, \tag{21b}$$

$$(\boldsymbol{a} \otimes \boldsymbol{b})\,(\boldsymbol{u} \otimes \boldsymbol{v}) = (\boldsymbol{b} \cdot \boldsymbol{u})\,(\boldsymbol{a} \otimes \boldsymbol{v}), \tag{21c}$$

$$\boldsymbol{A}\,(\boldsymbol{u} \otimes \boldsymbol{v}) = (\boldsymbol{Au}) \otimes \boldsymbol{v}, \tag{21d}$$

$$(\boldsymbol{u} \otimes \boldsymbol{v})\,\boldsymbol{A} = \boldsymbol{u} \otimes (\boldsymbol{A}^T\boldsymbol{v}), \tag{21e}$$

$$\boldsymbol{A}(\boldsymbol{u} \otimes \boldsymbol{v})\boldsymbol{B} = (\boldsymbol{Au}) \otimes (\boldsymbol{B}^T\boldsymbol{b}), \tag{21f}$$

$$\iota_1(\boldsymbol{u} \otimes \boldsymbol{v}) = \boldsymbol{u} \cdot \boldsymbol{v}, \tag{21g}$$

$$\iota_2(\boldsymbol{u} \otimes \boldsymbol{v}) = 0, \tag{21h}$$

$$\iota_3(\boldsymbol{u} \otimes \boldsymbol{v}) = 0, \tag{21i}$$

Proof. Here we will verify the identities (21b) and (21g), leaving the remaining as exercises.

To verify (21b), pick any $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{L}$, and then calculate

$$\boldsymbol{x} \cdot \left[(\boldsymbol{u} \otimes \boldsymbol{v})^T \boldsymbol{y}\right] \overset{\text{by (15)}}{=} \left[(\boldsymbol{u} \otimes \boldsymbol{v})\boldsymbol{x}\right] \cdot \boldsymbol{y} \overset{\text{by (20)}}{=} \left[(\boldsymbol{v} \cdot \boldsymbol{x})\, \boldsymbol{u}\right] \cdot \boldsymbol{y}$$
$$= (\boldsymbol{v} \cdot \boldsymbol{x})\,(\boldsymbol{u} \cdot \boldsymbol{y}) = \boldsymbol{x} \cdot \left[(\boldsymbol{u} \cdot \boldsymbol{y})\, \boldsymbol{v}\right] \overset{\text{by (20)}}{=} \boldsymbol{x} \cdot \left[(\boldsymbol{v} \otimes \boldsymbol{u})\, \boldsymbol{y}\right].$$

Therefore, $\boldsymbol{x} \cdot \left[(\boldsymbol{u} \otimes \boldsymbol{v})^T \boldsymbol{y} - (\boldsymbol{v} \otimes \boldsymbol{u})\, \boldsymbol{y}\right] = 0$ for all $\boldsymbol{x} \in \mathcal{V}$. It follows that $(\boldsymbol{u} \otimes \boldsymbol{v})^T \boldsymbol{y} - (\boldsymbol{v} \otimes \boldsymbol{u})\, \boldsymbol{y} = \boldsymbol{0}$, that is $\left[(\boldsymbol{u} \otimes \boldsymbol{v})^T - (\boldsymbol{v} \otimes \boldsymbol{u})\right] \boldsymbol{y} = 0$ for all $\boldsymbol{y} \in \mathcal{V}$, and thus $(\boldsymbol{u} \otimes \boldsymbol{v})^T - \boldsymbol{v} \otimes \boldsymbol{u} = \boldsymbol{0}$, whence $(\boldsymbol{u} \otimes \boldsymbol{v})^T = \boldsymbol{v} \otimes \boldsymbol{u}$, as asserted.

To verify (21g), recall the representation of $\iota_1(\boldsymbol{A})$ in (19a) and evaluate the first term on its right-hand side with $\boldsymbol{A} = \boldsymbol{u} \otimes \boldsymbol{v}$:

$$[(\boldsymbol{u} \otimes \boldsymbol{v})\, \boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3] = [(\boldsymbol{v} \cdot \boldsymbol{e}_1)\, \boldsymbol{u}, \boldsymbol{e}_2, \boldsymbol{e}_3] = v_1[\boldsymbol{u}, \boldsymbol{e}_2, \boldsymbol{e}_3]$$
$$= v_1[u_i \boldsymbol{e}_i, \boldsymbol{e}_2, \boldsymbol{e}_3] = u_i v_1 [\boldsymbol{e}_i, \boldsymbol{e}_2, \boldsymbol{e}_3] \overset{\text{by (12)}}{=} u_i v_1 \epsilon_{i23} = u_1 v_1.$$

The last step is based on the observation that ϵ_{i23} is nonzero only when $i = 1$.

Repeating the calculation with the second and third terms on the right-hand side of (19a) we arrive at

$$\iota_1(\boldsymbol{u} \otimes \boldsymbol{v}) = u_1 v_1 + u_2 v_2 + u_3 v_3 = u_i v_i \overset{\text{by (9)}}{=} \boldsymbol{u} \cdot \boldsymbol{v}.$$

□

6. Component form of vectors and tensors

Theorem 2. *Let $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ be a frame in $\mathcal{V}$. Then for an $\boldsymbol{A} \in \mathcal{L}$ we have*

$$\boldsymbol{A} = (\boldsymbol{e}_i \cdot \boldsymbol{A}\boldsymbol{e}_j)\, \boldsymbol{e}_i \otimes \boldsymbol{e}_j. \tag{22}$$

Proof. Recall that according to (8), any $\boldsymbol{u} \in \mathcal{V}$ may be expressed as $\boldsymbol{u} = (\boldsymbol{u} \cdot \boldsymbol{e}_j)\, \boldsymbol{e}_j$. Therefore

$$\boldsymbol{A}\boldsymbol{u} = \boldsymbol{A}\big((\boldsymbol{u} \cdot \boldsymbol{e}_j)\, \boldsymbol{e}_j\big) = (\boldsymbol{u} \cdot \boldsymbol{e}_j)\, \boldsymbol{A}\boldsymbol{e}_j.$$

Now, $\boldsymbol{A}\boldsymbol{e}_j$ is an element of $\mathcal{V}$, and again by (8) it may be expressed as $\boldsymbol{A}\boldsymbol{e}_j = (\boldsymbol{A}\boldsymbol{e}_j \cdot \boldsymbol{e}_i)\, \boldsymbol{e}_i$, and therefore

$$\boldsymbol{A}\boldsymbol{u} = (\boldsymbol{u} \cdot \boldsymbol{e}_j)\big[\,(\boldsymbol{A}\boldsymbol{e}_j \cdot \boldsymbol{e}_i)\, \boldsymbol{e}_i\big] = (\boldsymbol{A}\boldsymbol{e}_j \cdot \boldsymbol{e}_i)\,(\boldsymbol{u} \cdot \boldsymbol{e}_j)\, \boldsymbol{e}_i = (\boldsymbol{A}\boldsymbol{e}_j \cdot \boldsymbol{e}_i)\,(\boldsymbol{e}_i \otimes \boldsymbol{e}_j)\, \boldsymbol{u}.$$

Since $\boldsymbol{u}$ is arbitrary, it follows that $\boldsymbol{A} = (\boldsymbol{A}\boldsymbol{e}_j \cdot \boldsymbol{e}_i)\,(\boldsymbol{e}_i \otimes \boldsymbol{e}_j)$. □

Corollary 1. *Let $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ be a frame in $\mathcal{V}$. Then the set*

$$\mathcal{K} = \Big\{\, \boldsymbol{e}_i \otimes \boldsymbol{e}_j \;:\; i, j \in \{1, 2, 3\} \Big\} \tag{23}$$

is a basis for $\mathcal{L}$, and in particular, the space $\mathcal{L}$ is 9-dimensional.

Proof. According to Theorem 2, any $\boldsymbol{A} \in \mathcal{L}$ may be formed as a linear combination of the elements of $\mathcal{K}$. All that remains is to show that the set $\mathcal{K}$ is linearly independent. We leave that for an exercise. □

Equation (22) expresses the second order tensor $\boldsymbol{A}$ as a linear combination of the basis elements from $\mathcal{K}$. The nine (scalar) coefficients $a_{ij} = \boldsymbol{e}_i \cdot \boldsymbol{A}\boldsymbol{e}_j$ are *the components of the tensor $\boldsymbol{A}$ in the basis $\mathcal{K}$*. Thus, a tensor $\boldsymbol{A}$ may be expressed in terms of its components as $\boldsymbol{A} = a_{ij}\boldsymbol{e}_i \otimes \boldsymbol{e}_j$ in the same way a vector $\boldsymbol{u}$ is expressed in terms of its components $\boldsymbol{u} = u_i\boldsymbol{e}_i$. Let's observe that

$$\boldsymbol{A}\boldsymbol{u} = (a_{ij}\boldsymbol{e}_i \otimes \boldsymbol{e}_j)(u_p\boldsymbol{e}_p) = a_{ij}u_p\,(\boldsymbol{e}_i \otimes \boldsymbol{e}_j)\,\boldsymbol{e}_p = a_{ij}u_p\,(\boldsymbol{e}_j \cdot \boldsymbol{e}_p)\,\boldsymbol{e}_i = a_{ij}u_p\,\delta_{jp}\,\boldsymbol{e}_i = a_{ij}u_j\,\boldsymbol{e}_i.$$

We see that the components $a_{ij}u_j$ of the vector $\boldsymbol{A}\boldsymbol{u}$ are obtained by multiplying the matrix $[a_{ij}]$ and the vector $[u_j]$. This leads to the commutative diagram

$$\begin{tikzcd}
\boldsymbol{u} \arrow[r, "A"] \arrow[d, "\{\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3\}"'] & \boldsymbol{v} = \boldsymbol{A}\boldsymbol{u} \\
u_j \arrow[r, "a_{ij}"] & v_i = a_{ij}u_j \arrow[u, "\{\boldsymbol{e}_1,\boldsymbol{e}_2,\boldsymbol{e}_3\}"']
\end{tikzcd}$$

Example 1. Let a_{ij} and b_{ij} be the components of the second order tensors $\boldsymbol{A}$ and $\boldsymbol{B}$ in the basis $\mathcal{K}$. Show that the components of $\boldsymbol{C} = \boldsymbol{A}\boldsymbol{B}$ relative to $\mathcal{K}$ are $c_{ij} = a_{ip}b_{pj}$.

Solution. From (22) and (15) we have $\boldsymbol{A}\boldsymbol{B} = (\boldsymbol{e}_i \cdot \boldsymbol{A}\boldsymbol{B}\,\boldsymbol{e}_j)\,\boldsymbol{e}_i \otimes \boldsymbol{e}_j = (\boldsymbol{A}^T\boldsymbol{e}_i \cdot \boldsymbol{B}\,\boldsymbol{e}_j)\,\boldsymbol{e}_i \otimes \boldsymbol{e}_j$. However,

$$\begin{aligned}
\boldsymbol{A}^T\boldsymbol{e}_i &= (\boldsymbol{e}_p \cdot \boldsymbol{A}^T\boldsymbol{e}_i)\,\boldsymbol{e}_p = (\boldsymbol{A}\boldsymbol{e}_p \cdot \boldsymbol{e}_i)\,\boldsymbol{e}_p, = (\boldsymbol{e}_i \cdot \boldsymbol{A}\boldsymbol{e}_p)\,\boldsymbol{e}_p = a_{ip}\,\boldsymbol{e}_p,\\
\boldsymbol{B}\boldsymbol{e}_j &= (\boldsymbol{e}_q \cdot \boldsymbol{B}\boldsymbol{e}_j)\,\boldsymbol{e}_q = b_{qj}\,\boldsymbol{e}_q,
\end{aligned}$$

and therefore

$$\boldsymbol{A}^T\boldsymbol{e}_i \cdot \boldsymbol{B}\,\boldsymbol{e}_j = (a_{ip}\,\boldsymbol{e}_p)\,(b_{qj}\,\boldsymbol{e}_q) = a_{ip}b_{qj}\,\boldsymbol{e}_p \cdot \boldsymbol{e}_q = a_{ip}b_{qj}\,\delta_{pq} = a_{ip}b_{pj}.$$

We conclude that $\boldsymbol{C} = \boldsymbol{A}\boldsymbol{B} = a_{ip}b_{pj}\,\boldsymbol{e}_i \otimes \boldsymbol{e}_j$, whence $c_{ij} = a_{ip}b_{pj}$.

This slightly different but equivalent calculation does not invoke the components a_{ip} and b_{qj} directly:

$$\begin{aligned}
\boldsymbol{A}^T\boldsymbol{e}_i \cdot \boldsymbol{B}\,\boldsymbol{e}_j = \big[(\boldsymbol{e}_i \cdot \boldsymbol{A}\boldsymbol{e}_p)\,\boldsymbol{e}_p\big] \cdot \big[(\boldsymbol{e}_q \cdot \boldsymbol{B}\boldsymbol{e}_j)\,\boldsymbol{e}_q\big] &= (\boldsymbol{e}_i \cdot \boldsymbol{A}\boldsymbol{e}_p)\,(\boldsymbol{e}_q \cdot \boldsymbol{B}\boldsymbol{e}_j)\,(\boldsymbol{e}_p \cdot \boldsymbol{e}_q)\\
&= (\boldsymbol{e}_i \cdot \boldsymbol{A}\boldsymbol{e}_p)\,(\boldsymbol{e}_q \cdot \boldsymbol{B}\boldsymbol{e}_j)\,\delta_{pq} = (\boldsymbol{e}_i \cdot \boldsymbol{A}\boldsymbol{e}_p)\,(\boldsymbol{e}_p \cdot \boldsymbol{B}\boldsymbol{e}_j).
\end{aligned}$$

Choose whichever approach you prefer.

Note that $c_{ij} = a_{ip}b_{pj}$ states that the matrix $[c_{ij}]$ is the product of the matrices $[a_{ip}]$ and $[b_{pj}]$. □

What are the components of the tensor $\boldsymbol{u} \otimes \boldsymbol{v}$? Let's calculate:

$$\boldsymbol{u} \otimes \boldsymbol{v} = (u_i\,\boldsymbol{e}_i) \otimes (v_j\,\boldsymbol{e}_j) = u_iv_j\,\boldsymbol{e}_i \otimes \boldsymbol{e}_j.$$

We conclude that the components of the tensor $\boldsymbol{u} \otimes \boldsymbol{v}$ are u_iv_j.

Example 2. Let a_{ij} be the components of the tensor $\boldsymbol{A}$ in the frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$, that is, $\boldsymbol{A} = a_{ij}\,\boldsymbol{e}_i \otimes \boldsymbol{e}_j$. Show that $\iota_1(\boldsymbol{A}) = a_{ii}$.

Solution. Recall the representation of $\iota_1(\boldsymbol{A})$ in (19a). Let's evaluate the first term on the right-hand side:

$$[\boldsymbol{A}\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3] = [a_{ij}(\boldsymbol{e}_i \otimes \boldsymbol{e}_j)\,\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3] = [a_{ij}(\boldsymbol{e}_1 \cdot \boldsymbol{e}_j)\,\boldsymbol{e}_i, \boldsymbol{e}_2, \boldsymbol{e}_3]$$
$$= [a_{ij}\delta_{j1}\,\boldsymbol{e}_i, \boldsymbol{e}_2, \boldsymbol{e}_3] = [a_{i1}\,\boldsymbol{e}_i, \boldsymbol{e}_2, \boldsymbol{e}_3] \overset{\text{by (12)}}{=} a_{i1}\epsilon_{i23} = a_{11}.$$

The last step is based on the observation that ϵ_{i23} is nonzero only when $i = 1$.

Evaluating the remaining terms on the right-hand side of (19a), we conclude that $\iota_1(\boldsymbol{A}) = a_{11} + a_{22} + a_{33} = a_{ii}$. □

Remark 6. Let us observe that the expression a_{ii} is the *trace* of the matrix $[a_{ij}]$ of the components of $\boldsymbol{A}$ in the frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$. Since $\iota_1(\boldsymbol{A})$ is independent of the choice of frame, the equality $\iota_1(\boldsymbol{A}) = a_{ii}$ shows that the trace a_{ii} is frame-independent. That's despite the fact that the individual components a_{ij} vary with the frame. For this reason, the first principal invariant, $\iota_1(\boldsymbol{A})$, is called *the trace of* $\boldsymbol{A}$. We write this as

$$\iota_1(\boldsymbol{A}) = \operatorname{tr} \boldsymbol{A}. \tag{24a}$$

Example 3. Let a_{ij} be the components of the tensor $\boldsymbol{A}$ in the frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$, that is, $\boldsymbol{A} = a_{ij}\,\boldsymbol{e}_i \otimes \boldsymbol{e}_j$. Show that

$$\iota_2(\boldsymbol{A}) = \frac{1}{2}\Big(\big(\operatorname{tr} \boldsymbol{A}\big)^2 - \operatorname{tr}\big(\boldsymbol{A}^2\big)\Big), \tag{24b}$$

Solution. As in the previous example, we begin with calculating the first term on the right-hand side of (19b):

$$[\boldsymbol{e}_1, \boldsymbol{A}\boldsymbol{e}_2, \boldsymbol{A}\boldsymbol{e}_3] = [\boldsymbol{e}_1, a_{ij}(\boldsymbol{e}_i \otimes \boldsymbol{e}_j)\,\boldsymbol{e}_2, a_{pq}(\boldsymbol{e}_p \otimes \boldsymbol{e}_q)\,\boldsymbol{e}_3] = [\boldsymbol{e}_1, a_{ij}(\boldsymbol{e}_2 \cdot \boldsymbol{e}_j)\,\boldsymbol{e}_i, a_{pq}(\boldsymbol{e}_3 \cdot \boldsymbol{e}_q)\,\boldsymbol{e}_p]$$
$$= [\boldsymbol{e}_1, a_{ij}\delta_{2j}\,\boldsymbol{e}_i, a_{pq}\delta_{3q}\,\boldsymbol{e}_p] = [\boldsymbol{e}_1, a_{i2}\,\boldsymbol{e}_i, a_{p3}\,\boldsymbol{e}_p] = a_{i2}a_{p3}[\boldsymbol{e}_1, \boldsymbol{e}_i, \boldsymbol{e}_p] = a_{i2}a_{p3}\epsilon_{1ip}.$$

In the summation over the index i in the term $a_{i2}a_{p3}\epsilon_{1ip}$, the coefficient ϵ_{1ip} is zero when $i = 1$, therefore we need to consider only $i = 2$ and $i = 3$. Expanding the sum or i, we get:

$$[\boldsymbol{e}_1, \boldsymbol{A}\boldsymbol{e}_2, \boldsymbol{A}\boldsymbol{e}_3] = a_{22}a_{p3}\epsilon_{12p} + a_{32}a_{p3}\epsilon_{13p}.$$

Now, in the sum $a_{22}a_{p3}\epsilon_{12p}$, the only nonzero term is obtained for $p = 3$ since $\epsilon_{12p} = 1$ when $p = 3$ and is zero otherwise. Similarly, in the sum $a_{32}a_{p3}\epsilon_{13p}$, the only nonzero term is obtained for $p = 2$ since $\epsilon_{13p} = -1$ when $p = 2$ and is zero otherwise. We conclude that

$$[\boldsymbol{e}_1, \boldsymbol{A}\boldsymbol{e}_2, \boldsymbol{A}\boldsymbol{e}_3] = a_{22}a_{33} - a_{32}a_{23}.$$

Evaluating the remaining terms on the right-hand side of (19b) in a similar fashion, we arrive at

$$\iota_2(\boldsymbol{A}) = (a_{22}a_{33}a_{33}a_{11}a_{11}a_{22}) - (a_{32}a_{23}a_{13}a_{31}a_{21}a_{12})$$

Then it's a matter of some straightforward (but tedious) algebra to show that the expression above is equivalent to (24b). □

A calculation similar to those in the previous two examples shows that

$$\iota_3(\boldsymbol{A}) = [A\boldsymbol{e}_1, A\boldsymbol{e}_2, A\boldsymbol{e}_3] = a_{i1}a_{j2}a_{k3}\epsilon_{ijk}.$$

Expanding the summation as it was done in the previous example, leads to a sum of six terms. A close inspection of the term reveals that the sum is exactly the determinant of the matrix $[a_{ij}]$. For that reason, the third principal invariant, $\iota_3(\boldsymbol{A})$, is called *the determinant of $\boldsymbol{A}$*. We write this as

$$\iota_3(\boldsymbol{A}) = \det \boldsymbol{A}. \tag{24c}$$

Remark 7. An immediate consequence of (16a) is that

$$\operatorname{tr}(\alpha\boldsymbol{A} + \beta\boldsymbol{B}) = \alpha \operatorname{tr}\boldsymbol{A} + \beta \operatorname{tr}\boldsymbol{B} \quad \text{for all } \boldsymbol{A}, \boldsymbol{B} \in \mathcal{L},\ \alpha, \beta \in \mathbb{R}.$$

and thus, in particular, the trace is a linear function from $\mathcal{L}$ to $\mathbb{R}$. Immediate consequences of (16c) are that for any $\boldsymbol{A}, \boldsymbol{B} \in \mathcal{L}$ and $\alpha \in \mathbb{R}$:

$$\det(\alpha\boldsymbol{A}) = \alpha^3 \det\boldsymbol{A}, \tag{25a}$$

$$\det(\boldsymbol{A}\boldsymbol{B}) = \det\boldsymbol{A}\ \det\boldsymbol{B}, \tag{25b}$$

$$\det\boldsymbol{I} = 1. \tag{25c}$$

7. The scalar product of second order tensors

In Section 4 we made the set $\mathcal{V}$ of the second order tensors into a linear space by defining addition and multiplication by scalars in (14). In this section we equip $\mathcal{V}$ with a scalar product, thus making it an *inner product space*. Specifically, we define the scalar product $\boldsymbol{A}:\boldsymbol{B}$ of the second order tensors $\boldsymbol{A}, \boldsymbol{B} \in \mathcal{L}$ through

$$\boldsymbol{A}:\boldsymbol{B} = \operatorname{tr}(\boldsymbol{A}^T\boldsymbol{B}). \tag{26}$$

If we expresses $\boldsymbol{A}$ and $\boldsymbol{B}$ in terms of their components relative to a frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ as in $\boldsymbol{A} = a_{ij}\,\boldsymbol{e}_i \otimes \boldsymbol{e}_j$ and $\boldsymbol{B} = b_{pq}\,\boldsymbol{e}_p \otimes \boldsymbol{e}_q$, we get

$$\begin{aligned}\boldsymbol{A}:\boldsymbol{B} &= \operatorname{tr}\Big((a_{ij}\,\boldsymbol{e}_i \otimes \boldsymbol{e}_j)^T(b_{pq}\,\boldsymbol{e}_p \otimes \boldsymbol{e}_q)\Big) = a_{ij}b_{pq}\operatorname{tr}\Big((\boldsymbol{e}_j \otimes \boldsymbol{e}_i)(\boldsymbol{e}_p \otimes \boldsymbol{e}_q)\Big)\\ &\overset{\text{by (21c)}}{=} a_{ij}b_{pq}\operatorname{tr}\Big((\boldsymbol{e}_i \cdot \boldsymbol{e}_p)(\boldsymbol{e}_j \otimes \boldsymbol{e}_q)\Big) = a_{ij}b_{pq}(\boldsymbol{e}_i \cdot \boldsymbol{e}_p)\operatorname{tr}(\boldsymbol{e}_j \otimes \boldsymbol{e}_q) \overset{\text{by (21g)}}{=} a_{ij}b_{pq}\delta_{ip}\delta jq,\end{aligned}$$

and therefore

$$\boldsymbol{A}:\boldsymbol{B} = a_{ij}b_{ij}. \tag{27}$$

We conclude that $\boldsymbol{A}:\boldsymbol{B} = \boldsymbol{B}:\boldsymbol{A}$, and that $\boldsymbol{A}:\boldsymbol{A} = a_{ij}a_{ij} \geq 0$. One may verify along the same lines that the definition in (26) satisfies all the requirements of the inner product. We leave it for an exercise to verify that the basis $\mathcal{K}$ of $\mathcal{L}$ is an orthonormal set in terms of this scalar product.

8. Eigenvalues and spectral representation

Proposition 4. *Given a tensor $\boldsymbol{A} \in \mathcal{L}$, there exists a nonzero vector $\boldsymbol{u} \in \mathcal{V}$ such that $\boldsymbol{A}\boldsymbol{u} = \boldsymbol{0}$ if and only if* $\det\boldsymbol{A} = 0$.

Proof. (a) Suppose that $\det\boldsymbol{A} = 0$. We wish to prove that there exists a nonzero vector $\boldsymbol{u}$ so that $\boldsymbol{A}\boldsymbol{u} = \boldsymbol{0}$. By (24c) we have $\iota_3(\boldsymbol{A}) = 0$. Then by (19c), we have $[\boldsymbol{A}\boldsymbol{e}_1, \boldsymbol{A}\boldsymbol{e}_2, \boldsymbol{A}\boldsymbol{e}_3] = 0$ for any frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$. This implies that the set $\{\boldsymbol{A}\boldsymbol{e}_1, \boldsymbol{A}\boldsymbol{e}_2, \boldsymbol{A}\boldsymbol{e}_3\}$ is *not* linearly independent, and therefore there exist numbers $\alpha_i, \alpha_2, \alpha_3$, not all zero, so that $\alpha_i\boldsymbol{A}\boldsymbol{e}_i = \boldsymbol{0}$, that is, $\boldsymbol{A}(\alpha_i\boldsymbol{e}_i) = \boldsymbol{0}$. Then $\boldsymbol{u} = \alpha_i\boldsymbol{e}_i$ has the property that it is nonzero and $\boldsymbol{A}\boldsymbol{u} = \boldsymbol{0}$.

(b) Suppose that the unit vector $\boldsymbol{u}$ is such that $\boldsymbol{A}\boldsymbol{u} = \boldsymbol{0}$. We wish to show that $\det \boldsymbol{A} = 0$. Pick unit vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ so that $\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\}$ is orthonormal. Then by (24c) and (19c) we have $\det \boldsymbol{A} = \iota_3(\boldsymbol{A}) = [\boldsymbol{A}\boldsymbol{u}, \boldsymbol{A}\boldsymbol{v}, \boldsymbol{A}\boldsymbol{w}] = 0$ since $\boldsymbol{A}\boldsymbol{u} = \boldsymbol{0}$. □

A number $\lambda \in \mathbb{R}$ is said to be an *eigenvalue* of the tensor $\boldsymbol{A}$ if there exits a nonzero vector $\boldsymbol{v}$ so that $\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$, or equivalently, $(\lambda\boldsymbol{I} - \boldsymbol{A})\boldsymbol{v} = \boldsymbol{0}$, where $\boldsymbol{I}$ is the identity tensor. The vector $\boldsymbol{v}$ is called the *eigenvector* associated with the eigenvalue λ.

It follows from Proposition 4 that λ is an eigenvalue of $\boldsymbol{A}$ if and only if

$$\det(\lambda\boldsymbol{I} - \boldsymbol{A}) = 0. \tag{28}$$

This is called the *characteristic equation* of the tensor $\boldsymbol{A}$. The left-hand side of (28), which according to the following theorem is a cubic polynomial in λ, is called the *characteristic polynomial* of the tensor $\boldsymbol{A}$.

Theorem 3. *The characteristic polynomial* $\chi(\lambda) \equiv \det(\lambda\boldsymbol{I} - \boldsymbol{A})$ *of the tensor* $\boldsymbol{A}$ *is a cubic in* λ*, and its coefficients are the tensor's principal invariants:*

$$\chi(\lambda) = \lambda^3 - \iota_1(\boldsymbol{A})\,\lambda^2 + \iota_2(\boldsymbol{A})\,\lambda + \iota_3(\boldsymbol{A}). \tag{29}$$

Proof. In view of by (24c) and (19c) the characteristic polynomial may be expressed as

$$\chi(\lambda) = \det(\lambda\boldsymbol{I} - \boldsymbol{A}) = [(\lambda\boldsymbol{I} - \boldsymbol{A})\boldsymbol{e}_1, (\lambda\boldsymbol{I} - \boldsymbol{A})\boldsymbol{e}_2, (\lambda\boldsymbol{I} - \boldsymbol{A})\boldsymbol{e}_3]$$

for an arbitrary frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$. Let us expand that scalar triple product:

$$\begin{aligned}
&[\lambda\boldsymbol{e}_1 - \boldsymbol{A}\boldsymbol{e}_1, \lambda\boldsymbol{e}_2 - \boldsymbol{A}\boldsymbol{e}_2, \lambda\boldsymbol{e}_3 - \boldsymbol{A}\boldsymbol{e}_3] \\
&\qquad = [\lambda\boldsymbol{e}_1, \lambda\boldsymbol{e}_2 - \boldsymbol{A}\boldsymbol{e}_2, \lambda\boldsymbol{e}_3 - \boldsymbol{A}\boldsymbol{e}_3] - [\boldsymbol{A}\boldsymbol{e}_1, \lambda\boldsymbol{e}_2 - \boldsymbol{A}\boldsymbol{e}_2, \lambda\boldsymbol{e}_3 - \boldsymbol{A}\boldsymbol{e}_3].
\end{aligned}$$

The first term on the right-hand side expands to

$$\begin{aligned}
[\lambda\boldsymbol{e}_1, \lambda\boldsymbol{e}_2 - \boldsymbol{A}\boldsymbol{e}_2, \lambda\boldsymbol{e}_3 - \boldsymbol{A}\boldsymbol{e}_3] &= [\lambda\boldsymbol{e}_1, \lambda\boldsymbol{e}_2, \lambda\boldsymbol{e}_3 - \boldsymbol{A}\boldsymbol{e}_3] - [\lambda\boldsymbol{e}_1, \boldsymbol{A}\boldsymbol{e}_2, \lambda\boldsymbol{e}_3 - \boldsymbol{A}\boldsymbol{e}_3] \\
&= [\lambda\boldsymbol{e}_1, \lambda\boldsymbol{e}_2, \lambda\boldsymbol{e}_3] - [\lambda\boldsymbol{e}_1, \lambda\boldsymbol{e}_2, \boldsymbol{A}\boldsymbol{e}_3] - [\lambda\boldsymbol{e}_1, \boldsymbol{A}\boldsymbol{e}_2, \lambda\boldsymbol{e}_3] + [\lambda\boldsymbol{e}_1, \boldsymbol{A}\boldsymbol{e}_2, \boldsymbol{A}\boldsymbol{e}_3],
\end{aligned}$$

while the second term on the right-hand side expands to

$$\begin{aligned}
[\boldsymbol{A}\boldsymbol{e}_1, \lambda\boldsymbol{e}_2 - \boldsymbol{A}\boldsymbol{e}_2, \lambda\boldsymbol{e}_3 - \boldsymbol{A}\boldsymbol{e}_3] &= [\boldsymbol{A}\boldsymbol{e}_1, \lambda\boldsymbol{e}_2, \lambda\boldsymbol{e}_3 - \boldsymbol{A}\boldsymbol{e}_3] - [\boldsymbol{A}\boldsymbol{e}_1, \boldsymbol{A}\boldsymbol{e}_2, \lambda\boldsymbol{e}_3 - \boldsymbol{A}\boldsymbol{e}_3] \\
&= [\boldsymbol{A}\boldsymbol{e}_1, \lambda\boldsymbol{e}_2, \lambda\boldsymbol{e}_3] - [\boldsymbol{A}\boldsymbol{e}_1, \lambda\boldsymbol{e}_2, \boldsymbol{A}\boldsymbol{e}_3] - [\boldsymbol{A}\boldsymbol{e}_1, \boldsymbol{A}\boldsymbol{e}_2, \lambda\boldsymbol{e}_3] + [\boldsymbol{A}\boldsymbol{e}_1, \boldsymbol{A}\boldsymbol{e}_2, \boldsymbol{A}\boldsymbol{e}_3].
\end{aligned}$$

Putting it all together we arrive at

$$\begin{aligned}
\chi(\lambda) = {}& [\lambda\boldsymbol{e}_1, \lambda\boldsymbol{e}_2, \lambda\boldsymbol{e}_3] - [\lambda\boldsymbol{e}_1, \lambda\boldsymbol{e}_2, \boldsymbol{A}\boldsymbol{e}_3] - [\lambda\boldsymbol{e}_1, \boldsymbol{A}\boldsymbol{e}_2, \lambda\boldsymbol{e}_3] + [\lambda\boldsymbol{e}_1, \boldsymbol{A}\boldsymbol{e}_2, \boldsymbol{A}\boldsymbol{e}_3] \\
&- [\boldsymbol{A}\boldsymbol{e}_1, \lambda\boldsymbol{e}_2, \lambda\boldsymbol{e}_3] + [\boldsymbol{A}\boldsymbol{e}_1, \lambda\boldsymbol{e}_2, \boldsymbol{A}\boldsymbol{e}_3] + [\boldsymbol{A}\boldsymbol{e}_1, \boldsymbol{A}\boldsymbol{e}_2, \lambda\boldsymbol{e}_3] - [\boldsymbol{A}\boldsymbol{e}_1, \boldsymbol{A}\boldsymbol{e}_2, \boldsymbol{A}\boldsymbol{e}_3],
\end{aligned}$$

which we rearrange into

$$\begin{aligned}
\chi(\lambda) = {}& [\lambda\boldsymbol{e}_1, \lambda\boldsymbol{e}_2, \lambda\boldsymbol{e}_3] - [\lambda\boldsymbol{e}_1, \lambda\boldsymbol{e}_2, \boldsymbol{A}\boldsymbol{e}_3] - [\lambda\boldsymbol{e}_1, \boldsymbol{A}\boldsymbol{e}_2, \lambda\boldsymbol{e}_3] - [\boldsymbol{A}\boldsymbol{e}_1, \lambda\boldsymbol{e}_2, \lambda\boldsymbol{e}_3] \\
&+ [\lambda\boldsymbol{e}_1, \boldsymbol{A}\boldsymbol{e}_2, \boldsymbol{A}\boldsymbol{e}_3] + [\boldsymbol{A}\boldsymbol{e}_1, \lambda\boldsymbol{e}_2, \boldsymbol{A}\boldsymbol{e}_3] + [\boldsymbol{A}\boldsymbol{e}_1, \boldsymbol{A}\boldsymbol{e}_2, \lambda\boldsymbol{e}_3] - [\boldsymbol{A}\boldsymbol{e}_1, \boldsymbol{A}\boldsymbol{e}_2, \boldsymbol{A}\boldsymbol{e}_3]
\end{aligned}$$

and further simplify

$$\begin{aligned}
\chi(\lambda) = {}& \lambda^3 - \lambda^2\big([\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{A}\boldsymbol{e}_3] + [\boldsymbol{e}_1, \boldsymbol{A}\boldsymbol{e}_2, \boldsymbol{e}_3] + [\boldsymbol{A}\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3]\big) \\
&+ \lambda\big([\boldsymbol{e}_1, \boldsymbol{A}\boldsymbol{e}_2, \boldsymbol{A}\boldsymbol{e}_3] + [\boldsymbol{A}\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{A}\boldsymbol{e}_3] + [\boldsymbol{A}\boldsymbol{e}_1, \boldsymbol{A}\boldsymbol{e}_2, \boldsymbol{e}_3]\big) - [\boldsymbol{A}\boldsymbol{e}_1, \boldsymbol{A}\boldsymbol{e}_2, \boldsymbol{A}\boldsymbol{e}_3].
\end{aligned}$$

In view of equations (19), we recognize the coefficients of λ as $\boldsymbol{A}$'s principal invariants, and hence arrive at (29). □

Remark 8. Since the principal invariants $\iota_1(\boldsymbol{A})$, $\iota_2(\boldsymbol{A})$, $\iota_3(\boldsymbol{A})$ are real, the characteristic equation $\chi(\lambda) = 0$ has either one real root or three real roots. We conclude that an arbitrary second order tensor $\boldsymbol{A}$ has either one eigenvalue or three eigenvalues.[4]

Proposition 5. *Let $\boldsymbol{v}$ be an eigenvector of the tensor $\boldsymbol{A}$, and let λ be the corresponding eigenvector. Then, for any polynomial f, the vector $\boldsymbol{v}$ is an eigenvector of $f(\boldsymbol{A})$, and $f(\lambda)$ is the associated eigenvalue.*

Proof. From $\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$ we get

$$\boldsymbol{A}^2\boldsymbol{v} = \boldsymbol{A}(\boldsymbol{A}\boldsymbol{v}) = \boldsymbol{A}(\lambda\boldsymbol{v}) = \lambda\boldsymbol{A}\boldsymbol{v} = \lambda^2\boldsymbol{v}.$$

Then, by inductions, $\boldsymbol{A}^k\boldsymbol{v} = \lambda^k\boldsymbol{v}$, for all $k = 1, 2, \ldots$. But an nth degree polynomial is of the form $f(x) = \sum_{k=0}^{n} a_k x^k$. Therefore

$$f(\boldsymbol{A})\boldsymbol{v} = \sum_{k=0}^{n} a_k \boldsymbol{A}^k \boldsymbol{v} = \sum_{k=0}^{n} a_k \lambda^k \boldsymbol{v} = f(\lambda)\boldsymbol{v},$$

as asserted. □

Theorem 4 (Cayley-Hamilton)**.** *Any tensor $\boldsymbol{A}$ satisfies its own characteristic equation in the sense that*

$$\chi(\boldsymbol{A}) \equiv \boldsymbol{A}^3 - \iota_1(\boldsymbol{A})\,\boldsymbol{A}^2 + \iota_2(\boldsymbol{A})\,\boldsymbol{A} + \iota_3(\boldsymbol{A})\boldsymbol{I} = \boldsymbol{0}. \tag{30}$$

Proof. The proof in the general case is rather tedious, so here we present a proof in the special case where $\boldsymbol{A}$ has three eigenvalues, let's say λ_1, λ_2, λ_3, and that the corresponding eigenvectors $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\boldsymbol{v}_3$ form a basis for the vector space $\mathcal{V}$.

Since eigenvalues are roots of the characteristic equation, we have $\chi(\lambda) = (\lambda-\lambda_1)(\lambda-\lambda_2)(\lambda-\lambda_3)$, and therefore (30) is equivalent to

$$\chi(\boldsymbol{A}) = (\boldsymbol{A} - \lambda_1\boldsymbol{I})(\boldsymbol{A} - \lambda_2\boldsymbol{I})(\boldsymbol{A} - \lambda_3\boldsymbol{I}) = \boldsymbol{0}.$$

We observe that the order of the three parenthesized factors is immaterial since for any i and j we have

$$(\boldsymbol{A} - \lambda_i\boldsymbol{I})(\boldsymbol{A} - \lambda_j\boldsymbol{I}) = \boldsymbol{A}^2 - (\lambda_i + \lambda_j)\boldsymbol{A} + \lambda_i\lambda_j\boldsymbol{I},$$

which is unchanged under the interchange of i and j. Then we see that for any eigenvector $\boldsymbol{v}_i$ we have

$$\chi(\boldsymbol{A})\,\boldsymbol{v}_i = (\boldsymbol{A} - \lambda_1\boldsymbol{I})(\boldsymbol{A} - \lambda_2\boldsymbol{I})(\boldsymbol{A} - \lambda_3\boldsymbol{I})\,\boldsymbol{v}_i = \boldsymbol{0}$$

since we may rearrange the parenthesized factors to put $(\boldsymbol{A} - \lambda_i\boldsymbol{I})$ as the last term, and then $(\boldsymbol{A} - \lambda_i\boldsymbol{I})\,\boldsymbol{v}_i = \boldsymbol{0}$ since $\boldsymbol{v}_i$ is an eigenvector.

Now, an arbitrary $\boldsymbol{u} \in \mathcal{V}$ may be expressed as a linear combination $\boldsymbol{u} = \alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2 + \alpha_3\boldsymbol{v}_3$ due to the assumption that the eigenvectors form a basis. Then

$$\chi(\boldsymbol{A})\boldsymbol{u} = \chi(\boldsymbol{A})(\alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2 + \alpha_3\boldsymbol{v}_3) = \alpha_1\chi(\boldsymbol{A})\,\boldsymbol{v}_1 + \alpha_2\chi(\boldsymbol{A})\,\boldsymbol{v}_2 + \alpha_3\chi(\boldsymbol{A})\,\boldsymbol{v}_3 = \boldsymbol{0}.$$

We conclude that $\chi(\boldsymbol{A}) = \boldsymbol{0}$ since $\boldsymbol{u}$ is arbitrary. □

---

[4] Tacit in this statement is that we account for multiple roots. Thus, the polynomial equation $(\lambda-1)^2(\lambda-2) = 0$ has three roots, $\lambda = 1, 1, 2$.

9. Symmetric tensors

It's possible to show, but we skip the proof here, that a symmetric tensor, i.e., one with the property $\boldsymbol{A} = \boldsymbol{A}^T$, always has three eigenvalues, and the corresponding eigenvectors may be selected to be an orthonormal set.

For a symmetric tensor $\boldsymbol{A}$ we may express the principal invariants $\iota_1(\boldsymbol{A})$, $\iota_2(\boldsymbol{A})$, $\iota_3(\boldsymbol{A})$ in terms of $\boldsymbol{A}$'s eigenvalues λ_1, λ_2, λ_3 as follows.

Theorem 5.

$$\iota_1(\boldsymbol{A}) = \lambda_1 + \lambda_2 + \lambda_3, \tag{31a}$$

$$\iota_2(\boldsymbol{A}) = \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1, \tag{31b}$$

$$\iota_3(\boldsymbol{A}) = \lambda_1\lambda_2\lambda_3. \tag{31c}$$

Proof. The eigenvalues $\lambda_1, \lambda_2, \lambda_3$ are the roots of the characteristic equation (29). Therefore

$$\begin{aligned}\lambda^3 - \iota_1(\boldsymbol{A})\,\lambda^2 + \iota_2(\boldsymbol{A})\,\lambda - \iota_3(\boldsymbol{A}) &= (\lambda - \lambda_1)(\lambda - \lambda_2)(\lambda - \lambda_3) \\ &= \lambda^3 - (\lambda_1 + \lambda_2 + \lambda_3)\lambda^2 + (\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1)\lambda - \lambda_1\lambda_2\lambda_3.\end{aligned}$$

The theorem's assertion follows by comparing the left- and right-hand sides. □

Theorem 6 (Spectral representation)**.** *Let* $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ *be an orthonormal set of eigenvectors corresponding to the eigenvalues* $\lambda_1, \lambda_2, \lambda_3$ *of the symmetric tensor* $\boldsymbol{A}$*. Then*

$$\boldsymbol{A} = \sum_{i=1}^{3} \lambda_i \boldsymbol{e}_i \otimes \boldsymbol{e}_i. \tag{32}$$

Proof. In Exercise 11 you will show that $\boldsymbol{I} = \boldsymbol{e}_i \otimes \boldsymbol{e}_i$. It follows that

$$\boldsymbol{A} = \boldsymbol{A}\boldsymbol{I} = \boldsymbol{A}(\boldsymbol{e}_i \otimes \boldsymbol{e}_i) \overset{\text{by (21d)}}{=} (\boldsymbol{A}\boldsymbol{e}_i) \otimes \boldsymbol{e}_i = \sum_{i=1}^{3} (\lambda_i \boldsymbol{e}_i) \otimes \boldsymbol{e}_i. \qquad \square$$

A symmetric tensor $\boldsymbol{A}$ is said to be *positive semi-definite* if $\boldsymbol{u} \cdot \boldsymbol{A}\boldsymbol{u} \geq 0$ for all vectors $\boldsymbol{u}$, and it is said to be *positive definite* if $\boldsymbol{u} \cdot \boldsymbol{A}\boldsymbol{u} > 0$ for all nonzero vectors $\boldsymbol{u}$.

Let $\boldsymbol{A}$ as in the statement of Theorem 6. If $\boldsymbol{A}$ is positive semi-definite, then for any $\boldsymbol{u} \in \mathcal{V}$ we have

$$\boldsymbol{u} \cdot \boldsymbol{A}\boldsymbol{u} = \boldsymbol{u} \cdot \Big(\sum_{i=1}^{3} \lambda_i \boldsymbol{e}_i \otimes \boldsymbol{e}_i \Big) \boldsymbol{u} = \boldsymbol{u} \cdot \Big(\sum_{i=1}^{3} \lambda_i (\boldsymbol{e}_i \cdot \boldsymbol{u}) \boldsymbol{e}_i \Big) = \sum_{i=1}^{3} \lambda_i (\boldsymbol{e}_i \cdot \boldsymbol{u})(\boldsymbol{e}_i \cdot \boldsymbol{u}) = \sum_{i=1}^{3} \lambda_i (\boldsymbol{e}_i \cdot \boldsymbol{u})^2 \geq 0.$$

Since $\boldsymbol{u}$ is arbitrary, it follows that $\lambda_i \geq 0$ for $i = 1, 2, 3$. In other words, if $\boldsymbol{A}$ is symmetric and positive semi-definite, then its eigenvalues are nonnegative.

Remark 9. Let $\boldsymbol{A}$ be a symmetric positive definite tensor with an orthonormal set $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ of eigenvectors, and the corresponding eigenvalues λ_1, λ_2, λ_3, as in Theorem 6. Pick a point $\boldsymbol{o} \in \mathbb{E}_3$ as the origin, and identify any point $\boldsymbol{x} \in \mathbb{E}_3$ by the position vector that extends from $\boldsymbol{o}$ to $\boldsymbol{x}$. Then for any such $\boldsymbol{x}$ we have

$$\boldsymbol{A}\boldsymbol{x} = \sum_{i=1}^{3} \lambda_i\,(\boldsymbol{e}_i \otimes \boldsymbol{e}_i)\,\boldsymbol{x} = \sum_{i=1}^{3} \lambda_i\,(\boldsymbol{e}_i \cdot \boldsymbol{x})\,\boldsymbol{e}_i = \sum_{i=1}^{3} \lambda_i x_i\,\boldsymbol{e}_i,$$

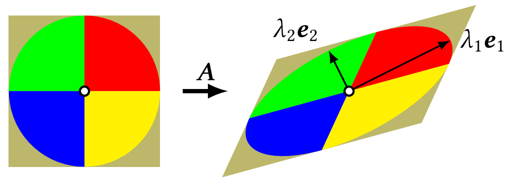


FIGURE 2. The application of the symmetric and positive definite tensor $\boldsymbol{A}$ to any neighborhood of the origin results in shrinking/stretching of the neighborhood in the eigenvector directions $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ by the factors λ_1 and λ_2.

where x_i is the component of $\boldsymbol{x}$ along $\boldsymbol{e}_i$. We see that the effect of applying $\boldsymbol{A}$ to the vector $\boldsymbol{x} = x_i\boldsymbol{e}_i$, amount to shrinking/stretching the components of $\boldsymbol{x}$ by factors of λ_i. In particular, $\boldsymbol{A}\boldsymbol{e}_i = \lambda_i\boldsymbol{e}_i$, that is, vectors that are aligned with the eigenvector directions simply get shrunk/stretched without a change in orientations. The overall effect is that under the action of $\boldsymbol{A}$, a neighborhood of the origin $\boldsymbol{o}$ deforms by shrinking/stretching by the factors λ_i in the directions $\boldsymbol{e}_i$. Figure 2 illustrates this effect in two dimensions.

Theorem 7. *A positive semi-definite tensor $\boldsymbol{A}$ has a unique positive semi-definite square root, $\boldsymbol{A}^{1/2}$. Specifically, if $\boldsymbol{A}$ has the spectral representation* (32)*, then*

$$\boldsymbol{A}^{1/2} = \sum_{i=1}^{3} \lambda_i^{1/2} \boldsymbol{e}_i \otimes \boldsymbol{e}_i. \tag{33}$$

Proof. Clearly $\boldsymbol{A}^{1/2}$ is positive semi-definite. Let us calculate:

$$\left(\boldsymbol{A}^{1/2}\right)^2 = \left(\sum_{i=1}^{3} \lambda_i^{1/2} \boldsymbol{e}_i \otimes \boldsymbol{e}_i\right)\left(\sum_{j=1}^{3} \lambda_j^{1/2} \boldsymbol{e}_j \otimes \boldsymbol{e}_j\right) = \sum_{i=1}^{3}\sum_{j=1}^{3} \lambda_i^{1/2}\lambda_j^{1/2} (\boldsymbol{e}_i \otimes \boldsymbol{e}_i)(\boldsymbol{e}_j \otimes \boldsymbol{e}_j)$$

$$\overset{\text{by (21c)}}{=} \sum_{i=1}^{3}\sum_{j=1}^{3} \lambda_i^{1/2}\lambda_j^{1/2} (\boldsymbol{e}_i \cdot \boldsymbol{e}_j)\, \boldsymbol{e}_i \otimes \boldsymbol{e}_j = \sum_{i=1}^{3}\sum_{j=1}^{3} \lambda_i^{1/2}\lambda_j^{1/2} \delta_{ij}\, \boldsymbol{e}_i \otimes \boldsymbol{e}_j = \sum_{i=1}^{3} \lambda_i\, \boldsymbol{e}_i \otimes \boldsymbol{e}_i = \boldsymbol{A},$$

which shows that $\boldsymbol{A}^{1/2}$ is a square root of $\boldsymbol{A}$. Indeed, it is the unique positive semi-definite root, for if B is another positive semi-definite tensor with the property $\boldsymbol{B}^2 = \boldsymbol{A}$, then according to Proposition 5, eigenvectors of $\boldsymbol{B}$ are eigenvectors of $\boldsymbol{A}$, and the squares of the eigenvalues of $\boldsymbol{B}$ are eigenvalues of $\boldsymbol{A}$. Therefore by the by the spectral representation theorem 6, $\boldsymbol{B}$ has the representation given in (33). □

10. Skew-symmetric tensors

Recall that a tensor $\boldsymbol{W}$ is said to be skew-symmetric if $\boldsymbol{W}^T = -\boldsymbol{W}$. Let us observe that for any two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ we have

$$\boldsymbol{u} \cdot \boldsymbol{W}\boldsymbol{v} = (\boldsymbol{W}^T\boldsymbol{u}) \cdot \boldsymbol{v} = -(\boldsymbol{W}\boldsymbol{u}) \cdot \boldsymbol{v} = -\boldsymbol{v} \cdot \boldsymbol{W}\boldsymbol{u}, \tag{34}$$

and in particular,

$$\boldsymbol{u} \cdot \boldsymbol{W}\boldsymbol{u} = 0. \tag{35}$$

A skew-symmetric tensor, like all tensors in $\mathcal{V}$, has at least one eigenvalue (see Remark 8). Let λ be that eigenvalue and the unit vector $\boldsymbol{e}_1$ be an associated eigenvector.

Thus, $\boldsymbol{W}\boldsymbol{e}_1 = \lambda\boldsymbol{e}_1$, whence $\boldsymbol{e}_1 \cdot \boldsymbol{W}\boldsymbol{e}_1 = \lambda\|\boldsymbol{e}_1\|^2$, which, in in view of (35), reduces to $\lambda\|\boldsymbol{e}_1\|^2 = 0$. Since $\|\boldsymbol{e}_1\| = 1$, we conclude that $\lambda = 0$. Thus, a skew-symmetric tensor has either one eigenvalue which is zero, or three eigenvalues, all being zeros. Let us note that the eigenvector $\boldsymbol{e}_1$ satisfies $\boldsymbol{W}\boldsymbol{e}_1 = \boldsymbol{0}$.

Let $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$ be vectors—not necessarily eigenvectors—so that $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ forms a right-handed frame $\mathcal{V}$. Then by Theorem 2 we have $\boldsymbol{W} = (\boldsymbol{e}_i \cdot \boldsymbol{W}\boldsymbol{e}_j)\,\boldsymbol{e}_i \otimes \boldsymbol{e}_j$ which expands into a sum of nine terms:

$$\begin{aligned}
\boldsymbol{W} &= (\boldsymbol{e}_1 \cdot \boldsymbol{W}\boldsymbol{e}_1)\,\boldsymbol{e}_1 \otimes \boldsymbol{e}_1 + (\boldsymbol{e}_2 \cdot \boldsymbol{W}\boldsymbol{e}_2)\,\boldsymbol{e}_2 \otimes \boldsymbol{e}_2 + (\boldsymbol{e}_3 \cdot \boldsymbol{W}\boldsymbol{e}_3)\,\boldsymbol{e}_3 \otimes \boldsymbol{e}_3 \\
&\quad + (\boldsymbol{e}_1 \cdot \boldsymbol{W}\boldsymbol{e}_2)\,\boldsymbol{e}_1 \otimes \boldsymbol{e}_2 + (\boldsymbol{e}_2 \cdot \boldsymbol{W}\boldsymbol{e}_1)\,\boldsymbol{e}_2 \otimes \boldsymbol{e}_1 \\
&\quad + (\boldsymbol{e}_2 \cdot \boldsymbol{W}\boldsymbol{e}_3)\,\boldsymbol{e}_2 \otimes \boldsymbol{e}_3 + (\boldsymbol{e}_3 \cdot \boldsymbol{W}\boldsymbol{e}_2)\,\boldsymbol{e}_3 \otimes \boldsymbol{e}_2 \\
&\quad + (\boldsymbol{e}_3 \cdot \boldsymbol{W}\boldsymbol{e}_1)\,\boldsymbol{e}_3 \otimes \boldsymbol{e}_1 + (\boldsymbol{e}_1 \cdot \boldsymbol{W}\boldsymbol{e}_3)\,\boldsymbol{e}_1 \otimes \boldsymbol{e}_3.
\end{aligned}$$

Each of the first three terms on the right-hand side is zero due to (35). In what remains, all terms that involve $\boldsymbol{W}\boldsymbol{e}_1$ are zero as noted in the previous paragraph. Furthermore, all terms of the form $\boldsymbol{e}_1 \cdot \boldsymbol{W}\boldsymbol{e}_k$ are zero since $\boldsymbol{e}_1 \cdot \boldsymbol{W}\boldsymbol{e}_k = (\boldsymbol{W}^T\boldsymbol{e}_1) \cdot \boldsymbol{e}_k = -(\boldsymbol{W}\boldsymbol{e}_1) \cdot \boldsymbol{e}_k = 0$ for the same reason. Thus, we are left with $\boldsymbol{W} = (\boldsymbol{e}_2 \cdot \boldsymbol{W}\boldsymbol{e}_3)\,\boldsymbol{e}_2 \otimes \boldsymbol{e}_3 + (\boldsymbol{e}_3 \cdot \boldsymbol{W}\boldsymbol{e}_2)\,\boldsymbol{e}_3 \otimes \boldsymbol{e}_2$. But $(\boldsymbol{e}_2 \cdot \boldsymbol{W}\boldsymbol{e}_3) = -(\boldsymbol{e}_3 \cdot \boldsymbol{W}\boldsymbol{e}_2)$ according to (34), and therefore

$$\boldsymbol{W} = \omega\,(\boldsymbol{e}_3 \otimes \boldsymbol{e}_2 - \boldsymbol{e}_2 \otimes \boldsymbol{e}_3), \quad \text{where } \omega = \boldsymbol{e}_3 \cdot \boldsymbol{W}\boldsymbol{e}_2. \tag{36}$$

Theorem 8. *Let $\boldsymbol{W}$ be a skew-symmetric tensor, $\boldsymbol{e}_1$ be an eigenvector of unit length, and ω defined as in* (36)*. Set $\boldsymbol{w} = \omega\,\boldsymbol{e}_1$. Then*

$$\boldsymbol{W}\boldsymbol{a} = \boldsymbol{w} \times \boldsymbol{a}, \quad \textit{for all } \boldsymbol{a} \in \mathcal{V}. \tag{37}$$

The vector $\boldsymbol{w}$ is called the axial vector *of the tensor $\boldsymbol{W}$.*

Proof. We have

$$\boldsymbol{W}\boldsymbol{a} = \omega\,(\boldsymbol{e}_3 \otimes \boldsymbol{e}_2 - \boldsymbol{e}_2 \otimes \boldsymbol{e}_3)\,\boldsymbol{a} = \omega\,\big((\boldsymbol{e}_2 \cdot \boldsymbol{a})\,\boldsymbol{e}_3 - (\boldsymbol{e}_3 \cdot \boldsymbol{a})\,\boldsymbol{e}_2\big).$$

Moreover, we have $\boldsymbol{a} = (\boldsymbol{a} \cdot \boldsymbol{e}_i)\,\boldsymbol{e}_i$, and therefore

$$\begin{aligned}
\boldsymbol{w} \times \boldsymbol{a} &= (\omega\,\boldsymbol{e}_1) \times (\boldsymbol{a} \cdot \boldsymbol{e}_i)\,\boldsymbol{e}_i = \omega\,(\boldsymbol{a} \cdot \boldsymbol{e}_i)\,\boldsymbol{e}_1 \times \boldsymbol{e}_i \\
&= \omega\,\big[(\boldsymbol{a} \cdot \boldsymbol{e}_2)\,\boldsymbol{e}_1 \times \boldsymbol{e}_2 + (\boldsymbol{a} \cdot \boldsymbol{e}_3)\,\boldsymbol{e}_1 \times \boldsymbol{e}_3\big] = \omega\,\big[(\boldsymbol{a} \cdot \boldsymbol{e}_2)\,\boldsymbol{e}_1 \times \boldsymbol{e}_2 - (\boldsymbol{a} \cdot \boldsymbol{e}_3)\,\boldsymbol{e}_3 \times \boldsymbol{e}_1\big].
\end{aligned}$$

Thus, we calculate

$$\boldsymbol{W}\boldsymbol{a} - \boldsymbol{w} \times \boldsymbol{a} = \omega\,\Big[(\boldsymbol{e}_2 \cdot \boldsymbol{a})(\boldsymbol{e}_3 - \boldsymbol{e}_1 \times \boldsymbol{e}_2) - (\boldsymbol{e}_3 \cdot \boldsymbol{a})(\boldsymbol{e}_2 - \boldsymbol{e}_3 \times \boldsymbol{e}_1)\Big] = \boldsymbol{0},$$

where we have appealed to $\boldsymbol{e}_3 = \boldsymbol{e}_1 \times \boldsymbol{e}_2$ and $\boldsymbol{e}_2 = \boldsymbol{e}_3 \times \boldsymbol{e}_1$ since $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ is a right-handed frame. □

Theorem 9. *Let $\boldsymbol{u}$ and $\boldsymbol{w}$ be arbitrary vectors. Then $\boldsymbol{W} = \boldsymbol{v} \otimes \boldsymbol{u} - \boldsymbol{u} \otimes \boldsymbol{v}$ is skew-symmetric, and $\boldsymbol{w} = \boldsymbol{u} \times \boldsymbol{v}$ is $\boldsymbol{W}$'s axial vector.*

Proof. By (21b) we have

$$\boldsymbol{W}^T = (\boldsymbol{v} \otimes \boldsymbol{u} - \boldsymbol{u} \otimes \boldsymbol{v})^T = (\boldsymbol{v} \otimes \boldsymbol{u})^T - (\boldsymbol{u} \otimes \boldsymbol{v})^T = \boldsymbol{u} \otimes \boldsymbol{v} - \boldsymbol{v} \otimes \boldsymbol{u} = -(\boldsymbol{v} \otimes \boldsymbol{u} - \boldsymbol{u} \otimes \boldsymbol{v}) = -\boldsymbol{W},$$

and therefore $\boldsymbol{W}$ is skew-symmetric. Now, let us calculate

$$\begin{aligned}\boldsymbol{W}(\boldsymbol{u}\times\boldsymbol{v}) &= (\boldsymbol{v}\otimes\boldsymbol{u}-\boldsymbol{u}\otimes\boldsymbol{v})(\boldsymbol{u}\times\boldsymbol{v})\\ &= \big(\boldsymbol{u}\cdot(\boldsymbol{u}\times\boldsymbol{v})\big)\,\boldsymbol{v} - \big(\boldsymbol{v}\cdot(\boldsymbol{u}\times\boldsymbol{v})\big)\,\boldsymbol{u} = [\boldsymbol{u},\boldsymbol{u},\boldsymbol{v}]\,\boldsymbol{v} - [\boldsymbol{v},\boldsymbol{u},\boldsymbol{v}]\,\boldsymbol{u}.\end{aligned}$$

The scalar triple products $[\boldsymbol{u},\boldsymbol{u},\boldsymbol{v}]$ and $[\boldsymbol{v},\boldsymbol{u},\boldsymbol{v}]$ are both zero, and therefore $\boldsymbol{W}(\boldsymbol{u}\times\boldsymbol{v}) = \boldsymbol{0}$. We conclude that $\boldsymbol{u}\times\boldsymbol{v}$ points along $\boldsymbol{W}$'s axial vector, and therefore $\boldsymbol{w} = c\,\boldsymbol{u}\times\boldsymbol{v}$ for some number c. We proceed to show that $c = 1$. Toward that end, let us observe that for any $\boldsymbol{a}\in\mathcal{V}$ we have $\boldsymbol{W}\boldsymbol{a} = \boldsymbol{w}\times\boldsymbol{a}$, that is

$$(\boldsymbol{v}\otimes\boldsymbol{u}-\boldsymbol{u}\otimes\boldsymbol{v})\,\boldsymbol{a} = (c\,\boldsymbol{u}\times\boldsymbol{v})\times\boldsymbol{a},$$

which expands to

$$(\boldsymbol{u}\cdot\boldsymbol{a})\,\boldsymbol{v} - (\boldsymbol{v}\cdot\boldsymbol{a})\,\boldsymbol{u} = c\,(\boldsymbol{u}\times\boldsymbol{v})\times\boldsymbol{a}. \tag{38}$$

In particular, setting $\boldsymbol{a}=\boldsymbol{u}$ this results in

$$(\boldsymbol{u}\cdot\boldsymbol{u})\,\boldsymbol{v} - (\boldsymbol{v}\cdot\boldsymbol{u})\,\boldsymbol{u} = c\,(\boldsymbol{u}\times\boldsymbol{v})\times\boldsymbol{u}.$$

Form the dot product of both sides with $\boldsymbol{v}$:

$$(\boldsymbol{u}\cdot\boldsymbol{u})(\boldsymbol{v}\cdot\boldsymbol{v}) - (\boldsymbol{u}\cdot\boldsymbol{v})^2 = c\,\big((\boldsymbol{u}\times\boldsymbol{v})\times\boldsymbol{u}\big)\cdot\boldsymbol{v}.$$

The coefficient of c is the scalar triple product $\big[\boldsymbol{u}\times\boldsymbol{v},\boldsymbol{u},\boldsymbol{v}\big]$, whose terms may be rotated into $\big[\boldsymbol{u},\boldsymbol{v},\boldsymbol{u}\times\boldsymbol{v}\big]$, which then evaluates to $(\boldsymbol{u}\times\boldsymbol{v})\cdot(\boldsymbol{u}\times\boldsymbol{v})$, that is, $\|\boldsymbol{u}\times\boldsymbol{v}\|^2$. Thus we have arrived at

$$\|\boldsymbol{u}\|^2\|\boldsymbol{v}\|^2 - (\boldsymbol{u}\cdot\boldsymbol{v})^2 = c\,\|\boldsymbol{u}\times\boldsymbol{v}\|^2. \tag{39}$$

Comparing this with (1) we conclude that $c = 1$. □

Remark 10. As we have shown that $c = 1$ is in the above proof, equation (38) implies the very useful identity

$$(\boldsymbol{u}\times\boldsymbol{v})\times\boldsymbol{a} = (\boldsymbol{u}\cdot\boldsymbol{a})\,\boldsymbol{v} - (\boldsymbol{v}\cdot\boldsymbol{a})\,\boldsymbol{u}, \tag{40a}$$

or equivalently

$$\boldsymbol{a}\times(\boldsymbol{u}\times\boldsymbol{v}) = (\boldsymbol{a}\cdot\boldsymbol{v})\,\boldsymbol{u} - (\boldsymbol{a}\cdot\boldsymbol{u})\,\boldsymbol{v}, \tag{40b}$$

which hold for all vectors $\boldsymbol{u},\boldsymbol{v},\boldsymbol{a}\in\mathcal{V}$. Here is a good way of remembering these identities. On the left-hand side of (40a), regard $\boldsymbol{v}$ as the "near neighbor" and $\boldsymbol{u}$ as the "far neighbor" of $\boldsymbol{a}$. Then the right-hand side of (40a) reads: "the (dot product of $\boldsymbol{a}$ with its far neighbor) times the near neighbor, minus the (dot product of $\boldsymbol{a}$ with its near neighbor) times the far neighbor". The quoted mnemonic applies word for word to (40b).

See Exercise 17 for alternative ways of expressing the identities (40a) and (40b).

Remark 11. The article [3] has several alternative derivations of the identities (40).

11. Orthogonal tensors

A tensor $\boldsymbol{Q}\in\mathcal{L}$ is said to be *orthogonal* if it preserves the dot product in the sense that

$$\boldsymbol{Q}\,\boldsymbol{u}\cdot\boldsymbol{Q}\,\boldsymbol{v} = \boldsymbol{u}\cdot\boldsymbol{v}, \quad \text{for all } \boldsymbol{u},\boldsymbol{v}\in\mathcal{V}. \tag{41}$$

Since $\boldsymbol{Q}\,\boldsymbol{u}\cdot\boldsymbol{Q}\,\boldsymbol{v} = \boldsymbol{u}\cdot\boldsymbol{Q}^T\boldsymbol{Q}\,\boldsymbol{v}$, the definition (41) may be written as

$$\boldsymbol{u}\cdot(\boldsymbol{Q}^T\boldsymbol{Q}-\boldsymbol{I})\boldsymbol{v} = \boldsymbol{0}, \quad \text{for all } \boldsymbol{u},\boldsymbol{v}\in\mathcal{V},$$

from which it follows that $\boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I}$. Conversely, if $\boldsymbol{Q} \in \mathcal{L}$ is such that $\boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I}$, then clearly (41) holds. We conclude that the condition

$$\boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I}. \tag{42}$$

is an equivalent definition of the orthogonality of a tensor $\boldsymbol{Q} \in \mathcal{L}$.

Let us observe that (41) implies that

$$\|\boldsymbol{Q}\boldsymbol{u}\| = \|\boldsymbol{u}\| \quad \text{for all } \boldsymbol{u} \in \mathcal{V}, \tag{43}$$

that is, an orthogonal tensor preserves vector lengths.

If $\boldsymbol{Q}$ is orthogonal, equation (25b) implies that $\det(\boldsymbol{Q}^T\boldsymbol{Q}) = (\det \boldsymbol{Q})^2 = 1$, that is, $\det \boldsymbol{Q} = \pm 1$, and therefore $\boldsymbol{Q}$ is invertible. It follows that

$$\boldsymbol{Q}^{-1} = \boldsymbol{Q}^T, \quad \boldsymbol{Q}\boldsymbol{Q}^T = \boldsymbol{I}. \tag{44}$$

The orthogonal tensor $\boldsymbol{Q}$ is said to be *proper orthogonal* is $\det \boldsymbol{Q} = 1$, and *improper orthogonal* if $\det \boldsymbol{Q} = -1$. If $\boldsymbol{Q}$ is improper orthogonal, then $-\boldsymbol{Q}$ is proper orthogonal. From now on we assume that our orthogonal tensors are proper orthogonal unless explicitly stated otherwise.

In view of (42), we have

$$\boldsymbol{Q}^T(\boldsymbol{Q} - \boldsymbol{I}) = \boldsymbol{Q}^T\boldsymbol{Q} - \boldsymbol{Q}^T = \boldsymbol{I} - \boldsymbol{Q}^T = (\boldsymbol{I} - \boldsymbol{Q})^T = -(\boldsymbol{Q} - \boldsymbol{I})^T.$$

Then $\det\big(\boldsymbol{Q}^T(\boldsymbol{Q} - \boldsymbol{I})\big) = -\det(\boldsymbol{Q} - \boldsymbol{I})$. But according to (25b), and since since $\det \boldsymbol{Q} = 1$, we have

$$\det\big(\boldsymbol{Q}^T(\boldsymbol{Q} - \boldsymbol{I})\big) = \det(\boldsymbol{Q}^T)\det(\boldsymbol{Q} - \boldsymbol{I}) = \det(\boldsymbol{Q} - \boldsymbol{I})$$

It follows that $\det(\boldsymbol{Q} - \boldsymbol{I}) = -\det(\boldsymbol{Q} - \boldsymbol{I})$, and therefore $\det(\boldsymbol{Q} - \boldsymbol{I}) = 0$. We conclude that $\boldsymbol{Q}$ has an eigenvalue equal to 1. Therefore, there exists and eigenvector, say $\boldsymbol{e}_1$, so that $\boldsymbol{Q}\boldsymbol{e}_1 = \boldsymbol{e}_1$. Let us point out in passing that by applying $\boldsymbol{Q}^T$ to this equation we get $\boldsymbol{e}_1 = \boldsymbol{Q}^T\boldsymbol{e}_1$, and therefore $\boldsymbol{e}_1$ is also and eigenvector of $\boldsymbol{Q}^T$ corresponding to the eigenvalue 1.

Theorem 10. *Let the unit vector $\boldsymbol{e}_1$ be an eigenvector of the proper orthogonal tensor $\boldsymbol{Q}$, and let $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$ be unit vectors so that $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ forms a right-handed orthonormal frame in $\mathcal{V}$. Then, there exists a $\theta \in [-\pi, \pi]$ so that*

$$\boldsymbol{Q} = \boldsymbol{e}_1 \otimes \boldsymbol{e}_1 + (\boldsymbol{e}_2 \otimes \boldsymbol{e}_2 + \boldsymbol{e}_3 \otimes \boldsymbol{e}_3)\cos\theta - (\boldsymbol{e}_2 \otimes \boldsymbol{e}_3 - \boldsymbol{e}_3 \otimes \boldsymbol{e}_2)\sin\theta, \tag{45a}$$

or equivalently, (due to Exercise 11):

$$\boldsymbol{Q} = \boldsymbol{e}_1 \otimes \boldsymbol{e}_1 + (\boldsymbol{I} - \boldsymbol{e}_1 \otimes \boldsymbol{e}_1)\cos\theta - (\boldsymbol{e}_2 \otimes \boldsymbol{e}_3 - \boldsymbol{e}_3 \otimes \boldsymbol{e}_2)\sin\theta. \tag{45b}$$

Proof. The representation of $\boldsymbol{Q}$ in (45a) is a particular application of the general tensor representation (22) in terms of the nine basis elements in (23). Deriving (45a) amounts to calculating the nine components $\boldsymbol{e}_i \cdot \boldsymbol{Q}\boldsymbol{e}_j$. The calculation of five of the components is straightforward, but that of the remaining four, whose values have been left blank in the following display, requires some effort:

$$\boldsymbol{e}_1 \cdot \boldsymbol{Q}\boldsymbol{e}_1 = \boldsymbol{e}_1 \cdot \boldsymbol{e}_1 = 1, \tag{46a}$$
$$\boldsymbol{e}_1 \cdot \boldsymbol{Q}\boldsymbol{e}_2 = (\boldsymbol{Q}^T\boldsymbol{e}_1) \cdot \boldsymbol{e}_2 = \boldsymbol{e}_1 \cdot \boldsymbol{e}_2 = 0, \tag{46b}$$
$$\boldsymbol{e}_1 \cdot \boldsymbol{Q}\boldsymbol{e}_3 = (\boldsymbol{Q}^T\boldsymbol{e}_1) \cdot \boldsymbol{e}_3 = \boldsymbol{e}_1 \cdot \boldsymbol{e}_3 = 0, \tag{46c}$$
$$\boldsymbol{e}_2 \cdot \boldsymbol{Q}\boldsymbol{e}_1 = \boldsymbol{e}_2 \cdot \boldsymbol{e}_1 = 0, \tag{46d}$$
$$\boldsymbol{e}_2 \cdot \boldsymbol{Q}\boldsymbol{e}_2 = \tag{46e}$$
$$\boldsymbol{e}_2 \cdot \boldsymbol{Q}\boldsymbol{e}_3 = \tag{46f}$$
$$\boldsymbol{e}_3 \cdot \boldsymbol{Q}\boldsymbol{e}_1 = \boldsymbol{e}_3 \cdot \boldsymbol{e}_1 = 0, \tag{46g}$$
$$\boldsymbol{e}_3 \cdot \boldsymbol{Q}\boldsymbol{e}_2 = \tag{46h}$$
$$\boldsymbol{e}_3 \cdot \boldsymbol{Q}\boldsymbol{e}_3 = \tag{46i}$$

To calculate the missing values, let us observe that according to (46b), $\boldsymbol{Q}\boldsymbol{e}_2$ is perpendicular to $\boldsymbol{e}_1$, and therefore co-planar with the vectors $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$. It follows that $\boldsymbol{Q}\boldsymbol{e}_2$ is a linear combination of $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$, that is, $\boldsymbol{Q}\boldsymbol{e}_2 = \alpha_2\boldsymbol{e}_2 + \alpha_3\boldsymbol{e}_3$ for some scalars α_2 and α_3. Since $\boldsymbol{e}_2$ is a unit vector, by (43) we have $\|\boldsymbol{Q}\boldsymbol{e}_2\| = 1$. But $\|\boldsymbol{Q}\boldsymbol{e}_2\|^2 = (\alpha_2\boldsymbol{e}_2 + \alpha_3\boldsymbol{e}_3) \cdot (\alpha_2\boldsymbol{e}_2 + \alpha_3\boldsymbol{e}_3) = \alpha_2^2 + \alpha_3^2$, and therefore $\alpha_2^2 + \alpha_3^2 = 1$.

Similarly, according to (46c), $\boldsymbol{Q}\boldsymbol{e}_3$ is perpendicular to $\boldsymbol{e}_1$, and therefore co-planar with the vectors $\boldsymbol{e}_2$ and $\boldsymbol{e}_3$. It follows that $\boldsymbol{Q}\boldsymbol{e}_3 = \beta_2\boldsymbol{e}_2 + \beta_3\boldsymbol{e}_3$ for some scalars β_2 and β_3, and $\beta_2^2 + \beta_3^2 = 1$.

Additionally, since $\boldsymbol{e}_2 \cdot \boldsymbol{e}_3 = 0$, from (41) we get $\boldsymbol{Q}\boldsymbol{e}_2 \cdot \boldsymbol{Q}\boldsymbol{e}_3 = 0$. It follows that

$$\boldsymbol{Q}\boldsymbol{e}_2 \cdot \boldsymbol{Q}\boldsymbol{e}_3 = (\alpha_2\boldsymbol{e}_2 + \alpha_3\boldsymbol{e}_3) \cdot (\beta_2\boldsymbol{e}_2 + \beta_3\boldsymbol{e}_3) = \alpha_2\beta_2 + \alpha_3\beta_3 = 0.$$

Finally, referring to (19c) and (24c), we have:

$$\begin{aligned}\det \boldsymbol{Q} = \iota_3(\boldsymbol{Q}) &= [\boldsymbol{Q}\boldsymbol{e}_1, \boldsymbol{Q}\boldsymbol{e}_2, \boldsymbol{Q}\boldsymbol{e}_3] = [\boldsymbol{e}_1, \alpha_2\boldsymbol{e}_2 + \alpha_3\boldsymbol{e}_3, \beta_2\boldsymbol{e}_2 + \beta_3\boldsymbol{e}_3] \\ &= \big(\boldsymbol{e}_1 \times (\alpha_2\boldsymbol{e}_2 + \alpha_3\boldsymbol{e}_3)\big) \cdot (\beta_2\boldsymbol{e}_2 + \beta_3\boldsymbol{e}_3) = (\alpha_2\boldsymbol{e}_3 - \alpha_3\boldsymbol{e}_2) \cdot (\beta_2\boldsymbol{e}_2 + \beta_3\boldsymbol{e}_3) \\ &= \alpha_2\beta_3 - \alpha_3\beta_2 = 1,\end{aligned}$$

where the final step is due to $\det \boldsymbol{Q} = 1$.

To summarize, we have the following four relationships among the four coefficients α_2, α_3, β_2, and β_3:

$$\alpha_2^2 + \alpha_3^2 = 1, \quad \beta_2^2 + \beta_3^2 = 1, \quad \alpha_2\beta_2 + \alpha_3\beta_3 = 0, \quad \alpha_2\beta_3 - \alpha_3\beta_2 = 1.$$

A close inspection leads to the following solution to the system:

$$\alpha_2 = \cos\theta, \quad \alpha_3 = \sin\theta, \quad \beta_2 = -\sin\theta, \quad \beta_3 = \cos\theta,$$

for some $-\pi < \theta \le \pi$. This enables us to fill in the missing parts in (46). We have:

$$\boldsymbol{e}_2 \cdot \boldsymbol{Q}\boldsymbol{e}_2 = \boldsymbol{e}_2 \cdot (\alpha_2\boldsymbol{e}_2 + \alpha_3\boldsymbol{e}_3) = \alpha_2 = \cos\theta, \tag{47a}$$
$$\boldsymbol{e}_2 \cdot \boldsymbol{Q}\boldsymbol{e}_3 = \boldsymbol{e}_2 \cdot (\beta_2\boldsymbol{e}_2 + \beta_3\boldsymbol{e}_3) = \beta_2 = -\sin\theta, \tag{47b}$$
$$\boldsymbol{e}_3 \cdot \boldsymbol{Q}\boldsymbol{e}_2 = \boldsymbol{e}_3 \cdot (\alpha_2\boldsymbol{e}_2 + \alpha_3\boldsymbol{e}_3) = \alpha_3 = \sin\theta, \tag{47c}$$
$$\boldsymbol{e}_3 \cdot \boldsymbol{Q}\boldsymbol{e}_3 = \boldsymbol{e}_3 \cdot (\beta_2\boldsymbol{e}_2 + \beta_3\boldsymbol{e}_3) = \beta_3 = \cos\theta. \tag{47d}$$

This, together with the equations in (46) indicates that only five of the nine components of the summation (22) survive and we get:

$$Q = e_1 \otimes e_1 + (e_2 \otimes e_2)\cos\theta - (e_2 \otimes e_3)\sin\theta + (e_3 \otimes e_2)\sin\theta + (e_3 \otimes e_3)\cos\theta,$$

which is equivalent to (45a). □

It is left to an exercise to show that

$$\iota_1(Q) = 1 + 2\cos\theta, \quad \iota_2(Q) = 1 + 2\cos\theta, \quad \iota_3(Q) = 1. \tag{48}$$

Remark 12. In view of (29) and the principal invariants found in (48), the characteristic polynomial of Q is

$$\begin{aligned}\chi(\lambda) &= \lambda^3 - (1 + 2\cos\theta)\lambda^2 + (1 + 2\cos\theta)\lambda - 1\\ &= (\lambda - 1)(\lambda^2 - 2\lambda\cos\theta + 1).\end{aligned}$$

The discriminant of the quadratic factor is $4(\cos^2\theta - 1)$, which is negative for any θ other than 0 or π. Therefore, under this restriction, the characteristic equation has only one real root, and Q has only one eigenvalue.

12. An orthogonal tensor as a rotation

Theorem 10 shows that any orthogonal tensor has the representation (45a). Here we wish to investigate that representation's geometric significance.

Recall that e_1 in (45a) is the sole eigenvector of Q. We are now going to show that that applying the orthogonal tensor Q to an arbitrary vector $u \in \mathcal{V}$ a amounts to rotating u by the angle θ about the axis spanned by e_1. More generally, applying Q to any subset $\mathcal{B} \subset \mathcal{V}$ amounts to rotating $\mathcal{B}$ about that axis by θ.

Figure 3 depicts the vectors of the frame $\{e_1, e_2, e_3\}$, where e_1 is the eigenvector of Q. It also shows:

- an arbitrary vector u;
- the result Qu of applying Q to u;
- the vectors a and b which are the orthogonal projections of u and Qu onto the "equatorial plane", that is, the plane spanned by e_2 and e_3.

We observe that the component of u along e_1 is $(u \cdot e_1)e_1$, whence the component of u on the equatorial plane is $a = u - (u \cdot e_1)e_1 = u - u_1 e_1$ in the component notation $u_j = u \cdot e_j$. To determine an expression for b, we apply (45a) to u:

$$\begin{aligned}Qu &= (e_1 \otimes e_1)u + (e_2 \otimes e_2 + e_3 \otimes e_3)u\cos\theta - (e_2 \otimes e_3 - e_3 \otimes e_2)u\sin\theta,\\ &= (e_1 \cdot u)e_1 + \big((e_2 \cdot u)e_2 + (e_3 \cdot u)e_3\big)\cos\theta - \big((e_3 \cdot u)e_2 - (e_2 \cdot u)e_3\big)\sin\theta\\ &= u_1 e_1 + (u_2 e_2 + u_3 e_3)\cos\theta - (u_3 e_2 - u_2 e_3)\sin\theta.\end{aligned}$$

The projection of Qu onto the equatorial plane is obtained by removing the e_1 component. Thus, $b = (u_2 e_2 + u_3 e_3)\cos\theta - (u_3 e_2 - u_2 e_3)\sin\theta$. Then, a straightforward calculation with

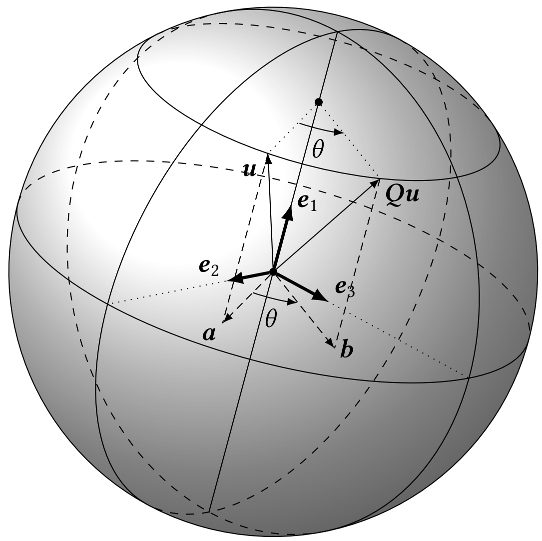


FIGURE 3. Applying the orthogonal tensor $\boldsymbol{Q}$ rotates the world by an angle θ about the axis spanned by the eigenvector $\boldsymbol{e}_1$. In particular, the (arbitrary) vector $\boldsymbol{u}$ goes into the vector $\boldsymbol{Qu}$. The projections of $\boldsymbol{u}$ and $\boldsymbol{Qu}$ onto the equatorial plane are $\boldsymbol{a}$ and $\boldsymbol{b}$.

the expression obtained for $\boldsymbol{a}$ and $\boldsymbol{b}$ shows that $\|\boldsymbol{a}\|^2 = \|\boldsymbol{b}\|^2 = u_2^2 + u_3^2$, and that

$$\boldsymbol{a}\cdot\boldsymbol{b} = (\boldsymbol{u} - u_1\boldsymbol{e}_1)\cdot\big((u_2\boldsymbol{e}_2 + u_3\boldsymbol{e}_3)\cos\theta - (u_3\boldsymbol{e}_2 - u_2\boldsymbol{e}_3)\sin\theta\big)$$
$$= (u_2^2 + u_3^2)\cos\theta. = \|\boldsymbol{a}\|\,\|\boldsymbol{b}\|\cos\theta,$$

from which we conclude that the angle between the vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ is θ, as asserted.

13. The exponential function

We define the exponential of a tensor $\boldsymbol{A}$ via

$$e^{\boldsymbol{A}} = \sum_{n=0}^{\infty}\frac{1}{n!}\boldsymbol{A}^n = \boldsymbol{I} + \boldsymbol{A} + \frac{1}{2}\boldsymbol{A}^2 + \cdots. \tag{49}$$

It can be shown (proof later) that this series converges for any $\boldsymbol{A}\in\mathcal{L}$, and therefore $e^{\boldsymbol{A}}$ is well-defined. Moreover, since $\boldsymbol{A}^n$ is in $\mathcal{L}$ for any n, then $e^{\boldsymbol{A}}\in\mathcal{L}$.

Let $\boldsymbol{X}(t) = e^{\boldsymbol{A}t}$, $t\in\mathbb{R}$. Then it is not difficult to show that $\boldsymbol{X}(t)$ is the unique solution of the initial value problem

$$\dot{\boldsymbol{X}}(t) = \boldsymbol{A}\boldsymbol{X}(t), \quad \boldsymbol{X}(0) = \boldsymbol{I}, \tag{50}$$

of the tensorial differential equation where $\dot{\boldsymbol{X}}(t)$ indicates the derivative $\frac{d}{dt}\boldsymbol{X}(t)$.

Proposition 6. *For any $\boldsymbol{A}\in\mathcal{L}$ we have*

$$\det e^{\boldsymbol{A}t} = e^{(\operatorname{tr}\boldsymbol{A})\,t}. \tag{51}$$

(proof will be provided later)

Proposition 7. *Let $\boldsymbol{W}$ be a skew-symmetric tensor. Then $\boldsymbol{Q} = e^{\boldsymbol{W}}$ is orthogonal.*

Proof. We know that $\boldsymbol{X}(t) = e^{\boldsymbol{W}t}$ is the solution of the initial value problem

$$\dot{\boldsymbol{X}}(t) = \boldsymbol{W}\boldsymbol{X}(t), \quad \boldsymbol{X}(0) = \boldsymbol{I},$$

Let $\boldsymbol{R}(t) = \boldsymbol{X}(t)\boldsymbol{X}(t)^T$. Then

$$\dot{\boldsymbol{R}} = \dot{\boldsymbol{X}}\boldsymbol{X}^T + \boldsymbol{X}\dot{\boldsymbol{X}}^T = (\boldsymbol{W}\boldsymbol{X})\boldsymbol{X}^T + \boldsymbol{X}(\boldsymbol{W}\boldsymbol{X})^T = \boldsymbol{W}(\boldsymbol{X}\boldsymbol{X}^T) + (\boldsymbol{X}\boldsymbol{X})^T)\boldsymbol{W}^T = \boldsymbol{W}\boldsymbol{R} - \boldsymbol{R}\boldsymbol{W}.$$

Since $\boldsymbol{R}(0) = \boldsymbol{X}(0)\boldsymbol{X}(0)^T = \boldsymbol{I}$, we see that $\boldsymbol{R}(t)$ is the (unique) solution of the initial value problem

$$\dot{\boldsymbol{R}} = \boldsymbol{W}\boldsymbol{R} - \boldsymbol{R}\boldsymbol{W}, \quad \boldsymbol{R}(0) = \boldsymbol{I}.$$

But we observe that $\boldsymbol{R}(t) \equiv \boldsymbol{I}$ is also a solution of that initial value problem. By uniqueness, we conclude that $\boldsymbol{R}(t) \equiv \boldsymbol{I}$. Consequently, $\boldsymbol{X}(t)\boldsymbol{X}(t)^T = \boldsymbol{I}$, that is, $\boldsymbol{X}(t)$ is an orthogonal tensor for all t. But $\boldsymbol{X}(t) = e^{\boldsymbol{W}t}$. Therefore, $e^{\boldsymbol{W}t}$ is an orthogonal tensor for all t. In particular particular, $e^{\boldsymbol{W}}$ is an orthogonal tensor. □

14. Polar decomposition

Lemma 1. *Let $\boldsymbol{R}$ be an orthogonal tensor and $\boldsymbol{V}$ be a symmetric positive definite tensor. Then $\boldsymbol{R}^T\boldsymbol{V}\boldsymbol{R}$ is symmetric and positive definite tensor.*

Proof. The tensor $\boldsymbol{R}^T\boldsymbol{V}\boldsymbol{R}$ is clearly symmetric. We need to show that the quadratic form $\boldsymbol{u}\cdot\boldsymbol{R}^T\boldsymbol{V}\boldsymbol{R}\boldsymbol{u} > 0$ for all nonzero vectors $\boldsymbol{u}$. To see that, we recall that according to Theorem 7, the tensor $\boldsymbol{V}$ has a unique symmetric positive definite square root which we write as $\boldsymbol{V}^{1/2}$. Then

$$\boldsymbol{u}\cdot\boldsymbol{R}^T\boldsymbol{V}\boldsymbol{R}\boldsymbol{u} = \boldsymbol{R}\boldsymbol{u}\cdot\boldsymbol{V}\boldsymbol{R}\boldsymbol{u} = \boldsymbol{R}\boldsymbol{u}\cdot\boldsymbol{V}^{1/2}\boldsymbol{V}^{1/2}\boldsymbol{R}\boldsymbol{u} = \boldsymbol{V}^{1/2}\boldsymbol{R}\boldsymbol{u}\cdot\boldsymbol{V}^{1/2}\boldsymbol{R}\boldsymbol{u} = \|\boldsymbol{V}^{1/2}\boldsymbol{R}\boldsymbol{u}\|^2 \geq 0,$$

which shows that $\boldsymbol{R}^T\boldsymbol{V}\boldsymbol{R}$ is positive semi-definite. Moreover, if the quadratic form is zero, then $\boldsymbol{V}^{1/2}\boldsymbol{R}\boldsymbol{u} = \boldsymbol{0}$, but since $\boldsymbol{V}^{1/2}$ is positive-definite, it is invertible, and therefore upon applying the inverse of $\boldsymbol{V}^{1/2}$ to the previous result, we see that $\boldsymbol{R}\boldsymbol{u} = \boldsymbol{0}$. Then applying $\boldsymbol{R}^T$ (that is, $\boldsymbol{R}^{-1}$) to this, we arrive at $\boldsymbol{u} = \boldsymbol{0}$. In conclusion, the quadratic form is zero only when $\boldsymbol{u}$ is zero, which indicates that $\boldsymbol{R}^T\boldsymbol{V}\boldsymbol{R}$ is positive-definite. □

Theorem 11 (Polar decomposition)**.** *An invertible tensor $\boldsymbol{A} \in \mathcal{L}$ admits a* right polar decomposition

$$\boldsymbol{A} = \boldsymbol{R}\boldsymbol{U}, \tag{52}$$

and a left polar decomposition

$$\boldsymbol{A} = \boldsymbol{V}\boldsymbol{R}, \tag{53}$$

where $\boldsymbol{R}$ is orthogonal, and $\boldsymbol{U}$ and $\boldsymbol{V}$ are symmetric and positive definite. The tensors $\boldsymbol{R}, \boldsymbol{U}, \boldsymbol{V}$ are uniquely determined by $\boldsymbol{A}$.

Proof. According to the exercises 15 and 16, the tensors $\boldsymbol{A}^T\boldsymbol{A}$ and $\boldsymbol{A}\boldsymbol{A}^T$ are symmetric and positive definite, and therefore by Theorem 7 they have symmetric positive-definite square roots, say $\boldsymbol{U}$ and $\boldsymbol{V}$, respectively. Thus

$$\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{U}^2, \quad \boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{V}^2.$$

Let

$$R = AU^{-1}, \quad P = V^{-1}A. \tag{54}$$

Let us show that both $\boldsymbol{R}$ and $\boldsymbol{P}$ are orthogonal tensors:

$$\begin{aligned}
\boldsymbol{R}^T\boldsymbol{R} &= \left(\boldsymbol{A}\boldsymbol{U}^{-1}\right)^T\left(\boldsymbol{A}\boldsymbol{U}^{-1}\right) = \left(\boldsymbol{U}^{-1}\boldsymbol{A}^T\right)\left(\boldsymbol{A}\boldsymbol{U}^{-1}\right) = \boldsymbol{U}^{-1}\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{U}^{-1} = \boldsymbol{U}^{-1}\boldsymbol{U}^2\boldsymbol{U}^{-1} = \boldsymbol{I},\\
\boldsymbol{P}\boldsymbol{P}^T &= \left(\boldsymbol{V}^{-1}\boldsymbol{A}\right)\left(\boldsymbol{V}^{-1}\boldsymbol{A}\right)^T = \left(\boldsymbol{V}^{-1}\boldsymbol{A}\right)\left(\boldsymbol{A}^T\boldsymbol{V}^{-1}\right) = \boldsymbol{V}^{-1}\boldsymbol{A}\boldsymbol{A}^T\boldsymbol{V}^{-1} = \boldsymbol{V}^{-1}\boldsymbol{V}^2\boldsymbol{V}^{-1} = \boldsymbol{I}.
\end{aligned}$$

The first equation in (54) is equivalent to (52). To complete the proof of (52) we need to show the uniqueness of the decomposition. Toward that end, let's suppose that $\boldsymbol{A}$ admits and alternative right polar decomposition $\boldsymbol{A} = \boldsymbol{R}'\boldsymbol{U}'$, where $\boldsymbol{R}'$ is orthogonal and $\boldsymbol{U}'$ is symmetric and positive definite. Then

$$\boldsymbol{A}^T\boldsymbol{A} = \left(\boldsymbol{R}'\boldsymbol{U}'\right)^T\left(\boldsymbol{R}'\boldsymbol{U}'\right) = \boldsymbol{U}'\boldsymbol{R}'^T\boldsymbol{R}'\boldsymbol{U}' = \boldsymbol{U}'^2,$$

that is, $\boldsymbol{U}'$ is a symmetric and positive definite square root of $\boldsymbol{A}^T\boldsymbol{A}$. But according to Theorem 7, such square root is unique, and therefore $\boldsymbol{U}' = \boldsymbol{U}$. Then

$$\boldsymbol{R}' = \boldsymbol{A}\boldsymbol{U}'^{-1} = \boldsymbol{A}\boldsymbol{U}^{-1} = \boldsymbol{R},$$

which proves the uniqueness of the right polar decomposition. The uniqueness of the left polar decomposition may be proved in the same way.

The second equation in (54) implies that $\boldsymbol{A} = \boldsymbol{V}\boldsymbol{P}$. In order to arrive at (53), we need to show that $\boldsymbol{P} = \boldsymbol{R}$. For that let's observe that $\boldsymbol{P}\boldsymbol{P}^T = \boldsymbol{I}$ and therefore

$$\boldsymbol{A} = \boldsymbol{V}\boldsymbol{P} = \left(\boldsymbol{P}\boldsymbol{P}^T\right)\boldsymbol{V}\boldsymbol{P} = \boldsymbol{P}\left(\boldsymbol{P}^T\boldsymbol{V}\boldsymbol{P}\right).$$

According to Lemma 1, the tensor $\boldsymbol{P}^T\boldsymbol{V}\boldsymbol{P}$ is symmetric and positive definite, therefore, $\boldsymbol{A} = \boldsymbol{P}\left(\boldsymbol{P}^T\boldsymbol{V}\boldsymbol{P}\right)$ is a right polar decomposition of $\boldsymbol{A}$. But we have already seen that the right polar decomposition is unique, and therefore it should agree with $\boldsymbol{A} = \boldsymbol{R}\boldsymbol{U}$ in (52). We conclude that $\boldsymbol{P} = \boldsymbol{R}$, completing the proof of (53). □

Remark 13. The name "polar decomposition" is associated with equations (52) and (53) in a loose analogy with the polar representation of complex numbers $z = re^{i\theta}$, where the rotation $e^{i\theta}$ in the complex plane in likened to the action of the orthogonal tensor $\boldsymbol{R}$ in the polar decomposition.

Remark 14. A *simple shear* is a deformation in the Cartesian x–y given by the mapping

$$\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} x + 2y\tan\gamma \\ y \end{pmatrix} = \begin{pmatrix} 1 & 2\tan\gamma \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}. \tag{55}$$

The mapping preserves the y coordinates of the points, therefore the points move in the horizontal direction. The horizontal displacement, $2y\tan\gamma$, is proportional to the y coordinate, therefore points with greater y are displaced by proportionally greater amounts. The constant of proportionality is taken to be $2\tan\gamma$ for at least two reasons. First, the algebra is simpler this way compared to what it would have been with a generic proportionality constant, say, c. Second, the angle γ turns out to have a geometric interpretation as we shall see toward the end of the calculations.

The gradient $\boldsymbol{F}$ of the mapping defined in (55), and the corresponding right Cauchy–Green strain tensor are

$$\boldsymbol{F} = \begin{pmatrix} 1 & 2\tan\gamma \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{C} = \boldsymbol{F}^T\boldsymbol{F} = \begin{pmatrix} 1 & 2\tan\gamma \\ 2\tan\gamma & 1 + 4\tan^2\gamma \end{pmatrix}.$$

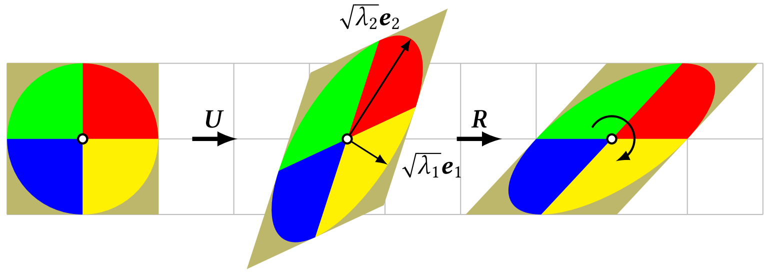


FIGURE 4. A homogeneous deformation with the deformation gradient $\boldsymbol{F} = \boldsymbol{RU}$ acts on any neighborhood of the origin by shrinking/stretching the neighborhood along the eigenvectors $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ by the factors $\sqrt{\lambda_1}$ and $\sqrt{\lambda_2}$, followed by a rigid rotation by the orthogonal tensor $\boldsymbol{R}$. The eigenvalues of $\boldsymbol{U}$ are the square roots of the eigenvalues λ_1 and λ_2 of the right Cauchy–Green strain tensor $\boldsymbol{C} = \boldsymbol{F}^T\boldsymbol{F}$.

The eigenvalues and eigenvectors of $\boldsymbol{C}$ are

$$\lambda_1 = 1/\sigma^2, \quad \boldsymbol{e}_1 = \begin{pmatrix} \sigma \\ 1 \end{pmatrix}, \qquad \lambda_2 = \sigma^2, \quad \boldsymbol{e}_2 = \begin{pmatrix} -1 \\ \sigma \end{pmatrix},$$

where $\sigma = (1 - \sin\gamma)/\cos\gamma$. Since $\boldsymbol{U}^2 = \boldsymbol{C}$, the eigenvectors of $\boldsymbol{U}$ coincide with those of $\boldsymbol{C}$, and its eigenvalues are $\sqrt{\lambda_1} = 1/\sigma$ and $\sqrt{\lambda_2} = \sigma$. Letting $\boldsymbol{P}$ be the matrix with columns $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$, and $\boldsymbol{L}$ be the diagonal matrix with the eigenvalues of $\boldsymbol{U}$ on the diagonal, we have $\boldsymbol{U} = \boldsymbol{PLP}^{-1}$, and therefore

$$\boldsymbol{U} = \begin{pmatrix} \sigma & -1 \\ 1 & \sigma \end{pmatrix} \begin{pmatrix} 1/\sigma & 0 \\ 0 & \sigma \end{pmatrix} \begin{pmatrix} \sigma & -1 \\ 1 & \sigma \end{pmatrix}^{-1} = \begin{pmatrix} \cos\gamma & \sin\gamma \\ \sin\gamma & \frac{1+\sin^2\gamma}{\cos\gamma} \end{pmatrix}.$$

Finally, we calculate the orthogonal tensor $\boldsymbol{R}$:

$$\boldsymbol{R} = \boldsymbol{FU}^{-1} = \begin{pmatrix} 1 & 2\tan\gamma \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\gamma & \sin\gamma \\ \sin\gamma & \frac{1+\sin^2\gamma}{\cos\gamma} \end{pmatrix}^{-1} = \begin{pmatrix} \cos\gamma & \sin\gamma \\ -\sin\gamma & \cos\gamma \end{pmatrix},$$

which represents a *clockwise* rotation by the angle γ. Figure 4 illustrates the effect of applying the deformation gradient $\boldsymbol{F}$ to a neighborhood of the origin. The neighborhood shrinks/stretches along the eigenvectors $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ of $\boldsymbol{U}$ by the factors $\sqrt{\lambda_1}$ and $\sqrt{\lambda_2}$, followed by clockwise rigid rotation by the angle γ.

15. More tensor algebra

Lemma 2. *Let $\boldsymbol{e}$ be a unit vector. Then the tensor $\boldsymbol{Q} = \boldsymbol{I} - 2\boldsymbol{e} \otimes \boldsymbol{e}$ is orthogonal.*

Proof. From (21b) we see that $\boldsymbol{Q} = \boldsymbol{Q}^T$. Therefore

$$\boldsymbol{Q}^T\boldsymbol{Q} = (\boldsymbol{I} - 2\boldsymbol{e} \otimes \boldsymbol{e})(\boldsymbol{I} - 2\boldsymbol{e} \otimes \boldsymbol{e}) = \boldsymbol{I} - 4\boldsymbol{e} \otimes \boldsymbol{e} + 4(\boldsymbol{e} \otimes \boldsymbol{e})(\boldsymbol{e} \otimes \boldsymbol{e}) = \boldsymbol{I}.$$

In the last step we have made use of the identity (21c). □

Lemma 3. *The tensor $\boldsymbol{Q}$ defined in the previous lemma is a reflection operator about the plane perpendicular to $\boldsymbol{e}$.*

Proof. We have:

$$\boldsymbol{Q}\boldsymbol{e} = (\boldsymbol{I} - 2\boldsymbol{e} \otimes \boldsymbol{e})\,\boldsymbol{e} = \boldsymbol{e} - 2(\boldsymbol{e} \cdot \boldsymbol{e})\,\boldsymbol{e} = -\boldsymbol{e}.$$

Moreover, for any vector $\boldsymbol{e}'$ such that $\boldsymbol{e}' \cdot \boldsymbol{e} = 0$ we have:

$$\boldsymbol{Q}\boldsymbol{e}' = (\boldsymbol{I} - 2\boldsymbol{e} \otimes \boldsymbol{e})\,\boldsymbol{e}' = \boldsymbol{e}' - 2(\boldsymbol{e}' \cdot \boldsymbol{e})\,\boldsymbol{e} = \boldsymbol{e}'.$$

□

Lemma 4. *Let the tensor $\boldsymbol{Q}$ be as in the previous lemma. If $\boldsymbol{Q}\boldsymbol{v} = -\boldsymbol{v}$ for a vector $\boldsymbol{v}$, then $\boldsymbol{v}$ must be a multiple of the vector $\boldsymbol{e}$.*

Proof. We have $\boldsymbol{Q}\boldsymbol{v} = (\boldsymbol{I} - 2\boldsymbol{e} \otimes \boldsymbol{e})\,\boldsymbol{v} = \boldsymbol{v} - 2(\boldsymbol{e} \cdot \boldsymbol{v})\,\boldsymbol{e}$. Setting this equal to $-\boldsymbol{v}$ we get $\boldsymbol{v} = (\boldsymbol{e} \cdot \boldsymbol{v})\,\boldsymbol{e}$. □

Lemma 5. *Let $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ and $\{\boldsymbol{e}'_1, \boldsymbol{e}'_2, \boldsymbol{e}'_3\}$ be two frames in $\mathbb{E}_3$, and let*

$$\boldsymbol{Q} = \boldsymbol{e}'_i \otimes \boldsymbol{e}_i. \qquad \textit{(summation over i!)}$$

Then $\boldsymbol{Q}$ is orthogonal and $\boldsymbol{Q}\boldsymbol{e}_i = \boldsymbol{e}'_i$ for each i, and thus, the orthogonal transformation $\boldsymbol{Q}$ rotates the frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ to the frame $\{\boldsymbol{e}'_1, \boldsymbol{e}'_2, \boldsymbol{e}'_3\}$.

Proof. From (21b) we see that $\boldsymbol{Q}^T = \boldsymbol{e}_i \otimes \boldsymbol{e}'_i$, therefore:

$$\boldsymbol{Q}^T\boldsymbol{Q} = (\boldsymbol{e}_j \otimes \boldsymbol{e}'_j)\,(\boldsymbol{e}'_i \otimes \boldsymbol{e}_i) = (\boldsymbol{e}'_j \cdot \boldsymbol{e}'_i)\,(\boldsymbol{e}_j \otimes \boldsymbol{e}_i) = \delta_{ij}\,(\boldsymbol{e}_j \otimes \boldsymbol{e}_i) = \boldsymbol{e}_i \otimes \boldsymbol{e}_i = \boldsymbol{I},$$

which shows that $\boldsymbol{Q}$ is orthogonal. Here we have applied the identity (21c) and the result of Exercise 11 that $\boldsymbol{e}_i \otimes \boldsymbol{e}_i = \boldsymbol{I}$. To complete the lemma's proof, we calculate

$$\boldsymbol{Q}\boldsymbol{e}_i = (\boldsymbol{e}'_j \otimes \boldsymbol{e}_j)\,\boldsymbol{e}_i = (\boldsymbol{e}_j \cdot \boldsymbol{e}_i)\,\boldsymbol{e}'_j = \delta_{ij}\,\boldsymbol{e}'_j = \boldsymbol{e}'_i.$$

□

The next two lemmas, due to Chao-Cheng Wang, are collectively known as *Wang's Lemma.*

Lemma 6 (Wang's Lemma, Part 1)**.** *Let $\boldsymbol{A} \in \mathcal{L}_{sym}$ have eigenvalues λ_1, λ_2, λ_3 and let $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ be a corresponding orthonormal set of eigenvectors. If the eigenvalues are distinct then:*

(a) $\boldsymbol{A} = \lambda_1\boldsymbol{e}_1 \otimes \boldsymbol{e}_1 + \lambda_2\boldsymbol{e}_2 \otimes \boldsymbol{e}_2 + \lambda_3\boldsymbol{e}_3 \otimes \boldsymbol{e}_3$;
(b) *The set $\{\boldsymbol{I}, \boldsymbol{A}, \boldsymbol{A}^2\}$ is linearly independent;*
(c) $\operatorname{span}\{\boldsymbol{I}, \boldsymbol{A}, \boldsymbol{A}^2\} = \operatorname{span}\{\boldsymbol{e}_1 \otimes \boldsymbol{e}_1, \boldsymbol{e}_2 \otimes \boldsymbol{e}_2, \boldsymbol{e}_3 \otimes \boldsymbol{e}_3\}$.

Proof. Part (a) is a restatement of the Spectral Theorem 6. For part (b), we need to show that $\alpha\boldsymbol{A}^2 + \beta\boldsymbol{A} + \gamma\boldsymbol{I} = \boldsymbol{0}$ implies that $\alpha = \beta = \gamma = 0$. To see this, apply the two sides of the equation $\alpha\boldsymbol{A}^2 + \beta\boldsymbol{A} + \gamma\boldsymbol{I} = \boldsymbol{0}$ to the eigenvector $\boldsymbol{e}_i$ to get $\alpha\lambda_i^2 + \beta\lambda_i + \gamma = 0$ for $i = 1, 2, 3$. We see that this quadratic equation has three distinct roots, and therefore its coefficients must be zero.

As to part (c), let $\mathcal{M} = \operatorname{span}\{\boldsymbol{e}_1 \otimes \boldsymbol{e}_1, \boldsymbol{e}_2 \otimes \boldsymbol{e}_2, \boldsymbol{e}_3 \otimes \boldsymbol{e}_3\}$. From Exercise 11 we know that $\boldsymbol{I} = \sum_i \boldsymbol{e}_i \otimes \boldsymbol{e}_i$. Furthermore, $\boldsymbol{A} = \sum_i \lambda_i\boldsymbol{e}_i \otimes \boldsymbol{e}_i$ and $\boldsymbol{A}^2 = \sum_i \lambda_i^2\boldsymbol{e}_i \otimes \boldsymbol{e}_i$. Therefore $\boldsymbol{I}, \boldsymbol{A}, \boldsymbol{A}^2 \in \mathcal{M}$. But by part (b), $\{\boldsymbol{I}, \boldsymbol{A}, \boldsymbol{A}^2\}$ is linearly independent, therefore it is a basis for the three-dimensional space $\mathcal{M}$. □

Lemma 7 (Wang's Lemma, Part 2)**.** *Let λ_1, λ_2, λ_3 be the eigenvalues of $\boldsymbol{A} \in \mathcal{L}_{sym}$, and suppose that $\lambda_3 = \lambda_2 \neq \lambda_1$. Let $\boldsymbol{e}$ be a unit eigenvector corresponding to the eigenvalue λ_1. Then we have:*

(a) $\boldsymbol{A} = \lambda_1 \boldsymbol{e} \otimes \boldsymbol{e} + \lambda_2(\boldsymbol{I} - \boldsymbol{e} \otimes \boldsymbol{e})$.
(b *The set* $\{\boldsymbol{I}, \boldsymbol{A}\}$ *is linearly independent;*
(c) $\operatorname{span}\{\boldsymbol{I}, \boldsymbol{A}\} = \operatorname{span}\{\boldsymbol{e} \otimes \boldsymbol{e}, \boldsymbol{I} - \boldsymbol{e} \otimes \boldsymbol{e}\}$.

The proof of this lemma is left for Exercise 31.

16. Isotropic functions

Much of this section's analysis is due to R. S. Rivlin and J. L. Ericksen [21]. The presentation here follows the significantly simplified version in [7].

A subset of $\mathcal{A}$ of $\mathcal{L}$ is said to be *invariant under* $\mathcal{L}_{orth}$ if $\boldsymbol{A} \in \mathcal{A}$ implies that $\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T \in \mathcal{A}$ for all $\boldsymbol{Q} \in \mathcal{L}_{\text{orth}}$. Some frequently occurring invariant subspaces are listed in the following lemma whose proof is left for Exercise 32.

Lemma 8. *The following subsets of* $\mathcal{L}$ *are invariant under* $\mathcal{L}_{orth}$*:*

$$\mathcal{L}, \quad \mathcal{L}_{sym}, \quad \mathcal{L}^+, \quad \mathcal{L}^+_{sym}, \quad \mathcal{L}_{orth}, \quad \mathcal{L}^+_{orth}.$$

Note: In the rest of this section, $\mathcal{A}$ signifies an invariant subset of $\mathcal{L}$ under $\mathcal{L}_{\text{orth}}$. We are interested in generally nonlinear functions $g : \mathcal{A} \to \mathbb{R}$ and $G : \mathcal{A} \to \mathcal{L}$.

A function $g : \mathcal{A} \to \mathbb{R}$ is said to be *isotropic* if

$$g(\boldsymbol{A}) = g(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T) \quad \text{for all } \boldsymbol{A} \in \mathcal{A} \text{ and all } \boldsymbol{Q} \in \mathcal{L}_{\text{orth}}. \tag{56}$$

A function $\boldsymbol{G} : \mathcal{A} \to \mathcal{L}$ is said to be *isotropic* if

$$\boldsymbol{Q}\, \boldsymbol{G}(\boldsymbol{A})\, \boldsymbol{Q}^T = \boldsymbol{G}(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T) \quad \text{for all } \boldsymbol{A} \in \mathcal{A} \text{ and all } \boldsymbol{Q} \in \mathcal{L}_{\text{orth}}. \tag{57}$$

Lemma 9. *Suppose* $g : \mathcal{A} \to \mathbb{R}$ *is such that*

$$g(\boldsymbol{A}) = g(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T) \quad \textit{for all } \boldsymbol{A} \in \mathcal{A} \textit{ and all } \boldsymbol{Q} \in \mathcal{L}^+_{orth}.$$

Then g *is isotropic.*

Proof. We need to show that $g(\boldsymbol{A}) = g(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T)$ for all $\boldsymbol{Q} \in \mathcal{L}_{\text{orth}}$. By the lemma's hypothesis we already know that this holds when $\boldsymbol{Q} \in \mathcal{L}^+_{\text{orth}}$. It remains to show that it also holds for $\boldsymbol{Q}$ in $\mathcal{L}_{\text{orth}} \backslash \mathcal{L}^+_{\text{orth}}$. But if $\boldsymbol{Q}$ is of the latter type, then $-\boldsymbol{Q}$ is in $\mathcal{L}^+_{\text{orth}}$, therefore for such a $\boldsymbol{Q}$ we calculate

$$g(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T) = g\big((-\boldsymbol{Q})\boldsymbol{A}(-\boldsymbol{Q}^T)\big) = g(\boldsymbol{A}). \qquad \square$$

Lemma 10. *Suppose* $G : \mathcal{A} \to \mathcal{L}$ *is such that*

$$\boldsymbol{Q}\, G(\boldsymbol{A})\, \boldsymbol{Q}^T = G(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T) \quad \textit{for all } \boldsymbol{A} \in \mathcal{A} \textit{ and all } \boldsymbol{Q} \in \mathcal{L}^+_{orth}.$$

Then G *is isotropic.*

We leave the proof of this lemma for Exercise 33.

16.1. **Scalar-valued isotropic functions.** Recall the definition a scalar-valued isotropic function in (56). The determinant, viewed as a (nonlinear) function from $\mathcal{L}$ to $\mathbb{R}$, is isotropic since $\det(\boldsymbol{QAQ}^T) = (\det \boldsymbol{Q})(\det \boldsymbol{A})(\det \boldsymbol{Q}^T) = \det \boldsymbol{A}$ for all $\boldsymbol{A} \in \mathcal{L}$ and $\boldsymbol{Q} \in \mathcal{L}_{\text{orth}}$. As another example, the trace, viewed as a (linear) function, from $\mathcal{L}$ to $\mathbb{R}$, is isotropic since

$$\operatorname{tr}\big(\boldsymbol{QAQ}^T\big) = \operatorname{tr}\big((\boldsymbol{QA})(\boldsymbol{Q}^T)\big) = \operatorname{tr}\big((\boldsymbol{Q}^T)(\boldsymbol{QA})\big) = \operatorname{tr} \boldsymbol{A}$$

for all $\boldsymbol{A} \in \mathcal{L}$ and $\boldsymbol{Q} \in \mathcal{L}_{\text{orth}}$. Here we have applied the identity $\operatorname{tr}(\boldsymbol{AB}) = \operatorname{tr}(\boldsymbol{BA})$ from Exercise 12.

Lemma 11. *The three principal invariants $\iota_i : \mathcal{L} \to \mathbb{R}$, $i = 1, 2, 3$, are isotropic functions.*

Proof. In the two examples above we have observed that $\iota_1 = \operatorname{tr}$ and $\iota_3 = \det$ are isotropic. It remains to verify that ι_2 is isotropic. For this, recall (24b) on page 13, and observe that $(\boldsymbol{QAQ}^T)^2 = \boldsymbol{QA}^2\boldsymbol{Q}^T$, hence $\operatorname{tr}\big((\boldsymbol{QAQ}^T)^2\big) = \operatorname{tr}(\boldsymbol{QA}^2\boldsymbol{Q}^T) = \operatorname{tr}(\boldsymbol{A}^2)$. $\square$

Lemma 12. *The function $g : \mathcal{A} \subset \mathcal{L}_{sym} \to \mathbb{R}$ is isotropic if and only if there exists a function $\tilde{g} : \mathbb{R}^3 \to \mathbb{R}$ such that*

$$g(\boldsymbol{A}) = \tilde{g}\big(\iota_1(\boldsymbol{A}), \iota_2(\boldsymbol{A}), \iota_3(\boldsymbol{A})\big), \quad \textit{for all } \boldsymbol{A} \in \mathcal{A}. \tag{58}$$

Proof. If g is of the form (58), then it is isotropic since by the previous lemma we have

$$\begin{aligned} g(\boldsymbol{QAQ}^T) &= \tilde{g}\big(\iota_1(\boldsymbol{QAQ}^T), \iota_2(\boldsymbol{QAQ}^T), \iota_3(\boldsymbol{QAQ}^T)\big) \\ &= \tilde{g}\big(\iota_1(\boldsymbol{A}), \iota_2(\boldsymbol{A}), \iota_3(\boldsymbol{A})\big) = g(\boldsymbol{A}) \quad \text{for all } \boldsymbol{A} \in \mathcal{A} \text{ and } \boldsymbol{Q} \in \mathcal{L}_{\text{orth}}. \end{aligned}$$

As to the converse, assume g is isotropic. We wish to show that $g(\boldsymbol{A})$ depends solely in the invariants of $\boldsymbol{A}$, that is, the invariants of $\boldsymbol{A}$ suffice to determine the value of $g(\boldsymbol{A})$. To put it in yet another way, if $\boldsymbol{A}$ and $\boldsymbol{B}$ have the same invariants, then $g(\boldsymbol{A}) = g(\boldsymbol{B})$.

Thus, take $\boldsymbol{A}$ and $\boldsymbol{B}$ in $\mathcal{L}_{\text{sym}}$, and let ι_1, ι_2, ι_3 denote their common invariants. Then $\boldsymbol{A}$ and $\boldsymbol{B}$ share the common characteristic equation:

$$\lambda^3 - \iota_1\lambda^2 + \iota_2\lambda - \iota_3 = 0,$$

and therefore they have the same eigenvalues. It follows that, their spectral decompositions look like these:

$$\boldsymbol{A} = \sum_i \lambda_i \boldsymbol{e}_i \otimes \boldsymbol{e}_i, \qquad \boldsymbol{B} = \sum_i \lambda_i \boldsymbol{e}_i' \otimes \boldsymbol{e}_i',$$

where $\{e_i'\}_{i=1}^3$ to $\{e_i\}_{i=1}^3$ are frames. Let $\boldsymbol{Q}$ be the orthogonal transformation that rotates the frame $\{e_i\}_{i=1}^3$ to $\{e_i'\}_{i=1}^3$ as in Lemma 5, that is, $\boldsymbol{e}_i' = \boldsymbol{Q}\boldsymbol{e}_i$, $i = 1, 2, 3$. Then by (21f) we have:

$$\boldsymbol{Q}\,(\boldsymbol{e}_l \otimes \boldsymbol{e}_i)\,\boldsymbol{Q}^T = (\boldsymbol{Q}\boldsymbol{e}_i) \otimes (\boldsymbol{Q}\boldsymbol{e}_i) = \boldsymbol{e}_i' \otimes \boldsymbol{e}_i'.$$

It follows that $\boldsymbol{QBQ}^T = \boldsymbol{A}$, whence:

$$g(\boldsymbol{A}) = g(\boldsymbol{QBQ}^T) = g(\boldsymbol{B}). \qquad \square$$

16.2. **Tensor-valued isotropic functions.**

Theorem 12 (Transfer Theorem)**.** *If $G : \mathcal{A} \subset \mathcal{L}_{sym} \to \mathcal{L}$ is isotropic, then any eigenvector of $\boldsymbol{A}$ is an eigenvector of $\boldsymbol{G}(\boldsymbol{A})$.*

Proof. Let λ_i, $i = 1, 2, 3$, be the eigenvalues, and $\boldsymbol{e}_i$ be an orthonormal set of the corresponding eigenvectors of $\boldsymbol{A}$. By the Spectral Theorem we have $\boldsymbol{A} = \sum_i \lambda_i \boldsymbol{e}_i \otimes \boldsymbol{e}_i$.

Let $\boldsymbol{Q} = \boldsymbol{I} - 2\boldsymbol{e}_1 \otimes \boldsymbol{e}_1$. By Lemma 2, $\boldsymbol{Q}$ is orthogonal, and thus, by (21f) we get:

$$\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T = \sum_i \lambda_i \boldsymbol{Q}(\boldsymbol{e}_i \otimes \boldsymbol{e}_i)\boldsymbol{Q}^T = \sum_i \lambda_i (\boldsymbol{Q}\boldsymbol{e}_i) \otimes (\boldsymbol{Q}\boldsymbol{e}_i) = \sum_i \lambda_i \boldsymbol{e}_i \otimes \boldsymbol{e}_i = \boldsymbol{A}.$$

In the last step we have used Lemma 3 whereby $\boldsymbol{Q}\boldsymbol{e}_i = -\boldsymbol{e}_i$ if $i = 1$ and $\boldsymbol{Q}\boldsymbol{e}_i = \boldsymbol{e}_i$ otherwise.

Now, since $\boldsymbol{G}$ is isotropic, we have $\boldsymbol{Q}\boldsymbol{G}(\boldsymbol{A})\boldsymbol{Q}^T = \boldsymbol{G}(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T) = \boldsymbol{G}(\boldsymbol{A})$, whence $\boldsymbol{Q}\boldsymbol{G}(\boldsymbol{A}) = \boldsymbol{G}(\boldsymbol{A})\boldsymbol{Q}$, therefore $\boldsymbol{Q}\boldsymbol{G}(\boldsymbol{A})\boldsymbol{e}_1 = \boldsymbol{G}(\boldsymbol{A})\boldsymbol{Q}\boldsymbol{e}_1 = -\boldsymbol{G}(\boldsymbol{A})\boldsymbol{e}_1$.

We see that $\boldsymbol{Q}$ maps the vector $\boldsymbol{G}(\boldsymbol{A})\boldsymbol{e}_1$ to its own negative, hence by Lemma 4 we conclude that $\boldsymbol{G}(\boldsymbol{A})\boldsymbol{e}_1$ is a multiple of $\boldsymbol{e}_1$, and therefore $\boldsymbol{e}_1$ is an eigenvector of $\boldsymbol{G}(\boldsymbol{A})$. □

The next theorem is central to all modern theories of continuum mechanics. The proof is somewhat long and technical and it's often omitted in textbooks on the subject. For instance, it is stated without proof in [5], page 31. The presentation in these notes is modeled after that in an appendix in [7].

Theorem 13 (Rivlin–Ericksen)**.** *The function $\boldsymbol{G} : \mathcal{A} \subset \mathcal{L}_{sym} \to \mathcal{L}_{sym}$ is isotropic if and only if*

$$\boldsymbol{G}(\boldsymbol{A}) = \alpha_0 \boldsymbol{I} + \alpha_1 \boldsymbol{A} + \alpha_2 \boldsymbol{A}^2 \quad \textit{for all } \boldsymbol{A} \in \mathcal{A}, \tag{59}$$

where α_0, α_1, and α_2 are scalar isotropic functions of $\boldsymbol{A}$.

Proof. If $\boldsymbol{G}$ has the form in (59), then it is trivial to verify that it is isotropic. Here we prove the converse. Thus, we pick an arbitrary $\boldsymbol{A} \in \mathcal{A}$ and consider three cases, as follows.

Case 1: $\boldsymbol{A}$ has three distinct eigenvalues.

According to the Spectral Theorem we have $\boldsymbol{A} = \sum_i \lambda_i \boldsymbol{e}_i \otimes \boldsymbol{e}_i$. By the Transfer Theorem (Theorem 12), $\boldsymbol{G}(\boldsymbol{A})$ has the same eigenvectors as $\boldsymbol{A}$, therefore it has a spectral decomposition of the form $\boldsymbol{G}(\boldsymbol{A}) = \sum_i \beta_i \boldsymbol{e}_i \otimes \boldsymbol{e}_i$. Then from Lemma 6 we deduce that

$$\boldsymbol{G}(\boldsymbol{A}) = \alpha_0(\boldsymbol{A})\,\boldsymbol{I} + \alpha_1(\boldsymbol{A})\,\boldsymbol{A} + \alpha_2(\boldsymbol{A})\,\boldsymbol{A}^2. \tag{60a}$$

Case 2: $\boldsymbol{A}$ has eigenvalues $\lambda_2 = \lambda_3 \neq \lambda_1$.

As in the Case 1, we have $\boldsymbol{A} = \sum_i \lambda_i \boldsymbol{e}_i \otimes \boldsymbol{e}_i$ and $\boldsymbol{G}(\boldsymbol{A}) = \sum_i \beta_i \boldsymbol{e}_i \otimes \boldsymbol{e}_i$, with $\beta_2 = \beta_3$. Therefore, with an appeal to Exercise 11 we have

$$\boldsymbol{G}(\boldsymbol{A}) = \beta_1 \boldsymbol{e}_1 \otimes \boldsymbol{e}_1 + \beta_2(\boldsymbol{e}_2 \otimes \boldsymbol{e}_2 + \boldsymbol{e}_3 \otimes \boldsymbol{e}_3) = \beta_1 \boldsymbol{e}_1 \otimes \boldsymbol{e}_1 + \beta_2(\boldsymbol{I} - \boldsymbol{e}_1 \otimes \boldsymbol{e}_1),$$

We see that $\boldsymbol{G}(\boldsymbol{A})$ is a linear combination of $\boldsymbol{e}_1 \otimes \boldsymbol{e}_1$ and $\boldsymbol{I} - \boldsymbol{e}_1 \otimes \boldsymbol{e}_1$. Therefore by Lemma 7 we have

$$\boldsymbol{G}(\boldsymbol{A}) = \alpha_0(\boldsymbol{A})\,\boldsymbol{I} + \alpha_1(\boldsymbol{A})\,\boldsymbol{A} + 0 \times \boldsymbol{A}^2. \tag{60b}$$

Case 3: $\boldsymbol{A}$ has eigenvalues $\lambda_1 = \lambda_2 = \lambda_3$.

We have $\boldsymbol{A} = \sum_i \lambda_i \boldsymbol{e}_i \otimes \boldsymbol{e}_i = \lambda_1(\boldsymbol{e}_1 \otimes \boldsymbol{e}_1 + \boldsymbol{e}_2 \otimes \boldsymbol{e}_2 + \boldsymbol{e}_3 \otimes \boldsymbol{e}_3) = \lambda_1 \boldsymbol{I}$, and therefore any nonzero vector in $\mathcal{V}$ is an eigenvector of $\boldsymbol{A}$, and by the Transfer Theorem, an eigenvector of $\boldsymbol{G}(\boldsymbol{A})$. We conclude that $\boldsymbol{G}(\boldsymbol{A})$ is a multiple of identity, and therefore

$$\boldsymbol{G}(\boldsymbol{A}) = \alpha_0(\boldsymbol{A})\,\boldsymbol{I} + 0 \times \boldsymbol{A} + 0 \times \boldsymbol{A}^2. \tag{60c}$$

We see that $\boldsymbol{G}(\boldsymbol{A})$ has the form (60a) in general, where the coefficients α_1 and α_2 may be zero in some special cases. It remains to show that the coefficients α_0, α_1 and α_2 are isotropic scalar functions, that is, they are of the form (12). Toward that end, pick an arbitrary $\boldsymbol{Q} \in \mathcal{L}_{\text{orth}}$ and compute:

$$\boldsymbol{G}(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T) = \alpha_0(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T)\,\boldsymbol{I} + \alpha_1(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T)\,\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T + \alpha_2(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T)\,(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T)^2.$$

The left hand side equals $\boldsymbol{Q}\boldsymbol{G}(\boldsymbol{A})\boldsymbol{Q}^T$ by isotropy. On the right hand side the expression $(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T)^2$ expands to $\boldsymbol{Q}\boldsymbol{A}^2\boldsymbol{Q}^T$. Therefore we get:

$$\boldsymbol{Q}\boldsymbol{G}(\boldsymbol{A})\boldsymbol{Q}^T = \alpha_0(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T)\,\boldsymbol{I} + \alpha_1(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T)\,\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T + \alpha_2(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T)\,\boldsymbol{Q}\boldsymbol{A}^2\boldsymbol{Q}^T,$$

which simplifies to

$$\boldsymbol{G}(\boldsymbol{A}) = \alpha_0(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T)\,\boldsymbol{I} + \alpha_1(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T)\,\boldsymbol{A} + \alpha_2(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T)\,\boldsymbol{A}^2.$$

By subtracting this from (60a) we obtain:

$$\big(\alpha_0(\boldsymbol{A}) - \alpha_0(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T)\big)\boldsymbol{I} + \big(\alpha_1(\boldsymbol{A}) - \alpha_1(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T)\big)\boldsymbol{A} + \big(\alpha_2(\boldsymbol{A}) - \alpha_2(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T)\big)\boldsymbol{A}^2 = \boldsymbol{0},$$

and thus, by the linear independence of the set $\{\boldsymbol{I}, \boldsymbol{A}, \boldsymbol{A}^2\}$ we conclude that

$$\alpha_0(\boldsymbol{A}) = \alpha_0(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T), \quad \alpha_1(\boldsymbol{A}) = \alpha_1(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T), \quad \alpha_2(\boldsymbol{A}) = \alpha_2(\boldsymbol{Q}\boldsymbol{A}\boldsymbol{Q}^T)$$

for all $\boldsymbol{Q} \in \mathcal{L}_{\text{orth}}$, indicating that the scalar-valued functions α_i are isotropic. □

Remark 15. By Lemma 12, the coefficients α_0, α_1 and α_2 in the Rivlin–Ericksen theorem are functions of the invariants of $\boldsymbol{A}$. Thus, a fully expanded version of (59) reads:

$$\begin{aligned} \boldsymbol{G}(\boldsymbol{A}) = {} & \tilde{\alpha}_0\big(\iota_1(\boldsymbol{A}), \iota_2(\boldsymbol{A}), \iota_3(\boldsymbol{A})\big)\boldsymbol{I} \\ & + \tilde{\alpha}_1\big(\iota_1(\boldsymbol{A}), \iota_2(\boldsymbol{A}), \iota_3(\boldsymbol{A})\big)\boldsymbol{A} \\ & + \tilde{\alpha}_2\big(\iota_1(\boldsymbol{A}), \iota_2(\boldsymbol{A}), \iota_3(\boldsymbol{A})\big)\boldsymbol{A}^2. \end{aligned} \tag{61}$$

17. Tensor calculus

This section presents an overview of the calculus of *scalar fields* $\phi : \mathbb{E}_3 \to \mathbb{R}$, *vector fields* $\boldsymbol{u} : \mathbb{E}_3 \to \mathcal{V}$, and *tensor fields* $\boldsymbol{A} : \mathbb{E}_3 \to \mathcal{L}$. To simplify the exposition, we assume that the function are defined everywhere in $\mathbb{E}_3$ and are differentiable as many times as needed. These assumptions are by no means vital and may be severely restricted in obvious ways.

17.1. **The gradient of a scalar field.** The *gradient* of a scalar field $\phi : \mathbb{E}_3 \to \mathbb{R}$ is the vector field $\operatorname{\mathbf{grad}} \phi : \mathbb{E}_3 \to \mathcal{V}$ with the property that its value at a point $\boldsymbol{x} \in \mathbb{E}_3$ satisfies

$$\operatorname{\mathbf{grad}} \phi(\boldsymbol{x}) \cdot \boldsymbol{q} = \frac{d}{d\epsilon} \phi(\boldsymbol{x} + \epsilon \boldsymbol{q})\Big|_{\epsilon=0}, \quad \text{for all } \boldsymbol{q} \in \mathcal{V}. \tag{62}$$

Given a constant (i.e., independent of $\boldsymbol{x}$) frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$, let us represent the point $\boldsymbol{x}$ and the vector $\boldsymbol{q}$ in terms of components in that frame, as in $\boldsymbol{x} = x_i\boldsymbol{e}_i$ and $\boldsymbol{q} = q_i\boldsymbol{e}_i$, and then introduce $\tilde{\phi} : \mathbb{R}^3 \to \mathbb{R}$ via $\tilde{\phi}(x_1, x_2, x_3) = \phi(x_p\boldsymbol{e}_p)$. Then

$$\frac{d}{d\epsilon}\phi(\boldsymbol{x} + \epsilon\boldsymbol{q}) = \frac{d}{d\epsilon}\tilde{\phi}(x_1 + \epsilon q_1, x_2 + \epsilon q_2, x_3 + \epsilon q_3) = \tilde{\phi}_{,i}(x_1 + \epsilon q_1, x_2 + \epsilon q_2, x_3 + \epsilon q_3)\, q_i,$$

where $\tilde{\phi}_{,i}$ is the partial derivative of $\tilde{\phi}$ with respect to its ith argument. Evaluating the result at $\epsilon = 0$ yields

$$\frac{d}{d\epsilon}\phi(\boldsymbol{x} + \epsilon\boldsymbol{q})\Big|_{\epsilon=0} = \tilde{\phi}_{,i}(x_1, x_2, x_3)\, q_i = \tilde{\phi}_{,i}(x_1, x_2, x_3)\, \boldsymbol{e}_i \cdot \boldsymbol{q},$$

which, in view of (62), implies that

$$\mathbf{grad}\, \phi(\boldsymbol{x}) = \tilde{\phi}_{,i}(x_1, x_2, x_3)\, \boldsymbol{e}_i. \tag{63}$$

Remark 16. The gradient defined in (62) is independent of any reference frame, therefore so is the component representation (63). The individual components, $\tilde{\phi}_{,i}(x_1, x_2, x_3)$, however, are very much dependent on the choice of the reference frame.

Remark 17. The notation $\tilde{\phi}_{,i}$ for the partial derivative of $\tilde{\phi}$ with respect to its ith argument is used quite widely within the context of tensor analysis in general, and in continuum mechanics in particular. We will adhere to that notation throughout the rest of these notes.

Remark 18. Beware that in our textbook the distinction between $\phi : \mathbb{E}_3 \to \mathbb{R}$ and $\tilde{\phi} : \mathbb{R}^3 \to \mathbb{R}$ is intentionally blurred—in page 47 this is called "a slight abuse of notation"—and $\phi_{,i}$ is used as the equivalent of $\tilde{\phi}_{,i}$.

17.2. **The gradient of a vector field.** The *gradient* of a vector field $\boldsymbol{v} : \mathbb{E}_3 \to \mathcal{V}$ is the tensor field $\mathbf{grad}\, \boldsymbol{v} : \mathbb{E}_3 \to \mathcal{L}$ with the property that its value at a point $\boldsymbol{x} \in \mathbb{E}_3$ satisfies

$$\big(\mathbf{grad}\, \boldsymbol{v}(\boldsymbol{x})\big)\boldsymbol{q} = \frac{d}{d\epsilon}\boldsymbol{v}(\boldsymbol{x} + \epsilon\boldsymbol{q})\Big|_{\epsilon=0}, \quad \text{for all } \boldsymbol{q} \in \mathcal{V}. \tag{64}$$

Given a constant (i.e., independent of $\boldsymbol{x}$) frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$, let's represent $\boldsymbol{v}$ in components, as in $\boldsymbol{v}(\boldsymbol{x}) = v_i(\boldsymbol{x})\, \boldsymbol{e}_i$. Moreover, let's represent the point $\boldsymbol{x}$ and the vector $\boldsymbol{q}$ in terms of their components in that frame, as in $\boldsymbol{x} = x_i\boldsymbol{e}_i$ and $\boldsymbol{q} = q_i\boldsymbol{e}_i$. Finally, for $i = 1, 2, 3$, let $\tilde{v}_i : \mathbb{R}^3 \to \mathbb{R}$ be defined trough $\tilde{v}_i(x_1, x_2, x_3) = v_i(x_p\boldsymbol{e}_p)$. Then

$$\frac{d}{d\epsilon}\boldsymbol{v}(\boldsymbol{x} + \epsilon\boldsymbol{q}) = \frac{d}{d\epsilon}\tilde{v}_i(x_1 + \epsilon q_1, x_2 + \epsilon q_2, x_3 + \epsilon q_3)\boldsymbol{e}_i = \tilde{v}_{i,j}(x_1 + \epsilon q_1, x_2 + \epsilon q_2, x_3 + \epsilon q_3)\, \boldsymbol{e}_i q_j,$$

where $\tilde{v}_{i,i}$ is the partial derivative of $\tilde{v}_i$ with respect to its jth argument. Evaluating the result at $\epsilon = 0$ yields

$$\frac{d}{d\epsilon}\boldsymbol{v}(\boldsymbol{x} + \epsilon\boldsymbol{q})\Big|_{\epsilon=0} = \tilde{v}_{i,j}(x_1, x_2, x_3)\, \boldsymbol{e}_i q_j = \tilde{v}_{i,i}(x_1, x_2, x_3)\, \boldsymbol{e}_i\, (\boldsymbol{e}_j \cdot \boldsymbol{q}) = \tilde{v}_{i,i}(x_1, x_2, x_3)\, (\boldsymbol{e}_i \otimes \boldsymbol{e}_j)\boldsymbol{q},$$

which, in view of (64), implies that

$$\mathbf{grad}\, \boldsymbol{v}(\boldsymbol{x}) = \tilde{v}_{i,j}(x_1, x_2, x_3)\, \boldsymbol{e}_i \otimes \boldsymbol{e}_j. \tag{65}$$

Thus, $\tilde{v}_{i,j}(x_1, x_2, x_3)$ are the components of $\mathbf{grad}\, \boldsymbol{v}(\boldsymbol{x})$ relative to the $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ frame.

17.3. **The divergence of vector and tensor fields.** The *divergence* of the vector field $\boldsymbol{v}$ is the scalar field $\operatorname{div}\boldsymbol{v} : \mathbb{E}_3 \to \mathbb{R}$ defined through

$$\operatorname{div}\boldsymbol{v} = \operatorname{tr}\mathbf{grad}\,\boldsymbol{v}. \tag{66}$$

If $\mathbf{grad}\,\boldsymbol{v}$ is expressed in terms of components relative to a frame as in (65), then (66) implies that

$$\operatorname{div}\boldsymbol{v} = \tilde{v}_{i,i}. \tag{67}$$

The divergence of the tensor field $\boldsymbol{A}$ is the vector field $\operatorname{div}\boldsymbol{A} : \mathbb{E}_3 \to \mathcal{V}$ with the property

$$(\operatorname{div}\boldsymbol{A})\cdot\boldsymbol{q} = \operatorname{div}(\boldsymbol{A}^T\boldsymbol{q}), \quad \text{for all } \boldsymbol{q} \in \mathcal{V}. \tag{68}$$

The presence of the transpose in that definition may look odd, but that is exactly what is needed for generalizing the Divergence Theorem, familiar from calculus, to tensors.

If $\boldsymbol{A}$ is expressed in components relative to a frame, as in $\boldsymbol{A} = \tilde{a}_{ij}\,\boldsymbol{e}_i \otimes \boldsymbol{e}_j$, then the components of $\operatorname{div}\boldsymbol{A}$ relative to the frame may be calculated as follows. From (21b) we have $\boldsymbol{A}^T = \tilde{a}_{ij}\,\boldsymbol{e}_j \otimes \boldsymbol{e}_i$, and therefore $\boldsymbol{A}^T\boldsymbol{q} = \tilde{a}_{ij}\,(\boldsymbol{e}_j \otimes \boldsymbol{e}_i)\boldsymbol{q} = \tilde{a}_{ij}\,(\boldsymbol{e}_i\cdot\boldsymbol{q})\,\boldsymbol{e}_j$, and therefore by (67),

$$\operatorname{div}(\boldsymbol{A}^T\boldsymbol{q}) = \tilde{a}_{ij,j}\,(\boldsymbol{e}_i\cdot\boldsymbol{q}) = (\tilde{a}_{ij,j}\,\boldsymbol{e}_i)\cdot\boldsymbol{q},$$

then in view of (68) we conclude that

$$\operatorname{div}\boldsymbol{A} = \tilde{a}_{ij,j}\,\boldsymbol{e}_i. \tag{69}$$

17.4. **The curl of a vector field.** The *curl* of the vector field $\boldsymbol{v}$ is the vector field $\mathbf{curl}\,\boldsymbol{v} : \mathbb{E}_3 \to \mathcal{V}$ defined through

$$(\mathbf{curl}\,\boldsymbol{v})\times\boldsymbol{q} = \left(\mathbf{grad}\,\boldsymbol{v} - (\mathbf{grad}\,\boldsymbol{v})^T\right)\boldsymbol{q} \quad \text{for all } \boldsymbol{q} \in \mathcal{V}. \tag{70}$$

If $\boldsymbol{v}$ is expressed in components relative to a frame as in $\boldsymbol{v} = v_i\,\boldsymbol{e}_i$, then

$$\mathbf{curl}\,\boldsymbol{v} = \epsilon_{ijk}\tilde{v}_{i,k}\,\boldsymbol{e}_j \tag{71a}$$

$$= (\tilde{v}_{3,2} - \tilde{v}_{2,3})\,\boldsymbol{e}_1 + (\tilde{v}_{1,3} - \tilde{v}_{3,1})\,\boldsymbol{e}_2 + (\tilde{v}_{2,1} - \tilde{v}_{1,2})\,\boldsymbol{e}_3. \tag{71b}$$

See page 53 of the textbook for details of the calculation.

17.5. **The Laplacian of scalar and vector fields.** The *Laplacian* of the scalar field ϕ is the scalar field $\Delta\phi : \mathbb{E}_3 \to \mathbb{R}$ defined through

$$\Delta\phi = \operatorname{div}\mathbf{grad}\,\phi. \tag{72}$$

If we express the points $\boldsymbol{x} \in \mathbb{E}_3$ in terms of components along a constant (i.e., independent of $\boldsymbol{x}$) frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ as $\boldsymbol{x} = x_i\boldsymbol{e}_i$, and let $\tilde{\phi}(x_1, x_2, x_3) = \phi(x_i\boldsymbol{e}_i)$ as we did earlier, then the Laplacian of ϕ may be expressed as

$$\Delta\phi = \tilde{\phi}_{,ii}. \tag{73}$$

See page 54 of the textbook for proof.

The *Laplacian* of the vector field $\boldsymbol{v}$ is the vector field $\Delta\boldsymbol{v} : \mathbb{E}_3 \to \mathcal{V}$ defined through

$$\Delta\boldsymbol{v} = \operatorname{div}\mathbf{grad}\,\boldsymbol{v}. \tag{74}$$

In terms of components this takes the form

$$\Delta\boldsymbol{v} = \tilde{v}_{i,jj}\,\boldsymbol{e}_i. \tag{75}$$

See page 54 of the textbook for proof.

18. Functions of tensors

Here we learn how to calculate derivatives of functions defined on $\mathcal{L}$.

18.1. **Scalar-valued functions of a tensor.**

The derivative of the scalar-valued function $\psi : \mathcal{L} \to \mathbb{R}$ is a tensor-valued function $D\psi : \mathcal{L} \to \mathcal{L}$ defined through

$$D\psi(\boldsymbol{A}) : \boldsymbol{B} = \frac{d}{d\epsilon}\psi(\boldsymbol{A} + \epsilon\boldsymbol{B})\Big|_{\epsilon=0}, \quad \text{for all } \boldsymbol{B} \in \mathcal{L}. \tag{76}$$

Let $\boldsymbol{A} = a_{ij}\boldsymbol{e}_i \otimes \boldsymbol{e}_j$ express the tensor $\boldsymbol{A}$ relative to a frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$. Then $\psi(\boldsymbol{A}) = \psi(a_{ij}\boldsymbol{e}_i \otimes \boldsymbol{e}_j)$ leads to

$$D\psi(\boldsymbol{A}) = \frac{\partial\psi(a_{ij}\boldsymbol{e}_i \otimes \boldsymbol{e}_j)}{\partial a_{ij}}\, \boldsymbol{e}_i \otimes \boldsymbol{e}_j. \tag{77}$$

See page 60 of the textbook for a proof.

Example 4. Let $\psi : \mathcal{L} \to \mathbb{R}$ be defined as $\psi(\boldsymbol{A}) = \frac{1}{2}\boldsymbol{A} : \boldsymbol{A}$. Let us show that $D\psi(\boldsymbol{A}) = \boldsymbol{A}$.

Let $\boldsymbol{A} = a_{ij}\boldsymbol{e}_i \otimes \boldsymbol{e}_j$ relative to some frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$. Then according to (27) we have $\psi(\boldsymbol{A}) = \frac{1}{2}a_{pq}a_{pq}$, and therefore

$$\frac{\partial}{\partial a_{ij}}\psi(\boldsymbol{A}) = \frac{\partial}{\partial a_{ij}}\Big(\frac{1}{2}a_{pq}a_{pq}\Big) = \frac{1}{2}\Big(\frac{\partial a_{pq}}{\partial a_{ij}}a_{pq} + a_{pq}\frac{\partial a_{pq}}{\partial a_{ij}}\Big) = \frac{\partial a_{pq}}{\partial a_{ij}}a_{pq} = \delta_{pi}\delta_{qj}a_{pq} = a_{ij}.$$

Then by (77) we have $D\psi(\boldsymbol{A}) = a_{ij}\, \boldsymbol{e}_i \otimes \boldsymbol{e}_j = \boldsymbol{A}$.

Theorem 14. *Let $\psi : \mathcal{L} \to \mathbb{R}$ be defined as $\psi(\boldsymbol{A}) = \det \boldsymbol{A}$. Then*

$$D\psi(\boldsymbol{A}) = (\det \boldsymbol{A})\, \boldsymbol{A}^{-T} \quad \textit{for all invertible } \boldsymbol{A} \in \mathcal{L}, \tag{78}$$

where $\boldsymbol{A}^{-T}$ is the transpose of the inverse of $\boldsymbol{A}$.

Proof. For any $\boldsymbol{B} \in \mathcal{L}$ and $\epsilon \in \mathbb{R}$ we have

$$\boldsymbol{A} + \epsilon\boldsymbol{B} = \epsilon\boldsymbol{A}\Big(\frac{1}{\epsilon}\boldsymbol{I} + \boldsymbol{A}^{-1}\boldsymbol{B}\Big) = -\epsilon\boldsymbol{A}\Big(-\frac{1}{\epsilon}\boldsymbol{I} - \boldsymbol{A}^{-1}\boldsymbol{B}\Big) = -\epsilon\boldsymbol{A}\Big(\lambda\boldsymbol{I} - \boldsymbol{A}^{-1}\boldsymbol{B}\Big),$$

where we have set $\lambda = -1/\epsilon$. Then by (25a) and (25b) we get

$$\det(\boldsymbol{A} + \epsilon\boldsymbol{B}) = -\epsilon^3(\det \boldsymbol{A})\det(\lambda\boldsymbol{I} - \boldsymbol{A}^{-1}\boldsymbol{B})$$

and then by (29)

$$\begin{aligned}
\psi(\boldsymbol{A} + \epsilon\boldsymbol{B}) &= -\epsilon^3(\det \boldsymbol{A})\Big(\lambda^3 - \iota_1(\boldsymbol{A}^{-1}\boldsymbol{B})\,\lambda^2 + \iota_2(\boldsymbol{A}^{-1}\boldsymbol{B})\,\lambda - \iota_3(\boldsymbol{A}^{-1}\boldsymbol{B}\Big) \\
&= (\det \boldsymbol{A})\Big(-\epsilon^3\lambda^3 + \iota_1(\boldsymbol{A}^{-1}\boldsymbol{B})\,\epsilon^3\lambda^2 - \iota_2(\boldsymbol{A}^{-1}\boldsymbol{B})\,\epsilon^3\lambda + \iota_3(\boldsymbol{A}^{-1}\boldsymbol{B}\,\epsilon^3\Big) \\
&= (\det \boldsymbol{A})\Big(1 + \iota_1(\boldsymbol{A}^{-1}\boldsymbol{B})\,\epsilon + \iota_2(\boldsymbol{A}^{-1}\boldsymbol{B})\,\epsilon^2 + \iota_3(\boldsymbol{A}^{-1}\boldsymbol{B})\,\epsilon^3\Big).
\end{aligned}$$

Then

$$\frac{d}{d\epsilon}\psi(\boldsymbol{A} + \epsilon\boldsymbol{B}) = (\det \boldsymbol{A})\Big(\iota_1(\boldsymbol{A}^{-1}\boldsymbol{B}) + \iota_2(\boldsymbol{A}^{-1}\boldsymbol{B})\,(2\epsilon) + \iota_3(\boldsymbol{A}^{-1}\boldsymbol{B})\,(3\epsilon^2)\Big),$$

whence

$$D\psi(\boldsymbol{A}) : \boldsymbol{B} = \frac{d}{d\epsilon}\psi(\boldsymbol{A} + \epsilon\boldsymbol{B})\Big|_{\epsilon=0} = (\det \boldsymbol{A})\, \iota_1(\boldsymbol{A}^{-1}\boldsymbol{B}).$$

Recalling (24a) we have $\iota_1(\boldsymbol{A}^{-1}\boldsymbol{B}) = \operatorname{tr}(\boldsymbol{A}^{-1}\boldsymbol{B})$, and recalling (26) this is the same as $\boldsymbol{A}^{-T} : \boldsymbol{B}$. We conclude that

$$D\psi(\boldsymbol{A}) : \boldsymbol{B} = (\det \boldsymbol{A})\,(\boldsymbol{A}^{-T} : \boldsymbol{B}) \quad \text{for all } \boldsymbol{B} \in \mathcal{L},$$

which is equivalent to (78) since $\boldsymbol{B}$ is arbitrary. □

Corollary 2. *Consider a time-dependent tensor $\boldsymbol{A}(t)$. We write $\dot{\boldsymbol{A}}$ for its derivative with respect to time. Then*

$$\frac{d}{dt} \det \boldsymbol{A} = (\det \boldsymbol{A})\ \operatorname{tr}(\boldsymbol{A}^{-1}\dot{\boldsymbol{A}}) = (\det \boldsymbol{A})\, \boldsymbol{A}^{-T} : \dot{\boldsymbol{A}}. \tag{79}$$

Proof. Let $\psi(\boldsymbol{A}) = \det \boldsymbol{A}$, and let $\boldsymbol{A}(t) = a_{ij}\, \boldsymbol{e}_i \otimes \boldsymbol{e}_j$ in some frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$, where $a_{ij} = a_{ij}(t)$ and $\dot{a}_{ij} = da_{ij}/dt$. Then by the chain rule

$$\begin{aligned} \frac{d}{dt} \det \boldsymbol{A} &= \frac{d}{dt}\psi(a_{ij}\, \boldsymbol{e}_i \otimes \boldsymbol{e}_j) = \frac{\partial \psi}{\partial a_{ij}} \dot{a}_{ij} = \frac{\partial \psi}{\partial a_{ij}} \delta_{ip}\delta_{jq}\dot{a}_{pq} \\ &= \frac{\partial \psi}{\partial a_{ij}} \Big((\boldsymbol{e}_i \otimes \boldsymbol{e}_j) : (\boldsymbol{e}_p \otimes \boldsymbol{e}_q) \Big) \dot{a}_{pq} = \Big(\frac{\partial \psi}{\partial a_{ij}}\, \boldsymbol{e}_i \otimes \boldsymbol{e}_j \Big) : \Big(\dot{a}_{pq}\, \boldsymbol{e}_p \otimes \boldsymbol{e}_q \Big) = D\psi(\boldsymbol{A}) : \dot{\boldsymbol{A}}. \end{aligned}$$

Substituting for $D\psi(\boldsymbol{A})$ from (78) we get

$$\frac{d}{dt} \det \boldsymbol{A} = (\det \boldsymbol{A})\, \boldsymbol{A}^{-T} : \dot{\boldsymbol{A}},$$

as asserted. □

(See page 62 of the textbook for an alternative proof.)

Remark 19. Since $\operatorname{tr}(\boldsymbol{AB}) = \operatorname{tr}(\boldsymbol{BA})$ and $\boldsymbol{A} : \boldsymbol{B} = \boldsymbol{B} : \boldsymbol{A}$, the equations (79) may be equivalently expressed as

$$\frac{d}{dt} \det \boldsymbol{A} = (\det \boldsymbol{A})\ \operatorname{tr}(\dot{\boldsymbol{A}}\boldsymbol{A}^{-1}) = (\det \boldsymbol{A})\, \dot{\boldsymbol{A}} : \boldsymbol{A}^{-T}. \tag{79-alt}$$

19. Integral Theorems

In the following, B is a "nice" domain in $\mathbb{E}_3$, and ∂B is the domain's boundary. We write dV for the infinitesimal volume element when integrating over B, and we write dA for the infinitesimal volume element when integrating over ∂B.

Here we are going to derive several forms of the Divergence Theorem starting from the basic theorem of Gauss for scalar functions defined in a Cartesian coordinate system in $\mathbb{R}^d$. Specifically, let $\phi\ := \Omega \subset R^d \to \mathbb{R}$ be a sufficiently smooth scalar field, and let $\boldsymbol{n} = \langle n_1, n_2, \ldots n_d \rangle$ be the outward unit normal vector at any point on the boundary of Ω. Then Gauss's theorem, which we state here without proof, is

$$\int_\Omega \frac{\partial \phi}{\partial x_i}\, dV = \int_{\partial\Omega} \phi n_i\, dA. \tag{80}$$

We restate this for functions $\phi\ :\ B \subset \mathbb{E}_3 \to \mathbb{R}$ as

Theorem 15 (Gauss)**.** *Let $\phi := B \subset \mathbb{E}_3 \to \mathbb{R}$ be a sufficiently smooth scalar field, and let $\boldsymbol{n}$ be the outward unit normal vector at any point on the boundary of B. Then*

$$\int_B \mathbf{grad}\,\phi\, dV = \int_{\partial B} \phi \boldsymbol{n}\, dA. \tag{81}$$

The well-known Divergence Theorem for vector fields is an immediate consequence of Gauss's theorem.

Theorem 16 (The Divergence Theorem for vector fields)**.** *For a sufficiently smooth vector field $\boldsymbol{v} : B \to \mathcal{V}$ we have*

$$\int_B \operatorname{div} \boldsymbol{v}\, dV = \int_{\partial B} \boldsymbol{v} \cdot \boldsymbol{n}\, dA, \tag{82}$$

where $\boldsymbol{n}$ is the outward unit normal to the boundary. The quantity on the right-hand side of (82) *is call the* flux *of the vector field $\boldsymbol{v}$ across the boundary.*

Proof. Recalling (67), that is, $\operatorname{div} \boldsymbol{v} = \tilde{v}_{i,i}$. Then according to (80) version of Gauss's theorem

$$\int_B \operatorname{div} \boldsymbol{v}\, dV = \int_B \tilde{v}_{i,i}\, dV = \int_{\partial B} \tilde{v}_i n_i\, dA = \int_{\partial B} \boldsymbol{v} \cdot \boldsymbol{n}\, dA,$$

as asserted. It's also possible to achieve the same result by applying the (81) version of Gauss's theorem, although the calculation would be a little bit longer. We have

$$\operatorname{div} \boldsymbol{v} = \tilde{v}_{i,i} = \tilde{v}_{i,j}\, \delta_{ij} = \tilde{v}_{i,j}\, \boldsymbol{e}_i \cdot \boldsymbol{e}_j = (\tilde{v}_{i,j}\, \boldsymbol{e}_j) \cdot \boldsymbol{e}_i = (\mathbf{grad}\, \tilde{v}_i) \cdot \boldsymbol{e}_i,$$

and therefore, by (81)

$$\int_B \operatorname{div} \boldsymbol{v}\, dV = \int_B (\mathbf{grad}\, \tilde{v}_i) \cdot \boldsymbol{e}_i\, dV = \int_{\partial B} \tilde{v}_i\, \boldsymbol{n} \cdot \boldsymbol{e}_i\, dA = \int_{\partial B} (\tilde{v}_i\, \boldsymbol{e}_i) \cdot \boldsymbol{n}\, dA = \int_{\partial B} \boldsymbol{v} \cdot \boldsymbol{n}\, dA. \qquad \square$$

Theorem 17 (The Divergence Theorem for tensor fields)**.** *For a sufficiently smooth tensor field $\boldsymbol{A} : B \to \mathcal{V}$ we have*

$$\int_B \operatorname{div} \boldsymbol{A}\, dV = \int_{\partial B} \boldsymbol{A}\boldsymbol{n}\, dA, \tag{83}$$

where $\boldsymbol{n}$ is the outward unit normal to the boundary.

Proof. Recalling the definition (68) of the divergence of a tensor, for any vector $\boldsymbol{q} \in \mathcal{V}$ we have

$$\int_B (\operatorname{div} \boldsymbol{A}) \cdot \boldsymbol{q}\, dV = \int_B \operatorname{div}(\boldsymbol{A}^T \boldsymbol{q})\, dV \overset{\text{by (82)}}{=} \int_{\partial B} \boldsymbol{A}^T \boldsymbol{q} \cdot \boldsymbol{n}\, dA \overset{\text{by (15)}}{=} \int_{\partial B} \boldsymbol{A}\boldsymbol{n} \cdot \boldsymbol{q}\, dA,$$

which is equivalent to (83) since $\boldsymbol{q}$ is arbitrary. □

Theorem 18 (The Generalized Divergence Theorem)**.** *Let $\boldsymbol{u}$ be a vector field and $\boldsymbol{A}$ a tensor field on a domain $B \subset \mathbb{E}_3$. Then we have*

$$\int_B \Big(\boldsymbol{u} \otimes \operatorname{div} \boldsymbol{A} + (\mathbf{grad}\, \boldsymbol{u})\, \boldsymbol{A}^T \Big)\, dV = \int_{\partial B} \boldsymbol{u} \otimes \boldsymbol{A}\boldsymbol{n}\, dA, \tag{84a}$$

or the equivalent transposed version:

$$\int_B \Big((\operatorname{div} \boldsymbol{A}) \otimes \boldsymbol{u} + \boldsymbol{A}\, (\mathbf{grad}\, \boldsymbol{u})^T \Big)\, dV = \int_{\partial B} (\boldsymbol{A}\boldsymbol{n}) \otimes \boldsymbol{u}\, dA. \tag{84b}$$

Proof. To prove the (84b) variant, pick any constant vector $\boldsymbol{q} \in \mathcal{V}$ and calculate

$$\int_{\partial B} (\boldsymbol{A}\boldsymbol{n} \otimes \boldsymbol{u})\,\boldsymbol{q}\,dA = \int_{\partial B} (\boldsymbol{u}\cdot\boldsymbol{q})\,\boldsymbol{A}\boldsymbol{n}\,dA \overset{\text{by (83)}}{=} \int_B \operatorname{div}\big((\boldsymbol{u}\cdot\boldsymbol{q})\,\boldsymbol{A}\big)\,dA.$$

Now, referring to equation (215) of Exercise 24, we have

$$\operatorname{div}\big((\boldsymbol{u}\cdot\boldsymbol{q})\,\boldsymbol{A}\big) = \boldsymbol{A}\,\mathbf{grad}(\boldsymbol{u}\cdot\boldsymbol{q}) + (\boldsymbol{u}\cdot\boldsymbol{q})\operatorname{div}\boldsymbol{A},$$

and then, by equation (216) of Exercise 25:

$$\mathbf{grad}(\boldsymbol{u}\cdot\boldsymbol{q}) = (\mathbf{grad}\,\boldsymbol{u})^T\boldsymbol{q} + (\mathbf{grad}\,\boldsymbol{q})^T\boldsymbol{u} = (\mathbf{grad}\,\boldsymbol{u})^T\boldsymbol{q}$$

since $\mathbf{grad}\,\boldsymbol{q} = \mathbf{0}$. Thus, we have

$$\operatorname{div}\big((\boldsymbol{u}\cdot\boldsymbol{q})\,\boldsymbol{A}\big) = \boldsymbol{A}\,(\mathbf{grad}\,\boldsymbol{u})^T\boldsymbol{q} + (\boldsymbol{u}\cdot\boldsymbol{q})\operatorname{div}\boldsymbol{A}, = \boldsymbol{A}\,(\mathbf{grad}\,\boldsymbol{u})^T\boldsymbol{q} + \big((\operatorname{div}\boldsymbol{A})\otimes\boldsymbol{u}\big)\,\boldsymbol{q}.$$

We conclude that

$$\int_{\partial B} (\boldsymbol{A}\boldsymbol{n} \otimes \boldsymbol{u})\,\boldsymbol{q}\,dA = \int_B \Big(\boldsymbol{A}\,(\mathbf{grad}\,\boldsymbol{u})^T\boldsymbol{q} + \big((\operatorname{div}\boldsymbol{A})\otimes\boldsymbol{u}\big)\,\boldsymbol{q}\Big)\,dV,$$

which is equivalent to (84b) since $\boldsymbol{q}$ is arbitrary. □

Remark 20. Let us take note of two special cases of (84a).

(1) If $\boldsymbol{u}$ is any constant vector field, then $\mathbf{grad}\,\boldsymbol{u} = \mathbf{0}$, and (84a) takes the form

$$\boldsymbol{u} \otimes \int_B \operatorname{div}\boldsymbol{A}\,dV = \boldsymbol{u} \otimes \int_{\partial B} \boldsymbol{A}\boldsymbol{n}\,dA,$$

which reduces to (83) since $\boldsymbol{u}$ is arbitrary.

(2) If $\boldsymbol{A}$ is the identity tensor, then $\operatorname{div}\boldsymbol{A} = \mathbf{0}$, and (84a) takes the form

$$\int_B \mathbf{grad}\,\boldsymbol{u}\,dV = \int_{\partial B} \boldsymbol{u}\otimes\boldsymbol{n}\,dA.$$

Forming the trace of each side and noting that $\operatorname{tr}\mathbf{grad}\,\boldsymbol{u} = \operatorname{div}\boldsymbol{u}$ and $\operatorname{tr}(\boldsymbol{u}\otimes\boldsymbol{n}) = \boldsymbol{u}\cdot\boldsymbol{n}$, this reduces to (82)

For future reference, let us summarize here the component forms of the previous three theorems.

Gauss's Theorem: $$\int_B \phi_{,j}\,dV = \int_{\partial B} \phi n_j\,dA \tag{85}$$

The Divergence Theorem for vectors: $$\int_B v_{i,j}\,dV = \int_{\partial B} v_j n_j\,dA \tag{86}$$

The Divergence Theorem for tensors: $$\int_B a_{ij,j}\,dV = \int_{\partial B} a_{ij} n_j\,dA \tag{87}$$

Theorem 19 (Stokes Theorem)**.** *Let S be a "nice" surface with a boundary curve C. Then for a sufficiently smooth vector field $\boldsymbol{v}$ defined in a neighborhood of S we have*

$$\int_S (\mathbf{curl}\,\boldsymbol{v})\cdot\boldsymbol{n}\,dA = \int_C \boldsymbol{v}\cdot\boldsymbol{t}\,ds, \tag{88}$$

where $\boldsymbol{n}$ a unit normal to S, and $\boldsymbol{t}$ is a unit tangent to C. See page 57 of the textbook for the details of how $\boldsymbol{n}$ and $\boldsymbol{t}$ are oriented.

20. Newtonian mechanics

Newton's law of motion, $\boldsymbol{f} = m\boldsymbol{a}$, relates the acceleration $\boldsymbol{a}$ of a point mass m due to the force $\boldsymbol{f}$ applied to it. In this section study the implications of Newton's law regarding the motion of system of interacting masses, and—through passing to a limit—the motion of a continuum.

> *Note:* In this section we suspend the summation convention. *There are no implicit summations over repeated indices!*

20.1. **The linear momentum.**

Consider a system of n particles P_i of masses m_i, $i = 1, 2, \dots, n$, tracked through the position vectors $\boldsymbol{r}_i$ relative to a fixed point $\boldsymbol{O}$ in $\mathbb{E}_3$.

Let $\widehat{\boldsymbol{f}}_i$ be the resultant of the forces acting on P_i. According to Newton's Law of Motion we have

$$m_i\dot{\boldsymbol{v}}_i = \widehat{\boldsymbol{f}}_i, \quad i = 1, 2, \dots, n. \tag{89}$$

The force $\widehat{\boldsymbol{f}}_i$ may be regarded as being derived from two sources:

(1) forces acting on P_i due to interaction with the other particles;
(2) forces acting on P_i by agents other that the particles P_j, $(j \neq i)$.

Forces of the first kind are called *internal forces.* Let $\boldsymbol{F}_{ij}$ be the force exerted on the particle P_i by the particle P_j. We assume that $\boldsymbol{F}_{ij}$ points in the direction P_iP_j. According to Newton's law of action/reaction, we have $\boldsymbol{F}_{ij} = -\boldsymbol{F}_{ji}$ for all i and j. In particular, $\boldsymbol{F}_{ii} = \boldsymbol{0}$ for all i.

Forces of the second kind are called *external forces.* We write $\boldsymbol{f}_i$ for the resultant of the external forces acting on P_i. Thus, Newton's law of motion (89) takes the form

$$m_i\dot{\boldsymbol{v}}_i = \boldsymbol{f}_i + \sum_{j=1}^{n} \boldsymbol{F}_{ij}, \quad i = 1, 2, \dots, n. \tag{90}$$

Summing up over all particles we get

$$\sum_{i=1}^{n} m_i\dot{\boldsymbol{v}}_i = \sum_{i=1}^{n} \boldsymbol{f}_i + \sum_{i=1}^{n}\sum_{j=1}^{n} \boldsymbol{F}_{ij},$$

but since $\boldsymbol{F}_{ij} + \boldsymbol{F}_{ji} = \boldsymbol{0}$, that double-sum is zero. Thus, the internal forces are eliminated and we are left with

$$\sum_{i=1}^{n} m_i\dot{\boldsymbol{v}}_i = \sum_{i=1}^{n} \boldsymbol{f}_i. \tag{91}$$

20.2. **The angular momentum.**

Continuing with the previous subsection's system of n particles, we take the cross product of $\boldsymbol{r}_i$ with (90)

$$\boldsymbol{r}_i \times m_i\dot{\boldsymbol{v}}_i = \boldsymbol{r}_i \times \boldsymbol{f}_i + \boldsymbol{r}_i \times \sum_{j=1}^{n} \boldsymbol{F}_{ij},$$

and sum over i:

$$\sum_{i=1}^{n} \boldsymbol{r}_i \times m_i\dot{\boldsymbol{v}}_i = \sum_{i=1}^{n} \boldsymbol{r}_i \times \boldsymbol{f}_i + \sum_{i=1}^{n}\Big(\boldsymbol{r}_i \times \sum_{j=1}^{n} \boldsymbol{F}_{ij}\Big). \tag{92}$$

The second term on the right-hand side is zero. That's the consequence of the identity

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} = \sum_{i=1}^{n}\sum_{j=i}^{n}(a_{ij}+a_{ji}) - \sum_{i=1}^{n} a_{ii},$$

which holds for any summand a_{ij}. In effect, it says that to add up all the entries of a square matrix, we

(1) add up the diagonal and below-diagonal entries;
(2) add up the diagonal and above-diagonal entries;
(3) subtract the diagonal entries because they have been double-counted.

Applying that identity with $a_{ij} = \boldsymbol{r}_i \times \boldsymbol{F}_{ij}$, we see that

$$\begin{aligned}\sum_{i=1}^{n}\boldsymbol{r}_i \times \sum_{j=1}^{n}\boldsymbol{F}_{ij} &= \sum_{i=1}^{n}\sum_{j=1}^{n}\boldsymbol{r}_i \times \boldsymbol{F}_{ij} = \sum_{i=1}^{n}\sum_{j=i}^{n}(\boldsymbol{r}_i \times \boldsymbol{F}_{ij} + \boldsymbol{r}_j \times \boldsymbol{F}_{ji}) - \sum_{i=1}^{n}\boldsymbol{r}_i \times \boldsymbol{F}_{ii} \\ &= \sum_{i=1}^{n}\sum_{j=i}^{n}(\boldsymbol{r}_i - \boldsymbol{r}_j)\times \boldsymbol{F}_{ij} - \sum_{i=1}^{n}\boldsymbol{r}_i \times \boldsymbol{F}_{ii} = \boldsymbol{0},\end{aligned}$$

because (i) $\boldsymbol{F}_{ij} = -\boldsymbol{F}_{ji}$; (ii) $\boldsymbol{F}_{ii} = \boldsymbol{0}$; and (iii) $(\boldsymbol{r}_i - \boldsymbol{r}_j)\times\boldsymbol{F}_{ij} = \boldsymbol{0}$ since $\boldsymbol{F}_{ij}$ is aligned with $\boldsymbol{r}_i - \boldsymbol{r}_j$. Thus, equation (92) reduces to

$$\sum_{i=1}^{n}\boldsymbol{r}_i \times m_i\dot{\boldsymbol{v}}_i = \sum_{i=1}^{n}\boldsymbol{r}_i \times \boldsymbol{f}_i. \tag{93}$$

20.3. **The continuum.** Consider the domain $B_t \subset \mathbb{E}_3$, parametrized by time t, as a representation of the motion and deformation of a continuum. Think of B_t as a collection of infinitesimal elements, resembling the system of n particles analyzed in the previous subsections. We saw that the internal forces drop out of the body's equations of motion, therefore here we concern ourselves with the external forces.

The external forces acting on B_t are classified into two types: *body forces* and *surface forces.* Body forces act on each infinitesimal element within the body. Typically that would be the force of gravity, measured as force per unit mass. Surface forces act on B_t's boundary, measured as force per unit area, called the *traction.* Writing $\boldsymbol{b}$ for the body forces, $\boldsymbol{t}$ for tractions, and ρ for the density, equations (91) and (93) take the form

$$\int_{B_t}\dot{\boldsymbol{v}}\rho\,dV = \int_{B_t}\boldsymbol{b}\rho\,dV + \int_{\partial B_t}\boldsymbol{t}\,dA, \tag{94}$$

$$\int_{B_t}\boldsymbol{r}\times\dot{\boldsymbol{v}}\rho\,dV = \int_{B_t}\boldsymbol{r}\times\boldsymbol{b}\rho\,dV + \int_{\partial B_t}\boldsymbol{r}\times\boldsymbol{t}\,dA, \tag{95}$$

where dV and dA are the infinitesimal volume and surface elements, respectively.

Remark 21. If all fields under consideration are bounded, then from (94) we get

$$\Big|\int_{\partial B_t}\boldsymbol{t}\,dA\Big| \le \int_{B_t}\big|\dot{\boldsymbol{v}}\rho - \boldsymbol{b}\rho\big|\,dV \le \kappa(t)\,\mathrm{vol}(B_t), \tag{96}$$

where $\kappa(t) = \max_{B_t}|\dot{\boldsymbol{v}}\rho - \boldsymbol{b}\rho|$, and $\mathrm{vol}(B_t)$ is the volume of B_t. It follows that

$$\Big|\int_{\partial B_t}\boldsymbol{t}\,dA\Big| \to 0 \quad \text{as } \mathrm{vol}(B_t)\to 0. \tag{97}$$

Moreover, by dividing (96) through by the area of B_t's boundary, area(∂B_t), we see that

$$\frac{1}{\operatorname{area}(\partial B_t)}\Big|\int_{\partial B_t} \boldsymbol{t}\,dA\Big| \le \frac{\kappa(t)\operatorname{vol}(B_t)}{\operatorname{area}(\partial B_t)},$$

If B_t represents of a family of nested self-similar domains of representative lengths δ, then vol(B_t) is proportional to δ^3 while area(∂B_t) is proportional to δ^2. We conclude that

$$\frac{1}{\operatorname{area}(\partial B_t)}\Big|\int_{\partial B_t} \boldsymbol{t}\,dA\Big| \to 0 \quad \text{as } \delta \to 0. \tag{98}$$

We will need both (97) and (98) in our proof of the existence of the Cauchy stress. In our textbook, the condition (98) is stated as an assumption, without proof, on page 80, equation (3.1).

21. The Cauchy stress tensor

Lemma 13. *Let B be the deformed state of a body at a fixed time t. For any interior point $\boldsymbol{x} \in B$, let $\boldsymbol{t}(\boldsymbol{x}, \boldsymbol{n})$ be the traction vector on an arbitrary surface patch through $\boldsymbol{x}$ having the unit normal vector $\boldsymbol{n}$. Then*

$$\boldsymbol{t}(\boldsymbol{x}, -\boldsymbol{n}) = -\boldsymbol{t}(\boldsymbol{x}, \boldsymbol{n}). \tag{99}$$

Proof. Figure 5(a) shows that patch through $\boldsymbol{x}$ as the disk D. We build a right cylinder of height on that encloses D as shown, and write S^+ and S^- for the cylinder's top and bottom surfaces, and S_h for its lateral surface. Applying (97) to the cylinder, we have

$$\int_{S^+} \boldsymbol{t}(\boldsymbol{\xi}, \boldsymbol{n})\,dA + \int_{S^-} \boldsymbol{t}(\boldsymbol{\xi}, -\boldsymbol{n})\,dA + \int_{S_h} \boldsymbol{t}(\boldsymbol{\xi}, \boldsymbol{m})\,dA \to \boldsymbol{0} \quad \text{as } h \to 0,$$

where $\boldsymbol{m}$ represents the unit outward normal at any point on the lateral surface S_h. As $h \to 0$, the integral over S_h vanishes, while the integrals over S^+ and S^- converge to integrals over the disk D, leading to

$$\int_D \Big(\boldsymbol{t}(\boldsymbol{\xi}, \boldsymbol{n}) + \boldsymbol{t}(\boldsymbol{\xi}, -\boldsymbol{n})\Big)\,dA = \boldsymbol{0}.$$

We wish to conclude that $\boldsymbol{t}(\boldsymbol{x}, \boldsymbol{n}) + \boldsymbol{t}(\boldsymbol{x}, -\boldsymbol{n}) = \boldsymbol{0}$. Indeed, if that were not the case, then $\Big(\boldsymbol{t}(\boldsymbol{\xi}, \boldsymbol{n}) + \boldsymbol{t}(\boldsymbol{\xi}, -\boldsymbol{n})\Big) \cdot \boldsymbol{a}$ would be positive in a small neighborhood of $\boldsymbol{x}$, for some vector $\boldsymbol{a}$. If D is small enough to be contained entirely within that neighborhood. then the integral above (dotted with $\boldsymbol{a}$) will evaluate to a positive quantity, contradicting the fact that it must be zero. □

Theorem 20 (The Cauchy Stress Theorem, Part 1)**.** *The traction tensor $\boldsymbol{t}(\boldsymbol{x}, \boldsymbol{n})$ of Lemma 13 is linear in $\boldsymbol{n}$, and thus, there exists a second order tensor field $\boldsymbol{S} : B \to \mathcal{L}$, so that*

$$\boldsymbol{t}(\boldsymbol{x}, \boldsymbol{n}) = \boldsymbol{S}(\boldsymbol{x})\boldsymbol{n}, \quad \textit{for all } \boldsymbol{x} \in B. \tag{100}$$

The tensor $\boldsymbol{S}(\boldsymbol{x})$ is called the Cauchy stress *at $\boldsymbol{x}$.*

Proof. Pick a point $\boldsymbol{x}$ in the interior of B and pass a surface through $\boldsymbol{x}$ with a unit normal $\boldsymbol{n}$ in an arbitrary direction. Let T_h be a triangle within that surface with edge lengths proportional to h. Construct a tetrahedron with the base T_h and the opposite vertex at some point, say $\boldsymbol{o}$, as in Figure 5(b)), whose edges are parallel to a frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$. Orient the tetrahedron so that $\boldsymbol{n} \cdot \boldsymbol{e}_i > 0$ for $i = 1, 2, 3$. Let's write $T_{h,1}, T_{h,2}, T_{h,3}$ for the tetrahedron's

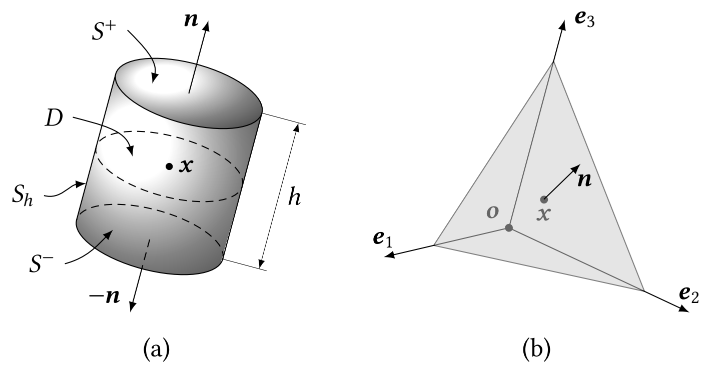


FIGURE 5. On the left, as the height h of the cylinder goes to zero, its top and bottom surfaces converge to the disk D. On the right, As the shaded triangle triangle shrinks toward $\boldsymbol{x}$, the entire tetrahedron shrinks toward $\boldsymbol{x}$.

faces that go through the vertex $\boldsymbol{o}$, and note that the outer normals to these faces are $-\boldsymbol{e}_i$, $i = 1, 2, 3$.

Applying (98) to the tetrahedron, we get

$$\frac{1}{\operatorname{area}(\partial\Delta_h)}\left(\int_{T_h} \boldsymbol{t}(\boldsymbol{n},\xi)\,dA + \sum_{i=1}^{n}\int_{T_{h,i}} \boldsymbol{t}(-\boldsymbol{e}_i,\xi)\,dA\right) \to \boldsymbol{0} \quad \text{as } h \to 0. \tag{101}$$

Let $P_i : T_h \to T_{h,i}$ be the orthogonal projection map of T_h onto $T_{h,i}$. This projects the area element dA of T_h to an elements of area $(\boldsymbol{n}\cdot\boldsymbol{e}_i)T_h$ on $T_{h,i}$, therefore an integration over the triangle $T_{h,i}$ may be performed on the triangle T_h according to

$$\int_{T_{h,i}} \boldsymbol{t}(-\boldsymbol{e}_i,\xi)\,dA = \int_{T_h} \boldsymbol{t}\big(-\boldsymbol{e}_i, P_i^{-1}(\xi)\big)\,(\boldsymbol{n}\cdot\boldsymbol{e}_i)dA,$$

and thus

$$\begin{aligned}\int_{T_h} \boldsymbol{t}(\boldsymbol{n},\xi)\,dA + \sum_{i=1}^{n}\int_{T_{h,i}} \boldsymbol{t}(-\boldsymbol{e}_i,\xi)\,dA &= \int_{T_h} \boldsymbol{t}(\boldsymbol{n},\xi)\,dA + \sum_{i=1}^{n}\int_{T_h} \boldsymbol{t}\big(-\boldsymbol{e}_i, P_i^{-1}(\xi)\big)\,(\boldsymbol{n}\cdot\boldsymbol{e}_i)\,dA,\\ &= \int_{T_h}\left(\boldsymbol{t}(\boldsymbol{n},\xi) + \sum_{i=1}^{n}(\boldsymbol{n}\cdot\boldsymbol{e}_i)\boldsymbol{t}\big(-\boldsymbol{e}_i, P_i^{-1}(\xi)\big)\right)dA.\end{aligned}$$

Furthermore, let us observe that $\operatorname{area}(T_{h,i}) = (\boldsymbol{n}\cdot\boldsymbol{e}_i)\operatorname{area}(T_h)$, and therefore

$$\operatorname{area}(\partial\Delta_h) = \operatorname{area}(T_h) + \sum_{i=1}^{n}(\boldsymbol{n}\cdot\boldsymbol{e}_i)\operatorname{area}(T_h) = \left(1 + \sum_{i=1}^{n}\boldsymbol{n}\cdot\boldsymbol{e}_i\right)\operatorname{area}(T_h) = c\operatorname{area}(T_h),$$

where we have let $c = 1 + \sum_{i=1}^{n}\boldsymbol{n}\cdot\boldsymbol{e}_i$. Multiplying (101) by the constant c we obtain

$$\begin{aligned}&\frac{c}{\operatorname{area}(\partial\Delta_h)}\left(\int_{T_h} \boldsymbol{t}(\boldsymbol{n},\xi)\,dA + \sum_{i=1}^{n}\int_{T_{h,i}} \boldsymbol{t}(-\boldsymbol{e}_i,\xi)\,dA\right)\\ &\qquad\qquad = \frac{1}{\operatorname{area}(T_h)}\int_{T_h}\left(\boldsymbol{t}(\boldsymbol{n},\xi) + \sum_{i=1}^{n}(\boldsymbol{n}\cdot\boldsymbol{e}_i)\boldsymbol{t}\big(-\boldsymbol{e}_i, P_i^{-1}(\xi)\big)\right)dA,\end{aligned}$$

which, according to (101) converges to zero as $h \to 0$. But as h goes to zero, the tetrahedron shrinks to $\boldsymbol{x}$, and therefore by the Mean Value Theorem the integral on the right-hand side converges to

$$\boldsymbol{t}(\boldsymbol{n}, \boldsymbol{x}) + \sum_{i=1}^{n} (\boldsymbol{n} \cdot \boldsymbol{e}_i) \boldsymbol{t}\big(-\boldsymbol{e}_i, P_i^{-1}(\boldsymbol{x})\big) = \boldsymbol{0}.$$

Since, $P_i^{-1}(\boldsymbol{x}) = \boldsymbol{x}$, we conclude that (resorting now to the summation convention):

$$\boldsymbol{t}(\boldsymbol{n}, \boldsymbol{x}) = -(\boldsymbol{n} \cdot \boldsymbol{e}_i)\, \boldsymbol{t}(-\boldsymbol{e}_i, \boldsymbol{x}) \overset{\text{by (99)}}{=} (\boldsymbol{n} \cdot \boldsymbol{e}_i)\, \boldsymbol{t}(\boldsymbol{e}_i, \boldsymbol{x}). = \Big(\boldsymbol{t}(\boldsymbol{e}_i, \boldsymbol{x}) \otimes \boldsymbol{e}_i\Big)\, \boldsymbol{n},$$

whereby $S(\boldsymbol{x}) = \boldsymbol{t}(\boldsymbol{e}_i, \boldsymbol{x}) \otimes \boldsymbol{e}_i$. □

Remark 22. The proof above concludes with $S(\boldsymbol{x}) = \boldsymbol{t}(\boldsymbol{e}_i, \boldsymbol{x}) \otimes \boldsymbol{e}_i$ which may give the false impression that $S(\boldsymbol{x})$ may depend on the choice of the frame. That it is not the case is evident from $\boldsymbol{t}(\boldsymbol{n}, \boldsymbol{x}) = S(\boldsymbol{x})\boldsymbol{n}$ since neither $\boldsymbol{n}$ nor $\boldsymbol{t}(\boldsymbol{n}, \boldsymbol{x})$ depend on the frame.

Theorem 21 (The Cauchy Stress Theorem, Part 2)**.** *The Cauchy tensor $S(\boldsymbol{x})$ established in Part 1 satisfies the equation*

$$\rho \dot{\boldsymbol{v}} = \operatorname{div} \boldsymbol{S} + \rho \boldsymbol{b}. \tag{102}$$

Proof. Recall the balance of linear momentum equation (94) and substitute for $\boldsymbol{t}$ from (100):

$$\int_{B_t} \dot{\boldsymbol{v}} \rho \, dV = \int_{B_t} \boldsymbol{b} \rho \, dV + \int_{\partial B_t} \boldsymbol{S}\boldsymbol{n} \, dA.$$

Apply the Divergence Theorem (83) to replace the surface integral by a volume integral

$$\int_{B_t} \dot{\boldsymbol{v}} \rho \, dV = \int_{B_t} \boldsymbol{b} \rho \, dV + \int_{B_t} \operatorname{div} \boldsymbol{S} \, dV,$$

and then collect the terms:

$$\int_{B_t} \Big(\rho \dot{\boldsymbol{v}} - \operatorname{div} \boldsymbol{S} - \rho \boldsymbol{b}\Big) \, dV = \boldsymbol{0}.$$

This holds for any domain B_t, and therefore the integrand is zero. □

Theorem 22 (The Cauchy Stress Theorem, Part 3)**.** *The Cauchy stress $S(\boldsymbol{x})$ established in Part 1 is a symmetric tensor at each $\boldsymbol{x}$, that is*

$$S(\boldsymbol{x}) = S(\boldsymbol{x})^T \quad \textit{for all } \boldsymbol{x} \in B. \tag{103}$$

Proof. Recall the balance of angular momentum equation (95). Form the cross product of that equation with an arbitrary constant vector $\boldsymbol{a}$:

$$\boldsymbol{a} \times \int_{B_t} \boldsymbol{r} \times \dot{\boldsymbol{v}} \rho \, dV = \boldsymbol{a} \times \int_{B_t} \boldsymbol{r} \times \boldsymbol{b} \rho \, dV + \boldsymbol{a} \times \int_{\partial B_t} \boldsymbol{r} \times \boldsymbol{t} \, dA,$$

then rearrange/regroup terms as

$$\int_{B_t} \rho \Big(\boldsymbol{a} \times (\boldsymbol{r} \times (\dot{\boldsymbol{v}} - \boldsymbol{b})\Big) \, dV = \int_{\partial B_t} \boldsymbol{a} \times (\boldsymbol{r} \times \boldsymbol{t}) \, dA.$$

Then apply $(210\text{b})_2$ to expand each of the integrands:

$$\int_{B_t} \rho \Big(\boldsymbol{r} \otimes (\dot{\boldsymbol{v}} - \boldsymbol{b}) - (\dot{\boldsymbol{v}} - \boldsymbol{b}) \otimes \boldsymbol{r}\Big)\, \boldsymbol{a} \, dV = \int_{\partial B_t} \Big(\boldsymbol{r} \otimes \boldsymbol{t} - \boldsymbol{t} \otimes \boldsymbol{r}\Big)\, \boldsymbol{a} \, dA.$$

This holds for all $\boldsymbol{a} \in \mathcal{V}$, therefore

$$\int_{B_t} \rho\Big(\boldsymbol{r} \otimes (\dot{\boldsymbol{v}} - \boldsymbol{b}) - (\dot{\boldsymbol{v}} - \boldsymbol{b}) \otimes \boldsymbol{r}\Big)\, dV = \int_{\partial B_t} \Big(\boldsymbol{r} \otimes \boldsymbol{t} - \boldsymbol{t} \otimes \boldsymbol{r}\Big)\, dA. \tag{104}$$

In view of (102), the left-hand side of (104), let's call it the LHS, reduces to

$$\text{LHS} = \int_{B_t} \Big(\boldsymbol{r} \otimes \operatorname{div} \boldsymbol{S} - (\operatorname{div} \boldsymbol{S}) \otimes \boldsymbol{r}\Big)\, dV.$$

As to the right-hand side of (104), we evaluate the integral of $\boldsymbol{r} \otimes \boldsymbol{t}$ by substituting $\boldsymbol{t} = \boldsymbol{S}\boldsymbol{n}$ from (100) and then applying the Generalized Divergence Theorem 84a from page 37:

$$\int_{\partial B_t} \boldsymbol{r} \otimes \boldsymbol{t}\, dA = \int_{\partial B_t} \boldsymbol{r} \otimes \boldsymbol{S}\boldsymbol{n}\, dA = \int_{B_t} \Big(\boldsymbol{r} \otimes \operatorname{div} \boldsymbol{S} + (\operatorname{\mathbf{grad}} \boldsymbol{r})\, \boldsymbol{S}^T\Big)\, dV = \int_{B_t} \Big(\boldsymbol{r} \otimes \operatorname{div} \boldsymbol{S} + \boldsymbol{S}^T\Big)\, dV,$$

where in the last step we have inserted $\operatorname{\mathbf{grad}} \boldsymbol{r} = \boldsymbol{I}$. Taking the transposes of all terms in this equation we see that

$$\int_{\partial B_t} \boldsymbol{t} \otimes \boldsymbol{r}\, dA = \int_{B_t} \Big((\operatorname{div} \boldsymbol{S}) \otimes \boldsymbol{r} + \boldsymbol{S}\Big)\, dV,$$

and therefore the right-hand side of (104), let's call it the RHS, evaluates to

$$\text{RHS} = \int_{B_t} \Big(\boldsymbol{r} \otimes \operatorname{div} \boldsymbol{S} + \boldsymbol{S}^T - (\operatorname{div} \boldsymbol{S}) \otimes \boldsymbol{r} - \boldsymbol{S}\Big)\, dV.$$

Then the equation LHS = RHS = $\mathbf{0}$ simplifies to

$$\int_{B_t} (\boldsymbol{S}^T - \boldsymbol{S})\, dV = \mathbf{0}.$$

Since B_t is arbitrary, it follows that $\boldsymbol{S}^T - \boldsymbol{S} = \mathbf{0}$. □

22. The interpretation of the Cauchy stress tensor

The Cauchy stress tensor $\boldsymbol{S}$ at a point $\boldsymbol{x}$ within a body B assigns to any unit vector $\boldsymbol{n}$ a traction vector $\boldsymbol{t} = \boldsymbol{S}\boldsymbol{n}$. The traction $\boldsymbol{t}$ measures the force per unit area acting on a plane perpendicular to $\boldsymbol{n}$ passing through the point $\boldsymbol{x}$. In this section we are concerned with a small neighborhood of $\boldsymbol{x}$ within which the tensor may be assumed to be essentially a constant.

Pick an arbitrary frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ and express $\boldsymbol{S}$ in components as $\boldsymbol{S} = \sigma_{ij}\, \boldsymbol{e}_i \otimes \boldsymbol{e}_j$.[5] From (22) on page 11 and the symmetric of $\boldsymbol{S}$ we know that $\sigma_{ij} = \boldsymbol{e}_i \cdot \boldsymbol{S}\boldsymbol{e}_j = \boldsymbol{S}\boldsymbol{e}_i \cdot \boldsymbol{e}_j$. But $\boldsymbol{S}\boldsymbol{e}_i$ is the traction on the plane perpendicular to $\boldsymbol{e}_i$ through $\boldsymbol{x}$, and the dot product of that vector with $\boldsymbol{e}_j$ is its component along $\boldsymbol{e}_j$. *We conclude that σ_{ij} is the is the $\boldsymbol{e}_j$ component of the traction acting on the plane perpendicular to $\boldsymbol{e}_i$.*

Figure 6 provides a visualization of this observation. There we see a cube positioned so that its faces are perpendicular to the vectors of the frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$. Imagine that the cube is so small that the stress $\boldsymbol{S}$ is effectively constant throughout. Then the tractions on the cube's colored faces are $\boldsymbol{S}\boldsymbol{e}_1$, $\boldsymbol{S}\boldsymbol{e}_2$, $\boldsymbol{S}\boldsymbol{e}_3$, each of which have been decomposed into vectors parallel to $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, and $\boldsymbol{e}_3$, revealing the nine components σ_{ij}. That said, we must

[5] Here we are following the tradition of using the symbols σ_{ij} for the components of $\boldsymbol{S}$.

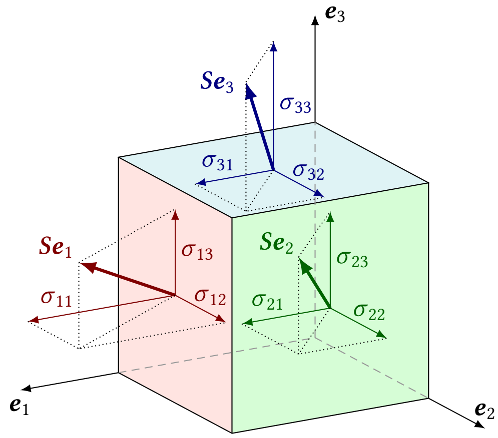


FIGURE 6. The vectors $\boldsymbol{Se}_1$, $\boldsymbol{Se}_2$, $\boldsymbol{Se}_3$ are true to-scale depictions of tractions on the faces of the cube due to the stress tensor given in (106), assuming that the $\boldsymbol{S}$ is constant within the cube.

recall that by Theorem 22, the stress tensor is symmetric, and consequently $\sigma_{ij} = \sigma_{ji}$. Therefore, in Figure 6 we have

$$\sigma_{12} = \sigma_{21}, \quad \sigma_{23} = \sigma_{32}, \quad \sigma_{31} = \sigma_{13}. \tag{105}$$

That figure is an accurate to-scale representation of the case where the components of $\boldsymbol{S}$ relative to the frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ are

$$\begin{pmatrix} 1.5 & 0.9 & 0.8 \\ 0.9 & 1.0 & 0.9 \\ 0.8 & 0.9 & 1.2 \end{pmatrix}. \tag{106}$$

Cauchy expresses the symmetry conditions (105) in his 1827 article [1] as:

> **Théorème II.** – *Si par un point quelconque d'un corps solide on mène deux axes qui se coupent à angles droits, et si l'on projette sur l'un de ces axes la pression ou tension supportée par un plan perpendiculaire à l'autre au point dont il s'agit, la projection ainsi obtenue ne variera pas quand on échangera entre eux ces mêmes axes.*
>
> English translation with help from Google Translate:
>
> **Theorem II.** – *If through any point of a solid body we pass two axes which intersect at right angles, and if we project onto one of these axes the pressure or tension supported by a plane perpendicular to the other at the point in question, the projection thus obtained will be equal when these axes are exchanged.*

Remark 23. In modern notation, the theorem quoted above says that $\boldsymbol{m} \cdot \boldsymbol{Sn} = \boldsymbol{n} \cdot \boldsymbol{Sm}$ for any two perpendicular vectors $\boldsymbol{m}$ and $\boldsymbol{n}$. Nowadays we know that $\boldsymbol{S}$ is symmetric and therefore $\boldsymbol{v} \cdot \boldsymbol{Su} = \boldsymbol{u} \cdot \boldsymbol{Sv}$ for *any* two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$. Those vectors need not be perpendicular!

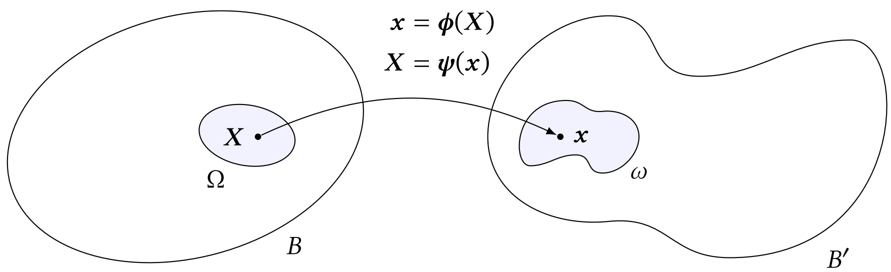


FIGURE 7. The deformation $\phi(X)$ maps the body $B \subset \mathbb{E}_3$ to $B' \subset \mathbb{E}_3$. The inverse of ϕ is ψ. A neighborhood Ω of X is mapped to a neighborhood ω of $x = \phi(X)$.

23. Kinematics

A *body* is an open domain in $\mathbb{E}_3$. A *deformation* of the body is a mapping $\phi : B \to \mathbb{E}_3$. We write X for a generic point in B, and $x = \phi(X)$ for the image of X under the mapping. We introduce the notation $B' = \phi(B)$ for the *deformed state* of B. We assume that the mapping ϕ is one-to-one, that is, distinct points of B remain distinct in B'. We also assume that ϕ is continuously differentiable, and call $F(X) = \mathbf{Grad}\,\phi(X)$ the *deformation gradient at* X. We further assume that $\det F(X) > 0$ for all $X \in B$. Finally, we call $u = x - X = \phi(X) - X$ the *displacement* of the point X. Figure 7 depicts such a deformation. ←

This comes too early; the '**Grad**' notation is defined a couple of pages later.

23.1. Measures of deformation.

Expanding ϕ into a Taylor series about a point $X_0 \in B$ we get

$$\phi(X) = \phi(X_0) + F(X_0)\,(X - X_0) + O(\|X - X_0\|^2),$$

and thus, we see that $F(X_0)$ acts as a linear operator in a small neighborhood of X_0, stretching/distorting/rotating the material about X_0.

The measure of distortion provided by F is not perfect since it's quite possible for $F(X_0)$ to be an orthogonal tensor which rotates but does not distort.

To provide true measure of distortion, we apply the Polar Decomposition Theorem (page 25) to isolate the rotation part, as in $F = QU$, where Q is orthogonal and U is symmetric and positive definite. We consider U as the true measure of the distortion. We recall from the discussion in Section 14 that $U^2 = F^T F$, and therefore calculating U calls for evaluating the square root of the tensor $F^T F$. That is somewhat a nontrivial task as it requires the knowledge of the eigenvalues and eigenvectors of $F^T F$. To save ourselves the trouble, we simply take $C = U^2$ as the measure of strain instead. This carries the same amount of information as U but is simpler to calculate as it involves merely composing F^T with F. The symmetric and positive definite tensor

$$C = F^T F, \tag{107}$$

is called the *right Cauchy–Green strain tensor*.

An alternative measure of distortion in provided by the left polar decomposition $\boldsymbol{F} = \boldsymbol{V}\boldsymbol{Q}$ (see (53)), whereby $\boldsymbol{V}^2 = \boldsymbol{F}\boldsymbol{F}^T$. This leads to the definition

$$\boldsymbol{B} = \boldsymbol{F}\boldsymbol{F}^T, \tag{108}$$

called the *left Cauchy–Green strain tensor.* The right Cauchy–Green strain tensor finds its use mostly in modeling fluids, while the left Cauchy–Green strain tensor finds its uses in modeling solids.

Remark 24. Let the deformation map $\boldsymbol{\phi}$ consist of rotation and parallel translation, as in $\boldsymbol{\phi}(\boldsymbol{X}) = \boldsymbol{c} + \boldsymbol{Q}\boldsymbol{X}$, where $\boldsymbol{x}$ is a constant vector and $\boldsymbol{Q}$ is a constant orthogonal tensor. Then $\boldsymbol{F} = \mathbf{Grad}\,\boldsymbol{\phi}(\boldsymbol{X}) = \boldsymbol{Q}$, and therefore

$$\boldsymbol{C} = \boldsymbol{F}^T\boldsymbol{F} = \boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I}.$$

$\boldsymbol{C}$, being an identity map, induces no distortion, as expected. We see that $\boldsymbol{B} = \boldsymbol{I}$ for the same reason.

23.2. **Motion.** The *motion* of a body B is a continuous deformation $\boldsymbol{\phi}_t$ of B parametrized by the time t. We call B the body's *reference configuration*, and $B_t = \boldsymbol{\phi}_t(B)$ its *current configuration.* We assume that $\boldsymbol{\phi}_0$ is the identity map, that is, $\boldsymbol{\phi}_0(\boldsymbol{X}) = \boldsymbol{X}$ for all $\boldsymbol{X} \in B$. We write $\boldsymbol{\psi}_t = \boldsymbol{\phi}_t^{-1}$. Thus, if $\boldsymbol{x} = \boldsymbol{\phi}_t(\boldsymbol{X})$, then $\boldsymbol{X} = \boldsymbol{\psi}_t(\boldsymbol{x})$. Also

$$\boldsymbol{X} = \boldsymbol{\psi}_t(\boldsymbol{x}) = \boldsymbol{\psi}_t\big(\boldsymbol{\phi}_t(\boldsymbol{X})\big), \quad \boldsymbol{x} = \boldsymbol{\phi}_t(\boldsymbol{X}) = \boldsymbol{\phi}_t\big(\boldsymbol{\psi}_t(\boldsymbol{x})\big).$$

23.3. **Material and spatial fields.** In view of the one-to-one mapping between B and B_t, any time-dependent scalar-, vector- or tensor-valued field $\Omega(\boldsymbol{X}, t)$ defined on B, induces a time-dependent field $\omega(\boldsymbol{x}, t)$ on B_t through

$$\omega(\boldsymbol{x}, t) = \Omega(\boldsymbol{X}, t), \quad \text{where } \boldsymbol{x} = \boldsymbol{\phi}_t(\boldsymbol{X}). \tag{109}$$

A field expressed in terms of the points $\boldsymbol{X}$ of B, such as $\Omega(\boldsymbol{X}, t)$, is called a *material field.* A field expressed in terms of the points $\boldsymbol{x}$ of B_t, such as $\omega(\boldsymbol{x}, t)$, is called a *spatial field.*

Instead of using distinct symbols such Ω and ω for the corresponding pairs of material and special descriptions of a field, we will often find it more economical to use a single symbol with the subscripts m for "material" and s for "special". Thus, given the material field Ω as above, the corresponding spatial field may be expressed as

$$\Omega_s(\boldsymbol{x}, t) = \Omega(\boldsymbol{X}, t), \quad \text{where } \boldsymbol{x} = \boldsymbol{\phi}_t(\boldsymbol{X}), \tag{110}$$

that is,

$$\Omega_s\big(\boldsymbol{\phi}_t(\boldsymbol{X}), t\big) = \Omega(\boldsymbol{X}, t), \quad \Omega_s(\boldsymbol{x}, t) = \Omega\big(\boldsymbol{\psi}_t(\boldsymbol{x}), t\big), \tag{(110-alt)}$$

Similarly, given the spatial field ω, the corresponding material field may be expressed as

$$\omega_m(\boldsymbol{X}, t) = \omega(\boldsymbol{x}, t), \quad \text{where } \boldsymbol{x} = \boldsymbol{\phi}_t(\boldsymbol{X}), \tag{111}$$

that is,

$$\omega_m\big(\boldsymbol{\psi}_t(\boldsymbol{x}), t\big) = \omega(\boldsymbol{x}, t), \quad \omega_m(\boldsymbol{X}, t) = \omega\big(\boldsymbol{\phi}_t(\boldsymbol{X}), t\big). \tag{(111-alt)}$$

Following [7] and [2], we write **Grad**, Div, **Curl** for the gradient, divergence, and curl of material fields, and **grad**, div, **curl** for the gradient, divergence, and curl of spatial fields. In [5], the '**Grad**', 'Div', and '**Curl**' are written

$$\nabla^{\boldsymbol{X}} \qquad \nabla^{\boldsymbol{X}} \cdot \qquad \nabla^{\boldsymbol{X}} \times$$

while '**grad**', 'div', and '**curl**' are written

$$\nabla^{\boldsymbol{x}} \qquad \nabla^{\boldsymbol{x}}\cdot \qquad \nabla^{\boldsymbol{x}}\times$$

Consider a material field $\Omega(\boldsymbol{X}, t)$ defined on the body B. For our current purposes, Ω make be scalar-, vector-, or tensor-valued. We write $\dot{\Omega}$ for the rate of change of Ω with respect to time. That is

$$\dot{\Omega}(\boldsymbol{X}, t) = \frac{\partial}{\partial t}\Omega(\boldsymbol{X}, t).$$

That holds no surprises. On the other hand, consider a spatial field $\omega(\boldsymbol{x}, t)$. What is its time derivative? The point $\boldsymbol{x}$ moves with time, therefore it needs to be differentiated too. We define the *total time derivative* of $\boldsymbol{\omega}$ as

$$\dot{\omega}(\boldsymbol{x}, t)\Big|_{\boldsymbol{x}=\boldsymbol{\phi}_t(\boldsymbol{X})} = \frac{\partial}{\partial t}\Big(\omega\big(\boldsymbol{\phi}_t(\boldsymbol{X}), t\big)\Big), \tag{112}$$

or equivalently, as

$$\dot{\omega}(\boldsymbol{x}, t) = \left[\frac{\partial}{\partial t}\Big(\omega\big(\boldsymbol{\phi}_t(\boldsymbol{X}), t\big)\Big)\right]_{\boldsymbol{X}=\boldsymbol{\psi}_t(\boldsymbol{x})}. \tag{(112-alt)}$$

As $\boldsymbol{\phi}_t(\boldsymbol{X})$ tracks the motion of the point that originates at $\boldsymbol{X}$, $\dot{\omega}$ measures the rate of change of ω as viewed by an observer who moves along with that point.

23.4. **Velocity and acceleration.** The point $\boldsymbol{X}$ of the body B undergoing the motion $\boldsymbol{\phi}_t$ is located at $\boldsymbol{x} = \boldsymbol{\phi}_t(\boldsymbol{X}) \in \mathbb{E}_3$ at time t. The *velocity* $\boldsymbol{V}(\boldsymbol{X}, t)$, and the *acceleration* $\boldsymbol{A}(\boldsymbol{X}, t)$ of that point are given by

$$\boldsymbol{V}(\boldsymbol{X}, t) = \frac{\partial}{\partial t}\boldsymbol{\phi}_t(\boldsymbol{X}), \quad \boldsymbol{A}(\boldsymbol{X}, t) = \frac{\partial^2}{\partial t^2}\boldsymbol{\phi}_t(\boldsymbol{X}), \tag{113}$$

These are expressed as material fields. The equivalent expression as spatial fields are

$$\boldsymbol{v}(\boldsymbol{x}, t) = \boldsymbol{V}(\boldsymbol{X}, t)\Big|_{\boldsymbol{X}=\boldsymbol{\psi}_t(\boldsymbol{x})} = \frac{\partial}{\partial t}\boldsymbol{\phi}_t(\boldsymbol{X})\Big|_{\boldsymbol{X}=\boldsymbol{\psi}_t(\boldsymbol{x})}, \tag{114a}$$

$$\boldsymbol{a}(\boldsymbol{x}, t) = \boldsymbol{A}(\boldsymbol{X}, t)\Big|_{\boldsymbol{X}=\boldsymbol{\psi}_t(\boldsymbol{x})} = \frac{\partial^2}{\partial t^2}\boldsymbol{\phi}_t(\boldsymbol{X})\Big|_{\boldsymbol{X}=\boldsymbol{\psi}_t(\boldsymbol{x})}. \tag{114b}$$

23.5. **Time derivatives.**

Theorem 23. *Let $\boldsymbol{\phi}_t$ be a motion of the body B, and let $\boldsymbol{v}(\boldsymbol{x}, t)$ be the spatial description of its velocity field. Consider an arbitrary scalar-valued spatial field $\omega(\boldsymbol{x}, t)$ defined on B_t. Then the total time derivative $\dot{\omega}$ of ω is given by*

$$\dot{\omega} = \frac{\partial \omega}{\partial t} + (\mathbf{grad}\,\omega)\cdot\boldsymbol{v}. \tag{115}$$

Proof. This is an immediate consequence of applying the chain rule of differentiation to the definition of the total derivative in (112), and the definition of the velocity in (114a):

$$\dot{\omega}(\boldsymbol{x}, t) = \left[\frac{\partial}{\partial t}\Big(\omega\big(\boldsymbol{\phi}_t(\boldsymbol{X}), t\big)\Big)\right]_{\boldsymbol{X}=\boldsymbol{\psi}_t(\boldsymbol{x})} = \mathbf{grad}\,\omega(\boldsymbol{x}, t)\cdot\frac{\partial}{\partial t}\boldsymbol{\phi}_t(\boldsymbol{X})\Big|_{\boldsymbol{X}=\boldsymbol{\psi}_t(\boldsymbol{x})} + \frac{\partial \omega}{\partial t}(\boldsymbol{x}, t). \qquad \square$$

Corollary 3. *Let $\boldsymbol{\phi}_t$ be a motion of the body B, and let $\boldsymbol{v}(\boldsymbol{x}, t)$ be the spatial description of its velocity field. Consider an arbitrary vector-valued spatial field $\mathbf{w}(\boldsymbol{x}, t)$ defined on B_t. Then the total time derivative $\dot{\mathbf{w}}$ of $\mathbf{w}$ is given by*

$$\dot{\boldsymbol{w}} = \frac{\partial \boldsymbol{w}}{\partial t} + (\mathbf{grad}\, \boldsymbol{w})\, \boldsymbol{v}. \tag{116}$$

In particular, the total time derivative of the the velocity field is given by

$$\dot{\boldsymbol{v}} = \frac{\partial \boldsymbol{v}}{\partial t} + (\mathbf{grad}\, \boldsymbol{v})\, \boldsymbol{v}. \tag{117}$$

Proof (version 1, without indices). Pick an arbitrary constant vector $\boldsymbol{a} \in \mathcal{V}$. Applying the result of Theorem 23 to the scalar field $\boldsymbol{w} \cdot \boldsymbol{a}$ we get

$$(\boldsymbol{w} \cdot \boldsymbol{a})^{\bullet} = \frac{\partial (\boldsymbol{w} \cdot \boldsymbol{a})}{\partial t} + \big(\mathbf{grad}(\boldsymbol{w} \cdot \boldsymbol{a})\big) \cdot \mathbf{grad}\, \boldsymbol{v}$$

From the definition of the total time derivative in ((112-alt) we see that $(\boldsymbol{w} \cdot \boldsymbol{a})^{\bullet} = \dot{\boldsymbol{w}} \cdot \boldsymbol{a}$. Furthermore, from the formula (216) of Exercise 25 and the fact that $\mathbf{grad}\, \boldsymbol{a} = \mathbf{0}$ we obtain $\mathbf{grad}(\boldsymbol{w} \cdot \boldsymbol{a}) = (\mathbf{grad}\, \boldsymbol{w})^T \boldsymbol{a}$, and thus we arrive at

$$\dot{\boldsymbol{w}} \cdot \boldsymbol{a} = \frac{\partial (\boldsymbol{w})}{\partial t} \cdot \boldsymbol{a} + \big((\mathbf{grad}\, \boldsymbol{w})^T \boldsymbol{a}\big) \cdot \boldsymbol{v} \overset{\text{by (15)}}{=} \frac{\partial (\boldsymbol{w})}{\partial t} \cdot \boldsymbol{a} + \big((\mathbf{grad}\, \boldsymbol{w})\big) \boldsymbol{v} \cdot \boldsymbol{a}.$$

Assertion (116) follows since this holds for all $\boldsymbol{a} \in \mathcal{V}$. □

Proof (version 2, with indices). Let $\boldsymbol{w} = w_i\, \boldsymbol{e}_i$ in some arbitrary but constant frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$. Expressing (115) in terms of components along that frame $\dot{\omega} = \frac{\partial \omega}{\partial t} + \omega_{,j} v_j$, and applying this to the component w_i, we get

$$\dot{w}_i = \frac{\partial w_i}{\partial t} + w_{i,j} v_j,$$

and therefore

$$\dot{w}_i\, \boldsymbol{e}_i = \frac{\partial w_i\, \boldsymbol{e}_i}{\partial t} + w_{i,j} v_j\, \boldsymbol{e}_i.$$

We conclude the proof by observing that

$$\begin{aligned} w_{i,j} v_j\, \boldsymbol{e}_i &= w_{i,j} v_k \delta_{kj}\, \boldsymbol{e}_i = w_{i,j} v_k\, (\boldsymbol{e}_k \cdot \boldsymbol{e}_j)\, \boldsymbol{e}_i \\ &= w_{i,j} v_k\, (\boldsymbol{e}_i \otimes \boldsymbol{e}_j)\, \boldsymbol{e}_k = \big(w_{i,j} (\boldsymbol{e}_i \otimes \boldsymbol{e}_j)\big)(v_k\, \boldsymbol{e}_k) = (\mathbf{grad}\, \boldsymbol{w})\, \boldsymbol{v}, \end{aligned}$$

where we have made use of (65). □

23.6. **The transformation of volume integrals.** Considering that the mapping $\boldsymbol{\phi}_t$ of the reference configuration B to the deformed configuration B_t is one-to-one, any integration over B may be expressed as an integration over B_t, and vice versa. We write dV_X and dV_x for the differential volume elements in B and B_t, respectively. In terms of an arbitrary frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$, dV_X is the volume of the parallelepiped formed by the vectors the three vectors $d\boldsymbol{X}_1 = dX_1\, \boldsymbol{e}1$, $d\boldsymbol{X}_2 = dX_2\, \boldsymbol{e}2$, and $d\boldsymbol{X}_3 = dX_3\, \boldsymbol{e}3$, which (see Figure 1 on page 4) is given by the scalar triple product

$$dV_X = [d\boldsymbol{X}_1, d\boldsymbol{X}_2, d\boldsymbol{X}_3].$$

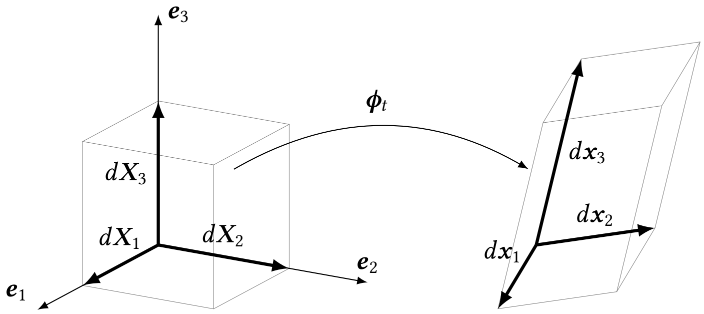


FIGURE 8. The motion ϕ_t takes the infinitesimal volume dV_X element spanned by the vectors dX_1, dX_2, dX_3 to the infinitesimal volume dV_x element spanned by the vectors dx_1, dx_2, dx_3.

Under the motion ϕ_t, the three vectors dX_1, dX_2, dX_3 are mapped to $dx_1 = F(X,t)\,dX_1$, $dx_2 = F(X,t)\,dX_2$, $dx_3 = F(X,t)\,dX_3$, where $F(X,t) = \mathbf{grad}\,\phi_t(X)$ as visualized in Figure 8. The volume of the deformed element it

$$dV_x = \big[F(X,t)\,dX_1, F(X,t)\,dX_2, F(X,t)\,dX_3\big] = \iota_3\big(F(X,t)\big)\,[dX_1, dX_2, dX_3]$$

according to (16a). We conclude that $dV_x = \iota_3\big(F(X,t)\big)\,dV_X = \det F(X,t)\,dV_X$ due to the definition of the determinant in (24c). This observation leads to

Theorem 24 (Transformation of volume integrals)**.** *Consider the motion ϕ_t of the body B, and let $B_t = \phi_t(B)$ as usual, and let $\omega(x,t)$ be any scalar field defined on B_t. For any subdomain $\Omega \subset B$, let $\Omega_t = \phi_t(\Omega)$. Then we have*

$$\int_{\Omega_t} \omega(x,t)\,dV_x = \int_{\Omega} \omega\big(\phi_t(X),t\big)\,\det F(X,t)\,dV_X. \tag{118}$$

Remark 25. The scalar field $\det F(X,t)$ that enters the previous considerations is called the *Jacobian* of the deformation ϕ_t. The equation $dV_x = \det F(X,t)\,dV_X$ indicates that the Jacobian provides a measure of the *local change of volume.* Thus, $\det F(X,t) > 1$ indicates a local expansion (dilation) of volume, while $\det F(X,t) < 1$ indicates a local contraction (compression) of volume. If the material is incompressible, then $\det F = 1$ throughout. In all cases, $\det F > 0$ as postulated in this section's opening paragraph.

23.7. **The time derivative of an integral over a moving domain.** We would be interested in calculating the derivative with respect to time of the integral on the left-hand side of equation (118). That would amount to calculating the time derivative of its right-hand side, and that would necessitate calculating the derivative of the Jacobian, $\det F(X,t)$. We break up the calculation into the following two lemmas.

Lemma 14. *Consider the motion ϕ_t of the body B. Let $v(x,t)$ be the spatial description of its velocity field, and let $F(X,t)$ be the material description of the deformation gradient. Then*

$$\frac{\partial}{\partial t} F(X,t) = \mathbf{grad}\,v(x,t)\Big|_{x=\phi_t(X)} F(X,t). \tag{119}$$

Proof. Let us calculate

$$\frac{\partial}{\partial t}\boldsymbol{F}(\boldsymbol{X},t) = \frac{\partial}{\partial t}\,\mathbf{Grad}\,\boldsymbol{\phi}_t(\boldsymbol{X}) = \mathbf{Grad}\,\frac{\partial}{\partial t}\boldsymbol{\phi}_t(\boldsymbol{X}) = \mathbf{Grad}\,\boldsymbol{V}(\boldsymbol{X},t)$$
$$= \mathbf{Grad}\Big(\boldsymbol{v}\big(\boldsymbol{\phi}_t(\boldsymbol{X}),t\big)\Big) = \mathbf{grad}\,\boldsymbol{v}(\boldsymbol{x},t)\Big|_{\boldsymbol{x}=\boldsymbol{\phi}_t(\boldsymbol{X})}\,\mathbf{Grad}\,\boldsymbol{\phi}_t(\boldsymbol{X}),$$

which is equivalent to (119) since $\mathbf{Grad}\,\boldsymbol{\phi}_t(\boldsymbol{X}) = \boldsymbol{F}(\boldsymbol{X},t)$. □

Lemma 15. *Under the assumptions and notation of of the previous lemma, we have*

$$\frac{\partial}{\partial t}\big(\det \boldsymbol{F}(\boldsymbol{X},t)\big) = \det \boldsymbol{F}(\boldsymbol{X},t)\;\operatorname{div}\boldsymbol{v}(\boldsymbol{x},t)\Big|_{\boldsymbol{x}=\boldsymbol{\phi}_t(\boldsymbol{X})}. \tag{120}$$

Proof. The formula (79-alt) on page 36 for the differentiation of a determinant tells us

$$\frac{\partial}{\partial t}\det \boldsymbol{F}(\boldsymbol{X},t) = \big(\det \boldsymbol{F}(\boldsymbol{X},t)\big)\operatorname{tr}\left(\Big(\frac{\partial}{\partial t}\boldsymbol{F}(\boldsymbol{X},t)\Big)\boldsymbol{F}^{-1}(\boldsymbol{X},t)\right).$$

But according to (119)

$$\Big(\frac{\partial}{\partial t}\boldsymbol{F}(\boldsymbol{X},t)\Big)\boldsymbol{F}^{-1}(\boldsymbol{X},t) = \mathbf{grad}\,\boldsymbol{v}(\boldsymbol{x},t)\Big|_{\boldsymbol{x}=\boldsymbol{\phi}_t(\boldsymbol{X})},$$

and therefore

$$\frac{\partial}{\partial t}\det \boldsymbol{F}(\boldsymbol{X},t) = \big(\det \boldsymbol{F}(\boldsymbol{X},t)\big)\operatorname{tr}\Big(\mathbf{grad}\,\boldsymbol{v}(\boldsymbol{x},t)\Big|_{\boldsymbol{x}=\boldsymbol{\phi}_t(\boldsymbol{X})}\Big).$$

Recalling the definition of the divergence in (66) leads to the desired result. □

The above lemma, along with the transformation Theorem 24, lead to *Reynolds Transport Theorem* which is an indispensable tool for applying laws of physics to continua.

Theorem 25 (Reynolds Transport Theorem)**.** *Consider the motion $\boldsymbol{\phi}_t$ of the body B, and let $\boldsymbol{v}(\boldsymbol{x},t)$ be the spatial description of its velocity field, and $\omega(\boldsymbol{x},t)$ be any scalar-valued spatial field defined on $B_t = \boldsymbol{\phi}_t(B)$. Then, for any subdomain $\Omega \subset B$ we have*

$$\frac{d}{dt}\int_{\Omega_t}\omega(\boldsymbol{x},t)\,dV_{\boldsymbol{x}} = \int_{\Omega_t}\Big(\frac{\partial\omega}{\partial t} + (\mathbf{grad}\,\omega)\cdot\boldsymbol{v} + \omega\operatorname{div}\boldsymbol{v}\Big)\,dV_{\boldsymbol{x}} \tag{121a}$$
$$= \int_{\Omega_t}\Big(\frac{\partial\omega}{\partial t} + \operatorname{div}(\omega\boldsymbol{v})\Big)\,dV_{\boldsymbol{x}} \tag{121b}$$
$$= \int_{\Omega_t}\frac{\partial\omega}{\partial t}\,dV_{\boldsymbol{x}} + \int_{\partial\Omega_t}\omega\,\boldsymbol{v}\cdot\boldsymbol{n}\,dA_{\boldsymbol{x}} \tag{121c}$$
$$= \int_{\Omega_t}(\dot{\omega} + \omega\operatorname{div}\boldsymbol{v})\,dV_{\boldsymbol{x}}, \tag{121d}$$

where $\Omega_t = \boldsymbol{\phi}_t(\Omega)$, $\partial\Omega_t$ is Ω_t's boundary, $\boldsymbol{n}$ is the outward unit normal to $\partial\Omega_t$, and $\dot{\omega}$ is the total time derivative of ω.

Proof. By the theorem of transformation of volume integrals (24) on page 50, we have

$$\frac{d}{dt}\int_{\Omega_t}\omega(\boldsymbol{x},t)\,dV_{\boldsymbol{x}} = \frac{d}{dt}\int_{\Omega}\omega\big(\boldsymbol{\phi}_t(\boldsymbol{X}),t\big)\;\det \boldsymbol{F}(\boldsymbol{X},t)\,dV_{\boldsymbol{X}}$$
$$= \int_{\Omega}\left[\frac{\partial}{\partial t}\Big(\omega\big(\boldsymbol{\phi}_t(\boldsymbol{X}),t\big)\Big)\;\det \boldsymbol{F}(\boldsymbol{X},t) + \omega\big(\boldsymbol{\phi}_t(\boldsymbol{X}),t\big)\frac{\partial}{\partial t}\det \boldsymbol{F}(\boldsymbol{X},t)\right]dV_{\boldsymbol{X}}.$$

Let us expand the individual terms within the square brackets. By the chain rule of differentiation we have:

$$\frac{\partial}{\partial t}\Big(\omega\big(\boldsymbol{\phi}_t(\boldsymbol{X}),t\big)\Big) = \mathbf{grad}\,\omega\Big|_{\left(\boldsymbol{\phi}_t(\boldsymbol{X}),t\right)} \cdot \frac{\partial}{\partial t}\boldsymbol{\phi}_t(\boldsymbol{X}) + \frac{\partial\omega}{\partial t}\Big|_{\left(\boldsymbol{\phi}_t(\boldsymbol{X}),t\right)}$$
$$= \mathbf{grad}\,\omega\Big|_{\left(\boldsymbol{\phi}_t(\boldsymbol{X}),t\right)} \cdot \boldsymbol{V}(\boldsymbol{X},t) + \frac{\partial\omega}{\partial t}\Big|_{\left(\boldsymbol{\phi}_t(\boldsymbol{X}),t\right)} = \Big((\mathbf{grad}\,\omega)\cdot\boldsymbol{v} + \frac{\partial\omega}{\partial t}\Big)\Big|_{\left(\boldsymbol{\phi}_t(\boldsymbol{X}),t\right)},$$

and by (120) we have

$$\omega\big(\boldsymbol{\phi}_t(\boldsymbol{X}),t\big)\frac{\partial}{\partial t}\det\boldsymbol{F}(\boldsymbol{X},t) = \omega\big(\boldsymbol{\phi}_t(\boldsymbol{X}),t\big)\det\boldsymbol{F}(\boldsymbol{X},t)\operatorname{div}\boldsymbol{v}\Big|_{\left(\boldsymbol{\phi}_t(\boldsymbol{X}),t\right)}$$
$$= (\omega\operatorname{div}\boldsymbol{v})\Big|_{\left(\boldsymbol{\phi}_t(\boldsymbol{X}),t\right)}\det\boldsymbol{F}(\boldsymbol{X},t).$$

We conclude that

$$\frac{d}{dt}\int_{\Omega_t}\omega(\boldsymbol{x},t)\,dV_{\boldsymbol{x}} = \int_{\Omega}\Big((\mathbf{grad}\,\omega)\cdot\boldsymbol{v} + \frac{\partial\omega}{\partial t} + \omega\operatorname{div}\boldsymbol{v}\Big)\Big|_{\left(\boldsymbol{\phi}_t(\boldsymbol{X}),t\right)}\det\boldsymbol{F}(\boldsymbol{X},t)\,dV_{\boldsymbol{X}},$$

which, by the theorem of transformation of volume, reduces to (121a). The variant (121b) is obtained by applying the identity (212) of Exercise 21 to (121a). The variant (121c) is obtained by applying the Divergence Theorem to (121b). The variant (121d) is obtained from (121a) and the property of the total derivative established in Theorem 23 on page 48. ☐

Remark 26. The representation (121c) of Reynolds Transport Theorem is a generalization to three dimensions of the one-dimensional Leibniz' differentiation formula:

$$\frac{d}{dt}\int_{a(t)}^{b(t)} f(x,t)\,dx = \int_{a(t)}^{b(t)}\frac{\partial}{\partial t}f(x,t)\,dx + f\big(b(t),t\big)b'(t) - f\big(a(t),t\big)a'(t).$$

We say the motion $\boldsymbol{\phi}_t$ of a body B is *volume-preserving* or *isochoric* if the volume of any part $\Omega\subset B$ remains unchanged in the course of the motion. In that regard we have:

Corollary 4. *Consider the motion $\boldsymbol{\phi}_t$ of the body B. $\boldsymbol{F}(\boldsymbol{X},t)$ be the material representation of the deformation gradient, and $\boldsymbol{v}(\boldsymbol{x},t)$ be the spatial description of the velocity field. If the motion isochoric, then* $\det\boldsymbol{F}(\boldsymbol{X},t)=1$ *and* $\operatorname{div}\boldsymbol{v}(\boldsymbol{x},t)=0$ *for all $\boldsymbol{X}\in B$, $\boldsymbol{x}\in B_t$, and all $t\geq 0$.*

Proof. In the transformation of volume integral formula (118), take $\omega(\boldsymbol{x},t)\equiv 1$. Then

$$\operatorname{vol}(\Omega) = \operatorname{vol}(\Omega_t) = \int_{\Omega_t} 1\,dV_{\boldsymbol{x}} = \int_{\Omega}\det\boldsymbol{F}(\boldsymbol{X},t)\,dV_{\boldsymbol{X}}.$$

Thus, $\int_\Omega \det\boldsymbol{F}(\boldsymbol{X},t)\,dV_{\boldsymbol{X}} = \operatorname{vol}(\Omega)$ for all Ω, and therefore $\det\boldsymbol{F}(\boldsymbol{X},t)$ is identically equal to 1.

As to the Corollary's second assertion, take $\omega(\boldsymbol{x},t)\equiv 1$ in Reynolds Transport Theorem. Then from (121c) for any $\Omega\subset B$ we get

$$0 = \frac{d}{dt}\operatorname{vol}(\Omega_t) = \frac{d}{dt}\int_{\Omega_t}1\,dV_{\boldsymbol{x}} = \int_{\Omega_t}\operatorname{div}\boldsymbol{v}\,dV_{\boldsymbol{x}}.$$

Since Ω is arbitrary, we must have $\operatorname{div}\boldsymbol{v}=0$. ☐

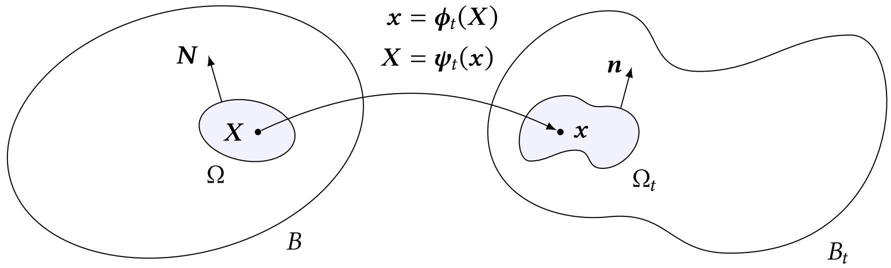


FIGURE 9. The deformation $\phi_t(X)$ maps the body $B \subset \mathbb{E}_3$ to $B_t \subset \mathbb{E}_3$ at time t. The inverse of ϕ_t at time t is ψ_t. An arbitrary part Ω of B is mapped to Ω_t.

23.8. **The transformation of surface integrals.**

Theorem 26. *Consider the motion ϕ_t of the body B, and let $F(X,t)$ be the corresponding deformation gradient. Let $\omega(x,t)$, $w(x,t)$, $A(x,t)$ be spacial descriptions of arbitrary scalar, vector, and tensor fields on B_t. For any subdomain $\Omega \subset B$, let $N(X)$ and $n(x)$ be the outward unit normals to the boundaries $\partial\Omega$ and $\partial\Omega_t$, respectively. Then*

$$\int_{\partial\Omega_t} \omega(x,t)\, n(x)\, dA_x = \int_{\partial\Omega} \omega\big(\phi_t(X),t\big)\, G(X,t)\, N(X)\, dA_X, \tag{122a}$$

$$\int_{\partial\Omega_t} w(x,t)\cdot n(x)\, dA_x = \int_{\partial\Omega} w\big(\phi_t(X),t\big)\cdot \big(G(X,t)\, N(X)\big)\, dA_X, \tag{122b}$$

$$\int_{\partial\Omega_t} A(x,t)\, n(x)\, dA_x = \int_{\partial\Omega} A\big(\phi_t(X),t\big)\, G(X,t)\, N(X)\, dA_X, \tag{122c}$$

where $G(X,t)$ is the second order tensor $G(X,t) = \big(\det F(X,t)\big)\, F^{-T}(X,t)$. Figure 9 provides a visualization aid.

Proof. Consider a patch $\mathcal{P}$ of the boundary of $\partial\Omega$ parametrized as $x = \chi(\xi_1,\xi_2)$, where ξ_1 and ξ_2 are Cartesian coordinates in the parameter space. The infinitesimal rectangle with sides $d\xi_1$ and $d\xi_2$ (therefore area $d\xi_1 d\xi_2$) in the parameter space is mapped to a parallelogram of area dA_X in a tangent plane of $\partial\Omega$ defined by the vectors $\dfrac{\partial\chi}{\partial\xi_1}\,d\xi_1$. and $\dfrac{\partial\chi}{\partial\xi_2}\,d\xi_2$. The cross product $\dfrac{\partial\chi}{\partial\xi_1}\,d\xi_1 \times \dfrac{\partial\chi}{\partial\xi_2}\,d\xi_2$ equals, in magnitude, the parallelogram's area and points along the unit normal N to the boundary. Thus:

$$N\, dA_X = \frac{\partial\chi}{\partial\xi_1}\,d\xi_1 \times \frac{\partial\chi}{\partial\xi_2}\,d\xi_2 = \Big(\frac{\partial\chi}{\partial\xi_1} \times \frac{\partial\chi}{\partial\xi_2}\Big)\, d\xi_1\, d\xi_2. \tag{123}$$

The parametrization in terms of (ξ_1,ξ_2) of the patch $\mathcal{P}$ carries over to the image $\mathcal{P}_t$ of $\mathcal{P}$ under the mapping $x = \phi_t(X)$. The parallelogram on $\mathcal{P}_t$ has area dA_x and is defined by the tangent vectors

$$\frac{\partial}{\partial\xi_1}\phi_t\big(\chi(\xi_1,\xi_2)\big)\, d\xi_1 \quad \text{and} \quad \frac{\partial}{\partial\xi_2}\phi_t\big(\chi(\xi_1,\xi_2)\big)\, d\xi_2.$$

The cross product of these then equals, in magnitude, to the parallelogram's area and points along the unit normal n to the surface. We apply the chain rule to these and write

$\boldsymbol{F}(\boldsymbol{X},t) = \textbf{Grad}\,\boldsymbol{\phi}_t(\boldsymbol{X})$, as usual, to simplify the resulting expression:

$$\begin{aligned}
\boldsymbol{n}\,dA_{\boldsymbol{x}} &= \Big(\frac{\partial}{\partial\xi_1}\boldsymbol{\phi}_t\big(\boldsymbol{\chi}(\xi_1,\xi_2)\big)\,d\xi_1\Big)\times\Big(\frac{\partial}{\partial\xi_2}\boldsymbol{\phi}_t\big(\boldsymbol{\chi}(\xi_1,\xi_2)\big)\,d\xi_2\Big)\\
&= \Big(\boldsymbol{F}\frac{\partial\boldsymbol{\chi}}{\partial\xi_1}\,d\xi_1\Big)\times\Big(\boldsymbol{F}\frac{\partial\boldsymbol{\chi}}{\partial\xi_2}\,d\xi_2\Big)\\
&= \Big(\boldsymbol{F}\frac{\partial\boldsymbol{\chi}}{\partial\xi_1}\Big)\times\Big(\boldsymbol{F}\frac{\partial\boldsymbol{\chi}}{\partial\xi_2}\Big)\,d\xi_1\,d\xi_2\\
&= (\det\boldsymbol{F})\,\boldsymbol{F}^{-T}\big(\frac{\partial\boldsymbol{\chi}}{\partial\xi_1}\times\frac{\partial\boldsymbol{\chi}}{\partial\xi_2}\big)\,d\xi_1\,d\xi_2,
\end{aligned}$$

where in the last step we have applied the identity (211) from Exercise 19. Comparing this with (123) we conclude that

$$\boldsymbol{n}\,dA_{\boldsymbol{x}} = (\det\boldsymbol{F})\,\boldsymbol{F}^{-T}\boldsymbol{N}\,dA_{\boldsymbol{X}} \tag{124}$$

which suffices to conclude the proof. □

24. Conservation laws

24.1. **The conservation of mass.** Consider the motion $\boldsymbol{\phi}_t$ of the body B, and let $B_t = \boldsymbol{\phi}_t(B)$ and $\boldsymbol{v}(\boldsymbol{x},t)$ be the velocity field associated with the motion. Let $\rho(\boldsymbol{x},t)$ be the density, i.e., mass per unit volume, of the body in the deformed state B_t. We assume that total mass of any subdomain Ω of B remains constant during the motion, that is

$$\frac{d}{dt}\int_{\Omega_t}\rho(\boldsymbol{x},t)\,dV_{\boldsymbol{x}} = 0,$$

where $\Omega_t = \boldsymbol{\phi}_t(\Omega)$. Then, according to equation (121b) of Reynolds Transport Theorem we have

$$\int_{\Omega_t}\Big(\frac{\partial\rho}{\partial t} + \operatorname{div}(\rho\boldsymbol{v})\Big)\,dV_{\boldsymbol{x}} = 0,$$

This holds for any subdomain Ω_t, and therefore the integrand is zero everywhere:

$$\frac{\partial\rho}{\partial t} + \operatorname{div}(\rho\boldsymbol{v}) = 0. \tag{125}$$

This expresses the *principle of conservation of mass* in our context.

24.2. **Integrals with respect to to density.** An interesting, and often useful, consequence of the conservation of mass equation (125) is

Theorem 27. *Let $\boldsymbol{\phi}_t$ be the motion of the body B as before, $B_t = \boldsymbol{\phi}_t(B)$, and let $\omega(\boldsymbol{x},t)$ be any scalar-, vector-, or tensor-valued field defined on B_t. Then for any subdomain Ω_t of B_t we have*

$$\frac{d}{dt}\int_{\Omega_t}\omega(\boldsymbol{x},t)\,\rho(\boldsymbol{x},t)\,dV_{\boldsymbol{x}} = \int_{\Omega_t}\dot{\omega}(\boldsymbol{x},t)\,\rho(\boldsymbol{x},t)\,dV_{\boldsymbol{x}}, \tag{126}$$

where $\dot{\omega}$ is the total time derivative of ω.

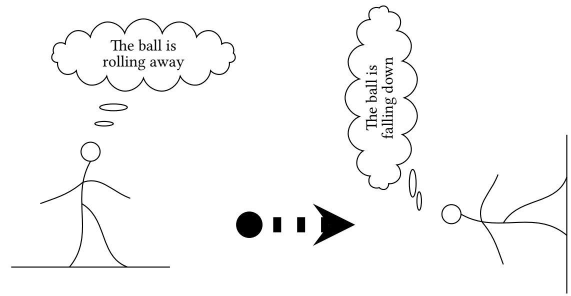


FIGURE 10. Two observers see the motion of the ball differently.

25. Frame-indifference

The axiom of *material frame-indifference* is a very basic tenet of continuum mechanics although it was formally identified as such by Oldroyd only in 1950 [10]. In the 1950's and 1960's the axiom was brought into the mainstream of continuum mechanics by Truesdell, Noll, and other researchers and became variously known as the principle of *frame-indifference, isotropy of space, material objectivity, material reference invariance.* As in [5], we will refer to it as the *principle of material frame-indifference*, or just *frame-indifference* for short.

Two observers equipped with independently moving reference frames will see a given motion differently; see the illustration in Figure 10. Their observations will be related through a possibly time-dependent translation and rotation that reflects their positions relative to each other.

To be precise, the positive vector $\boldsymbol{x}$ of a point $P \in \mathbb{E}_3$ viewed by one observer appears as a position vector

$$\boldsymbol{x}^* = \boldsymbol{q}(t) + \boldsymbol{Q}(t)\,\boldsymbol{x} \tag{127}$$

to the other, where $\boldsymbol{q}(t)$ is the translation vector and $\boldsymbol{Q}(t)$ is an orthogonal tensor. The motion $\boldsymbol{x} = \boldsymbol{\phi}_t(\boldsymbol{X})$ of a body B viewed by one observer appears as

$$\boldsymbol{x}^* = \boldsymbol{\phi}_t^*(\boldsymbol{X}) = \boldsymbol{q}(t) + \boldsymbol{Q}(t)\boldsymbol{\phi}_t(\boldsymbol{X}) \tag{128}$$

to the other. Applying the chain rule of differentiation, we see that the deformation gradients

$$\boldsymbol{F}(\boldsymbol{X},t) = \mathbf{Grad}\,\boldsymbol{\phi}_t(\boldsymbol{X}), \quad \boldsymbol{F}^*(\boldsymbol{X},t) = \mathbf{Grad}\,\boldsymbol{\phi}_t^*(\boldsymbol{X})$$

are related through $\boldsymbol{F}^*(\boldsymbol{X},t) = \boldsymbol{Q}(t)\boldsymbol{F}(\boldsymbol{X},t)$ or

$$\boldsymbol{F}^* = \boldsymbol{Q}\boldsymbol{F}, \tag{129}$$

for short.

25.1. **The transformation of the Cauchy–Green tensors.** We may obtain the relationship between other kinematic indicators as seen by the two observers. For instance, let

$$\boldsymbol{F} = \boldsymbol{R}\boldsymbol{U} = \boldsymbol{V}\boldsymbol{R}, \quad \boldsymbol{F}^* = \boldsymbol{R}^*\boldsymbol{U}^* = \boldsymbol{V}^*\boldsymbol{R}^* \tag{130}$$

be the polar decompositions of the deformation gradients. Then from (129) and $(130)_1$ we see that $\boldsymbol{F}^* = \boldsymbol{Q}\boldsymbol{F} = \boldsymbol{Q}(\boldsymbol{R}\boldsymbol{U}) = (\boldsymbol{Q}\boldsymbol{R})\boldsymbol{U}$. Since $\boldsymbol{Q}\boldsymbol{R}$ is an orthogonal tensor, what we have here is a polar decomposition of $\boldsymbol{F}^*$. On the other hand, the polar decomposition of $\boldsymbol{F}^*$ is given as $\boldsymbol{R}^*\boldsymbol{U}^*$ in $(130)_2$. By the uniqueness of the polar decomposition (see Theorem 11 on page 25) we conclude that $\boldsymbol{R}^* = \boldsymbol{Q}\boldsymbol{R}$ and $\boldsymbol{U}^* = \boldsymbol{U}$. The right Cauchy–Green strain tensors, as seen by the two observers, are $\boldsymbol{C} = \boldsymbol{U}^2$ and $\boldsymbol{C}^* = \boldsymbol{U}^{*2}$, and therefore

$$\boldsymbol{C}^* = \boldsymbol{C}. \tag{131}$$

Similarly, may may show (see Exercise 37) that the left Cauchy–Green tensors $\boldsymbol{B} = \boldsymbol{V}^2$ and $\boldsymbol{B}^* = \boldsymbol{V}^{*2}$ are related through

$$\boldsymbol{B}^* = \boldsymbol{Q}\boldsymbol{B}\boldsymbol{Q}^T. \tag{132}$$

25.2. **The transformation of the velocity gradient and the rate of strain.** The velocity gradients for the two observers may be obtained from (119) as

$$\dot{\boldsymbol{F}}(\boldsymbol{X},t) = \operatorname{\mathbf{grad}} \boldsymbol{v}(\boldsymbol{x},t)\Big|_{\boldsymbol{x}=\boldsymbol{\phi}_t(\boldsymbol{X})} \boldsymbol{F}(\boldsymbol{X},t), \quad \dot{\boldsymbol{F}}^*(\boldsymbol{X},t) = \operatorname{\mathbf{grad}} \boldsymbol{v}^*(\boldsymbol{x}^*,t)\Big|_{\boldsymbol{x}^*=\boldsymbol{\phi}^*_t(\boldsymbol{X})} \boldsymbol{F}^*(\boldsymbol{X},t).$$

We substitute for $\boldsymbol{F}^*$ from (129), and also for its time derivative from $\dot{\boldsymbol{F}}^* = (\boldsymbol{Q}\boldsymbol{F})^{\bullet} = \dot{\boldsymbol{Q}}\boldsymbol{F} + \boldsymbol{Q}\dot{\boldsymbol{F}}$ in the second of the two equations above, and obtain

$$\dot{\boldsymbol{Q}}\boldsymbol{F} + \boldsymbol{Q}\dot{\boldsymbol{F}} = \operatorname{\mathbf{grad}} \boldsymbol{v}^*(\boldsymbol{x}^*,t)\Big|_{\boldsymbol{x}^*=\boldsymbol{\phi}^*_t(\boldsymbol{X})} \boldsymbol{Q}\boldsymbol{F},$$

where $\boldsymbol{F} = \boldsymbol{F}(\boldsymbol{X},t)$ and $\boldsymbol{Q} = \boldsymbol{Q}(t)$. Then, substituting for $\dot{\boldsymbol{F}}$ from the fist equation we arrive at

$$\dot{\boldsymbol{Q}}\boldsymbol{F} + \boldsymbol{Q} \operatorname{\mathbf{grad}} \boldsymbol{v}(\boldsymbol{x},t)\Big|_{\boldsymbol{x}=\boldsymbol{\phi}_t(\boldsymbol{X})} \boldsymbol{F} = \operatorname{\mathbf{grad}} \boldsymbol{v}^*(\boldsymbol{x}^*,t)\Big|_{\boldsymbol{x}^*=\boldsymbol{\phi}^*_t(\boldsymbol{X})} \boldsymbol{Q}\boldsymbol{F},$$

We see that the $\boldsymbol{F}$ cancels from the two sides. We then solve for $\operatorname{\mathbf{grad}} \boldsymbol{v}^*$ and thus obtain the transformation rule for the velocity gradient:

$$\operatorname{\mathbf{grad}} \boldsymbol{v}^*(\boldsymbol{x}^*,t)\Big|_{\boldsymbol{x}^*=\boldsymbol{\phi}^*_t(\boldsymbol{X})} = \dot{\boldsymbol{Q}}\boldsymbol{Q}^T + \boldsymbol{Q} \operatorname{\mathbf{grad}} \boldsymbol{v}(\boldsymbol{x},t)\Big|_{\boldsymbol{x}=\boldsymbol{\phi}_t(\boldsymbol{X})} \boldsymbol{Q}^T. \tag{133}$$

The symmetric part of $\operatorname{\mathbf{grad}} \boldsymbol{v}$

$$\boldsymbol{D}(\boldsymbol{x},t) = \frac{1}{2}\Big(\operatorname{\mathbf{grad}} \boldsymbol{v}(\boldsymbol{x},t) + \operatorname{\mathbf{grad}} \boldsymbol{v}(\boldsymbol{x},t)^T\Big) \tag{134}$$

is called *the rate of strain*[6] and plays a central role in fluid mechanics. The rate of strain seen by the starred observer is

$$\boldsymbol{D}^*(\boldsymbol{x}^*,t) = \frac{1}{2}\Big(\operatorname{\mathbf{grad}} \boldsymbol{v}^*(\boldsymbol{x}^*,t) + \operatorname{\mathbf{grad}} \boldsymbol{v}^*(\boldsymbol{x}^*,t)^T\Big).$$

Substituting from (133) we see that

$$\boldsymbol{D}^*(\boldsymbol{x}^*,t) = \frac{1}{2}\bigg(\dot{\boldsymbol{Q}}\boldsymbol{Q}^T + \boldsymbol{Q}\dot{\boldsymbol{Q}}^T + \boldsymbol{Q}\big(\operatorname{\mathbf{grad}} \boldsymbol{v}(\boldsymbol{x},t) + \operatorname{\mathbf{grad}} \boldsymbol{v}(\boldsymbol{x},t)^T\big)\boldsymbol{Q}^T\bigg).$$

But since $\boldsymbol{Q}$ is orthogonal, we have $\boldsymbol{Q}\boldsymbol{Q}^T = \boldsymbol{I}$, and therefore $\dot{\boldsymbol{Q}}\boldsymbol{Q}^T + \boldsymbol{Q}\dot{\boldsymbol{Q}}^T = \boldsymbol{0}$. We conclude that

$$\boldsymbol{D}^*(\boldsymbol{x}^*,t) = \boldsymbol{Q}(t)\,\boldsymbol{D}(\boldsymbol{x},t)\,\boldsymbol{Q}(t)^T. \tag{135}$$

[6] In [5] (page 134) the rate of strain is denoted by $\boldsymbol{L}$. In all other literature that I have checked, the rate of strain is denoted by $\boldsymbol{D}$.

25.3. **The transformation of stress.** Consider the motion $\boldsymbol{\phi}_t$ of a body B and let $B_t = \boldsymbol{\phi}_t(B)$, and let $\boldsymbol{S}(\boldsymbol{x}, t)$ be the associated stress field. Thus, at any point $\boldsymbol{x} \in B_t$, the traction at $\boldsymbol{x}$ on a plain with a unit normal $\boldsymbol{n}$ is $\boldsymbol{t} = \boldsymbol{S}\boldsymbol{n}$.

A second observer moving relative to the first one according to (128), sees that traction as $\boldsymbol{t}^* = \boldsymbol{S}^*\boldsymbol{n}^*$, where $\boldsymbol{n}^* = \boldsymbol{Q}\boldsymbol{n}$ and $\boldsymbol{t}^* = \boldsymbol{Q}\boldsymbol{t} = \boldsymbol{Q}\boldsymbol{S}\boldsymbol{n}$. It follows that $\boldsymbol{Q}\boldsymbol{S}\boldsymbol{n} = \boldsymbol{S}^*\boldsymbol{Q}\boldsymbol{n}$ for all $\boldsymbol{n}$, and therefore

$$\boldsymbol{S}^* = \boldsymbol{Q}\,\boldsymbol{S}\,\boldsymbol{Q}^T. \tag{136}$$

25.4. **The principle of material frame-indifference.** The characteristics property of a material from the point of view of continuum mechanics is its *constitutive equation* which relates the stress in it to its deformation.

For instance, the stress resulting from a given deformation in a body made of steel would generally be very different from the stress in a body of the same shape subjected to the identical deformation, but made of rubber or water. There is a vast variety of materials, and correspondingly there is a vast variety of constitutive equations. These tend to be categorized into special classes, such as elastic materials, plastic materials, fluid, gases, etc.. In all cases, the constitutive equation expresses the stress $\boldsymbol{S}(\boldsymbol{x}, t)$ at a point $\boldsymbol{x}$ at time t in terms of the history of deformation:

$$\boldsymbol{S}(\boldsymbol{x}, t) = \boldsymbol{\mathcal{F}}(\boldsymbol{\phi}_s \,:\, s \leq t).$$

In general, $\boldsymbol{\mathcal{F}}$ is *functional* in the sense that it may depend on the history of the material's deformation up to the current time t. In the simplest cases $\boldsymbol{\mathcal{F}}$ depends on the current deformation $\boldsymbol{\phi}_t$ (elastic materials) or current time derivative $\dot{\boldsymbol{\phi}}_t$ (fluids). We will study the details of elastic materials and fluids later in these notes, but for now those distinctions are not essential.

The *principle of material frame-indifference* asserts that the a material's constitutive equation is independent of the observer. That is, two observers viewings the motion of a body, and acutely aware of the transformation rules (129), (131), (132), (135), and (136), should find that stresses and deformations are related through an identical rule for both of them.

The gist of the idea of the principle of material frame-indifference is illustrated by the following well-known and very simple example.

The simplest elastic system, often encountered in elementary courses, is a massless linear spring, modeled with the constitutive equation $F = kx$, known as *Hooke's Law*, where F is the force that stretches the spring, and x is the change in the spring's natural length due to the application of the force F. This constitutive equation is characterized by a single constant, k, called *Hooke's constant.* The linear dependence of F on x is immaterial for example that we are going to present; a nonlinear constitutive equation $F = f(x)$ will do just as well, and that's what we will use here.

Attach a mass to one end of the spring and attach the other end to the axle of a horizontal platter and spin the system with a constant angular velocity; see Figure 11(a). The spinning spring extends by some amount, say δ, due to the action of centrifugal force.

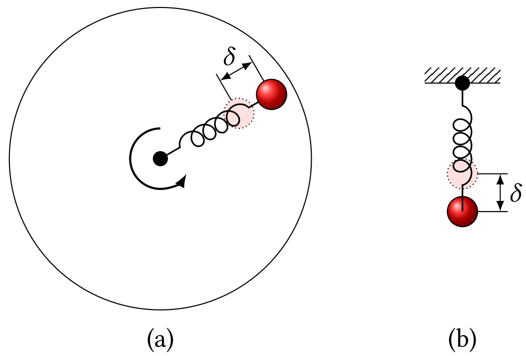


FIGURE 11. On the left, the mass-spring system spins on a horizontal platter and the spring stretches by δ. On the right, the mass-spring system is suspended from a fixed support and the spring stretches by δ. Are the tensile forces equal in the two experiments?

In a second experiment, suspend the same mass-spring system from a fixed support as in Figure 11(b). The spring now stretches due to gravity. Suppose that this extension also happens to be δ. It seems intuitive to conclude that the force of gravity equals to the centrifugal force.

What is the logic behind that conclusion? It's the principle of material frame-indifference! Consider an observer, Alice, riding on the spinning platter. From Alice's point of view, the spinning spring in Figure 11(a) is stationary, and so it looks identical to the suspended spring of Figure 11(b). The spring has been stretched by δ, therefore the force acting on it is $f(\delta)$ in both cases. From the point of view of Bob, a stationary observer on the lab floor, the spinning spring is stretched by the amount δ. Bob has no force measurement data for this experiment, but he accepts the principle of material frame-indifference, and therefore applies the same constitutive equation to the spinning spring and, noting that it has been stretched by δ, concludes that the tensile force in it is $f(\delta)$.

Admittedly, this instance of the application of the principle of material frame-indifference is so nearly trivial that it does not inspire confidence in its utility. In the following sections, however, we will see the principle's tremendous impact on the formulation of the constitutive equations.

26. Elastic materials

An *elastic material* is one whose Cauchy stress tensor $S(\boldsymbol{x}, t)$ at the point $\boldsymbol{x} = \boldsymbol{\phi}_t(\boldsymbol{X})$ is determined solely by the gradient of deformation at that point. Thus, its constitutive equation is

$$S(\boldsymbol{x}, t)\Big|_{\boldsymbol{x}=\boldsymbol{\phi}_t(\boldsymbol{X})} = \hat{S}\big(\boldsymbol{F}(\boldsymbol{X}, t), \boldsymbol{X}\big) \tag{137}$$

for some function $\hat{S} : \mathcal{L} \times B \to \mathcal{L}_{\text{sym}}$.

The explicit occurrence of $\boldsymbol{X}$ as the second argument of $\hat{S}$ in (137) allows for the constitutive equation to vary from point to point. In the special case where $\hat{S}$ is independent of the second argument, the material is said to be *homogeneous*. Otherwise the material is *inhomogeneous*.

Most of this section's discussion focuses on a single point $\boldsymbol{X}$ in the body, therefore there is no gain in emphasizing the dependence of $\hat{S}$ on $\boldsymbol{X}$. Thus we simply write the constitutive equation as:

$$\boldsymbol{S} = \hat{S}(\boldsymbol{F}). \tag{138}$$

26.1. **Frame-indifference.** Recall the two observers scenario introduced in section 25. The deformation gradients seen by the two observers are $\boldsymbol{F}$ and $\boldsymbol{F}^* = \boldsymbol{Q}\boldsymbol{F}$. Therefore the stresses are $\boldsymbol{S} = \hat{S}(\boldsymbol{F})$ and $\boldsymbol{S}^* = \hat{S}(\boldsymbol{F}^*) = \hat{S}(\boldsymbol{Q}\boldsymbol{F})$. Note that both observers use the same constitutive equation $\hat{S}$ in accordance with the principle of material frame-indifference. Thus from (136) we conclude that the constitutive equation of an elastic material subject to frame-indifference must satisfy:

$$\hat{S}(\boldsymbol{Q}\boldsymbol{F}) = \boldsymbol{Q}\hat{S}(\boldsymbol{F})\boldsymbol{Q}^T \quad \text{for all } \boldsymbol{F} \in \mathcal{L} \text{ and all } \boldsymbol{Q} \in \mathcal{L}^+_{\text{orth}}. \tag{139}$$

This may be put in a more symmetric form through the following trick. Consider the polar decomposition $\boldsymbol{F} = \boldsymbol{R}\boldsymbol{U}$ of $\boldsymbol{F}$ and let $\boldsymbol{Q} = \boldsymbol{R}^T$ in (139). Since $\boldsymbol{Q}\boldsymbol{F} = \boldsymbol{R}^T\boldsymbol{R}\boldsymbol{U} = \boldsymbol{U}$, we get $\hat{S}(\boldsymbol{U}) = \boldsymbol{R}^T\hat{S}(\boldsymbol{F})\boldsymbol{R}$, or equivalently:

$$\hat{S}(\boldsymbol{F}) = \boldsymbol{R}\hat{S}(\boldsymbol{U})\boldsymbol{R}^T, \quad \text{for all } \boldsymbol{F} = \boldsymbol{R}\boldsymbol{U} \in \mathcal{L}. \tag{140}$$

The *right Cauchy–Green strain tensor* is defined through $\boldsymbol{C} = \boldsymbol{F}^T\boldsymbol{F} = \boldsymbol{U}^2$. Therefore the preceding formula may be expressed as $\hat{S}(\boldsymbol{F}) = \boldsymbol{R}\hat{S}(\boldsymbol{C}^{1/2})\boldsymbol{R}^T$. Thus, we introduce the function $\tilde{S}$ through

$$\tilde{S}(\boldsymbol{C}) = \hat{S}(\boldsymbol{C}^{1/2}), \tag{141}$$

and arrive at:

$$\hat{S}(\boldsymbol{F}) = \boldsymbol{R}\tilde{S}(\boldsymbol{C})\boldsymbol{R}^T, \quad \text{for all } \boldsymbol{F} \in \mathcal{L}, \tag{142}$$

where $\boldsymbol{F} = \boldsymbol{R}\boldsymbol{U}$, and $\boldsymbol{C} = \boldsymbol{U}^2$. In summary, the principle of frame-indifference implies that the constitutive equation of an elastic material has to obey the identity (142).

26.2. **Symmetry groups and isotropy.** In this section we introduce the concepts of *symmetry group* and *isotropy* of an elastic material. These concepts are tied to a specific point of the material; the symmetry and isotropy properties may vary from point to point.

Suppose we rotate an elastic body B by an orthogonal tensor $\boldsymbol{Q}$ about a point $\boldsymbol{p} \in B$. If the material's elastic response[7] to arbitrary experiments is indistinguishable before and after the rotation at the point $\boldsymbol{p}$, then we say that $\boldsymbol{Q}$ a *symmetry transformation* of the material $\boldsymbol{p}$; see Figure 12. For instance, the ordinary salt, sodium chloride, is normally found in crystalline form, consisting of cubic crystals. Rotating a sample of salt by 90 degrees about an axis perpendicular to the crystal faces, or by 120 degrees about an axis along a diagonal of the crystals, results in a material which is mechanically indistinguishable from the original. Therefore those rotations are symmetry transformations of salt crystals. In contrast, the ordinary rubber's behavior is insensitive to orientation, therefore any rotation is a symmetry transformation.

[7] That is, stress induced by a deformation.

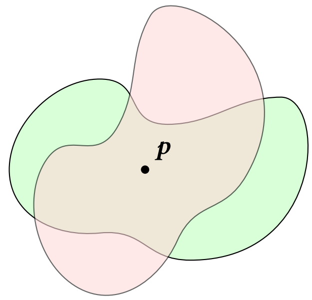


FIGURE 12. The body has been rotated about a point $\boldsymbol{p}$ by a an orthogonal tensor $\boldsymbol{Q}$. If the elastic behavior of the rotated material is indistinguishable from that of the non-rotated one *at the point* $\boldsymbol{p}$, then $\boldsymbol{Q}$ is a symmetry transformation at $\boldsymbol{p}$.

To quantify the effect of a symmetry transformation, let us consider an elastic body B, and without loss of generality, let's pick the rotation center $\boldsymbol{p}$ to be the origin. We subject B to two experiments as follows.

In the first experiment we subject B to a deformation $\boldsymbol{\phi}$. Thus, every point $\boldsymbol{X} \in \mathcal{B}$ moves to $\boldsymbol{x} = \boldsymbol{\phi}(\boldsymbol{X})$. The gradient of deformation is $\boldsymbol{F}(\boldsymbol{X}) = \mathbf{Grad}\,\boldsymbol{\phi}(\boldsymbol{X})$.

In the second experiment, we rotate B about the origin through an orthogonal tensor $\boldsymbol{Q} \in \mathcal{L}^+_{\text{orth}}$ and then apply the deformation $\boldsymbol{\phi}$. A point $\boldsymbol{X}$ moves to $\boldsymbol{QX}$ due to the rotation, and then it moves to $\boldsymbol{x} = \boldsymbol{\phi}(\boldsymbol{QX})$ due to the application of $\boldsymbol{\phi}$. The gradient of the deformation is

$$\tilde{\boldsymbol{F}}(\boldsymbol{X}) = \mathbf{Grad}\,\boldsymbol{\phi}\Big|_{\boldsymbol{QX}}\,\mathbf{Grad}(\boldsymbol{QX}) = \mathbf{Grad}\,\boldsymbol{\phi}\Big|_{\boldsymbol{QX}}\boldsymbol{Q} = \boldsymbol{F}(\boldsymbol{QX})\boldsymbol{Q}.$$

We conclude that $\tilde{\boldsymbol{F}}(\mathbf{0}) = \boldsymbol{F}(\mathbf{0})\boldsymbol{Q}$. If $\boldsymbol{Q}$ is a symmetry transformation of the material point at the origin, then the stress at the origin due to the deformation gradients $\boldsymbol{F}(\mathbf{0})$ and $\boldsymbol{F}(\mathbf{0})\boldsymbol{Q}$ should be identical, that is, $\hat{S}\big(\boldsymbol{F}(\mathbf{0})\boldsymbol{Q}\big) = \hat{S}\big(\boldsymbol{F}(\mathbf{0})\big)$, for all deformations $\boldsymbol{\phi}$, where $\hat{S}$ is the constitutive function of the material at the origin. Since $\boldsymbol{\phi}$ is arbitrary, the deformation gradient $\boldsymbol{F}(\mathbf{0})$ is arbitrary tensor in $\mathcal{L}^+$. Furthermore, since there is nothing special about the point selected as the origin, our findings applies to any point in B. We summarize this as:

$$\hat{S}(\boldsymbol{FQ}) = \hat{S}(\boldsymbol{F}) \quad \text{for all } \boldsymbol{F} \in \mathcal{L}^+. \tag{143}$$

The set of all symmetry transformation of a material at a point $\boldsymbol{p}$ is called the material's *symmetry group* at $\boldsymbol{p}$. Generally, a material's symmetry group is a subgroup of the proper orthogonal group, $\mathcal{L}^+_{\text{orth}}$, but if the symmetry group is the entire $\mathcal{L}^+_{\text{orth}}$, the material is said to be *isotropic*[8] at $\boldsymbol{p}$. Thus, a material is isotropic at a point $\boldsymbol{p}$ if

$$\hat{S}(\boldsymbol{FQ}, \boldsymbol{p}) = \hat{S}(\boldsymbol{F}, \boldsymbol{p}) \quad \text{for all } \boldsymbol{F} \in \mathcal{L}^+ \text{ and all } \boldsymbol{Q} \in \mathcal{L}^+_{\text{orth}}. \tag{144}$$

[8]From the Greek *isos*="equal" and *tropikos*="related to turn".

26.3. **Combining frame-indifference and isotropy.** We have seen that the principle of frame-indifference implies that and elastic material's constitutive function, $\hat{S}$, must satisfy the condition (139) or the equivalent (142). If, additionally, the material is isotropic, then the constitutive equation must satisfy the condition (144). The following proposition explores the consequences of these. Isotropic tensor-valued function were defined in section 16.2 on page 31.

Proposition 8. *Suppose that the elastic material is isotropic. Then the functions $\hat{S}$ and $\tilde{S}$ defined in* (138) *and* (141) *are* isotropic*, that is, for any rotation $\boldsymbol{Q}$ we have:*

$$\boldsymbol{Q}\hat{S}(\boldsymbol{F})\boldsymbol{Q}^T = \hat{S}(\boldsymbol{Q}\boldsymbol{F}\boldsymbol{Q}^T) \quad \text{for all } \boldsymbol{F} \in \mathcal{L}, \tag{145}$$

$$\boldsymbol{Q}\tilde{S}(\boldsymbol{C})\boldsymbol{Q}^T = \tilde{S}(\boldsymbol{Q}\boldsymbol{C}\boldsymbol{Q}^T) \quad \text{for all } \boldsymbol{C} \in \mathcal{L}^+_{sym}. \tag{146}$$

Proof. The identity (139), which expresses the frame-indifference of the stress, holds for all choices of $\boldsymbol{F} \in \mathcal{L}$ and rotations $\boldsymbol{Q}$. Thus we may replace $\boldsymbol{F}$ by $\boldsymbol{F}\boldsymbol{Q}^T$ in it. This gives $\hat{S}(\boldsymbol{Q}\boldsymbol{F}\boldsymbol{Q}^T) = \boldsymbol{Q}\hat{S}(\boldsymbol{F}\boldsymbol{Q}^T)\boldsymbol{Q}^T$. However $\hat{S}(\boldsymbol{F}\boldsymbol{Q}^T) = \hat{S}(\boldsymbol{F})$ by the isotropy condition (144). This proves (145).

To prove (146), pick an arbitrary positive-definite $\boldsymbol{C} \in \mathcal{L}^+_{\text{sym}}$, and let $\boldsymbol{U} \in \mathcal{L}^+_{\text{sym}}$ be its unique square root, that is, $\boldsymbol{C} = \boldsymbol{U}^2$. Then, according to the definition (141), for any rotation $\boldsymbol{Q}$ we have:

$$\begin{aligned}
\boldsymbol{Q}\tilde{S}(\boldsymbol{C})\boldsymbol{Q}^T &= \boldsymbol{Q}\hat{S}(\boldsymbol{U})\boldsymbol{Q}^T && \text{(by (141))}\\
&= \hat{S}(\boldsymbol{Q}\boldsymbol{U}\boldsymbol{Q}^T) && \text{(by (145))}\\
&= \tilde{S}\big((\boldsymbol{Q}\boldsymbol{U}\boldsymbol{Q}^T)^2\big) && \text{(by (141))}\\
&= \tilde{S}(\boldsymbol{Q}\boldsymbol{U}^2\boldsymbol{Q}^T)\\
&= \tilde{S}(\boldsymbol{Q}\boldsymbol{C}\boldsymbol{Q}^T).
\end{aligned}$$

□

Corollary 5. *Under the assumptions of the preceding proposition we have:*

$$\hat{S}(\boldsymbol{F}) = \tilde{S}(\boldsymbol{B})$$

where $\boldsymbol{B} = \boldsymbol{F}\boldsymbol{F}^T$ is the left Cauchy–Green strain tensor.

Proof. Consider the polar decomposition $\boldsymbol{F} = \boldsymbol{R}\boldsymbol{U}$. We have:

$$\boldsymbol{B} = \boldsymbol{F}\boldsymbol{F}^T = (\boldsymbol{R}\boldsymbol{U})(\boldsymbol{R}\boldsymbol{U})^T = (\boldsymbol{R}\boldsymbol{U})(\boldsymbol{U}\boldsymbol{R}) = \boldsymbol{R}\boldsymbol{U}^2\boldsymbol{R} = \boldsymbol{R}\boldsymbol{C}\boldsymbol{R},$$

therefore

$$\begin{aligned}
\hat{S}(\boldsymbol{F}) &= \boldsymbol{R}\tilde{S}(\boldsymbol{C})\boldsymbol{R}^T && \text{(by (142))}\\
&= \tilde{S}(\boldsymbol{R}\boldsymbol{C}\boldsymbol{R}^T) && \text{(by (146))}\\
&= \tilde{S}(\boldsymbol{B}).
\end{aligned}$$

□

Theorem 28. *The constitutive equations of an isotropic elastic material is necessarily of the form;*

$$\boldsymbol{S} = \hat{S}(\boldsymbol{F}) = \tilde{S}(\boldsymbol{B}) = \alpha_0\boldsymbol{I} + \alpha_1\boldsymbol{B} + \alpha_2\boldsymbol{B}^2, \tag{147a}$$

where $\boldsymbol{F}$ is the deformation gradient, $\boldsymbol{B} = \boldsymbol{F}\boldsymbol{F}^T$ is the left Cauchy–Green strain tensor, and the α_i, $i = 0, 1, 2$, are scalar-valued isotropic functions of $\boldsymbol{B}$, that is, they depend on $\boldsymbol{B}$'s principal invariants $\iota_1(\boldsymbol{B})$, $\iota_2(\boldsymbol{B})$, $\iota_3(\boldsymbol{B})$.

Proof. According to (146), $\tilde{S}$ is an isotropic function over $\mathcal{L}^+_{\mathrm{sym}}$. Then (147a) is an immediate consequence of Theorem 13 on page 31. □

Remark 27. The constitutive equation (147a) is not as simple as it may seem at first glance. Its coefficients are functions of $\boldsymbol{B}$'s invariants, therefore the fully expressed form of the equation is

$$\begin{aligned} \boldsymbol{S} &= \hat{S}(\boldsymbol{F}) = \tilde{S}(\boldsymbol{B}) \\ &= \alpha_0\Big(\iota_1(\boldsymbol{B}), \iota_2(\boldsymbol{B}), \iota_3(\boldsymbol{B})\Big)\boldsymbol{I} + \alpha_1\Big(\iota_1(\boldsymbol{B}), \iota_2(\boldsymbol{B}), \iota_3(\boldsymbol{B})\Big)\boldsymbol{B} + \alpha_2\Big(\iota_1(\boldsymbol{B}), \iota_2(\boldsymbol{B}), \iota_3(\boldsymbol{B})\Big)\boldsymbol{B}^2. \end{aligned}$$

Corollary 6. *The constitutive equation* (147a) *may be expressed equivalently as:*

$$\boldsymbol{S} = \beta_0\boldsymbol{I} + \beta_1\boldsymbol{B} + \beta_2\boldsymbol{B}^{-1}, \tag{147b}$$

where $\boldsymbol{B}$ is as before, and where the scalar-valued coefficients β_i, $i = 0, 1, 2$, depend on$\boldsymbol{B}$'s principal invariants $\iota_1(\boldsymbol{B})$, $\iota_2(\boldsymbol{B})$, $\iota_3(\boldsymbol{B})$.

Proof. According to the Caley-Hamilton theorem, $\boldsymbol{B}$ satisfies its own characteristic equation, that is:

$$\boldsymbol{B}^3 - \iota_1(\boldsymbol{B})\boldsymbol{B}^2 + \iota_2(\boldsymbol{B})\boldsymbol{B} + \iota_3(\boldsymbol{B})\boldsymbol{I} = \boldsymbol{0}.$$

Multiplying through by $\boldsymbol{B}^{-1}$ we get:

$$\boldsymbol{B}^2 - \iota_1(\boldsymbol{B})\boldsymbol{B} + \iota_2(\boldsymbol{B})\boldsymbol{I} + \iota_3(\boldsymbol{B})\boldsymbol{B}^{-1} = \boldsymbol{0}.$$

Thus, we may eliminate $\boldsymbol{B}^2$ in (147a) in favor of $\boldsymbol{B}^{-1}$:

$$\begin{aligned} \boldsymbol{S} &= \alpha_0\boldsymbol{I} + \alpha_1\boldsymbol{B} + \alpha_2\boldsymbol{B}^2 \\ &= \alpha_0\boldsymbol{I} + \alpha_1\boldsymbol{B} + \alpha_2\big(\iota_1\boldsymbol{B} - \iota_2\boldsymbol{I} - \iota_3\boldsymbol{B}^{-1}\big) \\ &= \big(\alpha_0 - \alpha_2\iota_2\big)\boldsymbol{I} + \big(\alpha_1 - \alpha_2\iota_1\big)\boldsymbol{B} - \alpha_2\iota_3\boldsymbol{B}^{-1}. \end{aligned}$$

□

27. Elasticity: Simple shear

A homogeneous elastic cube is placed in the Cartesian coordinate system with its faces parallel to the coordinate planes, as in Figure 13(a). Tractions are applied to the cube's surfaces to deform it into Figure 13(b) according to $\boldsymbol{x} = \boldsymbol{\phi}(\boldsymbol{X})$ defined by

$$\begin{aligned} x_1 &= X_1 + cX_2, \\ x_2 &= X_2, \\ x_3 &= X_3, \end{aligned}$$

where $c = \tan\theta$ is the angle of inclination of the deformed cube.

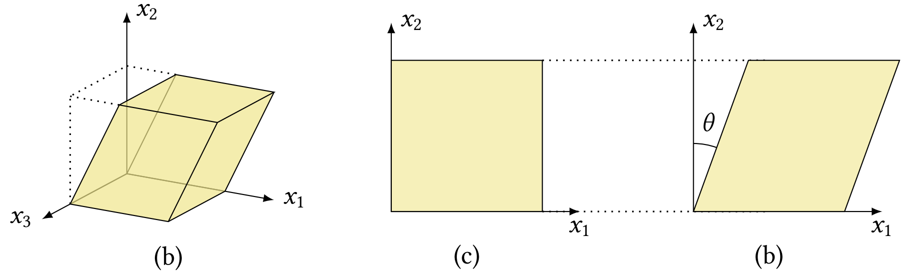


FIGURE 13. Subfigure (a) shows the deformation of the cube in a simple homogeneous shear in the x_1 direction. Subfigures (b) and (c) show the cube's before and after configurations with the x_3 axis pointing towards the viewer.

The components of the deformation gradient $\boldsymbol{F}$ are $F_{ij} = \phi_{i,j}$:

$$\boldsymbol{F} = \begin{pmatrix} 1 & c & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and therefore the left Cauchy–Green tensor and its inverse are

$$\boldsymbol{B} = \boldsymbol{F}\boldsymbol{F}^T = \begin{pmatrix} 1+c^2 & c & 0 \\ c & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \boldsymbol{B}^{-1} = \begin{pmatrix} 1 & -c & 0 \\ -c & 1+c^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The characteristic polynomial of $\boldsymbol{B}$ is

$$\det(\lambda \boldsymbol{I} - \boldsymbol{B}) = \lambda^3 - (3+c^2)\lambda^2 + (3+c^2)\lambda - 1,$$

therefore $\boldsymbol{B}$'s invariants are

$$\iota_1(\boldsymbol{B}) = 3 + c^2, \quad \iota_2(\boldsymbol{B}) = 3 + c^2, \quad \iota_3(\boldsymbol{B}) = 1.$$

Then from (147b) we get

$$\boldsymbol{S} = \beta_0(c^2)\boldsymbol{I} + \beta_1(c^2)\boldsymbol{B} + \beta_2(c^2)\boldsymbol{B}^{-1},$$

or in components:

$$\begin{aligned}\begin{pmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{pmatrix} &= \beta_0 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \beta_1 \begin{pmatrix} 1+c^2 & c & 0 \\ c & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \beta_2 \begin{pmatrix} 1 & -c & 0 \\ -c & 1+c^2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \beta_0 + \beta_1 + \beta_2 + c^2\beta_1 & (\beta_1 - \beta_2)c & 0 \\ (\beta_1 - \beta_2)c & \beta_0 + \beta_1 + \beta_2 + c^2\beta_2 & 0 \\ 0 & 0 & \beta_0 + \beta_1 + \beta_2. \end{pmatrix}\end{aligned}$$

From

$$S_{11} - S_{22} = c^2(\beta_1 - \beta_2), \quad S_{12} = (\beta_1 - \beta_2)c$$

we conclude that

$$S_{11} - S_{22} = cS_{12}.$$

This universal relationship between the stress components holds for any shearing experiment, independent of the type of elastic material.

28. Fluids

An *fluid* is a material whose Cauchy stress tensor $S(x,t)$ is determined solely by the gradient of the velocity $v(x,t)$:

$$S(x,t) = \hat{S}\big(\mathbf{grad}\, v(x,t)\big). \tag{148}$$

We begin with the analysis of the restrictions that frame-indifference places on the constitutive function $\hat{S}$.

Consider a fluid body B and a deformation $\phi_t : B \times [0,\infty) \to B_t \subset \mathbb{E}_3$ defined by

$$\phi_t(X) = X_0 + e^{At}(X - X_0) \tag{149}$$

for an arbitrarily fixed point $X_0 \in \mathbb{E}_3$ and an arbitrary constant tensor $A \in \mathcal{L}$. Let us study the properties of this deformation. We have:

$$\begin{aligned} \phi_0(X) &= X, & \mathbf{Grad}\, \phi_t(X) &= e^{At}, \\ \dot{\phi}_t(X) &= Ae^{At}(X - X_0), & \mathbf{Grad}\, \dot{\phi}_t(X) &= Ae^{At}. \end{aligned} \tag{150}$$

Let $v(x,t)$ be the corresponding spatial velocity field, that is,

$$v(x,t) = \dot{\phi}_t(X), \quad \text{where } x = \phi_t(X).$$

Applying **Grad** to this, we get via the chain rule

$$\mathbf{grad}\, v(x,t)\ \mathbf{Grad}\, \phi_t(X) = \mathbf{Grad}\, \dot{\phi}_t(X), \quad \text{where } x = \phi_t(X).$$

Substituting for $\mathbf{Grad}\, \phi_t$ and $\mathbf{Grad}\, \dot{\phi}_t$ from (150), this becomes $\mathbf{grad}\, v(x,t)e^{At} = Ae^{At}$ which simplifies to:

$$\mathbf{grad}\, v(x,t) = A. \tag{151}$$

This reveals the purpose of introducing that particular motion in (149)—we are subjecting the fluid to a motion with a prescribed velocity gradient.

Now consider two frames moving relative to each other, as in Section 25, whose motions are related through the transformation (127) on page 55. The gradient of velocity in the starred frame is given in (133). Substituting for $\mathbf{grad}\, v(x,t)$ from (151), we get

$$\mathbf{grad}\, v^*(x^*,t) = \dot{Q}Q^T + QAQ^T. \tag{152}$$

Let us observe that the tensor $\dot{Q}(t)Q(t)^T$ that appears above is skew-symmetric. This is because $Q(t)Q(t)^T = I$, therefore $\dot{Q}(t)Q(t)^T + Q(t)\dot{Q}(t)^T = \mathbf{0}$.

From the constitutive equation (148) and the calculation in (152) we see that the stress in the starred frame is given by:[9]

$$S^* = \hat{S}(\mathbf{grad}\, v^*) = \hat{S}\big(QAQ^T + \dot{Q}Q^T\big)$$

Therefore, according to the transformation of stress in (136) on page 57, we get:

$$\hat{S}\big(QAQ^T + \dot{Q}Q^T\big) = Q\hat{S}(A)Q^T \quad \text{for all } A \in \mathcal{L} \text{ and all } Q \in \mathcal{L}^+_{\text{orth}}. \tag{153}$$

[9]Note that both observers apply the same constitutive function, $\hat{S}$, to calculate stress. That's frame-invariance!

Let us split $\boldsymbol{A}$ into symmetric and skew-symmetric parts $\boldsymbol{A} = \boldsymbol{D} + \boldsymbol{W}$, where:

$$\boldsymbol{D} = \frac{1}{2}(\boldsymbol{A} + \boldsymbol{A}^T), \quad \boldsymbol{W} = \frac{1}{2}(\boldsymbol{A} - \boldsymbol{A}^T).$$

Now, let $\boldsymbol{Q}(t) = e^{-\boldsymbol{W}t}$. Then $\boldsymbol{Q}(t)^T = e^{-\boldsymbol{W}^T t} = e^{\boldsymbol{W}t}$. We note that the $\boldsymbol{Q}(t)\boldsymbol{Q}(t)^T = \boldsymbol{I}$, and therefore $\boldsymbol{Q}(t)$ is orthogonal. Furthermore, note that $\dot{\boldsymbol{Q}}(t) = -\boldsymbol{W}e^{-\boldsymbol{W}t}$, and in particular, $\boldsymbol{Q}(0) = \boldsymbol{I}$ and $\dot{\boldsymbol{Q}}(0) = -\boldsymbol{W}$. Substituting this choice of $\boldsymbol{Q}(t)$ in (153) and evaluating the result at $t = 0$ we see that $\hat{S}(\boldsymbol{A} - \boldsymbol{W}) = \hat{S}(\boldsymbol{A})$. But since $\boldsymbol{A} - \boldsymbol{W} = \boldsymbol{D}$, we conclude that:

$$\hat{S}(\boldsymbol{A}) = \hat{S}(\boldsymbol{D}), \quad \text{for all } \boldsymbol{A} \in \mathcal{L},$$

that is, $\hat{S}(\boldsymbol{A})$ depends only on the symmetric part of $\boldsymbol{A}$.

Going back to (153), we replace the arguments of $\hat{S}$ on both sides by their symmetric parts. The argument on the left hand side reduces to $\boldsymbol{QDQ}^T$ since $\dot{\boldsymbol{Q}}(t)\boldsymbol{Q}(t)^T$ is skew-symmetric as noted above. We conclude that:

$$\hat{S}(\boldsymbol{QDQ}^T) = \boldsymbol{Q}\hat{S}(\boldsymbol{D})\boldsymbol{Q}^T \quad \text{for all } \boldsymbol{D} \in \mathcal{L}_{\text{sym}} \text{ and all } \boldsymbol{Q} \in \mathcal{L}^+_{\text{orth}},$$

and thus, $\hat{S} : \mathcal{L}_{\text{sym}} \to \mathcal{L}_{\text{sym}}$ is an isotropic function in the sense defined in Section 16. Then according to Theorem 13 we have:

$$\boldsymbol{S} = \hat{S}(\mathbf{grad}\,\boldsymbol{v}) = \hat{S}(\boldsymbol{D}) = \alpha_0\boldsymbol{I} + \alpha_1\boldsymbol{D} + \alpha_2\boldsymbol{D}^2, \quad \text{where } \boldsymbol{D} = \frac{1}{2}\big(\mathbf{grad}\,\boldsymbol{v} + (\mathbf{grad}\,\boldsymbol{v})^T\big), \tag{154}$$

and where the scalar coefficients α_i, $i = 0, 1, 2$, are functions of the invariants of $\boldsymbol{D}$.

Remark 28. Although this result bears some resemblance to the constitutive equation (147a) for elastic materials, note that (147a) was obtained under the assumption of the isotropy of the material while we made no assumption on isotropy in deriving (154).

29. Internal constraints

Up to this point in these notes, we have considered materials that are capable of undergoing arbitrary motions $\boldsymbol{\phi}_t$ as long as the mapping $\boldsymbol{\phi}_t$ is one-to-one and the deformation gradient $\boldsymbol{F}(\boldsymbol{X}, t) = \mathbf{Grad}\,\boldsymbol{\phi}_t(\boldsymbol{X})$ is such that $\det \boldsymbol{F}(\boldsymbol{X}, t) > 0$ at all $\boldsymbol{X}$ and t. Realistic models of some materials, however, call for further restrictions on the class of deformations that they can sustain. The most prevalent of these are *incompressible materials* which are only capable of motions that do not change volume locally. We know that the deformation gradient $\boldsymbol{F}(\boldsymbol{X}, t)$ expresses expansion/contraction factor of volume at the point $\boldsymbol{X}$, therefore the motion of an incompressible material necessarily satisfies the additional constraint of $\det \boldsymbol{F}(\boldsymbol{X}, t) = 1$.

Other types of constraints are possible. For instance, automobile tires have nylon or steel cords embedded within rubber to stabilize the ride and enhance durability. As another example, a sheet of fabric can be be deformed easily in many ways, but the threads that make up the fabric are practically unstretchable.

The limitation of a material to deform in certain ways is called that material's *internal constraint*.[10] The internal constraint is expressed as $\zeta(\boldsymbol{F}) = 0$, where ζ is a scalar-valued

[10] A material may have more than one type of constraint, for instance it may be incompressible and at the same time be unstretchable along embedded fibers. To simplify the exposition, here we limit the discussion single constraints.

function of the deformation gradient $\boldsymbol{F}$. The frame-invariance of the constrain implies that

$$\zeta(\boldsymbol{Q}\boldsymbol{F}) = \zeta(\boldsymbol{F}) \quad \text{for all } \boldsymbol{F} \in \mathcal{L}^+ \text{ and } \boldsymbol{Q} \in \mathcal{L}^+_{\text{orth}}.$$

This may be simplified by replacing $\boldsymbol{F}$ by the right polar decomposition $\boldsymbol{R}\boldsymbol{U}$, and picking $\boldsymbol{Q} = \boldsymbol{R}^T$, whence

$$\zeta(\boldsymbol{F}) = \zeta(\boldsymbol{Q}\boldsymbol{F}) = \zeta(\boldsymbol{R}^T\boldsymbol{R}\boldsymbol{U}) = \zeta(\boldsymbol{U}) = \zeta(\boldsymbol{C}^{1/2}) \stackrel{\text{def}}{=} \lambda(\boldsymbol{C}),$$

where $\boldsymbol{C}$ is the right Cauchy–Green strain tensor. We conclude that the most general internal constraint has the form

$$\lambda(\boldsymbol{C}) = 0. \tag{155}$$

A deforming body resits forces that attempt to violate its internal constraint by developing internal reaction forces exhibited by a stress $\boldsymbol{N}$ which we call the *constraint reaction stress*, or just *reaction stress* for short. It can be shown (but we won't do this right now) that the reaction stress $\boldsymbol{N}$ and the rate of strain $\boldsymbol{D}$ (see (134) on page 56) are orthogonal in the scalar product of of the space $\mathcal{L}$ of second order tensors (see (26) on page 14), that is

$$\boldsymbol{N} : \boldsymbol{D} = 0. \tag{156}$$

This, along with (155), enables us to characterize $\boldsymbol{N}$. We state this as

Theorem 29. *The the constraint reaction stress $\boldsymbol{N}$ of a material with the internal constraint* (155) *is related to the deformation gradient $\boldsymbol{F}$ through*

$$\boldsymbol{N} = \alpha \boldsymbol{F}\, \lambda_C\, \boldsymbol{F}^T, \tag{157}$$

where α is a scalar, $\lambda_C = D\lambda(\boldsymbol{C})$ is the derivative of λ, and $\boldsymbol{C}$ is the right Cauchy–Green tensor corresponding to $\boldsymbol{F}$.

Proof. Let $\boldsymbol{C} = c_{ij}\, \boldsymbol{e}_i \otimes \boldsymbol{e}_j$ be the component representation of $\boldsymbol{C}$ in some frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$. We apply the chain rule of differentiation to calculate the time derivative of $\lambda(\boldsymbol{C})$:

$$\dot{\lambda}(\boldsymbol{C}) = \frac{\partial \lambda(c_{ij}\, \boldsymbol{e}_i \otimes \boldsymbol{e}_j)}{\partial c_{ij}}\, \dot{c}_{ij}.$$

According to (77) on page 35, the factors $\frac{\partial \lambda(c_{ij}\, \boldsymbol{e}_i \otimes \boldsymbol{e}_j)}{\partial c_{ij}}$ in the equation above are the components of the tensor λ_C that expresses the derivative of $\lambda(\boldsymbol{C})$. The factors $\dot{c}_{ij}$ are clearly the components of the tensor $\dot{\boldsymbol{C}}$. Then by (27), the time derivative of λ may be expressed as

$$\dot{\lambda}(\boldsymbol{C}) = \lambda_C : \dot{\boldsymbol{C}}. \tag{158}$$

In Exercise 39 you will show that $\dot{\boldsymbol{C}} = 2\boldsymbol{F}^T\boldsymbol{D}\boldsymbol{F}$, where $\boldsymbol{D}$ is the rate of strain. Therefore

$$\begin{aligned} \dot{\lambda}(\boldsymbol{C}) = 2\lambda_C : \boldsymbol{F}^T\boldsymbol{D}\boldsymbol{F} &\stackrel{\text{by (26)}}{=} 2\,\mathrm{tr}\Big((\boldsymbol{F}^T\boldsymbol{D}\boldsymbol{F})^T \lambda_C \Big) = 2\,\mathrm{tr}\Big(\boldsymbol{F}^T\boldsymbol{D}\boldsymbol{F}\lambda_C \Big) = 2\,\mathrm{tr}\Big(\big(\boldsymbol{F}^T\boldsymbol{D}\big)\big(\boldsymbol{F}\lambda_C\big) \Big) \\ &\stackrel{\text{Exer (12)}}{=} 2\,\mathrm{tr}\Big(\big(\boldsymbol{F}\lambda_C\big)\big(\boldsymbol{F}^T\boldsymbol{D}\big) \Big) = 2\,\mathrm{tr}\Big(\big(\boldsymbol{F}\lambda_C\boldsymbol{F}^T\big)\boldsymbol{D} \Big) \stackrel{\text{by (26)}}{=} 2\big(\boldsymbol{F}\lambda_C\boldsymbol{F}^T\big) : \boldsymbol{D}. \end{aligned}$$

It follows that the constraint $\lambda(\boldsymbol{C}) = 0$ is equivalent to the statement that the tensors $\boldsymbol{F}\lambda_C\boldsymbol{F}^T$ and $\boldsymbol{D}$ are orthogonal in the inner product of the space of second order tensors $\mathcal{L}$. But according to (156), the reaction stress $\boldsymbol{N}$ is also orthogonal to $\boldsymbol{D}$. Since this holds for all strain rates $\boldsymbol{D}$, it follows that $\boldsymbol{N}$ is a multiple of $\boldsymbol{F}\lambda_C\boldsymbol{F}^T$. □

Corollary 7. *The constraint reaction stress $\boldsymbol{N}$ of an incompressible material is a multiple of identity, that is, $\boldsymbol{N} = -p\boldsymbol{I}$. The factor p manifests itself as pressure.*

Proof. The deformation gradient $\boldsymbol{F}$ of a motion that preserves the local volume satisfies $\det \boldsymbol{F} = 1$, therefore $\det \boldsymbol{C} = \det(\boldsymbol{F}^T\boldsymbol{F}) = 1$. Thus, according to (155), the function λ is given by

$$\lambda(\boldsymbol{C}) = \det \boldsymbol{C} - 1.$$

Then by (79) on page 36 we have

$$\dot{\lambda}(\boldsymbol{C}) = (\det \boldsymbol{C})\ \operatorname{tr}(\boldsymbol{C}^{-1}\dot{\boldsymbol{C}}) = (\det \boldsymbol{C})\, \boldsymbol{C}^{-1} : \dot{\boldsymbol{C}} = \boldsymbol{C}^{-1} : \dot{\boldsymbol{C}}$$

since $\det \boldsymbol{C} = 1$. By comparison with (158), we conclude that

$$\lambda_C = \boldsymbol{C}^{-1} = (\boldsymbol{F}^T\boldsymbol{F})^{-1} = \boldsymbol{F}^{-1}\boldsymbol{F}^{-T},$$

whereby (157) reduces to

$$\boldsymbol{N} = \alpha \boldsymbol{F}\boldsymbol{F}^{-1}\boldsymbol{F}^{-T}\boldsymbol{F}^T = \alpha \boldsymbol{I}.$$

The corollary's assertion follows by setting $\alpha = -p$. □

29.1. **Incompressible isotropic elastic materials.** In equations (147) we have the general forms of constitutive equations for (unconstrained) isotropic elastic materials. The left Cauchy–Green strain tensor $\boldsymbol{B}$ in those equations can be any symmetric positive definite tensor. If the material is incompressible, however, then the possible deformations are limited to those with $\det \boldsymbol{B} = 1$ and the resulting stress has an added pressure term $-p\boldsymbol{I}$ according to Corollary 7. The pressure term merges with the $\alpha_0\boldsymbol{I}$ or $\beta\boldsymbol{I}$ terms in those equations and results in the following constitutive equations for incompressible isotropic elastic materials:

$$\boldsymbol{S} = -p\boldsymbol{I} + \alpha_1\boldsymbol{B} + \alpha_2\boldsymbol{B}^2, \tag{159a}$$

$$\boldsymbol{S} = -p\boldsymbol{I} + \beta_1\boldsymbol{B} + \beta_2\boldsymbol{B}^{-1}, \tag{159b}$$

where α_1, α_2, β_1, β_2 are functions of the invariants $\iota_1(\boldsymbol{B})$ and $\iota_2(\boldsymbol{B})$ only, since $\iota_3(\boldsymbol{B}) = \det \boldsymbol{B} = 1$.

29.2. **Incompressible fluids.** in equation (154) we have the general form of the constitutive equation for a general (unconstrained) fluid. The strain rate tensor $\boldsymbol{D}$ can be any symmetric tensor. If the fluid is incompressible, however, then the possible motions are limited to those with divergence-free velocities, i.e., $\operatorname{div} \boldsymbol{v} = 0$ (see Corollary 4 on page 52) and the resulting stress has an added pressure term $-p\boldsymbol{I}$ according to Corollary 7. The pressure term merges with the $\alpha_0\boldsymbol{I}$ in (154) and results in the following constitutive equations for incompressible fluids:

$$\boldsymbol{S} = -p\boldsymbol{I} + \alpha_1\boldsymbol{D} + \alpha_2\boldsymbol{D}^2, \tag{160}$$

where α_1 and α_2, are functions of the invariants $\iota_2(\boldsymbol{D})$ and $\iota_3(\boldsymbol{D})$ only, since

$$\iota_1(\boldsymbol{D}) = \operatorname{tr} \boldsymbol{D} = \frac{1}{2}\Big(\operatorname{tr} \mathbf{grad}\, \boldsymbol{v} + \operatorname{tr}(\mathbf{grad}\, \boldsymbol{v})^T\Big) = \operatorname{div} \boldsymbol{v} = 0.$$

Incompressible materials obeying the constitutive equation (160) are known as *Reiner–Rivlin fluids*, named after Markus Reiner and Ronald Rivlin who independently arrived at that representation in the 1940s.

29.3. **Newtonian fluids and the Navier-Stokes equations.** The special case of (160) where $\alpha_1 = 2\mu$ is a constant and $\alpha_2 = 0$, corresponds to what are called *Newtonian fluids*. All other fluids are called *non-Newtonian*. Thus, the constitutive equation of a Newtonian fluid is

$$\boldsymbol{S} = -p\boldsymbol{I} + 2\mu\boldsymbol{D}. \tag{161}$$

The coefficient μ which is the sole characteristic of a Newtonian fluid is called the fluid's *viscosity*.[11] Let us calculate the divergence of $\boldsymbol{S}$. The divergence of the pressure term $-p\boldsymbol{I}$ may be calculated by applying the result of Exercise 24, whereby

$$\operatorname{div}(-p\boldsymbol{I}) = \boldsymbol{I}\,\mathbf{grad}(-p) - p\,\mathbf{grad}\,\boldsymbol{I} = -\,\mathbf{grad}\,p.$$

As to the divergence of $\boldsymbol{D}$, we recall the definition of the Laplacian $\Delta\boldsymbol{v} = \operatorname{div}\mathbf{grad}\,\boldsymbol{v}$ of a vector field in (72). Additionally, in Exercise 26 you will show that $\operatorname{div}\big((\mathbf{grad}\,\boldsymbol{v})^T\big) = \mathbf{grad}\operatorname{div}\boldsymbol{v}$. Therefore

$$\operatorname{div}\boldsymbol{D} = \frac{1}{2}\Big(\operatorname{div}(\mathbf{grad}\,\boldsymbol{v}) + \operatorname{div}\big((\mathbf{grad}\,\boldsymbol{v})^T\big)\Big) = \frac{1}{2}\Big(\Delta\boldsymbol{v} + \mathbf{grad}\operatorname{div}\boldsymbol{v}\Big) = \frac{1}{2}\Delta\boldsymbol{v},$$

where the last step is due to $\operatorname{div}\boldsymbol{v} = 0$. We conclude that

$$\operatorname{div}\boldsymbol{S} = \mu\Delta\boldsymbol{v} - \mathbf{grad}\,p,$$

and therefore, the equation of motion (102) results in

$$\rho\dot{\boldsymbol{v}} = \mu\Delta\boldsymbol{v} - \mathbf{grad}\,p + \rho\boldsymbol{b},$$

where $\dot{\boldsymbol{v}}$ is the material derivative of the velocity field $\boldsymbol{v}$. This may be expressed in terms of spatial derivatives according to (117) on page 49. We conclude that the equations of motion of an Newtonian fluid are

$$\rho\Big(\frac{\partial\boldsymbol{v}}{\partial t} + (\mathbf{grad}\,\boldsymbol{v})\boldsymbol{v}\Big) = \mu\Delta\boldsymbol{v} - \mathbf{grad}\,p + \rho\boldsymbol{b}, \tag{162a}$$

$$\operatorname{div}\boldsymbol{v} = 0. \tag{162b}$$

This is the well-known *Navier–Stokes system of equations* of the motion of a Newtonian fluid. In the three-dimensional space, the vector equation (162a) and the scalar equation (162b) constitute a set of four equations in the four unknowns consisting of the three components of the velocity vector $\boldsymbol{v}$ and the scalar pressure p.

30. Fluids: Simple shear

Consider the steady-state shearing motion of Reiner–Rivlin fluid in the Cartesian coordinate system according to $\boldsymbol{x} = \boldsymbol{\phi}_t(\boldsymbol{X})$ defined by

$$x_1 = X_1 + cX_2t,$$
$$x_2 = X_2,$$
$$x_3 = X_3,$$

where c is a constant. This is intended to model the shearing of a fluid filling the gap between the parallel places $x_2 = 0$ and $x_2 = h$, where the lower plate is stationary while the upper plate slides in the x_1 direction with steady velocity U. See Figure 14. From the description of the equations of motion above it follows that $c = U/h$. The material

[11] One of the earliest statements regarding viscosity, i.e., internal friction in fluids, occurs in Book II of Newton's *Principia*, where in the modeling of fluids as continuous media, he hypothesizes: "the resistance arising from the want of slipperiness in the parts of a fluid is, other things being equal, proportional to the velocity with which the parts of the fluid separate one another." (Passage quoted from page 447 of Truesdell [23].).

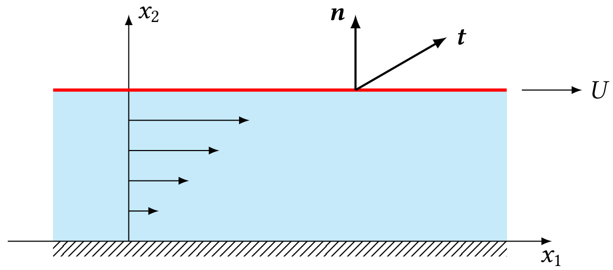


FIGURE 14. The top plate slides horizontally at a steady speed U, while the bottom plate is stationary, shearing the fluid that fills the space between the two. The traction $\boldsymbol{t}$ on the top place is not necessarily horizontal.

representation of the velocity vector $\boldsymbol{V} = \frac{\partial}{\partial t}\boldsymbol{\phi}_f$ relative to the Cartesian coordinates is $\boldsymbol{V}(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3) = \langle cX_2, 0, 0\rangle$. But $x_2 = X_2$, and therefore the spatial representation of the velocity field is $\boldsymbol{v}(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) = \langle cx_2, 0, 0\rangle$. The we calculate the velocity gradient, the rate of strain, and its square

$$\operatorname{\mathbf{grad}} \boldsymbol{v} = \begin{pmatrix} 0 & c & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \boldsymbol{D} = \frac{c}{2}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \qquad \boldsymbol{D}^2 = \frac{c^2}{4}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

We see that $\iota_1(\boldsymbol{D}) = \operatorname{tr}\boldsymbol{D} = 0$, which indicates that the motion is isochoric, and $\iota_3(\boldsymbol{D}) = \det \boldsymbol{D} = 0$. Moreover, we have $\operatorname{tr}\left(\boldsymbol{D}^2\right) = c^2/2$, and therefore, by applying (24b) we get

$$\iota_2(\boldsymbol{D}) = \frac{1}{2}\left(\left(\operatorname{tr}\boldsymbol{D}\right)^2 - \operatorname{tr}\left(\boldsymbol{D}^2\right)\right) = -\frac{c^2}{4}.$$

It follows that the coefficients α_1 and α_2 in (160) are functions of c^2, and therefore the matrix representation of the stress tensor is given by

$$\boldsymbol{S} = -p\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \alpha_1\left(c^2\right)\frac{c}{2}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \alpha_2\left(c^2\right)\frac{c^2}{4}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Letting $\kappa_1 = \alpha_1\left(c^2\right)\frac{c}{2}$ and $\kappa_2 = \alpha_2\left(c^2\right)\frac{c^2}{4}$, this takes the form

$$\boldsymbol{S} = -p\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} \kappa_2 & \kappa_1 & 0 \\ \kappa_1 & \kappa_2 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Note that κ_1 is an odd function of c while κ_2 is an even function of c.

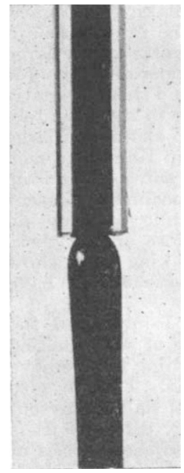
Remark 29. If the pressure p is constant throughout the fluid, the stress calculated above is constant (both in space and time), and therefore the equation of motion (102) is satisfied with zero body force.

Remark 30. The normal vector to the sliding plate is $\boldsymbol{n} = \langle 0, 1, 0\rangle$, therefore the traction on the plate is

$$\boldsymbol{t} = \boldsymbol{S}\boldsymbol{n} = \begin{pmatrix} \kappa_1 \\ -p + \kappa_2 \\ 0 \end{pmatrix}.$$

The traction's x_1 component, κ_1, is necessary to overcome the fluid's viscosity and maintain the plate's rightward motion. The traction's x_2 component, $-p+\kappa_2$, is an interesting feature. It indicates that a vertical force (beyond what is necessary to counteract the pressure) is needed to keep the plate at a constant elevation. That extra force, κ_2, which is entirely due to the third term in the Reiner–Rivlin constitutive equation, is notably absent in the case of a Newtonian fluid.

31. Fluids: Poiseuille flow



In the classical Poiseuille[12] flow, a Newtonian fluid is pushed down a long pipe of circular cross-section. The motion, described through the Navier–Stokes equations, leads to a parabolic velocity profile, a pressure that's constant over the tube's cross-sections, and drops linearly along the tube's axis. At the tube's outlet, the fluid's pressure equals atmospheric pressure, and the exiting fluid jet continues its motion while maintaining its cylindrical shape.

Merrington in the 1943 article [9] referring to "certain anomalous liquids" such as rubber solutions, pointed to a distinct swelling at the outflow of a tube through which a pressurized non-Newtonian fluid is forced to move. The adjacent photograph, illustrating the phenomenon, is taken from that article. The article does not attempt to provide quantitative justification for the phenomenon. The swelling is ascribed to the release of certain elastic strains developed under the pressure in the fluid.

Rathna's 1960 article [11] shows that the swelling can be explained in the context of Reiner–Rivlin fluids. In this section we will present a special case of her calculations in a self-contained form.

31.1. **The formulation of the problem.** Consider a steady flow of an incompressible Reiner–Rivlin fluid, whose constitutive equation is given in (160), through a vertical tube of radius R. We install a cylindrical (r, θ, z) coordinate system with the z axis aligned with the tube's axis and pointing up, and where r measures the perpendicular distance from that axis. We assume axial symmetry, and therefore none of the variables depends on the angular coordinate θ. We look for motion with the velocity $\boldsymbol{v} = \langle u_r, u_\theta, u_z \rangle$ parallel to the z axis. The velocity components in the (r, θ, z) coordinates are

$$u_r = 0, \quad u_\theta = 0, \quad u_z = \phi(r),$$

where, as we will find out, u_z is negative as the fluid moves downward. With the help of the formulas in Appending A, we calculate

$$\mathbf{grad}\, \boldsymbol{v} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \phi'(r) & 0 & 0 \end{pmatrix},$$

where a prime indicates derivative with respect to r. Then the rate of strain $\boldsymbol{D}$ and its square are:

$$\boldsymbol{D} = \frac{1}{2}\phi'(r)\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \boldsymbol{D}^2 = \frac{1}{4}\phi'(r)^2\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

We note that $\operatorname{tr} \mathbf{grad}\, \boldsymbol{v} = \operatorname{div} \boldsymbol{v} = 0$, indicating that the flow is isochoric, as it should be in an incompressible fluids. The principal invariants of $\boldsymbol{D}$ are $\iota_1(\boldsymbol{D}) = \operatorname{tr} \boldsymbol{D} = 0$, $\iota_3(\boldsymbol{D}) = \det \boldsymbol{D} = 0$, and with the help of equation (24b)

$$\iota_2(\boldsymbol{D}) = \frac{1}{2}\Big(\big(\operatorname{tr} \boldsymbol{D}\big)^2 - \operatorname{tr}\big(\boldsymbol{D}^2\big)\Big) = -\frac{1}{4}\phi'(r)^2.$$

[12] Pronounced **Poa'-zo-ee-ə**.

It follows that the coefficients α_1 and α_2 in the constitutive equation (160) are functions of $\phi'(r)^2$, and therefore the stress is

$$S = -p(r,z)\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix} + \frac{1}{2}\phi'(r)\alpha_1\big(\phi'(r)^2\big)\begin{pmatrix}0 & 0 & 1\\ 0 & 0 & 0\\ 1 & 0 & 0\end{pmatrix} + \frac{1}{4}\phi'(r)^2\alpha_2\big(\phi'(r)^2\big)\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 1\end{pmatrix}.$$

To simplify the notation, we introduce

$$\kappa_1(r) = \frac{1}{2}\phi'(r)\,\alpha_1\big(\phi'(r)^2\big), \quad \kappa_2(r) = \frac{1}{4}\phi'(r)^2\,\alpha_2\big(\phi'(r)^2\big), \tag{163}$$

whereby

$$S = -p(r,z)\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix} + \begin{pmatrix}\kappa_2(r) & 0 & \kappa_1(r)\\ 0 & 0 & 0\\ \kappa_1(r) & 0 & \kappa_2(r)\end{pmatrix}. \tag{164}$$

Then we apply the formulas (232) and (233) to calculate

$$(\mathbf{grad}\,\boldsymbol{v})\boldsymbol{v} = \begin{pmatrix}0\\0\\0\end{pmatrix}, \qquad \operatorname{div} S = -\begin{pmatrix}\frac{\partial p}{\partial r}\\ 0\\ \frac{\partial p}{\partial z}\end{pmatrix} + \begin{pmatrix}\kappa_2' + \frac{1}{r}\kappa_2\\ 0\\ \kappa_1' + \frac{1}{r}\kappa_1\end{pmatrix}.$$

Then, observing that $\kappa' + \frac{1}{r}\kappa = \frac{1}{r}(r\kappa)'$, the equation of motion (102) takes the form

$$0 = -\frac{\partial}{\partial r}p(r,z) + \frac{1}{r}\big(r\kappa_2\big)' \tag{165a}$$

$$0 = 0, \tag{165b}$$

$$0 = -\frac{\partial}{\partial z}p(r,z) + \frac{1}{r}\big(r\kappa_1\big)' - \rho g, \tag{165c}$$

where ρ is the fluid's (constant) density, and g is the gravitational acceleration.

Now we proceed to solve this set of partial differential equations. Differentiating (165a) with respect to z and (165c) with respect to r and subtracting the results we obtain $\left(\frac{1}{r}\big(r\kappa_1\big)'\right)' = 0$, whence

$$\frac{1}{r}\big(r\kappa_1\big)' = c, \tag{166}$$

for some constant c. Then $\big(r\kappa_1\big)' = cr$, and therefore $r\kappa_1 = \frac{c}{2}r^2 + \hat{c}$, where $\hat{c}$ is another constant. Evaluating this at $r = 0$ we see that $\hat{c} = 0$. We conclude that

$$\kappa_1(r) = \frac{c}{2}r. \tag{167}$$

Plugging this into (165c), we see that the equation reduces to $\frac{\partial}{\partial z}p(r,z) = c - \rho g$, and therefore

$$p(r,z) = (c - \rho g)z + f(r) \tag{168}$$

for some function $f(r)$. Then according to (165a) we have

$$f'(r) = \frac{1}{r}\big(r\kappa_2\big)'. \tag{169}$$

31.2. **Introducing the details of the constitutive equation.** This is as far as we can go without specifying the details of the fluid's properties. To continue, let us choose

$$\alpha_1(\xi) = c_1\xi^n, \quad \alpha_2(\xi) = c_2,$$

where n, c_1, and c_2 are positive constants.[13] Recalling the definitions of κ_1 and κ_2 in (163) $\phi'(r)\alpha_1\big(\phi'^2(r)\big) = cr$, we have

$$\kappa_1(r) = \frac{1}{2}\phi'(r)c_1\big(\phi'(r)^2\big)^n = \frac{c_1}{2}\phi'(r)^{2n+1}, \tag{170a}$$
$$\kappa_2(r) = \frac{c_2}{4}\phi'(r)^2. \tag{170b}$$

In particular, from (167) and (170a) we get $\frac{c}{2}r = \frac{c_1}{2}\phi'(r)^{2n+1}$, whence $\phi'(r)^{2n+1} = \frac{c}{c_1}r$ which we rewrite as

$$\phi'(r) = Ar^{m-1}, \tag{171}$$

where we have set

$$m = 1 + \frac{1}{2n+1}, \quad A = \Big(\frac{c}{c_1}\Big)^{m-1}.$$

Integrating (171) we get $\phi(r) = \frac{A}{m}r^m + K$. The constant of integration, K is determined by imposing the no-slip boundary condition $\phi(R) = 0$. We conclude that

$$\phi(r) = -\frac{A}{m}(R^m - r^m). \tag{172}$$

Recalling that $\phi(r)$ is the vertical component of the velocity, we see that the velocity is negative, and therefore the fluid moves downward, as intended.

Inserting for $\phi'(r)$ from (171) into (170b), we get

$$\kappa_2(r) = \frac{c_2}{4}A^2r^{2(m-1)}, \tag{173}$$

and therefore, in view of (169) we have

$$f'(r) = \frac{(2m-1)c_2}{4}A^2r^{2m-3},$$

which leads to

$$f(r) = \frac{(2m-1)c_2}{8(m-1)}A^2r^{2(m-1)} + K, \tag{174}$$

and then, from (168)

$$p(r,z) = \frac{(2m-1)c_2}{8(m-1)}A^2r^{2(m-1)} + (c - \rho g)z + K. \tag{175}$$

The integration constant K is to be determined.

[13] The parameter α_2 need not be a constant. The more general case of $\alpha_2(\xi) = c_0 + c_1\xi + c_2\xi^2 + \cdots$ is treated in [11].

31.3. **The stress components.** Having determined expressions for $p(r,z)$ in (175) and $\kappa_1(r)$ and $\kappa_2(r2)$ in (167) and (173), we evaluate the components of stress in (164) and obtain:

$$S_{rr} = S_{zz} = -p(r,z) + \kappa_2(r) = -\frac{c_2}{8(m-1)} A^2 r^{2(m-1)} - (c - \rho g)z - K, \tag{176a}$$

$$S_{\theta\theta} = -p(r,z) = -\frac{(2m-1)c_2}{8(m-1)} A^2 r^{2(m-1)} - (c - \rho g)z - K, \tag{176b}$$

$$S_{rz} = S_{zr} = \kappa_1(r) = \frac{c}{2} r, \tag{176c}$$

$$S_{r\theta} = S_{\theta r} = S_{\theta z} = S_{z\theta} = 0. \tag{176d}$$

In particular, the traction at the pipe's outlet, where the outward unit normal is $\boldsymbol{n} = \langle 0, 0, -1\rangle$, is determined from $\boldsymbol{t} = \boldsymbol{S}\boldsymbol{n} = \langle -S_{rr}, 0, -S_{zz}\rangle$.

Remark 31. It is significant to note that the vertical component of the traction, that is S_{zz}, calculated in (176a), generally varies with r. This is a characteristic of non-Newtonian fluids. The constitutive equation's D^2 term is absent in a Newtonian fluid, therefore $c_2 = 0$, and the r-dependent term in (176a) drops out.

If p_0 is the atmospheric pressure, then the force exerted by the atmosphere on the pipe's exit is $\pi R^2 p_0$. This is counteracted by the (variable) normal traction on the fluid's surface at the exit. The balance of forces is expressed through

$$\int_0^{2\pi}\int_0^R S_{zz}\, r\, dr\, d\theta = -\pi R^2 p_0.$$

Plugging S_{zz} from (176a) into this, setting $z = 0$, and carrying out the integration, we obtain the value of K:

$$K = -p_0 - \frac{c2}{8m(m-1)} A^2 R^{2(m-1)},$$

and thus we arrive at the final representation for

$$S_{rr} = S_{zz} = \frac{c_2}{8(m-1)} A^2 \left(\frac{1}{m} R^{2(m-1)} - r^{2(m-1)}\right) - (c - \rho g)z + p_0. \tag{177}$$

31.4. **Why does the fluid swell at the exit?** The S_{rr} calculated in (177) is the normal traction, that is, pressure, that the fluid exerts on the tube's walls. Note that S_{rr} varies with z. At the outlet we we have $z = 0$, and on the wall we have $r = R$, therefore pressure exerted on the tube's walls, let's call it P, is

$$P = S_{rr}\big|_{z=0,r=R} = p_0 - \frac{c_2}{8m} A^2 R^{2(m-1)}.$$

and therefore

$$P - p_0 = -\frac{c_2}{8m} A^2 R^{2(m-1)}.$$

I must have made a sign error somewhere, because this is supposed to be

$$P - p_0 = \frac{c_2}{8m} A^2 R^{2(m-1)}.$$

The excess, $P - p_0$, over the atmospheric pressure causes the swelling of the fluid as observed in Merringon's experiment [9]. The excess is proportional to c_2. In a Newtonian fluid, where we have $c_2 = 0$, there is no such swelling. Rathna [11] remarks that the Merringon's explanation of the cause of swelling as being due to elastic effects is not warranted. In this section we have seen that Merringon's effect can be explained solely in the context of non-Newtonian fluids.

31.5. **The mass flux.** The mass flux (mass of fluid passing through the tube per unit time) is

$$M = -\int_0^{2\pi}\int_0^R \rho\phi(r)\, r\, dr\, d\theta = \frac{\pi\rho A R^{m+2}}{m+2}.$$

Following Rathna [11], we introduce

$$\Gamma = \frac{M}{\rho\pi R^2} = \frac{A}{m+2}R^m,$$

which measures the volume of fluid passing through a unit cross-section of the tube, per unit time. We see that

$$\left(\frac{\Gamma}{R}\right)^2 = \frac{1}{(m+2)^2}A^2 R^{2(m-1)},$$

and therefore the excess pressure at the exit may be expressed as

$$P - p_0 = \frac{(m+1)^2 c_2}{8m}\left(\frac{\Gamma}{R}\right)^2.$$

We repeat Rathna's conclusion that the excess pressure, and therefore the amount of swelling, increases with higher flux and smaller pipe radius.

32. Fluids: Couette flow

This section on the Couette flow of non-Newtonian fluids is based on parts of Serrin's article [22]. The intent is to explain the *Weissenberg effect* [24] which, among other things highlights the tendency of a non-Newtonian fluid to climb a spinning rod immersed in it.

As in Serrin's work, we take the fluid's constitutive equation as $\boldsymbol{S} = -p\boldsymbol{I} + \alpha_1\boldsymbol{D} + \alpha_2\boldsymbol{D}^2$, where α_1 and α_2 are positive constants. Rathna [11] and Rathna and Bhatnagar [12] generalize the calculation to a fluids where α_1 and α_2 are certain functions of the second invariant $\iota_2(\boldsymbol{D})$.

32.1. **Solving the equations of motion.** In the cylindrical coordinates (r, θ, z) with the z axis pointing up, consider the concentric cylinders $r = r_1$ and $r = r_2$, $z \geq 0$, where $r_1 < r_2$. The space between the cylinders is filled with an incompressible fluid of prescribed volume V. The inner and outer cylinders are rotated at constant angular velocities ω_1 and ω_2, respectively. We wish to study the fluid's motion at steady-state.

Let us look for a solution where the velocity field is $\boldsymbol{v}(r, \theta, z) = r\omega(r)\boldsymbol{e}_\theta$, that is, fluid particles at distance r from the rotation axis move about the axis in horizontal circular paths at a steady angular velocity $\omega(r)$. Then according to (227), the velocity gradient, expressed in the cylindrical coordinates, is

$$\mathbf{grad}\,\boldsymbol{v} = \begin{pmatrix} 0 & -\omega(r) & 0 \\ \big(r\omega(r)\big)' & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and therefore the rate of strain D, and its square are

$$\boldsymbol{D} = \frac{1}{2}\begin{pmatrix} 0 & r\omega'(r) & 0 \\ r\omega'(r) & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \boldsymbol{D}^2 = \frac{1}{4}\begin{pmatrix} r^2\omega'(r)^2 & 0 & 0 \\ 0 & r^2\omega'(r)^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

We see that

$$\iota_1(\boldsymbol{D}) = \operatorname{tr}\boldsymbol{D} = 0, \quad \iota_2(\boldsymbol{D}) = \frac{1}{2}\left[(\operatorname{tr}\boldsymbol{D})^2 - \operatorname{tr}\boldsymbol{D}^2\right] = -frac14r^2\omega'(r)^2, \quad \iota_3(\boldsymbol{D}) = \det\boldsymbol{D} = 0$$

Since $\operatorname{div}\boldsymbol{v} = \operatorname{tr}\boldsymbol{D} = 0$, the velocity field $\boldsymbol{v}$ is consistent with the incompressibility constraint. Moreover, we note that

$$(\operatorname{\mathbf{grad}}\boldsymbol{v})\boldsymbol{v} = \begin{pmatrix} -r\omega(r)^2 \\ 0 \\ 0 \end{pmatrix}.$$

Applying the constitutive equation $\boldsymbol{S} = -p\boldsymbol{I} + \alpha_1\boldsymbol{D} + \alpha_2\boldsymbol{D}^2$, assuming α_1 and α_2 are constants, we calculate the stress

$$\boldsymbol{S} = -p(r,\theta,z)\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \frac{1}{2}\alpha_1\begin{pmatrix} 0 & r\omega'(r) & 0 \\ r\omega'(r) & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \frac{1}{4}\alpha_2\begin{pmatrix} r^2\omega'(r)^2 & 0 & 0 \\ 0 & r^2\omega'(r)^2 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{178}$$

Then we apply the formula (233) to calculate the divergence of $\boldsymbol{S}$:

$$\operatorname{div}\boldsymbol{S} = \begin{pmatrix} -\frac{\partial p}{\partial r} + \alpha_2\left(\frac{1}{4}r^2\omega'(r)^2\right)' \\ -\frac{1}{r}\frac{\partial p}{\partial r} + \alpha_1\frac{1}{r^2}\left(\frac{1}{2}r^3\omega'(r)\right)' \\ -\frac{\partial p}{\partial z} \end{pmatrix},$$

We substitute these into the equation of motion $\rho(\operatorname{\mathbf{grad}}\boldsymbol{v})\boldsymbol{v} = \operatorname{div}\boldsymbol{S} + \rho\boldsymbol{b}$, where the force per unit mass is $\boldsymbol{b} = -g\boldsymbol{e}_z$. We get:

$$-\rho r\omega(r)^2 = -\frac{\partial p}{\partial r} + \alpha_2\left(\frac{1}{4}r^2\omega'(r)^2\right)', \tag{179a}$$

$$0 = -\frac{1}{r}\frac{\partial p}{\partial \theta} + \alpha_1\frac{1}{r^2}\left(\frac{1}{2}r^3\omega'(r)\right)', \tag{179b}$$

$$0 = -\frac{\partial p}{\partial z} - \rho g. \tag{179c}$$

From (179c) we see that $p(r,\theta,z) = -\rho gz + f(r,\theta)$ where $f(r,\theta)$ is to be determined. Substituting this into (179b) leads to

$$\frac{\partial}{\partial\theta}f(r,\theta) = \alpha_1\frac{1}{r}\left(\frac{1}{2}r^3\omega'(r)\right)',$$

and therefore

$$f(r,\theta) = \alpha_1\frac{1}{r}\left(\frac{1}{2}r^3\omega'(r)\right)'\theta + h(r),$$

where $h(r)$ is to be determined. We note, however, that what we have obtained indicates that $f(r,\theta)$ increases linearly in θ, and therefore it cannot be 2π-periodic, as it should. We conclude that the coefficient of θ is zero, that is

$$f(r,\theta) = h(r), \quad \alpha_1 \frac{1}{r}\left(\frac{1}{2}r^3\omega'(r)\right)' = 0,$$

and therefore $\frac{1}{2}r^3\omega'(r) = A =$ a constant. It follows that $\omega'(r) = 2Ar^{-3}$, and therefore

$$\omega(r) = -Ar^{-2} + B,$$

where B is another constant.

The constants A and B may be determined by applying the boundary conditions $\omega(r_1) = \omega_1$ and $\omega(r_2) = \omega_2$. We get

$$A = \frac{\omega_2 - \omega_1}{r_2^2 - r_1^2} r_1^2 r_2^2, \quad B = \frac{r_2^2\omega_2 - r_1^2\omega_1}{r_2^2 - r_1^2}. \tag{180}$$

We plug the expression for $\omega(r)$ and $p(r,\theta,z) = -\rho g z + h(r)$ into (179a) and obtain

$$h'(r) = \rho r\Big(B - Ar^{-2}\Big)^2 - 4\alpha_2 A^2 r^{-5},$$

whence

$$h(r) = \frac{1}{2}\rho B^2 r^2 - \rho AB \ln r - \frac{1}{2}\rho A^2 r^{-2} + \alpha_2 A^2 r^{-4} + C$$

for some constant C. Thus, we arrive at

$$p(r,\theta,z) = -\rho g z + h(r).$$

32.2. **The analysis of the free surface.** The pressure on the fluid's open surface is the atmospheric pressure p_0. Letting $p = p_0$ in the preceding equation gives the equation of the height of the fluid's free surface z as a function of r:

$$z(r) = \frac{1}{\rho g}\big(h(r) - p_0\big).$$

Then we calculate the slope of the free surface $z'(r) = \frac{1}{\rho g}h'(r)$. For convenience, we rearrange previously calculated $h'(r)$ and express this as

$$z'(r) = \frac{A^2}{g r^5}\left[r^2\left(1 - \frac{B}{A}r^2\right)^2 - \frac{4\alpha_2}{\rho}\right]. \tag{181}$$

The α_2 that enters in this expression is what defined the fluid's non-Newtonian character. When α_2 is zero, the fluid in Newtonian. We see that effect of α_2 in the equation above is to reduce the slope of the free surface. To gain a better understanding, let us consider two special cases.

Case I: The inner cylinder stationary, that is, $\omega_1 = 0$, while $\omega_2 > 0$. Plugging $\omega_1 = 0$ in (180) we see that $B/A = 1/r_1^2$, and therefore (181) takes the form

$$z'(r)\Big|_{\omega_1=0} = \frac{A^2}{g r^5}\left[r^2\left(1 - \frac{r^2}{r_1^2}\right)^2 - \frac{4\alpha_2}{\rho}\right].$$

We see that in the case of a Newtonian fluid, that is, $\alpha_2 = 0$, we have $z'(r_1) = 0$ and $z'(r_2) > 0$, and the fluid's free surface takes the shape depicted in Figure 15(a).

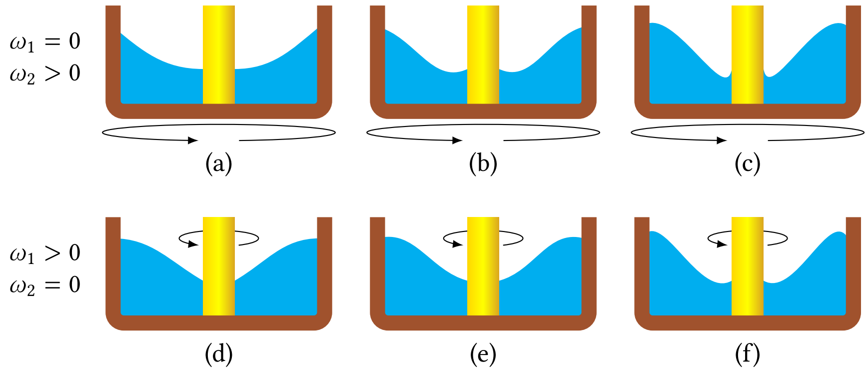


FIGURE 15. In the first row we have a Couette flow where the inner cylinder is stationary while the outer one rotates. The free surface of a Newtonian fluid (subfigure (a)) contacts the inner cylinder horizontally. In a non-Newtonian fluid, the fluid climbs up the inner cylinder (subfigures (b) and (c)). The angle of contact with the outer cylinder may be positive (b) or negative (c). In the second row we have a Couette flow where the outer cylinder is stationary while the inner one rotates. In a Newtonian fluid (subfigure (d)) the fluid is repelled from the inner cylinder. In a non-Newtonian fluid, the fluid may be repelled from (subfigure (e)) or climb up (subfigure (f)) the inner cylinder.

In the case a non-Newtonian fluid, that is $\alpha_2 > 0$, we see that $z'(r_1) < 0$, indicating the that fluid climbs the inner cylinder. The slope at the outer cylinder may be positive or negative, depending on the parameter values. Figures 15(b) and 15(c) depict the two possibilities.

Case II: The outer cylinder stationary, that is, $\omega_2 = 0$, while $\omega_1 > 0$. Plugging $\omega_2 = 0$ in (180) we see that $B/A = 1/r_2^2$, and therefore (181) takes the form

$$z'(r)\Big|_{\omega_2=0} = \frac{A^2}{gr^5}\left[r^2\left(1 - \frac{r^2}{r_2^2}\right)^2 - \frac{4\alpha_2}{\rho}\right].$$

We see that in the case of a Newtonian fluid, that is, $\alpha_2 = 0$, we have $z'(r_1) > 0$ and $z'(r_2) = 0$, and the fluid's free surface takes the shape depicted in Figure 15(d).

In the case a non-Newtonian fluid, that is $\alpha_2 > 0$, the value of $z'(r_1)$ may be positive or negative, depending on the parameter values. The fluid will climb up the inner cylinder if $z'(r_1) < 0$, that is,

$$z'(r_1)\Big|_{\omega_2=0} = \frac{A^2}{gr_1^5}\left[r_1^2\left(1 - \frac{r_1^2}{r_2^2}\right)^2 - \frac{4\alpha_2}{\rho}\right] < 0,$$

or equivalently,

$$r_1\left(1 - \frac{r_1^2}{r_2^2}\right) < 2\sqrt{\frac{\alpha_2}{\rho}},$$

which holds if r_1 is sufficiently small. If $r_1 \ll r_2$, this reduces to the particularly simple criterion

$$r_1 < 2\sqrt{\frac{\alpha_2}{\rho}}.$$

Figures 15(e) and 15(f) depict the two possibilities.

32.3. **What is wrong with the previous analysis?** We began the previous subsection by stating that

> The pressure on the fluid's open surface is the atmospheric pressure p_0. Letting $p = p_0$ in the preceding equation gives the equation of the height of the fluid's free surface z as a function of r.

That statement is not true. The correct statement is that at a point of the free surface with the outward unit normal $\boldsymbol{n}$, we should have $-p_0\boldsymbol{n} = \boldsymbol{S}\boldsymbol{n}$. To analyze the consequences of this, let the free surface be given by the equation $z = \phi(r)$. Then $\boldsymbol{n} = c\big\langle -\phi'(r), 0, 1\big\rangle$, with $c = 1/\sqrt{1+\phi'(r)^2}$, is the outward unit normal to the free surface, expressed in components along the cylindrical coordinate system. In view of the expression (178) calculated for the stress tensor, the equation $-p_0\boldsymbol{n} = \boldsymbol{S}\boldsymbol{n}$ takes the form

$$\begin{pmatrix} \Big(-p + \frac{1}{4}\alpha_2 r^2\omega'(r)^2\Big)\big(-\phi'(r)\big) \\ \Big(\frac{1}{2}\alpha_1 r\omega'(r)\Big)\big(-\phi'(r)\big) \\ p \end{pmatrix} = -p_0 \begin{pmatrix} -\phi'(r) \\ 0 \\ 1 \end{pmatrix}.$$

Since $\phi'(r)$ and $\omega'(r)$ are not identically zero, the middle equation requires that $\alpha_1 = 0$. The third equation requires that $p = p_0$, and therefore the first equation reduces to $\alpha_2 = 0$. Thus, we have reached an impasse. No fluid satisfies the balance of tractions on the free surface!

Where have we gone wrong? The problem lies at the very initial assumption that the velocity field is of the form $\boldsymbol{v}(r,\theta,z) = r\omega(r)\boldsymbol{e}_\theta$. This subsection's calculation indicates that that assumption is untenable if the fluid has free surface open to the atmosphere. In that connection, Serrin [22] remarks that "as long as the free surface remains relatively horizontal (i.e. at relatively low speeds of rotation) the discrepancy should not be serious". Unfortunately in practically all experimental results that demonstrate the Weissenberg effect, the fluid climbs dramatically up the inner cylinder, rendering the "relatively horizontal" assumption inapplicable. The articles [11, 12] make no mention of this calculation's limitation at all. Let us note that the issue raised here is not specific to non-Newtonian fluids. The problem persists even in the case of Newtonian fluids.

The proper determination of the fluid's free surface requires the analysis of a velocity field of the form $\boldsymbol{v}(r,z) = \eta(r,z)\boldsymbol{e}_r + r\omega(r,z)\boldsymbol{e}_\theta + \zeta(r,z)\boldsymbol{e}_z$. Even in the case of a Newtonian fluid where $\alpha_1 = 2\mu$ and $\alpha_2 = 0$, the equations of motion would be a nontrivial system

PDEs

$$\begin{aligned}
\rho\Big(\eta\eta_r + \zeta\eta_z - r\omega^2\Big) &= \mu\bigg(2\Big(\frac{1}{r}\big(r\eta\big)_r\Big)_r + \eta_{zz} + \zeta_{rz}\bigg) - p_r,\\
\rho\Big(\frac{1}{r}(r^2\omega)_r\eta + r\omega_z\zeta\Big) &= \mu\Big(\frac{1}{r^2}(r^3\omega_r)_r + r\omega_{zz}\Big),\\
\rho\Big(\eta\zeta_r + \zeta\zeta_z\Big) &= \mu\Big(\frac{1}{r}\big(r\zeta_r + r\eta_z\big)_r + 2\zeta_{zz} + \eta_{rz}\Big) - p_z - \rho g,\\
\frac{1}{r}(r\eta)_r + \zeta_z &= 0,
\end{aligned}$$

where an r or z subscript indicates a derivative with respect to that variable. The first three equations are the equations of motion, while the last equation expresses the incompressibility constraint $\operatorname{div} \boldsymbol{v} = 0$. This system of four couple PDEs in the four unknowns η, ω, ζ, p, is to be solved numerically with techniques applicable to free boundary problems. See Cuvelier and Schulkes [4] for a survey of such methods, and Zhang and Babuška [25] regarding convergence and error estimates for a very simple free boundary problem.

33. Elasticity: The spinning cylinder problem

This section is based on pages 158–160 of Chapter 4 of [2]. Also see Green and Zerna [6], page 100.

Consider an incompressible and isotropic elastic body in the form of a cylinder of length L and radius R in its undeformed configuration, and made of a *Mooney rubber*[14] whose constitutive equation is

$$\boldsymbol{S} = \hat{\boldsymbol{S}}(\boldsymbol{F}) = -p\boldsymbol{I} + \big[\alpha + \beta\,\iota_1(\boldsymbol{B})\big]\boldsymbol{B} - \beta\boldsymbol{B}^2,$$

where α and β are positive constants, and p is the pressure due to the constraint reaction stress.

Figure 16 depicts the spinning cylinder. We install a frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ so that $\boldsymbol{e}_1$ is aligned with the cylinder's axis. We spin the cylinder about $\boldsymbol{e}_1$ at a constant angular velocity ω while keeping the frame stationary, and allow the motion to reach steady-state. The cylinder will expand (become fatter) in the radial direction due to centrifugal forces and shrink in the axial direction in order to maintain a constant volume, as the material is incompressible. Let λL be the cylinder's length while spinning. Then the radius will be $\lambda^{-1/2}R$. We wish to calculate the measure of deformation, λ, as a function of the angular velocity of the spin, ω.

We begin with describing the motion $\boldsymbol{\phi}_t$. A material point at the location $\boldsymbol{X} = X_1\boldsymbol{e}_1 + X_2\boldsymbol{e}_2 + X_3\boldsymbol{e}_3$ goes to $\boldsymbol{x} = \boldsymbol{\phi}_t(\boldsymbol{X})$ where $\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + x_3\boldsymbol{e}_3$ is given by

$$\begin{aligned}
x_1 &= \lambda X_1,\\
x_2 &= \lambda^{-1/2}\big(X_2\cos\omega t - X_3\sin\omega t\big),\\
x_3 &= \lambda^{-1/2}\big(X_2\sin\omega t + X_3\cos\omega t\big).
\end{aligned} \tag{182}$$

We see that the deformation $\boldsymbol{\phi}_t(\boldsymbol{X})$ is linear in $\boldsymbol{X}$, and therefore it may be expressed as a tensor, say $\boldsymbol{M}(t)$, acting on $\boldsymbol{X}$, as in $\boldsymbol{\phi}_t(\boldsymbol{X}) = \boldsymbol{M}(t)\,\boldsymbol{X}$. Either by inspecting the equations

[14] Named after Melvin Mooney who was among the early investigators of modern continuum mechanics.

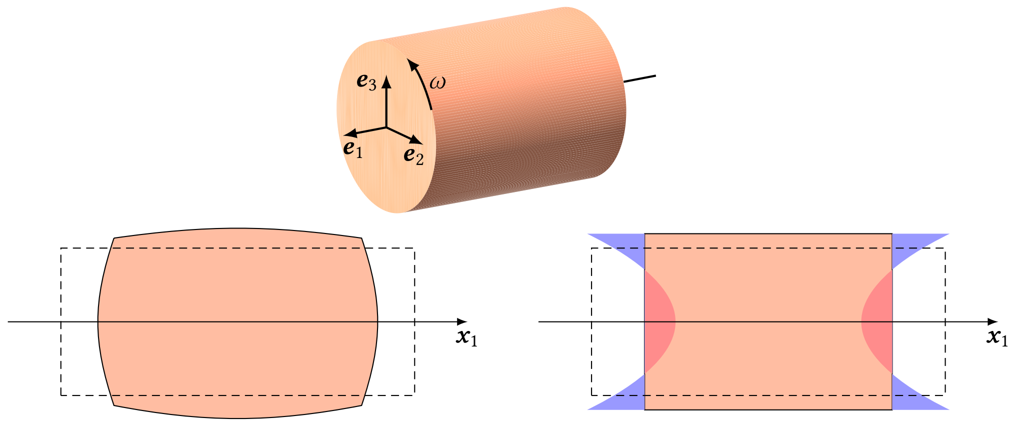


Figure 16. At the top, the rubber cylinder spinning about the $\boldsymbol{e}_1$ axis with angular velocity ω. On the bottom left, the dotted rectangle is the silhouette of the undeformed cylinder, while the filled region depicts the barreled shape of the spinning body when no surface tractions are applied. On the bottom right, self-equilibrated tractions, indicated by their distribution profiles, applied to the spinning cylinder's bases, force it into a cylindrical shape.

above, or by a direct appeal to the representation of the rotation tensor in (45a), we get

$$\boldsymbol{M}(t) = \lambda \boldsymbol{e}_1 \otimes \boldsymbol{e}_1 + \lambda^{-1/2}(\boldsymbol{e}_2 \otimes \boldsymbol{e}_2 + \boldsymbol{e}_3 \otimes \boldsymbol{e}_3)\cos\omega t - \lambda^{-1/2}(\boldsymbol{e}_2 \otimes \boldsymbol{e}_3 - \boldsymbol{e}_3 \otimes \boldsymbol{e}_2)\sin\omega t.$$

Then, since $\mathbf{Grad}\,\boldsymbol{X} = \boldsymbol{I}$, we have $\boldsymbol{F}(\boldsymbol{X}) = \mathbf{Grad}\,\boldsymbol{\phi}_t(\boldsymbol{X}) = \boldsymbol{M}(t)$, and therefore

$$\boldsymbol{B} = \boldsymbol{F}\boldsymbol{F}^T = \lambda^2 \boldsymbol{e}_1 \otimes \boldsymbol{e}_1 + \lambda^{-1}\boldsymbol{e}_2 \otimes \boldsymbol{e}_2 + \lambda^{-1}\boldsymbol{e}_3 \otimes \boldsymbol{e}_3.$$

According to the Spectral Representation Theorem, the coefficients $\lambda^2, \lambda^{-1}, \lambda^{-1}$ are the eigenvalues of $\boldsymbol{B}$, and therefore, by (31a) we conclude that $\iota_1(\boldsymbol{B}) = \lambda^2 + 2\lambda^{-1}$. We substitute this, along with the expression calculated for $\boldsymbol{B}$, into Mooney's constitutive equation and obtain

$$\boldsymbol{S} = \big(-p + \alpha\lambda^{-1} + \beta(\lambda^2 + 2\lambda^{-1})\big)\,\boldsymbol{I} + (\alpha + \beta\lambda^{-1})(\lambda^2 - \lambda^{-1})\,\boldsymbol{e}_1 \otimes \boldsymbol{e}_1.$$

The only spatial variable here is $p = p(\boldsymbol{x})$. Therefore we have

$$\operatorname{div}\boldsymbol{S} = -\,\mathbf{grad}\,p = -\frac{\partial p}{\partial x_1}\boldsymbol{e}_1 - \frac{\partial p}{\partial x_2}\boldsymbol{e}_2 - \frac{\partial p}{\partial x_3}\boldsymbol{e}_3.$$

With the goal of applying the equation of motion (102), we calculate the acceleration by differentiating each of the equations (182) twice with respect to t. We get

$$\ddot{x}_1 = 0, \quad \ddot{x}_2 = -\omega^2 x_2, \quad \ddot{x}_3 = -\omega^2 x_3,$$

and therefore $\ddot{\boldsymbol{x}} = -\omega^2(x_2\boldsymbol{e}_2 + x_3\boldsymbol{e}_3)$, and the equation of motion (102) in the absence of body forces reduces to

$$\frac{\partial p}{\partial x_1} = 0, \quad \frac{\partial p}{\partial x_2} = \rho\omega^2 x_2, \quad \frac{\partial p}{\partial x_3} = \rho\omega^2 x_3, \tag{183}$$

which may be readily integrated (Exercise 42) to yield

$$p = \frac{1}{2}\rho\omega^2 r^2 + c(t), \tag{184}$$

where $r = (x_2^2 + x_3^2)^{1/2}$ is the distance of the point $\boldsymbol{x}$ from the cylinder's axis, and $c(t)$ is to be determined.

Let $\boldsymbol{n}$ be a outward unit normal vector at an arbitrary point on the cylinder's curved surface. The traction at that point is

$$\begin{aligned}\boldsymbol{t} = \boldsymbol{S}\boldsymbol{n} &= \Big[\big(-p + \alpha\lambda^{-1} + \beta(\lambda^2 + 2\lambda^{-1})\big)\,\boldsymbol{I} + (\alpha + \beta\lambda^{-1})(\lambda^2 - \lambda^{-1})\,\boldsymbol{e}_1 \otimes \boldsymbol{e}_1\Big]\boldsymbol{n}\\ &= \big(-p + \alpha\lambda^{-1} + \beta(\lambda^2 + 2\lambda^{-1})\big)\boldsymbol{n}\end{aligned}$$

since $(\boldsymbol{e}_1 \otimes \boldsymbol{e}_1)\boldsymbol{n} = \boldsymbol{0}$ due to $\boldsymbol{n}$ being perpendicular to $\boldsymbol{e}_1$. If the boundary of the cylinder is free of obstructions, as we have implicitly assumed it is, then the traction on the boundary is zero, leading to

$$-p + \alpha\lambda^{-1} + \beta(\lambda^2 + 2\lambda^{-1}) = 0 \quad \text{at } r = \lambda^{-1/2}R,$$

as $\lambda^{-1/2}R$ is the radius of the spinning cylinder as noted earlier. Plugging this in (184) determines the unknown expression $c(t)$:

$$c(t) = \alpha\lambda^{-1} + \beta(\lambda^2 + 2\lambda^{-1}) - \frac{1}{2}\rho\omega^2\lambda^{-1}R.$$

With $c(t)$ thus determined, the expression for p in (184) takes the form

$$\begin{aligned}p &= \frac{1}{2}\rho\omega^2 r^2 + \alpha\lambda^{-1} + \beta(\lambda^2 + 2\lambda^{-1}) - \frac{1}{2}\rho\omega^2\lambda^{-1}R.\\ &= \frac{1}{2}\rho\omega^2(r^2 - \lambda^{-1}R) + \alpha\lambda^{-1} + \beta(\lambda^2 + 2\lambda^{-1}).\end{aligned}$$

Next, let us look at the tractions on the cylinder's flat faces. On the face with the unit normal $\boldsymbol{e}_1$, the traction is

$$\boldsymbol{t} = \boldsymbol{S}\boldsymbol{e}_1 = \Big[-p + \alpha\lambda^{-1} + \beta(\lambda^2 + 2\lambda^{-1}) + (\alpha + \beta\lambda^{-1})(\lambda^2 - \lambda^{-1})\Big]\boldsymbol{e}_1,$$

which, upon substitution of the expression for the pressure calculated earlier, takes the form

$$\boldsymbol{t} = \Big[-\frac{1}{2}\rho\omega^2(r^2 - \lambda^{-1}R) + (\alpha + \beta\lambda^{-1})(\lambda^2 - \lambda^{-1})\Big]\boldsymbol{e}_1.$$

We see that the traction on this face varies with r. Thus, unlike the case with the cylinder's curved surface, having zero traction there is infeasible. That is, in order to maintain a flat face at the cylinder's ends, we need to press against it just the right way. Equivalently, if we leave the cylinder's ends free, they won't remain flat; they will bulge inward or outward.

So suppose we intend to keep the cylinder's faces flat by exerting just the right amount of tractions on them. Moreover, we wish to do this in such a way that the resultant force applied to each face is zero. This is expressed as

$$2\pi\int_0^{\lambda^{-1/2}R}\Big[-\frac{1}{2}\rho\omega^2(r^2 - \lambda^{-1}R) + (\alpha + \beta\lambda^{-1})(\lambda^2 - \lambda^{-1})\Big]\,r\,dr = 0.$$

Carrying out the integration we arrive at the quartic equation

$$q(\lambda) \equiv \alpha\lambda^4 + \beta\lambda^3 - \Big(\alpha - \frac{1}{4}\rho\omega^2R^2\Big)\lambda - \beta = 0.$$

We see that $q(0) = -\beta$ and $q(1) = \frac{1}{4}\rho\omega^2R^2$. That sign change indicates that there is a root $\lambda = \lambda_\omega$ between 0 and 1. We mark the root by a subscript ω to emphasize the dependence of that root on the angular velocity ω. We claim that λ_ω is *the only positive root* of that equation. Indeed, a quartic equation has either no real roots, or two real roots, or four real roots (counting multiplicities). We have already established the existence of one real root. That leaves the possibility of having either two real roots or four real roots. We know that the quartic's zeroth degree term, $-\beta$, equals the product of its roots. That being negative means that if the quartic has only two real roots, then one is positive and one is negative, and we are done. If it has four real roots, then either one is positive and three are negative, which means we are done, or one is negative and three are positive, which requires further analysis.

If the quartic has three positive roots, then it must have an inflection point at some positive λ. The inflection points are readily obtained by calculating $q''(\lambda) = 6\lambda(2\alpha\lambda + \beta)$. We see that the only inflection points are at $\lambda = 0$ and $\lambda = -\frac{\beta}{2\alpha}$, eliminating the possibility of three positive roots.

To summarize: The quartic equation $q(\lambda) = 0$ has a unique positive root $\lambda = \lambda_\omega$ where $0 < \lambda_\omega < 1$. The root is determined in terms of the material parameters α, β, ρ, the cylinder's pre-deformation radius R, and the angular velocity of the spin, ω. The length of the spinning cylinder shrinks to $\lambda_\omega R$ and the radius expands to $\lambda_\omega^{-1/2} R$.

34. Elasticity: Twisting a cylinder

The previous section's rubber cylinder is subjected to another experiment. As before, it has length L and radius R in its undeformed configuration. We set up an XYZ Cartesian coordinate system so that the Z axis coincides with the cylinder's axis, and the cylinder's bases are at $Z = 0$ and $Z = L$.

We deform the cylinder so that a point at (X, Y, Z) goes to the point (x, y, z) according to

$$\begin{aligned} x &= X \cos \tau Z - Y \sin \tau Z, \\ y &= X \sin \tau Z + Y \cos \tau Z, \\ z &= Z. \end{aligned} \tag{185}$$

In words, we twist the cylinder about its axis so that each $Z = c$ section, which is a disk of radius R, remains at $Z = c$ and retains its shape and size, but rotates about the Z axis by an angle τZ, that is, proportional to that section's distance from the cylinder's base. We wish to do this by applying tractions to the cylinder's bases while keeping the cylinder's curved surface traction-free. Is that possible?

The answer is yes. We will calculate the traction needed to accomplish that. Figure 17 depicts the cylinder before and after the deformation.

Unlike the case in the previous section, this deformation is not linear in (X, Y, Z), and therefore the previous approach does not apply. We calculate the deformation gradient $\boldsymbol{F} = \boldsymbol{F}(X, Y, Z)$ directly by differentiating (185). We get:

$$\boldsymbol{F} = \begin{pmatrix} \cos \tau Z & -\sin \tau Z & -\tau(X \sin \tau Z + Y \cos \tau Z) \\ \sin \tau Z & \cos \tau Z & \tau(X \cos \tau Z - Y \sin \tau Z) \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos \tau Z & -\sin \tau Z & -\tau y \\ \sin \tau Z & \cos \tau Z & \tau x \\ 0 & 0 & 1 \end{pmatrix}.$$

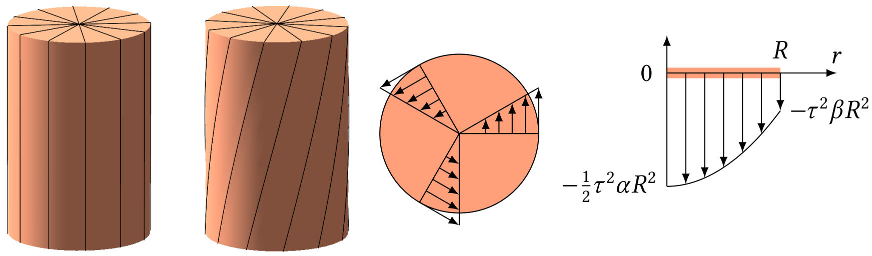


FIGURE 17. From the left: The first two figures depict the cylinder before and after twisting. The third and forth figures shows the horizontal and vertical components of the traction applied to the cylinder's top.

We see that $\det \boldsymbol{F} = 1$, therefore the deformation postulated in (185) is consistent with the material's incompressibility.

We calculate the left Cauchy–Green strain tensor $\boldsymbol{B} = \boldsymbol{F}\boldsymbol{F}^T$

$$\boldsymbol{B} = \begin{pmatrix} 1+\tau^2y^2 & -\tau^2xy & -\tau y \\ -\tau^2xy & 1+\tau^2x^2 & \tau x \\ -\tau y & \tau x & 1 \end{pmatrix},$$

and

$$\boldsymbol{B}^2 = \begin{pmatrix} 1+\tau^2(3+\tau^2r^2)y^2 & -\tau^2(3+\tau^2r^2)xy & -\tau(2+\tau^2r^2)y \\ -\tau^2(3+\tau^2r^2)xy & 1+\tau^2(3+\tau^2r^2)x^2 & \tau(2+\tau^2r^2)x \\ -\tau(2+\tau^2r^2)y & \tau(2+\tau^2r^2)x & 1+\tau^2r^2 \end{pmatrix},$$

where we have let $r = \sqrt{x^2+y^2} = \sqrt{X^2+Y^2}$. We also note that

$$\iota_1(\boldsymbol{B}) = \operatorname{tr}\boldsymbol{B} = 3+\tau^2r^2.$$

Plugging these into the constitutive equation $\boldsymbol{S} = -p\boldsymbol{I} + \big(\alpha + \beta\iota_1(\boldsymbol{B})\big)\boldsymbol{B} - \beta\boldsymbol{B}^2$, we get

$$\boldsymbol{S} = \begin{pmatrix} -p+\alpha+2\beta+\tau^2(\alpha y^2+\beta r^2) & -\alpha\tau^2xy & -(\alpha+\beta)\tau y \\ -\alpha\tau^2xy & -p+\alpha+2\beta+\tau^2(\alpha x^2+\beta r^2) & (\alpha+\beta)\tau x \\ -(\alpha+\beta)\tau y & (\alpha+\beta)\tau x & -p+\alpha+2\beta \end{pmatrix}. \tag{186}$$

Then we calculate

$$\operatorname{div}\boldsymbol{S} = \begin{pmatrix} -\frac{\partial p}{\partial x} + (2\beta-\alpha)\tau^2x \\ -\frac{\partial p}{\partial y} + (2\beta-\alpha)\tau^2y \\ -\frac{\partial p}{\partial z} \end{pmatrix},$$

and apply the equilibrium equation $\operatorname{div}\boldsymbol{S} = \boldsymbol{0}$ to get

$$-\frac{\partial p}{\partial x} + (2\beta-\alpha)\tau^2x = 0, \qquad -\frac{\partial p}{\partial y} + (2\beta-\alpha)\tau^2y = 0, \qquad -\frac{\partial p}{\partial z} = 0.$$

Upon integrating these equations we obtain

$$p = \frac{1}{2}(2\beta-\alpha)\tau^2r^2 + c, \tag{187}$$

where the constant c is to be determined.

Going back to the expression for $\boldsymbol{S}$, we evaluate the traction on the curved surface of the cylinder at the point $x = R, y = 0, z = 0$ where the outward unit normal vector is

$\boldsymbol{n} = \langle 1, 0, 0 \rangle$. We get

$$\boldsymbol{t} = \boldsymbol{S}\boldsymbol{n} = (-p + \alpha + 2\beta + \beta\tau^2 R^2)\,\boldsymbol{n}.$$

which shows that the traction, is perpendicular to the surface. By the problem's symmetry, the traction would be perpendicular to the curved surface everywhere, and will have the same magnitude. The assumption that the curved surface is traction-free leads to

$$-p + \alpha + 2\beta + \beta\tau^2 R^2 = 0,$$

or, substituting for p from (187)

$$\frac{1}{2}(2\beta - \alpha)\tau^2 r^2 + c + \alpha + 2\beta + \beta\tau^2 R^2 = 0,$$

which we solve for c and then substitute the result back into (187) to obtain

$$p = \alpha + 2\beta + \frac{1}{2}\tau^2\Big(\alpha(R^2 - r^2) + 2\beta r^2\Big).$$

Having thus obtained the pressure, we substitute it in (186) to obtain the finalized expression for $\boldsymbol{S}$.

That enables us to calculate the tractions on the cylinder's base which are needed to maintain the cylinder's deformed shape. The outward unit normal on the top face is $\boldsymbol{n} = \langle 0, 0, 1 \rangle$, and therefore the traction there is

$$\boldsymbol{t} = \boldsymbol{S}\boldsymbol{n} = \begin{pmatrix} -(\alpha+\beta)\tau y \\ (\alpha+\beta)\tau x \\ -p + \alpha + 2\beta \end{pmatrix} = \begin{pmatrix} -(\alpha+\beta)\tau y \\ (\alpha+\beta)\tau x \\ -\frac{1}{2}\tau^2\Big(\alpha(R^2 - r^2) + 2\beta r^2\Big) \end{pmatrix}.$$

The horizontal component of the traction, that is, the vector $\langle -(\alpha+\beta)\tau y, (\alpha+\beta)\tau x, 0\rangle$, is perpendicular to the position vector $\langle x, y, z\rangle$, and its magnitude is $(\alpha+\beta)\tau r$. This exerts the torque that twists the cylinder. The vertical component, $\frac{1}{2}\tau^2\Big(\alpha(R^2-r^2)+2\beta r^2\Big)$, is presses against the cylinder's top to keep it from bulging out. This is strictly a nonlinear elasticity effect. In the classical linear elasticity, which deals with infinitesimal deformations only, normal tractions are absent. We leave it for an exercise to show that the resultant torque is

$$T = \frac{(\alpha+\beta)\tau}{2\pi} A^2,$$

where A is the area of the cylinder's base.

Remark 32. The problem was introduced and solved in Rivlin [16] through rather ad hoc methods. He developed the solution further in [19] and [17]. In an earlier paper [13] he presents laboratory results of experiments performed on actual rubbers. The solution adopted here is close to that in [17] and it was suggested by Exercise 14 on page 164 of [2]. This is a rather odd method, as it uses Cartesian coordinates to solve a problem with cylindrical geometry. It is fortunate that things work out as they do in this instance, but it's not a general approach to such problems. A general approach to solving large deformation of elastic materials is developed in Green and Zerna [6]. They solve the problem of the twisting of a cylinder an illustration of the power of the tensor calculus of curvilinear coordinates which they introduce in great detail in the first chapter. Ideally that's the right approach to solving this problem but it requires some investment in learning the machinery of tensor calculus which we avoid in these notes.

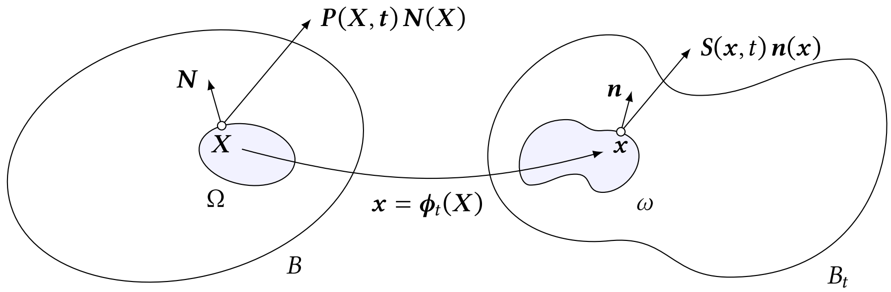


FIGURE 18. The tractions at X and x are $P(X,t)\,N(X)$ and $S(x,t)\,n(x)$, where P is the Piola–Kirchhoff stress tensor and S is the Cauchy stress tensor. The traction vectors $P\,N$ and $S\,n$ are parallel but generally of different lengths due to the local stretching/contraction of the area. If dA_X and dA_x are corresponding area elements at X and x, then $P\,N\,dA_X = S\,n\,dA_x$.

35. Elasticity: The Piola–Kirchhoff stress tensor

The Cauchy stress measures contact force per unit area in the *deformed* configuration. That's rather inconvenient in formulating many boundary value problems in elasticity since the deformed configuration may be unknown ahead of the time. The *Piola–Kirchhoff stress* introduced in this section sidesteps that difficulty since it measures contact force per unit area in the *reference* configuration. Figure 18 illustrates the idea. A part Ω of the body B_t is mapped to ω under the deformation ϕ_t. A point X on the boundary $\partial\Omega$ of Ω is mapped to a point x on the boundary $\partial\omega$ of ω. The outward unit normals to $\partial\Omega$ and $\partial\omega$ at X and x are N and n, respectively. The boundary traction at x is $S(x,t)n(x)$ where S is the Cauchy stress tensor. From the transformation of surface integral formula (122c) we know that

$$\int_{\partial\omega} S(x,t)\,n(x)\,dA_x = \int_{\partial\Omega} \big(\det F(X,t)\big)\,S\big(\phi_t(X),t\big)\,F^{-T}(X,t)\,N(X)\,dA_X.$$

Motivated by this, we introduce the *Piola–Kirchhoff stress* $P(X,t)$ via

$$P(X,t) = \big(\det F(X,t)\big)\,S\big(\phi_t(X),t\big)\,F^{-T}(X,t), \tag{188a}$$

or written compactly

$$P = (\det F)\,S\,F^{-T}, \tag{188b}$$

whereby the integration formula above takes the form

$$\int_{\partial\omega} S(x,t)\,n(x)\,dA_x = \int_{\partial\Omega} P(X,t)\,N(X)\,dA_X.$$

Therefore, the force $Sn\,dA_x$ acting on a surface element of size dA_x in the deformed configuration may be expressed as $PN\,dA_X$ in the reference configuration. Note that the vectors Sn and PN are parallel but generally of different lengths since the area elements dA_x and dA_X are not necessarily equal due to the stretching of the material.

In (147a) we have a representation of the most general constitutive equation for an isotropic elastic material. To express that equation in terms of the Piola–Kirchhoff stress tensor, we note that

$$BF^{-T} = FF^TF^{-T} = F,$$

and

$$B^2F^{-T} = B\big(BF^{-T}\big) = BF.$$

Therefore from (147a) and (188b) we get

$$P = (\det F)\big(\alpha_0 F^{-T} + \alpha_1 F + \alpha_2 BF\big),$$

where the α_i, $i = 0, 1, 2$, are scalar-valued isotropic functions of B, as before. The equivalent formula corresponding to the alternative representation (147b) is

$$P = (\det F)\big(\beta_0 F^{-T} + \beta_1 F + \beta_2\,(F^T B)^{-1}\big).$$

The equation of motion (102) on page 43 is expressed in in terms of the Cauchy stress S defined in spatial coordinates over the deformed configuration $\Omega_t = \phi_t(\Omega)$. We wish to obtain the equivalent equation of motion expressed in material coordinates over the reference configuration Ω. We already know how to transform the stress. Let us look at the transformation of the body force which appears as ρb in (102)

Theorem 30. *The equations of motion expressed in material coordinates are*

$$\rho_0 \dot{V} = \operatorname{Div} P + \rho_0 b, \tag{189a}$$

$$PF^T = FP^T. \tag{189b}$$

Proof. To do... □

36. Elasticity: Homogeneous deformations of a cube

This section is based on the original work of Rivlin [15, 18] and has been influenced by the presentation in Gurtin [8].

Consider a homogeneous, isotropic, incompressible elastic material defined through the constitutive equation

$$\hat{S}(F) = -pI + \beta B, \tag{190}$$

where S is the Cauchy stress, where $B = FF^T$ is the left Cauchy–Green strain tensor, and β is a positive constant. This is a very special case of the general constitutive equation (159b).

Then then corresponding Piola–Kirchhoff stress tensor, defined in (188b), is given by

$$P = -pF^{-T} + \beta F. \tag{191}$$

Consider a cube is made of such a material and placed in the Cartesian coordinate system so that its faces are parallel to the coordinates planes. We apply three pairs of uniformly distributed equal and opposite *dead forces* acting perpendicularly to cube's faces. The "dead force" adjective indicates that the force acting on a face remains the same as the cube deforms and the face's area changes. In other words, the surface traction $t = Sn$ as measured in the deformed configuration varies depending on how the cube deforms, but the surface traction $t = Pn$ as measured in the undeformed configuration remains fixed, regardless of how the cube deforms. We write $\eta\beta$ for the magnitude of the traction

$\boldsymbol{t} = \boldsymbol{P}\boldsymbol{n}$, viewing the dimensionless factor η as a measure of the applied force. A positive η expresses tension, while a negative η expresses compression.

We wish to investigate the possible deformations of the cube under such loads. We limit our search to homogeneous deformations, leaving open the question of whether other types of deformation may be possible. Thus, we present the Piola–Kirchhoff stress and the deformation gradient, relative to the coordinate system, as

$$\boldsymbol{P} = \begin{pmatrix} \eta\beta & 0 & 0 \\ 0 & \eta\beta & 0 \\ 0 & 0 & \eta\beta \end{pmatrix}, \quad \boldsymbol{F} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix},$$

where $\lambda_k > 0$ for $k = 1, 2, 3$. Plugging these into (191) we get

$$\begin{pmatrix} \eta\beta & 0 & 0 \\ 0 & \eta\beta & 0 \\ 0 & 0 & \eta\beta \end{pmatrix} = -p \begin{pmatrix} 1/\lambda_1 & 0 & 0 \\ 0 & 1/\lambda_2 & 0 \\ 0 & 0 & 1/\lambda_3 \end{pmatrix} + \beta \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix},$$

and therefore $\eta\beta = -p/\lambda_k + \beta\lambda_k$, $k = 1, 2, 3$, that is, $p = \beta(\lambda_k^2 - \eta\lambda_k)$. It follows that

$$\eta(\lambda_i - \lambda_j) = \lambda_i^2 - \lambda_j^2 \quad \text{for all } i, j \in \{1, 2, 3\},$$

or, in expanded form:

$$\eta(\lambda_1 - \lambda_2) = (\lambda_1 - \lambda_2)(\lambda_1 + \lambda_2) \tag{192a}$$

$$\eta(\lambda_2 - \lambda_3) = (\lambda_2 - \lambda_3)(\lambda_2 + \lambda_3) \tag{192b}$$

$$\eta(\lambda_3 - \lambda_1) = (\lambda_3 - \lambda_1)(\lambda_3 + \lambda_1). \tag{192c}$$

Additionally, the incompressibility constraint $\det \boldsymbol{F} = 1$ imposes the condition

$$\lambda_1\lambda_2\lambda_3 = 1. \tag{192d}$$

Thus, our quest reduces to finding one or more solutions to the system of four equations in (192) for the three unknowns $\lambda_1, \lambda_2, \lambda_3$ in terms of η. The choice $\lambda_1 = \lambda_2 = \lambda_3 = 1$ and arbitrary η certainly satisfy these equations. In other words, the undeformed cube is a solution for arbitrary load values η, positive or negative. To go beyond that trivial solution, we observe that our system of four equations is equivalent to

$$\big\{\lambda_1 = \lambda_2 \text{ or } \lambda_1 + \lambda_2 = \eta\big\} \text{ and } \big\{\lambda_2 = \lambda_3 \text{ or } \lambda_2 + \lambda_3 = \eta\big\}$$
$$\text{and } \big\{\lambda_3 = \lambda_1 \text{ or } \lambda_3 + \lambda_1 = \eta\big\} \text{ and } \lambda_1\lambda_2\lambda_3 = 1. \tag{193}$$

In Exercise 47 you will show that (193) cannot hold if $\lambda_1, \lambda_2, \lambda_3$ are all distinct. Therefore here we focus our attention on the case where $\lambda_1 = \lambda_2 \neq \lambda_3$, and write λ for λ_3 to simplify the notation. Then the first phrase of (193) is satisfied, the second and third phrases reduce to $\lambda + \lambda_1 = \eta$, and the fourth phrase reduces to $\lambda_1^2\lambda = 1$. Eliminating λ_1 between the latter two equations leads to $\phi(\lambda) \stackrel{\text{def}}{=} (\eta - \lambda)^2\lambda - 1 = 0$. The solution of the problem hinges on finding the roots of this cubic.

We have $\phi'(\lambda) = 3\lambda^2 - 4\eta\lambda + \eta^2 = (\lambda - \eta)(3\lambda - \eta)$, and therefore the critical points of ϕ are at $\eta/3$ and η. We also have $\phi''(\eta/3) = -2\eta < 0$ and $\phi''(\eta) = 2\eta > 0$, indicating that the critical points at ϕ are at $\eta/3$ and η are a local maximum and a local minimum, respectively. Equipped with the further observation that $\phi(-\infty) = -\infty$, $\phi(+\infty) = +\infty$, and $\phi(0) = \phi(\eta) = -1$, we are in a position to sketch the representative graphics of ϕ in Figure 19.

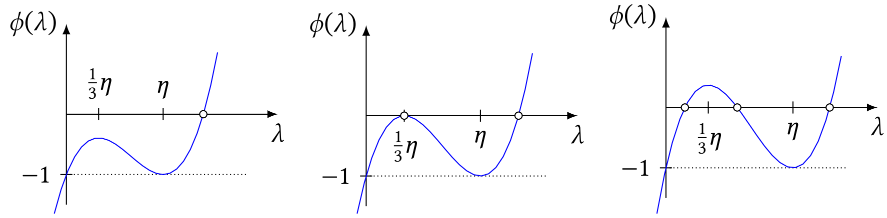


FIGURE 19. The graphs of $\phi(\lambda)$, from left to right, corresponding to $\eta = 1.6 < \eta_{\rm cr}$, $\eta = 1.89 \approx \eta_{\rm cr}$, $\eta = 2.1 > \eta_{\rm cr}$ show the cases where we have zero roots, one root, and two roots in the interval $0 < \lambda < \eta$.

The vertical location of the graph's local maximum, the "hump", may fall below, at, or above the horizontal axis, depending on the value of the parameter η. In those three cases, the function ϕ has one, or two, or three roots, respectively, as we seen the Figure 19. The middle graph, which corresponds to the case when the hump touches the horizontal axis, is obtained when $\phi(\eta/3) = \frac{4}{27}\eta^3 - 1 = 0$, that is $\eta = \eta_{\rm cr} \stackrel{\rm def}{=} \left(\frac{27}{4}\right)^{1/3}$.

Since $\phi(\eta) < 0$ and $\phi(+\infty) = +\infty$, there is always a root of ϕ with $\lambda > \eta$ but that root is of no interest in our context since $\lambda + \lambda_1 = \eta$ would imply that $\lambda_1 < 0$. Other roots, if any, would lie in the interval $(0, \eta)$ and would be admissible. The number of such roots, and consequently the number of solutions, depends on the value of η:

$$\begin{aligned}
&\text{if } \eta < \eta_{\rm cr}, && \text{then there are no solution,}\\
&\text{if } \eta = \eta_{\rm cr}, && \text{then there is one solution,}\\
&\text{if } \eta > \eta_{\rm cr}, && \text{then there are two solutions.}
\end{aligned}$$

The preceding analysis was carried out under the assumption that $\lambda_1 = \lambda_2 \neq \lambda_3$. By cyclic permutation of the indices we obtain solutions that favor the x_1 or x_2 directions, thus leading to a collection of as many as seven solutions of the type we have been searching for. We summarize this as follows:

- $\lambda_1 = \lambda_2 = \lambda_3 = 1$ is a solution (the "trivial" solution) for all η;
- if $\eta < \eta_{\rm cr}$, then there are no other solutions;
- if $\eta = \eta_{\rm cr}$, then there are three solution in addition to the trivial one;
- if $\eta > \eta_{\rm cr}$, then there are six solution in addition to the trivial one.

Figure 20 depicts representative sequence of the cube's deformations. Subfigure (a) is the undeformed configuration of the cube, which also corresponds to the trivial solution noted above. For load values $\eta < \eta_{\rm cr}$, that's all we get. As soon as the load η surpasses $\eta_{\rm cr} \approx 1.89$, the cube snaps to the configuration shown in subfigure (b). This can happen in any of the three coordinate directions, therefore subfigure (b), as well of each of the remaining subfigures, represents one of the three possibilities. Subfigures (c) and (d) depict the two possible deformation for $\eta = 1.92 > \eta_{\rm cr}$, corresponding to the two roots of ϕ. Subfigures (e) and (f) depict yet another such pair corresponding to a larger load $\eta = 2.3 > \eta_{\rm cr}$.

Remark 33. The preceding analysis indicates that the system of four equations (192) in the three unknowns λ_1, λ_2, λ_3, is *consistent* despite the mismatch between the number equations and the number of unknowns. In fact, by applying *Viète's formula* for solving

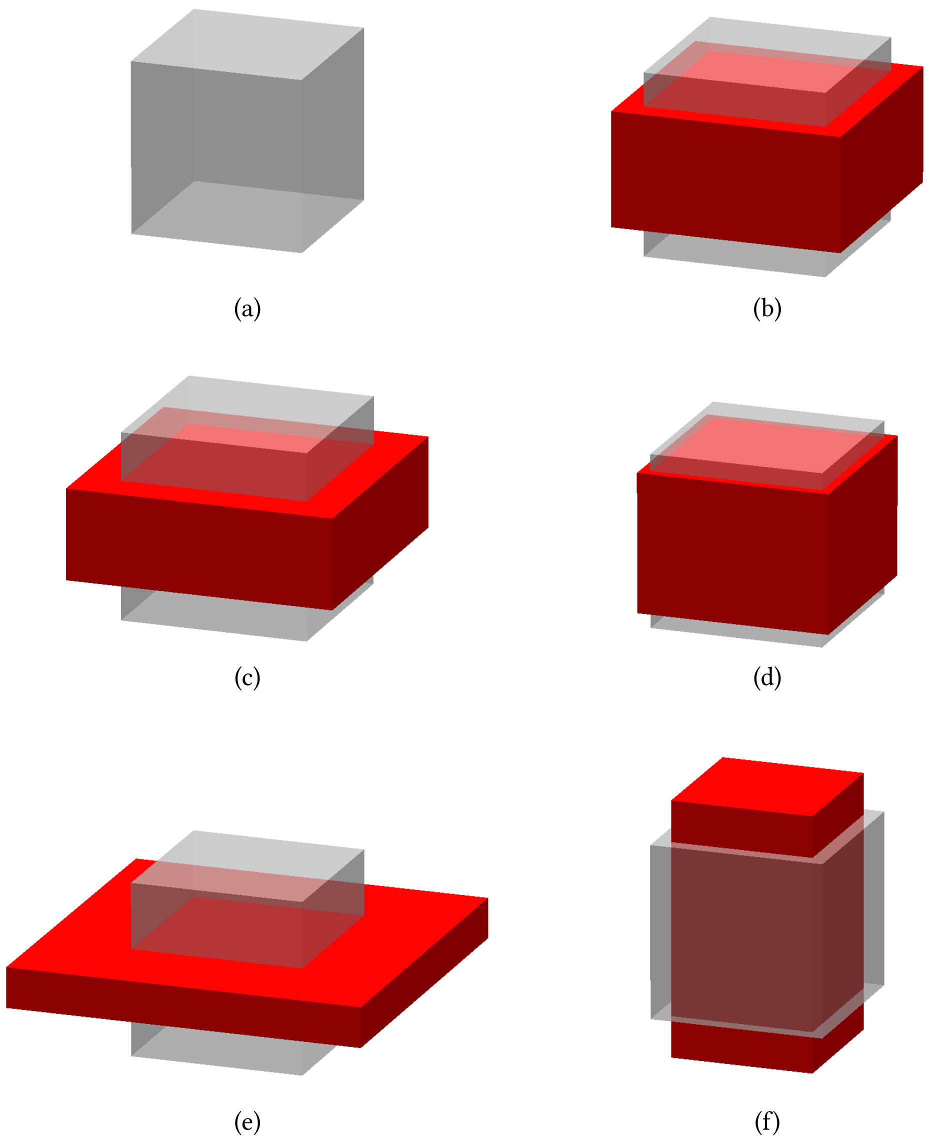


FIGURE 20. (a): the undeformed cube; (b): the cube snaps to this configuration when $\eta = \eta_{\text{cr}} \approx 1.89$; (c) and (d): the two solutions when $\eta = 1.92$; (e) and (f): the two solutions when $\eta = 2.3$;

cubics, we readily obtain a symbolic expression for the three roots of the cubic equation $\phi(\lambda) = (\eta - \lambda)^2\lambda - 1 = 0$:

$$\lambda = \frac{2}{3}\eta\left[1 + \cos\left(\frac{2\pi k}{3} + \frac{1}{3}\arccos\zeta\right)\right], \quad \text{where } k = -1, 0, 1, \text{ and } \zeta = \frac{27}{2\eta^3} - 1.$$

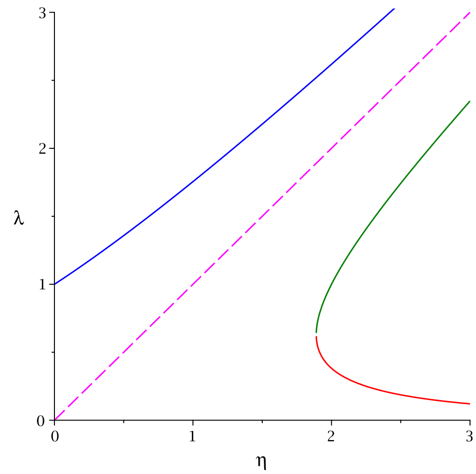


FIGURE 21. The solid lines depict the three roots of the equation $(\eta - \lambda)^2\lambda - 1 = 0$ plotted against η. The roots plotted in red and green lie below the line $\lambda = \eta$, plotted in dashed magenta. The root plotted in blue lies above that line and is not of interest in the context of our application. The leftmost point of the red and green graphs is at $\eta = \eta_{\text{cr}} \approx 1.89$.

We will have real roots if $-1 \le \zeta \le 1$, which is equivalent to $\eta \ge \eta_{\text{cr}}$. The root corresponding to $k = 0$ is grater than η, which, as was remarked earlier, is inadmissible in the context of this elasticity problem. The roots produced by to $k = \pm 1$ lie between 0 and η and are admissible. Figure 21 shows the graphs of the three roots versus η.

Remark 34. Not all solutions obtained in this section are stable. Unstable solutions may be practically impossible to produce in experiments. Rivlin's article [20] contains a survey of the literature on the stability issue up to 1974, and provides a concise summary of the results, as follows.

The trivial solution, that is, the one with $\lambda_1 = \lambda_2 = \lambda_3$, exists for all η. The trivial solution is *unstable* if $\eta < 0$ (compression) or $\eta > 2$. It is *stable* for $0 < \eta < 2$.

As we have seen, non-trivial solutions exist if $\eta > \eta_{\text{cr}}$. In that case the cubic $\phi(\lambda)$ has three roots. The root that is greater than η is inadmissible as we have noted before. Let us name the smaller of the remaining two roots $\lambda^{(1)}$, and the other one $\lambda^{(2)}$. Referring to Figure 19, we see that

$$0 < \lambda^{(1)} < \frac{1}{3}\eta < \lambda^{(2)} < \eta.$$

Rivlin proves that the deformations corresponding to $\lambda^{(1)}$ are *stable* and those corresponding to $\lambda^{(2)}$ are *unstable*. Therefore, the configurations (c) and (e) in Figure 20 are stable, while (d) and (f) are unstable.

Remark 35. The problem discussed in this section was introduced in Rivlin [14]. In [15] he analyzed the more general problem where the force pairs applied to cube's faces are not necessarily of equal magnitudes. In [18] he extended the analysis of [14] to Mooney-Rivlin materials. The stability of the solutions discussed in these articles is in the limited context of stability relative to infinitesimal superimposed deformations along the coordinate axes. The analysis under the stronger criterion of stability relative to arbitrary infinitesimal superimposed deformations is done in [20].

37. Elasticity: Hyperelastic materials

Material based on [7, 8]

Definitions from page 65 of Green & Zerna: Mooney material: $W = C_1(I_1-3)+C_2(I_2-3)$; Neo-Hookean material: when $C_2 = 0$.

38. Linear Elasticity

The theory of *linear elasticity* was developed in the early to mid 1800s by Cauchy, Poisson, Navier, and many other contributors. The goal of this section is introduce the basic ideas of linear elasticity and to relate these to the general theory of elasticity that we have studied up to this point in these lecture notes.

As we shall see, linear elasticity is *not a special case* of the general theory of elasticity. Rather, in a sense that we will make precise, linear elasticity is a *linear approximation* to the general theory. As such, we should be aware that not everything that we have learned about the general theory is applicable to this approximation. The most glaring discrepancy is that constitutive equations of linear elasticity *are not* frame-invariant.

Historically, “theory of elasticity” has meant “theory of linear elasticity” and the general theory of the previous sections is referred to as the “theory of nonlinear elasticity”. Nevertheless, we shall continue referring to the general theory as “elasticity”, and refer to the linearized version as “linear elasticity”.

This section’s presentation is a blend of the presentation on pages 288–302 of Gonzalez and Stuart [5], and Gurtin [7].

38.1. **Fourth order tensors.** Up to this point in these lecture notes, we have used the word *tensor* to refer to *second order tensors*. These are linear mappings from $\mathcal{V}$ to $\mathcal{V}$. We wrote $\mathcal{L}$ for the linear space of the second order tensors, and introduced a scalar product $\boldsymbol{A}:\boldsymbol{B}$ on $\mathcal{L}$. In this section it will be necessary to study linear mappings from $\mathcal{L}$ to $\mathcal{L}$. These are called *fourth order tensors*.

Following the tradition in linear elasticity, we will indicate the application of the fourth order tensor G on the second order tensor $\boldsymbol{H}$ by placing $\boldsymbol{H}$ in square brackets, as in $\mathsf{G}[\boldsymbol{H}]$.

Consider the fourth order tensor G, and let $\boldsymbol{T} = \mathsf{G}[\boldsymbol{H}]$. Writing $\boldsymbol{H} = h_{kl}\boldsymbol{e}_k \otimes \boldsymbol{e}_l$ and $\boldsymbol{T} = t_{ij}\boldsymbol{e}_k \otimes \boldsymbol{e}_l$ relative to a frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ in $\mathcal{V}$, we have

$$t_{ij} = \boldsymbol{e}_i \cdot \boldsymbol{T}\boldsymbol{e}_j = \boldsymbol{e}_i \cdot \mathsf{G}[\boldsymbol{H}]\boldsymbol{e}_j = \boldsymbol{e}_i \cdot \mathsf{G}[h_{kl}\boldsymbol{e}_k \otimes \boldsymbol{e}_l]\boldsymbol{e}_j = \boldsymbol{e}_i \cdot \mathsf{G}[\boldsymbol{e}_k \otimes \boldsymbol{e}_l]\boldsymbol{e}_j\, h_{kl}$$

We define the *components of* G *relative to the* $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ *frame* as

$$G_{ijkl} = \boldsymbol{e}_i \cdot \mathsf{G}[\boldsymbol{e}_k \otimes \boldsymbol{e}_l]\boldsymbol{e}_j,$$

whereby the component representation of $\boldsymbol{T} = \mathbf{G}[\boldsymbol{H}]$ is expressed as

$$t_{ij} = G_{ijkl} h_{kl}.$$

Thus, a fourth order tensor $\mathbf{G}$ is determined by the 81 components G_{ijkl} relative to a frame.

38.2. **More on derivatives.** Consider a generally nonlinear function $\boldsymbol{G} : \mathcal{L} \to \mathcal{L}$. We say $\boldsymbol{G}$ *is differentiable at* $\boldsymbol{F}$ if there exists a fourth order tensor $\boldsymbol{G}'$ such that[15]

$$\lim_{\|\boldsymbol{H}\| \to 0} \frac{\left\| \boldsymbol{G}(\boldsymbol{F} + \boldsymbol{H}) - \boldsymbol{G}(\boldsymbol{F}) - \boldsymbol{G}'(\boldsymbol{F})[\boldsymbol{H}] \right\|}{\|\boldsymbol{H}\|} = 0, \tag{194}$$

and we refer to $\boldsymbol{G}'(\boldsymbol{F})$ as the derivative of $\boldsymbol{G}$ at $\boldsymbol{F}$. If such a $\boldsymbol{G}'(\boldsymbol{F})$ exists, then it is unique. If there is no such $\boldsymbol{G}'(\boldsymbol{F})$, then one says that $\boldsymbol{G}$ *is not differentiable at* $\boldsymbol{F}$. In what follows, we will tacitly assume that our functions are differentiable wherever necessary.

Remark 36. The definition of the derivative in (194) implies that

$$\boldsymbol{G}'(\boldsymbol{F})[\boldsymbol{H}] = \frac{d}{d\epsilon} \boldsymbol{G}(\boldsymbol{F} + \epsilon \boldsymbol{H}) \bigg|_{\epsilon = 0}. \tag{195}$$

Theorem 31. *Let* $\boldsymbol{G} : \mathcal{L} \to \mathcal{L}$ *be isotropic. Then*

$$\boldsymbol{Q}\boldsymbol{G}'(\boldsymbol{F})[\boldsymbol{H}]\boldsymbol{Q}^T = \boldsymbol{G}'(\boldsymbol{Q}\boldsymbol{F}\boldsymbol{Q}^T)[\boldsymbol{Q}\boldsymbol{H}\boldsymbol{Q}^T]. \quad \textit{for all } \boldsymbol{H} \in \mathcal{L} \textit{ and all } \boldsymbol{Q} \in \mathcal{L}_{orth}. \tag{196}$$

Proof. Lengthy proof, page 237 of Gurtin [7], to be added later □

38.3. **The elasticity tensor.** Let $\boldsymbol{P} = \hat{\boldsymbol{P}}(\boldsymbol{F})$ be the constitutive equation at a generic point $\boldsymbol{p}$ of an elastic material, expressed in terms of the Piola-Kirchhoff stress tensor $\boldsymbol{P}$. When dealing with linear elasticity, we will assume that the stress in the undeformed material is zero, that is, $\hat{\boldsymbol{P}}(\boldsymbol{I}) = \mathbf{0}$. The derivative $\hat{\boldsymbol{P}}'(\boldsymbol{F})$ evaluated at $\boldsymbol{F} = \boldsymbol{I}$ is called the material's *elasticity tensor* at $\boldsymbol{p}$. We write

$$\mathbf{C} = \hat{\boldsymbol{P}}'(\boldsymbol{I}).$$

Proposition 9. *Let* $\boldsymbol{S} = \hat{\boldsymbol{S}}(\boldsymbol{F})$ *be the constitutive equation of the elastic material expressed in terms of the Cauchy stress tensor. Then* $\mathbf{C} = \hat{\boldsymbol{S}}'(\boldsymbol{I})$*. That is, the elasticity tensor could have been equally defined as the derivative of the Cauchy stress evaluated at* $\boldsymbol{F} = \boldsymbol{I}$*.*

Proof. From the definition of the Piola-Kirchhoff tensor in (188b) we have

$$\hat{\boldsymbol{P}}(\boldsymbol{F})\boldsymbol{F}^T = (\det \boldsymbol{F})\, \hat{\boldsymbol{S}}(\boldsymbol{F})$$

Differentiating this with the help of the product rule, for any $\boldsymbol{H} \in \mathcal{L}$ we get

$$\hat{\boldsymbol{P}}'(\boldsymbol{F})[\boldsymbol{H}]\, \boldsymbol{F}^T + \hat{\boldsymbol{P}}(\boldsymbol{F})\boldsymbol{H}^T = (\det \boldsymbol{F})'[\boldsymbol{H}]\, \hat{\boldsymbol{S}}(\boldsymbol{F}) + (\det \boldsymbol{F})\, \hat{\boldsymbol{S}}'(\boldsymbol{F})[\boldsymbol{H}]$$

Substituting $(\det \boldsymbol{F})' = (\det \boldsymbol{F})\boldsymbol{F}^{-T}$ according to (78), and then evaluating the result at $\boldsymbol{F} = \boldsymbol{I}$, we obtain

$$\hat{\boldsymbol{P}}'(\boldsymbol{I})[\boldsymbol{H}]\, \boldsymbol{I}^T + \hat{\boldsymbol{P}}(\boldsymbol{I})\boldsymbol{H}^T = (\det \boldsymbol{I})\boldsymbol{I}^{-T}[\boldsymbol{H}]\, \hat{\boldsymbol{S}}(\boldsymbol{I}) + (\det \boldsymbol{I})\, \hat{\boldsymbol{S}}'(\boldsymbol{I})[\boldsymbol{H}],$$

which, due to the assumption $\hat{\boldsymbol{P}}(\boldsymbol{I}) = \hat{\boldsymbol{S}}(\boldsymbol{I}) = \mathbf{0}$, reduces to

$$\hat{\boldsymbol{P}}'(\boldsymbol{I})[\boldsymbol{H}] = \hat{\boldsymbol{S}}'(\boldsymbol{I})[\boldsymbol{H}] \quad \text{for all } \boldsymbol{H} \in \mathcal{L}.$$ □

[15]This definition of the derivative generalizes to nonlinear mapping on Banach spaces. In that context, $\boldsymbol{G}'$ is called the *Fréchet derivative* of $\boldsymbol{G}$.

Proposition 10 (Properties of the elasticity tensor)**.** *We have*

(1) $\mathsf{C}[\boldsymbol{H}] \in \mathcal{L}_{sym}$ *for all* $\boldsymbol{H} \in \mathcal{L}$*;*
(2) $\mathsf{C}[\boldsymbol{W}] = \boldsymbol{0}$ *for all* $\boldsymbol{W} \in \mathcal{L}_{skew}$*.*

Proof. Form (195) we have

$$\mathsf{C}[\boldsymbol{H}] = \hat{S}'(\boldsymbol{I})[\boldsymbol{H}] = \frac{d}{d\epsilon}\hat{S}(\boldsymbol{I} + \epsilon\boldsymbol{H})\bigg|_{\epsilon=0}.$$

But $\hat{S}(\boldsymbol{F} \in \mathcal{L}_{\text{sym}}$ for every $\boldsymbol{F} \in \mathcal{L}$. It follows that $\mathsf{C}[\boldsymbol{H}] \in \mathcal{L}_{\text{sym}}$ for all $\boldsymbol{H} \in \mathcal{L}$, thus proving part (1) of the proposition.

As to part (2), pick any $\boldsymbol{W} \in \mathcal{L}_{\text{skew}}$, and let $\boldsymbol{Q}(t) = e^{\boldsymbol{W}t}$. We have seen (where?) that $\boldsymbol{Q}(t) \in \mathcal{L}_{\text{orth}}$. Then plugging this $\boldsymbol{Q}$ in the frame-invariance requirement (139), we get

$$\hat{S}\big(\boldsymbol{Q}(t)\boldsymbol{F}\big) = \boldsymbol{Q}(t)\hat{S}(\boldsymbol{F})\boldsymbol{Q}(t)^T,$$

which, upon the substitution $\boldsymbol{F} = \boldsymbol{I}$, and recalling $\hat{S}(\boldsymbol{I}) = \boldsymbol{0}$, reduces to $\hat{S}\big(\boldsymbol{Q}(t)\big) = \boldsymbol{0}$. Differentiating this with respected to t we obtain $\hat{S}'\big(\boldsymbol{Q}(t)\big)[\dot{\boldsymbol{Q}}(t)] = \boldsymbol{0}$. In particular, at $t = 0$ we have $\hat{S}'\big(\boldsymbol{Q}(0)\big)[\dot{\boldsymbol{Q}}(0)] = \boldsymbol{0}$. But $\boldsymbol{Q}(0) = \boldsymbol{I}$, and $\dot{\boldsymbol{Q}}(0) = \boldsymbol{W}$. We conclude that $\hat{S}'(\boldsymbol{I})[\boldsymbol{W}] = \boldsymbol{0}$, that is, $\mathsf{C}[\boldsymbol{W}] = \boldsymbol{0}$. □

Pick any $\boldsymbol{H} \in \mathcal{L}$ let

$$\boldsymbol{E} = \frac{1}{2}\big(\boldsymbol{H} + \boldsymbol{H}^T\big), \quad \boldsymbol{W} = \frac{1}{2}\big(\boldsymbol{H} - \boldsymbol{H}^T\big).$$

Then $\boldsymbol{E} \in \mathcal{L}_{\text{sym}}$ and $\boldsymbol{W} \in \mathcal{L}_{\text{skew}}$, and $\boldsymbol{H} = \boldsymbol{E} + \boldsymbol{W}$. From Proposition 10 and the linearity of C it follows that $\mathsf{C}[H] = \mathsf{C}[\boldsymbol{E} + \boldsymbol{W}] = \mathsf{C}[\boldsymbol{E}]$, that is, the values of C are completely determined by its restriction to $\mathcal{L}_{\text{sym}}$. From now on, we will regard C as a linear mapping of $\mathcal{L}_{\text{sym}}$ to $\mathcal{L}_{\text{sym}}$.

Proposition 11. *Suppose the elastic material is isotropic. Then*

$$\boldsymbol{Q}\,\mathsf{C}[\boldsymbol{H}]\boldsymbol{Q}^T = \mathsf{C}[\boldsymbol{Q}\boldsymbol{H}\boldsymbol{Q}^T] \quad \textit{for all } \boldsymbol{H} \in \mathcal{L} \textit{ and all } \boldsymbol{Q} \in \mathcal{L}_{orth}, \tag{197}$$

that is, C *is an isotropic function.*

Proof.

$$\boldsymbol{Q}\,\mathsf{C}[\boldsymbol{H}]\boldsymbol{Q}^T = \boldsymbol{Q}\,\hat{S}'(\boldsymbol{I})[\boldsymbol{H}]\boldsymbol{Q}^T \overset{\text{by (196)}}{=} \hat{S}'(\boldsymbol{Q}\boldsymbol{I}\boldsymbol{Q}^T)[\boldsymbol{Q}\boldsymbol{H}\boldsymbol{Q}^T] = \hat{S}'(\boldsymbol{I})[\boldsymbol{Q}\boldsymbol{H}\boldsymbol{Q}^T] = \mathsf{C}[\boldsymbol{Q}\boldsymbol{H}\boldsymbol{Q}^T].$$

□

Theorem 32. *Suppose that* $\mathsf{G} : \mathcal{L}_{sym} \to \mathcal{L}_{sym}$ *is linear and isotropic. Then, there exist scalar constants* μ *and* λ *such that*

$$\mathsf{G}[\boldsymbol{E}] = 2\mu\boldsymbol{E} + \lambda(\operatorname{tr}\boldsymbol{E})\boldsymbol{I} \quad \textit{for all } \boldsymbol{E} \in \mathcal{L}_{sym}. \tag{198}$$

Proof. According to Theorem 13 (page 31), any (not necessarily linear) isotropic function on $\mathcal{L}_{\text{sym}}$ is of the form

$$G(\boldsymbol{E}) = \alpha_0\boldsymbol{I} + \alpha_1\boldsymbol{E} + \alpha_2\boldsymbol{E}^2 \quad \text{for all } \boldsymbol{E} \in \mathcal{L}_{\text{sym}}$$

where the coefficients α_0, α_1, α_2 are scalar-valued functions of the invariants of $\boldsymbol{E}$:

$$\iota_1(\boldsymbol{E}) = \operatorname{tr}\boldsymbol{E}, \quad \iota_2(\boldsymbol{E}) = \frac{1}{2}\Big((\operatorname{tr}\boldsymbol{E})^2 - \operatorname{tr}\big(\boldsymbol{E}^2\big)\Big), \quad \iota_3(\boldsymbol{E}) = \det\boldsymbol{E}.$$

But in the current case, $\boldsymbol{G}$ is also linear, and therefore the choices of ι_i are limited to

$$\alpha_0 = c_0 \operatorname{tr}\boldsymbol{E} + c_1, \quad \alpha_1 = c_2, \quad \alpha_2 = 0,$$

where c_0, c_1, and c_2 are scalar constants. This reduces the representation of $\boldsymbol{G}$ to

$$\mathbf{G}[\boldsymbol{E}] = (c_0 \operatorname{tr}\boldsymbol{E} + c_1)\boldsymbol{I} + c_2\boldsymbol{E} \quad \text{for all } \boldsymbol{E} \in \mathcal{L}_{\text{sym}}.$$

The linearity of $\mathbf{G}$ implies that $\mathbf{G}(\mathbf{0}) = \mathbf{0}$, and therefore $c_1 = 0$. The theorem's assertion follows by setting $c_0 = \lambda$ and $c_2 = 2\mu$. □

Corollary 8. *The elasticity tensor* $\mathbf{C}$ *of an isotropic material has the form*

$$\mathbf{C}[\boldsymbol{E}] = 2\mu\boldsymbol{E} + \lambda(\operatorname{tr}\boldsymbol{E})\boldsymbol{I} \quad \textit{for all } \boldsymbol{E} \in \mathcal{L}_{sym}. \tag{199}$$

The coefficients μ and λ are called the material's ***Lamé moduli****.*[16]

Viewing the elasticity tensor $\mathbf{C}$ as a linear operator from $\mathcal{L}$ to $\mathcal{L}$, we say that $\mathbf{C}$ *is symmetric* if it is self-adjoint, that is,

$$\boldsymbol{A} : \mathbf{C}[\boldsymbol{B}] = \boldsymbol{B} : \mathbf{C}[\boldsymbol{A}] \quad \text{for all } \boldsymbol{A}, \boldsymbol{B} \in \mathcal{L},$$

where the colon is the scalar product on $\mathcal{L}$ defined in Section 7.

According to Proposition 10, $\mathbf{C}[\boldsymbol{W}] = \mathbf{0}$ for all skew-symmetric tensors $\boldsymbol{W}$. Therefore $\mathbf{C}$ as an operator on $\mathcal{L}$ cannot be positive definite. We have seen, however, that $\mathbf{C}$ is completely defined by its restriction to $\mathcal{L}_{\text{sym}}$, therefore we call $\mathbf{C}$ *positive definite* if

$$\boldsymbol{E} : \mathbf{C}[\boldsymbol{E}] > 0 \quad \text{for all nonzero } \boldsymbol{E} \in \mathcal{L}_{\text{sym}}. \tag{200}$$

In the same vein, $\mathbf{C}$ is said to be *strongly elliptic* if

$$\boldsymbol{H} : \mathbf{C}[\boldsymbol{H}] > 0 \quad \text{for all } \boldsymbol{H} = \boldsymbol{a} \otimes \boldsymbol{b}, \text{ where } \boldsymbol{a}, \boldsymbol{b} \in \mathcal{V}. \tag{201}$$

We will leave of for an exercise to show that if $\mathbf{C}$ is positive definite, then it is strongly elliptic (but not vice versa.)

Proposition 12. *If the material is isotropic, then* $\mathbf{C}$ *is symmetric.*

Proof. Pick any $\boldsymbol{E}, \boldsymbol{H} \in \mathcal{L}_{\text{sym}}$. We know that $\boldsymbol{H} : \boldsymbol{I} = \operatorname{tr}\boldsymbol{H}$. Therefore

$$\boldsymbol{H} : \mathbf{C}[\boldsymbol{E}] = 2\mu\boldsymbol{H} : \boldsymbol{E} + \lambda(\operatorname{tr}\boldsymbol{E})(\operatorname{tr}\boldsymbol{H}),$$

which is symmetric in $\boldsymbol{E}$ and $\boldsymbol{H}$. Therefore

$$\boldsymbol{H} : \mathbf{C}[\boldsymbol{E}] = \boldsymbol{E} : \mathbf{C}[\boldsymbol{H}],$$

proving that $\mathbf{C}$ is symmetric. □

Theorem 33. *If the material is isotropic, then* $\mathbf{C}$ *is positive definite if and only if*

$$\mu > 0, \quad 2\mu + 3\lambda > 0, \tag{202}$$

and $\mathbf{C}$ *is strongly elliptic if and only if*

$$\mu > 0, \quad 2\mu + \lambda > 0. \tag{203}$$

[16] If the material in inhomogeneous, the Lamé moduli may vary from point to point.

Proof. Suppose (202) holds. Pick any $\boldsymbol{E} \in \mathcal{L}_{\text{sym}}$, let $\alpha = \frac{1}{3}\operatorname{tr}\boldsymbol{E}$, and define $\boldsymbol{E}_0 = \boldsymbol{E} - \alpha\boldsymbol{I}$. Since $\operatorname{tr}\boldsymbol{I} = 3$, we see that $\operatorname{tr}\boldsymbol{E}_0 = 0$, and therefore

$$\boldsymbol{E} = \boldsymbol{E}_0 + \alpha\boldsymbol{I}, \quad \operatorname{tr}\boldsymbol{E} = 3\alpha, \quad \boldsymbol{E}_0 : \boldsymbol{I} = 0.$$

Now we calculate

$$\begin{aligned}\boldsymbol{E} : \mathbf{C}[\boldsymbol{E}] = 2\mu\boldsymbol{E} : \boldsymbol{E} + \lambda(\operatorname{tr}\boldsymbol{E})^2 &= 2\mu(\boldsymbol{E}_0 + \alpha\boldsymbol{I}) : (\boldsymbol{E}_0 + \alpha\boldsymbol{I}) + (3\alpha)^2\lambda \\ &= 2\mu\big(\|\boldsymbol{E}_0\|^2 + 3\alpha^2\big) + 9\alpha^2\lambda = 2\mu\|\boldsymbol{E}_0\|^2 + 3\alpha^2(2\mu + 3\lambda),\end{aligned}$$

which shows that $\mathbf{C}$ is positive definite. The remaining proofs are left as exercises. □

Let us return to the constitutive equation $\boldsymbol{P} = \hat{\boldsymbol{P}}(\boldsymbol{F})$ of the general elasticity, where $\boldsymbol{F}$ is the deformation gradient and $\boldsymbol{P}$ is the Piola–Kirchhoff stress tensor. Given the deformation $\boldsymbol{x} = \boldsymbol{\phi}_t(\boldsymbol{X})$, the vector $\boldsymbol{u} = \boldsymbol{x} - \boldsymbol{X}$ expresses the *displacement* of the point $\boldsymbol{X}$. We define $\boldsymbol{H} = \mathbf{Grad}\,\boldsymbol{u} = \boldsymbol{F} - \boldsymbol{I}$. Linear elasticity is concerned with deformations where $\boldsymbol{H}$ is small. In view of the definition of the elasticity tensor $\mathbf{C}$, we have

$$\hat{S}(\boldsymbol{F}) = \hat{S}(\boldsymbol{I}) + \mathbf{C}[\boldsymbol{F} - \boldsymbol{I}] + o\big(\|\boldsymbol{F} - \boldsymbol{I}\|^2\big) = \mathbf{C}[\boldsymbol{H}] + o\big(\|\boldsymbol{H}\|^2\big)$$

Letting

$$\boldsymbol{E} = \frac{1}{2}\big(\boldsymbol{H} + \boldsymbol{H}^T\big)$$

we have

$$\boldsymbol{S} = \mathbf{C}[\boldsymbol{E}] + o\big(\|\boldsymbol{H}\|^2\big)$$

In *linear elasticity* we drop the $o\big(\|\boldsymbol{H}\|^2\big)$ term above, and express the constitutive equation as

$$\boldsymbol{S} = \mathbf{C}[\boldsymbol{E}], \quad \boldsymbol{E} = \frac{1}{2}\big(\mathbf{grad}\,\boldsymbol{u} + (\mathbf{grad}\,\boldsymbol{u})^T\big).$$

The the equation of motion takes the form

$$\rho\ddot{\boldsymbol{u}} = \operatorname{div}\boldsymbol{S} + \rho\boldsymbol{b}.$$

If the material is isotropic, then $\boldsymbol{S}$ is given by (199). Moreover, if the material is homogeneous, that is, the Lamé moduli are constants, then

$$\begin{aligned}\operatorname{div}\boldsymbol{S} = 2\mu\operatorname{div}\boldsymbol{E} + \lambda\operatorname{div}\big((\operatorname{tr}\boldsymbol{E})\boldsymbol{I}\big) &= \mu\operatorname{div}\big(\mathbf{grad}\,\boldsymbol{u} + (\mathbf{grad}\,\boldsymbol{u})^T\big) + \lambda\,\mathbf{grad}\operatorname{tr}\boldsymbol{E} \\ &= \mu\Delta\boldsymbol{u} + (\lambda + \mu)\,\mathbf{grad}\operatorname{div}\boldsymbol{u},\end{aligned}$$

and therefore the equation of motion becomes

$$\rho\ddot{\boldsymbol{u}} = \mu\Delta\boldsymbol{u} + (\lambda + \mu)\,\mathbf{grad}\operatorname{div}\boldsymbol{u} + \rho\boldsymbol{b}. \tag{204}$$

Remark 37. It can be shown that the initial value problems corresponding to (204) are well-posed systems of hyperbolic PDEs if $\mathbf{C}$ is *strongly elliptic*, that is, $\mu > 0$ and $2\mu + \lambda > 0$.

In static equilibrium, the equation of motion (204) reduces to

$$\mu\Delta\boldsymbol{u} + (\lambda + \mu)\,\mathbf{grad}\operatorname{div}\boldsymbol{u} + \rho\boldsymbol{b} = \mathbf{0}. \tag{205}$$

Remark 38. It can be shown that boundary value problems corresponding to (205) are well-posed systems of elliptic PDEs if $\mathbf{C}$ is positive definite, that is, $\mu > 0$ and $2\mu + 3\lambda > 0$.

38.4. **Waves in isotropic linear elastic materials.** In this section we consider isotropic elastic materials, therefore the equation of motion is (204).

A displacement field of the form

$$\boldsymbol{u}(\boldsymbol{x}, t) = \boldsymbol{a} \sin(\boldsymbol{x} \cdot \boldsymbol{n} - ct) \tag{206}$$

is called a *sinusoidal progressive wave* traveling in the direction $\boldsymbol{n}$ at velocity c with amplitude $\boldsymbol{a}$. If $\boldsymbol{a}$ and $\boldsymbol{n}$ are collinear, we say that the wave is *longitudinal.* If $\boldsymbol{a}$ and $\boldsymbol{n}$ are perpendicular, we say that the wave is *transverse.*

We plug the displacement field (206) into the equation of motion (204) to see what conclusions we may derive from it. Applying the result of Exercise 22 and the chain rule, we see that

$$\mathbf{grad}\, \boldsymbol{u} = \boldsymbol{a} \otimes \boldsymbol{n} \cos(\boldsymbol{x} \cdot \boldsymbol{n} - ct),$$

and therefore

$$\operatorname{div} \boldsymbol{u} = \operatorname{tr} \mathbf{grad}\, \boldsymbol{u} = \boldsymbol{a} \cdot \boldsymbol{n} \cos(\boldsymbol{x} \cdot \boldsymbol{n} - ct), \qquad \mathbf{curl}\, \boldsymbol{u} = \boldsymbol{n} \times \boldsymbol{a} \cos(\boldsymbol{x} \cdot \boldsymbol{n} - ct).$$

If (206) is a longitudinal wave, then $\boldsymbol{a}$ and $\boldsymbol{n}$ are collinear, and therefore $\boldsymbol{n} \times \boldsymbol{a} = \mathbf{0}$, and consequently $\mathbf{curl}\, \boldsymbol{u} = \mathbf{0}$. If (206) is a transverse wave, then $\boldsymbol{a}$ and $\boldsymbol{n}$ are perpendicular, and therefore $\boldsymbol{a} \cdot \boldsymbol{n} = 0$, and consequently $\operatorname{div} \boldsymbol{u} = 0$.

Furthermore, it can be shown (Exercise 50) that

$$\Delta \boldsymbol{u} \overset{\text{def}}{=} \operatorname{div} \mathbf{grad}\, \boldsymbol{u} = -\boldsymbol{a} \sin(\boldsymbol{x} \cdot \boldsymbol{n} - ct), \tag{207a}$$

$$\mathbf{grad} \operatorname{div} \boldsymbol{u} = -(\boldsymbol{a} \cdot \boldsymbol{n})\, \boldsymbol{n}\, \sin(\boldsymbol{x} \cdot \boldsymbol{n} - ct) = -(\boldsymbol{n} \otimes \boldsymbol{n})\, \boldsymbol{a}\, \sin(\boldsymbol{x} \cdot \boldsymbol{n} - ct), \tag{207b}$$

$$\ddot{\boldsymbol{u}} = -c^2 \boldsymbol{a} \sin(\boldsymbol{x} \cdot \boldsymbol{n} - ct). \tag{207c}$$

Plugging these into (204), with the body force $\boldsymbol{b}$ set to zero, we arrive at

$$\rho c^2 \boldsymbol{a} = \mu \boldsymbol{a} + (\lambda + \mu)(\boldsymbol{n} \otimes \boldsymbol{n})\boldsymbol{a},$$

or equivalently,

$$\frac{1}{\rho}\Big[\mu \boldsymbol{I} + (\lambda + \mu)\boldsymbol{n} \otimes \boldsymbol{n}\Big]\, \boldsymbol{a} = c^2 \boldsymbol{a}.$$

That motivates the introduction of the *acoustic tensor* $\boldsymbol{A}(\boldsymbol{n})$:

$$\boldsymbol{A}(\boldsymbol{n}) = \frac{1}{\rho}\Big[\mu \boldsymbol{I} + (\lambda + \mu)\boldsymbol{n} \otimes \boldsymbol{n}\Big] = \frac{\lambda + 2\mu}{\rho} \boldsymbol{n} \otimes \boldsymbol{n} + \frac{\mu}{\rho}\,(\boldsymbol{I} - \boldsymbol{n} \otimes \boldsymbol{n}), \tag{208}$$

whereby the equation of motion reduces to

$$\boldsymbol{A}(\boldsymbol{n})\boldsymbol{a} = c^2 \boldsymbol{a}, \tag{209}$$

and thus, c^2 is an eigenvalue of $\boldsymbol{A}(\boldsymbol{n})$, and $\boldsymbol{a}$ is the corresponding eigenvector.

Evidently the expression on the right-hand side of (208) is the spectral decomposition of $\boldsymbol{A}(n)$. We see that

- $(\lambda + 2\mu)/\rho$ is an eigenvalue, and the corresponding eigenvector is $\boldsymbol{n}$. Therefore a longitudinal wave travels at the speed $c = \sqrt{(\lambda + 2\mu)/\rho}$;
- μ/ρ is an eigenvalue, and the corresponding eigenvector is orthogonal to $\boldsymbol{n}$. Therefore a transverse wave travels at the speed $c = \sqrt{\mu/\rho}$.

Note that these speeds are real if $\mathbf{C}$ is strongly elliptic.

39. Exercises

1. The proof of Proposition 2 shows that
$$[\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}] = -[\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{w}], \quad [\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}] = -[\boldsymbol{u}, \boldsymbol{w}, \boldsymbol{v}]$$
for all vectors $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$. Conclude that the scalar triple product is invariant under cyclic permutations, that is
$$[\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}] = [\boldsymbol{v}, \boldsymbol{w}, \boldsymbol{u}], \quad [\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}] = [\boldsymbol{w}, \boldsymbol{u}, \boldsymbol{v}].$$

2. Prove (16b) and verify that $\iota_2(\boldsymbol{A})$ is given by (19b).

3. Prove (16c) and verify that $\iota_3(\boldsymbol{A})$ is given by (19c).

4. Verify (21d).

5. Verify (21e).

6. Verify (21f).

7. Verify (21h).

8. Verify (21i).

9. Complete the proof of Corollary 1 by showing that the set of tensors $\mathcal{K}$ defined in (23) is linearly independent. *Hint:* Show that $\alpha_{ij}\boldsymbol{e}_i \otimes \boldsymbol{e}_j = \boldsymbol{0}$ implies that the coefficients α_{ij} are zero.

10. Show that the basis $\mathcal{K}$ defined in (23) is orthonormal with respect to the scalar product (26).

11. Show that $\boldsymbol{I} = \boldsymbol{e}_i \otimes \boldsymbol{e}_i$.

12. Show that $\operatorname{tr} \boldsymbol{A}^T = \operatorname{tr} \boldsymbol{A}$ and $\operatorname{tr}(\boldsymbol{AB}) = \operatorname{tr}(\boldsymbol{BA})$ for all $\boldsymbol{A}, \boldsymbol{B} \in \mathcal{L}$.

13. If $\boldsymbol{W}$ is skew-symmetric, show that
$$\iota_1(\boldsymbol{W}) = 0, \quad \iota_2(\boldsymbol{W}) = \|\boldsymbol{w}\|^2, \quad \iota_3(\boldsymbol{W}) = 0,$$
where $\boldsymbol{w}$ is $\boldsymbol{W}$'s axial vector. *Hint:* Let $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ be a right-handed frame where $\boldsymbol{e}_1$ is an eigenvector of $\boldsymbol{W}$. We know that $\boldsymbol{w} = \omega\, \boldsymbol{e}_1$ and $\boldsymbol{W}\boldsymbol{e}_1 = \boldsymbol{0}$. Apply (37) to calculate Calculate $\boldsymbol{W}\boldsymbol{e}_2$ and $\boldsymbol{W}\boldsymbol{e}_3$. Then apply (19) to calculate the principal invariants.

14. Derive the equations (48).

15. For any $\boldsymbol{A} \in \mathcal{L}$, show that $\boldsymbol{A}^T\boldsymbol{A}$ is symmetric and positive semi-definite. Moreover, if $\boldsymbol{A}$ is invertible, show that $\boldsymbol{A}^T\boldsymbol{A}$ is positive definite. *Hint:* A tensor $\boldsymbol{A}$ is invertible if $\boldsymbol{A}\boldsymbol{u} = \boldsymbol{0}$ implies that $\boldsymbol{u} = \boldsymbol{0}$.

16. For any $\boldsymbol{A} \in \mathcal{L}$, show that $\boldsymbol{A}\boldsymbol{A}^T$ is symmetric and positive semi-definite. Moreover, if $\boldsymbol{A}$ is invertible, show that $\boldsymbol{A}\boldsymbol{A}^T$ is positive definite. *Hint:* You may refer the following without proof in your solution. A tensor $\boldsymbol{A}$ is invertible if $\boldsymbol{A}^T\boldsymbol{u} = \boldsymbol{0}$ implies that $\boldsymbol{u} = \boldsymbol{0}$. This is a consequence of the facts that a tensor is invertible if its determinant is nonzero, and $\det \boldsymbol{A} = \det \boldsymbol{A}^T$ for any tensor $\boldsymbol{A}$.

17. Show that the identities (40a) and (40b) may be expressed as
$$(\boldsymbol{u} \times \boldsymbol{v}) \times \boldsymbol{a} = (\boldsymbol{v} \otimes \boldsymbol{u} - \boldsymbol{u} \otimes \boldsymbol{v})\, \boldsymbol{a}, \tag{210a}$$
$$\boldsymbol{a} \times (\boldsymbol{u} \times \boldsymbol{v}) = (\boldsymbol{u} \otimes \boldsymbol{v} - \boldsymbol{v} \otimes \boldsymbol{u})\, \boldsymbol{a}. \tag{210b}$$
Note that (210a) says that $\boldsymbol{u}\times\boldsymbol{v}$ is the axial vector of the skew-symmetric tensor $\boldsymbol{v}\otimes\boldsymbol{u}-\boldsymbol{u}\otimes\boldsymbol{v}$.

18. Let $\boldsymbol{Q}(t)$ be an orthogonal tensor that depends on time t, and let $\dot{\boldsymbol{Q}}$ be its derivative. Show that $\boldsymbol{W} = \dot{\boldsymbol{Q}}^T\boldsymbol{Q}$ is skew-symmetric.

19. Show that for any invertible tensor $\boldsymbol{A}$ and any pair of vectors $\boldsymbol{u}$, $\boldsymbol{v}$, we have

$$(\boldsymbol{A}\boldsymbol{u}) \times (\boldsymbol{A}\boldsymbol{v}) = (\det \boldsymbol{A})\boldsymbol{A}^{-T}(\boldsymbol{u} \times \boldsymbol{v}). \tag{211}$$

20. Show that the basis $\mathcal{K}$ in (23) is orthonormal with respect to the scalar product defined in (26).

21. Consider the scalar field ϕ and the vector field $\boldsymbol{v}$. Use the index notation to show that

$$\operatorname{div}(\phi\boldsymbol{v}) = (\mathbf{grad}\,\phi) \cdot \boldsymbol{v} + \phi \operatorname{div} \boldsymbol{v}. \tag{212}$$

22. Consider the scalar field ϕ and the vector field $\boldsymbol{v}$. Use the index notation to show that

$$\mathbf{grad}(\phi\boldsymbol{v}) = \boldsymbol{v} \otimes \mathbf{grad}\,\phi + \phi\,\mathbf{grad}\,\boldsymbol{v}. \tag{213}$$

23. Consider the vector fields $\boldsymbol{u}$ and $\boldsymbol{v}$. Use the index notation to show that

$$\operatorname{div}(\boldsymbol{u} \otimes \boldsymbol{v}) = (\mathbf{grad}\,\boldsymbol{u})\,\boldsymbol{v} + \boldsymbol{u}\ \operatorname{div} \boldsymbol{v}. \tag{214}$$

24. Consider the scalar field ϕ and the tensor field $\boldsymbol{A}$. Show that

$$\operatorname{div}(\phi\boldsymbol{A}) = \boldsymbol{A}\,\mathbf{grad}\,\phi + \phi \operatorname{div} \boldsymbol{A}. \tag{215}$$

Suggestion: Expand the left-hand and right-hand sides into components and verify that the two sides are the same.

25. Show that the following identity holds for any pair of vector fields $\boldsymbol{u}$, $\boldsymbol{v}$:

$$\mathbf{grad}(\boldsymbol{u} \cdot \boldsymbol{v}) = (\mathbf{grad}\,\boldsymbol{u})^T\boldsymbol{v} + (\mathbf{grad}\,\boldsymbol{v})^T\boldsymbol{u}. \tag{216}$$

Suggestion: Express $\boldsymbol{u}$ and $\boldsymbol{v}$ in components as in Sections 17.1 and 17.2, and then evaluate the left- and right-hand sides separately to show that they are identical.

26. Show that for any vector field $\boldsymbol{v}$ we have

$$\operatorname{div}\big((\mathbf{grad}\,\boldsymbol{v})^T\big) = \mathbf{grad}(\operatorname{div} \boldsymbol{v}).$$

27. Show that for any tensor field $\boldsymbol{S}$ and vector field $\boldsymbol{u}$, we have

$$\operatorname{div}(\boldsymbol{S}^T\boldsymbol{v}) = (\operatorname{div} \boldsymbol{S}) \cdot \boldsymbol{v} + \boldsymbol{S} : \mathbf{grad}\,\boldsymbol{v}.$$

(This is essentially the same as the problem 2.11(b) on page 67 of [5].)

28 (Signorini's Theorem)**.** Consider a body B in equilibrium under the stress field $\boldsymbol{S}$ and the body force field $\boldsymbol{b}$, that is, $\operatorname{div} \boldsymbol{S} + \rho\boldsymbol{b} = \boldsymbol{0}$. Show that

$$\int_B \boldsymbol{S}\,dV = \int_B \rho\boldsymbol{b} \otimes \boldsymbol{x}\,dV + \int_{\partial B} (\boldsymbol{S}\boldsymbol{n}) \otimes \boldsymbol{x}\,dA, \tag{217}$$

where $\boldsymbol{x}$ is the variable of integration, and $\boldsymbol{n}$ is the outward unit normal to the boundary of B. This result is commonly presented in terms of the the average stress $\overline{\boldsymbol{S}} = \frac{1}{\operatorname{vol}(B)} \int_B \boldsymbol{S}\,dV$ as:

$$\overline{\boldsymbol{S}} = \frac{1}{\operatorname{vol}(B)}\left[\int_B \rho\boldsymbol{b} \otimes \boldsymbol{x}\,dV + \int_{\partial B} (\boldsymbol{S}\boldsymbol{n}) \otimes \boldsymbol{x}\,dA\right]. \tag{218}$$

Hint: Show that for any constant vector $\boldsymbol{a}$ we have $\big((\boldsymbol{S}\boldsymbol{n}) \otimes \boldsymbol{x}\big)\,\boldsymbol{a} = \big((\boldsymbol{a} \cdot \boldsymbol{x})\,\boldsymbol{S}\big)\boldsymbol{n}$, and therefore

$$\int_{\partial B} \big((\boldsymbol{S}\boldsymbol{n}) \otimes \boldsymbol{x}\big)\,\boldsymbol{a}\,dA = \int_B \operatorname{div}\big((\boldsymbol{a} \cdot \boldsymbol{x})\,\boldsymbol{S}\big)\,dV.$$

Then expand $\operatorname{div}\big((\boldsymbol{a} \cdot \boldsymbol{x})\,\boldsymbol{S}\big)$ by applying (215).

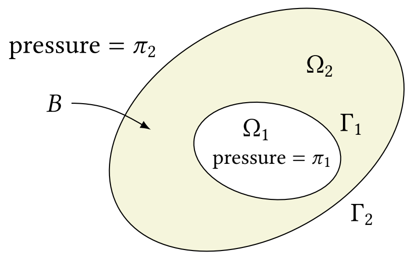


FIGURE 22. (Exercise 29) The domain $B = \Omega_1 \backslash \Omega_1$ is pressurized by π_1 and π_2 from the inside and the outside.

29. Consider two closed surfaces Γ_1 and Γ_2 enclosing the domains Ω_1 and Ω_2 in $\mathbb{E}_3$, where $\Omega_1 \subset \Omega_2$ as in Figure 22. let $B = \Omega_2 \backslash \Omega_1$, that is, the region between those two surfaces, represent a body at rest, subject to uniform pressures π_1 and π_2 applied to its boundary surfaces Γ_1 and Γ_2, and no body forces. Apply Signorini's Theorem (Exercise 28) to show that the average stress $\overline{S}$ within B is given by

$$\overline{S} = -\frac{\pi_2 v_2 - \pi_1 v_1}{v_2 - v_1} \boldsymbol{I},$$

where v_1 and v_2 are the volumes of Ω_1 and Ω_2, and $\boldsymbol{I}$ is the identity tensor. Thus, the average stress is a *pressure* of the amount $\frac{\pi_2 v_2 - \pi_1 v_1}{v_2 - v_1}$.

Hint: Apply (218) to an arbitrary $\boldsymbol{a} \in \mathcal{V}$, and then convert the resulting surface integrals to volume integrals via Gauss's theorem (page 37).

30. The Reynolds Transport Theorem 25 on page 51 applies to scalar fields $\omega(\boldsymbol{x}, t)$. Show the the following counterpart of (121b) holds for vector fields $\boldsymbol{w}(\boldsymbol{x}, t)$:

$$\frac{d}{dt} \int_{\Omega_t} \boldsymbol{w}(\boldsymbol{x}, t)\, dV_{\boldsymbol{x}} = \int_{\Omega_t} \Big(\frac{\partial \boldsymbol{w}}{\partial t} + \operatorname{div}(\boldsymbol{w} \otimes \boldsymbol{v}) \Big)\, dV_{\boldsymbol{x}}.$$

Hint: Apply (121b) to $\boldsymbol{w} \cdot \boldsymbol{a}$, where $\boldsymbol{a}$ is an arbitrary constant vector.
Note: This identity is not needed anywhere in these notes but it makes for a good exercise.

31. Prove Part 2 of Wang's Lemma (page 28).

32. Prove Lemma 8 on page 29.

33. Prove Lemma 10 on page 29.

34. A function $\phi : \mathcal{V} \to \mathbb{R}$ is said to be *isotropic* if

$$\phi(\boldsymbol{Q}\boldsymbol{v}) = \phi(\boldsymbol{v}) \quad \text{for all } \boldsymbol{v} \in \mathcal{V} \text{ and } \boldsymbol{Q} \in \mathcal{L}_{\text{orth}}.$$

Show that ϕ is isotropic if and only if there exists a function $\hat{\phi} : \mathbb{R} \to \mathbb{R}$ so that $\phi(\boldsymbol{v}) = \hat{\phi}\big(\|\boldsymbol{v}\|\big)$ for all $\boldsymbol{v} \in \mathcal{V}$.

Hint: Show that if $\|\boldsymbol{u}\| = \|\boldsymbol{u}'\|$, then $\phi(\boldsymbol{u}) = \phi(\boldsymbol{u}')$. One way of doing this is to pick frames $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ and $\{\boldsymbol{e}'_1, \boldsymbol{e}'_2, \boldsymbol{e}'_3\}$ so that $\boldsymbol{u} = \alpha \boldsymbol{e}_1$ and $\boldsymbol{u}' = \alpha \boldsymbol{e}'_1$, and then let $\boldsymbol{Q}$ be the orthogonal tensor that rotates the frame $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ to frame $\{\boldsymbol{e}'_1, \boldsymbol{e}'_2, \boldsymbol{e}'_3\}$ as in Lemma 5 on page 28.

35. A function $\phi : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ is said to be *isotropic* if

$$\phi(\boldsymbol{Q}\boldsymbol{u}, \boldsymbol{Q}\boldsymbol{v}) = \phi(\boldsymbol{u}, \boldsymbol{v}) \quad \text{for all } \boldsymbol{u}, \boldsymbol{v} \in \mathcal{V} \text{ and } \boldsymbol{Q} \in \mathcal{L}_{\text{orth}}. \tag{219}$$

Show that ϕ is isotropic if and only if there exists a function $\hat{\phi} : \mathbb{R}^3 \to \mathbb{R}$ so that

$$\phi(\boldsymbol{u}, \boldsymbol{v}) = \hat{\phi}\big(\|\boldsymbol{u}\|, \boldsymbol{u} \cdot \boldsymbol{v}, \|\boldsymbol{v}\|\big). \tag{220}$$

Hint: Consider two pairs of vectors, $(\boldsymbol{u}, \boldsymbol{v})$ and $(\boldsymbol{u}', \boldsymbol{v}')$, so that $\|\boldsymbol{u}'\| = \|\boldsymbol{u}\|$, $\boldsymbol{u}' \cdot \boldsymbol{v}' = \boldsymbol{u} \cdot \boldsymbol{v}$, and $\|\boldsymbol{v}'\| = \|\boldsymbol{v}\|$. Install frames $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ and $\{\boldsymbol{e}_1', \boldsymbol{e}_2', \boldsymbol{e}_3'\}$ as in the hint to Exercise 34, and conclude that $\phi(\boldsymbol{u}', \boldsymbol{v}') = \phi(\boldsymbol{u}, \boldsymbol{v})$.

36. A function $\boldsymbol{h} : \mathcal{V} \to \mathcal{V}$ is said to be *isotropic* if

$$\boldsymbol{h}(\boldsymbol{Q}\boldsymbol{v}) = \boldsymbol{Q}\boldsymbol{h}(\boldsymbol{v}) \quad \text{for all } \boldsymbol{v} \in \mathcal{V} \text{ and } \boldsymbol{Q} \in \mathcal{L}_{\text{orth}}.$$

Show that $\boldsymbol{h}$ is isotropic if and only if there exists a function $\hat{h} : \mathbb{R} \to \mathbb{R}$ so that

$$\boldsymbol{h}(\boldsymbol{v}) = \hat{h}\big(\|\boldsymbol{v}\|\big)\boldsymbol{v} \quad \text{for all } \boldsymbol{v} \in \mathcal{V}.$$

Hint: Define $\phi : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ as $\phi(\boldsymbol{u}, \boldsymbol{v}) = \boldsymbol{u} \cdot \boldsymbol{h}(\boldsymbol{v})$. Show that ϕ is isotropic (see Exercise 35) and then simplify the result observing that ϕ is linear in $\boldsymbol{u}$.

37. Referring to (132) on page 56, show that the left Cauchy–Green strain tensors $\boldsymbol{B}$ and $\boldsymbol{B}^*$ of a motion observed by two observers are related through $\boldsymbol{B}^* = \boldsymbol{Q}\boldsymbol{B}\boldsymbol{Q}^T$.

38. Consider the motion $\boldsymbol{\phi}_t$, and let $\boldsymbol{F}(\boldsymbol{X}, t)$ and $\boldsymbol{v}(\boldsymbol{x}, t)$ be the corresponding deformation gradient and the spatial velocity fields, respectively. Show that

$$\dot{\boldsymbol{F}}(\boldsymbol{X}, t) = \operatorname{\mathbf{grad}} \boldsymbol{v}(\boldsymbol{x}, t)\, \boldsymbol{F}(\boldsymbol{X}, t) \quad \text{where } \boldsymbol{x} = \boldsymbol{\phi}_t(\boldsymbol{X}). \tag{221}$$

Hint: Recall the equation $\operatorname{\mathbf{grad}} \boldsymbol{v}(\boldsymbol{x}, t)\ \operatorname{\mathbf{Grad}} \boldsymbol{\phi}_t(\boldsymbol{X}) = \operatorname{\mathbf{Grad}} \dot{\boldsymbol{\phi}}_t(\boldsymbol{X})$ from Section 28.

39. Consider the motion $\boldsymbol{\phi}_t$, and let $\boldsymbol{F}$, $\boldsymbol{C}$, and $\boldsymbol{D}$ be the motion's deformation gradient, the right Cauchy–Green strain tensor, and the strain rate, respectively. Show that

$$\dot{\boldsymbol{C}} = 2\boldsymbol{F}^T\boldsymbol{D}\boldsymbol{F}. \tag{222}$$

Hint: Apply (221).

40. Is the constitutive equation $\boldsymbol{S} = \hat{\boldsymbol{S}}(\boldsymbol{F}) = \boldsymbol{F} + \boldsymbol{F}^T$ frame-indifferent? Here $\boldsymbol{S}$ is the Cauchy stress associated with the deformation gradient $\boldsymbol{F}$.

41. A *visco-elastic material* combines the behavioral characteristics of both an elastic and a viscous material. Its constitutive equation $\boldsymbol{S} = \hat{\boldsymbol{S}}\big(\boldsymbol{F}, \dot{\boldsymbol{F}}\big)$ relates the stress $\boldsymbol{S}$ to the deformation gradient $\boldsymbol{F}$ and its rate of change, $\dot{\boldsymbol{F}}$. Show that frame-indifference restricts the constitutive equation to

$$\boldsymbol{S} = \boldsymbol{R}\hat{\boldsymbol{S}}\big(\boldsymbol{U}, \dot{\boldsymbol{U}}\big)\boldsymbol{R}^T,$$

where $\boldsymbol{F} = \boldsymbol{R}\boldsymbol{U}$ is the deformation gradient's right polar decomposition.

42. Show how equation (184) follows from equation (183).

43. The classical Poiseuille flow concerns the steady-state motion of an incompressible Newtonian fluid pumped through an open-ended cylindrical pipe. That is a special case of the problem studied in Section 31. Here, the constitutive equation is $\boldsymbol{S} = -p\boldsymbol{I} + 2\mu\boldsymbol{D}$, where the constant μ is the fluid's viscosity.

Repeat the calculations of Section 31 for this special case and conclude that

$$p(z) = (c - \rho g)z + p_0, \qquad \phi(r) = -\frac{c}{4\mu}(R^2 - r^2), \tag{223}$$

where c is a (positive) constant, and p_0 is the atmospheric pressure. We note that the pressure varies linearly in z, and that the velocity profile is parabolic. See Exercise 44 on how to determine the constant c.

Hint: Derive the equations of motion and observe that the pressure $p(r, z)$ is independent of r, that is, $p = p(z)$. Set $p(0) = p_0$ to determine one of the integration constants.

44. In Exercise 43 show that the pipe's volumetric flow rate Q, that is, the volume of the fluid passing through a pipe's cross-section per unit time, is given by $Q = \frac{\pi c}{8\mu} R^4$, or equivalently, $Q = \frac{c}{8\pi\mu} A^2$, where A is the pipe's cross-sectional area. This enables us to calculate the constant c in (223) in terms of the easily measurable quantities Q, R, or A.

45. Calculate the resultant $\boldsymbol{T}$ of the forces acting on the slanted face of the cube in Figure 13 of Section 27. Use the Cauchy stress tensor that was calculated in that section. Assume that it is a unit cube.

46. Repeat the previous exercise, but this time use the Piola–Kirchhoff stress tensor to do the calculations.

47. Show that (193) cannot hold if $\lambda_1, \lambda_2, \lambda_3$ are all distinct.

48. Suppose that the elasticity tensor C of Theorem 33 (page 95) is positive definite. Show that $\mu > 0$ and $2\mu + 3\lambda > 0$.

49. Complete the proof of Theorem 33 (page 95) by showing that the elasticity tensor C of an isotropic material is strongly elliptic if and only if $\mu > 0$ and $2\mu + \lambda > 0$.

50. Verify the equations (207). *Hint:* Apply the results of Exercises 21 and 23.

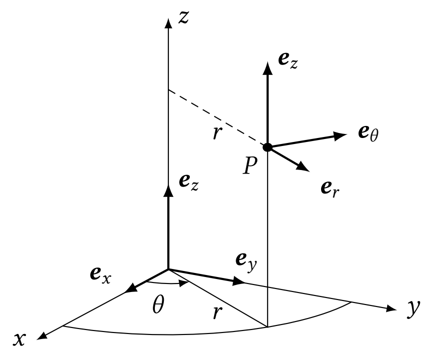


FIGURE 23. This is an illustration of the relationship between a Cartesian and cylindrical coordinates, and the corresponding frames $\{\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z\}$ and $\{\boldsymbol{e}_r, \boldsymbol{e}_\theta, \boldsymbol{e}_z\}$. The point P has cylindrical coordinates r, θ, z.

APPENDIX A. FORMULAS IN CYLINDRICAL COORDINATES

The cylindrical coordinates (r, θ, z) and the corresponding Cartesian coordinates (x, y, z) are related through

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

Vectors and tensors in the Cartesian coordinates are expressed in terms of the frame $\{\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z\}$ which is aligned with the coordinate axes. Vectors and tensors in the cylindrical coordinates are expressed in terms of the frame $\{\boldsymbol{e}_r, \boldsymbol{e}_\theta, \boldsymbol{e}_z\}$ which is related to the Cartesian frame through

$$\boldsymbol{e}_x = \boldsymbol{e}_r\cos\theta - \boldsymbol{e}_\theta\sin\theta, \quad \boldsymbol{e}_y = \boldsymbol{e}_r\sin\theta + \boldsymbol{e}_\theta\cos\theta, \quad \boldsymbol{e}_z = \boldsymbol{e}_z.$$

Figure 23 illustrates the relationship between a Cartesian and cylindrical coordinates, and the corresponding frames.

Thus, a vector $\boldsymbol{v}$ in relative to the frame $\{\boldsymbol{e}_r, \boldsymbol{e}_\theta, \boldsymbol{e}_z\}$ is expressed as

$$\boldsymbol{v} = v_r\boldsymbol{e}_r + v_\theta\boldsymbol{e}_\theta + v_z\boldsymbol{e}_z. \tag{224}$$

The components of a tensor $\boldsymbol{A}$ has relative to the frame $\{\boldsymbol{e}_r, \boldsymbol{e}_\theta, \boldsymbol{e}_z\}$ are expressed as A_{rr}, $A_{r\theta}$, etc. Here is the full matrix of the components of $\boldsymbol{A}$:

$$\boldsymbol{A} = \begin{pmatrix} A_{rr} & A_{r\theta} & A_{rz} \\ A_{\theta r} & A_{\theta\theta} & A_{\theta z} \\ A_{zr} & A_{z\theta} & A_{zz} \end{pmatrix}. \qquad \text{(tensor components relative to the } \{\boldsymbol{e}_r, \boldsymbol{e}_\theta, \boldsymbol{e}_z\} \text{ frame)} \tag{225}$$

In the following list of differentiation formulas, ϕ, $\boldsymbol{v}$, and $\boldsymbol{A}$ are generic scalar, vector, and tensor fields, respectively.

$$\mathbf{grad}\,\phi = \frac{\partial\phi}{\partial r}\boldsymbol{e}_r + \frac{1}{r}\frac{\partial\phi}{\partial\theta}\boldsymbol{e}_\theta + \frac{\partial\phi}{\partial z}\boldsymbol{e}_z \tag{226}$$

$$\mathbf{grad}\,\boldsymbol{v} = \begin{pmatrix} \dfrac{\partial v_r}{\partial r} & \dfrac{1}{r}\dfrac{\partial v_r}{\partial\theta} - \dfrac{v_\theta}{r} & \dfrac{\partial v_r}{\partial z} \\ \dfrac{\partial v_\theta}{\partial r} & \dfrac{1}{r}\dfrac{\partial v_\theta}{\partial\theta} + \dfrac{v_r}{r} & \dfrac{\partial v_\theta}{\partial z} \\ \dfrac{\partial v_z}{\partial r} & \dfrac{1}{r}\dfrac{\partial v_z}{\partial\theta} & \dfrac{\partial v_z}{\partial z} \end{pmatrix} \qquad \text{(tensor components relative to the } \{\boldsymbol{e}_r, \boldsymbol{e}_\theta, \boldsymbol{e}_z\} \text{ frame)} \tag{227}$$

$$\operatorname{div}\boldsymbol{v} = \frac{1}{r}\frac{\partial(rv_r)}{\partial r} + \frac{1}{r}\frac{\partial v_\theta}{\partial\theta} + \frac{\partial v_z}{\partial z} \tag{228}$$

$$\mathbf{curl}\,\boldsymbol{v} = \left(\frac{1}{r}\frac{\partial v_z}{\partial\theta} - \frac{\partial v_\theta}{\partial z}\right)\boldsymbol{e}_r + \left(\frac{\partial v_r}{\partial z} - \frac{\partial v_z}{\partial r}\right)\boldsymbol{e}_\theta + \frac{1}{r}\left(\frac{\partial(rv_\theta)}{\partial r} - \frac{\partial v_r}{\partial\theta}\right)\boldsymbol{e}_z \tag{229}$$

$$\Delta\phi = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial\phi}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2\phi}{\partial\theta^2} + \frac{\partial^2\phi}{\partial z^2} \tag{230}$$

$$\Delta\boldsymbol{v} = \left(\Delta v_r - \frac{1}{r^2}v_r - \frac{2}{r^2}\frac{\partial v_\theta}{\partial\theta}\right)\boldsymbol{e}_r + \left(\Delta v_\theta - \frac{1}{r^2}v_\theta + \frac{2}{r^2}\frac{\partial v_r}{\partial\theta}\right)\boldsymbol{e}_\theta + \left(\Delta v_z\right)\boldsymbol{e}_z \tag{231}$$

$$\begin{aligned}(\mathbf{grad}\,\boldsymbol{v})\boldsymbol{v} = {} & \left(v_r\frac{\partial v_r}{\partial r} + \frac{v_\theta}{r}\frac{\partial v_r}{\partial\theta} + v_z\frac{\partial v_r}{\partial z} - \frac{v_\theta^2}{r}\right)\boldsymbol{e}_r \\ & + \left(v_r\frac{\partial v_\theta}{\partial r} + \frac{v_\theta}{r}\frac{\partial v_\theta}{\partial\theta} + v_z\frac{\partial v_\theta}{\partial z} + \frac{v_r v_\theta}{r}\right)\boldsymbol{e}_\theta \\ & + \left(v_r\frac{\partial v_z}{\partial r} + \frac{v_\theta}{r}\frac{\partial v_z}{\partial\theta} + v_z\frac{\partial v_z}{\partial z}\right)\boldsymbol{e}_z \end{aligned} \tag{232}$$

$$\begin{aligned}\operatorname{div}\boldsymbol{A} = {} & \left(\frac{\partial A_{rr}}{\partial r} + \frac{1}{r}\frac{\partial A_{r\theta}}{\partial\theta} + \frac{\partial A_{rz}}{\partial z} + \frac{1}{r}\left(A_{rr} - A_{\theta\theta}\right)\right)\boldsymbol{e}_r \\ & + \left(\frac{\partial A_{\theta r}}{\partial r} + \frac{1}{r}\frac{\partial A_{\theta\theta}}{\partial\theta} + \frac{\partial A_{\theta z}}{\partial z} + \frac{1}{r}\left(A_{\theta r} + A_{r\theta}\right)\right)\boldsymbol{e}_\theta \\ & + \left(\frac{\partial A_{zr}}{\partial r} + \frac{1}{r}\frac{\partial A_{z\theta}}{\partial\theta} + \frac{\partial A_{zz}}{\partial z} + \frac{A_{zr}}{r}\right)\boldsymbol{e}_z \end{aligned} \tag{233}$$

References

[1] Augustin-Louis Cauchy. De la pression ou tension dans un corps solide (published 1827). In *Exercices de Mathématiques, Œuvres Complètes D'Augustin Chaucy*, volume VII, pages 60–78. Gauthier-Villars, Paris, 1889. (Cited on p. 45)

[2] P. Chadwick. *Continuum mechanics*. Dover Publications, Inc., Mineola, NY, second edition, 1999. Concise theory and problems. (Cited on pp. 1, 47, 80, 85)

[3] S. Chapman and E. A. Milne. The proof of the formula for the vector triple product. *The Mathematical Gazette*, 23(253):35–38, 1939. (Cited on p. 20)

[4] C. Cuvelier and R. M. S. M. Schulkes. Some numerical methods for the computation of capillary free boundaries governed by the Navier-Stokes equations. *SIAM Rev.*, 32(3):355–423, 1990. (Cited on p. 80)

[5] Oscar Gonzalez and Andrew M. Stuart. *A first course in continuum mechanics*. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 2008. (Cited on pp. 1, 31, 47, 55, 56, 92, 99)

[6] A. E. Green and W. Zerna. *Theoretical elasticity*. Clarendon Press, Oxford, second edition, 1968. (Cited on pp. 80, 85)

[7] Morton E. Gurtin. *An introduction to continuum mechanics*, volume 158 of *Mathematics in Science and Engineering*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1981. (Cited on pp. 1, 29, 31, 47, 92, 93)

[8] Morton E. Gurtin. *Topics in finite elasticity*, volume 35 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, Pa., 1981. (Cited on pp. 87, 92)

[9] A. C. Merrington. Flow of visco-elastic materials in capillaries. *Nature*, 152:663, December 1943. (Cited on pp. 71, 74)

[10] J. G. Oldroyd. On the formulation of rheological equations of state. *Proc. Roy. Soc. London Ser. A*, 200:523–541, 1950. (Cited on p. 55)

[11] S. L. Rathna. Coulette and Poiseuille flow in non-Newtonian fluids. *Proc. Nat. Inst. Sci. India Part A*, 26:392–399, 1960. (Cited on pp. 71, 73, 74, 75, 79)

[12] S. L. Rathna and P. L. Bhatnagar. Weissenberg and Merrington effects in non-Newtonian fluids. *J. Indian Inst. Sci.*, 45:57–82, 1963. (Cited on pp. 75, 79)

[13] R. S. Rivlin. Torsion of a rubber cylinder. *Journal of Applied Physics*, 18:444–449, May 1947. (Cited on p. 85)

[14] R. S. Rivlin. Large elastic deformations of isotropic materials. I. Fundamental concepts. *Philos. Trans. Roy. Soc. London Ser. A*, 240:459–490, 1948. (Cited on p. 91)

[15] R. S. Rivlin. Large elastic deformations of isotropic materials. II. Some uniqueness theorems for pure, homogeneous deformation. *Philos. Trans. Roy. Soc. London Ser. A*, 240:491–508, 1948. (Cited on pp. 87, 91)

[16] R. S. Rivlin. Large elastic deformations of isotropic materials. III. Some simple problems in cylindrical polar coordinates. *Philos. Trans. Roy. Soc. London Ser. A*, 240:509–525, 1948. (Cited on p. 85)

[17] R. S. Rivlin. Large elastic deformations of isotropic materials. IV. Further developments of the general theory. *Philos. Trans. Roy. Soc. London Ser. A*, 241:379–397, 1948. (Cited on p. 85)

[18] R. S. Rivlin. A uniqueness theorem in the theory of highly-elastic materials. *Proc. Cambridge Philos. Soc.*, 44:595–597, 1948. (Cited on pp. 87, 91)

[19] R. S. Rivlin. A note on the torsion of an incompressible highly-elastic cylinder. *Proc. Cambridge Philos. Soc.*, 45:485–487, 1949. (Cited on p. 85)

[20] R. S. Rivlin. Stability of pure homogeneous deformations of an elastic cube under dead loading. *Quarterly of Applied Mathematics*, 32(3):265–271, 1974. (Cited on pp. 91, 92)

[21] R. S. Rivlin and J. L. Ericksen. Stress-deformation relations for isotropic materials. *J. Rational Mech. Anal.*, 4:323–425, 1955. (Cited on p. 29)

[22] James Serrin. Poiseuille and Couette flow on non-Newtonian fluids. *Z. Angew. Math. Mech.*, 39:295–299, 1959. (Cited on pp. 75, 79)

[23] C. Truesdell. Notes on the history of the general equations of hydrodynamics. *Amer. Math. Monthly*, 60:445–458, 1953. (Cited on p. 68)

[24] K. Weissenberg. A continuum theory of rheological phenomena. *Nature*, 159:310–311, March 1947. (Cited on p. 75)

[25] Zhimin Zhang and Ivo Babuška. A numerical method for steady state free boundary problems. *SIAM J. Numer. Anal.*, 33(6):2184–2214, 1996. (Cited on p. 80)

Email address: `rostamian@umbc.edu`

Department of Mathematics and Statistics, UMBC