Honors Theses

Ogden Honors College

5-2015

Modeling Deformations of Solids at the Continuum Scale

Paxton Turner

Modeling Deformations of Solids at the Continuum Scale

by

Paxton Turner

Undergraduate honors thesis under the direction of

Dr. Blaise Bourdin

Department of Mathematics

Submitted to the LSU Honors College in partial fulfillment of
the Upper Division Honors Program.

May, 2015

Louisiana State University
& Agricultural and Mechanical College
Baton Rouge, Louisiana

Modeling Deformations of Solids at the Continuum Scale

Paxton Turner
Advisor: Blaise Bourdin

1 Introduction

In this thesis we study continuum mechanics, which comprises modeling the mechanical response of a body when subjected to applied forces. In particular, our focus is on elastic materials, which behave like springs in the sense that the restoring force is proportional to the size of the displacement. For this reason, in the literature they are sometimes referred to as Neo-Hookean materials, in reference to the classical case of an idealized spring studied by the British physicist Robert Hooke. An essential property of elastic materials is that their material response or *stress* is completely determined by the applied forces or *strain*. Moreover, the past history of deformations is independent of the material response. The goal of this thesis is to formalize these intuitive physical ideas mathematically in terms of a series of constitutive equations. Section 2 describes the background and notation needed to understand this thesis, consisting mainly of linear algebra and multivariable calculus. Our main section, Section 3, derives the constitutive equations for elasticity from the classical momentum balance laws and formulates an optimization problem that is equivalent to a linearized elasticity problem. The latter is known as a *variational formulation* of linearized elasticity, and its presentation culminates this thesis.

Of course there are many other possiblities for modeling material behavior. A suitable model for future inquiry is elastoplaticity. Elastoplatic materials behave elasticity until the loading exceeds a certain threshold, at which point the chemical bonds in the molecular lattice of the material have been restructured, and the stress-strain relationship of the material is no longer the same as before.

As a simple example, consider a paperclip. If it is bent a small amount, then the paperclip snaps back into its original configuration. That is, it behaves like a spring. However, suppose that the inner ring of the paperclip is bent out $180°$, and then forcibly bent back into the original configuration. If the original small deformation is reapplied, the paperclip will no longer snap back into its original configuration. Rather, the large loading applied to the paperclip has permanently changed the structure of its molecular bonds, and thus, its original stress-strain relationship has been permanently altered.

As in the case of elasticity, there is a classical and variational formulation of the elastoplastic theory. Optimization algorithms for solving such problems would serve as nice segue into numerical analysis and convexity. These are tasks for a future project.

2 Mathematical Preliminaries

2.1 Linear Algebra: Vectors, Tensors, Trace, Determinant

This section follows Chapter 1 and the Appendix of Gurtin's text [Gur81a] closely.

The fundamental objects of this study are *vectors* $v \in \mathbb{R}^3$, tuples of real numbers, and *linear transformations* or *second-order tensors* $L : \mathbb{R}^3 \to \mathbb{R}^3$, which can be represented as 3×3 matrices with real entries. For brevity, throughout this section, second-order tensors will be referred to as *tensors*[1]. Moreover, $\mathbb{R}^3$ is equipped with the standard Euclidean dot product $\cdot$ (also called inner product). If $u = (u_1, u_2, u_3), v = (v_1, v_2, v_3) \in \mathbb{R}^3$, then the dot product of u and v is

$$u \cdot v := u_1 v_1 + u_2 v_2 + u_3 v_3.$$

Suppose V is a vector space. In general, an inner product $\cdot$ on V must satisfy three properties. Suppose $u, v, w \in V$.

1. Positive definiteness: $v \cdot v \geq 0$ and $v \cdot v = 0$ iff $v = 0$.
2. Bilinearity: Given $a, b \in \mathbb{R}$, $av + bw \cdot u = a(v \cdot u) + b(w \cdot u)$
3. Symmetry: $v \cdot w = w \cdot v$.

This inner product gives rise to the Euclidean norm:

$$|u| := \sqrt{u \cdot u}.$$

A *linear functional* on V is a map $L : V \to \mathbb{R}$ that is linear. The following Riesz representation theorem of linear functionals extends to much more general spaces (Hilbert spaces and even Banach spaces), but here we only need it for finite dimensional inner product spaces.

Theorem 2.1. *Suppose L is a linear functional on a finite dimensional inner product space V. Then there exists a unique vector $a \in V$ such that*

$$L(v) = a \cdot v \tag{2.1}$$

for all $v \in V$.

Proof. First, suppose that such an a exists. Then we show that a must be unique. Supposing a' also satisfies Equation 2.1, then using the bilinearity of the dot product, we have

$$0 = a \cdot v - a' \cdot v = (a - a') \cdot v$$

for all $v \in V$. In particular, setting $v = a - a'$ and applying the positive-definiteness of the inner product implies $a - a' = 0$. Uniqueness is proved.

Now for existence. Using the Graham-Schmidt process, construct an orthonormal basis $\{v_1, \ldots, v_n\}$ of V. Set $a = \sum_{i=1}^n L(v_i)v_i$. Suppose $w \in V$ and

[1] In Section 3 when introducing the generalized Hooke's law, we will be more precise to indicate the order of a given tensors

express it with respect to the selected basis: $w = \sum_{i=1}^n w_i v_i$. Then using the orthonormality of $\{v_i\}_{i=1}^n$, we have

$$L(w) = \sum_{i=1}^n w_i L(v_i) = \left(\sum_{i=1}^n w_i v_i\right) \cdot \left(\sum_{i=1}^n L(v_i) v_i\right) = a \cdot w$$

as desired.

$\square$

The above theorem is especially useful for constructing the gradient of a smooth function, as defined in Section 2.3.

Let $\mathrm{Lin} = \mathrm{Lin}(V)$ denote the space of tensors on V. Another consequence of Theorem 2.1 is the existence and uniqueness of the transpose, which we state without proof.

Definition 1. *Let $S \in \mathrm{Lin}$. Let S^{T} denote the unique tensor such that*

$$Su \cdot v = u \cdot S^{\mathrm{T}} v$$

for all $u, v \in V$. The map S^{T} is known as the transpose *of S.*

The transpose enjoys several nice properties:

1. $(S + T)^{\mathrm{T}}$
2. $(ST)^{\mathrm{T}} = T^{\mathrm{T}} S^{\mathrm{T}}$
3. $(S^{\mathrm{T}})^{\mathrm{T}}$.

A tensor is called *symmetric* if $S = S^{\mathrm{T}}$ and is called *skew* if $S = -S^{\mathrm{T}}$. Moreover, any tensor S can be uniquely decomposed into a skew part W and symmetric part E as follows:

$$S = E + W$$

where

$$E = \frac{1}{2}(S + S^{\mathrm{T}})$$
$$W = \frac{1}{2}(S - S^{\mathrm{T}}).$$

Skew tensors W on $\mathbb{R}^3$ also have the nice property that they can be regarded as a cross product. That is, there exists a unique vector u such that, for any W,

$$Wv = u \times v$$

for all $v \in \mathbb{R}^3$. This correspondence can be written explicitly as follows. If $W = \begin{pmatrix} 0 & -u_3 & u_2 \\ u_2 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix}$, then $u = (u_1, u_2, u_3)$. The vector u is known as the axis of the skew tensor, and later we will see that it represents the infinitesimal axis of rotation of a deforming rigid body.

Suppose $e_1, e_2, \ldots, e_n$ is an orthonormal basis on V (for example, the standard basis on $\mathbb{R}^3$). Define a tensor $e_i \otimes e_j$ by

$$(e_i \otimes e_j)v = (e_j \cdot v)e_i$$

for all $v \in V$. It is a simple task to show that the collection $\{e_i \otimes e_j\}$ forms a basis of Lin and that given a tensor S, if we write

$$S = \sum_{i,j} S_{ij} e_i \otimes e_j,$$

then the S_{ij} are the matrix entries of then tensor S when it is written with respect to the standard basis. Observe that $(e_i \otimes e_i)v$ is the projection of v onto the standard basis vector e_i.

The *trace* $tr(S)$ of a tensor S is defined to be $tr(S) := \sum_{i=1}^n S_{ii}$. One can also check that the trace satisfies $tr(u \otimes v) = u \cdot v$. Using the trace, one obtains an inner product structure on Lin given by the *contraction* operator also denoted by $\cdot$ and defined as follows

$$S \cdot T := tr(S^{\mathrm{T}} T).$$

In terms of coordinates, this can be written

$$S \cdot T = \sum_{i,j} S_{ij} T_{ij}.$$

Let Skw denote the skew tensors, and let Sym denote the symmetric tensors. Then the following proposition can be used to verify that Lin = Skw $\oplus$ Sym is an orthogonal decomposition, meaning that if $S \in$ Sym and $T \in$ Skw, then $S \cdot T = 0$.

Proposition 2.2. *Skew and symmetric tensors have the following properties.*

1. *If* $S \in \mathrm{Sym}$*, then* $S \cdot T = S \cdot T^{\mathrm{T}} = S \cdot \{\frac{1}{2}(T + T^{\mathrm{T}})\}$.
2. *If* $W \in \mathrm{Skw}$*, then* $W \cdot T = -W \cdot T^{\mathrm{T}} = W \cdot \{\frac{1}{2}(T - T^{\mathrm{T}})\}$.
3. *If* $S \in \mathrm{Sym}$ *and* $W \in \mathrm{Skw}$*, then* $S \cdot W = 0$.

Another important function on tensors is the determinant. In general, this operation is a bit complicated to define. Let S_n denote the symmetric group on n elements. An element $\sigma \in S_n$ is a permutation or bijection $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$. It is a fact of elementary group theory that any $\sigma \in S_n$ can be written as a composition of transpositions, where a transposition (ij) is a bijection leaving all elements of $\{1, \ldots, n\}$ fixed except for i and j, which are swapped. If $\sigma \in S_n$ can be written as a composition of an odd (respectively, even) number of transpositions, then we say that σ is odd (respectively, even). Moreover, it is known that no permuation can be both even and odd.

With this background, we define the sign $\mathrm{sgn}(\sigma)$ of a permutation $\sigma \in S_n$:

$$\mathrm{sgn}(\sigma) = \begin{cases} 1 & : \sigma \text{ even} \\ -1 & : \sigma \text{ odd} \end{cases}$$

$$\det S := \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) \prod_{i \in \sigma} S_{\sigma(i)i}.$$

While this definition looks complicated, for our purposes we are generally interested in 3×3 determinants, which can be computed explicitly using Sarrus's rule:

$$\det\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = aei + bfg + cdh - (bdi + afh + ceg).$$

Some fundamental properties of the determinant are stated below.

Proposition 2.3. *Let $S, T \in \mathrm{Lin}$. Then*

1. $\det(ST) = \det(S)\det(T)$
2. $\det(S) = \det(S^{\mathrm{T}})$
3. *S is invertible if and only if* $\det(S) \neq 0$

Proof. Below is a brute-force proof of the first property. The main idea needed is the simple fact that the number of odd permutations in S_n is the same as the number of even permutations in S_n.

$$det(S)det(T) = \left(\sum_{\alpha\in S_n} sgn(\alpha) S_{\sigma(1)1} \dots S_{\sigma(n)n}\right)\left(\sum_{\beta\in S_n} sgn(\beta) T_{\beta(1)1} \dots T_{\beta(n)n}\right)$$

Next, by matrix multiplication and definition of the determinant:

$$det(ST) = \sum_{\sigma\in S_n} sgn(\sigma)\left(\sum_{i_1} S_{\sigma(1)i_1} T_{i_1 1}\right)\cdots\left(\sum_{i_n} S_{\sigma(n)i_n} T_{i_n n}\right) = \sum_{\sigma\in S_n} sgn(\sigma) \sum_{(i_1,\dots,i_n)\in[0,n]^n} S_{\sigma(1)i_1}\cdots S_{\sigma(n)i_n} T_{i_1 1}\cdots T_{i_n n}.$$

There are a lot of cancellations in the final sum. This enables us to show that the two expressions are equal.

Fix a tuple $\{i_1, \dots, i_n\}$ with not all terms distinct and consider the sum

$$\sum_{\sigma\in S_n} sgn(\sigma) S_{\sigma(1)i_1}\cdots S_{\sigma(n)i_n} T_{i_1 1}\cdots T_{i_n n}.$$

We claim that this sum is 0. Write the above as:

$$\sum_{\sigma even} S_{\sigma(1)i_1}\cdots S_{\sigma(n)i_n} T_{i_1 1}\cdots T_{i_n n} - \sum_{\theta odd} S_{\theta(1)i_1}\cdots S_{\theta(n)i_n} T_{i_1 1}\cdots T_{i_n n} \tag{2.2}$$

Choose $i_k = i_\ell$ s.t. $k \neq \ell$. For σ even, act by t_σ, the transposition swapping $\sigma(k)$ and $\sigma(\ell)$. We get an odd permutation $t_\sigma\sigma$ satisfying:

$$S_{\sigma(1)i_1}\cdots S_{\sigma(n)i_n} T_{i_1 1}\cdots T_{i_n n} - S_{t_\sigma\sigma(1)i_1}\cdots S_{t_\sigma\sigma(n)i_n} T_{i_1 1}\cdots T_{i_n n} = 0$$

This process is clearly reversible so we have a bijection from the even permutations to odd permutations. We thus write the above sum 2.2 as

$$\sum_{\sigma even} S_{\sigma(1)i_1}\cdots S_{\sigma(n)i_n}T_{i_1 1}\cdots T_{i_n n} - \sum_{\sigma even} S_{t_\sigma \sigma(1)i_1}\cdots S_{t_\sigma \sigma(n)i_n}T_{i_1 1}\cdots T_{i_n n}$$

Thus, all terms with not all of the i_k distinct cancel out in the total sum.

Hence, we can write the previous sum as a sum over all $\{i_1, \ldots, i_n\}$ where all the i_k are distinct.

$$\begin{aligned}
det(ST) &= \sum_{\sigma\in S_n} sgn(\sigma) \sum_{(i_1,\ldots,i_n)\in[0,n]^n} S_{\sigma(1)i_1}\cdots S_{\sigma(n)i_n}T_{i_1 1}\cdots T_{i_n n}\\
&= \sum_{\sigma\in S_n} sgn(\sigma) \sum_{\substack{(i_1,\ldots,i_n)\in[0,n]^n\\ i_k \text{ distinct}}} S_{\sigma(1)i_1}\cdots S_{\sigma(n)i_n}T_{i_1 1}\cdots T_{i_n n}\\
&= \sum_{\sigma\in S_n} sgn(\sigma) \sum_{\theta\in S_n} S_{\sigma(1)\theta(1)}\cdots S_{\sigma(n)\theta(n)}T_{\theta(1)1}\cdots T_{\theta(n)n}\\
&= \sum_{\sigma\in S_n} sgn(\sigma) \sum_{\theta\in S_n} S_{\sigma\theta^{-1}(1)1}\cdots S_{\sigma\theta^{-1}(n)n}T_{\theta(1)1}\cdots T_{\theta(n)n}\\
&= \sum_{\beta\in S_n} sgn(\beta\theta) \sum_{\theta\in S_n} S_{\beta(1)1}\cdots S_{\beta(n)n}T_{\theta(1)1}\cdots T_{\theta(n)n}\\
&= \sum_{\beta\in S_n}\sum_{\theta\in S_n} sgn(\beta)sgn(\theta)S_{\beta(1)1}\cdots S_{\beta(n)n}T_{\theta(1)1}\cdots T_{\theta(n)n}\\
&= det(T) * det(S)
\end{aligned}$$

where in the third to last line we changed variables: $\beta = \sigma\theta^{-1}$.

□

There is a subspace of Lin that we would like to define, namely Orth, the set of all tensors Q satisfying $QQ^{\mathrm{T}} = I$, where I is the identity. These tensors preserve the inner product in the sense that

$$Qu \cdot Qv = u \cdot v.$$

A *rotation* is an orthogonal tensor whose determinant is positive. Later, we will see that rigid deformations of a material have a gradient that is an rotation. In that sense, such a transformation does not put any *energy* into the system.

2.2 Linear Algebra: Spectral Theorem, Polar Decomposition, and Isotropic Tensor Functions

The fundamental theorems needed for continuum mechanics are the spectral theorem for symmetric tensors and the polar decomposition theorem, which states, informally, that any linear transformation can be described as a rotation composed with what is known as a *positive definite* and symmetric tensor. A tensor T is positive definite if it satisfies $Tu \cdot u > 0$ for all $u \neq 0$. The latter theorem provides an analogy at the level of tensors to the polar form of complex numbers, where a complex number can be represented as a magnitude one complex number (a direction, essentially) and a real number (its magnitude).

To describe the spectral theorem, we must cover some preliminaries on eigenvalues.

Definition 2. *A complex number* λ *is said to be an* eigenvalue *of a tensor* S *if* $Sv = \lambda v$ *for some* $v \neq 0$. *In this case,* v *is known as an* eigenvector *of* S. *The subspace* E_λ *of all vectors* v *satisfying* $Sv = \lambda v$ *is known as the* eigenspace *corresponding to* λ.

It is also easy to show that the eigenvalues of S are precisely the real roots in λ of the polynomial equation $\det(S - \lambda I) = 0$. The polynomial $\det(S - \lambda I)$ is known as the characteristic polynomial of S.

Proposition 2.4. *The eigenvalues of a positive definite symmetric tensor* S *are real and positive. Moreover, if* λ, λ' *are eigenvalues of* S *and* $\lambda \neq \lambda'$, *then* $E_\lambda \perp E_{\lambda'}$.

Proof. If λ is an eigenvalue of S, let v be a corresponding eigenvector. Then $Sv \cdot v = \lambda(v \cdot v) > 0$, which, by the positive-definiteness of the inner product implies that $\lambda \in \mathbb{R}$ and $\lambda > 0$. For the second part, suppose $\lambda' \neq \lambda$ is another eigenvalue with eigenvector v'. Then by the symmetry of S,

$$Sv \cdot v' = v \cdot Sv' = \lambda(v \cdot v') = \lambda'(v \cdot v') \Rightarrow v \cdot v' = 0$$

since $\lambda \neq \lambda'$.

$\square$

Theorem 2.5. *Suppose* S *is a symmetric linear transformation on* $\mathbb{R}^n$. *Then there exists a basis* $e_1, e_2, \ldots, e_n$ *of mutually orthogonal eigenvectors with corresponding eigenvalues* $\lambda_1, \lambda_2, \ldots, \lambda_n$, *respectively, such that*

$$S = \sum_{i=1}^{n} \lambda_i e_i \otimes e_i.$$

Proof. We offer a sketch of the proof of this result. If S is symmetric, then it can be shown that S has n eigenvalues (including multiplicity). That is, when S is naturally regarded as an operator on $\mathbb{C}^n$, all of its complex eigenvalues turn out to be real. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of S. Then the eigenspaces $E_{\lambda_1}, \ldots, E_{\lambda_m}$ are mutually orthogonal. Select an orthonormal basis B_{λ_i} for each eigenspace E_{λ_i}. Then $B = \cup B_{\lambda_i}$ is an orthonormal basis for $\mathbb{R}^n$ consisting entirely of eigenvectors of S. If we write S with respect to the basis B, then we see that S has precisely the desired form. $\square$

In dimension 3, the previous theorem takes a particularly nice form.

Corollary 2.6. *Let* S *be a symmetric tensor on* $\mathbb{R}^3$. *Then*

1. *S has three distinct eigenvalues if and only if the eigenspaces of S are three mutually perpendicular one-dimensional subspaces.*

2. *S has exactly two distinct eigenvalues if and only if S admits a representation*

$$S = \lambda_1 e \otimes e + \lambda_2(I - e \otimes e), \; |e| = 1, \lambda_1 \neq \lambda_2.$$

Observe that $I - e \otimes e$ is the orthogonal projection onto the plane perpendicular to e.

3. S has one distinct eigenvalue if and only if $S = \lambda I$.

To formulate the polar decomposition theorem, we need a notion of the square-root of a positive definite symmetric tensor.

Lemma 2.7. *Let S be positive definite and symmetric. Then there is a unique positive definite and symmetric tensor U such that*

$$U^2 = S.$$

Notationally, we write that $U = \sqrt{S}$.

Proof. We only show existence. Use Theorem 2.5 to write $S = \sum_{i=1}^n \lambda_i e_i \otimes e_i$ where the λ_i are (positive) eigenvalues and the e_i are orthonormal eigenvectors. Then set $U = \sum_{i=1}^n \sqrt{\lambda_i} e_i \otimes e_i$. □

Now we can state the polar decomposition theorem. It is not difficult to show, but we omit the proof.

Theorem 2.8. *Let $M \in \mathrm{Lin}$ and suppose M satisfies $\det M > 0$. Then there exists positive definite symmetric tensors U, V and a rotation R such that*

$$F = RU = VR.$$

This decomposition is also unique: one can show that $U = \sqrt{F^{\mathrm{T}} F}$ and $V = \sqrt{F F^{\mathrm{T}}}$.

Physically, the above theorem will later allow us to represent any deformation of a body as rotation (rigid deformation) and a series of stretches along the three directions given by the eigenvectors.

Remark 1. *For the rest of this section, all tensors are assumed to be on $\mathbb{R}^3$. Thus, Orth refers to the 3×3 orthogonal matrices.*

The *spectrum* of a positive definite and symmetric tensor S is an ordered list of eigenvalues $\{\lambda_1, \lambda_2, \lambda_3\}$ where $\lambda_1 \leq \lambda_2 \leq \lambda_3$ and each eigenvalue is repeated in this list a number of times equal to its multiplicity. The *principle invariants* of a tensor are given by

$$\mathcal{I}_S = \{\iota_1(S), \iota_2(S), \iota_3(S)\} = \{tr(S), \frac{1}{2}(tr(S)^2 - tr(S^2)), \det(S)\}.$$

It can be shown that

$$\begin{aligned} tr(S) &= \lambda_1 + \lambda_2 + \lambda_3 \\ \frac{1}{2}(tr(S)^2 - tr(S^2)) &= \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3 \\ \det(S) &= \lambda_1\lambda_2\lambda_3. \end{aligned}$$

In fact, these are just the coefficients of λ in the characteristic polynomial $\det(S - \lambda I)$. Therefore, there is a one-to-one correspondence between the spectrum of a tensor and its principle invariants. These notions are central to the theory of continuum mechanics in the case when the material is *isotropic*, as we shall see. Now we plunge into the mathematical notion of isotropy, which is, roughly speaking, a sort of stability under the action of Orth.

Definition 3. *Let $\mathcal{G} \subset \mathrm{Orth}$. A set $\mathcal{A} \subset \mathrm{Sym}$ is* invariant under $\mathcal{G}$ *if $QAQ^{\mathrm{T}} \in \mathcal{A}$ whenever $A \in \mathcal{A}$ and $Q \in \mathcal{G}$. We are primarily interested in the case where $\mathcal{G} = \mathrm{Orth}$, in which case we say that A is* isotropic. *Moreover a scalar function $\phi : \mathcal{A} \to \mathbb{R}$ (respectively, tensor function $G : \mathcal{A} \to \mathrm{Lin}$) is said to be* invariant under $\mathcal{G}$ *if $\phi(A) = \phi(QAQ^{\mathrm{T}})$ (respectively, $QG(A)Q^{\mathrm{T}} = G(QAQ^{\mathrm{T}})$) for every $A \in \mathcal{A}$ and $Q \in \mathcal{G}$. Likewise, we are interested in the case when $\mathcal{G} = \mathrm{Orth}$ in which case we say that ϕ (respectively, G) is* isotropic.

Moreover, we are generally interested in the case where $\mathcal{A}$ is precisely Sym, the set of symmetric tensors. Note that Sym is invariant under Orth because

$$(QAQ^{\mathrm{T}})^{\mathrm{T}} = (Q^{\mathrm{T}})^{\mathrm{T}} A^{\mathrm{T}} Q^{\mathrm{T}} = QAQ^{\mathrm{T}}.$$

We use a series of lemmas to come up with a simple form for isotropic functions on Sym. Let $\mathcal{I}(\mathcal{A}) = \{\mathcal{I}_A : A \in \mathcal{A}\}$ denote the set of principle invariants of all tensors in $\mathcal{A}$. Clearly, $\mathcal{I}(\mathcal{A}) \subset \mathbb{R}^3$. Moreover, for simplicity, assume that $\mathcal{A}$ is invariant under Orth for the rest of this section.

Lemma 2.9. *A function $\phi : \mathcal{A} \to \mathbb{R}$ is isotropic if and only if there exists a function $\tilde{\phi} : \mathcal{I}(\mathcal{A}) \to \mathbb{R}$ such that*

$$\phi(A) = \tilde{\phi}(\mathcal{I}_A)$$

for all $A \in \mathcal{A}$.

Proof. For the backward implication, if $\phi(A) = \tilde{\phi}(\mathcal{I}_A)$, then ϕ is isotropic because the eigenvalues (and hence, principle invariants) of a tensor A are invariant under orthogonal transformations. Now suppose ϕ is isotropic. Then we need to show that $\phi(A) = \phi(B)$ whenever $\mathcal{I}_A = \mathcal{I}_B$. However, if two tensors have the same spectrum, then one can use the spectral theorem, Theorem 2.5, to show that there exists an orthogonal tensor Q such that $QAQ^{\mathrm{T}} = B$. Simply set Q to be the unique tensor mapping an orthonormal basis of eigenvectors of A injectively to an orthonormal basis of eigenvectors of B. The orthonormality guarantees that $Q \in \mathrm{Orth}$, and ϕ being isotropic guarantees that $\phi(A) = \phi(QAQ^{\mathrm{T}}) = \phi(B)$, as desired.

□

The next lemma shows that the eigenvectors of A are preserved by isotropic tensor functions.

Lemma 2.10. *Suppose $G : \mathcal{A} \to \mathrm{Lin}$ is isotropic. Then the eigenvectors of A are the same as the eigenvectors of $G(A)$.*

Proof. Use the spectral theorem to write $A = \sum_{i=1}^{3} \lambda_i e_i \otimes e_i$ for some orthonormal set of eigenvectors $\{e_i\}$. Let Q be the unique transformation so that $e_1 \mapsto -e_1$, $e_2 \mapsto e_2$, and $e_3 \mapsto e_3$. It follows that Q is orthonormal and preserves the eigenspaces of A. This implies Q commutes with A. (Reason: If v_i is an eigenvector of A with corresponding eigenvector λ_i, then $QAv_i = Q(\lambda_i v_i) = \lambda_i Qv_i$, and since $Qv_i \in E_{\lambda_i}$ by assumption, $\lambda_i Qv_i = AQv_i$, as desired). Therefore, $QAQ^{\mathrm{T}} = A$. By isotropy of G, we see that $G(A) = G(QAQ^{\mathrm{T}}) = QG(A)Q^{\mathrm{T}}$.

Next,

$$G(A)e_1 = QG(A)Q^{\mathrm{T}}e_1 = QG(A)(-e_1) = -QG(A)e_1,$$

which implies that $G(A)e_1$ is an eigenvector of Q sent to its negative. By construction of Q, this occurs only if $G(A)e_1$ lies in the span of e_1. That is,

$$G(A)e_1 = re_1$$

for some scalar $r \in \mathbb{R}$. Indeed, e_1 is an eigenvector of $G(A)$. $\square$

The following lemma due to Wang gives a criterion for the independence of powers of a positive definite symmetric matrix.

Lemma 2.11. *Let $A \in \mathrm{Sym}$, and suppose $A = \sum_{i=1}^{3} \lambda_i e_i \otimes e_i$ is a spectral decomposition.*

1. *If all the λ_i are distinct, then $\{I, A, A^2\}$ is linearly independent and*

$$\mathrm{Span}\{I, A, A^2\} = \mathrm{Span}\{e_1 \otimes e_1, e_2 \otimes e_2, e_3 \otimes e_3\}.$$

2. *If A has two distinct eigenvalues, then one can write $A = \lambda_1 e \otimes e + \lambda_2 (I - e \otimes e)$, $\{I, A\}$ is a linearly independent set, and*

$$\mathrm{Span}\{I, A\} = \mathrm{Span}\{e \otimes e, I - e \otimes e\}.$$

3. *If A has only one distinct eigenvalue λ, then $A = \lambda I$.*

Combining Wang's lemma with the Cayley-Hamilton theorem, which states that any tensor satisfies its own characteristic polynomial, we have a general representation theorem for isotropic tensor functions. Later we shall see how this provides with an explicit representation of the finite elasticity tensor.

Theorem 2.12. *Suppose $\mathcal{A}$ is a subset of invertible, symmetric tensors. Then $G : \mathcal{A} \to \mathrm{Sym}$ is isotropic if and only if there exist scalar functions $\beta_0, \beta_1, \beta_2 : \mathcal{I}(\mathcal{A}) \to \mathbb{R}$ such that*

$$G(A) = \beta_0(\mathcal{I}_A)I + \beta_1(\mathcal{I}_A)A + \beta_2(\mathcal{I}_A)A^{-1}$$

for every $A \in \mathcal{A}$.

In the special case where G is linear, we can construct an even simpler form for G. Later we shall see how this provides us with a simple representation of the linearized elasticity tensor.

Theorem 2.13. *A function $G : \mathrm{Sym} \to \mathrm{Lin}$ is isotropic if and only if there exist scalars μ and λ such that*

$$G(A) = \lambda(tr(A))I + 2\mu A$$

for every $A \in \mathrm{Sym}$.

A final fact we will need when dealing with the elasticity tensor is the fact that the derivative of an isotropic function is isotropic.

Theorem 2.14. *Suppose G is isotropic and that $\mathcal{A}$ is an open subset of a subspace $U \subset \mathrm{Lin}$. If $Q \in \mathrm{Orth}^+$,*

$$\mathbf{Q}D\mathbf{G}(\mathbf{A})[\mathbf{U}]\mathbf{Q}^{\mathrm{T}} = D\mathbf{G}(\mathbf{Q}\mathbf{A}\mathbf{Q}^{\mathrm{T}})[\mathbf{Q}\mathbf{U}\mathbf{Q}^{\mathrm{T}}].$$

2.3 Multivariable Calculus: Derivatives, Gradient, Curl, Divergence

In this section, we lay out the basics of differentiation and several useful identities from multivariable calculus. Again, we follow [Gur81a] closely, this time Chapter 2.

Definition 4 (Little-oh notation)**.** *Suppose $f : V \to W$ is a function between normed vector spaces V and W. Then we write $f(v) = o(v)$ as $v \to 0$ if*

$$\lim_{v \to 0} \frac{||f(v)||}{||v||} = 0.$$

Intuitively, the above definition states that $f(v)$ goes to 0 faster than v. Moreover, we write $f(v) = g(v) + o(v)$ to signify that $(f - g)(v) = o(v)$. This notion allows us to define the derivative of a function, which can be regarded as the best linear approximation to a function at a given point.

Definition 5. *Let $g : U \to W$ be a function on an open set $U \subset V$. Then g is differentiable at a point $x \in V$ if there exists a linear function $Dg(x)$ called the* derivative of g at x *satisfying*

$$g(x + u) - [g(x) + Dg(x)[u]] = o(u)$$

as $u \to 0$.

If g is differentiable at each x in U, then we say g is *differentiable*. In this case, we regard Dg as a function on U to linear transformations between U and W. Hence, it makes sense to talk about the continuity and differentiability of Dg. We say that g is C^1 on U if g is differentiable and Dg is continuous, and say that g is C^k on U is g is C^{k-1} and Dg is C^{k-2}.

Lemma 2.15. *The derivative $Dg(x)$ at x of a function g, if it exists, is unique.*

Proof. If g is differentiable at x, then

$$g(x + u) - g(x) - Dg(x)[u] = o(u).$$

Fixing $v \in U$ to be a unit vector, we have

$$g(x + \alpha v) - g(x) - Dg(x)[\alpha v] = o(\alpha v).$$

This implies, by the linearity of Dg that

$$\lim_{\alpha \to 0} \frac{||g(x + \alpha v) - g(x) - \alpha Dg(x)[v]||}{\alpha ||v||} = \lim_{\alpha \to 0} ||\frac{1}{\alpha}[g(x + \alpha v) - g(x)] - Dg(x)[v]|| = 0.$$

We conclude that $\lim_{\alpha \to 0} \frac{1}{\alpha}[g(x + \alpha v) - g(x)] = Dg(x)[v]$. Linearity implies that the previous equation is true for all $v \in V$ (not just unit vectors).

For a function $g : \mathbb{R} \to \mathbb{R}$, the above computation yields the familiar difference quotient:

$$g'(t) = \lim_{h \to 0} \frac{g(t+h) - g(t)}{h}.$$

□

As an example, we compute the derivative of the determinant, a fact that will be useful later.

Proposition 2.16. *Define φ on the set of all invertible tensors T by $\varphi(T) = \det T$. Then φ is smooth and*

$$D\varphi(T)[U] = (\det T) tr(UT^{-1})$$

for all $U \in \mathrm{Lin}$.

Proof. Consider $\det(S - \lambda I)$, the characteristic polynomial $p(\lambda)$ of S. Then by properties of the determinant, the coefficient p_k of λ^k in $p(\lambda)$ is a sum of terms of the form $\prod_{j=1}^{k} S_{a_j b_j}$ for some sequences a_j, b_j. By the rearrangement inequality

$$p_k \leq C_k \sum_{i,j} |S_{ij}|^k.$$

Thus, if $k \geq 2$ and S is contained in a ball of radius $\delta < 1$, $p_k \leq C_k \sum_{i,j} S_{ij}^2$. This implies that if $k \geq 2$, then $p_k = o(S)$ as $S \to 0$ because

$$\frac{p_k}{||S||} \leq \frac{C_k \sum_{i,j} S_{ij}^2}{\sqrt{\sum_{i,j} S_{ij}^2}} = C_k ||S|| \to 0.$$

It follows that we may write

$$p(-1) = \det(S + I) = 1 + tr(S) + o(S).$$

Next, suppose A is invertible and $U \in \mathrm{Lin}$. Then

$$\begin{aligned}
\det(A + U) &= \det[(I + UA^{-1})A] \\
&= (\det A) \det(I + UA^{-1}) \\
&= (\det A)[1 + tr(UA^{-1} + o(U)]
\end{aligned}$$

as $U \to 0$. Observe that the map $U \to (\det(A)) tr(UA^{-1})$ is a composition of linear functions, and hence also linear. By continuity of the inverse, trace, multiplication, and determinant operations, $D\varphi$ is continuous.

□

Now we define some operators related to differentiation: the gradient, divergence, and curl.

Definition 6. *Suppose $U \subset V$ and $\varphi : U \to \mathbb{R}$ is smooth. Observe that $D\varphi(x) : V \to \mathbb{R}$ is a linear function, by definition. Using the representation theorem for linear forms, for each $x \in V$, there exists a unique vector $\nabla\varphi(x)$, known as the* gradient *of φ at x such that $D\varphi(x)[u] = \nabla\varphi(x) \cdot u$.*

When $v : V \to V$ is a vector field on V, then $Dv(x)$ is a tensor. For such functions, we replace $Dv(x)[u]$ with $\nabla v(x)\, u$, and refer to the tensor $\nabla v(x)$ as the gradient of v.

Definition 7. *If v is a smooth vector field on V, then the* divergence $\operatorname{div} v$ *of v is the scalar field defined by*

$$\operatorname{div} v = \operatorname{tr}\nabla v.$$

If S is a smooth tensor field on V, we define the divergence $\operatorname{div} S$ *of S to be the unique vector field satisfying*

$$(\operatorname{div} S) \cdot u = \operatorname{div}(S^{\mathrm{T}} u)$$

for all $u \in V$.

Definition 8. *The* curl *of v, denoted by* $\operatorname{curl} v$*, is axial vector of the skew part of ∇v. Concretely,* $\operatorname{curl} v$ *is the unique vector field satisfying*

$$(\nabla v - \nabla v^{\mathrm{T}})u = (\operatorname{curl} v) \times u.$$

Finally, less important for our purposes, given a scalar or vector field $\Phi \in C^2$, we define the *laplacian* $\Delta\Phi$ to be the vector field

$$\Delta\Phi = \operatorname{div}\nabla\Phi$$

The above definitions can be made concrete by the following useful list of coordinate representations of the specified operations. If $v \in \mathbb{R}^n$ is a vector, let v_i denote that i-th component of v in the standard basis. As before, φ is scalar valued, v is vector valued, and S is tensor valued.

$$(\nabla\varphi)_i = \frac{\partial v_i}{\partial x_i}, \qquad (\nabla v)_{ij} = \frac{\partial v_i}{\partial x_j}, \qquad \operatorname{div} v = \sum_i \frac{\partial v_i}{\partial x_i}$$

$$(\operatorname{div} S)_i = \sum_j \frac{\partial S_{ij}}{\partial x_j}, \qquad \Delta\varphi = \sum_i \frac{\partial^2\varphi}{\partial x_i^2}, \qquad (\Delta v)_j = \Delta(v_j)$$

$$\operatorname{curl} v = (\frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3}, \frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1}, \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2}).$$

Another important fact that we will need is the generalized product rule for bilinear operations.

Proposition 2.17. *Let f and g be differentiable functions at $x_0 \in U \subset V$, and π a bilinear form on $V \times V$. Then the function $h = fg$ is differentiable at x_0 and its derivative is given by*

$$Dh(x_0)[u] = \pi(f(x_0), Dg(x_0)[u]) + \pi(Df(x_0)[u], g(x_0)).$$

Equivalently, define functions $h_1(x) = f_0(x)g(x)$ and $h_2(x) = f(x)g_0(x)$. Then

$$Dh(x_0) = Dh_1(x_0) + Dh_2(x_0).$$

Using this proposition and some manipulation, we can derive a host of useful identities.

Proposition 2.18. *Let $\phi, v, w,$ and S be differentiable fields with ϕ scalar valued, v and w vector valued, and S tensor valued. Then*

$$\begin{aligned}
\nabla(\phi v) &= \phi \nabla v + v \otimes \nabla \phi \\
\operatorname{div}(\phi v) &= \phi \operatorname{div} v + v \cdot \nabla \phi \\
\nabla(v \cdot w) &= (\nabla w)^{\mathrm{T}} v + (\nabla v)^{\mathrm{T}} w \\
\operatorname{div}(v \otimes w) &= v \operatorname{div} w + (\nabla v) w \\
\operatorname{div}(S^{\mathrm{T}} v) &= S \cdot \nabla v + v \cdot \operatorname{div} S \\
\operatorname{div}(\phi S) &= \phi \operatorname{div} S + S \nabla \phi.
\end{aligned}$$

We close with the statement of the *chain rule*.

Proposition 2.19. *Let $f : V \to W$ and $g : U \to V$. Suppose g is differentiable at x and that f is differentiable at $g(x)$. Then the composition*

$$h = f \circ g$$

is differentiable at x and

$$\nabla[f \circ g](x) = \nabla f(g(x)) \nabla g(x).$$

2.4 More Multivariable Calculus

A *region* $\mathcal{R}$ is a subset of $\mathbb{R}^n$ whose boundary can be parametrized as with piecewise smooth maps. A region $\mathcal{R}$ is said to be *connected* if for all points $x, y \in \mathcal{R}$, there exists a smooth curve $\gamma : [0,1] \to \mathcal{R}$ such that $\gamma(0) = x$ and $\gamma(1) = y$. Naturally, a region is said to be *closed* (respectively, open) if $\mathcal{R}$ is an closed (respectively, open) subset of $\mathbb{R}^n$. A region is said to be *bounded* if it is contained in a ball of finite radius. The *boundary* $\partial\mathcal{R}$ of a region is given by the set of points that are in the closure of $\mathcal{R}$ and $\mathbb{R}^n \backslash \mathcal{R}$. The interior $\mathring{\mathcal{R}}$ of a region consists of all $p \in \mathcal{R}$ such that there exists an open set U_p satisfying $p \in U_p \subset \mathcal{R}$. A *subregion* U or $\mathcal{R}$ is a subset of $\mathcal{R}$ that is also a region.

To illustrate connectedness and the chain rule, we prove that a function with constant gradient on a connected set is affine-linear.

Proposition 2.20. *Suppose $f : U \to V$ is a vector field on $U \subset V$, and that $\nabla f(u) = \nabla f$ is constant for all $u \in U$. Then*

$$f(x) = f(y) + \nabla f(x - y).$$

Proof. Given $x, y \in \mathcal{R}$, connect them with a path γ. By the fundamental theorem of calculus and the chain rule

$$f(x) - f(y) = \int_0^1 \frac{d}{ds} f(\gamma(s)) ds = \int_0^1 \nabla(f(\gamma(s))\gamma'(s) ds = \nabla f \int_0^1 \gamma'(s) dx = \nabla f(x - y).$$

The third equality is true by the linearity of integration.

□

The previous proof motivates us to define the *line integral* $\int_\gamma v \cdot dx$ of a vector field v along a curve $\gamma : [0, 1] \to V$.

$$\int_\gamma v \cdot dx := \int_0^1 \gamma'(s) \cdot v(\gamma(s)) ds.$$

The next statement, the localization theorem, provides a method of extracting local information of regular solutions to PDEs. It will also provide a geometric interpretation of the divergence and curl theorems. The proof is quite simple.

Theorem 2.21. *Let Φ be a continuous scalar or vector field on $U \subset V$. Let $B_\delta(x)$ denote the ball of radius δ about x. Then*

$$\Phi(x_0) = \lim_{\delta \to 0} \frac{1}{\mathrm{Vol}(B_\delta(x_0))} \Phi dV.$$

Proof. Let $v(\delta) = \mathrm{Vol}(B_\delta(x_0))$ and $B_\delta = B_\delta(x_0)$.

$$\begin{aligned} |\Phi(x_0) - \frac{1}{v(\delta)} \int_{B_\delta} \Phi dV| &= |\frac{1}{v(\delta)} \left(\int_{B_\delta} \Phi(x_0) - \Phi dV \right) | \\ &\leq \frac{1}{v(\delta)} \left(\int_{B_\delta} |\Phi(x_0) - \Phi dV| \right) \leq \sup_{x \in B_\delta} |\Phi(x_0) - \Phi(x)| \to 0 \end{aligned}$$

as $\delta \to 0$, by continuity.

□

Stokes's theorem is one of the hallmarks of calculus on manifolds. For our purposes, we need two statements that can be derived from this general principle: the divergence theorem and the curl theorem.

Theorem 2.22 (Divergence)**.** *Suppose $\mathcal{R}$ is a bounded regular region, let φ be a smooth scalar field, v a smooth vector field, and S a smooth tensor field on $\mathcal{R}$. Let $n : \mathcal{R} \to V$ denote the unit outward-pointing normal vector field. We have*

$$\int_{\partial \mathcal{R}} \varphi n dS = \int_{\mathcal{R}} \nabla \varphi dV.$$
$$\int_{\partial \mathcal{R}} v \cdot n dA = \int_{\mathcal{R}} \mathrm{div}\, v dV$$
$$\int_{\partial \mathcal{R}} S n dA = \int_{\mathcal{R}} \mathrm{div}\, S dV.$$

This result with Proposition 2.18 has the following useful corollary that can be thought of as a higher-dimensional version of integration by parts.

Corollary 2.23. *Let $\mathcal{R}$ be a bounded regular region and $n : \mathcal{R} \to V$ be its unit outward-pointing normal field as before. Suppose that u and v are smooth scalar fields. Then*

$$\int_{\mathcal{R}} \nabla u \cdot \nabla v \, dV = \int_{\partial \mathcal{R}} u \, \nabla v \cdot n \, dA - \int_{\mathcal{R}} u \, \mathrm{div} v \, dV$$

The next result is also referred to as Stokes's theorem, even though it is a very specific case of the more general result on manifolds of the same name.

Theorem 2.24 (Curl)**.** *Suppose v is a smooth vector field on a region $\mathcal{R} \subset \mathbb{R}^3$. Let D_δ denote a two-dimensional disc in $\mathcal{R}$, and denote by $\gamma : [0, 1] \to \partial D_\delta$ a counterclockwise parametrization of the circle on the boundary of D_δ. Again, n denotes the unit outward-pointing normal on D_δ. Then*

$$\int_{D_\delta} (\operatorname{curl} v) \cdot n \, dA = \int_\gamma v \cdot dx.$$

Using the localization theorem, the divergence at x of a vector field v can be interpreted as the net flux through an infinitesimal ball about x (see Equation 2.3). Likewise, one can see from Equation 2.4 that the curl at x of v is the vector field pointing in the direction where the circulation per unit area about is maximized.

$$\operatorname{div} v(x_0) = \lim_{\delta \to 0} \frac{1}{v(\delta)} \int_{\partial B_\delta(x_0)} v \cdot n dA. \tag{2.3}$$

$$n \cdot \operatorname{curl} v(x_0) = \lim_{\delta \to 0} \frac{\int_\gamma v \cdot dx}{\operatorname{Area}(D_\delta)} \tag{2.4}$$

We close with some basic but extremely useful results on change of variables formulas for integrals. The first statement concerns volume integrals.

Theorem 2.25. *Let $f : \mathcal{R} \to \mathbb{R}^n$ be a smooth invertible transformation. Let φ be a continuous scalar field on $f(\mathcal{R})$. Then*

$$\int_{f(\mathcal{R})} \varphi(x) \, dV_x = \int_{\mathcal{R}} \varphi(f(p)) \det \nabla f(p) \, dV_p.$$

For linear transformations $T : \mathbb{R}^3 \to \mathbb{R}^3$, this theorem has the interesting version below. Let $u, v, w \in \mathbb{R}^3$ be linearly independent. Then

$$\det T = \frac{Tu \cdot (Tv \times Tw)}{u \cdot (v \times w)}$$

since $u \cdot (v \times w)$ represents the volume of the parallelpiped formed by u, v, and w. By localizing, we see that Theorem 2.25 implies

$$\det \nabla f(p) = \lim_{\delta \to 0} \frac{\operatorname{Vol}(f(B_\delta(p)))}{\operatorname{Vol}(B_\delta(p))}.$$

That is, the determinant of f at p represents the relative change in volume locally at p induced by f. A function satisfying $\operatorname{Vol}(f(U)) = \operatorname{Vol}(U)$ for all subregions $U \subset \mathcal{R}$ is said to be *isochoric*. Such functions are pivotal when studying conservation of mass. The above localization implies the following corollary.

Corollary 2.26. *A deformation is isochoric if and only if*

$$\det \nabla f \equiv 1.$$

A lesser-known theorem, but necessary for our purposes when we want to pull back the constitutive relations to the reference configuration in Section 3, is a version of change of variable for surface integrals. It involves a transformation $F \to (\det F)F^{-\mathrm{T}}$ known as the *Piola transform.*

Theorem 2.27. *Let $f : \mathcal{R} \to \mathbb{R}^n$ be a smooth invertible tranformation. Suppose φ is a continuous scalar field, v is a continuous vector field, and T is a continuous tensor field on $\mathcal{R}$. Let m and n be the outward-pointing unit normal vector fields on $\partial f(\mathcal{R})$ and $\partial\mathcal{R}$, respectively.*

$$\begin{aligned}
\int_{\partial f(\mathcal{R})} \varphi(x)m(x)dA_x &= \int_{\partial\mathcal{R}} \varphi(f(p))(\det \nabla f(p))(\nabla f(p))^{-\mathrm{T}}n(p)dA_p \\
\int_{\partial f(\mathcal{R})} v(x)\cdot m(x)dA_x &= \int_{\partial\mathcal{R}} v(f(p))\cdot \left[(\det \nabla f(p))(\nabla f(p))^{-\mathrm{T}}n(p)\right] \\
\int_{\partial f(\mathcal{R})} T(x)m(x)dA_x &= \int_{\partial\mathcal{R}} T(f(p))(\det \nabla f(p))(\nabla f(p))^{-\mathrm{T}}n(p)dA_p \\
\int_{\partial f(\mathcal{R})} (x-\vec{0})\times T(x)m(x)dA_x &= \int_{\partial\mathcal{R}} (f(p)-\vec{0})\times T(f(p))(\det \nabla f(p))(\nabla f(p))^{-\mathrm{T}}n(p)dA_p.
\end{aligned}$$

3 Elasticity

This section is built from the treatment of solid mechanics in [Gur81a].

3.1 Deformations, Strain, and Motions

From this point on, we will restrict our focus to connected regular regions in $\mathbb{R}^3$, referred to as *bodies.*

Given a body $\mathcal{B}$, a function $f : \mathcal{B} \to \mathcal{R}^3$ is said to be a *deformation* if it is smooth, injective, and satisfies $\det \nabla f > 0$. The latter requirement ensures that f is orientation-preserving and that the inverse of f is also smooth. In general, we write $F = \nabla f$. Moreover, the domain of f is also referred to as the *reference configuration.*

It will also be useful later for the purposes of linearization to consider the *displacement* given by

$$u(p) = f(p) - p.$$

If F is constant on $\mathcal{B}$, then we showed that $f(p) = f(q) + F(p-q)$. In this case, we say that f is *homogeneous.* This can be used to show the following proposition.

Proposition 3.1. *Let f be a homogeneous deformation. Then given any point $q \in \mathbb{R}^3$, we can uniquely decompose f as follows:*

$$f = d_1 \circ g = g \circ d_2$$

where q is a fixed point of g and d_1, d_2 are translations.

Proof. We show only the first decomposition. Define

$$g(p) = q + F(p - q).$$

Then set $d_1 = g^{-1} \circ f$. Note that g is certainly invertible because $\nabla g = \nabla f = F \neq 0$. Moreover, $\nabla(g^{-1} \circ f) = (\nabla g)^{-1}(\nabla f) = F^{-1}F = I$. Therefore,

$$g^{-1} \circ f(p) = g^{-1} \circ f(q) + I(p - q),$$

so d_1 is a translation. To see uniqueness, if $d_1 \circ g_1 = d_1' \circ g_1'$, then $(d_1^{-1} d_1')g_1' = g_1$, which implies $(d_1^{-1} d_1')g_1'(q) = (d_1^{-1} d_1')(q) = q$. But $d_1^{-1} d_1'$ is a translation, and hence has a fixed point only if $d_1 = d_1'$. By left-cancellation, $g_1' = g_1$ as well.

□

By the Polar Decomposition Theorem, F can be uniquely written as $F = RU = VR$ where R is a rotation and U, V are positive definite and symmetric. If U has the form

$$U = I + (\lambda - 1)e \otimes e$$

with $\lambda \in \mathbb{R}$ and $|e| = 1$, then U is referred to as a *stretch*. In fact, any positive definite symmetric matrix U can be written as a series of stretches $U = U_1 U_2 U_3$, where U_i is a stretch. To see this, let $U = \sum_i \lambda_i e_i \otimes e_i$ be the spectral decomposition of U. Then set $U_i := I + (\lambda_i - 1)e_i \otimes e_i$. We see

$$\begin{aligned} U_1 U_2 U_3 &= (I + (\lambda_1 - 1)e_1 \otimes e_1)(I + (\lambda_2 - 1)e_2 \otimes e_2)(I + (\lambda_3 - 1)e_3 \otimes e_3) \\ &= (I + (\lambda_1 - 1)e_1 \otimes e_1 + (\lambda_2 - 1)e_2 \otimes e_2)(I + (\lambda_3 - 1)e_3 \otimes e_3) \\ &= (I + \sum_i (\lambda_i - 1)e_i \otimes e_i) = \sum_i \lambda_i e_i \otimes e_i = U. \end{aligned}$$

since $(e_i \otimes e_i)(e_j \otimes e_j) = 0$ when $i \neq j$.

In light of this structure, the eigenvalues λ_i of U are known as the *principal stretches*, since they describe how much the eigenspaces are scaled or extended by U.

Now consider a deformation f that is not necessarily homogeneous. Then since

$$f(p) = f(q) + F(q)(p - q) + o(p - q),$$

so that f behaves like a homogeneous deformation in a neighborhood about p. It makes sense to decompose $F = RU = VR$. Letting p vary, we get tensors R, U, and V called the *rotation tensor*, *left stretch tensor*, and *right stretch tensor*, respectively. Since U and V are often times difficult to compute, it is useful to define the *left* and *right* Cauchy-Green strain tensors, B and C as follows:

$$C = U^2 = F^{\mathrm{T}}F \quad B = V^2 = FF^{\mathrm{T}}.$$

These tensors are very important for the small deformations considered in linearized elasticity.

This section closes with a characterization of *rigid deformations*, distance-preserving deformations. That is, $f : \mathcal{B} \to \mathbb{R}$ is a rigid deformation if

$$|f(p) - f(q)| = |p - q|$$

for all $p, q \in \mathcal{B}$.

Proposition 3.2. *The following are equivalent.*

1. *f is a rigid deformation*
2. *f is homogeneous and has a constant gradient that is a rotation*
3. *$F(p)$ is a rotation for all $p \in \mathcal{B}$*
4. *Any of U, V, B or C is the identity*
5. *f preserves the lengths of curves.*

It is also necessary to introduce the concept of *infinitesimal rigid displacement* u, which is a deformation that arises from truncating the higher-order terms in the displacement of a rigid deformation f. These concepts are central to linearized elasticity.

Given a deformation f, we write $f(p) = p+u(p)$, where u is the displacement and see that

$$F = I + \nabla u.$$

Therefore, the Cauchy-Green strain tensors can be written

$$B = I + \nabla u + \nabla u^{\mathrm{T}} + \nabla u \nabla u^{\mathrm{T}}$$
$$C = I + \nabla u + \nabla u^{\mathrm{T}} + \nabla u^{\mathrm{T}} \nabla u$$

Observe that if f is rigid, $B = C = I$, which implies, using the above equations, that as $u \to 0$,

$$\nabla u = -\nabla u^{\mathrm{T}} + o(u).$$

Intuitively speaking, the gradient of the displacement of f is *infinitesimally* skew. Therefore, it is natural to define a vector field u on $\mathcal{B}$ to be an *infinitesimal rigid displacement* if $\nabla u(p)$ is constant skew for all $p \in \mathcal{B}$. Interestingly enough, if ∇u is skew everywhere, that implies that it is constant, as seen in the following proposition.

Proposition 3.3. *Let u be a smooth vector field on $\mathcal{B}$. Then the following are equivalent*

1. *u is an infinitesimal rigid displacement.*
2. *u satisfies*
$$(p - q) \cdot [u(p) - u(q)] = 0$$
for all $p, q \in \mathcal{B}$.
3. *$\nabla u(p)$ is skew for all $p \in \mathcal{B}$.*

Proof. We show that 1. $\Leftrightarrow$ 2., and that 3. $\Leftrightarrow$ 1.

First, 1. $\Rightarrow$ 2. because

$$(p-q)\cdot[u(p)-u(q)] = (p-q)\cdot W(p-q) = W^{\mathrm{T}}(p-q)\cdot(p-q)$$

by definition of the transpose. However, $W^{\mathrm{T}}(p-q)\cdot(p-q) = -W(p-q)\cdot(p-q)$.

Thus, $W(p-q)\cdot(p-q) = 0$, as desired. For the opposite direction, let $g(p) = (p-q)\cdot[u(p)-u(q)]$. Then by the product rule and assumption 2.,

$$\begin{aligned}\nabla g(p)[v] &= Iv\cdot[u(p)-u(q)] + (p-q)\cdot\nabla u(p)[v]\\ &= v\cdot[u(p)-u(q)] + \nabla u^{\mathrm{T}}(p-q)\cdot v\\ &= v\cdot[u(p)-u(q)+\nabla u^{\mathrm{T}}(p-q)] = 0\end{aligned}$$

for all $v\in\mathbb{R}^3$. Therefore,

$$h(q) := u(p)-u(q)+\nabla u^{\mathrm{T}}(p)(p-q) = 0.$$

Taking gradient of $h(q)$ implies

$$\nabla h(q)[v] = -\nabla u(q)[v] + \nabla u^{\mathrm{T}}(p)(-Iv) = (-\nabla u(q) - \nabla u^{\mathrm{T}}(p))v = 0.$$

Since this holds for all p, q, v, we have

$$-\nabla u(q) = \nabla u^{\mathrm{T}}(q) = -\nabla u(p).$$

Indeed, ∇u is skew and constant $\Rightarrow$ u is an infinitesimal rigid displacement. It is clear by assumption that 1. $\Rightarrow$ 3. For the opposite direction, let $p\in\mathcal{B}$ and B_δ a small ball about p. We show that ∇u is constant in B_δ. By connectedness, this will imply that ∇u is constant on $\mathcal{B}$ and complete the proof.

Given $p, q\in B_\delta$, construct the line segment $\gamma(t) := q + t(p-q)$ from p to q. Then

$$\begin{aligned}u(p)-u(q) &= \int_0^1 \frac{d}{dt}(u\circ\gamma(s))ds\\ &= \int_0^1 \nabla u(\gamma(s))\gamma'(s)ds\\ &= \int_0^1 \nabla u(\gamma(s))(p-q)ds.\end{aligned}$$

Therefore,

$$(p-q)[u(p)-u(q)] = \int_0^1 (p-q)\cdot\nabla u(\gamma(s))(p-q)ds = 0$$

since ∇u is skew everywhere. Therefore, u has property 2. on B_δ, which implies u is an infinitesimal rigid displacement on B_δ. The proposition follows. $\square$

The previous developments are essential for establishing the concept of *motions*, defined below. Later we will generalize the concept of rigid deformation to that of a rigid motion, and use our analysis of inifitesimal rigid displacement to classify the velocity gradients of rigid motions.

Definition 9. *Given a body $\mathcal{B}$, a* motion *of $\mathcal{B}$ is a C^3 function*

$$x : \mathcal{B} \times \mathbb{R} \to \mathbb{R}^3$$

such that $x(\cdot, t)$ is a deformation of $\mathcal{B}$ for each fixed t. The point $x = x(p, t)$ is referred to as the place occupied by the material point p at time t, and $\mathcal{B}_t := x(\mathcal{B}, t)$ is used to denote the region of space occupied by the body at time t.

Another natural object is the trajectory $\mathcal{T}$ that describes all the points and times passed through by the motion:

$$\mathcal{T} = \{(x, t) | x \in \mathcal{B}_t, t \in \mathbb{R}\}.$$

The *reference map* $p : \mathcal{T} \to \mathcal{B}$ is the unique map satisfying $p(\cdot, t) : \mathcal{B}_t \to \mathcal{B}$ is the inverse to $x(\cdot, t)$. Using the inverse function theorem (omitted in these notes), one can easily show that $p(x, t)$ is C^3.

Now we define several quantities intrinsic to the concept of motion. The *velocity* is

$$\dot{x}(p, t) = \frac{d}{dt} x(p, t),$$

and the acceleration is

$$\ddot{x}(p, t) = \frac{d^2}{dt^2} x(p, t).$$

The *spatial description of velocity* is

$$v(x, t) = \dot{x}(p(x, t), t).$$

In general, the reference map provides a road map between functions defined on $\mathcal{B} \times \mathbb{R}$, known as *material fields*, and functions defined on $\mathcal{T}$, known as *spatial fields.* The *spatial description* Φ_s of a material field $(p, t) \to \Phi(p, t)$ is

$$\Phi_s(x, t) = \Phi(p(x, t), t).$$

Likewise, the *material description* Ψ_m of a spatial field $(x, t) \to \Psi(x, t)$ is

$$\Psi_m(p, t) = \Psi(x(p, t), t).$$

For a material field Φ,

$$\dot{\Phi}(p, t) = \frac{d}{dt} \Phi(p, t)$$

and the material gradient is

$$\nabla \Phi(p, t) = \nabla_p \Phi(p, t),$$

the gradient only with respect to the p variable. We write $F = \nabla x$ for the *deformation gradient*, the material gradient of the motion. Similarly,

$$\Psi'(x,t) = \frac{d}{dt}\Psi(x,t)$$
$$\text{grad}\Psi(x,t) = \nabla_x \Psi(x,t)$$

are the *spatial time derivative* and *spatial gradient* of Ψ. In the same manner, one can define the spatial gradient and curl, written simply divΨ and curlΨ, respectively. The material divergence and curl denoted by Div and Curl, respectively.

Next is the *material time derivative* $\dot{\Psi}$ of a *spatial field*:

$$\dot{\Psi}(x,t) = ((\dot{\Psi})_m)_s = \frac{d}{dt}\Psi(x(p,t),t)\big|_{p=p(x,t)}.$$

The next proposition lets us compute this sort of derivative explicitly.

Proposition 3.4. *Let φ and u be smooth spatial fields φ scalar valued and u vector valued. Then*

$$\dot{\varphi} = \varphi' + v \cdot \text{grad}\varphi$$
$$\dot{u} = u' + (\text{grad}u)v.$$

The next two propositions rely on the chain rule, and are useful for later computations. Let $L = \text{grad}\, v$ denote the *velocity gradient.*

Proposition 3.5. *Let u be a smooth spatial vector field. Then*

$$\nabla(u_m) = (\text{grad}u)_m \nabla x.$$

Proposition 3.6.

$$\dot{F} = L_m F$$
$$\ddot{F} = (\text{grad}\, \dot{v})_m F.$$

A motion x on $\mathcal{B}$ is said to be *rigid* if

$$\frac{d}{dt}|x(p,t) - x(q,t)| = 0$$

for all $p, q \in \mathcal{B}$. The following proposition characterizes rigid motions by their velocity gradients.

Proposition 3.7. *Let x be a motion with velocity vector v. Then x is a rigid motion if and only if $v(\cdot,t)$ is an infinitesimal rigid displacement for each time t.*

Proof. Fix distinct points $p, q \in \mathcal{B}$. Let $\delta(t) = |x(p,t) - x(q,t)|$. Then

$$\begin{aligned}\frac{d}{dt}(\delta(t)^2) &= 2\dot{\delta}(t)d(t) = 2[x(p,t) - x(q,t)] \cdot [\dot{x}(p,t) - \dot{x}(q,t)] \\ &= 2[x(p,t) - x(q,t)] \cdot [v(x(p,t),t) - v(x(q,t),t)] = 0\end{aligned}$$

The last equality implies that for all $x, y \in \mathcal{T}$,

$$0 = (x - y) \cdot (v(x,t) - v(y,t)),$$

so by 2. of Proposition 3.3, $v(\cdot, t)$ is an infinitesimal rigid displacement, as desired. To get the opposite direction, follow this argument backwards. $\square$

By the previous result, the velocity gradient $L(x,t)$ of a rigid motion x can be written $L(x,t) = W(t)$, where W is a smoothly varying skew tensor.

For a general motion x, write $L(x,t) = W(x,t) + D(x,t)$ where D and W are the skew and symmetric parts, respectively, of L. Since D is a symmetric matrix, it can be written as a series of stretches, and is referred to as the *rate of stretching.* If ω is the axial vector of W, then $\mathrm{curl} v = 2\omega$, by definition. The vector ω can be thought of as the infinitesimal axis of rotation of the motion. We recover the following proposition involving the material time derivative of v.

Proposition 3.8. *The following properties hold for a velocity field v.*

1. $\dot{v} = v' + \frac{1}{2}\mathrm{grad}(v^2) + 2Wv$.
2. $\dot{v} = v' + \frac{1}{2}\mathrm{grad}(v^2) + (\mathrm{curl}\, v) \times v$

Proof. Equation 2. follows from equation 1. because $\mathrm{curl} v = 2\omega$ and $Wv = \omega \times v$. To see Equation 1., note from Proposition 3.4 that

$$\dot{v} = v' + (\mathrm{grad}\, v)v.$$

Also, applying the product rule,

$$\mathrm{grad}(v \cdot v)(x,t)[u] = 2v \cdot \mathrm{grad}\, v\, u = [2(\mathrm{grad} v^{\mathrm{T}})v] \cdot u$$

so that $\frac{1}{2}\mathrm{grad}\, v^2 = (\mathrm{grad}\, v^{\mathrm{T}})v$. Since $2W = \mathrm{grad}\, v - \mathrm{grad}\, v^{\mathrm{T}}$,

$$(2W + \frac{1}{2}\mathrm{grad}\, v^2)v = (\mathrm{grad} v)v.$$

This proves the proposition. $\square$

We close this section on kinematics with an interesting interpretation of the stretch $D = \frac{1}{2}(L + L^{\mathrm{T}})$.

Proposition 3.9. *Let $x \in \mathring{\mathcal{B}}_{t_0}$ with t_0 fixed, let n be a vector with $|n| = 1$, $p, q \in \mathcal{B}$ satisfying $x(p,t_0) = x(q,t_0) + \alpha e$, and $\delta_\alpha(t) = |x(p,t) - x(q,t)|$.*

$$\frac{\dot{\delta}_\alpha(t_0)}{\delta_\alpha(t_0)} = e \cdot D(x,t_0)e.$$

Intuitively, the relative change in the distance function is given by the stretched norm generated by the symmetric matrix D. The proof is simple.

Proof. By the set-up, $\delta_\alpha(t_0) = \alpha$. Using the computation of $(\delta_\alpha \dot{(t)}^2)$ from the characterization of rigid motions,

$$\delta_\alpha(t_0)\dot{\delta}_\alpha(t_0) = (x + \alpha e - x) \cdot [v(x + \alpha e, t_0) - v(x, t_0)] \Rightarrow$$
$$\frac{\dot{\delta}_\alpha(t_0)}{\delta_\alpha(t_0)} = \frac{e \cdot [v(x + \alpha e, t_0) - v(x, t_0)]}{\alpha}.$$

Take limits and use the computation of the uniqueness of the derivative from Section 2.3 to conclude

$$\lim_{\alpha \to 0} \frac{[v(x + \alpha e, t_0) - v(x, t_0)]}{\alpha} = L(x, t_0)e.$$

Finally, $e \cdot L(x, t_0)e = e \cdot D(x, t_0)e$ because $e \cdot W(x, t_0)e = 0$. These last facts combine to prove the theorem.

□

3.2 Transport Laws and Conservation of Mass

The goal of this section is to take time derivatives of volume integrals, which intuitively, describe the net instantaneous flow through a subregion $U \subset \mathcal{B}$ of a body. Then we develop the notion of *mass* and *density*, a scalar field with the property that its integral over a subregion U_t, known as *mass*, is constant over time. That is, mass is *conserved.*

The key result is Reynold's transport theorem.

Theorem 3.10. *Let Φ be a smooth scalar or vector field. Then for any subregion U and time t*

$$\begin{aligned}\frac{d}{dt}\int_{U_t} \Phi dV &= \int_{U_t} (\dot{\Phi} + \Phi \mathrm{div} v) dV \\ &= \int_{U_t} \Phi' dV + \int_{\partial U_t} \Phi v \cdot n dA\end{aligned}$$

where n denotes the normal vector on U_t.

Proof. The obvious thing to do is pull back the integral to the reference configuration and differentiate underneath the integral sign.

$$\begin{aligned}\frac{d}{dt}\int_{U_t} \Phi dV &= \frac{d}{dt}\int_{U} \Phi_m \det F dV \\ &= \int_U (\Phi_m \det F)^{\cdot} dV.\end{aligned}$$

Now, we want to use the product rule on the above expression. First, note that by the chain rule and our computation of the derivative of the determinant $\varphi(A) := \det A$

$$(\det F)^{\cdot} = \nabla\varphi(F(p, t))[\dot{F}(p, t)] = (\det F)\mathrm{tr}(\dot{F}F^{-1}).$$

Recall from the previous section that $\dot{F} = L_m F$. Therefore,

$$(\det F)^{\cdot} = (\det F)\mathrm{tr}(L_m) = (\det F)(\mathrm{tr}L)_m = (\det F)(\mathrm{div}\, v)_m.$$

Tracing back our steps, we have

$$\begin{aligned} \frac{d}{dt}\int_{U_t} \Phi dV &= \int_U (\dot{\Phi} + \Phi \mathrm{div}\, v)_m \det F dV \\ &= \int_{U_t} (\dot{\Phi} + \Phi \mathrm{div}\, v) dV \end{aligned}$$

where the last line follows from changing variables again. This proves Part 1. For Part 2., observe that

$$\dot{\Phi} + \Phi\mathrm{div}\, v = \Phi' + (\mathrm{grad}\Phi)\cdot v + \Phi\mathrm{div}\, v = \mathrm{div}\, v,$$

where the last equality follows from Proposition 2.18. Now apply the divergence theorem.

$$\begin{aligned} \int_{U_t} (\dot{\Phi} + \Phi\mathrm{div}\, v) dV &= \int_{U_t} \Phi' + \mathrm{div}(\Phi v) dV \\ &= \int_{U_t} \Phi' dV + \int_{\partial U_t} \Phi v \cdot n dA \end{aligned}$$

□

Setting $\Phi \equiv 1$ yields a statement about volume transport.

Corollary 3.11.

$$\frac{d}{dt}\mathrm{Vol}(U_t) = \int_{U_t} \mathrm{div}\, v dV = \int_{\partial U_t} v \cdot n dA$$

A motion is said to be isochoric[2] if

$$\frac{d}{dt}\mathrm{Vol}(U_t) = 0.$$

Corollary 3.12. *The following are equivalent.*

1. *x is isochoric.*
2. $\mathrm{div}\, v = 0$
3. *For every subregion $U \subset \mathcal{B}$*

$$\int_{\partial U_t} v \cdot n dA = 0$$

[2]Question for the reader: How is this notion intrinsically distinct from that of isochoric deformations?

The final condition says, intuitively, that the net flux on any subregion is 0. Hence, we expect volume to be preserved throughout. These considerations naturally lead us to the topic of *conservation of mass*.

For any deformation f there is a density field ρ_f such that the mass $m(U)$ of a subregion U is given by

$$m(U) = \int_{f(U)} \rho_f dV.$$

Also, the RHS is independent of the f chosen. We also require mass to be positive, that is, $\rho_f > 0$. Let ρ_0 denote the *initial density* of $\mathcal{B}$, that is, the density field corresponding to the identity deformation. The the localization theorem and changing variables easily yields the following.

Proposition 3.13. *Let f be a deformation of $\mathcal{B}$. Then*

$$\rho_f(f(p)) \det \nabla f(p) = \rho_0(p).$$

Given a motion x, we write $\rho_{x(\cdot,t)} = \rho(x,t)$ for $x \in \mathcal{B}_t$. Observe that, by change of variables and conservation of mass,

$$\frac{d}{dt}\int_{U_t} \rho dV = 0.$$

The above equation combined with Reynold's transport theorem yields the following *local* conservation of mass laws.

$$\dot{\rho} + \rho \operatorname{div} v = 0$$
$$\rho' + \operatorname{div}(\rho v) = 0$$

Moreover, change of variables guarantees the following.

Lemma 3.14. *let Φ be a continuous spatial field. The given any subregion U,*

$$\int_{U_t} \Phi(x,t)\rho(x,t)dV_x = \int_U \Phi_m(p,t)\rho_0(p)dV_p.$$

This lemma has the fascinating consequence

$$\frac{d}{dt}\int_{U_t} \Phi\rho dV = \int_{U_t} \dot{\Phi}\rho dV$$

that enables us to differentiate under the integral sign, even though the domain is not fixed, with respect to the *mass measure ρdV*.

3.3 Momentum Balance Laws and Existence of a Stress

In this Section, we describe the effects of three different types of forces on a body $\mathcal{B}$:

1. contact forces between two subregions $U_1, U_2 \subset \mathcal{B}$,
2. surface forces caused by interaction between $\mathcal{B}$ and its environment, and

3. forces exerted on interior points of $\mathcal{B}$ by the environment (eg gravity).

Cauchy's fundamental work in continuum mechanics was inspired by his work on Cherbourg Harbor and the Canal de l'Ourcq. He made the following insightful assumption.

Assumption. *There exists a surface force density $s(n, x, t)$ for all $(x, t) \in \mathcal{T}$ with the property that if $S \subset \mathcal{B}$ is a two-dimensional subregion of $\mathcal{B}$ (i.e. a* surface*) with outward-pointing unit normal vector n at x, $s(n, x, t)$ describes the force per unit area exerted on the negative side of S by the positive side.*

As a consequence of this assumption, two surfaces S_1 and S_2 intersecting at some point x and sharing the same normal vector n will experience the same force per unit area at the point of intersection.

This motivates the following definition. Let $\mathcal{N}$ denote the set of unit vectors in $\mathbb{R}^3$.

Definition 10. *A* system of forces *for a body $\mathcal{B}$ during a motion x with trajectory $\mathcal{T}$ is a pair (s, b) of functions*

1. *$s : \mathcal{N} \times \mathcal{T} \to \mathbb{R}^3$ smooth for fixed n, t*
2. *$b : \mathcal{T} \to \mathbb{R}^3$ continuous for fixed t*

known as the surface force *and* body force*, respectively.*

These concepts can be extended to compute the *force* f and *moment* m, respectively, on a subregion U:

$$
\begin{aligned}
f(U, t) &:= \int_{\partial U_t} s(n) dA + \int_{U_t} b dV \\
m(U, t) &:= \int_{\partial U_t} r \times s(n) dA + \int_{U_t} r \times b,
\end{aligned}
$$

where n is the outward-pointing normal to ∂U_t and $r(x) := x - \vec{0}$ is the position vector.

Briefly, we define the *linear momentum* ℓ and *angular momentum* a, respectively, of a subregion $U \subset \mathcal{B}$.

$$
\begin{aligned}
\ell(U, t) &= \int_{U_t} v \rho dV \\
a(U, t) &= \int_{U_t} r \times v \rho dV
\end{aligned}
$$

Here, $r(x) := x - \vec{0}$ is the position vector. These can be differentiated with respect to time using the last result of Section 3.2.

$$
\begin{aligned}
\dot{\ell}(U, t) &= \int_{U_t} \dot{v} \rho dV \\
\dot{a}(U, t) &= \int_{U_t} r \times \dot{v} \rho dV.
\end{aligned}
$$

To get the second equation, use the product rule and the fact that $\dot{r} = v$.

This background established, we can state the basic axioms of motion and force: the momentum balance laws. Observe that the first law is a reformulation of Newton's Second Law: *force* is equal to *mass* times *acceleration*.

Axiom. *The* momentum balance laws *are said to be satisfied if*

1. $f(U, t) = \dot{\ell}(U, t)$
2. $m(U, t) = \dot{a}(U, t)$.

The next result, Cauchy's existence of a stress, we would like to establish is fundamental to continuum mechanics. It states that a system of forces satisfies the momentum balance laws if and only if the surface traction $s(n)$, to be thought of as the material response to deformation, is linear. First we need a preliminary result, the *Principle of Virtual Work*.

Lemma 3.15. *Let (s, b) be a system of forces on $\mathcal{B}$ during a motion. A necessary and sufficient condition that the momentum balance laws be satisfied is that for every subregion $U \subset \mathcal{B}$ and time t,*

$$\int_{\partial U_t} s(n) \cdot w \, dA + \int_{U_t} (b - \rho \dot{v}) \cdot w \, dV = 0$$

for every infinitesimal rigid displacement w

This statement aligns with the intution that an approximately rigid deformation should not put any energy into the system.

Proof. Recall that ∇w is skew and constant. Write w uniquely as

$$w = w_0 + \omega \times r$$

where ω is the axial vector corresponding to ∇w. Using the symmetry of the vector triple-product, we have

$$\begin{aligned} s \cdot w &= w_0 \cdot s + s \cdot (\omega \times r) = w_0 \cdot s + \omega \cdot (r \times s) \\ (b - \rho \dot{v}) \cdot w &= w_0 \cdot s + s \cdot (\omega \times (b - \rho \dot{v})) = w_0 \cdot s + \omega \cdot (r \times (b - \rho \dot{v})). \end{aligned}$$

Now integrate the two previous equations and add them:

$$\begin{aligned} &\int_{\partial U_t} s(n) \cdot w \, dA + \int_{U_t} (b - \rho \dot{v}) \cdot w \, dV = \\ &w_0 \cdot \left(\int_{\partial U_t} s(n) dA + \int_{U_t} b - \rho \dot{v} dV \right) \\ &\qquad + \omega \cdot \left(\int_{\partial U_t} r \times s(n) dA + \int_{U_t} r \times (b - \rho \dot{v}) dV \right) \end{aligned}$$

For the above to vanish for all vectors w_0, ω (that is, for all infinitesimal rigid displacements), we must have

$$f(U,t) = \int_{\partial U_t} s(n)dA + \int_{U_t} bdV = \int_{U_t} \dot{v}\rho dV = \dot{\ell}(U,t)$$
$$m(U,t) = \int_{\partial U_t} r \times s(n)dA + \int_{U_t} r \times bdV = \int_{U_t} r \times \dot{v}\rho dV = \dot{a}(U,t).$$

These are precisely the momentum balance laws.

□

Now we are ready to state Cauchy's *Existence of a Stress.*

Theorem 3.16. *Let (s, b) be a system of forces for $\mathcal{B}$ during a motion x. The momentum laws are necessary and sufficient for the existence of a spatial tensor field T (referred to as the* Cauchy stress*) satisfying*

(1) $s(n) = Tn$ *for all* $n \in \mathcal{N}$,

(2) T *is symmetric, and*

(3) $\operatorname{div} T + b = \rho\dot{v}$.

Property (3) is known as the *equation of motion.* This is also one of the constitutive equations for the PDE governing elasticity, as we shall see later on.

Proof. First, we need to guarantee that taking divergence of T makes sense, that is, that T is smooth. This can be done by supposing that T satisfies (1). Choose an orthonormal basis e_1, e_2, e_3 of $\mathbb{R}^3$. Write n with respect to this basis: $n = \sum_i n_i e_i$. Then

$$(Te_i \otimes e_i)n = (Te_i \otimes e_i)(\sum_j n_j e_j) =$$
$$\sum_j n_j (Te_i \otimes e_i)e_j = \sum_j n_j (e_i \cdot e_j)Te_i = Tn_i e_i$$

by orthonormality. Hence,

$$\sum_i [(Te_i) \otimes e_i](\sum_j n_j e_j) = \sum_i T(n_i e_i) = Tn.$$

So $T = \sum_i (Te_i) \otimes e_i = \sum_i s(e_i) \otimes e_i$ by (1). Therefore, $T(x,t) = \sum_i s(e_i, x, t) \otimes e_i$, which implies, by the smoothness of the surface traction s, that the components of T with respect to the basis $e_i \otimes e_j$ are smooth. Hence, T is smooth if (1) holds.

Part 1: Necessity

The proof of (1) will go in three steps.

Step 1: Select an arbitrary orthonormal basis e_1, e_2, e_3, x in the interior of $\mathcal{B}_t$, and k a unit vector with $k \cdot e_i > 0$ (*i.e.* k lies in the positive orthant). Without loss of generality, assume $x = \vec{0}$. Construct a small tetrahedron P_δ whose vertices are all on the coordinate axes and whose face F_δ opposite to x

is a distance δ from x. We also require the span of k to be perpendicular to F_δ so that k is the outward-pointing normal vector field on F_δ.

Next, $\mathrm{Vol}(P_\delta) = C_1\delta^3$ for some constant $C_1 > 0$. Moreover, $\mathrm{Area}(\partial P_\delta) = C_2\delta^2$ for some constant $C_2 > 0$. Label the remaining faces $F_{i\delta}$ such that $-e_i$ is the outward-pointing unit normal vector field to $F_{i\delta}$.

For convenience, we use the notation from [Gur81a] and set $b_* = b - \rho\dot{v}$. Observe that b_* is continuous on $\mathcal{T}$. By the momentum balance law for force, and the fact that b is bounded,

$$\int_{\partial P_\delta} s(n)\, dA = \int_{P_\delta} b_*\, dV \Rightarrow$$
$$|\int_{\partial P_\delta} s(n)\, dA| \leq \kappa \mathrm{vol}(P_\delta)$$

By our analysis of the surface area and volume of P_δ,

$$\frac{1}{\mathrm{Area}(\partial P_\delta)} \int_{\partial P_\delta} s(n)\, dA \to 0$$

as $\delta \to 0$. Also,

$$\int_{\partial P_\delta} s(n)\, dA = \int_{F_\delta} s(k)\, dA + \sum_i \int_{F_{i\delta}} s(-e_i)\, dA. \tag{3.1}$$

Now we need a bit of elementary analytic geometry. The equation of the plane containing F_δ can be written as the set of vectors v satisfying

$$k \cdot v = \delta.$$

Let $k_i = k \cdot e_i$. Then the coordinates of the vertices of P_δ in the given basis are $(\delta/k_1, 0, 0), (0, \delta/k_2, 0)$, and $(0, 0, \delta/k_3)$. It follows that

$$\mathrm{Vol}(P_\delta) = \frac{\delta^3}{2k_1k_2k_3}, \quad \mathrm{Area}(F_\delta) = \frac{k_i\delta^2}{2k_1k_2k_3}$$

Also, by construction,

$$\mathrm{Vol}(P_\delta) = \delta \mathrm{Area}(F_\delta).$$

This implies that $\mathrm{Area}(F_{i\delta}) = (k \cdot e_i)\mathrm{Area}(F_\delta)$. Using this relation and the continuity of $s(n)$, we can localize to get

$$\frac{1}{\mathrm{Area}(F_\delta)} \int_{F_\delta} s(k)\, dA \to s(k, x)$$
$$\frac{1}{\mathrm{Area}(F_\delta)} \int_{F_{i\delta}} s(-e_i) dA \to (k \cdot e_i)s(-e_i, x).$$

as $\delta \to 0$. Then, Equation 3.1 implies that

$$s(k, x) = -\sum_i (k \cdot e_i)s(-e_i, x). \tag{3.2}$$

In the next two steps, we show that Equation 3.2 holds for all k.

Step 2:

We show that Newton's second law of action-reaction holds, that is

$$s(k, x) = -s(-k, x)$$

for all unit vectors k. Select an orthonormal basis $e_1 = k, e_2, e_3$ of $\mathbb{R}^3$. Select a sequence of vectors k_n satisfying $k_n \cdot e_i > 0$ such that $k_n \to k$ as $n \to \infty$. By Step 1, Equation 3.1 holds for each k_n. Since both sides of Equation 3.1 are continuous on $\mathcal{N}$,

$$s(k, x) = \lim_{n\to\infty} s(k_n, x) = \lim_{n\to\infty} -\sum_i (k \cdot e_i) s(-e_i, x) = -s(-k, x),$$

as desired.

Step 3:

Let e_1, e_2, e_3 be an orthonormal basis and k an arbitrary unit vector not contained in any of the coordinate planes formed by our choice of basis ($k \cdot e_i \neq 0$). Define a new orthonormal basis that $\{e_i'\}$ that satisfy $k \cdot e_i' > 0$ as follows:

$$e_i' := (\mathrm{sgn}(k \cdot e_i))e_i.$$

Thus, we can apply Equation 3.1 and the reflection principle from Step 2 to see

$$s(k, x) = -\sum_i (k \cdot e_i') s(-e_i', x) = -\sum_i (k \cdot e_i) s(e_i, x).$$

Therefore, Equation 3.1 holds for all $n \in \mathcal{N}$ and all $x \in \mathcal{B}$ by continuity. Indeed, $s(k, x)$ is a tensor field. The proof of (1) is complete.

To see that (2) is true, apply balance of linear momentum:

$$\int_{\partial U_t} T n dA + \int_{U_t} b dV = \int_{U_t} \dot{v}\rho \, dV,$$

the divergence theorem,

$$\int_{U_t} \mathrm{div}\, T + b \, dV = \int_{U_t} \dot{v}\rho \, dV,$$

and localize $\Rightarrow \mathrm{div} T + b - \rho\dot{v} = 0$.

Now we prove (3), the symmetry of T. By the divergence theorem and Proposition 2.18, if w is any smooth vector field on $\mathcal{T}$,

$$\begin{aligned}\int_{\partial U_t} Tn \cdot w = \int_{\partial U_t} (T^{\mathrm{T}} w) \cdot n dA &= \int_{U_t} \mathrm{div}(T^{\mathrm{T}} w) dV \\ &= \int_{U_t} (w \cdot \mathrm{div}\, T + T \cdot \mathrm{grad}\, w) \, dV.\end{aligned}$$

By part (2), $\mathrm{div}\, T = -b + \rho\dot{v}$, so

$$\int_{\partial U_t} s(n) \cdot w dA + \int_{U_t} (b - \rho\dot{v}) dV = \int_{U_t} T \cdot \mathrm{grad}\, w \, dV. \tag{3.3}$$

If w is any infinitesimal rigid displacement, we conclude from the necessity of the principle of virtual work that the LHS vanishes. In this case, $\operatorname{grad} w$ is skew. Hence, we can conclude $T \cdot W$ vanishes for any W skew. This implies that T is symmetric, because $\mathrm{Sym} \perp \mathrm{Skw}$.

Part 2: Sufficiency If there exists a spatial tensor field satisfying (1) - (3). If w is any infinitesimal rigid displacement, then $\operatorname{grad} w$ is skew, so that the RHS of Equation 3.3 vanishes. Applying the sufficiency of the principle of virtual work, we conclude that the momentum balance laws are satisfied.

□

We close this section with what [Gur81a] refers to as the *Theorem of Power Expended*. This computation will appear in various forms throughout our study of elasticity. Observe that physically, this law confirms that the aggregate of the mechanical and kinetic energy is equal to the work performed by external forces.

Theorem 3.17. *Recall that $D = \frac{1}{2}(\operatorname{grad} v + \operatorname{grad} v^{\mathrm{T}})$ is the stretching of the gradient. For every subregion $U \subset \mathcal{B}$ and time t,*

$$\int_{\partial U_t} s(n) \cdot v \, dA + \int_{U_t} b \cdot v \, dV = \int_{U_t} T \cdot D \, dV + \frac{d}{dt} \int_{U_t} \frac{v^2}{2} \rho \, dV$$

Proof. By the symmetry of T we see

$$T \cdot \operatorname{grad} v = T \cdot D.$$

Moreover, we can differentiate the second term on the RHS with respect to the mass measure

$$\frac{d}{dt} \int_{U_t} \frac{v^2}{2} \rho \, dV = \int_{U_t} v \cdot \dot{v} \rho \, dV.$$

Recalling that $b_* = b - \rho \dot{v}$,

$$\int_{U_t} b_* \cdot v \, dV = \int_{U_t} b \cdot v \, dV - \frac{d}{dt} \int_{U_t} \frac{v^2}{2} \rho \, dV.$$

Now apply Equation 3.3 with $w = v$ to get the proposition above.

□

3.4 Nonlinear Elasticity

At this point, we have all of the set-up needed to define an *elastic body*. First we need the notion of a *dynamical process*.

Definition 11. *A* dynamical process *is a pair (x, T) such that*

1. *x is a motion*
2. *T is a symmetric tensor field on the trajectory $\mathcal{T}$ of x, and*
3. *$T(x, t)$ is a smooth function of x on $\mathcal{B}_t$.*

For example, any force system (s, b) satisfying the momentum balance laws produces T such that (x, T) is a dynamical process.

We say that a *material body* is a body $\mathcal{B}$ with a mass density ρ and a family $\mathcal{D}$ of dynamical processes. These processes can be viewed as the permissible deformations and surface tractions that one can apply. Now we come to the central definition.

Definition 12. *An elastic body is a material body whose constitutive class consists of those pairs* (x, T) *where there exists a function* $\hat{T}$ *such that*

$$T(x(p,t),t) = \hat{T}(F(p,t),p) \text{ and}$$
$$Q\hat{T}(F)Q^{\mathrm{T}} = \hat{T}(QF)$$

for every rotation $Q \in \mathrm{Orth}^{+}$.

The first equation tells us that the material response is independent of the *past history* of the material; all the information that is needed to predict this response of the body is the material point p and the deformation gradient, a local measure of displacement, at time t. This is analogous to Hooke's law where the reaction force of a spring only depends on the magnitude of the displacement.

The second equation guarantees that the material response is *independent of observer*, a natural condition from physics.

Hence, one can think of the material response as a smooth function

$$\hat{T} : \mathrm{Lin}^{+} \times \mathcal{B} \to \mathrm{Sym}.$$

In this way, the material response is completely determined by time-independent, that is, *homogeneous*, deformations. The independence of observer also guarantees that $\hat{T}$ is completely determined by the stretching factor U of $F = RU$, where the RHS is a polar decomposition for F. Indeed

$$\hat{T}(F) = \hat{T}(RU) = Q\hat{T}(F)Q^{\mathrm{T}}.$$

It is a natural concept to consider the symmetries of the tensor function $\hat{T}$. A *symmetry transformation* at p is a rotation $Q \in \mathrm{Orth}^{+}$ such that

$$\hat{T}(F,p) = \hat{T}(FQ,p).$$

One can think of this definition physically as follows. The LHS describes the response of a material after running an experiment. The RHS describes changing one's frame of reference first, running the same experiment, and measuring the response. We have the following observation. Let $\mathcal{G}_p \subset \mathrm{Orth}^{+}$ be the set of symmetry transformations of $\hat{T}$.

Proposition 3.18. $\mathcal{G}_p$ *is a subgroup of* Orth^{+}.

Proof. We show that $\mathcal{G}_p$ is closed under multiplication and inversion. Suppose $Q \in \mathcal{G}_p$, then for all $F \in \mathrm{Lin}^{+}$,

$$\hat{T}(F) = \hat{T}(FQ).$$

Since $FQ^{\mathrm{T}} = FQ^{-1} \in \mathrm{Lin}^{+}$,

$$\hat{T}(FQ^{\mathrm{T}}) = \hat{T}(F)$$

so that $Q^{-1} \in \mathcal{G}_p$. Similarly if $Q, R \in \mathcal{G}_p$,

$$\hat{T}\left((FQ)R\right) = \hat{T}(FQ) = \hat{T}(F)$$

so that $QR \in \mathcal{G}_p$. $\square$

A material is said to be *isotropic* at p if it has the property that $\mathcal{G}_p = \mathrm{Orth}^+$. In this case, if $F = VR$ is a polar decomposition of F,

$$\hat{T}(F) = \hat{T}(VR) = \hat{T}(V)$$

so that $\hat{T}$ can be interpreted as a function $\hat{T} : \mathrm{Psym} \to \mathrm{Sym}$. Using Theorem 2.13, this implies

$$\hat{T}(F) = \nu_0(\mathcal{I}_V)I + \nu_1(\mathcal{I}_V)V + \nu_2(\mathcal{I}_V)V^{-1}$$

where the ν_i are scalar functions of the principal invariants of V. Another nice property for a body to satisfy is the following.

Definition 13. *A body $\mathcal{B}$ is said to be homogeneous if the material response $\hat{T}(F,p)$ is independent of the point p.*

Now we list the constitutive equations for nonlinear elasticity[3]. Observe that everything is defined in the *deformed configuration*, that is, in the trajectory $\mathcal{T}$ as opposed to the reference configuration.

Definition 14 (Nonlinear Elasticity in the Deformed Configuration)**.** *Let $\mathcal{B}$ be n elastic body with material response $\hat{T}$ and density field ρ. Suppose x is a motion with $\nabla x = F$, and (x,T) is in the constitutive class $\mathcal{D}$ of $\mathcal{B}$. Let b be the body force defined on $\mathcal{T}$. Then*

1. $T(x,t) = \hat{T}(F,p)$ *(elastic material response)*
2. $\operatorname{div} T + b = \rho \dot{v}$ *(equation of motion)*
3. $\rho \det F = \rho_0$ *(conservation of mass)*

From the perspective of solving PDE, working in the deformed configuration is not useful because the trajectory $\mathcal{T}$ is not known in advance. For the remainder of this section, we will describe how to pull back the elasticity problem to the *reference configuration* by what is known as the *Piola–Kirchhoff stress.*

Using Theorem 2.27, we have

$$\int_{\partial U_t} Tm\,dA = \int_{\partial U} (\det F)T_m F^{-\mathrm{T}} n\,dA.$$

This motivates the definition of the *Piola–Kirchhoff stress*, a material tensor field $S : \mathcal{B} \times \mathbb{R} \to \mathrm{Lin}$ such that

$$S := (\det F)T_m F^{-\mathrm{T}}.$$

[3]Nonlinear elasticity is also referred to in the literature as *finite elasticity.*

Let's also pull back the body force b using change of variables Theorem 2.25:

$$\int_{U_t} b\,dV = \int_U b_m(\det F),$$

so naturally we define $b_0 := b_m(\det F)$. Now we recover the balance laws satisfied in the reference configuration involving S.

Proposition 3.19. *The Piola-Kirchhoff stress tensor satisfies*

$$\int_{\partial U} Sn\,dA + \int_U b_0\,dV = \int_U \ddot{x}\rho_0\,dV$$
$$\int_{\partial U} (x-\vec{0}) \times Sn\,dA + \int_U (x-\vec{0}) \times b_0\,dV = \int_U (x-\vec{0}) \times \ddot{x}\rho_0\,dV$$

Proof. We prove the second part. Apply Theorem 2.27 to the balance of angular momentum.

$$\begin{aligned}
\int_{\partial U_t} (r-\vec{0}) \times Tn\,dA + \int_{U_t} (r-\vec{0}) \times b\,dV = \int_{U_t} (r-\vec{0}) \times \dot{v}\rho\,dV \Rightarrow \\
\int_{\partial U} (x-\vec{0}) \times T_m(\det F)F^{-\mathrm{T}}m\,dA + \int_U (x-\vec{0}) \times b_m \det F\,dV \\
= \int_U (x-\vec{0}) \times (\dot{v})_m\rho_m \det F\,dV \Rightarrow \\
\int_{\partial U} (x-\vec{0}) \times Sm\,dA + \int_U (x-\vec{0}) \times b_0\,dV = \int_U (x-\vec{0}) \times \ddot{x}\rho_0\,dV,
\end{aligned}$$

as desired. □

Localize the first equation in the proposition above and use the symmetry of T_m to see

$$\begin{aligned}
\operatorname{Div} S + b_0 &= \rho_0\ddot{x} \\
SF^{\mathrm{T}} &= FS^{\mathrm{T}}.
\end{aligned}$$

Observe that in this way, S is *not* generally symmetric. This is in contrast to the linearized theory where we will have symmetry. Now we derive a version of the Theorem of Power Expended (see Theorem 3.17) in the reference configuration.

Theorem 3.20. *Given a subregion $U \subset \mathcal{B}$,*

$$\int_{\partial U} Sn \cdot \dot{x}\,dA + \int_U b_0 \cdot \dot{x}\,dV = \int_U S \cdot \dot{F}\,dV + \frac{d}{dt}\int_U \frac{\ddot{x}^2}{2}\rho_0\,dV.$$

Proof. We take the original result Theorem 3.17 and tranform each piece one by one. The first term can be written, using the symmetry of T and Theorem 2.27

$$\int_{\partial U_t} Tm \cdot v\,dA = \int_{\partial U_t} Tv \cdot m\,dA = \int_{\partial U} Sn \cdot \dot{x}\,A.$$

Using the standard change of variables in three dimensions and bilinearity of $\cdot$,

$$\int_{U_t} b \cdot v \, dV = \int_U b_m \cdot v_m \det F \, dV = \int_U b_0 \cdot \dot{x}.$$

and

$$\int_{U_t} \frac{v^2}{2} \rho \, dV = \int_U \frac{\dot{x}}{2} \rho_0 \, dV.$$

Now, using the symmetry of T and the fact that $\dot{F} = L_m F$,

$$T \cdot D = T \cdot L = T \cdot \dot{F}_s F_s^{-1} = T F_s^{-\mathrm{T}} \cdot \dot{F}_s$$

which gives

$$\int_{U_t} T \cdot D \, dV = \int_U (\det F) T_m F^{-\mathrm{T}} \cdot \dot{F} = \int_U S \cdot \dot{F} \, dV.$$

□

This established, we can now state the constitutive equations for nonlinear elasticity in the reference configuration. In this case, although the domain of the PDE $\mathcal{B} \times \mathbb{R}$ is fixed, the equations themselves have become much more complicated.

Definition 15 (Nonlinear Elasticity in the Reference Configuration)**.** *Given an elastic body $\mathcal{B}$ and an initial density field ρ_0, a dynamical process (x, T) is in the constitutive class for $\mathcal{B}$ if the following equations are satisfied.*

$$\begin{aligned} F &= \nabla x \\ S = \hat{S}(F) &= (\det F) \hat{T}_m(F) F^{-\mathrm{T}} \\ \operatorname{Div} S + b_0 &= \rho_0 \ddot{x} \end{aligned}$$

Observe that the equation for conservation of mass is not required; the evolution of the body from the reference configuration is assured to satisfy conservation of mass by the second equation. Or, more simply stated, only the initial density ρ_0 shows up in the PDE, so it is not necessary to constrain ρ.

3.5 Small Deformations,Linearized Elasticity, and Variational Formulation

The nonlinear formulations of elasticity have their disadvantages of being quite complicated. Linearized elasticity provides a simplified model where the constitutive equations are more tractible. They are a valid approximation in the case where the displacement gradient ∇u is small. In this case the asymptotics

$$\begin{aligned} C &= I + \nabla u + \nabla u^{\mathrm{T}} + o(\nabla u) \\ B &= I + \nabla u + \nabla u^{\mathrm{T}} + o(\nabla u) \end{aligned}$$

guarantee that the *infinitesimal strain* (or *symmetrized gradient*) $E := \frac{1}{2}(\nabla u + \nabla u^{\mathrm{T}}$ is approximately $\frac{1}{2}(C - I)$ or $\frac{1}{2}(B - I)$. The symmetrized gradient makes an important appearance in the PDE's to follow.

In the rest of this work, we make the natural physical assumption that there is no stress in the reference configuration, *i.e.* $\hat{S}(I) = 0$. Next we use this assumption and smoothness to linearize the Piola-Kirchhoff stress.

$$
\begin{aligned}
\hat{S}(F) = \hat{S}(I + \nabla u) &= \hat{S}(I) + D\hat{S}(I)[\nabla u] + o(\nabla u) \\
&= D\hat{S}(I)[\nabla u] + o(\nabla u)
\end{aligned}
$$

To go further, it will be necessary to get more information about the linear transformation $\hat{S}(I)$, referred to as the *elasticity tensor* and written

$$C[H] := D\hat{S}(I)[H].$$

This *fourth-order tensor* C enjoys several nice properties.

Theorem 3.21 (Properties of the Elasticity Tensor)**.** *Let $C = D\hat{S}(I)$. Then*

(1) $C = D\hat{T}(I)$

(2) $C[H] \in \mathrm{Sym}$ *for all* $H \in \mathrm{Lin}$ *and* $C[W] = 0$ *if* $W \in \mathrm{Skw}$

(3) C is isotropic at p if $\mathcal{B}$ is isotropic at p

(4) If C is isotropic at p, there exist scalars μ and λ on $\mathcal{B}$ known as the Lamé moduli *such that at p,*

$$C[H] = 2\mu H + \lambda(\mathrm{tr}H)I$$

whenever H is symmetric.

Suppose $\mathcal{B}$ is isotropic at p. Then C is symmetric (*i.e.* $H \cdot C[G] = G \cdot C[H]$).Also, C is positive definite (*i.e.* $H \cdot C[H] > 0$ for all *symmetric* tensors H) if $\mu > 0$ and $2\mu + 3\lambda > 0$.

The above properties imply that we can rewrite the initial asymptotic as

$$\hat{S}(F) = C[H] + o(H) = C[E] + o(H)$$

where, recall, $H = \nabla u$ and $E = \frac{1}{2}(\nabla u + \nabla u^{\mathrm{T}})$ is the *infinitesimal strain.* This motivates us to truncate the Piola-Kirchhoff tensor and formulate a linear system of PDE as follows.

Definition 16 (Constitutive Equations for Linearized Elasticity)**.** *Given a material body $\mathcal{B}$, a motion x, material body force b_0, and an initial density ρ_0, the following equations form the basis of linearized elasticity. Recall that $u(p,t) = x(p,t) - p$ is the displacement.*

1. $E = \frac{1}{2}(\nabla u + \nabla u^{\mathrm{T}})$ *(Infinitesimal Strain)*

2. $S = C[E]$ *(Linearization of $\hat{S}$)*

3. $\mathrm{Div}\, S + b_0 = \rho_0 \ddot{u}$ *(Equation of Motion)*

These assumptions lead to a simpler theory. For example, observe that S is now a symmetric matrix, in contrast to the nonlinear theory. We close our study of elasticity with a derivation of a variational formulation for linearized elasticity: this is an optimization problem whose solution satisfies the constitutive equations above. For our purposes, we will pose this problem for *static* linearized elasticity; this is the case when the motion x is independent of time so that $\dot{u} = 0$. Physically, the body is deformed from the reference configuration in a way that preserves the momentum balance laws. However, for the case of linearized elasticity (as opposed to the nonlinear case), the existence of a solution *requires* that momentum balance laws be satisfied in the reference configuration. To make progress toward this destination, we begin with the definition of an elastic state.

Definition 17. *An* elastic state *is a triple of smooth fields* (u, E, S) *satisfying conditions (1) - (3) from Definition 16 above for a given body force* b.

This allows to state the boundary value problem for linearized elastostatics.

Definition 18 (Boundary value problems)**.** *Let* B_1 *and* B_2 *be subregions of* $\mathcal{B}$ *such that* $\partial\mathcal{B} = B_1 \cup B_2$ *and* $B_1 \cap B_2 = \emptyset$*, so that the union is disjoint. The mixed problem is stated below.*

Given the following data: a body $\mathcal{B}$*, boundary decomposition* B_1, B_2*, an elasticity tensor* C*, a body force* b*, a surface displacement (a deformation)* $\hat{u}$ *on* B_1*, and surface traction (a vector field)* $\hat{s}$ *on* B_2*, find an elastic state* $[u, E, S]$ *corresponding to* b *that is compatible with the boundary conditions in the sense that:*

$$u|_{B_1} = \hat{u}, \quad S|_{B_2} n = \hat{s}.$$

Such a triple is called a solution *to the mixed problem. If* $B_1 = \partial\mathcal{B}$*, then* $[u, E, S]$ *is a solution to the* displacement problem*. If* $B_2 = \mathcal{B}$*, then* $[u, E, S]$ *is a solution to the* traction problem*.*

A *kinematically admissible state* is a triple of smooth fields $[u, E, S]$ satisfying properties (1) and (2) from Definition 16. Our goal is to prove that a solution of the mixed problem satisfies a minimization property over all kinematically admissible states.

Theorem 3.22 (Principle of Minimum Potential Energy)**.** *Suppose* C *is symmetric and positive definite. Let* $s = [u_1, E_1, S_1]$ *be a solution of the mixed problem from Definition 18. Define a functional* Φ *on the set of all kinematically admissible states* $s' = [u', E', S']$ *by*

$$\Phi(s') = \frac{1}{2}\int_{\mathcal{B}} E' \cdot C[E']\, dV - \int_{\mathcal{B}} b \cdot u'\, dV - \int_{B_2} \hat{s} \cdot u'\, dA. \tag{3.4}$$

Then

$$\Phi(s) \leq \Phi(s')$$

for every kinematically admissible state s'*, and equality holds if and only if* $u' = u + w$ *for some infinitesimal rigid displacement* w*.*

The first term on the RHS of Equation (3.4) is referred to as the strain energy and written

$$\mathcal{U}(E) := \frac{1}{2}\int_{\mathcal{B}} E \cdot C[E].$$

If we think of the stress and strain as *dual* in a functional analytic sense (as is the case in the convex-analytic perspective on plasticity), then the above is simply a rescaling of the natural action of the strain on the stress.

To prove Theorem 3.22, we need the following easy lemma that makes use of the symmetry of S, a fact that did not hold in the nonlinear case.

Lemma 3.23. *Suppose $[u, E, S]$ is a triple of smooth fields satisfying conditions (1) and (3) from Definition 16. Then*

$$\int_{\partial\mathcal{B}} Sn \cdot u \, dA + \int_{\mathcal{B}} b \cdot u \, dV = \int_{\mathcal{B}} S \cdot E \, dV.$$

Proof. Apply the divergence theorem, the symmetry of S, and a divergence identity from Proposition 2.18, in that order:

$$\begin{aligned}\int_{\partial\mathcal{B}} Sn \cdot u \, dA &= \int_{\partial\mathcal{B}} Su \cdot n \, dA \\ &= \int_{\mathcal{B}} \mathrm{Div}(Su) \, dV \\ &= \int_{\mathcal{B}} (u \cdot \mathrm{Div}\, S + S \cdot \nabla u) \, dV.\end{aligned}$$

Now, $\mathrm{Div}\, S = -b$ by assumption and $S \cdot \nabla u = S \cdot E$ by the symmetry of S. Regrouping terms yields the lemma statement. □

The proof of the previous lemma can be slightly modified to yield the following.

Lemma 3.24. *Suppose S satisfies*

$$\mathrm{Div}\, S + b = 0.$$

Let u' be a smooth field, and define

$$E' = \frac{1}{2}(\nabla u + \nabla u^{\mathrm{T}})$$

to be the symmetrized gradient. Then

$$\int_{\partial\mathcal{B}} Sn \cdot u' \, dA + \int_{\mathcal{B}} b \cdot u' \, dV = \int_{\mathcal{B}} S \cdot E'.$$

This lemma has the interesting consequence of placing restrictions on the initial data for a traction problem. Namely, the reference configuration must satisfy the momentum balance laws.

Proposition 3.25. *A necessary condition for the* traction problem *to have a solution is that*

$$\int_{\partial\mathcal{B}} \hat{s}\, dA + \int_{\mathcal{B}} b\, dV = 0$$
$$\int_{\partial\mathcal{B}} r \times \hat{s}\, dA + \int_{\mathcal{B}} r \times b\, dV = 0.$$

Proof. The first statement immediately follows from the divergence theorem and the equation of motion. For the second statement, let u be a linear infinitesimal rigid displacement on $\mathcal{B}$ given by

$$u(p) = \omega \times r(p).$$

Then by the previous lemma, the fact that ∇u is skew, and the triple product identity

$$0 = \int_{\partial\mathcal{B}} \hat{s} \cdot u\, dA + \int_{\mathcal{B}} b \cdot u\, dV = \int_{\partial\mathcal{B}} \omega \cdot (r \times \hat{s})\, dA + \int_{\mathcal{B}} \omega \cdot (r \times b)\, dV \Rightarrow$$
$$0 = \omega \cdot \left(\int_{\partial\mathcal{B}} r \times \hat{s}\, dA + \int_{\mathcal{B}} r \times b\, dV \right).$$

The above is true for any choice of ω, so we conclude the second equation holds from proposition statement.

□

The lemma above also has the immediate consequence (letting $S = C[E]$)

$$\int_{\partial\mathcal{B}} Sn \cdot u\, dA + \int_{\mathcal{B}} b \cdot u\, dV = 2\mathcal{U}(E).$$

This fact is used to prove uniqueness of solutions up to infinitesimal rigid displacement.

Proposition 3.26. *Suppose C is positive definite and that $[u_1, E_1, S_1]$ and $[u_2, E_2, S_2]$ solve the same mixed problem. Then $E_1 = E_2$, $S_1 = S_2$, and $u_1 = u_2 + w$, where w is some infinitesimal rigid displacement.*

Proof. The idea is to superimpose the two solutions. Let

$$w = u_1 - u_2, \quad E = E_1 - E_2, \quad S = S_1 - S - 2.$$

By linearity of the PDE, $[w, E, S]$ is a solution to the mixed problem with $w = 0$ on B_1 and $Sn = 0$ on B_2. Also,

$$\operatorname{Div} S = \operatorname{Div}(S_1 - S_2) = 0,$$

so that $b = 0$ in this new mixed problem. Therefore,

$$\int_{\partial\mathcal{B}} Sn \cdot w\, dV + \int_{\mathcal{B}} b \cdot w = \int_{B_1} Sn \cdot w\, dV + \int_{B_2} Sn \cdot w\, dV + 0 = 0.$$

Hence,

$$\int_{\mathcal{B}} E \cdot C[E] = 0,$$

so by the positive definiteness of C,

$$E \equiv 0.$$

This implies ∇w is skew, so that w is indeed an infinitesimal rigid displacement. □

Now we close this section with a proof of Theorem 3.22.

Proof of Theorem 3.22: Principle of Minimum Potential Energy. Define

$$w = u' - u, \quad \bar{E} = E' - E,$$

so that $\bar{E} = \frac{1}{2}(\nabla w + \nabla w^{\mathrm{T}})$ and $w|_{B_1} = 0$ because both s and s' are kinematically admissible. By the symmetry of C,

$$E' \cdot C[E'] = E \cdot C[E] + 2C[E] \cdot \bar{E} + \bar{E} \cdot C[\bar{E}].$$

Integrating both sides, we have

$$\mathcal{U}(E') - \mathcal{U}(E) = \mathcal{U}(\bar{E}) + \int_{\mathcal{B}} S \cdot \bar{E}.$$

By Lemma 3.24,

$$\int_{\mathcal{B}} S \cdot \bar{E}\, dV = \int_{\partial\mathcal{B}} Sn \cdot w\, dA + \int_{\mathcal{B}} b \cdot w\, dV = \int_{B_2} \hat{s} \cdot w\, dA + \int_{\mathcal{B}} b \cdot w\, dV.$$

Since $w = u' - u$, the last two equations yield

$$\Phi(s') - \Phi(s) = \mathcal{U}(\bar{E}) \geq 0$$

Moreover, equality is attained only if $\bar{E} = 0 \Leftrightarrow w$ is an infinitesimal rigid displacement. The theorem is proved. □

We close with the statement of the *integral* or *weak form* of the PDE for linearized elasticity. Observe that, using Lemma 3.23, the strong formulation in Definition 18 implies that the weak formulation will be satisfied.

Definition 19 (Weak formulation of elasticity)**.** *Given a decomposition*

$$\mathcal{B}_1 \sqcup \mathcal{B}_2 = \partial\mathcal{B}$$

and $\hat{s}$ a smooth displacement on $\mathcal{B}_2$, find a smooth displacement $u : \mathcal{B} \to \mathbb{R}^3$ such that

$$\int_{\mathcal{B}} S \cdot \nabla v\, dV = \int_{\mathcal{B}} v \cdot b\, dV + \int_{\mathcal{B}_2} v \cdot \hat{s}\, dA$$

for all smooth vector fields v vanishing on $\mathcal{B}_1$ and having integrable gradient on $\mathcal{B}$.

This concludes our discussion of elasticity.

References

[AP10] Basile Audoly and Yves Pomeau. *Elasticity and Geometry.* Oxford University Press, 198 Madison Avenue New York, NY 10016 USA, 2010.

[Gur81a] Morton E. Gurtin. *An Introduction to Continuum Mechanics.* Academic Press, Inc., 111 Fifth Avenue, New York, New York, 10003, 1981.

[Gur81b] Morton E. Gurtin. *Topics in Finite Elasticity.* Society for Industrial and Applied Mathematics, 3600 Market Street, Philadelphia, PA 19104-2688, 1981.

[Roc99] R. Rockafellar. *Convex Analysis.* Springer, 233 Spring Street New York, NY 10013-1578 USA, 1999.

[RW99] Daya Reddy and Han Weimin. *Plasticity: Mathematical Theory and Numerical Analysis.* Springer, 233 Spring Street New York, NY 10013-1578 USA, 1999.