Continuum mechanics

Ross Grogan-Kaylor, Sandra Hager

This presentation adapts material mostly from Bonet and Wood's *Nonlinear Continuum Mechanics for Finite Element Analysis* and somewhat from McGinty's `continuummechanics.org`.

Linear algebra review

Why linear algebra?

1. Matrices are used to simply store data that is inherently multidimensional. At the same time, matrices also describe linear functions, which arise from natural physical assumptions. The Cauchy stress σ and the strains $\mathbf{E}$ and $\mathbf{e}$ are notably symmetric linear functions.
2. Linear algebra formalizes ideas about the coordinate systems that we use to describe deformations. The most important formalization for this presentation is the idea of *invariance under choice of basis*, or *objectivity*.

Outline of linear algebra review

- Two perspectives on matrices
- Matrices as properties
- Matrices as linear functions
  - Identity and inverse matrices
  - Matrix-matrix multiplication
  - Symmetric matrices
  - Orthogonal matrices
- Invariants
  - Definition of invariant
  - Determinants
  - Eigenstuff
  - Double contraction
  - Invariants of a mmtrix

Two perspectives on matrices

A *matrix* is a grid of numbers. Here is a 3×3 matrix:

$$\begin{pmatrix} 1 & -2 & \pi \\ 0 & 10 & 8 \\ -5 & e & 14 \end{pmatrix}.$$

We use 3×3 matrices to describe *3D properties* of 3D objects, and to describe how these properties are *interrelated* and *change* over time.

When we want to describe *3D properties*, use 3×3 matrices that are thought of as *simply storing data*.

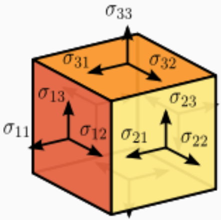
When we want to talk about how these properties are interrelated or change, we use 3×3 matrices that are interpreted as descriptive of a special type of function, a *linear function*.

Both interpretations are used simultaneously when considering stress and strain tensors!

In engineering contexts, matrices are called "tensors" when they are intended to be thought of as simply "arrays of numbers." The word "tensor" can actually mean "special type of vector," "array of numbers," and "linear function" all at once! See the section on $\binom{p}{q}$ tensors in the appendix.

Tensors: matrices as properties

In continuum mechanics, a total description of the stress-state at a point is given by a matrix of numbers. *A single number is not enough information!* The stress *tensor*, a 3×3 matrix of numbers, is what describes the stress-state at a point.



$$\boldsymbol{\sigma} = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix}$$

- σ_{ii} is the instantaneous normal stress in the direction of $\hat{\mathbf{x}}_i$.
- σ_{ij} is the instantaneous shear stress in the plane spanned by $\hat{\mathbf{x}}_i$ and $\hat{\mathbf{x}}_j$.

Matrices as linear functions

The various stress and strain tensors not only store data, but are also linear functions .

A function $T : \mathbb{R}^3 \to \mathbb{R}^3$ is said to be *linear* if and only if T *preserves the composition of any input vector*; that is, T is linear if and only if, for all vectors $\mathbf{v} \in \mathbb{R}^3$,

$$T(\mathbf{v}) = T(v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + v_3\mathbf{e}_3) = v_1 T(\mathbf{e}_1) + v_2 T(\mathbf{e}_2) + v_3 T(\mathbf{e}_3).$$

But if we know what happens to the $\mathbf{e}_i$, then we know what happens to any $\mathbf{v}$! This implies that if you know the matrix

$$\begin{pmatrix} T(\mathbf{e}_1) & T(\mathbf{e}_2) & T(\mathbf{e}_3) \end{pmatrix},$$

then you know what T is. So, every linear function is completely specified if you know the matrix that corresponds to it. We can go the “other way,” too: every matrix corresponds to a linear function.

The *identity* matrix $\mathbf{I}$ is the matrix for which

$$\mathbf{IA} = \mathbf{A} = \mathbf{AI},$$

for all matrices $\mathbf{A}$.

A matrix is *invertible* if and only if the linear function it describes is one-to-one. (This condition is analogous to the horizontal line test for functions $\mathbb{R} \to \mathbb{R}$).

If a matrix is $\mathbf{A}$ invertible, then it has an *inverse matrix* $\mathbf{A}^{-1}$ for which

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{I} = \mathbf{AA}^{-1}.$$

As was noted above, not all matrices have inverses.

The matrix-matrix product of matrices **B** and **A** returns a matrix **BA** that describes the linear function which performs **A**, and then **B**. That is, **BA** is the matrix for which $(\mathbf{BA})\mathbf{v} = \mathbf{B}(\mathbf{A}\mathbf{v})$.

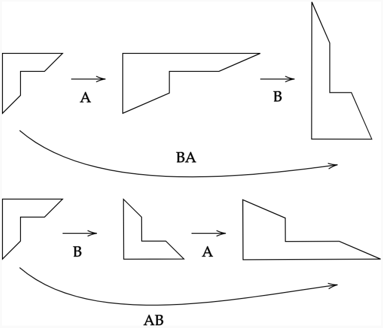
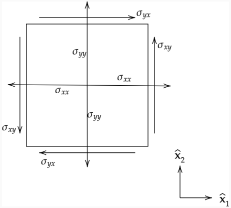


Figure: Here we compare **BA** with **AB**, where **A** is a horizontal stretch and **B** is a 90° counterclockwise rotation. Note that the order of composition is "backwards": the matrix on the right is applied before the one on the left. Most importantly, the figure demonstrates that $\mathbf{BA} \neq \mathbf{AB}$ in general .

The *transpose* of a matrix $\mathbf{A}$ is denoted $\mathbf{A}^\top$ and has *ij* entry a_{ji} .

A matrix $\mathbf{A}$ is *symmetric* if and only if $\mathbf{A} = \mathbf{A}^\top$. Example: Cauchy stress tensor $\boldsymbol{\sigma}$.



The top and bottom stresses and the left and right stresses are equal due to translational equilibrium. Then $\sigma_{xy} = \sigma_{yx}$ due to rotational equilibrium. In fact, if Cauchy's postulate is assumed, then the stress tensor is always symmetric in realistic[1] conditions .

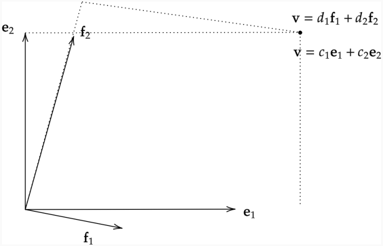
$$\boldsymbol{\sigma} = \begin{pmatrix} \sigma_{xx} & \sigma_{yx} & \sigma_{zx} \\ \sigma_{xy} & \sigma_{yy} & \sigma_{zy} \\ \sigma_{xz} & \sigma_{yz} & \sigma_{zz} \end{pmatrix} = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix}.$$

[1] when the stress tensor is continuous and only depends on the surface normal

- A matrix $\mathbf{A}$ is *orthogonal* if and only if $\mathbf{A}$ is length-preserving, that is, if and only if $||\mathbf{v}|| = ||\mathbf{A}\mathbf{v}||$ for all $\mathbf{v}$.
- The inverse of any orthogonal matrix $\mathbf{A}$ is $\mathbf{A}^{-1} = \mathbf{A}^{\top}$. So orthogonal matrices are easily invertible.
- Example: rotation matrices are orthogonal.

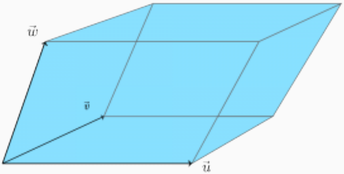
Consider a basis $\mathbf{F} = \begin{pmatrix} \mathbf{f}_1 & \dots & \mathbf{f}_n \end{pmatrix}$ of $\mathbb{R}^n$.

If $\mathbf{v} = d_1\mathbf{f}_1 + \ldots + d_n\mathbf{f}_n$, then we say that $d_1, \ldots, d_n$ are the *coordinates of* $\mathbf{v}$ *relative to the basis* $\mathbf{F}$.



We say that a function (not necessarily linear) accepting input from $\mathbb{R}^n$ is *basis-independent*, *invariant*, or *objective* if and only if the output of the function does not depend on the basis used for $\mathbb{R}^n$.

The determinant of a matrix is the oriented volume of the parallelapiped spanned by the matrix's columns.



$$\det(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \det \begin{pmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{pmatrix}.$$

If **A** and **B** represent the same linear function[2], but the bases chosen for **A** and **B** differ, then $\det(\mathbf{A}) = \det(\mathbf{B})$. Therefore, the determinant is invariant with respect to the bases chosen to represent its matrix argument.

[2] See the "transformation in a basis" in the appendix for the formalization of this idea.

Consider a linear function $T : \mathbb{R}^n \to \mathbb{R}^m$ that has the $m \times n$ matrix $\mathbf{A}$.

If a vector $\mathbf{v}$ and a scalar λ are such that

$$T(\mathbf{v}) = \lambda \mathbf{v},$$

then we say $\mathbf{v}$ is an *eigenvector* of T with *eigenvalue* λ.

Eigenvectors and eigenvalues are both invariant under change of basis.

The eigenvectors and eigenvalues of a stress or strain tensor are called the *principal directions* and *principal values* of that tensor.

Double contraction Invariants

Given two rank 2 tensors (matrices) **A** and **B**, the *double contraction* of **A** and **B** produces a scalar:

$$\mathbf{A} : \mathbf{B} = \sum_{i,j=1}^{n} a_{ij} b_{ij}.$$

Note that

$$\mathbf{A} : \mathbf{B} = \operatorname{tr}(\mathbf{A}^\top \mathbf{B}),$$

where the trace tr of a matrix is the sum of its diagonal entries,

$$\operatorname{tr}\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{11} & a_{22} & a_{23} \\ a_{11} & a_{32} & a_{33} \end{pmatrix} = a_{11} + a_{22} + a_{33}.$$

Invariants of a matrix

An *invariant* of a linear function $T : \mathbb{R}^n \to \mathbb{R}^m$ is any function (i.e. property) of T that does not depend on the basis chosen for $\mathbb{R}^n$.

Recall that eigenvalues (principal values) are invariant under change of basis. For any 3×3 matrix $\mathbf{A}$, the following are invariant because they are functions of eigenvalues:

$$
\begin{aligned}
I_{\mathbf{A}} &= \mathrm{tr}(\mathbf{A}) = \lambda_1 + \lambda_2 + \lambda_3 \\
II_{\mathbf{A}} &= \mathbf{A} : \mathbf{A} = \mathrm{tr}(\mathbf{A}^\top \mathbf{A}) = \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3 \\
III_{\mathbf{A}} &= \det(\mathbf{A}) = \lambda_1\lambda_2\lambda_3.
\end{aligned}
$$

Use the Jordan canonical form to justify the eigenvalue formula for $I_{\mathbf{A}}$. For $III_{\mathbf{A}}$, this argument is actually circular because we only know $\det(\mathbf{A}) = \lambda_1\lambda_2\lambda_3$ due to the basis-independence of $\det$.

Kinematics

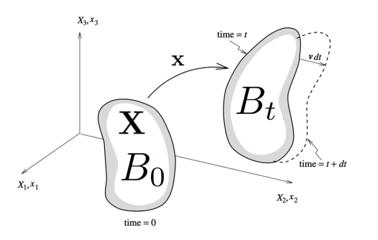
Outline of kinematics concepts

- Deformation, Lagrangian and Eulerian quantities
- Deformation gradient
- Polar decomposition

Deformation, Lagrangian vs. Eulerian

$$\mathbf{x} : \underbrace{B_0}_{\text{space}, \subseteq \mathbb{R}^3} \times \underbrace{\mathbb{R}}_{\text{time}} \to \underbrace{B_t}_{\text{space}, \subseteq \mathbb{R}^3} \times \underbrace{\mathbb{R}}_{\text{time}}$$

$$\mathbf{x} = \mathbf{x}(\mathbf{X}, t)$$



- "Lagrangian" refers to the reference configuration ; uppercase letters (e.g. $\mathbf{X}$) used for this
- "Eulerian" refers to the deformed configuration ; lowercase letters (e.g. $\mathbf{x}$) used for this

The *deformation gradient* **F** is defined as the following Jacobian matrix:

$$\mathbf{F} := \frac{\partial \mathbf{x}}{\partial \mathbf{X}}$$

$$F_{ij} = \frac{\partial x_i}{\partial X_j}, \mathbf{x}(\mathbf{X}) = (x_1(\mathbf{X}), ..., x_n(\mathbf{X})), \mathbf{X} = (X_1, ..., X_n)$$

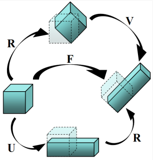
F is a Lagrangian-Eulerian hybrid!

F acts on a vector displacement $\Delta\mathbf{X}$ in the reference configuration to produce a vector displacement in the deformed configuration:

$$\Delta\mathbf{x} = \frac{\partial \mathbf{x}}{\partial \mathbf{X}} \Delta\mathbf{X} = \mathbf{F}\Delta\mathbf{X}.$$

Use the mnemonic that $\partial\mathbf{X} = \Delta\mathbf{X}$ in the "denominator" cancels with the other $\Delta\mathbf{X}$ to remember.

Polar decomposition



$$\mathbf{F} = \mathbf{V}\mathbf{R} = \mathbf{R}\mathbf{U},$$

where **R** is a rotation matrix . Note that **R** is the same in both decompositions of **F**.

$$\mathbf{F}^\top\mathbf{F} = \mathbf{U}^\top\mathbf{R}^\top\mathbf{R}\mathbf{U} = \mathbf{U}\mathbf{R}^{-1}\mathbf{R}\mathbf{U} = \mathbf{U}^2$$
$$\mathbf{F}\mathbf{F}^\top = \mathbf{V}\mathbf{R}\mathbf{R}^\top\mathbf{V}^\top = \mathbf{V}\mathbf{R}\mathbf{R}^{-1}\mathbf{V} = \mathbf{V}^2.$$

$\mathbf{F}^\top\mathbf{F} = \mathbf{U}^2$ is called the *right Cauchy-Green deformation tensor*, and $\mathbf{F}\mathbf{F}^\top = \mathbf{V}^2$ is called the *left Cauchy-Green deformation tensor*. $\mathbf{F}^\top\mathbf{F} = \mathbf{U}^2$ is Lagrangian and $\mathbf{F}\mathbf{F}^\top = \mathbf{V}^2$ is Eulerian. Both are symmetric, rotation-independent matrices.

Strain

Outline of strain concepts

- Engineering strains (normal and shear)
- Lagrangian (Green) and Eulerian (Almansi) strains

Overleaf

Strain is a measure of how much a material has stretched.

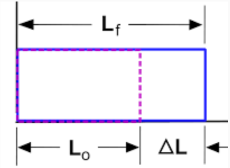


Figure: Engineering normal strain in one dimension

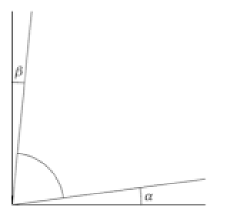


Figure: Engineering shear strain in two dimensions

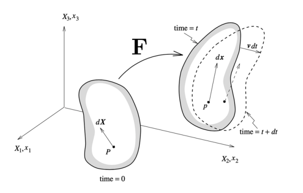
$$\epsilon_{\text{eng}} := \frac{\Delta L}{L_0} \quad \text{(one-dimensional case)}$$

$$\gamma_{\text{eng}} := \tan(\alpha + \beta) \quad \text{(two-dimensional case)}$$

Lagrangian strain

To derive the Lagrangian strain, we need to know the following:

$$
\begin{aligned}
\Delta\mathbf{x} \cdot \Delta\mathbf{x} &= (\mathbf{F}\Delta\mathbf{X}) \cdot (\mathbf{F}\Delta\mathbf{X}) \\
&= (\mathbf{F}\Delta\mathbf{X})^{\top}\mathbf{F}\Delta\mathbf{X} = (\Delta\mathbf{X})^{\top}\mathbf{F}^{\top}\mathbf{F}\Delta\mathbf{X} \\
&= \Delta\mathbf{X} \cdot \mathbf{F}^{\top}\mathbf{F}\Delta\mathbf{X}.
\end{aligned}
$$



The Lagrangian strain tensor **E** is discovered as follows:

$$
\begin{aligned}
\Delta\mathbf{x} \cdot \Delta\mathbf{x} - \Delta\mathbf{X} \cdot \Delta\mathbf{X} &= \Delta\mathbf{X} \cdot \mathbf{F}^{\top}\mathbf{F}\Delta\mathbf{X} - \Delta\mathbf{X} \cdot \Delta\mathbf{X} \\
&= \Delta\mathbf{X} \cdot (\mathbf{F}^{\top}\mathbf{F}\Delta\mathbf{X} - \Delta\mathbf{X}) \\
&= \Delta\mathbf{X} \cdot (\mathbf{F}^{\top}\mathbf{F} - \mathbf{I})\Delta\mathbf{X} \\
&= \Delta\mathbf{X} \cdot (\underbrace{2\mathbf{E}}_{\text{def}})\Delta\mathbf{X}.
\end{aligned}
$$

$$
\mathbf{E} := \frac{1}{2}(\mathbf{F}^{\top}\mathbf{F} - \mathbf{I}) = \frac{1}{2}(\mathbf{U}^2 - \mathbf{I}).
$$

Eulerian strain

To derive the Eulerian strain, we again consider $\Delta\mathbf{x} \cdot \Delta\mathbf{x} - \Delta\mathbf{X} \cdot \Delta\mathbf{X}$.

The Eulerian strain tensor $\mathbf{e}$ is discovered as follows:

$$\begin{aligned}
\Delta\mathbf{x} \cdot \Delta\mathbf{x} - \Delta\mathbf{X} \cdot \Delta\mathbf{X} &= \Delta\mathbf{x} \cdot \Delta\mathbf{x} - (\mathbf{F}^{-1}\Delta\mathbf{x})^\top \mathbf{F}^{-1}\Delta\mathbf{x} \\
&= \Delta\mathbf{x} \cdot \Delta\mathbf{x} - (\Delta\mathbf{x})^\top \mathbf{F}^{-\top}\mathbf{F}^{-1}\Delta\mathbf{x} \\
&= \Delta\mathbf{x} \cdot (\Delta\mathbf{x} - \mathbf{F}^{-\top}\mathbf{F}^{-1}\Delta\mathbf{x}) = \Delta\mathbf{x} \cdot (\mathbf{I} - \mathbf{F}^{-\top}\mathbf{F}^{-1})\Delta\mathbf{x} \\
&= \Delta\mathbf{x} \cdot (\underbrace{2\mathbf{e}}_{\text{def}})\Delta\mathbf{x}.
\end{aligned}$$

$$\mathbf{e} = \frac{1}{2}(\mathbf{I} - (\mathbf{F}\mathbf{F}^\top)^{-1}) = \frac{1}{2}(\mathbf{I} - \mathbf{V}^{-2}).$$

Note: $(\mathbf{A}^\top)^{-1} = (\mathbf{A}^{-1})^\top$, so we denote both by $\mathbf{A}^{-\top}$.

Lagrangian and Eulerian strains

We've defined

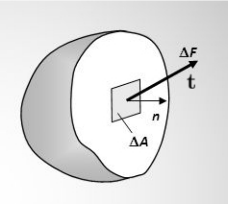
$$\begin{aligned} \mathbf{E} &= \frac{1}{2}(\mathbf{F}^\top \mathbf{F} - \mathbf{I}) = \frac{1}{2}(\mathbf{U}^2 - \mathbf{I}) \\ \mathbf{e} &= \frac{1}{2}(\mathbf{I} - (\mathbf{F}\mathbf{F}^\top)^{-1}) = \frac{1}{2}(\mathbf{I} - \mathbf{V}^{-2}). \end{aligned}$$

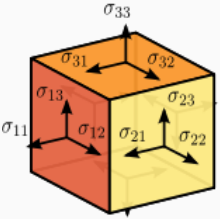
- The Lagrangian and Eulerian strains **E** and **e** are more commonly referred to as the *Green* and *Almansi* strains.
- Both **E** and **e** are symmetric and independent of rotation.
- (When strains are small enough for quadratic terms to be neglected, then $\mathbf{E} \approx \mathbf{e} \approx \epsilon$, where $\epsilon = \frac{1}{2}(\mathbf{F} + \mathbf{F}^\top) - \mathbf{I}$ is the small strain tensor. See the appendix.)

Stress

Outline of stress concepts

- Traction vector and Cauchy stress tensor
- Balance equations and constitutive models
  - Example: Hooke's law
- First Piola-Kirchhoff stress
- Second Piola-Kirchhoff stress
- Prelude to hyperelasticity: internal energy





$$\mathbf{t} := \frac{\Delta \mathbf{F}}{\Delta A}$$

The traction vector linearly depends on the surface normal in equilibrium.

$$\mathbf{t} = \underbrace{\boldsymbol{\sigma}}_{3\times 3}\hat{\mathbf{n}}$$

If we assume that the traction vector only depends on the surface normal (Cauchy's postulate) in a continuous manner, then this always holds .

Balance of mass:

$$\underbrace{\int_v \rho dv}_{\text{change in mass in time } dt} = \underbrace{-\int_{\partial v} \rho(\mathbf{v} \cdot \hat{\mathbf{n}} dt) da}_{\text{flux of mass out of boundary in time } dt}$$

$$\frac{\partial \rho}{\partial t} = -\mathrm{div}(\rho \mathbf{v}).$$

Balance of linear momentum, in differential form:

$$\underbrace{\mathrm{div}(\boldsymbol{\sigma})}_{\text{volume density of surface force}} + \underbrace{\rho \mathbf{f}}_{\text{volume-density of body force}} = \underbrace{\rho \mathbf{a}}_{\text{volume-density of force}}.$$

Balance of angular momentum:

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}^\top.$$

ρ : volume-density of mass, $\mathbf{v}$: velocity, $\mathbf{a}$: acceleration, $\hat{\mathbf{n}}$: surface normal, $\boldsymbol{\sigma}$: Cauchy stress, $\mathbf{f}$: mass-density of body force, v : deformed volume, dv : deformed volume element, da : deformed area element.

Example: Hooke's law

Hooke's law, a constitutive relation between stress and strain , can be written using the double contraction : as

$$\boldsymbol{\sigma} = \mathcal{C} : \boldsymbol{\epsilon}$$

$$\sigma_{ij} = \mathcal{C}_{ijk\ell}\epsilon_{kl}.$$

(Here $\mathcal{C}$, a "fourth order tensor," is the *Lagrangian elasticity tensor*, defined as as follows:

$$\frac{\partial \boldsymbol{\sigma}_{\mathrm{PK2}}}{\partial \mathbf{v}} = \underbrace{\frac{\partial \boldsymbol{\sigma}_{\mathrm{PK2}}}{\partial \mathbf{E}}}_{\mathcal{C}} \frac{\partial \mathbf{E}}{\partial \mathbf{v}}$$

$$\mathcal{C} := \frac{\partial \boldsymbol{\sigma}_{\mathrm{PK2}}}{\partial \mathbf{E}}$$

$$\mathcal{C}_{ijk\ell} = \frac{\partial (\boldsymbol{\sigma}_{\mathrm{PK2}})_{ij}}{\partial E_{k\ell}}).$$

First Piola-Kirchhoff stress

The *first Piola-Kirchhoff stress tensor* (linearly) transforms an oriented area in the reference configuration into the force that acts on the coorresponding area in the deformed configuration.

Consider a differential element of a surface force in the deformed configuration $d\mathbf{p}$:

$$
\begin{aligned}
d\mathbf{p} = \mathbf{t}da = \boldsymbol{\sigma}\hat{\mathbf{n}}da &= \boldsymbol{\sigma}d\mathbf{a} \\
&= \boldsymbol{\sigma}(\det(\mathbf{F})\mathbf{F}^{-\top}d\mathbf{A}) \\
&= \underbrace{\det(\mathbf{F})\boldsymbol{\sigma}\mathbf{F}^{-\top}}_{\boldsymbol{\sigma}_{\mathrm{PK1}}}d\mathbf{A} \\
&= \boldsymbol{\sigma}_{\mathrm{PK1}}d\mathbf{A}.
\end{aligned}
$$

(See "Volume and area change" in the appendix to see why the second line is true).

$$\boldsymbol{\sigma}_{\mathrm{PK1}} := \det(\mathbf{F})\boldsymbol{\sigma}\mathbf{F}^{-\top}.$$

The *second Piola-Kirchhoff stress tensor* (linearly) transforms an oriented area in the reference configuration into the force in the reference configuration that acts on that area.

To define this tensor, “pull back” a change in force $\Delta\mathbf{p}$ from the deformed configuration with use of the deformation gradient to obtain a change in force $\Delta\mathbf{P}$ in the reference configuration:

$$\Delta\mathbf{P} = \mathbf{F}^{-1}\Delta\mathbf{p} = \underbrace{\mathbf{F}^{-1}\boldsymbol{\sigma}_{\mathrm{PK1}}}_{\boldsymbol{\sigma}_{\mathrm{PK2}}}\Delta\mathbf{A}.$$

We have defined

$$\boldsymbol{\sigma}_{\mathrm{PK2}} := \mathbf{F}^{-1}\boldsymbol{\sigma}_{\mathrm{PK1}} = \det(\mathbf{F})\mathbf{F}^{-1}\boldsymbol{\sigma}\mathbf{F}^{-T}.$$

Note that $\boldsymbol{\sigma}_{\mathrm{PK2}}$ is symmetric.

In equilibrium, the *principle of virtual work* follows from the momentum balance equation:

$$\Delta W = \underbrace{\int_v \boldsymbol{\sigma} : \Delta \boldsymbol{\ell} dv}_{\Delta W_{\text{int}}} - \int_v \mathbf{f} \cdot \Delta \mathbf{v} dv - \int_{\partial v} \mathbf{t} \cdot \Delta \mathbf{v} d(\partial v).$$

Here, the lowercase v denotes deformed (rather than reference) volume.

It can be shown that

$$\Delta W_{\text{int}} = \int_V \boldsymbol{\sigma}_{\text{PK1}} : \dot{\mathbf{F}} dV,$$

where V is the reference volume. (See “Energetic conjugates” in the appendix for more on this).

Hyperelasticity

Outline of hyperelasticity concepts

- Definitions of elasticity, hyperelasticity
- Formulas for first and second Piola-Kirchhoff stresses
- Isotropic hyperelasticity

In an *elastic* material, the Cauchy stress only depends on the reference configuration **X**,

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}(\mathbf{F}(\mathbf{X}), \mathbf{X}).$$

In a *hyperelastic* material, we also have the "stored strain energy" function

$$\Psi(\mathbf{F}(\mathbf{X}), \mathbf{X}) = \int_{t_0}^{t_f} \boldsymbol{\sigma}_{\mathrm{PK1}} : \dot{\mathbf{F}} dt.$$

Note that $\dot{\Psi} = \boldsymbol{\sigma}_{\mathrm{PK1}} : \dot{\mathbf{F}}$ so

$$\Delta W_{\mathrm{int}} = \int_V \dot{\Psi} dV.$$

This means that $\dot{\Psi}$ is volume-density of internal work, so $\psi = \int_{t_0}^{t_f} \dot{\Psi} dt$ is the volume-density of internal work that accumulates over the deformation. (So the name "stored strain energy" for Ψ is misleading).

Formulas for first and second Piola-Kirchhoff stresses

$$(\boldsymbol{\sigma}_{\mathrm{PK1}})_{ij} = \frac{\partial \Psi(\mathbf{F}(\mathbf{X}), \mathbf{X})}{\partial F_{ij}}$$

$$\boldsymbol{\sigma}_{\mathrm{PK1}} = \frac{\partial \Psi(\mathbf{F}(\mathbf{X}), \mathbf{X})}{\partial \mathbf{F}}$$

$$(\boldsymbol{\sigma}_{\mathrm{PK2}})_{ij} = \frac{\partial \Psi}{\partial E_{ij}} = 2\frac{\partial \Psi}{\partial C_{ij}}$$

$$\boldsymbol{\sigma}_{\mathrm{PK2}} = \frac{\partial \Psi}{\partial \mathbf{E}} = 2\frac{\partial \Psi}{\partial \mathbf{C}}$$

Some notational convenience for hyperelastic materials

Since Ψ is invariant under rotation and since $\mathbf{F} = \mathbf{RU}$, then

$$\Psi(\mathbf{F}(\mathbf{X}), \mathbf{X}) = \Psi(\mathbf{U}(\mathbf{X}), \mathbf{X}) = \Psi(\mathbf{U}^2(\mathbf{X}), \mathbf{X}) = \Psi(\mathbf{F}(\mathbf{X})^\top \mathbf{F}(\mathbf{X}), \mathbf{X}) = \Psi(\mathbf{C}(\mathbf{X}), \mathbf{X}),$$

where $\mathbf{C} := \mathbf{F}^\top \mathbf{F}$.

Isotropic hyperelasticity

In a hyperelastic material that is *isotropic*,

$$\Psi = \Psi(I_{\mathbf{C}}, II_{\mathbf{C}}, III_{\mathbf{C}})$$

$$I_{\mathbf{C}} = \mathrm{tr}(\mathbf{C}), II_{\mathbf{C}} = \mathbf{C} : \mathbf{C}, III_{\mathbf{C}} = \det(\mathbf{C}) = \det(\mathbf{F})^2$$

Computing $\boldsymbol{\sigma}_{\mathrm{PK2}} = 2\frac{\partial \Psi}{\partial \mathbf{C}}$ with the chain rule leads to

$$\boldsymbol{\sigma}_{\mathrm{PK2}} = 2\frac{\partial \Psi}{\partial I_{\mathbf{C}}} + 4\frac{\partial \Psi}{\partial II_{\mathbf{C}}}\mathbf{C} + 2\det(\mathbf{F})^2\frac{\partial \Psi}{\partial III_{\mathbf{C}}}\mathbf{C}^{-1}.$$

Use $\boldsymbol{\sigma} = \det(\mathbf{F})^{-1}\mathbf{F}\boldsymbol{\sigma}_{\mathrm{PK2}}\mathbf{F}^\top$ to get the Cauchy stress as a function of the stored strain energy :

$$\boldsymbol{\sigma} = 2\det(\mathbf{F})^{-1}\frac{\partial \Psi}{\partial I_{\mathbf{C}}}\mathbf{F}\mathbf{F}^\top + 4\det(\mathbf{F})^{-1}\frac{\partial \Psi}{\partial II_{\mathbf{C}}}(\mathbf{F}\mathbf{F}^\top)^2 + 2\det(\mathbf{F})\frac{\partial \Psi}{\partial III_{\mathbf{C}}}.$$

Takeaways

Overleaf

1. (Linear algebra review). A matrix can be interpreted as descriptive of a linear function, or as simply storing data. Linear functions come up when dealing with coordinate systems.
2. (Kinematics). Most concepts in continuum mechanics are either Lagrangian (referring to the reference configuration) or Eulerian (referring to the deformed configuration) in flavor.
3. (Strain). In one dimension, strain is a unitless “stretch ratio.” The Lagrangian strain tensor is useful in solid mechanics, and the Eulerian strain tensor is useful in fluid mechanics.
4. (Stress). Under realistic assumptions, the traction vector linearly depends on the surface normal. The first and second Piola-Kirchhoff stress tensors relate undeformed oriented areas to deformed and undeformed forces, respectively.
5. (Hyperelasticity). An elastic material is one whose stress tensors only depend on the reference configuration. A hyperelastic material is one which has a “stored strain energy” function.

Appendix

Contents of appendix

- In depth review of linear functions and matrices
- $\binom{p}{q}$ tensors
- Volume and area change
- Additional kinematics (required reading for the following two bullet points)
- The small strain tensor
- Energetic conjugates

Appendix

In depth review of linear functions and matrices

Linear functions

A function $T : \mathbb{R}^n \to \mathbb{R}^m$ is said to be *linear* if and only if

$$T(\mathbf{v}_1 + \mathbf{v}_2) = T(\mathbf{v}_1) + T(\mathbf{v}_2)$$
$$T(c\mathbf{v}) = cT(\mathbf{v})$$

for all vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}$ in $\mathbb{R}^n$ and all scalars c in $\mathbb{R}$.

An equivalent (probably more intuitive) condition is that

$$T(v_1\mathbf{e}_1 + ... + v_n\mathbf{e}_n) = v_1T(\mathbf{e}_1) + ... + v_nT(\mathbf{e}_n).$$

This illustrates that a linear function is one that *respects the composition of its vector input*. A weighted sum of vectors goes in, and a weighted sum of transformed vectors comes out, with the weights being the same as before.

We just saw that if $T : \mathbb{R}^n \to \mathbb{R}^m$ is linear, then for $\mathbf{v} = \sum_{i=1}^n v_i \mathbf{e}_i$, we have

$$T(\mathbf{v}) = T(\sum_{i=1}^n v_i \mathbf{e}_i) = \sum_i v_i T(\mathbf{e}_i).$$

This leads us to define a *matrix-vector product* as follows,

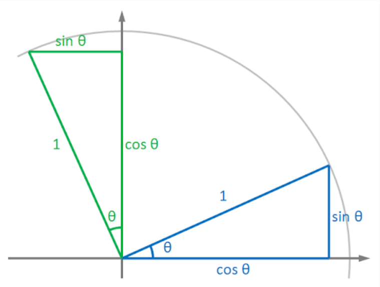
$$\mathbf{A}\mathbf{v} = \begin{pmatrix} \mathbf{A}_{:1} & \dots & \mathbf{A}_{:n} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} := v_1 \mathbf{A}_{:1} + \ldots + v_n \mathbf{A}_{:n},$$

so that

$$T(\mathbf{v}) = \underbrace{\begin{pmatrix} T(\mathbf{e}_1) & \dots & T(\mathbf{e}_n) \end{pmatrix}}_{\mathbf{A}} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \mathbf{A}\mathbf{v}.$$

Thus, every linear function T corresponds to a unique matrix $\mathbf{A}$.

Example: rotation matrices



$$\mathbf{R}_\theta = \begin{pmatrix} R_\theta(\mathbf{e}_1) & R_\theta(\mathbf{e}_2) \end{pmatrix} = \begin{pmatrix} R_\theta \begin{pmatrix} 1 \\ 0 \end{pmatrix} & R_\theta \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{pmatrix}$$
$$= \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

Matrix-matrix multiplication

Suppose T, S are composable linear functions, $T : \mathbb{R}^n \to \mathbb{R}^m, T(\mathbf{v}) = \underbrace{\mathbf{A}}_{m\times n}\mathbf{v}$ and $S : \mathbb{R}^m \to \mathbb{R}^p, S(\mathbf{v}) = \underbrace{\mathbf{B}}_{p\times m}\mathbf{v}$.

Then the matrix of $S \circ T : \mathbb{R}^n \to \mathbb{R}^p$ is the $p \times n$ matrix

$$\begin{aligned} \begin{pmatrix} (S\circ T)(\mathbf{e}_1) & \dots & (S\circ T)(\mathbf{e}_n) \end{pmatrix} &= \begin{pmatrix} S(T(\mathbf{e}_1)) & \dots & S(T(\mathbf{e}_n)) \end{pmatrix} \\ &= \begin{pmatrix} S(\mathbf{A}_{:1}) & \dots & S(\mathbf{A}_{:n}) \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{BA}_{:1} & \dots & \mathbf{BA}_{:n} \end{pmatrix}. \end{aligned}$$

We define the *matrix-matrix* product **BA** to be

$$\mathbf{BA} = \mathbf{B}\begin{pmatrix} \mathbf{A}_{:1} & \dots & \mathbf{A}_{:n} \end{pmatrix} := \begin{pmatrix} \mathbf{BA}_{:1} & \dots & \mathbf{BA}_{:n} \end{pmatrix}$$

so that

$$(S\circ T)(\mathbf{v}) = (\mathbf{BA})\mathbf{v} = \mathbf{B}(\mathbf{A}\mathbf{v}).$$

We defined the matrix-vector product

$$\mathbf{A}\mathbf{v} = \begin{pmatrix} \mathbf{A}_{:1} & \dots & \mathbf{A}_{:n} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} := v_1 \mathbf{A}_{:1} + \dots + v_n \mathbf{A}_{:n} = \sum_{i=1}^{n} v_i \mathbf{a}_i .$$

The *i*th entry of $\mathbf{v}$ is then $v_i = \mathbf{A}_{i:} \cdot \mathbf{v}$. Because

$$\mathbf{BA} = \mathbf{B} \begin{pmatrix} \mathbf{A}_{:1} & \dots & \mathbf{A}_{:n} \end{pmatrix} := \begin{pmatrix} \mathbf{BA}_{:1} & \dots & \mathbf{BA}_{:n} \end{pmatrix},$$

the *ij* entry of **BA** is

$$(\mathbf{BA})_{ij} = (\mathbf{BA}_{:j})_i = \mathbf{B}_{i:} \cdot \mathbf{A}_{:j} = \sum_{k=1}^{n} B_{ik} A_{kj} .$$

Consider a basis $\mathbf{F} = \begin{pmatrix} \mathbf{f}_1 & \dots & \mathbf{f}_n \end{pmatrix}$ of $\mathbb{R}^n$.

If $\mathbf{v} = d_1\mathbf{f}_1 + ... + d_n\mathbf{f}_n$, then we say that $d_1, ..., d_n$ are the *coordinates of* $\mathbf{v}$ *relative to the basis* $\mathbf{F}$.

Since

$$d_1\mathbf{f}_1 + ... + d_n\mathbf{f}_n = \mathbf{F} \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix} = \mathbf{v},$$

we define $[\mathbf{v}]_{\mathbf{F}}$ to be the vector for which $\mathbf{F}[\mathbf{v}]_{\mathbf{F}} = \mathbf{v}$.

$$[\mathbf{v}]_{\mathbf{F}} = \mathbf{F}^{-1}\mathbf{v}.$$

Transformation in a basis

Given a linear function $T : \mathbb{R}^n \to \mathbb{R}^m$, we define $T_{\mathbf{F}} : \mathbb{R}^n \to \mathbb{R}^m$, known as T *with respect to the basis* $\mathbf{F}$:

$$\begin{aligned} T_{\mathbf{F}}([\mathbf{v}]_{\mathbf{F}}) &= [T(\mathbf{v})]_{\mathbf{F}} \\ &= \mathbf{F}^{-1} T(\mathbf{F}[\mathbf{v}]_{\mathbf{F}}). \end{aligned}$$

Therefore if T has matrix $\mathbf{A}$ then $T_{\mathbf{F}}$ has matrix

$$\mathbf{F}^{-1}\mathbf{A}\mathbf{F} = \mathbf{F}^{-1}\left(T(\mathbf{f}_1) \dots T(\mathbf{f}_n)\right) = \begin{pmatrix} [T(\mathbf{f}_1)]_{\mathbf{F}} & \dots & [T(\mathbf{f}_n)]_{\mathbf{F}} \end{pmatrix}.$$

Appendix

$\binom{p}{q}$ tensors

Decomposing a linear function

Earlier, we decomposed **v** using the basis $\{\mathbf{e}_1, \mathbf{e}_2, ..., \mathbf{e}_n\}$ for $\mathbb{R}^n$. Let's see what happens when we similarly decompose a linear function's matrix.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + a_{12}\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + a_{13}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$
$$+ a_{21}\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + a_{22}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + a_{23}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$
$$+ a_{31}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} + a_{32}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} + a_{33}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$\binom{p}{q}$ tensors: overview

This slide provides an overview of the insights that the following slides in this section go over in-depth. Note: you won't be able to understand this terminology until you read the following slides. Still glance over this.

- Two main concepts underlie the idea of a tensor: the idea of a *tensor product space*, and the idea of a *vector space's dual space*. Tensor product spaces are relevant because composition of linear functions (as is seen in the composition $\mathbf{v}\phi$) is itself a multilinear function. Dual spaces are relevant because arbitrary linear functions decompose into basis sums of rank one linear functions, each of which is a composition of a linear function sending a scalar to a vector and an element of a dual space.
    - We can consider of vector space to be to be "linear elements" because they are respected by linear functions. Similarly, we can consider elements of tensor product spacesto be "multilinear elements" because they are respected by multilinear functions.
- We discover $V^* \otimes W \cong \{\text{linear functions} : W^* \to V\}$ by decomposing a linear function into a basis sum of rank one linear functions.
- Linear functions can be represented in a basis with matrices. Tensors (which are the generalization of linear functions) can be represented in a basis as multidimensional matrices.
- In particular, matrices ($\binom{1}{1}$ tensors) are "collections of data" (weighted sums of rank one tensors), descriptive of linear functions, and multilinear elements, all at the same time.

Decomposing a linear function, cont'd

In other words,

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \sum_{i=1}^{3} \sum_{j=1}^{3} a_{ij} \mathbf{e}_i \mathbf{e}_j^\top,$$

since $\mathbf{e}_i \mathbf{e}_j^\top$ is the 3×3 matrix with a 1 in the ij entry and 0's elsewhere.

What happens when we change the basis of the matrices (the $\mathbf{e}_i \mathbf{e}_j^\top$) in this sum, and send $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ to an arbitrary basis $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ for $\mathbb{R}^3$?

We obtain that

$$\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} = \sum_{i=1}^{3} \sum_{j=1}^{3} b_{ij} \mathbf{u}_i \mathbf{u}_j^\top,$$

where (b_{ij}) is the matrix of $\mathbf{A}$ with respect to the basis $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$.

(The ith column of (b_{ij}) is the ith column of (a_{ij}) after it has been converted to the basis $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$. See "transformation in a basis," discussed earlier in this appendix).

Rephrasing the decomposition: get rid of matrices

Any matrix may be obtained as the result of taking some matrix and changing the basis from $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ to $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$. This means that any matrix **A** can be expressed as

$$\mathbf{A} = \sum_{i=1}^{3} \sum_{j=1}^{3} a_{ij} \mathbf{u}_i \mathbf{u}_j^\top .$$

Therefore, if T is the linear function corresponding to **A**, then

$$T = \sum_{i=1}^{3} \sum_{j=1}^{3} a_{ij} \mathbf{u}_i \phi_{\mathbf{u}_j} ,$$

where $\phi_{\mathbf{v}_0} : \mathbb{R}^3 \to \mathbb{R}$ is a linear function defined by $\phi_{\mathbf{v}_0}(\mathbf{v}) = \mathbf{v}_0^\top \mathbf{v}$.

Overleaf

From the last slide,

$$T = \sum_{i=1}^{3} \sum_{j=1}^{3} a_{ij} \mathbf{u}_i \phi_{\mathbf{u}_j}.$$

Note that the dimension of the image of each term in the sum is 1. So, we have shown any linear function is a sum of "rank 1" linear functions. (Warning: the rank of a linear function is not the same thing as the rank of a tensor!)

Each $\mathbf{u}_i \phi_{\mathbf{u}_j}$ can be associated to a symbol of the form $\mathbf{u}_i \otimes \mathbf{u}_j$. *This* $\otimes$ *is just a symbol; we will* derive *the rules it must follow*!

In general, we send $\mathbf{v}\phi_{\mathbf{w}} \mapsto \mathbf{v} \otimes \phi_{\mathbf{w}}$. Let's forget about the $\mathbf{w}$ subscript on $\phi_{\mathbf{w}}$, and say we're sending $\mathbf{v}\phi \mapsto \mathbf{v} \otimes \phi$. Here, $\mathbf{v}$ is in $\mathbb{R}^3$, and ϕ is in $\{\text{linear functions} : \mathbb{R}^3 \to \mathbb{R}\}$, which we denote by $(\mathbb{R}^3)^*$. The space $(\mathbb{R}^3)^*$ is called the *dual vector space* to $\mathbb{R}^3$.

Check

Can you show that $(\mathbb{R}^3)^* := \{\text{linear functions} : \mathbb{R}^3 \to \mathbb{R}\}$ is a vector space under the operations of function addition and scalar multiplication?

The “zero element condition” is not too bad. Next, we need to show that if ϕ_1, ϕ_2 are linear functions $\mathbb{R}^3 \to \mathbb{R}^3$, then $\phi_1 + \phi_2$ is also a linear function $\mathbb{R}^3 \to \mathbb{R}^3$. Lastly, we must show that if $c \in \mathbb{R}$ and ϕ is a linear function $\mathbb{R}^3 \to \mathbb{R}$, then $c\phi$ is also a linear function $\mathbb{R}^3 \to \mathbb{R}$.

We were saying that we want to send $\mathbf{v}\phi \mapsto \mathbf{v} \otimes \phi$.

Note that elements of the form $\mathbf{v}\phi$ respect addition and scalar multiplication in the following way:

$$
\begin{aligned}
(\mathbf{v}_1 + \mathbf{v}_2)\phi &= \mathbf{v}_1\phi + \mathbf{v}_2\phi \\
\mathbf{v}(\phi_1 + \phi_2) &= \mathbf{v}\phi_1 + \mathbf{v}\phi_2 \\
(c\mathbf{v})\phi &= c(\mathbf{v}\phi) = \mathbf{v}(c\phi).
\end{aligned}
$$

In terms of matrices, this is the same as

$$
\begin{aligned}
(\mathbf{v}_1 + \mathbf{v}_2)\mathbf{w}^\top &= \mathbf{v}_1\mathbf{w} + \mathbf{v}\mathbf{w}^\top \\
\mathbf{v}(\mathbf{w}_1 + \mathbf{w}_2)^\top &= \mathbf{v}\mathbf{w}_1^\top + \mathbf{v}\mathbf{w}_2^\top \\
(c\mathbf{v})\mathbf{w}^\top &= c(\mathbf{v}\mathbf{w}^\top) = \mathbf{v}(c\mathbf{w}^\top).
\end{aligned}
$$

If we want the elements of the form $\mathbf{v} \otimes \phi$ to respect addition and scalar multiplication in the same way that elements of the form $\mathbf{v}\phi$ do, then it must be the case that

$$
\begin{aligned}
(\mathbf{v}_1 + \mathbf{v}_2) \otimes \phi &= \mathbf{v}_1 \otimes \phi + \mathbf{v}_2 \otimes \phi \\
\mathbf{v} \otimes (\phi_1 + \phi_2) &= \mathbf{v} \otimes \phi_1 + \mathbf{v} \otimes \phi_2 \\
(c\mathbf{v}) \otimes \phi &= c(\mathbf{v} \otimes \phi) = \mathbf{v} \otimes (c\phi).
\end{aligned}
$$

The isomorphism we've been looking for

The space spanned by elements of the form $\mathbf{v} \otimes \phi$ is called a *tensor product space*; this particular tensor product space is denoted by $\mathbb{R}^3 \otimes (\mathbb{R}^3)^*$. Recall that $(\mathbb{R}^3)^* = \{\text{linear functions} : \mathbb{R}^3 \to \mathbb{R}\}$.

In general, the tensor product space between two vector spaces V and W (over $\mathbb{R}$) is defined as

$$V \otimes W := \{\sum_{ij} a_{ij} \mathbf{v}_i \otimes \mathbf{w}_i \mid \mathbf{v}_i \text{ in } V, \mathbf{w}_i \text{ in } W, a_{ij} \in \mathbb{R}\}.$$

We have seen that

$$\mathbb{R}^3 \otimes (\mathbb{R}^3)^* \cong \{\text{linear functions: } \mathbb{R}^3 \to \mathbb{R}^3\}.$$

The $\cong$ denotes *isomorphism*, and is the formal notation for stating that elements of $\mathbb{R}^3 \otimes (\mathbb{R}^3)^*$ interact with each other in the same way as elements of $\{\text{linear functions: } \mathbb{R}^3 \to \mathbb{R}^3\}$.

The isomorphism we were looking for

We have seen

$$\mathbb{R}^3 \otimes (\mathbb{R}^3)^* \cong \{\text{linear functions: } \mathbb{R}^3 \to \mathbb{R}^3\}.$$

Secretly, we have shown that this isomorphism is obtained by sending a rank 1 linear function (a linear function whose image is 1-dimensional) to an element of $\mathbb{R}^3 \otimes (\mathbb{R}^3)^*$. We actually constructed the definition of a "tensor product space" as we did so in order to *force* an isomorphism to happen.

More generally, for vector spaces V and W, the tensor product space $V \otimes W$ is defined as the space of all symbols of the form $\mathbf{v} \otimes \mathbf{w}$, where $\otimes$ respects addition and scalar multiplication in the required ways for an isomorphism. Then we have

$$V \otimes W^* \cong \{\text{linear functions: } W \to V\}.$$

Definition of a $\binom{p}{q}$ tensor

Without further ado, we define a (p, q) *tensor on* $\mathbb{R}^3$ to be an element of the space $T^p_q(\mathbb{R}^3)$:

$$T^p_q(\mathbb{R}^3) := \underbrace{\mathbb{R}^3 \otimes \ldots \otimes \mathbb{R}^3}_{p \text{ times}} \otimes \underbrace{(\mathbb{R}^3)^* \otimes \ldots \otimes (\mathbb{R}^3)^*}_{q \text{ times}}.$$

With this new definition, the isomorphism we proved is restated as

$$T^1_1(\mathbb{R}^3) \cong \{\text{linear functions: } \mathbb{R}^3 \to \mathbb{R}^3\}.$$

In general,

$$T^1_1(V) \cong \{\text{linear functions: } V \to V\}.$$

Tensor terminology

Let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ be a basis for $\mathbb{R}^3$ and $\{\epsilon_1, \epsilon_2, \epsilon_3\}$ be a basis for $(\mathbb{R}^3)^*$. Then any (p, q) tensor T from T^p_q can be expressed as

$$T = \sum_{i_1 \dots i_p, j_1 \dots j_q \in \{1,2,3\}} T^{i_1 \dots i_p}_{j_1 \dots j_q} \mathbf{e}_{i_1} \otimes \dots \otimes \mathbf{e}_{i_p} \otimes \epsilon_{j_1} \otimes \dots \otimes \epsilon_{j_q}.$$

The scalars $T^{i_1 \dots i_p}_{j_1 \dots j_q}$ are called the *coordinates* of T. Note that a (p, q) tensor on $\mathbb{R}^3$ has $3^p 3^q$ components. Thus a $(1, 1)$ tensor, i.e., a linear function, or matrix, has 9 components.

When associated with its components, a (p, q) tensor on $\mathbb{R}^3$ can be interpreted as a "$(p+q)$-dimensional matrix". For example, a $(1, 1)$ tensor is just a plain old $1 + 1 = 2$ dimensional matrix.

In continuum mechanics, the phrase "second order tensor" means "$(1, 1)$ tensor." (Actually, under certain very common assumptions $(1, 1)$ tensors are isomorphic to $(2, 0)$ and $(0, 2)$ tensors, so there is no need to worry!)

What really are tensor product spaces?

Given vector spaces V, W over $\mathbb{R}$, we defined the tensor product space $V \otimes W$ to be the space spanned by elements of the form $\otimes$, where $\in V$ and $\in W$, such that

$$\begin{aligned} (\mathbf{v}_1 + \mathbf{v}_2) \otimes \mathbf{w} &= \mathbf{v}_1 \otimes \mathbf{w} + \mathbf{v}_2 \otimes \mathbf{w} \\ \mathbf{v} \otimes (\mathbf{w}_1 + \mathbf{w}_2) &= \mathbf{v} \otimes \mathbf{w}_1 + \mathbf{v} \otimes \mathbf{w}_2 \\ (c\mathbf{v}) \otimes \mathbf{w} &= c(\mathbf{v} \otimes \mathbf{w}) = \mathbf{v} \otimes (c\mathbf{w}), \end{aligned}$$

for all $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v} \in V$, $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w} \in W$, and $c \in \mathbb{R}$.

In other words, $V \otimes W$ is the space spanned by elements of the form $\otimes$, where $\otimes$ *appears* to be a *multilinear function*. (A multilinear function is a function of that is linear in each argument).

In this sense, just as vectors are “linear elements” that are respected by linear functions, elements of tensor product spaces are “multilinear elements” that are respected by multilinear functions, in the following sense...

Universal property of tensor product spaces

Here we show that every multilinear function coorresponds to a linear function on a tensor product space. In this sense, elements of tensor product spaces are respected by multilinear functions.

For simplicity, consider a bilinear (rather than k-linear) function $f : V \times W \to Y$. Then applying f to a tuple $(\mathbf{v}, \mathbf{w}) \in V \times W$ is the same as "converting" this tuple to $\mathbf{v} \otimes \mathbf{w} \in V \otimes W$, and then plugging $\mathbf{v} \otimes \mathbf{w}$ into a linear function $h : V \otimes W \to Y$ defined on basis elements of $V \otimes W$ by $h(\mathbf{e}_i \otimes \mathbf{f}_j) = f(\mathbf{e}_i, \mathbf{f}_j)$. (Note: since we require h to be linear, it is enough to specify how it acts on basis elements).

$$(\mathbf{v}, \mathbf{w}) \mapsto \mathbf{v} \otimes \mathbf{w} \overset{h}{\mapsto} h(\mathbf{v}, \mathbf{w})$$
$$h(\mathbf{e}_i \otimes \mathbf{f}_j) = f(\mathbf{e}_i, \mathbf{f}_j)$$

Tensor product and double contraction revisited

In continuum mechanics, the "tensor product" of two vectors is defined as

$$\mathbf{v} \otimes \mathbf{w} := \mathbf{v}\mathbf{w}^\top .$$

This should really be called the *outer product* of **v** and **w**. We saw earlier that this arises from decomposing a matrix into a sum of rank 1 linear functions.

The *double contraction* of two $(2,2)$ tensors $T = \sum_{ij} T^{i_1 i_2}_{j_1 j_2} \mathbf{e}_{i_1} \otimes \mathbf{e}_{i_2} \otimes \phi_{j_1} \otimes \phi_{j_2}$ and $S = \sum_{ij} S^{i_1 i_2}_{j_1 j_2} \mathbf{e}_{i_1} \otimes \mathbf{e}_{i_2} \otimes \phi_{j_1} \otimes \phi_{j_2}$ is the following $(2,2)$ tensor:

$$T : S = \sum_{j_1, j_2 \in \{1,2,3\}} T^{i_1 i_2}_{j_1 j_2} S^{j_1 j_2}_{k_1 k_2} \mathbf{e}_{i_1} \otimes \mathbf{e}_{i_2} \otimes \phi_{k_1} \otimes \phi_{k_2} .$$

That is, $C = T : S$ is the $(2,2)$ tensor with components

$$C^{i_1 i_2}_{k_1 k_2} = T^{i_1 i_2}_{j_1 j_2} S^{j_1 j_2}_{k_1 k_2} .$$

Appendix

Area and volume change

$$d\mathbf{a} = \det(\mathbf{F})\mathbf{F}^{-\top}d\mathbf{A}.$$

Proof.

$$\begin{aligned}
dv &= \det(\mathbf{F})dV \\
d\mathbf{x}\cdot d\mathbf{a} &= \det(\mathbf{F})d\mathbf{X}\cdot d\mathbf{A} \\
(\mathbf{F}d\mathbf{X})\cdot d\mathbf{a} &= \det(\mathbf{F})d\mathbf{X}\cdot d\mathbf{A} \\
(d\mathbf{X})^\top\mathbf{F}^\top d\mathbf{a} &= \det(\mathbf{F})d\mathbf{X}\cdot d\mathbf{A} \\
d\mathbf{X}\cdot\mathbf{F}^\top d\mathbf{a} &= d\mathbf{X}\cdot(\det(\mathbf{F})d\mathbf{A}) \\
\mathbf{F}^\top d\mathbf{a} &= \det(\mathbf{F})d\mathbf{A} \\
d\mathbf{a} &= \det(\mathbf{F})\mathbf{F}^{-\top}d\mathbf{A}.
\end{aligned}$$

Volume change

$$dv = \det(\mathbf{F})dV$$
$$\rho = \det(\mathbf{F})\rho_0 \text{ when mass is time-independent.}$$

Proof of first line:

$$\begin{aligned} dv &= \det(d\mathbf{x}_1, d\mathbf{x}_2, d\mathbf{x}_3) \\ &= \det(\mathbf{F}d\mathbf{X}_1, \mathbf{F}d\mathbf{X}_2, \mathbf{F}d\mathbf{X}_3) \\ &= \det(\mathbf{F}(d\mathbf{X}_1, d\mathbf{X}_2, d\mathbf{X}_3)) \\ &= \det(\mathbf{F})\det(d\mathbf{X}_1, d\mathbf{X}_2, d\mathbf{X}_3) \\ &= \det(\mathbf{F})dV. \end{aligned}$$

Proof of second line. When mass is time-independent, we have $dm = \rho_0 dV = \rho_0 \det(\mathbf{F})dv$ and $dm = \rho dv$. Therefore $\rho_0 \det(\mathbf{F})dv = \rho dv$, so $\rho = \det(\mathbf{F})\rho_0$.

Appendix

Additional kinematics

(required reading for the following two sections)

Deformation gradient in terms of displacement field

The *displacement field* **u** is

$$\mathbf{u} = \mathbf{x} - \mathbf{X}.$$

We will show that

$$\mathbf{F} = \frac{\partial \mathbf{u}}{\partial \mathbf{X}} + \mathbf{I}.$$

Proof: since $= \mathbf{x} - \mathbf{X}$, then $\mathbf{F} = \frac{\partial}{\partial \mathbf{X}}\mathbf{x} = \frac{\partial}{\partial \mathbf{X}}(\mathbf{X} + \mathbf{u}) = \mathbf{I} + \frac{\partial \mathbf{u}}{\partial \mathbf{X}}$.

Eulerian velocity gradient

$$\dot{\mathbf{F}} = \frac{d\mathbf{F}}{dt} = \underbrace{\frac{d}{dt}\frac{\partial \mathbf{x}}{\partial \mathbf{X}} = \frac{\partial}{\partial \mathbf{X}}\frac{d\mathbf{x}}{dt}}_{\text{Examine } ij \text{ entry of each}} = \frac{\partial \mathbf{v}}{\partial \mathbf{X}} = \frac{\partial \mathbf{v}}{\partial \mathbf{x}}\frac{\partial \mathbf{x}}{\partial \mathbf{X}} = \boldsymbol{\ell}\mathbf{F}.$$

The *ij* entry of $\frac{d}{dt}\frac{\partial \mathbf{x}}{\partial \mathbf{X}}$ is $\frac{d}{dt}\frac{\partial x_i}{\partial X_j} = \frac{\partial}{\partial X_j}\frac{dx_i}{dt}$. (Commuting $\frac{d}{dt}$ with $\frac{\partial}{\partial x_i}$ here is valid).

Here, we have defined the *Eulerian velocity gradient*

$$\boldsymbol{\ell} := \frac{\partial \mathbf{v}}{\partial \mathbf{x}}$$

and shown that

$$\ell = \dot{\mathbf{F}}\mathbf{F}^{-1}.$$

Appendix
The small strain tensor

Deriving ϵ from E and e

We can derive a *small strain tensor* ϵ by considering simultaneous engineering normal and shear strains in the case of small strains, i.e. when quadratic terms are neglected (see `https://en.wikipedia.org/wiki/Deformation_(physics)`).

Use $\mathbf{F} = \frac{\partial \mathbf{u}}{\partial \mathbf{X}} + \mathbf{I}$ with

$$\mathbf{E} = \frac{1}{2}(\mathbf{F}^\top \mathbf{F} - \mathbf{I})$$
$$\mathbf{e} = \frac{1}{2}(\mathbf{I} - (\mathbf{F}\mathbf{F}^\top)^{-1})$$

to show... (next slide)

Deriving ϵ from E and e, cont'd

$$\mathbf{E} = \frac{1}{2}\left(\frac{\partial \mathbf{u}}{\partial \mathbf{X}} + \left(\frac{\partial \mathbf{u}}{\partial \mathbf{X}}\right)^\top + \left(\frac{\partial \mathbf{u}}{\partial \mathbf{X}}\right)^\top \frac{\partial \mathbf{u}}{\partial \mathbf{X}}\right) = \epsilon + \frac{1}{2}\left(\frac{\partial \mathbf{u}}{\partial \mathbf{X}}\right)^\top \frac{\partial \mathbf{u}}{\partial \mathbf{X}}$$

$$\mathbf{e} = \frac{1}{2}\left(\frac{\partial \mathbf{u}}{\partial \mathbf{X}} + \left(\frac{\partial \mathbf{u}}{\partial \mathbf{X}}\right)^\top - \frac{\partial \mathbf{u}}{\partial \mathbf{X}}\left(\frac{\partial \mathbf{u}}{\partial \mathbf{X}}\right)^\top\right) = \epsilon - \frac{1}{2}\frac{\partial \mathbf{u}}{\partial \mathbf{X}}\left(\frac{\partial \mathbf{u}}{\partial \mathbf{X}}\right)^\top.$$

When the entries of $\frac{\partial \mathbf{u}}{\partial \mathbf{X}}$, which are quadratic, are small enough so that the product terms can be neglected, then

$$\mathbf{E} \approx \mathbf{e} \approx \epsilon,$$

where we've defined

$$\epsilon := \frac{1}{2}\left(\frac{\partial \mathbf{u}}{\partial \mathbf{X}} + \left(\frac{\partial \mathbf{u}}{\partial \mathbf{X}}\right)^\top\right).$$

A simpler formula for the small strain tensor

We showed that

$$\boldsymbol{\epsilon} := \frac{1}{2}\left(\frac{\partial \mathbf{u}}{\partial \mathbf{X}} + \left(\frac{\partial \mathbf{u}}{\partial \mathbf{X}}\right)^{\top}\right),$$

where $\mathbf{u} = \mathbf{x} - \mathbf{X}$.

Note that since $\mathbf{F} = \frac{\partial \mathbf{u}}{\partial \mathbf{X}} + \mathbf{I} \iff \frac{\partial \mathbf{u}}{\partial \mathbf{X}} = \mathbf{F} - \mathbf{I}$, then it follows that

$$\boldsymbol{\epsilon} = \frac{1}{2}(\mathbf{F} + \mathbf{F}^{\top}) - \mathbf{I}.$$

ϵ approximates a rotation-independent tensor

It's also worth noting that when *rotations* are small, the small strain tensor is approximately equal to the following *rotation*-independent matrix:

$$\epsilon \approx \mathbf{U} - \mathbf{I}.$$

So the small strain tensor can be thought of as approximating $\mathbf{U} - \mathbf{I}$!

Proof. We show $\mathbf{U} \approx \frac{1}{2}(\mathbf{F} + \mathbf{F}^\top)$ when rotations are small. If rotations are small, then $\mathbf{R} \approx \mathbf{I}$ so $\mathbf{F} = \mathbf{RU} \approx \mathbf{U}$. Using the symmetry of $\mathbf{U}$, we have $\frac{1}{2}(\mathbf{F} + \mathbf{F}^\top) \approx \frac{1}{2}(\mathbf{U} + \mathbf{U}^\top) = \mathbf{U}$.

Appendix

Energetic conjugates

Stored internal energy

Principle of virtual work:

$$\Delta W = \underbrace{\int_v \boldsymbol{\sigma} : \Delta \boldsymbol{\ell} dv}_{\Delta W_{\text{int}}} - \int_v \mathbf{f} \cdot \Delta \mathbf{v} dv - \int_{\partial v} \mathbf{t} \cdot \Delta \mathbf{v} d(\partial v).$$

We say that linear functions **P** and **Q** are *energetic conjugates* in the volume U if and only if

$$\Delta W_{\text{int}} = \int_U \mathbf{P} : \Delta \mathbf{Q} dU.$$

So $\boldsymbol{\sigma}$ and $\boldsymbol{\ell}$ are energetic conjugates in the deformed volume v.

Internal work in the reference volume

Focus on ΔW_{int}:

$$\Delta W_{\text{int}.} = \int_v \boldsymbol{\sigma} : \Delta \ell dv.$$

Use $dv = \det(\mathbf{F}) dV$ to transform the above equation from a statement about the deformed volume v into a statement about the reference volume V:

$$\Delta W_{\text{int}} = \int_V \det(\mathbf{F}) \boldsymbol{\sigma} : \Delta \ell dV.$$

We see $\det(\mathbf{F})\boldsymbol{\sigma}$ is energetically conjugate to $\Delta \ell$ in the reference volume V.

Energetic conjugate to $\boldsymbol{\sigma}_{\text{PK1}}$

Overleaf

Use $\mathbf{A} : \mathbf{B} = \operatorname{tr}(\mathbf{A}^\top \mathbf{B})$.

$$\begin{aligned}
\Delta W_{\text{int}} &= \int_V \det(\mathbf{F})\boldsymbol{\sigma} : \Delta\boldsymbol{\ell} dV = \int_V \det(\mathbf{F})\boldsymbol{\sigma} : \dot{\mathbf{F}}\mathbf{F}^{-1} dV \\
&= \int_V \operatorname{tr}\left((\det(\mathbf{F})\boldsymbol{\sigma})^\top \dot{\mathbf{F}}\mathbf{F}^{-1} \right) dV \\
&= \int_V \det(\mathbf{F}) \operatorname{tr}\left(\boldsymbol{\sigma}\dot{\mathbf{F}}\mathbf{F}^{-1} \right) dV \\
&= \int_V \det(\mathbf{F}) \operatorname{tr}\left(\mathbf{F}^{-1}\boldsymbol{\sigma}\dot{\mathbf{F}} \right) dV \\
&= \int_V \underbrace{\det(\mathbf{F})\boldsymbol{\sigma}\mathbf{F}^{-\top}}_{\boldsymbol{\sigma}_{\text{PK1}}} : \dot{\mathbf{F}} dV.
\end{aligned}$$

$$\boldsymbol{\sigma}_{\text{PK1}} := \det(\mathbf{F})\boldsymbol{\sigma}\mathbf{F}^{-\top}$$

$\boldsymbol{\sigma}_{\text{PK1}}$ is energetically conjugate with $\dot{\mathbf{F}}$ in the reference volume V.

Rate of deformation tensor

(Content of this slide is required to obtain that $\boldsymbol{\sigma}_{\text{PK2}}$ and $\dot{\mathbf{E}}$ are energetic conjugates in the reference volume).

The Eulerian velocity gradient $\boldsymbol{\ell}$ can be expressed as $\boldsymbol{\ell} = \mathbf{d} + \mathbf{w}$, where $\mathbf{d}$ is the symmetric matrix $\mathbf{d} := \frac{1}{2}(\boldsymbol{\ell} + \boldsymbol{\ell}^\top)$ and $\mathbf{w}$ is the skew-symmetric matrix $\mathbf{w} := \frac{1}{2}(\boldsymbol{\ell} - \boldsymbol{\ell}^\top)$. $\mathbf{d}$ is an Eulerian strain rate tensor, and is called the *rate of deformation tensor* in literature.

We use $\boldsymbol{\ell} = \dot{\mathbf{F}}\mathbf{F}^{-1}$ with the definition of Lagrangian strain to derive a relationship between $\dot{\mathbf{E}}$ and $\mathbf{d}$: time-differentiate the equation $\mathbf{E} = \frac{1}{2}(\mathbf{F}^\top\mathbf{F} - \mathbf{I})$ and then use $\dot{\mathbf{F}} = \boldsymbol{\ell}\mathbf{F}$ to obtain $\dot{\mathbf{E}} = \mathbf{F}^\top\mathbf{d}\mathbf{F} \iff \mathbf{d} = \mathbf{F}^{-\top}\dot{\mathbf{E}}\mathbf{F}^{-1}$. We say that $\mathbf{d}$ is the “push-forward” of $\dot{\mathbf{E}}$ and that $\dot{\mathbf{E}}$ is the “pull-back” of $\mathbf{d}$.

Energetic conjugate to σ_{PK2}

We derived $\boldsymbol{\sigma}_{\mathrm{PK1}}$ using the fact $\Delta\boldsymbol{\ell} = \dot{\mathbf{F}}\mathbf{F}$. To show that $\boldsymbol{\sigma}_{\mathrm{PK2}}$ is energetically conjugate to $\dot{\mathbf{E}}$ in the reference volume, we analogously use two facts: firstly, that $\boldsymbol{\sigma} : \Delta\boldsymbol{\ell} = \boldsymbol{\sigma} : \Delta\mathbf{d}$ where $\Delta\boldsymbol{\ell} = \Delta\mathbf{d} + \Delta\mathbf{w}$, and secondly, that $\dot{\mathbf{E}}$ is the "push-forward" of $\mathbf{d}$.

Again, use $\mathbf{A} : \mathbf{B} = \mathrm{tr}(\mathbf{A}^\top\mathbf{B})$.

$$
\begin{aligned}
\Delta W_{\mathrm{int}} &= \int_v \boldsymbol{\sigma} : \Delta\boldsymbol{\ell}\, dv = \int_v \boldsymbol{\sigma} : \Delta\mathbf{d}\, dv = \int_V \det(\mathbf{F})\boldsymbol{\sigma} : \Delta\mathbf{d}\, dV \\
&= \int_V \det(\mathbf{F})\boldsymbol{\sigma} : \mathbf{F}^{-\top}\Delta\dot{\mathbf{E}}\mathbf{F}^{-1} = \int_V \det(\mathbf{F})\mathrm{tr}(\boldsymbol{\sigma}^\top\mathbf{F}^{-\top}\Delta\dot{\mathbf{E}}\mathbf{F}^{-1}) \\
&= \int_V \det(\mathbf{F})\mathrm{tr}(\boldsymbol{\sigma}\mathbf{F}^{-\top}\Delta\dot{\mathbf{E}}\mathbf{F}^{-1}) = \int_V \det(\mathbf{F})\mathrm{tr}(\mathbf{F}^{-1}\boldsymbol{\sigma}\mathbf{F}^{-\top}\Delta\dot{\mathbf{E}}) \\
&= \int_V \mathrm{tr}(\boldsymbol{\sigma}_{\mathrm{PK2}}\Delta\dot{\mathbf{E}}) = \int_V \boldsymbol{\sigma}_{\mathrm{PK2}} : \Delta\dot{\mathbf{E}}.
\end{aligned}
$$

List of energetic conjugates

Reminder: **P** and **Q** are said to be *energetic conjugates* in the volume U if and only if

$$\Delta W_{\text{int}} = \int_U \mathbf{P} : \Delta \mathbf{Q} dU.$$

List of energetic conjugates:

- $\boldsymbol{\sigma}$ and $\{\mathbf{d} \text{ or } \boldsymbol{\ell}\}$, in the deformed volume
- $\det(\mathbf{F})\boldsymbol{\sigma}$ and $\{\mathbf{d} \text{ or } \boldsymbol{\ell}\}$, in the reference volume
- $\boldsymbol{\sigma}_{\text{PK1}}$ and $\dot{\mathbf{F}}$, in the reference volume
- $\boldsymbol{\sigma}_{\text{PK2}}$ and $\dot{\mathbf{E}}$, in the reference volume