Charles University

Faculty of Mathematics and Physics

Thermodynamics and mechanics of solids

Based on the lectures by prof. RNDr. Martin Kružík, Ph.D., DSc.

Compiled and typeset by Kamil Belán

Academic Year 2024/2025

Last updated: June 20, 2025

Contents

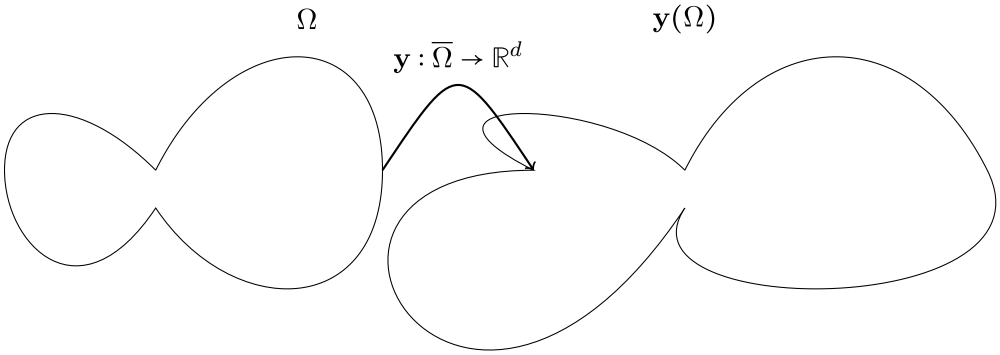
1 TODO

- include missing lecture about potential forces
- include missing lecture about rank one convexity

2 Geometry

2.1 Deformation

Suppose we are given an abstract body $\Omega \subset \mathbb{R}^d, d = 2, 3$. Choosing a particular state, we denote it the **reference configuration**. After the acting of forces, the body deforms into the **current, deformed configuration**.



The mapping that produces the deformation is called the **deformation**, denoted $\mathbf{y}$, i.e.

$$\mathbf{y} : \overline{\Omega} \to \mathbb{R}^d.$$

Of large interest will be the **deformation gradient**

$$\mathbb{F}(\mathbf{x}) = \nabla y(\mathbf{x}), (\nabla y)_{ij} = \frac{\partial y^i}{\partial x^j},$$

on which we put some physically sound restrictions, such as

$$\det \mathbb{F} > 0.$$

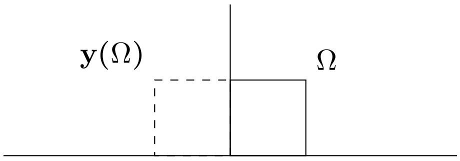
This means in particular that the determinant is nonzero, but also that the deformation preserves the orientation of bases:

$$(\mathbf{e}_1 \times \mathbf{e}_2) \cdot \mathbf{e}_3 > 0 \Rightarrow (\mathbb{F}\mathbf{e}_1 \times \mathbb{F}\mathbf{e}_2) \cdot \mathbb{F}\mathbf{e}_3 > 0.$$

Example. Suppose the deformation is given as

$$\mathbf{y}(\mathbf{x}) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{x},$$

i.e., $\mathbb{F} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$, $\det \mathbb{F} = -1$. This is an example of a deformation that is *forbidden*.



Imagine it is a sheet of paper in a plane - you cannot reflect it without lifting it from the plane.

2.2 Displacement

Another useful way of describing the deformation is by using the **displacement vector u**:

$$\mathbf{u}(\mathbf{x}) = \mathbf{y}(\mathbf{x}) - \mathbf{x},$$

so taking the gradient gives

$$\nabla \mathbf{u}(\mathbf{x}) = \mathbb{F}(\mathbf{x}) - \mathbb{I},$$

or in other words

$$\mathbb{F}(\mathbf{x}) = \mathbb{I} + \nabla \mathbf{u}(\mathbf{x}).$$

Remark. It is interesting that we do not place any restrictions on the determinant of the displacement gradient.

2.3 Changes of measures

We need to examine how the volume, area and lengths change under deformation. In what follows, for a set $\omega \subset \mathbb{R}^d$ in the reference configuration we denote $\omega^y \subset \mathbb{R}^d$ to be the deformed body in the current configuration, i.e.

$$\omega^y = \mathbf{y}(\omega).$$

2.3.1 Change of volume

Using the change of variable theorem we obtain (realize $\det \mathbb{F} > 0$)

$$\lambda(\omega^y) = \int_{\omega^y} 1 \,\mathrm{d}\mathbf{x}^y = \int_{\omega} \det \mathbb{F}(\mathbf{x}) \,\mathrm{d}\mathbf{x}\,,$$

so we write $\mathrm{d}\mathbf{x}^y = \det \mathbb{F}\,\mathrm{d}\mathbf{x}$. This motivates "our" definition of the determinant of the deformation gradient:[1]

$$\det \mathbb{F}(\mathbf{x}) = \lim_{r \to 0^+} \frac{\lambda(\mathbf{y}(B(\mathbf{x}, r))}{\lambda(B(\mathbf{x}, r))}, \tag{1}$$

where $B(\mathbf{x}, r)$ is a (closed) ball centered at $\mathbf{x}$ of radius r.

[1] This is in fact just the Lebesgue differentiation theorem.

2.3.2 Change of lengths

Suppose the line segment $\mathbf{x} + \Delta\mathbf{x}$ undergoes deformation. How does its length change? Taylor expansion yields:

$$\mathbf{y}(\mathbf{x} + \Delta\mathbf{x}) = \mathbf{y}(\mathbf{x}) + \mathbb{F}(\mathbf{x})\Delta\mathbf{x} + \text{h.o.t.},$$

where h.o.t. stands for higher order terms. Using the expansion, we can write

$$|\mathbf{y}(\mathbf{x} + \Delta\mathbf{x}) - \mathbf{y}(\mathbf{x})|^2 = (\Delta\mathbf{x})^\top \mathbb{F}^\top \mathbb{F} \Delta\mathbf{x} = (\Delta\mathbf{x})^\top \mathbb{C}(\mathbf{x}) \Delta\mathbf{x},$$

where we have denoted

$$\mathbb{C}(\mathbf{x}) = \mathbb{F}^\top(\mathbf{x})\mathbb{F}(\mathbf{x}),$$

as the **Right Cauchy Green tensor**. Realize that in fact

$$\mathbb{C} : \overline{\Omega} \to \overline{\Omega}, \mathbb{C} : \mathbf{x} \mapsto \mathbb{F}^\top(\mathbb{F}\mathbf{x}),$$

and recall that $\mathbb{C}$ is in fact the metric tensor on $\mathbf{y}(\omega)$ (for admissable $\mathbf{y}$.)

Example. Let the deformation $\mathbf{y}$ be given as $\mathbf{y}(\mathbf{x}) = \mathbb{R}\mathbf{x} + \mathbf{v}, \mathbf{v} \in \mathbb{R}^d, \mathbb{R} \in \mathrm{SO(d)}$, where [2]

$$\mathrm{SO(d)} = \left\{\mathbb{A} \in \mathbb{R}^{d\times d}, \mathbb{A}^\top\mathbb{A} = \mathbb{A}\mathbb{A}^\top = \mathbb{I}, \det\mathbb{A} = 1 \det\mathbb{A} > 0\right\}.$$

Then $\mathbb{F} = \mathbb{R}$, $\mathbb{C} = \mathbb{I}$.

2.3.3 Change of surfaces

For $\mathbb{A} \in \mathbb{R}^{d\times d}$ regular we define its **cofactor matrix** $\operatorname{cof}\mathbb{A}$ as

$$\operatorname{cof}\mathbb{A} = (\det\mathbb{A})\mathbb{A}^{-\top},$$

which is an interesting quantity whatsoever; we will be quite often using the following theorem

Theorem 1 (Piola's identity)**.** *Let* $\mathbf{y} \in C^2(\Omega; \mathbb{R}^d)$*, then* $\forall \mathbf{x} \in \Omega$:

$$\nabla \cdot (\operatorname{cof}\nabla\mathbf{y}(\mathbf{x})) = \mathbf{0}.$$

For a regular matrix $\mathbb{A}$, we also have the identity

$$\mathbb{A}^{-1} = \frac{1}{\det\mathbb{A}}(\operatorname{cof}\mathbb{A})^\top, \tag{2}$$

What about the determinant of the cofactor? Clearly

$$\det\operatorname{cof}\mathbb{A} = (\det\mathbb{A})^d \det\mathbb{A}^{-\top} = (\det\mathbb{A})^{d-1},$$

so we can also express equation 2 in a different way

$$\mathbb{A}^{-1} = \frac{(\operatorname{cof}\mathbb{A})^\top}{(\det\operatorname{cof}\mathbb{A})^{\frac{1}{d-1}}}. \tag{3}$$

From geometry, recall the change of variables for surface integration:

$$\int_{\partial\omega^y} \mathbf{n}^y(\mathbf{x}^y)\,\mathrm{d}S^y = \int_{\partial\omega} \operatorname{cof}\mathbb{F}(\mathbf{x})\mathbf{n}(\mathbf{x})\,\mathrm{d}S,$$

[2] From the fact $\mathbb{A}$ is orthogonal automatically follows $\det\mathbb{A} = \pm 1$.

where $\mathbf{n}^y(\mathbf{x}^y)$ is the outward unit normal to the deformed boundary ω^y at the point $\mathbf{x}^y \in \omega^y$. Informally, we write $\mathbf{n}^y \, \mathrm{d}S^y = \operatorname{cof} \mathbb{F} \mathbf{n} \, \mathrm{d}S$. We can also explicitely express the normal to the deformed boundary as

$$\mathbf{n}^y(\mathbf{x}^y) = \frac{\operatorname{cof} \mathbb{F}(\mathbf{x})\mathbf{n}(\mathbf{x})}{|\operatorname{cof} \mathbb{F}(\mathbf{x})\mathbf{n}(\mathbf{x})|}, \mathbf{x} \in \partial\omega, \mathbf{y}(\mathbf{x}) \in \partial\omega^y. \tag{4}$$

Notice we are silently assuming

$$\mathbf{y}(\partial\omega) = \partial\mathbf{y}(\omega) = \partial\omega^y.$$

2.4 Affine transformations

An example of a deformation is the so called **affine transformation**.

Example. Suppose the folowing deformation:

$$\mathbf{y}(\mathbf{x}) = \mathbb{A}\mathbf{x} + \mathbf{v}, \mathbb{A} \in \mathbb{R}^{d \times d}, \mathbf{v} \in \mathbb{R}^d, \det \mathbb{A} > 0.$$

Clearly then $\mathbb{F}(\mathbf{x}) = \mathbb{A}$.

It is crucial to realize how $\mathbb{F}, \mathbb{F}^\top, \mathbb{F}^{-1}\mathbb{F}^{-\top}$ work.

- $\mathbb{F}$ takes a vector $\mathbf{x} - \mathbf{0}$ from the *reference configuration* and maps it to the vector $\mathbb{F}\mathbf{x} - \mathbb{F}\mathbf{0}$ in the *current configuration*
- $\mathbb{F}^{-1}$ takes the vector $\mathbb{F}\mathbf{x} - \mathbb{F}\mathbf{0}$ from the *current configuration* and maps it to the vector $\mathbf{x} - \mathbf{0}$ from the *reference configuration*
- $\mathbb{F}^\top$ is defined through: $\mathbb{F}\mathbf{u} \cdot \mathbf{w} = \mathbf{u} \cdot \mathbb{F}^\top \mathbf{w}$, and since $\mathbb{F}$ is defined on the reference configuration, $\mathbb{F}^\top$ must take something from the *current configuration* and return something from the *reference configuration.*
- $\mathbb{F}^{-\top}$ consequently takes something from the *reference configuration* and maps it to something from the *current configuration.*

What when $\mathbb{C} = \mathbb{I}$? Can we say something about $\mathbb{F}$? Write $\mathbb{C} = \mathbb{F}^\top\mathbb{F} = \mathbb{I}$, so $\mathbb{F}^\top = \mathbb{F}^{-1}, \det \mathbb{F} > 0$, meaning $\mathbb{F} \in \mathrm{SO(d)}$,*i.e.*, $\mathbb{F}(\mathbf{x}) = \mathbb{R}(\mathbf{x}), \mathbf{x} \in \Omega$, where $\mathbb{R}$ is a rotation tensor. Investigate the cofactor of the deformation gradient:

$$\operatorname{cof} \mathbb{F} = \det \mathbb{F}\mathbb{F}^{-\top} = \operatorname{cof} \mathbb{R} = 1\mathbb{R}(\mathbf{x}) = \mathbb{F}(\mathbf{x}),$$

and so $\operatorname{cof} \mathbb{F} = \mathbb{F}$. Recall Piola's identity:

$$\mathbf{0} = \nabla \cdot \operatorname{cof} \mathbb{F} = \nabla \cdot \mathbb{F}(\mathbf{x}) = \nabla \cdot \nabla\mathbf{y}(\mathbf{x}) = \triangle\, \mathbf{y}(\mathbf{x}).$$

We have the identity:

$$\frac{1}{2} \triangle |\nabla\mathbf{y}|^2 = |\nabla\nabla\mathbf{y}|^2 + \nabla\mathbf{y} : \nabla \triangle \mathbf{y},$$

but since $\triangle\, \mathbf{y} = 0$ from the above and $\triangle |\nabla\mathbf{y}|^2 = \triangle \operatorname{tr}(\mathbb{F}^\top\mathbb{F}) = \triangle \operatorname{tr}(\mathbb{I}) = 0$, we also have $|\nabla\nabla\mathbf{y}|^2 = 0$, meaning the deformation must have have the form

$$\mathbf{y}(\mathbf{x}) = \mathbb{R}\mathbf{x} + \mathbf{v} \,\text{locally}\,,$$

meaning that $\mathbf{y}$ is only piecewise affine. We will show that it must however be globally affine. Let $\mathbf{y}$ be piecewise affine. Since $\mathbf{y}$ is continous on the whole $\overline{\Omega}$, it must be continuous across the faces of the partition and in particular

$$\mathbb{R}_1\mathbf{x} + \mathbf{v}_1 = \mathbb{R}_2\mathbf{x} + \mathbf{v}_2,$$

with $\mathbb{R}_1, \mathbb{R}_2 \in \mathrm{SO(d)}$ being rotations, $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^d$ some constant vectors and $\mathbf{x} \in \{\mathbf{x}\cdot\mathbf{n} = c\}$ is a vector from the interface. Denoting $\mathbf{n}, \mathbf{t}$ to be the normal and tangential vector to the interface, one has (realize $\mathbf{x} = (\mathbf{x}\cdot\mathbf{n})\mathbf{n} + (\mathbf{x}\cdot\mathbf{t})\mathbf{t}$)

$$(\mathbb{R}_1 - \mathbb{R}_2)\mathbf{n}\mathbf{x}\cdot\mathbf{n} + (\mathbb{R}_1 - \mathbb{R}_2)\mathbf{t}\mathbf{x}\cdot\mathbf{t} = \mathbf{v}_2 - \mathbf{v}_1, \forall \mathbf{x} \in \{\mathbf{x}\cdot\mathbf{n} = c\},$$
$$c(\mathbb{R}_1 - \mathbb{R}_2)\mathbf{n} + (\mathbb{R}_1 - \mathbb{R}_2)\mathbf{t} = \mathbf{v}_2 - \mathbf{v}_1.$$

Looking on the LHS, we see that the first term in constant on the hyperplane, but the second one is not - the tangential vector depends on the position. However, the RHS is constant, which must mean

$$(\mathbb{R}_1 - \mathbb{R}_2)\mathbf{t} = 0,$$

or

$$\mathrm{Ker}\,(\mathbb{R}_1 - \mathbb{R}_2) = \mathrm{span}\,\{\mathbf{t}\} = \mathbf{n}^\perp.$$

But if that is true, all nontrivial bussiness is happening only for vectors that are perpendicular to $\mathbf{t}$,*i.e.*, that are in the direction of $\mathbf{n}$, so it must hold

$$\mathbb{R}_1 - \mathbb{R}_2 = (\mathbb{R}_1 - \mathbb{R}_2)(\mathbf{n} \otimes \mathbf{n}).$$

But that must mean $\mathbb{R}_1 = \mathbb{R}_2$: we see the rank of the RHS is 1, but the rank of the LHS is at least 2: one has

$$\dim\,\mathrm{Ker}\,(\mathbb{R}_1 - \mathbb{R}_2) + \mathrm{rank}\,(\mathbb{R}_1 - \mathbb{R}_2) = d,$$

and the dimension of the kernel is 1, as derived above. So, $\mathrm{rank}\,(\mathbb{R}_1 - \mathbb{R}_2) = d - 1 \geq 2$, for[3] $d \geq 3$. Finally, using one of the original equations

$$c(\mathbb{R}_1 - \mathbb{R}_2)\mathbf{n} + (\mathbb{R}_1 - \mathbb{R}_2)\mathbf{t} = \mathbf{v}_2 - \mathbf{v}_1,$$

and our fresh information $(\mathbb{R}_1 - \mathbb{R}_2)\mathbf{t} = \mathbf{0}, \mathbb{R}_1 = \mathbb{R}_2$, we conclude that also

$$\mathbf{v}_2 = \mathbf{v}_1.$$

In total, we have obtained $\mathbb{R}_1 = \mathbb{R}_2 \equiv \mathbb{R}, \mathbf{v}_1 = \mathbf{v}_2 \equiv \mathbf{v}$ and the transformation is affine.

Definition 1 (Types of deformation)**.** The deformation $\mathbf{y} : \overline{\Omega} \to \mathbb{R}^d, \mathbb{F}(\mathbf{x}) = \nabla\mathbf{y}(\mathbf{x})$ is called

- homogenous, if $\mathbb{F}$ is constant in $\overline{\Omega}$,
- rigid, if it is homogenous and $\mathbb{F} = \mathbb{R} \in \mathrm{SO(d)}$,
- incompressible, if $\det \mathbb{F} = 1$.

3 Forces

Naturally, the deformation is caused by the presence of some forces. To capture the evolution of the shape of the body throughout the deformation even when the forces are known is kind of

[3] In the case $d = 2$ we have to argue differently, but the assertion is true.

complicated, as it is nonlinear from its very nature (we will see this in the upcoming chapter). Classical physics gives us some balance laws, that usually hold in inertial frames, *i.e.*, in time-constant frames. Hence, formulating fundamental balance laws will always be easier in the deformed, current configuration, rather then in the reference configuration. Let us so begin with the study of forces in the currrent configuration.

3.1 Forces in the deformed configuration

Recall that our assumption always is $\mathbf{y}(\overline{\Omega}) = \overline{\Omega^y} = \overline{\Omega}^y$, for $\Omega^y = \mathbf{y}(\Omega)$. We are thus able to define the **volume density of applied body forces**

$$\mathbf{f}^y : \overline{\Omega}^y \to \mathbb{R}^3$$

(in newtons per cubic meters, e.g. gravity). The same on the boundary

$$\mathbf{g}^y : \Gamma_N^y \to \mathbb{R}^3$$

(**surface density of applied contact forces** (in newtons per square meters = Pascals, e.g. hydrostatic pressure.)

3.1.1 Cauchy stress tensor

Lemma 1 (Stress principle of Euler and Cauchy)**.** *There exists a (Cauchy) stress vector function* $\mathbf{t}^y : \overline{\Omega}^y \times \mathcal{S}^{d-1} \to \mathbb{R}^d$ *with the following properties.*

1. *If* $\mathbf{x}^y \in \Gamma_N^y$*, then* $\mathbf{t}^y(\mathbf{x}^y, \mathbf{n}^y) = \mathbf{g}^y(\mathbf{x}^y)$*, where* $\mathbf{n}^y$ *is the unit outer normal vector to* $\partial\Omega^y$ *at* $\mathbf{x}^y$*.*
2. $\forall \omega^y \subset \Omega^y$ *it holds that*
$$\int_{\omega^y} \mathbf{f}(\mathbf{x}^y)\,\mathrm{d}\mathbf{x}^y + \int_{\partial\omega^y} \mathbf{t}^y(\mathbf{x}^y, \mathbf{n}^y)\,\mathrm{d}S^y = \mathbf{0}.$$
(Balance of forces in static equilibrium.)
3. $\forall \omega^y \subset \Omega^y$ *it holds that*
$$\int_{\omega^y} \mathbf{x}^y \times \mathbf{f}^y(\mathbf{x}^y)\,\mathrm{d}\mathbf{x}^y + \int_{\partial\omega^y} \mathbf{x}^y \times \mathbf{t}^y(\mathbf{x}^y, \mathbf{n}^y)\,\mathrm{d}S^y = \mathbf{0}.$$
(Balance of momenta of forces in static equilibrium.)

Euler says (while thinking of the Newton's 3rd law) that the direct consequence of this is the existence of $\mathbb{T}^y(\mathbf{x}^y)$ *such that*

$$\mathbf{t}^y(\mathbf{x}^y, \mathbf{n}^y) = \mathbb{T}^y(\mathbf{x}^y)\mathbf{n}^y, \tag{5}$$

where the tensorial quantity (field) $\mathbb{T}^y(\mathbf{x}^y)$ *is called the* ***Cauchy stress tensor****.*

3.1.2 Balance equations in the deformed configuration

Classical physics gives us 2 fundamental relations: Newtons second law for momenta and for angular momenta. We examine these in the continuum mechanics setting.

Using the second property together with 5 gives

$$\int_{\omega^y} \mathbf{f}^y(\mathbf{x}^y)\,\mathrm{d}\mathbf{x}^y + \int_{\partial\omega^y} \mathbb{T}^y(\mathbf{x}^y)\mathbf{n}^y\,\mathrm{d}S^y = \int_{\omega^y} \mathbf{f}^y(\mathbf{x})^y\,\mathrm{d}\mathbf{x}^y + \int_{\omega^y} \nabla\cdot\mathbb{T}^y(\mathbf{x}^y)\,\mathrm{d}\mathbf{x}^y = 0, \tag{6}$$

so using the localization theorem we get

$$\mathbf{f}^y(\mathbf{x}^y) + \nabla\cdot\mathbb{T}^y(\mathbf{x}^y) = \mathbf{0}, \forall \mathbf{x}^y \in \Omega^y. \tag{7}$$

From the third property it follows

$$\begin{aligned}
\int_{\omega^y} \mathbf{x}^y \times \mathbf{f}^y(\mathbf{x}^y)\,\mathrm{d}\mathbf{x}^y + \int_{\partial\omega^y} \mathbf{x}^y \times \mathbb{T}^y(\mathbf{x}^y)\mathbf{n}^y\,\mathrm{d}S^y &= \int_{\omega^y} \varepsilon_{ijk} x_j^y f_k^y \mathbf{e}_i\,\mathrm{d}\mathbf{x}^y + \int_{\omega^y} \varepsilon_{ijk} x_j^y (T_{km}^y n_m^y)\mathbf{e}_i\,\mathrm{d}S^y = \\
&= \int_{\omega^y} \varepsilon_{ijk} x_j^y f_k^y \mathbf{e}_i\,\mathrm{d}\mathbf{x}^y + \int_{\omega^y} \varepsilon_{ijk} \frac{\partial(x_j^y T_{km}^y)}{\partial x_m^y}\mathbf{e}_i\,\mathrm{d}\mathbf{x}^y = \\
&= \int_{\omega^y} \varepsilon_{ijk} x_j^y f_k^y \mathbf{e}_i\,\mathrm{d}\mathbf{x}^y + \\
&+ \int_{\omega^y} \varepsilon_{ijk} x_j^y \frac{\partial T_{km}^y}{\partial x_m^y}\mathbf{e}_i\,\mathrm{d}\mathbf{x}^y + \int_{\omega^y} \varepsilon_{ijk}\delta_{jm} T_{km}^y \mathbf{e}_i\,\mathrm{d}\mathbf{x}^y = \mathbf{0},
\end{aligned}$$

realize now

$$\int_{\omega^y} \varepsilon_{ijk}\left(x_j^y f_k^y + x_j^y \frac{\partial T_{km}^y}{\partial x_m^y}\right)\mathbf{e}_i\,\mathrm{d}\mathbf{x}^y = \int_{\omega^y} \varepsilon_{ijk} x_j^y \left(f_k^y + \underbrace{\frac{\partial T_{km}^y}{\partial x_m^y}}_{=(\nabla\cdot\mathbb{T}^y)_k}\right)\mathbf{e}_i\,\mathrm{d}\mathbf{x}^y = \mathbf{0},$$

because the balance of forces 7 holds. The balance of angular momenta thus reduces to

$$\int_{\omega^y} \varepsilon_{ijk} T_{kj}^y \mathbf{e}_i\,\mathrm{d}\mathbf{x}^y = \mathbf{0},$$

and using the localization theorem, we obtain

$$T_{kj}^y(\mathbf{x}^y) = T_{jk}^y(\mathbf{x}^y),\ i.e.\, \mathbb{T}^y(\mathbf{x}^y) = (\mathbb{T}^y(\mathbf{x}^y))^\top. \tag{8}$$

The **Cauchy stress tensor is symmetric.**

3.2 Forces in the undeformed configuration

We have obtained the equations in the deformed configuration, where they are easily formulated. On the other hand, that is a bit unconvenient - we solve the equations to find the deformed configuration, *i.e.*, the equations hold in the domain that is obtained as a solution to the equations themselves. This brings us to to the need to find a new way to write the down the balance laws - in the reference configuration. The construction is a bit synthetic, our intuition will be guided by the requirement to obtain similair equations as in the current configuration.

3.2.1 Piola-Kirchhoff stresses

Definition 2 (First Piola-Kirchhoff stress tensor)**.** Given the Cauchy stress tensor $\mathbb{T}^y(\mathbf{x}^\mathbf{y})$, we define the **First Piola Kirchhoff stress tensor**

$$\mathbb{T} : \overline{\Omega} \to \mathbb{R}^{3\times 3}, \mathbb{T}(\mathbf{x}) = \mathbb{T}^y(\mathbf{y}(\mathbf{x}))\,\mathrm{cof}\,\mathbb{F}(\mathbf{x}) = \det\mathbb{F}(\mathbf{x})\mathbb{T}^y(\mathbf{y}(\mathbf{x}))\mathbb{F}^{-\top}(\mathbf{x}).$$

Definition 3 (Second Piola-Kirchhoff stress tensor)**.** The quantity

$$\mathbb{S}(\mathbf{x}) = \mathbb{F}^{-1}(\mathbf{y}(\mathbf{x}))\mathbb{T}(\mathbf{x}),$$

is called the **second Piola-Kirchhoff stress tensor.**

Remark. The first Piola-Kirchhoff stress tensor $\mathbb{T}$ is *not symmetric in general*, but the second

$$\mathbb{S} = \det \mathbb{F}\mathbb{F}^{-1}\mathbb{T}^y\mathbb{F}^{-\top}$$

is symmetric. Also, we see that not every matrix can serve as $\mathbb{T}$; it must hold $\mathbb{T}(\operatorname{cof}\mathbb{F})^{-1}$ is symmetric.

Remark. We have the following identity (contrary to the appearance, this is not a trivial computation and one has to use the Piola's identity):

$$\nabla_{\mathbf{x}} \cdot \mathbb{T}(\mathbf{x}) = \det \mathbb{F}(\mathbf{x}) \left(\nabla_{\mathbf{x}^y} \cdot \mathbb{T}^y(\mathbf{x}^y)\right)\Big|_{\mathbf{x}^y=\mathbf{y}(\mathbf{x})}. \tag{9}$$

3.2.2 Balance equations in the deformed configuration

Recall the balance of momentum

$$-\nabla \cdot \mathbb{T}^y(\mathbf{x}^y) = \mathbf{f}^y(\mathbf{x}^y) \,\mathrm{in}\, \Omega^y,$$

multiplying by $\det \mathbb{F} > 0$ yields

$$\det \mathbb{F}(\mathbf{x})\nabla \cdot (\mathbb{T}^y(\mathbf{x}^y)) = \det \mathbb{F}(\mathbf{x})\mathbf{f}^y(\mathbf{x}^y).$$

Using the Piola's identity we have shown 9, so the left hand side actually is $-\nabla \cdot \mathbb{T}(\mathbf{x})$. Seeing the similarity with the balance of forces in the current configuration, we are tempted to denote

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}^y(\mathbf{y}(\mathbf{x}))\det \mathbb{F}(\mathbf{x}),$$

as the volume density of body forces in the *reference configuration*; after this definition, the balance of forces becomes

$$-\nabla \cdot \mathbb{T}(\mathbf{x}) = \mathbf{f}(\mathbf{x}),$$

and since this equation is formulated in the reference configuration, we call it the **balance of forces in the reference configuration.**[4]

Viewed from an "integral" perspective, the total applied body force on the body can be written as

$$\int_{\omega^y} \mathbf{f}^y(\mathbf{x}^y)\,\mathrm{d}\mathbf{x}^y = \int_{\omega} \mathbf{f}^y(\mathbf{y}(\mathbf{x}))\det \mathbb{F}(\mathbf{x})\,\mathrm{d}x = \int_{\omega} \mathbf{f}(\mathbf{x})\,\mathrm{d}\mathbf{x}\,.$$

We can do the same for the balance of surface forces; the total contact force acting on the body is

$$\begin{aligned}\int_{\partial\omega^y} \mathbf{g}^y(\mathbf{x}^y)\,\mathrm{d}S^y &= \int_{\partial\omega^y} \mathbf{t}^y(\mathbf{x}^y, \mathbf{n}^y)\,\mathrm{d}S^y = \int_{\partial\omega^y} \mathbb{T}^y(\mathbf{x}^y)\mathbf{n}^y(\mathbf{x}^y)\,\mathrm{d}S^y = \\ &= \int_{\partial\omega} \mathbb{T}^y(\mathbf{y}(\mathbf{x}))\operatorname{cof}\mathbb{F}(\mathbf{x})\mathbf{n}(\mathbf{x})\,\mathrm{d}S = \int_{\partial\omega} \mathbb{T}(\mathbf{x})\mathbf{n}(\mathbf{x})\,\mathrm{d}S = \int_{\partial\omega} \mathbf{g}^y(\mathbf{y}(\mathbf{x}))|\operatorname{cof}\mathbb{F}(\mathbf{x})\mathbf{n}(\mathbf{x})|\,\mathrm{d}S\,,\end{aligned}$$

[4] This title is used even though it is slightly misleading - the forces are still acting in the current configuration, but are expressed through the reference configuration.

so if we define

$$\mathbf{g}(\mathbf{x}) = \mathbb{T}(\mathbf{x})\mathbf{n}(\mathbf{x}) = \mathbf{g}^y(\mathbf{y}(\mathbf{x}))|\mathrm{cof}\,\mathbb{F}(\mathbf{x})\mathbf{n}(\mathbf{x})|,$$

as the contact force in the **reference configuration**, we formally have a similiar expression.

Remark. Let us summarize our stress measures.

- the Cauchy stress tensor: $\mathbb{T}^y = \mathbb{T}^y(\mathbf{x}^y)$. The Cauchy stress tensor is connected with the surface density of the contact forces in the *current configuration* given a point in the *current configuration*. In concrete terms

$$\mathbf{g}^y(\mathbf{x}^y) = \mathbb{T}^y(\mathbf{x}^y)\mathbf{n}^y(\mathbf{x}^y),$$

- the First Piola-Kirchhoff stress tensor: $\mathbb{T} = \mathbb{T}(\mathbf{x}), \mathbb{T}(\mathbf{x}) = \det\mathbb{F}(\mathbf{x})\mathbb{T}^y(\mathbf{y}(\mathbf{x}))\,\mathrm{cof}\,\mathbb{F}(\mathbf{x})$. The First Piola-Kirchhoff stress tensor is connected with the surface density of the contact forces in the *current configuration* given a ppint in the *reference configuration*. In concrete terms

$$\mathbf{g}(\mathbf{x}) = \mathbb{T}(\mathbf{x})\mathbf{n}(\mathbf{x}),$$

- the Second Piola-Kirchhoff stress tensor $\mathbb{S} = \mathbb{S}(\mathbf{x}), \mathbb{S}(\mathbf{x}) = \mathbb{F}^{-1}(\mathbf{y}(\mathbf{x}))\mathbb{T}(\mathbf{x})$ works similairly to the First Piola-Kirchhoff tensor, but is symmetric.

Note carefully that in both cases we are talking about forces in the *current configuration*, but these can be expressed either through the *reference configuration* (with the help of $\mathbb{T}$) or through the *current configuration* itself (by using $\mathbb{T}^y$).

3.3 Conservative forces

Forces that are conservative play an imporant role in physics. One should keep in mind the Emmy Noether theorem: "something is conserved ⇔ the system has a symmetry." In the setting of continuum mechanics, things get mathematicaly more involved. See the definition

Definition 4 (Conservative forces, potential)**.** The (body and contact) forces $\mathbf{f}, \mathbf{g}$ are called conservative, provided there exists a functional

$$\mathcal{P} : \left\{\mathbf{y} \in C^1\left(\overline{\Omega}\right)\right\} \to \mathbb{R},$$

such that it holds

$$\delta\mathcal{P}(\mathbf{y}, \boldsymbol{\varphi}) = \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}(\mathbf{y} + t\boldsymbol{\varphi})\Big|_{t=0} = \int_\Omega \mathbf{f}(\mathbf{x})\cdot\boldsymbol{\varphi}(\mathbf{x})\,\mathrm{d}\mathbf{x} + \int_{\Gamma_N} \mathbf{g}(\mathbf{x})\cdot\boldsymbol{\varphi}(\mathbf{x})\,\mathrm{d}S\,, \forall\boldsymbol{\varphi} \in X,$$

where X is some space of test functions on Ω. The functional $\mathcal{P}$ is called the potential of the forces.

Remark. The forces are in reference configuration.

Example (Hydrostatic force)**.** Let us show that the hydrostatic force

$$\mathbf{g}_h = -p\mathbf{n},$$

with p being the hydrostatic pressure, is conservative and that the potential is

$$\mathcal{P} = -\int_{\Omega^y} p(\mathbf{x}^y)\,\mathrm{d}\mathbf{x}^y = -\int_\Omega p(\mathbf{y}(\mathbf{x}))\det\mathbb{F}(\mathbf{x})\,\mathrm{d}\mathbf{x}\,.$$

Calculate the Gateuax derivative

$$\delta\mathcal{P}(\mathbf{y},\boldsymbol{\varphi}) = -\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega p(\mathbf{y}+t\boldsymbol{\varphi})\det\left(\nabla_x\mathbf{y}+t\nabla_x\boldsymbol{\varphi}\right)\mathrm{d}\mathbf{x}\bigg|_{t=0} = -\int_\Omega \frac{\mathrm{d}}{\mathrm{d}t}p(\mathbf{y}+t\boldsymbol{\varphi})\det\left(\nabla_x\mathbf{y}+t\nabla_x\boldsymbol{\varphi}\right)\bigg|_{t=0}\mathrm{d}x =$$
$$= -\int_\Omega \nabla_y p(\mathbf{y})\cdot\boldsymbol{\varphi}\det\mathbb{F} + p(\mathbf{y})\frac{\partial\det\mathbb{F}}{\partial\mathbb{F}} : \nabla_x\boldsymbol{\varphi}\,\mathrm{d}x = -\int_\Omega \nabla_y p(\mathbf{y})\cdot\boldsymbol{\varphi}\det\mathbb{F} + p(\mathbf{y})\,\mathrm{cof}\,\mathbb{F} : \nabla_x\boldsymbol{\varphi}\,\mathrm{d}x =$$
$$= -\int_\Omega \nabla_y p(\mathbf{y})\cdot\boldsymbol{\varphi}\det\mathbb{F} + \nabla_x\cdot\left(p(\mathbf{y})\,\mathrm{cof}\,\mathbb{F}\boldsymbol{\varphi}\right) - \nabla_x p(\mathbf{y})\cdot\mathrm{cof}\,\mathbb{F}\boldsymbol{\varphi}\,\mathrm{d}x =$$
$$= -\int_\Omega \nabla_y p(\mathbf{y})\cdot\boldsymbol{\varphi}\det\mathbb{F} - \nabla_x p(\mathbf{y})\cdot\mathrm{cof}\,\mathbb{F}\boldsymbol{\varphi}\,\mathrm{d}x - \int_{\partial\Omega} p(\mathbf{y})\boldsymbol{\varphi}\cdot\mathrm{cof}\,\mathbb{F}\mathbf{n}\,\mathrm{d}S\,,$$

where we have used the Piola's identity. Realize now

$$\det\mathbb{F}\nabla_y p(\mathbf{y}) = \det\mathbb{F}\mathbf{e}_i\frac{\partial p(\mathbf{y}(\mathbf{x}))}{\partial y_i} = \det\mathbb{F}\mathbf{e}_i\frac{\partial p(\mathbf{y}(\mathbf{x}))}{\partial x_k}\frac{\partial x_k}{\partial y_i} = \det\mathbb{F}\mathbf{e}_i\nabla_x p(\mathbf{y}(\mathbf{x}))_k\left(\mathbb{F}^{-1}\right)_{ki} =$$
$$= \det\mathbb{F}\mathbb{F}^{-1}\nabla_x p(\mathbf{y}(\mathbf{x})),$$

so

$$-\int_\Omega \nabla_y p(\mathbf{y})\cdot\boldsymbol{\varphi}\det\mathbb{F} - \nabla_x p(\mathbf{y})\cdot\mathrm{cof}\,\mathbb{F}\boldsymbol{\varphi}\,\mathrm{d}x - \int_{\partial\Omega} p(\mathbf{y})\boldsymbol{\varphi}\cdot\mathrm{cof}\,\mathbb{F}\mathbf{n}\,\mathrm{d}S =$$
$$= -\int_\Omega \det\mathbb{F}\mathbb{F}^{-1}\nabla_x p(\mathbf{y})\cdot\boldsymbol{\varphi} - \nabla_x p(\mathbf{y})\cdot\mathrm{cof}\,\mathbb{F}\boldsymbol{\varphi}\,\mathrm{d}x - \int_{\partial\Omega} p(\mathbf{y})\boldsymbol{\varphi}\cdot\mathrm{cof}\,\mathbb{F}\mathbf{n}\,\mathrm{d}S =$$
$$= -\int_\Omega \nabla_x p(\mathbf{y})\cdot\left(\det\mathbb{F}\mathbb{F}^{-\top}\boldsymbol{\varphi}\right) - \nabla_x p(\mathbf{y})\cdot\mathrm{cof}\,\mathbb{F}\boldsymbol{\varphi}\,\mathrm{d}x - \int_{\partial\Omega} p(\mathbf{y})\boldsymbol{\varphi}\cdot\mathrm{cof}\,\mathbb{F}\mathbf{n}\,\mathrm{d}S =$$
$$= -\int_{\partial\Omega} p(\mathbf{y}(\mathbf{x}))\boldsymbol{\varphi}(\mathbf{x})\cdot\mathrm{cof}\,\mathbb{F}\mathbf{n}\,\mathrm{d}S =$$
$$= -\int_{\partial\Omega^y} p(\mathbf{x}^y)\mathbf{n}^y\cdot\boldsymbol{\varphi}(\mathbf{y}(\mathbf{x}))\,\mathrm{d}S^y\,,$$

and so we have shown the hydrostatic force $-p\mathbf{n}$ is conservative with the potential $-\int_{\Omega^y} p(\mathbf{x}^y)\,\mathrm{d}\mathbf{x}^y$.

4 Elasticity

So far, our excursion to the world of solid mechanics has been fairly general. We have formulated fundamental mechanical laws independetly of the material that undergoes the deformation. In this chapter, we will discuss a class of materials known as **elastic solids.**

Definition 5 (Elasticity)**.** We say that a material is **elastic (or Cauchy elastic)** if the Cauchy stress tensor is determined only by the current configuration. In morece concrete terms, the material is elastic provided there is a response function $\tilde{\mathbb{T}}^D : \Omega\times\mathbb{R}^{3\times3}_+ \to \mathbb{R}^{3\times3}_{\mathrm{sym}}$ such that (Kružík and Roubíček, 2019a)

$$\mathbb{T}^y(\mathbf{y}(\mathbf{x})) = \tilde{\mathbb{T}}^D(\mathbf{x},\mathbb{F}).$$

The response function is also called the **constitutive law**. Realize that this consideration is quite restrictive

- the response does not depend on any thermodynamical quantities (explicitely): temperature, dissipation, etc.,
- the response depends only on the gradient of the deformation, not on the deformation itself,
- the response function *does depend* on the reference configuration, but the stress *does not.*

Remark. If we know the material is elastic, we also have the information about the First Piola-Kirchhoff stress, as $\mathbb{T}(\mathbf{x}) = \mathbb{T}^y(\mathbf{x}^y)\operatorname{cof}\mathbb{F}(\mathbf{x})$, so

$$\mathbb{T}(\mathbf{x}) = \mathbb{T}^y(\mathbf{y}(\mathbf{x}))\operatorname{cof}\mathbb{F}(\mathbf{x}) = \tilde{\mathbb{T}}^D(\mathbf{x},\mathbb{F})\operatorname{cof}\mathbb{F}(\mathbf{x}) := \tilde{\mathbb{T}}(\mathbf{x},\mathbb{F}). \tag{10}$$

We will be using both $\tilde{\mathbb{T}}^D, \tilde{\mathbb{T}}$ from now on.

4.1 Frame invariance principle

The principle of (material) frame invariance, or (material) frame indifference is closely connected with the notion of **objectivity**. Those terms concern the change of observer: a transformation $\mathbf{x} \mapsto \mathbb{R}\mathbf{x}$, for some $\mathbb{R} \in \mathrm{SO}(3)$. A certain invariance is fundamental to (classical) physics - recall the Galileo's principle of relativity. At the moment, we will deal ourselves with "objective vectors and tensors." That is, suppose we are given the stress vector $\mathbf{t}^y(\mathbf{x}^y, \mathbf{n}^y)$ and a different deformation

$$\mathbf{z}(\mathbf{x}) = \mathbb{R}\mathbf{y}(\mathbf{x}), \mathbb{R} \in \mathrm{SO}(3),$$

(this deformation is only a rotation, so it can in fact be seen as a change of the frame in the current configuration.) Denote as expected $\mathbf{x}^z = \mathbf{z}(\mathbf{x})$; the priniciple of frame invariance then states

$$\mathbf{t}^z(\mathbf{x}^z, \mathbf{n}^z) = \mathbf{t}^z(\mathbb{R}\mathbf{x}^y, \mathbb{R}\mathbf{n}^y) \equiv \mathbb{R}\mathbf{t}^y(\mathbf{x}^y, \mathbf{n}^y), \forall \mathbb{R} \in \mathrm{SO}(3),$$

i.e., if we just rotate the current configuration, the traction vector also only rotates. Realize that this in fact means

$$\mathbf{t}^z(\mathbf{x}^z, \mathbf{n}^z) = \mathbf{t}^z(\mathbf{x}^z, \mathbb{R}\mathbf{n}^y) = \mathbb{T}^z(\mathbf{x}^z)\mathbb{R}\mathbf{n}^y = \mathbb{R}\mathbf{t}^y(\mathbf{x}^y, \mathbf{n}^y) = \mathbb{R}\mathbb{T}^y(\mathbf{x}^y)\mathbf{n}^y,$$

where we have denoted $\mathbb{T}^z$ to be the Cauchy stress tensor in the configuration after the deformation $\mathbf{z}$. This however means

$$\mathbb{T}^z(\mathbf{x}^z) = \mathbb{T}^z(\mathbb{R}\mathbf{x}^y) = \mathbb{R}\mathbb{T}^y(\mathbf{x}^y)\mathbb{R}^\top.$$

This is the transformation of the Cauchy stress tensor under the change of observer. It can be shown [5] that the deformation gradient transforms as $\mathbb{R}\mathbb{F}$. On the level of constitutive laws, the frame invariance principle states:

$$\tilde{\mathbb{T}}^D(\mathbf{x}, \mathbb{R}\mathbb{F}) = \mathbb{R}\tilde{\mathbb{T}}^D(\mathbf{x}, \mathbb{F})\mathbb{R}^\top, \forall \mathbb{R} \in \mathrm{SO}(3), \forall \mathbb{F} \in \mathbb{R}^{3\times 3}_+, \forall \mathbf{x} \in \overline{\Omega},$$

which is closely related to the transformation properties of the Cauchy stress tensor. See that we are not transorming $\mathbf{x}$ - that is a vector from the *reference configuration*, which remains unchanged in the change of observer transformation. For $\tilde{\mathbb{T}}$ (defined in 10) one has

$$\tilde{\mathbb{T}}(\mathbf{x}, \mathbb{R}\mathbb{F}) = \det(\mathbb{R}\mathbb{F})\tilde{\mathbb{T}}^D(\mathbf{x}, \mathbb{R}\mathbb{F})(\mathbb{R}\mathbb{F})^{-\top} = \det(\mathbb{R}\mathbb{F})\mathbb{R}\tilde{\mathbb{T}}^D(\mathbf{x}, \mathbb{F})\mathbb{R}^\top\mathbb{R}\mathbb{F}^{-\top} = \det\mathbb{F}\mathbb{R}\tilde{\mathbb{T}}^D(\mathbf{x}, \mathbb{F})\mathbb{F}^{-\top} = \mathbb{R}\tilde{\mathbb{T}}(\mathbf{x}, \mathbb{F}),$$

thus

$$\tilde{\mathbb{T}}(\mathbf{x}, \mathbb{R}\mathbb{F}) = \mathbb{R}\tilde{\mathbb{T}}(\mathbf{x}, \mathbb{F}), \forall \mathbb{R} \in \mathrm{SO}(3), \forall \mathbb{F} \in \mathbb{R}^{3\times 3}_+, \forall x \in \overline{\Omega},$$

Notice that the primary formulation of the frame indifference principle is through the constitutive law for the Cauchy stress tensor, an objective tensor, from which the transformation of the First Piola-Kirchhoff tensor is derived.

[5] Let $\mathbf{x}^z = \mathbb{R}\mathbf{x}^y$, then

$$F^z_{ij} = \frac{\partial x^z_i}{\partial x_j} = \frac{\partial x^z_i}{\partial x^y_k}\frac{\partial x^y_k}{\partial x_j} = \frac{\partial}{\partial x^y_k}(R_{im}x^y_m)F^y_{kj} = R_{ik}F^y_{kj}.$$

4.2 Isotropic material

The principle of material frame indifference is a consequence of the fundamental invariance of laws of classical physics - it holds for any materials whatsoever. Still, some materials have further important properties that should be captured in the constitutive law. We examine now the property of **material symmetry**.

Take $\mathbf{x}_0 \in \overline{\Omega}$ general but fixed, take

$$\mathbf{v}(\mathbf{z}) = \mathbf{x}_0 + \mathbb{R}^\top(\mathbf{z} - \mathbf{x}_0)$$

for some $\mathbb{R} \in \mathrm{SO}(3)$, so ($\tilde{\mathbf{x}} := \mathbf{v}(\mathbf{z})$)

$$\mathbf{v}^{-1}(\tilde{\mathbf{x}}) = \mathbf{x}_0 + \mathbb{R}(\tilde{\mathbf{x}} - \mathbf{x}_0),$$

and define a *new deformation* $\tilde{\mathbf{y}} = \mathbf{y} \circ \mathbf{v}^{-1} : \mathbf{v}(\overline{\Omega}) \to \mathbf{y}(\overline{\Omega})$, as

$$\tilde{\mathbf{y}}(\tilde{\mathbf{x}}) = \mathbf{y}(\mathbf{x}_0 + \mathbb{R}(\tilde{\mathbf{x}} - \mathbf{x}_0)).$$

Compute

$$\tilde{\mathbb{F}}(\tilde{\mathbf{x}}) = \nabla_{\tilde{\mathbf{x}}}\tilde{\mathbf{y}} = \mathbb{F}(\tilde{\mathbf{x}})\mathbb{R},$$

and notice that $\mathbf{x}_0^{\tilde{y}} = \tilde{\mathbf{y}}(\mathbf{x}_0) = \mathbf{y}(\mathbf{x}_0) = \mathbf{x}_0^{y}$, from which it follows

$$\mathbb{T}^y\left(\mathbf{x}_0^y\right) = \tilde{\mathbb{T}}^D(\mathbf{x}_0, \mathbb{F}) = \mathbb{T}^{\tilde{y}}(\mathbf{x}_0^{\tilde{y}}) = \tilde{\mathbb{T}}^D\left(\mathbf{x}_0, \tilde{\mathbb{F}}(\mathbf{x}_0)\right) = \tilde{\mathbb{T}}^D(\mathbf{x}_0, \mathbb{F}(\mathbf{x}_0)\mathbb{R}).$$

We see that at the point $\mathbf{x}_0$, the response function has the property

$$\tilde{\mathbb{T}}^D\left(\mathbf{x}_0, \tilde{\mathbb{F}}(\mathbf{x}_0)\right) = \tilde{\mathbb{T}}^D\left(\mathbf{x}_0, \tilde{\mathbb{F}}(\mathbf{x}_0)\mathbb{R}\right),$$

which motivates our definition

Definition 6 (Isotropic material)**.** We cal the material **isotropic** if it holds

$$\tilde{\mathbb{T}}^D(\mathbf{x}, \mathbb{F}) = \tilde{\mathbb{T}}^D(\mathbf{x}, \mathbb{F}\mathbb{R}), \forall \mathbb{R} \in \mathrm{SO}(3), \forall \mathbb{F} \in \mathbb{R}_+^{3\times 3}, \forall \mathbf{x} \in \overline{\Omega}.$$

Remark. For the First Piola-Kirchhoff we obtain:

$$\tilde{\mathbb{T}}(\mathbf{x}, \mathbb{F}\mathbb{R}) = \tilde{\mathbb{T}}^D(\mathbf{x}, \mathbb{F}\mathbb{R}) \operatorname{cof}(\mathbb{F}\mathbb{R}) = \tilde{\mathbb{T}}^D(\mathbf{x}, \mathbb{F})\mathbb{F}^{-\top}\mathbb{R} \det \mathbb{F} = \tilde{\mathbb{T}}(\mathbf{x}, \mathbb{F})\mathbb{R},$$

In total, for a isotropic elastic material the constitutive law for the First Piola-Kirchhoff stress tensor has the property

$$\tilde{\mathbb{T}}(\mathbf{x}, \mathbb{R}_1\mathbb{F}\mathbb{R}_2) = \mathbb{R}_1\tilde{\mathbb{T}}(\mathbf{x}, \mathbb{F})\mathbb{R}_2, \forall \mathbb{R}_1, \mathbb{R}_2 \in \mathrm{SO}(3), \forall \mathbb{F} \in \mathbb{R}_+^{3\times 3}, \forall \mathbf{x} \in \overline{\Omega},$$

so in particular

$$\tilde{\mathbb{T}}\left(\mathbf{x}, \mathbb{R}\mathbb{F}\mathbb{R}^\top\right) = \mathbb{R}\tilde{\mathbb{T}}(\mathbf{x}, \mathbb{F})\mathbb{R}^\top,$$

which will prove very useful later on.

Remark. We stress that instead of material indifference, that granted us $\tilde{\mathbb{T}}(\mathbf{x}, \mathbb{R}\mathbb{F}) = \mathbb{R}\tilde{\mathbb{T}}(\mathbf{x}, \mathbb{F})$, and which is valid for all materials, the second property $\tilde{\mathbb{T}}(\mathbf{x}, \mathbb{F}\mathbb{R}) = \tilde{\mathbb{T}}(\mathbf{x}, \mathbb{F})\mathbb{R}$ holds *only for isotropic materials.*

4.3 Hyperelastic materials

Definition 7. We say that a material is hyperelastic (sometimes called Green elastic) if there is a function $W : \overline{\Omega} \times \mathbb{R}^{3\times3}_{+} \to \mathbb{R}$ such that (Kružík and Roubíček, 2019a)

$$\mathbb{T}(\mathbf{x}) = \tilde{\mathbb{T}}(\mathbf{x}, \mathbb{F}) = \frac{\partial W(\mathbf{x}, \mathbb{F})}{\partial \mathbb{F}}.$$

The function $W, [W] = \frac{J}{m^3} = \frac{Nm}{m^3} = \mathrm{Pa}$, is called **stored energy density.**

Remark. See that this simple is the formula for the First Piola-Kirchhoff stress; for the Cauchy stress tensor, it gets more complicated.

4.4 Properties of W

Physics puts some assumption on W :

1. $W \geq 0$,
2. $W(\mathbb{I}) = 0$,
3. $W(\mathbf{x}, \mathbb{R}\mathbb{F}) = W(\mathbf{x}, \mathbb{F}), \forall \mathbb{R} \in \mathrm{SO}(3), \forall \mathbf{x} \in \overline{\Omega}, \forall \mathbb{F} \in \mathbb{R}^{3\times3}_{+}$,
4. $W(\mathbf{x}, \underbrace{\mathbb{R}\mathbb{U}}_{=\mathbb{F}}) = W(\mathbf{x}, \mathbb{U}), \mathbb{U} = \sqrt{\mathbb{C}}$, (matrices are from the polar decomposition)
5. $W(\mathbf{x}, \mathbb{F}) \to \infty$ if $\det \mathbb{F} \to 0_{+}$,
6. $\exists \alpha > 0, \exists p, q, r \geq 1 \; s.t. \; W(\mathbf{x}, \mathbb{F}) \geq \alpha(\|\mathbb{F}\|^p + \|\mathrm{cof}\, \mathbb{F}\|^q + (\det \mathbb{F})^r) - d, \forall \mathbf{x} \in \overline{\Omega}, \forall \mathbb{F} \in \mathbb{R}^{3\times3}_{+}$.

Let us comment on these briefly.

1. means really that energy is nonnegative,
2. means that no energy is needed to not deform the material,
3. states that energy does not change under the change of observer; see below for details. This in particular also means $W(\mathbb{R}) = 0 \forall \mathbb{R} \in \mathrm{SO(d)}$.[6]
4. energy changes only when the the geometry of the domain changes,
5. it takes infinite energy to deform the body to a point,
6. mostly a mathematical assumption - this assures *coercivity*

The principle of frame indifference told us

$$\tilde{\mathbb{T}}(\mathbf{x}, \mathbb{R}\mathbb{F}) = \mathbb{R}\tilde{\mathbb{T}}(\mathbf{x}, \mathbb{F}),$$

so

$$\mathbb{R}^{\top} \frac{\partial W(\mathbf{x}, \mathbb{R}\mathbb{F})}{\partial \mathbb{F}} = \frac{\partial W(\mathbf{x}, \mathbb{F})}{\partial \mathbb{F}},$$

[6]We could also show that if the hyperelastic solid is also isotropic, we also have $W(\mathbf{x}, \mathbb{F}\mathbb{R}) = W(\mathbf{x}, \mathbb{F})$, and so $W(\mathbf{x}, \mathbb{R}\mathbb{F}\mathbb{R}^{\top}) = W(\mathbf{x}, \mathbb{F}), \forall \mathbb{R} \in \mathrm{SO(d)}$, so W is an isotropic scalar function, and from the fact $W(\mathbf{x}, \mathbb{F}) = \hat{W}(\mathbf{x}, \mathbb{C})$, in fact follows $W(\mathbf{x}, \mathbb{F}) = \hat{W}(\mathbf{x}, \mathbb{C}) = \hat{W}(\mathbf{x}, \mathbb{R}\mathbb{C}\mathbb{R}^{\top})$, and so $\hat{W}$ is a scalar isotropic function of a symmetric positive definite tensorial argument. This allows one to represent W using only the invariants of $\mathbb{C}$.

we *suppose we can Taylor expand*:

$$W\big(\mathbf{x}, \mathbb{R}(\mathbb{F}+\tilde{\mathbb{F}})\big) = W\big(\mathbf{x}, \mathbb{R}\mathbb{F}+\mathbb{R}\tilde{\mathbb{F}}\big) = W(\mathbf{x}, \mathbb{R}\mathbb{F}) + \frac{\partial W(\mathbf{x}, \mathbb{R}\mathbb{F})}{\partial \mathbb{F}} : \big(\mathbb{R}\tilde{\mathbb{F}}\big) + \text{h.o.t.}$$
$$= W(\mathbf{x}, \mathbb{R}\mathbb{F}) + \mathbb{R}^{\intercal}\frac{\partial W(\mathbf{x}, \mathbb{R}\mathbb{F})}{\partial \mathbb{F}} : \tilde{\mathbb{F}} + \text{h.o.t.}\,,$$

where we have used

$$\mathbb{A} : (\mathbb{B}\mathbb{C}) = \sqrt{\operatorname{tr}(\mathbb{A}^{\intercal}\mathbb{B}\mathbb{C})} = \sqrt{\operatorname{tr}(\mathbb{C}^{\intercal}\mathbb{B}^{\intercal}\mathbb{A})} = \sqrt{\operatorname{tr}(\mathbb{B}^{\intercal}\mathbb{A}\mathbb{C}^{\intercal})} = \big(\mathbb{B}^{\intercal}\mathbb{A}\big) : \mathbb{C}.$$

Using the principle of frame indifference, we in fact have shown

$$W\big(\mathbf{x}, \mathbb{R}(\mathbb{F}+\tilde{\mathbb{F}})\big) = W(\mathbf{x}, \mathbb{R}\mathbb{F}) + \frac{\partial W(\mathbf{x}, \mathbb{F})}{\partial \mathbb{F}} : \tilde{\mathbb{F}}.$$

If we expand now just $W\big(\mathbf{x}, \mathbb{F}+\tilde{\mathbb{F}}\big)$, we obtain

$$W\big(\mathbf{x}, \mathbb{F}+\tilde{\mathbb{F}}\big) = W(\mathbf{x}, \mathbb{F}) + \frac{\partial W(\mathbf{x}, \mathbb{F})}{\partial \mathbb{F}} : \tilde{\mathbb{F}},$$

so upon subtraction we obtain

$$W\big(\mathbf{x}, \mathbb{R}\big(\mathbb{F}+\tilde{\mathbb{F}}\big)\big) - W\big(\mathbf{x}, \mathbb{F}+\tilde{\mathbb{F}}\big) = W(\mathbf{x}, \mathbb{R}\mathbb{F}) - W(\mathbf{x}, \mathbb{F}),$$

and above we have shown this quantity has a zero first order expansion,*i.e.*

$$\frac{\partial}{\partial \mathbb{F}}\left(W(\mathbf{x}, \mathbb{R}\mathbb{F}) - W(\mathbf{x}, \mathbb{F})\right) = \mathbb{0}.$$

from which it follows [7]

$$W(\mathbf{x}, \mathbb{R}\mathbb{F}) = W(\mathbf{x}, \mathbb{F}) + k(\mathbb{R}).$$

If we first take $\mathbb{F} = \mathbb{I}$, then

$$W(\mathbf{x}, \mathbb{R}) = W(\mathbf{x}, \mathbb{I}) + k(\mathbb{R}),$$

and then $\mathbb{F} = \mathbb{R}$,

$$W\big(\mathbf{x}, \mathbb{R}^2\big) = W(\mathbf{x}, \mathbb{R}\mathbb{R}) = W(\mathbf{x}, \mathbb{R}) + k(\mathbb{R}) = W(\mathbf{x}, \mathbb{I}) + 2k(\mathbb{R}),$$

so

$$W(\mathbf{x}, \mathbb{R}^n) = W(\mathbf{x}, \mathbb{I}) + nk(\mathbb{R}), \forall \mathbb{R} \in \mathrm{SO(d)}$$

Since the set of rotations is closed and bounded, it is compact, so there exists a convergent subsequence of $\{\mathbb{R}^n\}$. Moreover, we assume W to be continuous (we took the derivative...), so $\lim_{n\to\infty} W(\mathbf{x}, \mathbb{R}^n)$ exists and from the properties of W we get it is finite. But then $k(\mathbb{R}) = 0$, as otherwise $nk(\mathbb{R}) \to \infty$. All in all, we have shown

$$W(\mathbf{x}, \mathbb{R}\mathbb{F}) = W(\mathbf{x}, \mathbb{F}), \forall \mathbb{R} \in \mathrm{SO(d)}\,, \forall \mathbb{F} \in \mathbb{R}_+^{d\times d}, \forall x \in \overline{\Omega}. \tag{11}$$

This also implies the property

$$W(\mathbf{x}, \mathbb{F}) = W(\mathbf{x}, \mathbb{R}\mathbb{U}) = W(\mathbf{x}, \mathbb{U}),$$

as from the properties of the polar decomposition one has $\mathbb{R} \in \mathrm{SO(d)}\,, \mathbb{U} \in \mathbb{R}_+^{d\times d}$.

[7] The set of matrices with positive determinant is connected.

Remark. From its very definition, W has a potential.

Definition 8 (Natural state of the body)**.** The natural state of the body is the state in which

$$W(\mathbf{x},\mathbb{F}) = 0 \wedge \frac{\partial W}{\partial \mathbb{F}}(\mathbf{x},\mathbb{F}) = \mathbb{0}. \tag{12}$$

Remark (Unnatural vegetable). Not all materials (bodies) have its natural states. Zum Beispiel, a carrot does not have a natural state.

4.5 Variational approach to hyperelasticity

The motion is governed by the laws discussed in the section 3.2, and so in theory the deformation $\mathbf{y}$ can be solved for. In the case of hyperelasticity, there is another, and quite elegant, approach to obtain the sought deformation - the deformation will be a minimizer to a certain functional. This is not surprising at all - a connection between minimizers and weak solutions to Euler-Lagrange equations has been established in the courses on PDE's.

Definition 9 (Energy functional)**.** Let us have $\partial\Omega = \Gamma_N \cup \Gamma_D, \Gamma_N \cap \Gamma_D = \varnothing$, where the parts of the boundary are those where Neumann/Dirichlet boudary conditions are prescribed. The energy functional of the material is the functional

$$I(\mathbf{y}) = \int_\Omega W(\mathbf{x},\nabla\mathbf{y}(\mathbf{x}))\,\mathrm{d}\mathbf{x} - \int_\Omega \mathbf{f}(\mathbf{x})\cdot\mathbf{y}(\mathbf{x})\,\mathrm{d}\mathbf{x} - \int_{\Gamma_N} \mathbf{g}(\mathbf{x})\cdot\mathbf{y}(\mathbf{x})\,\mathrm{d}S\,,$$

where the first part corresponds to the stored energy and the remaining terms are the work done by the external loads.

Remark. This is exactly the notion from PDE's: if one looks for the deformation in a Sobolev space $\mathbf{y}_0 + \mathrm{W}^{1,r}_0\big(\Omega;\mathbb{R}^d\big)$, for some $r \in (1,\infty)$ and some Dirichlet boundary condition $\mathbf{y}_0 \in \mathrm{W}^{1,r}\big(\Omega;\mathbb{R}^d\big)$ and the sources are at least from $\mathbf{f} \in \big(\mathrm{W}^{1,r}\big(\Omega;\mathbb{R}^d\big)\big)^*, \mathbf{g} \in \mathrm{L}_2\big(\Gamma_N;\mathbb{R}^d\big)$, then the above functional exactly has the form

$$\int_\Omega W(\mathbf{x},\nabla\mathbf{y}(\mathbf{x}))\,\mathrm{d}x - <\mathbf{f},\mathbf{y}>_{\left(\mathrm{W}^{1,r}_0(\Omega;\mathbb{R}^d)\right)^*} - <\mathbf{g},\mathbf{y}>_{\mathrm{L}_2(\Gamma_N;\mathbb{R}^d)},$$

Let us repeat the process to establish the connection between PDE's and finding minimizers. If $\mathbf{y}$ is the minimizer of I, then $I(t\boldsymbol{\varphi}+\mathbf{y}) \geq I(\mathbf{y}), \forall t, \boldsymbol{\varphi}$. If we denote

$$\iota(t) := I(t\boldsymbol{\varphi}+\mathbf{y}),$$

then it most hold (this corresponds to the calculation of $\delta I(\mathbf{y},\varphi)$)

$$0 = \iota'(0) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\int_\Omega W(\mathbb{F}+t\nabla\boldsymbol{\varphi})\,\mathrm{d}\mathbf{x} - \int_\Omega \mathbf{f}(\mathbf{x})\cdot(\mathbf{y}(\mathbf{x})+t\boldsymbol{\varphi}(\mathbf{x}))\,\mathrm{d}\mathbf{x} - \int_{\Gamma_N}\mathbf{g}(\mathbf{x})\cdot(\mathbf{y}(\mathbf{x})+t\boldsymbol{\varphi}(\mathbf{x}))\,\mathrm{d}x\right)\bigg|_{t=0},$$

calculating the derivatives yields

$$\begin{aligned}
0 &= \int_\Omega \frac{\partial W(\mathbb{F})}{\partial\mathbb{F}} : \nabla\boldsymbol{\varphi}\,\mathrm{d}\mathbf{x} - \int_\Omega \mathbf{f}\cdot\boldsymbol{\varphi}\,\mathrm{d}\mathbf{x} - \int_{\Gamma_N}\mathbf{g}\cdot\boldsymbol{\varphi}\,\mathrm{d}S = \\
&= \int_\Omega \frac{\partial}{\partial x_j}\left(\frac{\partial W(\mathbb{F})}{\partial F_{ij}}\varphi_i\right)\mathrm{d}\mathbf{x} - \int_\Omega \frac{\partial}{\partial x_j}\left(\frac{\partial W(\mathbb{F})}{\partial F_{ij}}\right)\varphi_i\,\mathrm{d}\mathbf{x} - \int_\Omega \mathbf{f}\cdot\boldsymbol{\varphi}\,\mathrm{d}\mathbf{x} - \int_{\Gamma_N}\mathbf{g}\cdot\boldsymbol{\varphi}\,\mathrm{d}S = \\
&= \int_{\Gamma_N} \frac{\partial W(\mathbb{F})}{\partial F_{ij}}\varphi_i n_j\,\mathrm{d}S - \int_\Omega \frac{\partial}{\partial x_j}\left(\frac{\partial W(\mathbb{F})}{\partial F_{ij}}\right)\varphi_i\,\mathrm{d}\mathbf{x} - \int_\Omega \mathbf{f}\cdot\boldsymbol{\varphi}\,\mathrm{d}\mathbf{x} - \int_{\Gamma_N}\mathbf{g}\cdot\boldsymbol{\varphi}\,\mathrm{d}S\,,
\end{aligned}$$

so it must hold

$$-\frac{\partial}{\partial x_j}\left(\frac{\partial W(\mathbb{F})}{\partial F_{ij}}\right) = f_i \text{ in}\,\Omega,\ \frac{\partial W(\mathbb{F})}{\partial F_{ij}} n_j = g_i \,\text{on}\,\Gamma_N.$$

This is exactly

$$-\nabla\cdot\mathbb{T} = \mathbf{f}\,\text{in}\,\Omega, \mathbb{T}\mathbf{n} = \mathbf{g},\ \text{on}\,\Gamma_N.$$

This implies that $\mathbf{y}$ minimizes energy $\Rightarrow$ $\mathbf{y}$ solves the equations of classical mechanics.

4.5.1 Nonconvexity of W

A wonderful result would be to show $\mathbf{y}\mapsto W(x,\nabla\mathbf{y})$ is convex, as then I would be (with some additional conditions like Caratheodority of $\mathbf{y}$) weakly sequentially lower semicontinuous and our analysis would be much simpler. Is it however possible?

Recall

$$W(\mathbb{I}) = 0 \Rightarrow W(\mathbb{R}) = 0, \forall \mathbb{R}\in \mathrm{SO}(3)$$

and $W(\mathbb{F}) > 0$ whenever $\mathbb{F}\notin \mathrm{SO}(3)$ Assume now

$$\mathbb{R}_1 = \begin{bmatrix}1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & -1\end{bmatrix}, \mathbb{R}_2 = \mathbb{I},$$

and realize

$$\frac{1}{4}\mathbb{R}_1 + \frac{3}{4}\mathbb{R}_2 = \frac{1}{4}\begin{bmatrix}4 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 2\end{bmatrix},$$

so $\det\left(\frac{1}{4}\mathbb{R}_1 + \frac{3}{4}\mathbb{R}_2\right) = \frac{1}{4} \neq 1$. But this means

$$0 \neq W\left(\frac{1}{4}\mathbb{R}_1 + \frac{3}{4}\mathbb{R}_2\right) > \frac{1}{4}W(\mathbb{R}_1) + \frac{3}{4}W(\mathbb{R}_2) = 0,$$

as both terms on the RHS are zero. Hence, we have checked W is *not* convex and that complicates things:

Example (Minimizer does not exist)**.** Assume

$$J(u) = \int_0^1 \left(1-\left(u'(x)\right)^2\right)^2 + u(x)^2\,\mathrm{d}x\,, u\in \mathrm{W}^{1,4}(0,1), u(0) = u(1) = 0,$$

and find the minimum of J. First of all, $J > 0$, so $\inf J >\geq 0 \neq -\infty$. I can take u_k such that

$$u_k'(x) = \begin{cases}1, & \text{on}\,(0,1/2^k)\\ -1, & \text{on}\,(1/2^k, 1/2^{k-1})\end{cases},$$

for $k\in\mathbb{N}$ (the function looks like a saw). Then $J(u_k)\to 0 \Rightarrow \inf J = 0$ but there is no minimizer, because $\{u\}_k$ does not converge[8] in $\mathrm{W}^{1,4}((0,1))$.

Not everything is lost...

[8] We see $\lim_{k\to\infty}\|u_k\|_{\mathrm{L}_4((0,1))} = 0$, but $\lim_{k\to\infty}\|u_k'\|_{\mathrm{L}_4((0,1))} = 1$.

4.5.2 Polyconvexity

Definition 10 (Polyconvexity, 1977 J.M. Ball)**.** $W : \mathbb{R}^{3\times3} \to \mathbb{R} \cup \{\infty\}$ is polyconvex provided there exists convex and lower-semicontinuous function $h : \mathbb{R}^{3\times3} \times \mathbb{R}^{3\times3} \times \mathbb{R} \to \mathbb{R} \cup \{\infty\}$:

$$W(\mathbb{A}) = h(\mathbb{A}, \operatorname{cof}\mathbb{A}, \det\mathbb{A}).$$

Example.
- If W is convex and lower-semicontinuous then W is polyconvex.
- $W(\mathbb{A}) = \det\mathbb{A}$ is polyconvex but not convex.

Remark (Weak convergence in $\mathrm{L_p}(\Omega;\mathbb{R}^3)$). Let $1 < p < \infty$ and $\{\mathbf{u_k}\} \subset \mathrm{L_p}(\Omega;\mathbb{R}^3)$. We say $\{\mathbf{u_k}\}$ converges weakly to $\mathbf{u}$ in $\mathrm{L_p}(\Omega;\mathbb{R}^3)$ provided

$$\int_\Omega \mathbf{u_k}\cdot\boldsymbol{\varphi}\,\mathrm{d}\mathbf{x} \to \int_\Omega \mathbf{u}\cdot\boldsymbol{\varphi}\,\mathrm{d}\mathbf{x}\,, \forall\boldsymbol{\varphi} \in \mathrm{L}_{p'}(\Omega;\mathbb{R}^3)\,.$$

Theorem 2 (Magic)**.** *Assume that* $\mathbf{y}_k$ *converges weakly to* $\mathbf{y}$ *in* $W^{1,p}(\Omega;\mathbb{R}^3)\,, \Omega \subset \mathbb{R}^3 \in C^{0,1}, p > 3$. *Then* $\det\nabla\mathbf{y}_k$ *converges weakly to* $\det\nabla\mathbf{y}$ *in* $L_{\frac{p}{3}}(\Omega)$. *Moreover* $\operatorname{cof}\nabla\mathbf{y}_l$ *converges weakly to* $\operatorname{cof}\nabla\mathbf{y}$ *in* $L_{\frac{p}{2}}(\Omega;\mathbb{R}^{3\times3})$.

Proof. Only in 2 dimensions. The determinant can be written as:

$$\det\nabla\mathbf{y} = \frac{\partial y_1}{\partial x_1}\frac{\partial y_2}{\partial x_2} - \frac{\partial y_1}{\partial x_2}\frac{\partial y_2}{\partial x_1} = \frac{\partial}{\partial x_1}\left(y_1\frac{\partial y_2}{\partial x_2}\right) - \frac{\partial}{\partial x_2}\left(y_1\frac{\partial y_2}{\partial x_1}\right) = \nabla\cdot\left(y_1, -\frac{\partial y_2}{\partial x_1}\right),$$

so then

$$\int_\Omega \det\nabla\mathbf{y}^k\varphi\,\mathrm{d}x = \int_\Omega \frac{\partial}{\partial x_1}\left(y_1^k\frac{\partial y_2^k}{\partial x_2}\right)\varphi\,\mathrm{d}x - \int_\Omega \frac{\partial}{\partial x_2}\left(y_1^k\frac{\partial y_2^k}{\partial x_1}\right)\varphi\,\mathrm{d}x = -\int_\Omega y_1^k\frac{\partial y_2^k}{\partial x_2}\frac{\partial\varphi}{\partial x_1}\,\mathrm{d}x + \int_\Omega \frac{\partial\varphi}{\partial x_2}y_1^k\frac{\partial y_2^k}{\partial x_1}\,\mathrm{d}x\,,$$

and the result follows from the embedding theorems and strong convergence (strong times weak gives weak convergence). □

The above theorem is truly remarkable, as in general nonlinearities destroy weak convergence.

Let us recall some concepts from analysis

Definition 11 ((Weak) sequential lower semicontinuity)**.** A function $f : X \to \mathbb{R}, X$ a Banach space, is said to be

- sequentially lower semicontious, provided

$$f(y) \le \liminf_{k\to\infty} f(y_k), \forall\{y_k\} \subset X : u_k \to u \in X,$$

- weakly sequentially lower semicontinuous, provided

$$f(y) \le \liminf_{k\to\infty} f(y_k), \forall\{y_k\} \subset X : u_k \rightharpoonup u \in X.$$

Definition 12 (Direct method)**.** The direct method of the calculus of variation is the following scheme: suppose $\Omega \subset \mathbb{R}^d, \Omega \in C^{0,1}, p \in (1,\infty)$ and the functional $I : \mathrm{W}^{1,p}(\Omega;\mathbb{R}^d) \to \mathbb{R} \cup \{\infty\}$ is weakly sequentially lower semicontinuous and coercive.

1. from coercivity one has $\inf I > -\infty$, so $\exists\{\mathbf{y}_k\} \subset \mathrm{W}^{1,p}(\Omega;\mathbb{R}^d)$ such that
$$\inf I = \lim_{k\to\infty} I(\mathbf{y}_k),$$
2. using coercivity, one shows $\|\mathbf{y}_k\|_{\mathrm{W}^{1,p}(\Omega;\mathbb{R}^d)} \leq C$, an uniform bound,
3. for $p \in (1,\infty)$ the space $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^d)$ is a reflexive Banach space - $\exists\{\mathbf{y}_{n_k}\} \subset \{\mathbf{y}_k\}$ weakly convergent,
4. from WSLSC one infers:
$$\inf I = \lim_{k\to\infty} I(\mathbf{y}_k) \geq \liminf_{k\to\infty} I(\mathbf{y}_k) \geq I(\mathbf{y}),$$
so
$$\inf I = I(\mathbf{y}).$$

One important sufficient condition for the functional
$$I(\mathbf{y}) = \int_\Omega W(\nabla\mathbf{y})\,\mathrm{d}x\,,$$
to be WSLSC is when W is (sequentially) lower semicontinuous and convex. In the case of polyconvexity, one has
$$I(\mathbf{y}) = \int_\Omega W(\nabla\mathbf{y})\,\mathrm{d}x = \int_\Omega h(\nabla\mathbf{y},\det\nabla\mathbf{y},\operatorname{cof}\nabla\mathbf{y})\,\mathrm{d}x\,,$$
and from our above result, we also know that if $\mathbf{y}_k \rightharpoonup \mathbf{y}$ in some Sobolev space, also $\det\nabla\mathbf{y}_k \rightharpoonup \det\nabla\mathbf{y}, \operatorname{cof}\nabla\mathbf{y}_k \rightharpoonup \operatorname{cof}\nabla\mathbf{y}$ (in some Lebesgue spaces) and h is convex, which is enough for WSLSC of I, so one can use the direct method; details are presented in Kružík and Roubíček, 2019b.

4.5.3 Rank-one convexity

Our last remark will be about an even weaker notion of convexity - the so called **rank-one convexity**. The definition is

Definition 13. The function $W : \mathbb{R}^{n\times m} \to \mathbb{R}$ is said to be rank-1 convex, provided $\forall\lambda \in [0,1], \forall\mathbb{A},\mathbb{B} \in \mathbb{R}^{n\times m} : \operatorname{rank}(\mathbb{A}-\mathbb{B}) \leq 1$ it holds
$$W(\lambda\mathbb{A} + (1-\lambda)\mathbb{B}) \leq \lambda W(\mathbb{A}) + (1-\lambda)W(\mathbb{B}).$$

Remark. One could say a rank-1 convex function is a function that is convex on a restricted set of arguments.

The set of matrices that have rank at most 1 is isomorphic with the set $\{\mathbf{a}\otimes\mathbf{b}|\mathbf{a},\mathbf{b} \in \mathbb{R}^n\}$. But this means[9]
$$W \text{ is rank-1 convex} \Leftrightarrow t \mapsto W(\mathbb{A} + t\mathbf{a}\otimes\mathbf{b}) \text{ is convex}\,, \forall\mathbb{A} \in \mathbb{R}^n, \mathbf{a},\mathbf{b} \in \mathbb{R}^n,$$
and since $\mathbb{A},\mathbf{a},\mathbf{b}$ are arbitrary, this condition can be rephrased as that the mapping is convex at $t = 0$. Using some analysis, this means

[9] The first implication is easy, as of course the argument of $W(t_1 + -t_2\mathbf{a}\otimes\mathbf{b})$ has rank at most 1. For the other direction, realize that the mapping between the sets truly is an isomorphism and that we require the convexity for all matrices $\mathbb{A}$ and all vectors $\mathbf{a},\mathbf{b}$.

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}(W(\mathbb{A}+t\mathbf{a}\otimes\mathbf{b}))\bigg|_{t=0} = \left(\frac{\partial^2 W}{\partial \mathbb{A}^2}\mathbf{a}\otimes\mathbf{b}\right):\mathbf{a}\otimes\mathbf{b} = \frac{\partial^2 W}{\partial A_{ij}\partial A_{kl}}a_i b_j a_k b_l \geq 0,$$

which is known as the **Legendre-Hadamard** conditions.

In general, only rank-one convexity is not enough to obtain existence of minimizers. Moreover, one has the following implications and none of those can be reversed

$$\text{convexity} \Rightarrow \text{polyconvexity} \Rightarrow \text{rank-one convexity}\,.$$

4.5.4 Ensuring injectivity

Finally, the fact $\det \mathbb{F} > 0$ is not our only constraint; we also require $\mathbf{y}$ is injective (almost everywhere). To see this need not be implied we present the example:

Example. Assume the following domain: $\Omega = (1,2)\times(0,4\pi)\times(1,2)$ and the deformation $\mathbf{y}:\overline{\Omega}\to\mathbb{R}^3, \mathbf{y}(x_1,x_2,x_3) = (x_1\cos x_2, x_1\sin x_2, x_3), \mathbb{F} = \begin{bmatrix}\cos x_2 & -x_1\sin x_2 & 0\\ \sin x_2 & x_1\cos x_2 & 0\\ 0 & 0 & 1\end{bmatrix}$. We can calculate $\det \mathbb{F} = x_1\cos^2 x_2 + x_1\sin^2 x_2 = x_1$. But even though the deformation has positive determinant, we still face self-penetration issues, i.e., $\mathbf{y}$ *is not injective.*

There exists a simple condition under which the mapping $\mathbf{y}$ is injective *a.e.*.

Theorem 3 (Ciarlet-Nečas condition)**.** *Let* $p>3$ *and let* $\det \mathbb{F} > 0$ *a.e. in* $\Omega\subset\mathbb{R}^3, \mathbf{y}\in W^{1,p}\big(\Omega;\mathbb{R}^3\big)$. *If*

$$\int_\Omega \det \mathbb{F}(\mathbf{x})\,\mathrm{d}\mathbf{x} \leq \lambda(\mathbf{y}(\Omega))$$

then $\mathbf{y}$ *is injective almost everywhere in* Ω*, i.e.,* $\exists\omega\subset\Omega:\lambda(\omega)=0, \mathbf{y}|_{\Omega/\omega}$ *is injective.*

Is the determinant condition of any use? Let us compute, assuming $\mathbf{y}=\mathbf{0}$ on $\partial\Omega$.

$$\int_\Omega \det \mathbb{F}\,\mathrm{d}\mathbf{x} = \int_\Omega \frac{\partial}{\partial x_1}\left(y_1\frac{\partial y_2}{\partial x_2}\right) - \frac{\partial}{\partial x_2}\left(y_1\frac{\partial y_2}{\partial x_1}\right)\mathrm{d}\mathbf{x} = \int_{\partial\Omega} y_1\frac{\partial y_2}{\partial x_2}n_1 - y_1\frac{\partial y_2}{\partial x_1}n_2\,\mathrm{d}S \underbrace{\Rightarrow}_{y=0\text{ on }\partial\Omega} \int_\Omega \det \mathbb{F}(\mathbf{x})\,\mathrm{d}\mathbf{x} = 0.$$

This is powerful! Assume that $\mathbf{y}(\mathbf{x}) = \mathbb{F}\mathbf{x}$ on $\partial\Omega$, then

$$\int_\Omega \det \mathbb{F}(\mathbf{x})\,\mathrm{d}\mathbf{x} = \lambda(\Omega)\det \mathbb{F}.$$

Now let

$$I(\mathbf{y}) = \int_\Omega \det \mathbb{F}(\mathbf{x})\,\mathrm{d}\mathbf{x}\,, \mathbf{y}\in \mathrm{W}^{1,p}\big(\Omega;\mathbb{R}^3\big)\,, \mathbf{y}(\mathbf{x}) = \mathbb{F}\mathbf{x}.$$

Then I is constant[10] and it holds

$$I(\mathbf{y}) = \lambda(\Omega)\det \mathbb{F}.$$

5 Linearized elasticity

Recall the Right Cauchy-Green tensor: $\mathbb{C} = \mathbb{F}^\top\mathbb{F}$. Using it, we can define

Theorem 4 (Green-Lagrange strain tensor/Green-St.-Venaint strain tensor)**.** *Let* $\mathbb{C}$ *be the Right Cauchy-Green tensor. We define the Green-Lagrange strain tensor/Green-St.-Venaint strain*

[10] All constant functionals are convex.

tensor as

$$\mathbb{E} = \frac{1}{2}(\mathbb{C} - \mathbb{I}).$$

Remark. The Green-St.-Venaint strain tensor can be rewritten as:

$$\mathbb{E} = \frac{1}{2}\Big((\mathbb{I} + \nabla\mathbf{u})^\top(\mathbb{I} + \nabla\mathbf{u}) - \mathbb{I}\Big) = \frac{1}{2}\Big(\nabla\mathbf{u} + (\nabla\mathbf{u})^\top\Big) + \frac{1}{2}(\nabla\mathbf{u})^\top\nabla\mathbf{u} = \mathbb{e}(\mathbf{u}) + \frac{1}{2}\mathbb{C}(\nabla\mathbf{u}),$$

where we have denoted

$$\mathbb{e}(\mathbf{u}) = \frac{1}{2}\Big(\nabla\mathbf{u} + (\nabla\mathbf{u})^\top\Big) = \operatorname{sym}(\nabla\mathbf{u}),$$

and

$$\mathbb{C}(\nabla\mathbf{u}) = (\nabla\mathbf{u})^\top\nabla\mathbf{u}.$$

This quantities will prove useful to us when linearizing.

We have shown the stored energy density can in fact be represented as a function of the Right Cauchy-Green tensor:

$$W(\mathbb{F}) = \overline{W}(\mathbb{C}),$$

and since $\mathbb{C} = 2\mathbb{E} + \mathbb{I}$, one can also think of the stored energy density as a function of the Green-St.Venaint strain tensor

$$W(\mathbb{F}) = \hat{W}(\mathbb{E}) = \hat{W}\Big(\mathbb{e}(\mathbf{u}) + \frac{1}{2}\mathbb{C}(\nabla\mathbf{u})\Big).$$

The theory of linearized elasticity is the theory in which one supposes

$$\mathbb{E} \approx \mathbb{e} = \operatorname{sym}(\nabla\mathbf{u}),$$

meaning that the term $\mathbb{C}(\nabla\mathbf{u}) = (\nabla\mathbf{u})^\top\nabla\mathbf{u}$ that is quadratic in the displacement gradient *can be considered small*. Seen as this, the theory of linearized elasticity is in fact **a theory of small strain** or of **small displacement gradient.**[11] From now on, the energy density will hence be taken just as a function of the symmetric part of the strain tensor, *i.e.*, the linearized strain tensor:

$$W(\mathbb{F}) \equiv \hat{W}(\mathbb{e}(\mathbf{u})).$$

Before we linearize our stored energy density, let us investigate the case $\mathbb{E} = \mathbb{O}$. Then $\mathbb{O} = \frac{1}{2}(\mathbb{C} - \mathbb{I})$, so $\mathbb{C} = \mathbb{I}$ and from what we have seen in the section 2.4, this means

$$\mathbb{F} = \mathbb{R} \in \mathrm{SO(d)}\,.$$

But from the property of W it then follows

$$\hat{W}(\mathbb{E} = \mathbb{O}) = W(\mathbb{R}) = 0,$$

and that also $\hat{W}(\mathbb{e}) > 0$ for $\mathbb{e} \neq \mathbb{O}$. *Stored energy density has a (unstrict) global minimum at zero.* We are ready to use Taylor expansion:

$$\hat{W}(\mathbb{e}(\mathbf{u})) = \hat{W}(\mathbb{O}) + \frac{\partial\hat{W}}{\partial\mathbb{e}}(\mathbb{O}) : \mathbb{e}(\mathbf{u}) + \frac{1}{2}\frac{\partial^2\hat{W}}{\partial\mathbb{e}^2}(\mathbb{O})\mathbb{e}(\mathbf{u}) : \mathbf{e}(\mathbf{u}) + \text{h.o.t.}\,,$$

[11] This is not a bad consideration - the displacement gradient (or strain) is dimensionless, so it makes sense to talk about smallness independent of the physical units (consider for example the "theory of small displacement" - but the unit of displacement is an unit of length, and the notion of smallness would thus depend on the choice of units.

since $\hat{W}(\mathbb{0}) = 0$, and $\frac{\partial \hat{W}}{\partial \mathbb{e}}(\mathbb{0}) = \mathbb{0}$ because $\hat{W}$ has a minimum at $\mathbb{e} = \mathbb{0}$. The above (formal) manipulation leads us to the definition

Definition 14 (Tensor of elastic constants)**.**

$$\mathcal{C} = \frac{\partial^2 \hat{W}}{\partial \mathbb{e}^2}(\mathbb{0}), C_{ijkl} = \frac{\partial^2 \hat{W}}{\partial e_{ij} \partial e_{kl}}.$$

Remark. Since we assume $\hat{W}$ is smooth, we have some symmetries, and from the general 81 components of C_{ijkl} only 21 are unique.

With the notion of a tensor of elastic constants, we can then write the (linearized) stored energy density as

$$w(\mathbb{e}) = \frac{1}{2}(\mathcal{C}\mathbb{e}) : \mathbb{e},$$

(we are using small letters to distinguish the linearized quantites). Following our definition $\mathbb{T} = \frac{\partial W}{\partial \mathbb{F}}$ we have

$$\mathbb{T} = \frac{\partial W}{\partial \mathbb{F}} = \frac{\partial \mathbb{e}}{\partial \mathbb{F}} \frac{\partial w(\mathbb{e})}{\partial \mathbb{e}},$$

with

$$\frac{\partial \mathbb{e}}{\partial \mathbb{F}} = \frac{\partial}{\partial \mathbb{F}} \left(\frac{1}{2} \left(\mathbb{F} + \mathbb{F}^\top \right) - \mathbb{I} \right) = \mathcal{I}_{\text{sym}},$$

where $\mathcal{I}_{\text{sym}}$ has the property[12] $\mathcal{I}_{\text{sym}} \mathbb{A} = \text{sym}(\mathbb{A})$. Also

$$\frac{\partial w(\mathbb{e})}{\partial \mathbb{e}} = \frac{1}{2} \frac{\partial}{\partial \mathbb{e}} (\mathcal{C}\mathbb{e} : \mathbb{e}) = \frac{1}{2} (\mathcal{C}\mathcal{I} : \mathbb{e} + \mathcal{C}\mathbb{e} : \mathcal{I}) = \mathcal{C}\mathbb{e}$$

and finally

$$\mathcal{I}_{\text{sym}} \mathcal{C}\mathbb{e} = \mathcal{C}\mathbb{e},$$

because $\mathcal{C}\mathbb{e}$ is symmetric thanks to the property

$$C_{ijkl} = \frac{\partial^2 W}{\partial e_{ij} \partial e_{kl}} = \frac{\partial^2 W}{\partial e_{ji} \partial e_{kl}} = C_{jikl}.$$

So we have obtained $\mathbb{T} = \frac{\partial w}{\partial \mathbb{e}} = \mathcal{C}\mathbb{e}$ which is usually denoted as the *Cauchy stress* $\mathbb{\sigma}$

$$\mathbb{\sigma} = \frac{\partial w(\mathbb{e})}{\partial \mathbb{e}} = \mathcal{C}\mathbb{e},$$

or in components

$$\sigma_{ij} = C_{ijkl} e_{kl}.$$

5.1 Equations

Rewritting the equations in the linearized elasticity setting we obtain the system

$$\begin{aligned} -\nabla \cdot \mathbb{\sigma} = -\nabla \cdot (\mathcal{C}\mathbb{e}) &= \mathbf{f} \text{ in } \Omega \\ \mathbb{\sigma}\mathbf{n} &= \mathbf{g} \text{ on } \Gamma_N, \\ \mathbf{u} &= \mathbf{0} \text{ on } \Gamma_D. \end{aligned}$$

[12] $\mathcal{I}_{\text{sym}\,ijkl} = \frac{1}{2}(\delta_{ik}\delta_{jl} + \delta_{jk}\delta_{il})$.

The weak formulation can be obtained in the following way:

$$\int_\Omega -\nabla\cdot(\mathcal{C}\mathbb{e})\cdot\boldsymbol{\varphi}\,\mathrm{d}\mathbf{x} = \int_\Omega \mathbf{f}\cdot\boldsymbol{\varphi}\,\mathrm{d}\mathbf{x}\,,\forall\boldsymbol{\varphi}\in X = \left\{\mathrm{W}^{1,2}\left(\Omega;\mathbb{R}^d\right), u = 0\,\mathrm{on}\,\Gamma_D\right\},$$

so integrating per partes in the first term gives us

$$\int_\Omega -\nabla\cdot(\mathcal{C}\mathbb{e})\cdot\boldsymbol{\varphi}\,\mathrm{d}\mathbf{x} = -\int_{\Gamma_N}\boldsymbol{\varphi}\cdot\mathcal{C}\mathbb{e}\mathbf{n}\,\mathrm{d}S + \int_\Omega \mathcal{C}\mathbb{e}:\nabla\boldsymbol{\varphi}\,\mathrm{d}\mathbf{x}\,,$$

since the test function are zero on the ”Dirichlet part of the boundary.” Using the Neumann BC and using also the fact $\mathrm{sym}\,(\mathcal{C}\mathbb{e}) = \mathcal{C}\mathbb{e}$ leads to

$$\int_\Omega -\nabla\cdot(\mathcal{C}\mathbb{e})\cdot\boldsymbol{\varphi}\,\mathrm{d}\mathbf{x} = -\int_{\Gamma_N}\mathbf{g}\cdot\boldsymbol{\varphi}\,\mathrm{d}S + \int_\Omega \mathcal{C}\mathbb{e}:\mathrm{sym}\,(\nabla\boldsymbol{\varphi})\,\mathrm{d}\mathbf{x}\,,$$

and so our weak formulation indeed becomes

$$\forall\boldsymbol{\varphi}\in X:\int_\Omega \mathcal{C}\mathbb{e}:\mathrm{sym}\,(\nabla\boldsymbol{\varphi})\,\mathrm{d}\mathbf{x} = \int_\Omega \mathbf{f}\cdot\boldsymbol{\varphi}\,\mathrm{d}\mathbf{x} + \int_{\Gamma_N}\mathbf{g}\cdot\boldsymbol{\varphi}\,\mathrm{d}S\,.$$

which has a nice structure if one realizes $\mathbb{e} = \mathrm{sym}\,(\nabla\mathbf{u})$, and we are solving for $\mathbf{u}$:

$$\int_\Omega \mathcal{C}\,\mathrm{sym}\,(\nabla\mathbf{u}):\mathrm{sym}\,(\nabla\boldsymbol{\varphi})\,\mathrm{d}\mathbf{x} = \int_\Omega \mathbf{f}\cdot\boldsymbol{\varphi}\,\mathrm{d}\mathbf{x} + \int_{\Gamma_N}\mathbf{g}\cdot\boldsymbol{\varphi}\,\mathrm{d}S\,.$$

This is a linear elliptic equation - to show existence & uniqueness, it suffices to show that

- the bilinear form $B: X\times X\to\mathbb{R}$
$$B(\mathbf{u},\boldsymbol{\varphi}) = \int_\Omega \mathcal{C}\,\mathrm{sym}\,(\nabla\mathbf{u}):\mathrm{sym}\,(\nabla\boldsymbol{\varphi})\,\mathrm{d}\mathbf{x}\,,$$
is X−bounded and X−coercive,
- and the right hand side $F: X\to\mathbb{R}$
$$F(\boldsymbol{\varphi}) = \int_\Omega \mathbf{f}\cdot\boldsymbol{\varphi}\,\mathrm{d}\mathbf{x} + \int_{\Gamma_N}\mathbf{g}\cdot\boldsymbol{\varphi}\,\mathrm{d}S$$
is an element of X^*,

as then Lax-Milgram does the whole job for us. However, in order to show this, we would have to be able to *control the norm* of $\nabla\mathbf{u}$ just by the norm of $\mathrm{sym}\,(\nabla\mathbf{u})$. Is that even possible?

5.1.1 Korn’s inequality

Example. Suppose $\mathbf{u}\in\mathcal{D}(\Omega;\mathbb{R}^n)$. Then the question is:

$$\exists C>0:\int_\Omega|\,\mathrm{sym}\,(\nabla\mathbf{u})\,|^2\,\mathrm{d}\mathbf{x}\geq C\int_\Omega|\nabla\mathbf{u}|^2\,\mathrm{d}\mathbf{x}?$$

Can this hold? Make a quick test: Take $\mathbf{u}$ such that $\nabla\mathbf{u} = \mathrm{asym}\,(\nabla\mathbf{u})$, *i.e.*

$$\nabla\mathbf{u} = -(\nabla\mathbf{u})^\top.$$

This in particular must mean that

$$\mathrm{tr}\,\nabla\mathbf{u} = \nabla\cdot\mathbf{u} = 0,$$

in Ω. But then

$$0 = \int_\Omega \varphi(\nabla\cdot\mathbf{u})\,\mathrm{d}\mathbf{x} = -\int_\Omega \mathbf{u}(\nabla\cdot\varphi)\,\mathrm{d}\mathbf{x}\,,$$

for all $\varphi \in \mathcal{D}\big(\Omega;\mathbb{R}^d\big)$, and so

$$\mathbf{u} = \mathbf{0}\ a.e.\ \text{in}\,\Omega.$$

which *does not disprove the above inequality* - it in fact shows

$$\text{sym}\,(\nabla\mathbf{u}) = \mathbb{0} \Rightarrow \mathbf{u} = 0,$$

meaning that even in the "worst case" when the symmetric part of the gradient is zero, the whole norm can still *be controlled by it.*

Let us try to understand why this is the case; compute

$$\begin{aligned}\int_\Omega |\mathbb{e}(\mathbf{u})|^2\,\mathrm{d}x &= \frac{1}{4}\int_\Omega\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\right)\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\right)\mathrm{d}x\\ &= \frac{1}{4}\int_\Omega\left(\frac{\partial u_i}{\partial x_j}\right)^2+2\frac{\partial u_i}{\partial x_j}\frac{\partial u_j}{\partial x_i}+\left(\frac{\partial u_j}{\partial x_i}\right)^2\mathrm{d}x = \frac{1}{2}\int_\Omega\left(\frac{\partial u_i}{\partial x_j}\right)^2+\frac{\partial u_i}{\partial x_j}\frac{\partial u_j}{\partial x_i}\,\mathrm{d}x\,,\end{aligned}$$

where we used the symmetry property. Integrating by parts two times to obtain "$\partial_i u_i \partial_j u_j = (\partial_j u_j)^2$ [13]. All in all

$$\frac{1}{2}\int_\Omega\left(\frac{\partial u_i}{\partial x_j}\right)^2+\left(\frac{\partial u_i}{\partial x_i}\right)^2\mathrm{d}x \geq 0.$$

Theorem 5 (Korn's inequality)**.** *Let $\Omega \subset \mathbb{R}^n$ be a bounded Lipschitz domain. Then there exists $C > 0$ such that $\forall\mathbf{u} \in W^{1,2}((\Omega;\mathbb{R}^n))$ it holds*

$$\Big(\|\,sym\,(\nabla\mathbf{u})\|^2_{L_2\left((\Omega;\mathbb{R}^{n\times n})\right)}+\|\mathbf{u}\|^2_{L_2((\Omega;\mathbb{R}^n))}\Big) \geq C\|\mathbf{u}\|^2_{W^{1,2}((\Omega;\mathbb{R}^n))}.$$

Remark. Trivially, we have $\forall\mathbf{u} \in \mathrm{W}^{1,2}((\Omega;\mathbb{R}^n))$

$$\|\,\text{sym}\,(\nabla\mathbf{u})\|^2_{\mathrm{L}_2((\Omega;\mathbb{R}^n))}+\|\mathbf{u}\|^2_{\mathrm{L}_2((\Omega;\mathbb{R}^n))} \leq \|\mathbf{u}\|^2_{\mathrm{W}^{1,2}((\Omega;\mathbb{R}^n))}.$$

So Korn's inequality in fact tells us

$$\mathbf{u} \mapsto \|\,\text{sym}\,(\nabla\mathbf{u})\|^2_{\mathrm{L}_2((\Omega;\mathbb{R}^n))}+\|\mathbf{u}\|^2_{\mathrm{L}_2((\Omega;\mathbb{R}^n))}$$

is an equivalent norm on $\mathrm{W}^{1,2}((\Omega;\mathbb{R}^n))$.

We can in fact obtain something even nicer - as in the case of Poincare inequality, when under some conditions the seminorm $u \mapsto \|\nabla u\|^p_{\mathrm{L}_p(\Omega)}$ was a norm on $\mathrm{W}^{1,p}(\Omega)$, we would like that map (we in fact have shown it is a seminorm already) $\mathbf{u} \mapsto \|\,\text{sym}\,(\nabla\mathbf{u})\|_{\mathrm{L}_p(\Omega)}$ is a norm on $\mathrm{W}^{1,p}(\Omega;\mathbb{R}^n)$.

Definition 15 (Axial vectors)**.** Let $\mathbb{A} = -\mathbb{A}^\top, \mathbb{A} \in \mathbb{R}^{n\times n}$. Then there is $b \in \mathbb{R}^n$ such that $\mathbb{A}\mathbf{v} = \mathbf{b}\times\mathbf{v}, \forall\mathbf{v} \in \mathbb{R}^n$. The vector $\mathbf{b}$ is called the axial vector of $\mathbb{A}$.

Remark ($\mathbb{R}^n$)*.* This truly holds in $\mathbb{R}^n$, not only in $\mathbb{R}^3$. We only have to replace $\times$ by $\wedge$, the outter product.

Assume that $\mathbf{u} \in C^2\big(\Omega;\mathbb{R}^3\big)$. Then

$$\frac{\partial^2 u_i}{\partial x_j \partial x_k} = \frac{\partial e_{ik}}{\partial x_j}(\mathbf{u})+\frac{\partial e_{ij}}{\partial x_k}(\mathbf{u})-\frac{\partial e_{jk}}{\partial x_i}(\mathbf{u}).$$

[13] Sign does not change as we integrate 2 times. Also, we have homogenous Dirichlet

If now $\mathbb{e}(\mathbf{u}) = \mathbb{0}$, then $\mathbf{u}$ is an affine function, because

$$\frac{\partial^2 u_i}{\partial x_j \partial x_k} = 0, \forall i, j, k \in \{1, 2, 3\}.$$

[14] It must thus hold

$$u_i(x) = a_i + b_{ij} x_j,$$

and $\frac{\partial u_i}{\partial x_j} = b_{ij} = -b_{ji}$, because $\mathbb{e}(\mathbf{u}) = \mathbb{0}$, so it must be skew symmetric. The skew-symmetry also means it can be written as

$$\mathbf{u}(\mathbf{x}) = \mathbf{a} + \mathbf{d} \times \mathbf{x}.$$

If additionaly we assume that $\mathbf{u} = \mathbf{0}$ on some $\Gamma_D \subset \partial\Omega, \mathcal{H}(\Gamma_D) > 0$ and $\mathbf{u}(\mathbf{x}) = \mathbf{a} + \mathbf{d} \times \mathbf{x}$, then $\mathbf{u} = \mathbf{0}$ identically in Ω: the function is determined by two parameters $\mathbf{a}, \mathbf{d}$, which are however fixed on Γ_D. For example, pick $\mathbf{x}, \mathbf{y} \in \Gamma_D$, then $\mathbf{0} = \mathbf{a} + \mathbf{d} \times \mathbf{y} = \mathbf{a} + \mathbf{d} \times \mathbf{x}$, so

$$\mathbf{d} \times (\mathbf{x} - \mathbf{y}) = 0.$$

But since $\mathbf{x}, \mathbf{y}$ can be chosen arbitrarily (linearly indepndelty), this must mean $\mathbf{d} = \mathbf{0}$ and consequently $\mathbf{a} = \mathbf{0}$.

Finally, we have shown that

$$\mathbf{u} \mapsto \|\mathbb{e}(\mathbf{u})\|_{\mathrm{L}_2((\Omega;\mathbb{R}^{n\times n}))}$$

is a norm on

$$V = \{\mathbf{w} \in \mathrm{W}^{1,2}\left((\Omega; \mathbb{R}^3)\right), \mathbf{w} = \mathbf{0} \,\mathrm{on}\, \Gamma_D\}$$

which is equivalent to the norm of $\mathrm{W}^{1,2}\left((\Omega; \mathbb{R}^3)\right)$.

Coming back to our equation

$$B(\mathbf{u}, \boldsymbol{\varphi}) = F(\boldsymbol{\varphi}), \forall \boldsymbol{\varphi} \in X,$$

we have showed everything to use Lax-Milgram $\Rightarrow \exists! u \in V$. This also means the functional

$$I(\mathbf{u}) = \int_\Omega \frac{1}{2} \mathcal{C} \,\mathrm{sym}\,(\nabla \mathbf{u}) : \mathrm{sym}\,(\nabla \mathbf{u}) - F(\mathbf{u})\, \mathrm{d}\mathbf{x}, \forall u \in X.$$

has an unique minimizer.

6 Convex analysis

Convex functions have some remarkable qualities, which we will explore a bit. Throughout the chapter, we work with $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ that are convex. All the theory could be developed above a general Banach space. In our applications however we will be dealing mostly with the space of matrices $\mathbb{R}^{n\times n}$, so the notation is acommodated to that.

Definition 16 (Epigraph of a set)**.** The epigraph of a function f is the set

$$\mathrm{epi}\, f = \{(x, y) : y \geq f(x)\}$$

Remark. With the notion of $\mathrm{epi}\, f$ we can work with sets instead of functions. Moreover, it holds

- $\mathrm{epi}\, f$ is closed $\Leftrightarrow$ f is lower-semicontinuous,
- f is convex $\Leftrightarrow$ $\mathrm{epi}\, f$ is convex

[14] Recall that Ω is simply connected.

From one of the consequences of Hahn-Banach theorem (oddělovací věty), we obtain the existence of such $\xi \in \mathbb{R}^n$ (dependent of x) that for fixed $x \in \mathbb{R}^n$ it yields

$$f(z) \geq f(x) + \xi(x) \cdot (z - x), \forall z \in \mathbb{R}^n.$$

If f is differentiable at x, then

$$\xi(x) = \nabla f(x).$$

But in general the derivative does not have to exist. This motivates the following definition.

Definition 17 (Subgradient, subdifferential)**.** Let $f : \mathbb{R}^{n\times m} \to \mathbb{R} \cup \{\infty\}$ be convex and lower semicontinuous and let $\mathbb{X}$ be fixed, but arbitrary. The function $\mathbb{\xi}(\mathbb{X}) : \mathbb{R}^{n\times m} \to \mathbb{R}^{n\times m}$ such that it holds

$$f(\mathbb{Y}) \geq f(\mathbb{X}) + \mathbb{\xi}(\mathbb{X}) : (\mathbb{Y} - \mathbb{X}), \forall \mathbb{Y} \in \mathbb{R}^{n\times m},$$

is called the subgradient of f at $\mathbb{X}$. The **set** of all subgradients of f at $\mathbb{X}$ is called the subdifferential and denoted $\partial f(\mathbb{X})$.

Remark.
- If $\partial f(\mathbb{X})$ is a singleton, then $\nabla f(\mathbb{X})$ exist.
- $\partial f(\mathbb{X})$ is convex
- $\mathbb{0} \in \partial f(\mathbb{X}) \forall \mathbb{X} \in \mathbb{R}^{n\times m}$ is a condition for the minimizer.

Definition 18 (Indicator function)**.** Let $K \subset \mathbb{R}^{n\times m}$ be a closed convex nonempty set. The function $I_K(\mathbb{X})$ given as

$$I_K(\mathbb{X}) = \begin{cases} 0, & \text{if} \, \mathbb{X} \in K \\ +\infty, & \text{otherwise} \end{cases},$$

is called the indicator function of K

The indicator function is hepful for constraint minimization. If f is reasonable (at least finitely valued on K), then it holds:

$$\min_K f = \min_{\mathbb{R}^{n\times m}} (f + I_K).$$

Example (Unit interval)**.** Let $K = [0, 1]$. What is $\partial I_K(x)$?

If $x \in (0, 1)$, then $I_K(x) = 0$ so the only ξ such that $0 = I_K(y) \geq 0 + \xi(y - x)$ holds $\forall y \in [0, 1]$ is $\xi = 0$.

If $x = 0, x = 1$ then $\partial I_K(0) = (-\infty, 0], \partial I_K(1) = [0, \infty)$. This resembles a normal "vector", but in fact it is not a single vector and more a "cone" of vectors.

Definition 19 (Normal cone to a set)**.** Let K be closed convex nonempty set. The subdifferential of the indicator function I_K is called the normal cone to the set K and it is denoted by N_K.

Example. Minimize x^2 on $[1, 2]$. We are looking for

$$\min_{[1,2]} x^2 = \min_{\mathbb{R}} \left(x^2 + I_{[1,2]}(x)\right).$$

It must hold at the minimum

$$0 \in \partial\left(x^2 + I_{[1,2]}(x)\right) \Leftrightarrow -\partial I_{[1,2]}(x) \supset \partial x^2 \Leftrightarrow \left(x^2\right)' \in -N_{[1,2]}(x)$$

Example. Take a square $K = [0, 1] \times [0, 1] \subset \mathbb{R}^2$. We know $K \in C^{0,1}$ so the outter normal exist at a.a. points on the boundary. The outter normal does not exist in the corners, but the normal cone does.

Definition 20 (Fenchel (convex) conjugate — Legendre transformation)**.** Let $f : \mathbb{R}^{n\times m} \to \mathbb{R}$ and $\mathbb{X}$ be fixed. Let $\mathbb{X}^*$ be a "slope" such that

$$f(\mathbb{X}) \geq \mathbb{X}^* : \mathbb{X} - k, \forall \mathbb{X} \in \mathbb{R}^{n\times m},$$

which means $k \geq \mathbb{X}^* : \mathbb{X} - f(\mathbb{X}), \forall \mathbb{X} \in \mathbb{R}^{n\times m}$.

We set

$$f^*(\mathbb{X}^*) := \sup_{\mathbb{X}\in\mathbb{R}^{n\times m}} (\mathbb{X}^* : \mathbb{X} - f(\mathbb{X})).$$

Remark. f^* is always convex even if f is not. But when f is convex and lower-semicontinuous, then

$$f^{**} = f, \text{ (biconjugate)}.$$

Theorem 6 (Fenchel identity)**.** *Let* $\mathbb{X}^* \in \partial f(\mathbb{X})$*. Then*

$$\mathbb{X}^* : \mathbb{X} = f(\mathbb{X}) + f^*(\mathbb{X}^*).$$

Proof. Let us assume that $\mathbb{X}^* \in f(\mathbb{X})$. Then it must hold

$$f(\mathbb{Y}) \geq f(\mathbb{X}) + \mathbb{X}^* : (\mathbb{Y} - \mathbb{X}), \forall \mathbb{Y} \in \mathbb{R}^{n\times m},$$

so rearranging yields

$$\mathbb{X}^* : \mathbb{X} - f(\mathbb{X}) \geq \mathbb{X}^* : \mathbb{Y} - f(\mathbb{Y}),$$

so taking the supremum gives exactly [15]

$$\mathbb{X}^* : \mathbb{X} - f(\mathbb{X}) \geq \sup_{\mathbb{Y}} (\mathbb{X}^* : \mathbb{Y} - f(\mathbb{Y})) = f^*(\mathbb{X}^*).$$

We have thus obtained

$$\mathbb{X}^* : \mathbb{X} \geq f^*(\mathbb{X}^*) + f(\mathbb{X}).$$

□

Remark (Minimization of $f \Leftrightarrow$ minimization of f^*)*.* We see that it holds:

$$x^* \in \partial f(x) \Leftrightarrow x \in \partial f^*(x^*).$$

7 (von Mieses) elastoplasticity

The previous mathematical apparatus now enables us to develop a theory for modelling more complicated, *fundamentally unsmooth*, phenomena. It all begins with a man.

7.1 Problem of a man...

Assume a person is pulling a box of weight m by a spring. If he is pulling just a little, the box does not move, only the spring is deformed - but in a reversible, elastic way. To move the box,

[15]The inequality becomes equality, as it can be reached by taking $y = x$.

the man needs to pull at least with the force $\sigma_0 = mgc$, where c is some friction coefficient. When he us pulling with force greater than σ_0, the box is moving and does not require any extra force to be moved (the system to be deformed). The deformation can be decomposed as

$$\mathbb{e} = \tilde{\mathbb{e}} + \mathbb{p},$$

where $\mathbb{e}$ is the total strain, $\tilde{\mathbb{e}}$ is the elastic strain and $\mathbb{p}$ is the plastic strain.

The elasticity part is described as

$$\begin{cases} -\nabla \cdot \mathbb{\sigma} = \mathbf{f}, & \text{in the bulk} \\ \mathbb{\sigma}\mathbf{n} = \mathbf{g}, & \text{on the boundary} \end{cases},$$

with some constituive relation

$$\mathbb{\sigma} = \mathcal{C}\tilde{\mathbb{e}} = \mathcal{C}(\mathbb{e} - \mathbb{p}).$$

What about the plastic part? The material does not deform when subjected to some (small enough) stress $\mathbb{\sigma}$, which can be modelled as

$$\begin{cases} \dot{\mathbb{p}}(t) \in N_K(\mathbb{\sigma}), \\ \mathbb{p}(0) = \mathbb{p}_0, \end{cases}$$

where $K \subset \mathbb{R}^{n\times n}_{\mathrm{sym}}$ is a convex closed subset of stresses $\mathbb{\sigma}$ such that $\mathbb{0} \in K$. This means that the plastic deformation is zero inside K, i.e. for some stresses.

Remark. Very often, the deformation is considered "incompressible", i.e.,

$$\det \mathbb{F} = 1,$$

which in linear case translates into

$$\operatorname{tr} \mathbb{e} = 0.$$

In most cases, the set K is given as

$$K = \{\mathbb{\sigma} : \varphi(\mathbb{\sigma}) \leq 0\},$$

where φ is the **yield function.** The set

$$\{\mathbb{\sigma} | \varphi(\mathbb{\sigma}) = 0\}$$

is called the **yield surface.** Very often we have

$$\varphi(\mathbb{\sigma}) = |\mathbb{\sigma}^D| - c_0,$$

where $|\cdot|$ denotes the Frobenius norm and

$$\mathbb{\sigma}^D = \mathbb{\sigma} - \frac{1}{3}(\operatorname{tr}\mathbb{\sigma})\mathbb{I},$$

is the *deviatoric part of the stress tensor.* Notice in particular that

$$\operatorname{tr} \mathbb{\sigma}^D = 0.$$

7.1.1 Plastic evolution

From the previous we have

$$\dot{\mathbb{p}} = \begin{cases} \mathbb{0}, & \text{if } \varphi(\sigma) < 0, \\ \frac{\lambda}{|\sigma^D|}\sigma^D, & \text{if } \varphi(\sigma) = 0, \lambda \geq 0 \end{cases}.$$

Also $\dot{\mathbb{p}} \in N_K(\sigma) = \partial I_K(\sigma)$ so

$$\sigma \in \partial I_K^*(\dot{\mathbb{p}}),$$

where

$$I_K^*(\dot{\mathbb{p}}) = \sup_{\mathbb{q} \in \mathbb{R}^{3\times 3}} (\dot{\mathbb{p}} : \mathbb{q} - I_K(\mathbb{q})) = \sup_{\mathbb{q} \in K} \dot{\mathbb{p}} : \mathbb{q},$$

is the Fenchel transformation of I_K, also called the **supporting function** of $\dot{\mathbb{p}}$. We are able to rewrite the supremum to take the form[16]

$$I_K^*(\dot{\mathbb{p}}) = \dot{\mathbb{p}} : \frac{c_0}{|\dot{\mathbb{p}}|}\dot{\mathbb{p}},$$

if however the second term lies in K. Realize now that if $\operatorname{tr} \dot{\mathbb{p}} = 0$ then

$$I_K^*(\dot{\mathbb{p}}) = c_0 |\dot{\mathbb{p}}|,$$

and if $\operatorname{tr} \dot{\mathbb{p}} \neq 0$, then $I_K^*(\dot{\mathbb{p}}) = +\infty$. If we now define the **dissipation potential** D as

$$D(\dot{\mathbb{p}}) = \begin{cases} c_0|\dot{\mathbb{p}}|, & \text{if } \operatorname{tr} \dot{\mathbb{p}} = 0 \\ +\infty, & \text{otherwise} \end{cases},$$

we get the following condition

$$\sigma \in \partial D(\dot{\mathbb{p}}).$$

Let us summarise a bit. For the stress tensor we have

$$\sigma = \mathcal{C}(\mathbb{e} - \mathbb{p}) \in \partial D(\dot{\mathbb{p}}).$$

The general relation also yields

$$\sigma = \frac{\partial w(\tilde{\mathbb{e}})}{\partial \tilde{\mathbb{e}}} = \frac{\partial w(\mathbb{e} - \mathbb{p})}{\partial \tilde{\mathbb{e}}},$$

where $w(\tilde{\mathbb{e}}) = \frac{1}{2}\mathcal{C}\tilde{\mathbb{e}} : \tilde{\mathbb{e}}$ is the free energy density. Using the chain rule we obtain the condition

$$-\frac{\partial w(\mathbb{e} - \mathbb{p})}{\partial \mathbb{p}} \in \partial D(\dot{\mathbb{p}}).$$

In total, we are solving the following system

$$\begin{cases} 0 \in \frac{\partial w(\mathbb{e} - \mathbb{p})}{\partial \mathbb{p}} + \partial D(\dot{\mathbb{p}}), & \text{in } \Omega (\textit{flow rule}) \\ \mathbb{p}(0) = \mathbb{p}_0, & \text{in } \Omega \\ -\nabla \cdot (\mathcal{C}(\mathbb{e} - \mathbb{p})) = \mathbf{f}, & \text{in } \Omega \\ \text{boundary conditions}, & \text{on } \partial\Omega \end{cases}.$$

How to solve the system?

[16] To utilize Cauchy-Schwarz later.

7.1.2 Discrete time setting

Let us take $t \in [0,T]$ and fix $\tau = \frac{T}{N}, N \in \mathbb{N}$ for some $N >> 1$. Assume that using some discrete scheme, we are able to calculate $\mathbb{p}$ at a certain time. Then we must solve

$$\begin{cases} 0 \in \frac{\partial w(\mathbb{e}-\mathbb{p})}{\partial \mathbb{p}} + \partial D\left(\frac{\mathbb{p}-\mathbb{p}_{k-1}}{\tau}\right), & \text{in } \Omega \\ -\nabla \cdot (\mathcal{C}(\mathbb{e}_k - \mathbb{p}_k)) = \mathbf{f}_k, & \text{in } \Omega \end{cases}.$$

Which are the E-L equations of the functional [17]

$$I(\mathbf{u}, \mathbb{p}) = \int_\Omega w(\mathbb{e}(\mathbf{u}) - \mathbb{p})\,\mathrm{d}x + \tau \int_\Omega D\left(\frac{1}{\tau}(\mathbb{p} - \mathbb{p}_{k-1})\right)\mathrm{d}x - \int_\Omega \mathbf{f}_k \cdot \mathbf{u}\,\mathrm{d}x - \int_{\Gamma_N} \mathbf{g}_k \cdot u\,\mathrm{d}S\,.$$

Really, taking the variaton with respect to $\mathbf{u}$ gives us

$$-\nabla \cdot (\mathcal{C}(\mathbb{e}_k - \mathbb{p}_k)) = \mathbf{f}_k,$$

and the variation with respect to $\mathbb{p}$ gives us

$$0 \in -\sigma + \partial D\left(\frac{1}{\tau}(\mathbb{p} - \mathbb{p}_{k-1})\right).$$

If we want to minimize this functional, *i.e.*, solve the equations, it must hold [18] $D(\mathbb{q}) \neq +\infty$ (for $\mathbb{q}$ being the argument). From our assumptions on the dissipation potential this however implies.

$$D(\mathbb{q}) = c_0|\mathbb{q}|, \operatorname{tr}\mathbb{q} = 0,$$

and we say the evolution is **rate-independent**. We see that D is 1-homogenous:

$$D(\alpha\mathbb{q}) = \alpha D(\mathbb{q}).$$

Rewritting the functional now yields:

$$I(\mathbf{u}, \mathbb{p}) = \frac{1}{2}\int_\Omega \mathcal{C}(\mathbb{e}(\mathbf{u}) - \mathbb{p}) : (\mathbb{e}(\mathbf{u}) - \mathbb{p})\,\mathrm{d}x + \int_\Omega c_0|\mathbb{p} - \mathbb{p}_{k-1}|\,\mathrm{d}x - L_k(\mathbf{u}), \mathbb{p}(0) = \mathbb{p}_0,$$

where $L_k(\mathbf{u})$ is the loading (at the k-th time step.) The sought solution is the pair $(\mathbf{u}_k, \mathbb{p}_k)$ which satisfies

$$I(\mathbf{u}_k, \mathbb{p}_k) = \min_{\mathbf{u}, \mathbb{p}} I(\mathbf{u}, \mathbb{p}).$$

7.2 Rheological models

7.2.1 Dashpots

Or *tlumič* in Czech. The plastic stress is assumed to take the form

$$\sigma_p = \mathcal{D}\dot{\mathbb{e}}(\nabla\mathbf{u}), \sigma_{ij} = D_{ijkl}\dot{e}_{kl}(\nabla\mathbf{u}),$$

where $\mathcal{D}$ is the **tensor of viscosity constants.** [19]

[17] We have guessed it.
[18] If not, we have no chance of minimizing it.
[19] People say viscosity stresses or viscous stress. This is used, but nonetheless it is wrong.

7.2.2 Kelvin-Voigt mateial

The response of some materials can be modelled as a ”parallel composition of a spring and a dashpot.” Then, the total stress is

$$\mathbb{\sigma} = \mathbb{\sigma}_p + \mathbb{\sigma}_e,$$

that is the sum of the plastic and the elastic stresses. The strain is of course the same:

$$\mathbb{e} = \mathbb{e}_p = \mathbb{e}_e.$$

The governing equations thus are

$$\begin{aligned}
-\nabla\cdot(\mathcal{C}\mathbb{e}(\mathbf{u}) + \mathcal{D}\dot{\mathbb{e}}(\mathbf{u})) &= \mathbf{f},\ \text{in}\,\Omega\\
(\mathcal{C}\mathbb{e} + \mathcal{D}\dot{\mathbb{e}})\mathbf{n} &= \mathbf{0},\ \text{on}\,\Gamma_N\\
\mathbf{u} &= \mathbf{0},\ \text{on}\,\Gamma_D\\
\mathbb{e}(t=0) &= \mathbb{e}_0,\ \text{in}\,\Omega.
\end{aligned}$$

Let us *formally* write down the energy balance. As usual, multiply the first equation by $\dot{\mathbf{u}}$ and integrate $\int_\Omega \mathrm{d}\mathbf{x}$.

$$\int_\Omega -\nabla\cdot(\mathcal{C}\mathbb{e} + \mathcal{D}\dot{\mathbb{e}})\cdot\dot{\mathbf{u}}\,\mathrm{d}x = \int_\Omega \mathbf{f}\cdot\dot{\mathbf{u}}\,\mathrm{d}x\,,$$

using Gauss

$$\int_\Omega (\mathcal{C}\mathbb{e} + \mathcal{D}\dot{\mathbb{e}}) : \nabla\dot{\mathbf{u}}\,\mathrm{d}x = \int_\Omega \mathbf{f}\cdot\dot{\mathbf{u}}\,\mathrm{d}x = {}^{20} \int_\Omega \mathcal{C}\mathbb{e} : \dot{\mathbb{e}}\,\mathrm{d}x + \int_\Omega \mathcal{D}\dot{\mathbb{e}} : \dot{\mathbb{e}}\,\mathrm{d}x = \int_\Omega \mathbf{f}\cdot\mathbf{u}\,\mathrm{d}x\,,$$

and now we rewrite

$$= \int_\Omega \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{2}\mathcal{C}\mathbb{e}(\mathbf{u}) : \mathbb{e}(\mathbf{u})\Big)\,\mathrm{d}x + \int_\Omega \mathcal{D}\dot{\mathbb{e}}(\mathbf{u}) : \dot{\mathbb{e}}(\mathbf{u})\,\mathrm{d}x = \int_\Omega \mathbf{f}\cdot\dot{\mathbf{u}}\,\mathrm{d}x\,,$$

and integrate in time:

$$\int_0^T \int_\Omega \frac{\mathrm{d}}{\mathrm{d}t}\Big(\frac{1}{2}\mathcal{C}\mathbb{e}(\mathbf{u}) : \mathbb{e}(\mathbf{u})\Big)\,\mathrm{d}x\,\mathrm{d}t + \int_0^T \int_\Omega \mathcal{D}\dot{\mathbb{e}}(\mathbf{u}) : \dot{\mathbb{e}}(\mathbf{u})\,\mathrm{d}x\,\mathrm{d}t = \int_0^T \int_\Omega \mathbf{f}\cdot\dot{\mathbf{u}}\,\mathrm{d}x\,\mathrm{d}t\,.$$

Remember that

$$w(\mathbb{e}(\mathbf{u})) = \frac{1}{2}\mathcal{C}\mathbb{e}(\mathbf{u}) : \mathbb{e}(\mathbf{u}),$$

so we have obtained

$$\int_\Omega w(\mathbb{e}(\mathbf{u}(T)))\,\mathrm{d}x - \int_\Omega w(\mathbb{e}(\mathbf{u}(0)))\,\mathrm{d}x + \int_0^T \int_\Omega \mathcal{D}\dot{\mathbb{e}}(\mathbf{u}) : \dot{\mathbb{e}}(\mathbf{u})\,\mathrm{d}x\,\mathrm{d}t = \int_0^T \int_\Omega \mathbf{f}\cdot\dot{\mathbf{u}}\,\mathrm{d}x\,\mathrm{d}t\,.$$

[20] It holds $\dot{\mathbb{e}}(\mathbf{u}) = \mathbb{e}(\dot{\mathbf{u}})$.

7.2.3 Maxwell material

This is the case when we "put the spring and the dashpot in serial composition". The total stress is

$$\sigma = \sigma_p = \sigma_e,$$

and the total strain is

$$\varepsilon = \mathbb{e}_p + \mathbb{e}_e.$$

7.3 Internal parameters

A lot of materials can be described using some internal parameters $\mathbb{z}$ (scalars, vectos, tensors; we take the tensor case for generality); for example, plastic strain, fatique, damage, length of a crack, delamination.

The model

$$\sigma = \partial_{\dot{\mathbb{e}}}\zeta(\dot{\mathbb{e}}, \dot{\mathbb{z}}) + \partial_{\mathbb{e}}\varphi(\mathbb{e}, \mathbb{z}),$$

with the flow rule

$$0 \in \partial_{\dot{\mathbb{z}}}\zeta(\dot{\mathbb{e}}, \dot{\mathbb{z}}) + \partial_{\mathbb{z}}\varphi(\mathbb{e}, \mathbb{z}).$$

is called the **generalized Kelvin-Voigt** model/material. From now on, we will be using φ for the stored energy density. There is some analogy:

- φ is the stored energy density = potential of elastic stress, reversible procceses
- ζ is the (pseudo)potential of dissipative forces., irreveresible proccesses

To do anything, we need to obtain some energy balance, so test by $\dot{\mathbf{u}}$. Investigate the terms:

$$\sigma : \dot{\mathbb{e}} = \partial_{\dot{\mathbb{e}}}\zeta(\dot{\mathbb{e}}, \dot{\mathbb{z}}) : \dot{\mathbb{e}} + \partial_{\mathbb{e}}\varphi(\mathbb{e}, \mathbb{z}) : \dot{\mathbb{e}},$$

realize now that from the flow rule it follows

$$(\partial_{\dot{\mathbb{z}}}\zeta(\dot{\mathbb{e}}, \dot{z}) + \partial_{\mathbb{z}}\varphi(\mathbb{e}, \mathbb{z})) : \dot{\mathbb{z}} = 0,$$

so i can add it to the previous term and obtain

$$\sigma : \dot{\mathbb{e}} = \partial_{\dot{\mathbb{e}}}\zeta(\dot{\mathbb{e}}, \dot{\mathbb{z}}) : \dot{\mathbb{e}} + \partial_{\mathbb{e}}\varphi(\mathbb{e}, \mathbb{z}) : \dot{\mathbb{e}} + \partial_{\dot{\mathbb{z}}}\zeta(\dot{\mathbb{e}}, \dot{\mathbb{z}}) : \dot{\mathbb{z}} + \partial_{\mathbb{z}}\varphi(\mathbb{e}, \mathbb{z}) : \dot{\mathbb{z}},$$

Realize now that we have obtained

$$\partial_{\mathbb{e}}\varphi(\mathbb{e}, \mathbb{z}) : \dot{\mathbb{e}} + \partial_{\mathbb{z}}\varphi(\mathbb{e}, \mathbb{z}) : \dot{\mathbb{z}} = \frac{\mathrm{d}}{\mathrm{d}t}\varphi(\mathbb{e}, \mathbb{z}),$$

and denoting the quantity

$$\xi(\dot{\mathbb{e}}, \dot{\mathbb{z}}) := \partial_{\dot{\mathbb{e}}}\zeta(\dot{\mathbb{e}}, \dot{\mathbb{z}}) : \dot{\mathbb{e}} + \partial_{\dot{\mathbb{z}}}\zeta(\dot{\mathbb{e}}, \dot{\mathbb{z}}) : \dot{\mathbb{z}},$$

as the *rate of the dissipation* we obtain

$$\sigma : \dot{\mathbb{e}} = \frac{\mathrm{d}}{\mathrm{d}t}\varphi(\mathbb{e}, \mathbb{z}) + \xi(\dot{\mathbb{e}}, \dot{\mathbb{z}}).$$

What are the properties of ξ? First of all, we require

$$\xi \geq 0.$$

Assume ζ is a covex function:

$$\zeta(0,0) \geq \zeta(\dot{\mathbb{e}},\dot{\mathbb{z}}) + \partial_{\dot{\mathbb{e}}}\zeta(\dot{\mathbb{e}},\dot{\mathbb{z}}) : (-\dot{\mathbb{e}}) + \partial_{\dot{\mathbb{z}}}\zeta(\dot{\mathbb{e}},\dot{\mathbb{z}}) : (-\dot{\mathbb{z}}).$$

Moreover, assume now $\zeta(0,0) = 0$. We have

$$\xi(\dot{\mathbb{e}},\dot{\mathbb{z}}) = \partial_{\dot{\mathbb{e}}}\zeta(\dot{\mathbb{e}},\dot{\mathbb{z}}) : \dot{\mathbb{e}} + \partial_{\dot{\mathbb{z}}}\zeta(\dot{\mathbb{e}},\dot{\mathbb{z}}) : \dot{\mathbb{z}} \geq \zeta(\dot{\mathbb{e}},\dot{\mathbb{z}}) \geq 0.$$

Finally, the total power balance becomes

$$\int_\Omega \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\rho|\dot{\mathbf{u}}|^2\,\mathrm{d}x + \int_\Omega \frac{\mathrm{d}}{\mathrm{d}t}\varphi(\mathbb{e},z)\,\mathrm{d}x + \int_\Omega \xi(\dot{\mathbb{e}},\dot{z})\,\mathrm{d}x = \int_\Omega \mathbf{f}\cdot\dot{\mathbf{u}}\,\mathrm{d}x\,,$$

and the total energy balance becomes

$$\int_\Omega \frac{1}{2}\rho\big(|\dot{\mathbf{u}}(T)|^2 - |\dot{\mathbf{u}}(0)|^2\big)\,\mathrm{d}x + \int_\Omega (\varphi(\mathbb{e}(T),z(T)) - \varphi(\mathbb{e}(0),z(0)))\,\mathrm{d}x + \int_0^T\int_\Omega \xi(\dot{\mathbb{e}},\dot{z})\,\mathrm{d}x\,\mathrm{d}t = \int_0^T\int_\Omega \mathbf{f}\cdot\dot{\mathbf{u}}\,\mathrm{d}x\,\mathrm{d}t\,.$$

8 Thermodynamics in the framework of GSM (generalized standard materials)

Having obtained some knowledge of thermodynamical quantites, we are ready to generalize the theory. We will see that the evolution of a specimen can be acquired by the knowledge of the stored energy density ψ and the dissipation "potential" ζ

Denote

$$\psi = \psi(\mathbb{e},\mathbb{z},\theta),\ \zeta = \zeta(\dot{\mathbb{e}},\dot{\mathbb{z}})$$

to be *the stored energy* and the *dissipation potential.* Here $\theta > 0$ denotes the absolute thermodynamic temperature. Let us denote

$$\mathbb{\sigma}_{el} = \frac{\partial\psi}{\partial\mathbb{e}},\ \mathbb{\sigma}_{in} = \frac{\partial\psi}{\partial\mathbb{z}},\ s = -\frac{\partial\psi}{\partial\theta},$$

as the elastic and inelastic stress and the entropy density. Moreover, define

$$w(\mathbb{e},\mathbb{z},\theta,s) = \psi(\mathbb{e},\mathbb{z},\theta) + \theta s$$

as the **internal energy density**. If we calculate the time derivative of the internal energy density we obtain:

$$\dot{w} = \frac{\partial}{\partial t}(\psi(\mathbb{e},z,\theta) + \theta s) = \frac{\partial\psi}{\partial\mathbb{e}} : \dot{\mathbb{e}} + \frac{\partial\psi}{\partial\mathbb{z}} : \dot{\mathbb{z}} + \underbrace{\frac{\partial\psi}{\partial\theta}\dot{\theta} + \dot{\theta}s}_{=-s\dot{\theta}+\dot{\theta}s=0} + \theta\dot{s}.$$

We *postulate*:

$$\dot{w} = \sigma_{el} : \dot{\mathbb{e}} + \sigma_{in} : \dot{\mathbb{z}} + \xi(\dot{\mathbb{e}},\dot{\mathbb{z}}) - \nabla\cdot\mathbf{j},$$

where $\mathbf{j}$ is the heat flux. From this postulate, we obtain

$$\xi(\dot{\mathbb{e}},\dot{\mathbb{z}}) = \theta\dot{s} + \nabla\cdot\mathbf{j}. \tag{13}$$

A common modelling choice is the dependency

$$\mathbf{j} = \mathbf{j}(\theta, \mathbb{e}, \mathbb{z}, \nabla\theta) = -\mathbb{K}(\mathbb{e}, \mathbb{z}, \theta)\nabla\theta,$$

known as the *Fourier law.* Here

$$\mathbb{K} \in \{\mathbb{A} \in \mathbb{R}^{3\times3} | \mathbb{A} > 0\},$$

is the *matrix of heat flux coefficients.* This is a classical example of a constitutive law.

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega s(t,\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_\Omega \frac{1}{\theta}(\xi - \nabla\cdot\mathbf{j})\,\mathrm{d}\mathbf{x} &= \int_\Omega \frac{\xi}{\theta}\,\mathrm{d}\mathbf{x} + \int_\Omega \frac{\nabla\cdot(\mathbb{K}\nabla\theta)}{\theta}\,\mathrm{d}\mathbf{x} = \\ &= \int_{\partial\Omega} \frac{\mathbb{K}\nabla\theta}{\theta}\cdot\mathbf{n}\,\mathrm{d}S - \int_\Omega \mathbb{K}\nabla\theta\cdot\nabla\left(\frac{1}{\theta}\right)\mathrm{d}\mathbf{x} + \int_\Omega \frac{\xi}{\theta}\,\mathrm{d}\mathbf{x} = \\ &= \int_\Omega \left(\frac{\xi}{\theta} + \frac{\mathbb{K}\nabla\theta\cdot\nabla\theta}{\theta^2}\right)\mathrm{d}\mathbf{x} - \int_{\partial\Omega} \frac{\mathbf{j}}{\theta}\cdot\mathbf{n}\,\mathrm{d}S\,.\end{aligned}$$

This relation is known as the *Clausius-Duhem inequality.*[21]

From the definition of s

$$s = -\frac{\partial\psi}{\partial\theta}(\theta, \mathbb{e}, \mathbb{z}),$$

it follows

$$\dot{s} = -\frac{\partial^2\psi}{\partial\theta^2}\theta - \frac{\partial^2\psi}{\partial\theta\partial\mathbb{e}} : \dot{\mathbb{e}} - \frac{\partial^2\psi}{\partial\theta\partial\mathbb{z}} : \dot{\mathbb{z}},$$

and so

$$\theta\dot{s} = \underbrace{-\frac{\partial^2\psi}{\partial\theta^2}\theta}_{:=C_V}\dot{\theta} - \frac{\partial^2\psi}{\partial\theta\partial\mathbb{e}} : (\dot{\mathbb{e}}\theta) - \frac{\partial^2\psi}{\partial\theta\partial\mathbb{z}} : (\dot{\mathbb{z}}\theta) = C_V\dot{\theta} - \frac{\partial^2\psi}{\partial\theta\partial\mathbb{e}} : (\dot{\mathbb{e}}\theta) - \frac{\partial^2\psi}{\partial\theta\partial\mathbb{z}} : (\dot{\mathbb{z}}\theta),$$

where we have identified

$$C_V = -\theta\frac{\partial^2\psi}{\partial\theta^2},$$

as the *heat capacity at the constant volume.* Coming back to 13, we read

$$C_V\dot{\theta} + \nabla\cdot\mathbf{j} = \xi(\dot{\mathbb{e}}, \dot{\mathbb{z}}) + \theta\frac{\partial^2\psi}{\partial\theta\partial\mathbb{e}} : \dot{\mathbb{e}} + \theta\frac{\partial^2\psi}{\partial\theta\partial\mathbb{z}} : \dot{\mathbb{z}}.$$

This is our *heat equation,* the right hand side are the sources. We could identify the derivatives of the potential with lets say some derivative of σ_{el}, but let us keep the "thermodynamics and mechanics separated."; although it does not really make sense. In total

$$\begin{aligned}C_V\dot{\theta} - \nabla\cdot(\mathbb{K}\nabla\theta) &= \xi + \theta\frac{\partial^2\psi}{\partial\theta\partial\mathbb{e}} : \dot{\mathbb{e}} + \theta\frac{\partial^2\psi}{\partial\theta\partial\mathbb{z}} : \dot{\mathbb{z}}, \\ \rho\ddot{\mathbf{u}} - \nabla\cdot(\sigma_{el} + \sigma_{in}) &= \mathbf{f}, \\ 0 &\in \partial_{\dot{\mathbb{z}}}\zeta(\dot{\mathbb{e}}, \dot{\mathbb{z}}) + \partial_{\mathbb{z}}\psi(\mathbb{e}, \mathbb{z}, \theta),\end{aligned}$$

plus of course some initial and boundary conditions.

[21] Although inequalty, there appears only the equality sign "=". I do not actually know what that means.

9 Summary

At the end, the lecture is summarized.

It began with deformation:

$$\mathbf{y} : \overline{\Omega} \to \mathbb{R}^3, \nabla \mathbf{y} = \mathbb{F}, \mathbb{C} = \mathbb{F}^\top \mathbb{F}, \det \mathbb{F} > 0.$$

and some quantities associated with these. A little excursion allowed as to define

$$W = W(\nabla \mathbf{y}) = W(\mathbb{F}),$$

to be the stored energy density. Note that later on, we have called it ψ. Coming back to deformation, we have defined various stress measures:

$$\mathbb{T}^y, \mathbb{T} = \mathbb{T}^y \operatorname{cof} \mathbb{F}, \mathbb{S} = \mathbb{F}^{-1}\mathbb{T}.$$

Wanting to show existence of solutions, we needeed the convexity of some functionals. A problem with rotations however meant we needed to lower our expectations and we had to discover polyconvexity and rank-1 convexity. This included *e.g.* Legendre-Hadamard condition.

Realizing we are stuck in full theory, we began exploring linearized elasticity. To show existence, we refreshed the Korn's inequality. And because that all seemed easy, a question about time dependence has been asked: is everything truly stationary?

No, it is not; that lead us to von Mises elastoplasticity and to a class of materials, such as Kelvin-Voigt or Maxwell materials. Generalizing this framework and also including some internal variables, we have given the foundations of (the thermodynamics of)gave generalized standard materials: this was especially elegant, as from the Helmholtz free energy and the dissipation potential, we were able to derive evolution equations for the important thermodynamical quantites. This included some energy/power estimates, balances and the notion of entropy and its rate.

10 (Some) tutorials

10.1 Change of observer

The requirement of material frame indifference yields:

$$W = \hat{W}(\mathbb{F}) = \hat{W}(\mathbb{Q}\mathbb{F}), \forall \mathbb{Q} \in \operatorname{orth}.$$

10.2 Change of reference configuration

The requiremenent of material symmetry yields:

$$W = \hat{W}(\mathbb{F}) = \hat{W}(\mathbb{F}\mathbb{P}), \forall \mathbb{P} \in \mathcal{G},$$

where $\mathcal{G}$ is the symmetry group of the material.

10.3 Consequences of isotropic hyperelastic solid

Remark (Groups unim, orth). The ”biggest sensible” symmetry group is the unimodular group:

$$\operatorname{unim} = \{\mathbb{P}, \det \mathbb{P} = \pm 1\}.$$

There exists another common group:

$$\text{orth}\,\{\mathbb{Q}, \mathbb{Q}\mathbb{Q}^\top = \mathbb{Q}^\top\mathbb{Q} = \mathbb{I}\} \subset \text{unim}\,.$$

We thus have $W = \hat{W}(\mathbb{F}) = \hat{W}(\mathbb{Q}\mathbb{F}) = \hat{W}(\mathbb{F}\mathbb{Q}), \forall \mathbb{Q} \in \text{orth}\,, \forall \mathbb{F}$.

Use *polar decomposition*: $\mathbb{F} = \mathbb{R}\mathbb{U} = \mathbb{V}\mathbb{R}, \mathbb{R} \in \text{orth}\,, \mathbb{U}, \mathbb{V}\,\text{positively definite}\,, \mathbb{U} = \sqrt{\mathbb{C}}, \mathbb{V} = \sqrt{\mathbb{B}}$.

To make use of it, we first use m.f.i. and then isotropy and then first use isotropy and then m.f.i. So from material frame indifference

$$W = \hat{F}(\mathbb{F}) = \hat{W}(\mathbb{Q}\mathbb{F}) = \hat{W}\left(\mathbb{R}^\top\mathbb{R}\mathbb{U}\right) = \hat{W}(\mathbb{U}) = \overline{W}(\mathbb{C}),$$

where we have taken $\mathbb{Q} = \mathbb{R}^\top$. Note that this works universaly (without the need of isotropy), as it comes from the objectivity consideration.

From isotropy

$$W = \hat{W}(\mathbb{F}) = \hat{W}(\mathbb{F}\mathbb{Q}) = \overline{W}(\mathbb{C}) = \overline{W}\left((\mathbb{F}\mathbb{Q})^\top(\mathbb{F}\mathbb{Q})\right) = \overline{W}\left(\mathbb{Q}^\top\mathbb{F}^\top\mathbb{F}\mathbb{Q}\right) = \overline{W}\left(\mathbb{Q}^\top\mathbb{C}\mathbb{Q}\right), \forall Q \in \text{orth}\,, \forall \mathbb{C}\,\text{admissable}\,.$$

Now backwards using the second polar decomposition:

$$W = \hat{W}(\mathbb{F}) = \hat{W}(\mathbb{F}\mathbb{Q}) = \hat{W}(\mathbb{V}\mathbb{R}\mathbb{Q}) = \hat{W}\left(\mathbb{V}\mathbb{R}\mathbb{R}^\top\right) = \hat{W}\left(\sqrt{\mathbb{B}}\right) = \tilde{W}(\mathbb{B}),$$

$$W = \tilde{W}(\mathbb{B}) = \tilde{W}\left(\mathbb{Q}\mathbb{F}(\mathbb{Q}\mathbb{F})^\top\right) = \tilde{W}\left(\mathbb{Q}\mathbb{B}\mathbb{Q}^\top\right).$$

So far, we have shown

$$\begin{aligned} W(t, \mathbf{X}) &= \overline{W}(\mathbb{C}(t, \mathbf{X}), \mathbf{X}) = \overline{W}\left(\mathbb{Q}\mathbb{C}\mathbb{Q}^\top\right), \\ W(t, \mathbf{X}) &= \tilde{W}(\mathbb{B}(t, \mathbf{X}), \mathbf{X}) = \tilde{W}\left(\mathbb{Q}\mathbb{B}\mathbb{Q}^\top\right), \end{aligned}$$

In HW, we will know

$$\mathbb{T} = 2\frac{\partial \hat{W}(\mathbb{F})}{\partial \mathbb{F}}, T_{iJ} = 2\frac{\partial \hat{W}}{\partial F_{iJ}}$$

and we can show

$$\mathbb{T} = 2\frac{\partial \hat{W}(\mathbb{F})}{\partial \mathbb{F}} = 2\mathbb{F}\frac{\partial \tilde{W}(\mathbb{B})}{\partial \mathbb{B}}\mathbb{F}, \mathbb{T}^y = \ldots, \mathbb{S} = 2\frac{\partial \overline{W}(\mathbb{C})}{\partial \mathbb{C}}.$$

Definition 21 (Isotropic functions)**.** We say the functions $\hat{a}(y_\alpha, \mathbf{y}_\alpha, \mathbb{Y}_\alpha), \hat{\mathbf{a}}(y_\alpha, \mathbf{y}_\alpha, \mathbb{Y}_\alpha), \hat{\mathbb{A}}(y_\alpha, \mathbf{y}_\alpha, \mathbb{Y}_\alpha), \alpha = 1, \ldots, N$ are isotropic functions (of their respective arguments) if it holds

$$\begin{aligned} \hat{a}(y_\alpha, \mathbf{y}_\alpha, \mathbb{Y}_\alpha) &= \hat{a}\left(y_\alpha, \mathbb{Q}\mathbf{y}_\alpha, \mathbb{Q}\mathbb{Y}_\alpha\mathbb{Q}^\top\right), \\ \mathbb{Q}\hat{\mathbf{a}}(y_\alpha, \mathbf{y}_\alpha, \mathbb{Y}_\alpha) &= \hat{\mathbf{a}}\left(y_\alpha, \mathbb{Q}\mathbf{y}_\alpha, \mathbb{Q}\mathbb{Y}_\alpha\mathbb{Q}^\top\right), \\ \mathbb{Q}\hat{\mathbb{A}}\mathbb{Q}^\top &= \hat{\mathbb{A}}\left(y_\alpha, \mathbb{Q}\mathbf{y}_\alpha, \mathbb{Q}\mathbb{Y}_\alpha\mathbb{Q}^\top\right), \end{aligned}$$

.

So we see that $\overline{W}(\mathbb{C}), \tilde{W}(\mathbb{B})$ are **scalar isotropic functions of 1 tensorial (symmetric) argument.**

Theorem 7 (Representation theorem for scalar isotropic functions)**.** *Let* $\psi = \hat{\psi}(\mathbb{A}) = \hat{\psi}(\mathbb{Q}\mathbb{A}\mathbb{Q}^\top)$

be a scalar isotropic function of a single symmetric tensorial variable. Then it must hold

$$\hat{\psi}(\mathbb{A}) \equiv \hat{\psi}(\mathrm{I}_1(\mathbb{A}), \mathrm{I}_2(\mathbb{A}), \mathrm{I}_3(\mathbb{A})),$$

where

$$\begin{aligned}
\mathrm{I}_1(\mathbb{A}) &= \operatorname{tr}\mathbb{A}, \\
\mathrm{I}_2(\mathbb{A}) &= \frac{1}{2}\left((\operatorname{tr}\mathbb{A})^2 - \operatorname{tr}\mathbb{A}^2\right), \\
\mathrm{I}_3(\mathbb{A}) &= \det\mathbb{A},
\end{aligned}$$

are the invariants of $\mathbb{A}$*.*

Proof. $\det(\mathbb{A} - \lambda\mathbb{I}) = -\lambda^3 + \lambda^2 \mathrm{I}_1 - \lambda \mathrm{I}_2 + \mathrm{I}_3 = p_\lambda(\mathbb{A})$ We will prove a different assertion:

$\mathbb{A}, \mathbb{B}$ are symmetric with the same invariants $\Leftrightarrow \exists \mathbb{Q} : \mathbb{A} = \mathbb{Q}\mathbb{B}\mathbb{Q}^\top$ " $\Leftarrow$ " is trivial, as then the matrices are similliar, so they have the same char. polynomial, so they have the same invariants. $\Rightarrow$ have same eigenvalues, so if i write the spectral decomposiiton, i can write

$$\mathbb{A} = \mathbb{Q}\mathbb{\Lambda}\mathbb{Q}^\top, \mathbb{B} = \mathbb{Q}\mathbb{\Lambda}\mathbb{R}^\top = \mathbb{R}\mathbb{Q}^\top\mathbb{A}\mathbb{Q}\mathbb{R}^\top.$$

Now suppose that the function is not a function of the invariants: $\hat{\psi} \neq \tilde{\psi}(\mathrm{I}_1, \mathrm{I}_2, \mathrm{I}_3)$. That means $\exists \mathbb{A}_,, \mathbb{A}_2$ such that $\mathrm{I}_1(\mathbb{A}_1) = \mathrm{I}_1(\mathbb{A}_2)$ and the same for the remaining invariants. Using the previous assertion, we have

$$\exists \mathbb{Q} : \mathbb{A}_1 = \mathbb{Q}\mathbb{A}_2\mathbb{Q}^\top \Rightarrow \hat{\psi}(\mathbb{A}_1) = \hat{\psi}\left(\mathbb{Q}\mathbb{A}_2\mathbb{Q}^\top\right) = \hat{\psi}(\mathbb{A}_2), \text{ but } \hat{\psi}(\mathbb{A}_1) \neq \hat{\psi}(\mathbb{A}_2).$$

□

Since using polar decomposition it can be shown the invariants of $\mathbb{B}, \mathbb{C}$ are the same we recieve

$$W = \tilde{W}(\mathrm{I}_1(\mathbb{B}), \mathrm{I}_2(\mathbb{B}), \mathrm{I}_3(\mathbb{B})) = \overline{W}(\mathrm{I}_1(\mathbb{C}), \mathrm{I}_2(\mathbb{C}), \mathrm{I}_3(\mathbb{C})).$$

10.4 Representation in terms of principial stresses

... in terms of the eigenvalues $\mathbb{U}, \mathbb{V}$. The invariants can be expressed as

$$\begin{aligned}
\mathrm{I}_1 &= \lambda_1 + \lambda_2 + \lambda_3, \\
\mathrm{I}_2 &= \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_1\lambda_3, \\
\mathrm{I}_3 &= \lambda_1\lambda_2\lambda_3.
\end{aligned}$$

Often in materials science the quantites can be expressed in these variables:

Example (Ogden materials)**.**

$$\hat{W}(\lambda_1, \lambda_2, \lambda_3) = \sum_{k=1}^{n} \frac{\mu_k}{\alpha_k}\left(\lambda_1^{\alpha_k} + \lambda_2^{\lambda_k} + \lambda_3^{\alpha_k} - 3\right)$$

How to calculate e.g. $\mathbb{T}$ in this representation?

$$\mathbb{T} = 2\frac{\partial W(\mathrm{I}_1, \mathrm{I}_2, \mathrm{I}_3)}{\partial \mathbb{B}}\mathbb{F} = 2\frac{\partial \hat{W}}{\partial \mathbb{B}}(\lambda_1(\mathbb{B}), \lambda_2(\mathbb{B}), \lambda_3(\mathbb{B}))2\frac{\partial \hat{W}}{\partial \lambda_i}\frac{\partial \lambda_i(\mathbb{B})}{\partial \mathbb{B}}\mathbb{F}.$$

What (in the hell) is $\frac{\partial \lambda_i(\mathbb{B})}{\partial \mathbb{B}}$? [22]

$$\mathbb{B}(s) = \sum_{\alpha=1}^{3} \omega_\alpha(s)\mathbf{g}_\alpha(s) \otimes \mathbf{g}_\alpha(s), \forall s \in I$$

where I is some open interval and $\{\mathbf{g}_\alpha\}$ is an ON eigenbasis of $\mathbb{B}$. Next, realize

$$\omega_1(s) = \mathbf{g}_1(s) \cdot \mathbb{B}(s)\mathbf{g}_1(s),$$

and differentiate this:

$$\frac{\mathrm{d}\omega(s)}{\mathrm{d}s} = \frac{\mathrm{d}\mathbf{g}_1}{\mathrm{d}s} \cdot \mathbb{B}\mathbf{g}_1 + \mathbf{g}_1 \frac{\mathrm{d}\mathbb{B}}{\mathrm{d}s}\mathbf{g}_1 + \mathbf{g}_1 \cdot \mathbb{B}\frac{\mathrm{d}\mathbf{g}}{\mathrm{d}s} = \frac{1}{2} + +0.$$

10.5 Hyperelasticity with constraints

Very often, some other considerations have taken to be in account when describing some materials. Examples include

- *incompressibility*: $\det \mathbb{F} = 1$,
- *inextensibility*:$\mathbf{l} \cdot \mathbb{C}\mathbf{l} = 1$, for some $\mathbf{l} \in \mathbb{R}^3$ (*i.e.*, T.A. materials)

10.6 Rational thermodynamics

In rational thermodynamics, we *postulate*:

$$\mathbf{J}_\eta = \frac{1}{\theta_R}\mathbf{Q}, R_\eta = \frac{R}{\theta}, \tag{14}$$

i.e., the flux/production of entropy is the flux/production of heat divided by temperature. This *makes sense to assume.*

10.6.1 Clausius-Duhem inequality

Recall *balance of mass*:

$$\rho(t, \mathbf{x}) \det \mathbb{F} = \rho_R(\mathbf{X}),$$

balance of momentum:

$$\rho_r(\mathbf{x})\frac{\partial^2 \boldsymbol{\chi}}{\partial t^2} = \nabla \cdot \mathbb{T} + \rho_R \mathbf{B}(t, \mathbf{X}),$$

balance of internal energy:

$$\rho_r \frac{\partial e_r}{\partial t}(t, \mathbf{X}) = \underbrace{-\nabla \cdot \mathbf{Q}}_{=-\nabla\cdot(\det \mathbb{F}\mathbb{F}^{-1}\mathbf{q}(t,\mathbf{x}))} + \dot{\mathbb{F}} : \mathbb{T} + \rho_R R(t, \mathbf{X}),$$

where also

$$\dot{\mathbb{F}} : \mathbb{T} = \frac{1}{2}\mathbb{S} : \dot{\mathbb{C}},$$

and the *balance of entropy*:

[22] Recall the Daleckii-Krein theorem:

$$\rho_R \frac{\partial \eta_R}{\partial t} = -\nabla \cdot \mathbf{J}_\eta + \rho_R R_\eta + \xi_R.$$

The definition of the Helmholtz free energy is: $\psi = e - \theta\eta$, or in the reference configuration:

$$\psi_R = e_R - \theta_R \eta_R.$$

Take the time derivative and calculate

$$\begin{aligned}\rho_R \dot{\psi}_R = \rho_R\big(\dot{e}_R - \dot{\theta}\eta - \theta\dot{\eta}\big) &= \nabla \cdot \mathbf{Q} + \dot{\mathbb{F}} : \mathbb{T} + \rho_R\big(R - \dot{\theta}\eta\big) - \theta(-\nabla \cdot \mathbf{J}_\eta + \rho_R R_\eta + \xi_R) = \\ &= -\nabla \cdot \mathbf{Q} + \dot{\mathbb{F}} : \mathbb{T} - \rho_R \dot{\theta_R}\eta + \frac{\theta_R}{\theta_R}\nabla \cdot (\mathbf{Q}) - \theta_R \xi_R + \theta_R \nabla\left(\frac{1}{\theta_R}\right) \cdot \mathbf{Q},\end{aligned}$$

where we have used 14 to cancel some terms. In total we obtain

$$\boxed{\rho_R\big(\dot{\psi_R} + \eta_R \dot{\theta_R}\big) - \dot{\mathbb{F}} : \mathbb{T} - \theta_R \mathbf{Q} \cdot \nabla\left(\frac{1}{\theta_R}\right) = -\theta_R \xi_R \leq 0. \qquad (15)}$$

Rational thermodynamics also *postulates* Clausius-Duhem inequality holds for all thermodynamically admissable proccesses.

10.6.2 Isothermal setting

Let us consider a special case - an *isothermal setting*, meaning $0 = \dot{\theta} = \nabla\left(\frac{1}{\theta_R}\right) = 0$. The Clausius-Duhem inequality then becomes:

$$\rho_R \dot{\psi_R} - \dot{\mathbb{F}} : \mathbb{T} \leq 0,$$

for all admissable proccesses. Realize that this is the same as:

$$\frac{\partial(\rho_R \psi_R)}{\partial t} - \frac{1}{2}\mathbb{S} : \dot{\mathbb{C}},$$

and that $\rho_R \psi_R = W$. By hyperelasticity $W = W(\mathbb{F})$ and so from material frame indifference $W = \overline{W}(\mathbb{C})$. Taking the time derivative means:

$$\left(\frac{\partial W}{\partial \mathbb{C}} - \frac{1}{2}\dot{\mathbb{T}}\right) : \dot{\mathbb{C}} \leq 0, \forall \mathbb{C}, \dot{\mathbb{C}}$$

and that can only be met when

$$2\frac{\partial W}{\partial \mathbb{C}} = \mathbb{S},$$

since $\mathbb{C}, \dot{\mathbb{C}}$ are independent.

Really, define the motion

$$\boldsymbol{\chi}(t, \mathbf{X}) = \mathbf{X}_0 = \exp\left((t - t_0)\mathbb{D}\right)\mathbb{F}_0(\mathbf{X} - \mathbf{X}_0),$$

then

$$\mathbb{F}(t, \mathbf{X}) = \exp\left((t - t_0)\mathbb{D}\right)\mathbb{F}_0,$$

so $\mathbb{F}(t_0, \mathbf{X}_0) = \mathbb{F}_0$. Time derivative can be computed to be:

$$\dot{\mathbb{F}}(t, \mathbf{X}) = \mathbb{D} \exp\left((t - t_0)\right)\mathbb{F},$$

and $\dot{\mathbb{F}}(t_0, \mathbf{X}_0) = \mathbb{D}\mathbb{F}_0$. This means

$$\mathbb{C}(t_0, \mathbf{X}_0) = \mathbb{F}_0^\top \mathbb{F}_0,$$

and

$$\dot{\mathbb{C}}(t_0, \mathbf{X}_0) = 2\mathbb{F}_0^\top \mathbb{D}\mathbb{F}_0,$$

We see that we can choose $\mathbb{C}(t_0, \mathbf{X}_0)$ and $\dot{\mathbb{C}}(t_0, \mathbf{X}_0)$ independetly.

Suppose we are given the constraint

$$f(\mathbb{C}) = 0,$$

which fits the conditions

$$e.g. \, \mathbf{l} \cdot \mathbb{C}\mathbf{l} - 1 = 0, \det \mathbb{C} - 1 = 0.$$

Thus the Clausis-Duhem inequality with constraints reduces to:

$$\left(\frac{\partial W}{\partial \mathbb{C}} - \frac{1}{2}\dot{\mathbb{S}}\right) : \dot{\mathbb{C}} \le 0, \forall \mathbb{C}, \dot{\mathbb{C}} \, s.t. \, f(\mathbb{C}) = 0. \tag{16}$$

The condition is "almost equivalent" to

$$\frac{\partial f}{\partial \mathbb{C}} : \dot{\mathbb{C}} = 0,$$

which is convenient, as we have the following theorem.

Theorem 8. *Let $\mathbb{A} \in \mathbb{R}^{n\times m}, \mathbf{b} \in \mathbb{R}^{n\times 1}, \boldsymbol{\alpha} \in \mathbb{R}^{m\times 1}, \beta \in \mathbb{R}$ such that*

$$\mathbb{A}\mathbf{x} + \mathbf{b} = 0,$$
$$\boldsymbol{\alpha} \cdot \mathbf{x} + \beta \le 0,$$

for some $\mathbf{x} \in \mathbb{R}^m$. Let S be the set of solutions of the equation and assume it is nonempty. Then the following are equivalent:

- *$\forall \mathbf{x} \in S$ the equation holds*
- *$\exists \lambda \in \mathbb{R}^n \ne 0 \, s.t. \, \boldsymbol{\alpha}^\top - \lambda^\top \mathbb{A} = 0, \beta - \lambda \cdot \mathbf{b} \le 0.$*

Remark. In our case, we have

$$\mathbf{b} = 0, \mathbb{A} = \frac{\partial f}{\partial \mathbb{C}}, \mathbf{x} = \dot{\mathbb{C}}.$$

Using this it can be shown that under this constraint the Cauchy stress must take the form[23]

$$\mathbb{T}^y = \lambda \mathbb{I} + 2\mathbb{F}\frac{\partial W}{\partial \mathbb{C}}\mathbb{F}^\top.$$

We usually identify $\lambda = p_{\text{th}}$, with the thermodynamically determined stress.

Theorem 9. *The followig statements are equivalent:*

- *$\forall \mathbf{x} \in S = \{\mathbf{x} : \mathbb{A}\mathbf{x} + \mathbf{b} = \mathbf{0}\} : \boldsymbol{\alpha} \cdot \mathbf{x} + \beta \ge 0,$*
- *$\exists \boldsymbol{\lambda} \ne \mathbf{0} \, s.t. \, \boldsymbol{\alpha}^\top - \boldsymbol{\lambda}^\top \mathbb{A} = \mathbf{0} \wedge b - \lambda \cdot \beta \ge 0$*

[23] use $\mathbb{T}^y = \det \mathbb{F}\mathbb{F}^{-1}\mathbb{T}^y\mathbb{F}^{-\top}$, differentiate and realize $\frac{\partial f}{\partial \mathbb{C}} = \det \mathbb{C} rcg^{-\top}, \det \mathbb{F} = 1 = \det \mathbb{C}$, and plug this in.

Proof. first ii) $\Rightarrow$ i): multiply the first row by $\mathbf{x}$: $\boldsymbol{\alpha}\cdot\mathbf{v}-\boldsymbol{\lambda}\cdot\mathbb{A}\mathbf{x}=0$, sum it up with the second inequality and obtain

$$\boldsymbol{\alpha}\cdot\mathbf{x}+\beta-\boldsymbol{\lambda}\cdot(\mathbb{A}\mathbf{x}+\mathbf{b})\geq 0,$$

so when $\mathbf{x}\in S, \mathbb{A}\mathbf{x}+\mathbf{b}=\mathbf{0}$, and we obtain

$$\boldsymbol{\alpha}\cdot\mathbf{x}+\beta\geq 0.$$

Now i) $\Rightarrow$ ii). It suffices to show i)

$$\Rightarrow \exists\boldsymbol{\lambda}\neq\mathbf{0}\, s.t.\, \boldsymbol{\alpha}^\top-\boldsymbol{\lambda}^\top\mathbb{A}=0,$$

since if $\boldsymbol{\alpha}\cdot\mathbf{x}+\beta\geq 0\forall\mathbf{x}\in S$, then $\boldsymbol{\alpha}\cdot\mathbf{x}-\boldsymbol{\lambda}\cdot\mathbb{A}\mathbf{x}=0$, where $\mathbb{A}\mathbf{x}=-\mathbf{b}\forall\mathbf{x}\in S$. This immediately implies the sought result. This proof is by contradiction: suppose

$$(\mathbb{A}\mathbf{x}+\mathbf{b}=0\Rightarrow\boldsymbol{\alpha}\cdot\mathbf{x}+\beta\geq 0)\wedge\exists\mathbf{x}_0\, s.t.\, \mathbb{A}\mathbf{x}_0=\mathbf{0}\wedge\boldsymbol{\alpha}\cdot\mathbf{x}_0\neq 0.$$

We can now take

$$\mathbb{A}(\mathbf{x}+\delta\mathbf{x}_0)+\mathbf{b}=\mathbf{0}\Rightarrow\boldsymbol{\alpha}\cdot(\mathbf{x}+\delta\mathbf{x}_0)+\beta\geq 0,$$

for an arbitrary $\delta\in\mathbb{R}$. But this is clearly not possible, as we can take $\delta<0, |\delta|>>1$ and surely the second relation will not be met. ☐

10.7 Inflation of a hyperelastic balloon

To prepare ourselves, first we examine the *biaxial deformation of a incompressible hyperelastic sheet.*

10.7.1 Biaxial deformation of a incompressible hyperelastic sheet

The deformation gradient is

$$\mathbb{F}=\begin{bmatrix}\lambda_1 & 00 & \\ 0 & \lambda_2 & 0\\ 0 & 0 & \lambda_3\end{bmatrix}, \mathbb{B}=\mathbb{F}\mathbb{F}^\top=\begin{bmatrix}\lambda_1^2 & 00 & \\ 0 & \lambda_2^2 & 0\\ 0 & 0 & \lambda_3^2\end{bmatrix}.$$

Moreover, assume the material is the incompressible Ogden:

$$W(\lambda_1,\lambda_2,\lambda_3)=\sum_{k=1}^{N}\frac{\mu_k}{\alpha_k}\left(\lambda_1^{\alpha_k}+\lambda_2^{\alpha_k}+\lambda_3^{\alpha_k}-3\right),$$
$$\lambda_1\lambda_2\lambda_3=1.$$

Cauchy stress then would be

$$\mathbb{T}^y=-p\mathbb{I}+\frac{2}{J}\frac{\partial W(\mathbb{B})}{\partial\mathbb{B}}\mathbb{B}=-p\mathbb{I}+2\sum_{j=1}^{3}\frac{\partial W}{\partial\lambda_j}\frac{\partial\lambda_j}{\partial\mathbb{B}}\mathbb{B},$$

recall that

$$J=1, \frac{\partial\lambda_j}{\partial\mathbb{B}}=\frac{1}{2\lambda_j}\mathbf{g}_j\otimes\mathbf{g}_j, \mathbb{B}=\sum_{j=1}^{3}\lambda_j^2(\mathbf{g}_j\otimes\mathbf{g}_j.)$$

We must calculate

$$\frac{\partial W}{\partial\lambda_j}=\sum_{k=1}^{N}\frac{\mu_k}{\lambda_k}\alpha_k\lambda_j^{\alpha_k-1},$$

and so

$$\mathbb{T}^y = -p\mathbb{I} + 2\sum_{j=1}^{3}\sum_{k=1}^{N}\mu_k\frac{1}{2}\lambda_j^{\alpha_k-2}(\mathbf{g}_j\otimes\mathbf{g}_j)\sum_{l=1}^{3}\lambda_l^2(\mathbf{g}_l\otimes\mathbf{g}_l) = -p\mathbb{I} + \sum_{j=1}^{3}\sum_{k=1}^{N}\mu_k\lambda_j^{\alpha_k}(\mathbf{g}_j\otimes\mathbf{g}_j).$$

We now assume

$$T_{33} = 0,$$

called the *thin sheet assumption*, *i.e.*, plane-stress problem. This means

$$0 = -p + \sum_{k=1}^{N}\mu_k\lambda_3^{\alpha_k},$$

since $\mathbf{g}_j = \mathbf{e}_j$. The pressure thus is

$$p = \sum_{k=1}^{N}\mu_k\lambda_3^{\alpha_k}.$$

The remaining stresses are

$$T_{11} = -\sum_{k=1}^{N}\mu_k\lambda_3^{\alpha_k} + \sum_{k=1}^{N}\mu_k\lambda_1^{\alpha_k},$$

where

$$\lambda_3 = \frac{1}{\lambda_1\lambda_2},$$

so

$$T_{11} = \sum_{k=1}^{N}\mu_k\left(\lambda_1^{\alpha_k} - (\lambda_1\lambda_2)^{-\alpha_k}\right),$$

and similiarly

$$T_{22} = \sum_{k=1}^{N}\mu_k\left(\lambda_2^{\alpha_k} - (\lambda_1\lambda_2)^{-\alpha_k}\right).$$

Alltogether,

$$\mathbb{T}^y = T_{11}(\mathbf{e}_1\otimes\mathbf{e}_1) + T_{22}(\mathbf{e}_2\otimes\mathbf{e}_2) \equiv \sigma_1(\mathbf{e}_1\otimes\mathbf{e}_2) + \sigma_2(\mathbf{e}_2\otimes\mathbf{e}_2).$$

10.7.2 Simplified approach for a balloon

Assume now

$$\sigma_1 = \sigma_2 \equiv \sigma,$$

meaning the the balloon is being stretched the same way in both directions. This is equivalent to

$$\lambda_1 = \lambda_2 \equiv \lambda,$$

i.e.,

$$\sigma = \sum_{k=1}^{N}\mu_k\left(\lambda^{\alpha_k} - \lambda^{-2\alpha_k}\right).$$

Let the thickness of the baloon be h and assume $h << 1$. We also define

$$\sigma_T = \sigma h,$$

as in fact the *surface tension.* The virtual work principle states

$$p_0 \delta V = \sigma_T \delta S,$$

where p_0 is the overpressure. This can be manioulated into

$$p_0 \frac{4}{3}\pi 3\pi r^2 \delta r = \sigma_T 2\pi r \delta r$$
$$p_0 r^2 = 2r\sigma_T,$$

and so

$$p_0 = \frac{2\sigma_T}{r} = 2\sigma_T K,$$

which is the *Laplace-Young* condition. The pressure in the baloon is the greater the less the radius the balloon has, or the greater the curvature K gets.

Subsituting for σ_T yields

$$p_0 = \frac{2\sigma h}{r} = 2\sigma \underbrace{\left(\frac{h}{H}\right)}_{=\lambda_3=\frac{1}{\lambda^2}} \frac{H}{R} \underbrace{\left(\frac{R}{r}\right)}_{=\frac{1}{\lambda}=\frac{2\sigma}{\lambda^3}} = \frac{2H}{R\lambda^3} \sum_{k=1}^{N} \mu_k \left(\lambda^{\alpha_k} - \lambda^{-2\lambda_k}\right),$$

where H, R are the reference thickness and radius and h, r are the thickness and radius in the deformed configuration. so finally

$$p_0 = \frac{2H}{R} \sum_{k=1} \mu_k \left(\lambda^{\alpha_k - 3} - \lambda^{-2\lambda_k - 3}\right).$$

Plotting this for a rubber-like material, the dependency $p_0(\lambda)$ shows that first, starting from 0, p_0 is very steep, but suddenly at a one time the material expands very rapidly.

10.7.3 Exact solution

Denote now A, B, H to be the inner radius, outer radius and the thickness, the same for a, b, h. It will be advantegous to use spherical coordinates:

$$R \in [A, B],$$
$$\Theta \in [0, \pi),$$
$$\Phi \in [0, 2\pi),$$

The deformation is

$$r = f(R)R,$$
$$\theta = \Theta,$$
$$\varphi = \Phi$$

This gets sophisticated now, as

$$\mathbb{F} = \frac{\partial \xi^i}{\partial X^J} \mathbf{g}_i \otimes \mathbf{G}^J, \mathbf{g}_i = \frac{\partial \mathbf{x}\left(\xi^1, \xi^2, \xi^3\right)}{\partial \xi^i},$$

in curvilinear coordinates. It can be obtained:

$$\mathbb{F} = \frac{\partial(f(R)R)}{\partial R}\big(\mathbf{g}_r \otimes \mathbf{G}^R\big) + \mathbf{g}_\theta \otimes \mathbf{G}^\Theta + \mathbf{g}_\varphi \otimes \mathbf{G}^\Phi,$$

but every decent person works in coordinate *s.t.* $\|\mathbf{g}_r\| = 1$ etc. Calculation gives

$$\|\mathbf{G}_R\| = 1, \|\mathbf{G}_\Theta\| = R, \|\mathbf{G}_\Phi\| = R\sin\Theta,$$

and the inverse for the forms. Now we write the deformation gradient in the "normalized" coordinates, without writing things like $\mathbf{g}_{\hat{r}}$.

$$\mathbb{F} = Rf'(R)\big(\mathbf{g}_r \otimes \mathbf{G}^R\big) + f(R)\mathbb{I}.$$

After long and complicated calculations, it can be shown

$$p_0 = \int_{\lambda_a}^{\lambda_b} \frac{\tilde{W}'(\lambda)}{\lambda^3 - 1}\,\mathrm{d}\lambda\,,$$

where $\tilde{W} = W\big(\frac{1}{\lambda^2}, \lambda, \lambda\big)$.

References

Kružík, Martin and Tomáš Roubíček (2019a). "Elastic Materials". In: *Mathematical Methods in Continuum Mechanics of Solids*. Ed. by Martin Kružík and Tomáš Roubíček. Cham: Springer International Publishing, pp. 25–50. ISBN: 978-3-030-02065-1. DOI: 10.1007/978-3-030-02065-1_2. URL: https://doi.org/10.1007/978-3-030-02065-1_2 (visited on 06/18/2025).

– (2019b). "Polyconvex Materials: Existence of Energy-Minimizing Deformations". In: *Mathematical Methods in Continuum Mechanics of Solids*. Ed. by Martin Kružík and Tomáš Roubíček. Cham: Springer International Publishing, pp. 51–86. ISBN: 978-3-030-02065-1. DOI: 10.1007/978-3-030-02065-1_3. URL: https://doi.org/10.1007/978-3-030-02065-1_3 (visited on 06/18/2025).