

Part I
Nonlinear Continuum
Mechanics

COPYRIGHTED MATERIAL

KINEMATICS OF DEFORMABLE BODIES

Continuum mechanics models the physical universe as a collection of “deformable bodies,” a concept that is easily accepted from our everyday experiences with observable phenomena. Deformable bodies occupy regions in three-dimensional Euclidean space \mathcal{E} , and a given body will occupy different regions at different times. The subsets of \mathcal{E} occupied by a body \mathcal{B} are called its *configurations*. It is always convenient to identify one configuration in which the geometry and physical state of the body are known and to use that as the *reference configuration*; then other configurations of the body can be characterized by comparing them with the reference configuration (in ways we will make precise later).

For a given body, we will assume that the reference configuration is an open, bounded, connected subset Ω_0 of \mathbb{R}^3 with a smooth boundary $\partial\Omega_0$. The body is made up of physical points called *material points*. To identify these points, we assign each a vector \mathbf{X} and we identify the components of \mathbf{X} as the coordinates of the place occupied by the material point when the body is in its reference configuration relative to a fixed Cartesian coordinate system.

It is thus important to understand that the body \mathcal{B} is a *non-denumerable* set of material points \mathbf{X} . This is the fundamental hypoth-

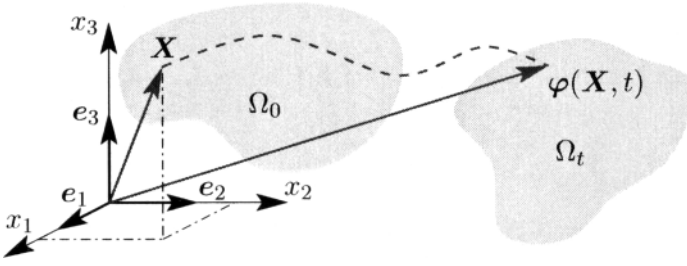


Figure 1.1: Motion from the reference configuration Ω_0 to the current configuration Ω_t .

esis of continuum mechanics: Matter is not discrete; it is continuously distributed in one-to-one correspondence with points in some subset of \mathbb{R}^3 . Bodies are thus “continuous media”: The components of \mathbf{X} with respect to some basis are real numbers. Symbolically, we could write

$$\mathcal{B} = \{\mathbf{X}\} \sim \{\mathbf{X} = X_i \mathbf{e}_i \in \overline{\Omega_0}\}$$

for some orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ and origin $\mathbf{0}$ chosen in three-dimensional Euclidean space and, thus, identified with \mathbb{R}^3 . Hereafter, repeated indices are summed throughout their ranges; i.e. the “summation convention” is employed.

Kinematics is the study of the motion of bodies, without regard to the causes of the motion. It is purely a study of geometry and is an exact science within the hypothesis of a *continuum* (a continuous media).

1.1 Motion

We pick a point $\mathbf{0}$ in \mathbb{R}^3 as the origin of a fixed coordinate system $(x_1, x_2, x_3) = \mathbf{x}$ defined by orthonormal vectors \mathbf{e}_i , $i = 1, 2, 3$. The system (x_1, x_2, x_3) is called the *spatial* coordinate system. When the physical body \mathcal{B} occupies its reference configuration Ω_0 at, say, time $t = 0$, the material point \mathbf{X} occupies a position (place) corresponding to the vector $\mathbf{X} = X_i \mathbf{e}_i$. The spatial coordinates (X_1, X_2, X_3) of \mathbf{X}

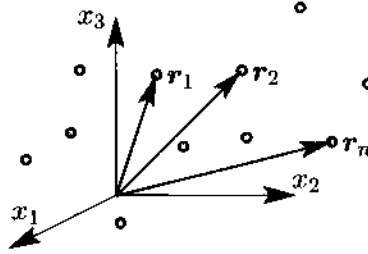


Figure 1.2: A discrete set of material particles.

are *labels* that identify the material point. The coordinate labels X_i are sometimes called *material coordinates* (see Fig. 1.1).

Remark Notice that if there were a *countable* set of discrete material points, such as one might use in models of molecular or atomistic dynamics, the particles (discrete masses) could be labeled using natural numbers $n \in \mathbb{N}$, as indicated in Fig. 1.2. But the particles (material points) in a continuum are not countable, so the use of a label of three real numbers for each particle corresponding to the coordinates of their position (at $t = 0$) in the reference configuration seems to be a very natural way to identify such particles. \square

The body moves through \mathcal{E} over a period of time and occupies a configuration $\Omega_t \subset \mathbb{R}^3$ at time t . Thus, material points X in $\overline{\Omega_0}$ (the closure of Ω_0) are mapped into positions x in $\overline{\Omega_t}$ by a smooth vector-valued mapping (see Fig. 1.1)

$$\boxed{x = \varphi(X, t).} \quad (1.1)$$

Thus, $\varphi(X, t)$ is the spatial position of the material point X at time t . The one-parameter family $\{\varphi(X, t)\}$ of positions is called the *trajectory* of X . We demand that φ be differentiable, injective, and orientation preserving. Then φ is called the *motion* of the body:

1. Ω_t is called the current configuration of the body.

2. φ is injective (except possibly at the boundary $\partial\Omega_0$ of Ω_0).
3. φ is *orientation preserving* (which means that the physical material cannot penetrate itself or reverse the orientation of material coordinates, which means that $\det \nabla\varphi(\mathbf{X}, t) > 0$).

Hereafter we will not explicitly show the dependence of φ and other quantities on time t unless needed; this time dependency is taken up later.

The vector field

$$\boxed{\mathbf{u} = \varphi(\mathbf{X}) - \mathbf{X}} \quad (1.2)$$

is the *displacement* of point \mathbf{X} . Note that

$$d\mathbf{x} = \nabla\varphi(\mathbf{X}) d\mathbf{X} \quad \left(\text{i.e., } dx_i = \frac{\partial\varphi_i}{\partial X_j} dX_j \right).$$

The tensor

$$\boxed{\mathbf{F}(\mathbf{X}) = \nabla\varphi(\mathbf{X})} \quad (1.3)$$

is called the *deformation gradient*. Clearly,

$$\boxed{\mathbf{F}(\mathbf{X}) = \mathbf{I} + \nabla\mathbf{u}(\mathbf{X})}, \quad (1.4)$$

where \mathbf{I} is the identity tensor and $\nabla\mathbf{u}$ is the *displacement gradient*.

Some Definitions

- A *deformation* is homogeneous if $\mathbf{F} = \mathbf{C} = \text{constant}$.
- A motion is *rigid* if it is the sum of a translation \mathbf{a} and a rotation \mathbf{Q} :

$$\varphi(\mathbf{X}) = \mathbf{a} + \mathbf{Q}\mathbf{X},$$

where $\mathbf{a} \in \mathbb{R}^3$, $\mathbf{Q} \in \mathbb{O}_+^3$, with \mathbb{O}_+^3 the set of orthogonal matrices of order 3 with determinant equal to +1.

- As noted earlier, the fact that the motion is orientation preserving means that

$$\det \nabla \varphi(\mathbf{X}) > 0 \quad \forall \mathbf{X} \in \overline{\Omega}_0.$$

- Recall that

$$\text{Cof } \mathbf{F} = \text{cofactor matrix (tensor) of } \mathbf{F} = \det \mathbf{F} \mathbf{F}^{-T}.$$

For any matrix $\mathbf{A} = [A_{ij}]$ of order n and for each row i and column j , let \mathbf{A}'_{ij} be the matrix of order $n - 1$ obtained by deleting the i th row and j th column of \mathbf{A} . Let $d_{ij} = (-1)^{i+j} \det \mathbf{A}'_{ij}$. Then the matrix

$$\text{Cof } \mathbf{A} = [d_{ij}]$$

is the cofactor matrix of \mathbf{A} and d_{ij} is the (i, j) cofactor of \mathbf{A} . Note that

$$\mathbf{A}(\text{Cof } \mathbf{A})^T = (\text{Cof } \mathbf{A})^T \mathbf{A} = (\det \mathbf{A})\mathbf{I}. \quad (1.5)$$

1.2 Strain and Deformation Tensors

A differential material line segment in the reference configuration is

$$dS_0^2 = d\mathbf{X}^T d\mathbf{X} = dX_1^2 + dX_2^2 + dX_3^2,$$

while the same material line in the current configuration is

$$dS^2 = d\mathbf{x}^T d\mathbf{x} = d\mathbf{X}^T \mathbf{F}^T \mathbf{F} d\mathbf{X}.$$

The tensor

$$\mathbf{C} = \mathbf{F}^T \mathbf{F} = \text{the right Cauchy–Green deformation tensor}$$

is thus a measure of the change in dS_0^2 due to (gradients of) the motion

$$dS^2 - dS_0^2 = d\mathbf{X}^T \mathbf{C} d\mathbf{X} - d\mathbf{X}^T d\mathbf{X}.$$

\mathbf{C} is symmetric, positive definite. Another deformation measure is simply

$$dS^2 - dS_0^2 = d\mathbf{X}^T(2\mathbf{E})d\mathbf{X},$$

where

$$\mathbf{E} = \frac{1}{2}(\mathbf{C} - \mathbf{I}) = \text{the Green-St. Venant strain tensor.} \quad (1.6)$$

Since $\mathbf{F} = \mathbf{I} + \nabla\mathbf{u}$ and $\mathbf{C} = \mathbf{F}^T\mathbf{F}$, we have

$$\mathbf{E} = \frac{1}{2}(\nabla\mathbf{u} + \nabla\mathbf{u}^T + \nabla\mathbf{u}^T\nabla\mathbf{u}). \quad (1.7)$$

The tensor

$$\mathbf{B} = \mathbf{F}\mathbf{F}^T = \text{the left Cauchy-Green deformation tensor}$$

is also symmetric and positive definite, and we can likewise define

$$dS^2 - dS_0^2 = d\mathbf{X}^T\mathbf{F}^T(2\hat{\mathbf{A}})\mathbf{F}d\mathbf{X} = d\mathbf{x}^T(2\hat{\mathbf{A}})d\mathbf{x},$$

where

$$\hat{\mathbf{A}} = \frac{1}{2}(\mathbf{I} - \mathbf{B}^{-1}) = \text{the Almansi-Hamel strain tensor} \quad (1.8)$$

or

$$\hat{\mathbf{A}} = \frac{1}{2}(\text{grad } \mathbf{u} + \text{grad } \mathbf{u}^T - \text{grad } \mathbf{u} \text{ grad } \mathbf{u}^T), \quad (1.9)$$

where $\text{grad } \mathbf{u}$ is the spatial gradient $\text{grad } \mathbf{u} = \partial\mathbf{u}/\partial\mathbf{x}$ (i.e., $(\text{grad } \mathbf{u})_{ij} = \partial u_i(\mathbf{X}(\mathbf{x}), t)/\partial x_j$); see also Sec. 1.3.

Interpretation of E Take $dS_0 = dX_1$ (i.e., $d\mathbf{X} = (dX_1, 0, 0)^T$). Then

$$dS^2 - dS_0^2 = dS^2 - dX_1^2 = (1 + 2E_{11}) dX_1^2,$$

so

$$E_{11} = \frac{1}{2} \left(\left(\frac{ds}{dX_1} \right)^2 - 1 \right) = \begin{cases} \text{a measure of the stretch of a} \\ \text{material line originally oriented} \\ \text{in the } X_1 \text{ direction in } \Omega_0. \end{cases}$$

We call e_1 the extension in the X_1 direction at \mathbf{X} (which is a dimensionless measure of change in length per unit length)

$$e_1 \stackrel{\text{def}}{=} \frac{dS - dX_1}{dX_1} = \sqrt{1 + 2E_{11}} - 1,$$

or

$$2E_{11} = (1 + e_1)^2 - 1.$$

Similar definitions apply to E_{22} and E_{33} .

Now take $d\mathbf{X} = (dX_1, dX_2, 0)^T$ and

$$\cos \theta = \frac{d\mathbf{x}_1 \cdot d\mathbf{x}_2}{\|d\mathbf{x}_1\| \|d\mathbf{x}_2\|} = \frac{C_{12}}{\sqrt{1 + 2E_{11}} \sqrt{1 + 2E_{22}}} \quad (\text{Exercise}).$$

The *shear* (or *shear strain*) in the X_1 - X_2 plane is defined by the angle change (see Fig. 1.3),

$$\gamma_{12} \stackrel{\text{def}}{=} \frac{\pi}{2} - \theta.$$

Therefore

$$\boxed{\sin \gamma_{12} = \frac{2E_{12}}{\sqrt{1 + 2E_{11}} \sqrt{1 + 2E_{22}}}.} \quad (1.10)$$

Thus, E_{12} (and, analogously, E_{13} and E_{23}) is a measure of the shear in the X_1 - X_2 (or X_1 - X_3 and X_2 - X_3) plane.

Small strains The tensor

$$\mathbf{e} = \frac{1}{2} (\nabla \mathbf{u} + \nabla \mathbf{u}^T), \quad (1.11)$$

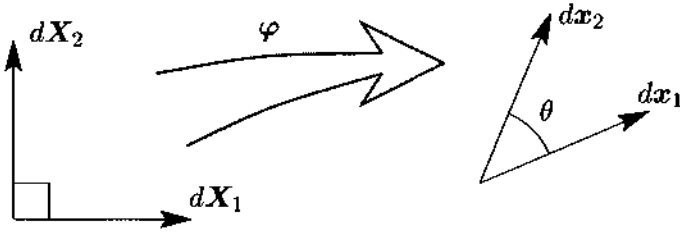


Figure 1.3: Change of angle through the motion φ .

is called the *infinitesimal* or *small* or *engineering strain tensor*. Clearly

$$\mathbf{E} = \mathbf{e} + \frac{1}{2} \nabla \mathbf{u}^T \nabla \mathbf{u}. \quad (1.12)$$

Note that if \mathbf{E} is “small” (i.e., $|E_{ij}| \ll 1$), then we obtain

$$\begin{aligned} e_1 &= (1 + 2E_{11})^{1/2} - 1 \\ &= 1 + E_{11} - 1 + 0(E_{11}^2) \\ &\approx E_{11} = e_{11}, \end{aligned}$$

that is,

$$e_{11} = e_1 = \frac{dS - dX_1}{dX_1}, \quad \text{etc.},$$

and

$$2e_{12} = \sin \gamma_{12} \approx \gamma_{12}, \quad \text{etc.}$$

Thus, small strains can be given the classical textbook interpretation: e_n is the change in length per unit length and e_{12} is the change in the right angle between material lines in the X_1 and X_2 directions. In the case of small strains, the Green–St. Venant strain tensor and the Almansi–Hamel strain tensor are indistinguishable.

1.3 Rates of Motion

If $\varphi(\mathbf{X}, t)$ is the motion (of \mathbf{X} at time t), i.e.,

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

then

$$\dot{\mathbf{x}} = \dot{\mathbf{x}}(\mathbf{X}, t) \stackrel{\text{def}}{=} \frac{\partial \varphi(\mathbf{X}, t)}{\partial t} \quad (1.13)$$

is the velocity and

$$\ddot{\mathbf{x}} = \ddot{\mathbf{x}}(\mathbf{X}, t) \stackrel{\text{def}}{=} \frac{\partial^2 \varphi(\mathbf{X}, t)}{\partial t^2} \quad (1.14)$$

is the acceleration. Since φ is (in general) bijective, we can also describe the velocity as a function of the place \mathbf{x} in \mathbb{R}^3 and time t :

$$\mathbf{v} = \mathbf{v}(\mathbf{x}, t) = \dot{\mathbf{x}}(\varphi^{-1}(\mathbf{x}, t), t).$$

This is called the *spatial description* of the velocity.

This leads to two different ways to interpret the rates of motion of continua:

- The *material* description (functions are defined on *material points* \mathbf{X} in the body \mathcal{B} in correspondence to points in \mathbb{R}^3);
- The *spatial* description (functions are defined on (spatial) *places* \mathbf{x} in \mathbb{R}^3).

When the equations of continuum mechanics are written in terms of the material description, the collective equations are commonly referred to as the *Lagrangian* form (formulation) of the equations (see Fig. 1.4). When the spatial description is used, the term *Eulerian* form (formulation) is used (see Fig. 1.5).

There are differences in the way rates of change appear in the Lagrangian and Eulerian formulations.

- In the Lagrangian case: Given a field $\psi_m = \psi_m(\mathbf{X}, t)$ (the subscript m reminding us that we presume ψ is a function of the material coordinates),

$$\frac{d\psi_m(\mathbf{X}, t)}{dt} = \frac{\partial \psi_m(\mathbf{X}, t)}{\partial t} + \frac{\partial \psi_m(\mathbf{X}, t)}{\partial \mathbf{X}} \cdot \frac{\partial \mathbf{X}}{\partial t},$$

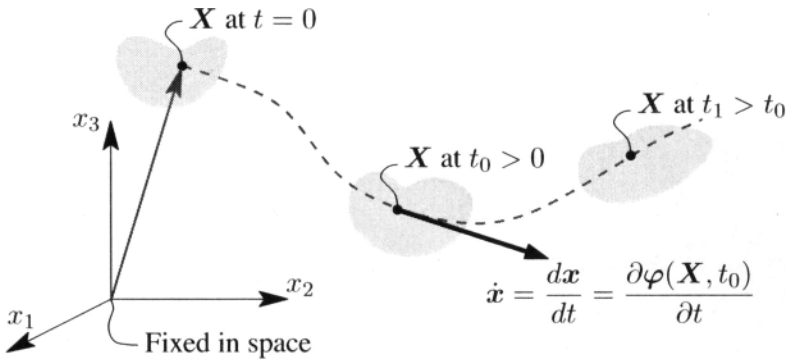


Figure 1.4: Lagrangian (material) description of velocity. The velocity of a material point is the time rate of change of the position of the point as it moves along its path (its trajectory) in \mathbb{R}^3 .

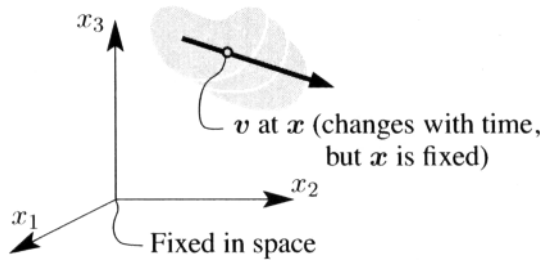


Figure 1.5: Eulerian (spatial) description of velocity. The velocity at a fixed place \mathbf{x} in \mathbb{R}^3 is the speed and direction (at time t) of particles flowing through the place \mathbf{x} .

but $\partial \mathbf{X} / \partial t = 0$ because \mathbf{X} is simply a label of a material point. Thus,

$$\boxed{\frac{d\psi_m(\mathbf{X}, t)}{dt} = \frac{\partial \psi_m(\mathbf{X}, t)}{\partial t}} \tag{1.15}$$

- In the Eulerian case: Given a field $\psi = \psi(\mathbf{x}, t)$,

$$\frac{d\psi(\mathbf{x}, t)}{dt} = \left. \frac{\partial\psi(\mathbf{x}, t)}{\partial t} \right|_{\mathbf{x} \text{ fixed}} + \frac{\partial\psi(\mathbf{x}, t)}{\partial \mathbf{x}} \cdot \frac{\partial \mathbf{x}}{\partial t},$$

but $\frac{\partial \mathbf{x}}{\partial t} = \mathbf{v}(\mathbf{x}, t)$ is the velocity at position \mathbf{x} and time t . Thus,

$$\boxed{\frac{d\psi(\mathbf{x}, t)}{dt} = \frac{\partial\psi(\mathbf{x}, t)}{\partial t} + \mathbf{v}(\mathbf{x}, t) \cdot \frac{\partial\psi(\mathbf{x}, t)}{\partial \mathbf{x}}.} \quad (1.16)$$

Notation We distinguish between the gradient and divergence of fields in the Lagrangian and Eulerian formulation as follows:

Lagrangian:	$\frac{\partial}{\partial \mathbf{X}} = \nabla = \text{Grad},$	$\frac{\partial}{\partial \mathbf{X}} \cdot \boldsymbol{\varphi} = \nabla \cdot \boldsymbol{\varphi} = \text{Div } \boldsymbol{\varphi};$
Eulerian:	$\frac{\partial}{\partial \mathbf{x}} = \text{grad},$	$\frac{\partial}{\partial \mathbf{x}} \cdot \mathbf{v} = \text{div } \mathbf{v}.$

In classical literature, some authors write

$$\boxed{\frac{D\psi}{Dt} = \frac{\partial\psi}{\partial t} + \mathbf{v} \cdot \text{grad } \psi} \quad (1.17)$$

as the “material time derivative” of a scalar field ψ , giving the rate of change of ψ at a fixed described place \mathbf{x} at time t . Thus, in the Eulerian formulation, the acceleration is

$$\mathbf{a} = \frac{D\mathbf{v}}{Dt} = \frac{\partial\mathbf{v}}{\partial t} + (\mathbf{v} \cdot \text{grad}) \mathbf{v},$$

\mathbf{v} being the velocity.

1.4 Rates of Deformation

The spatial (Eulerian) field

$$\boxed{\mathbf{L} = \mathbf{L}(\mathbf{x}, t) \stackrel{\text{def}}{=} \frac{\partial}{\partial \mathbf{x}} \mathbf{v}(\mathbf{x}, t) = \text{grad } \mathbf{v}(\mathbf{x}, t)} \quad (1.18)$$

is the *velocity gradient*. The time rate of change of the deformation gradient \mathbf{F} is

$$\begin{aligned}\dot{\mathbf{F}} &\equiv \frac{\partial}{\partial t} \nabla \varphi(\mathbf{X}, t) = \nabla \frac{\partial \varphi}{\partial t}(\mathbf{X}, t) \\ &= \frac{\partial}{\partial \mathbf{X}} \mathbf{v}(\mathbf{x}, t) = \frac{\partial \mathbf{v}}{\partial \mathbf{x}} \frac{\partial \mathbf{x}}{\partial \mathbf{X}} = \text{grad } \mathbf{v} \mathbf{F},\end{aligned}$$

or

$$\boxed{\dot{\mathbf{F}} = \text{grad } \mathbf{v} \mathbf{F} = \mathbf{L}_m \mathbf{F}}, \quad (1.19)$$

where $\mathbf{L}_m = \mathbf{L}$ is written in material coordinates, so

$$\mathbf{L}_m = \dot{\mathbf{F}} \mathbf{F}^{-1}. \quad (1.20)$$

It is standard practice to write \mathbf{L} in terms of its symmetric and skew-symmetric parts:

$$\boxed{\mathbf{L} = \mathbf{D} + \mathbf{W}}. \quad (1.21)$$

Here

$$\begin{aligned}\mathbf{D} &= \frac{1}{2}(\mathbf{L} + \mathbf{L}^T) = \text{the deformation rate tensor}, \\ \mathbf{W} &= \frac{1}{2}(\mathbf{L} - \mathbf{L}^T) = \text{the spin tensor}.\end{aligned} \quad (1.22)$$

We can easily show that if \mathbf{v} is the velocity field,

$$\mathbf{W} \mathbf{v} = \frac{1}{2} \boldsymbol{\omega} \times \mathbf{v}, \quad (1.23)$$

where $\boldsymbol{\omega}$ is the vorticity

$$\boldsymbol{\omega} = \text{curl } \mathbf{v}. \quad (1.24)$$

Recall (cf. Exercise 2.6) that

$$D(\det \mathbf{A}) : \mathbf{V} = (\det \mathbf{A}) \mathbf{V}^T : \mathbf{A}^{-1},$$

for any invertible tensor \mathbf{A} and arbitrary $\mathbf{V} \in L(V, V)$. Also, if $\mathbf{f}(\mathbf{g}(t)) = \mathbf{f} \circ \mathbf{g}(t)$ denotes the composition of functions \mathbf{f} and \mathbf{g} , the chain rule of differentiation leads to

$$\frac{d\mathbf{f}(\mathbf{g}(t))}{dt} = d\mathbf{f}(\mathbf{g}(t)) \cdot \frac{d\mathbf{g}(t)}{dt} = D\mathbf{f}(\mathbf{g}(t)) : \dot{\mathbf{g}}(t).$$

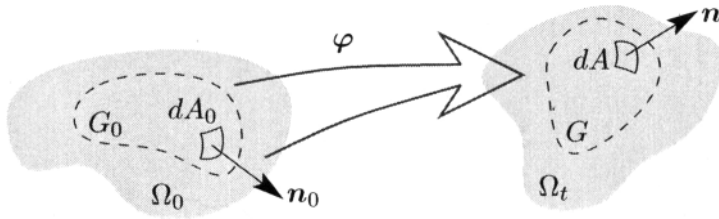


Figure 1.6: Mapping from reference configuration into current configuration.

Combining these expressions, we have

$$\begin{aligned} \frac{\dot{\det \mathbf{F}}}{\det \mathbf{F}} &= \frac{\partial \det \mathbf{F}}{\partial t} = D(\det \mathbf{F}) : \dot{\mathbf{F}} = \det \mathbf{F} \dot{\mathbf{F}}^T : \mathbf{F}^{-1} \\ &= \det \mathbf{F} \operatorname{tr} \mathbf{L}_m = \det \mathbf{F} \operatorname{div} \mathbf{v} \end{aligned}$$

(since $\dot{\mathbf{F}}^T : \mathbf{F}^{-1} = \operatorname{tr} \dot{\mathbf{F}} \mathbf{F}^{-1} = \operatorname{tr} \mathbf{L}_m$, where $\operatorname{tr} \mathbf{L} = \operatorname{tr} \operatorname{grad} \mathbf{v} = \operatorname{div} \mathbf{v}$).

Summing up:

$$\boxed{\frac{\dot{\det \mathbf{F}}}{\det \mathbf{F}} = \det \mathbf{F} \operatorname{div} \mathbf{v}.} \quad (1.25)$$

There is a more constructive way of deriving (1.25) using the definitions of determinant and cofactors of \mathbf{F} ; see Exercise 4 in Set I.2.

1.5 The Piola Transformation

The situation is this: A subdomain $G_0 \subset \Omega_0$ of the reference configuration of a body, with boundary ∂G_0 and unit exterior vector n_0 normal to the surface-area element dA_0 , is mapped by the motion φ into a subdomain $G = \varphi(G_0) \subset \Omega_t$ of the current configuration with boundary ∂G with unit exterior vector n normal to the “deformed” surface area dA (see Fig. 1.6).

Let $\mathbf{T} = \mathbf{T}(\mathbf{x}) = \mathbf{T}(\varphi(\mathbf{X}))$ denote a tensor field defined on G and $\mathbf{T}(\mathbf{x}) n(\mathbf{x})$ the flux of \mathbf{T} across ∂G , $n(\mathbf{x})$ being a unit normal to ∂G . Here Ω_t is fixed so t is held constant and not displayed. Corresponding to \mathbf{T} , a tensor field $\mathbf{T}_0 = \mathbf{T}_0(\mathbf{X})$ is defined on G_0 that associates the

flux $\mathbf{T}_0(\mathbf{X}) \mathbf{n}_0(\mathbf{X})$ through ∂G_0 , $\mathbf{n}_0(\mathbf{X})$ being the unit normal to ∂G_0 . We seek a relationship between $\mathbf{T}_0(\mathbf{X})$ and $\mathbf{T}(\mathbf{x})$ that will result in the same total flux through the surfaces ∂G_0 and ∂G , so that

$$\int_{\partial G_0} \mathbf{T}_0(\mathbf{X}) \mathbf{n}_0(\mathbf{X}) dA_0 = \int_{\partial G} \mathbf{T}(\mathbf{x}) \mathbf{n}(\mathbf{x}) dA, \quad (1.26)$$

with $\mathbf{x} = \varphi(\mathbf{X})$. This relationship between \mathbf{T}_0 and \mathbf{T} is called the *Piola transformation*.

Proposition 1.1 (Piola Transformation) *The above correspondence holds if*

$$\mathbf{T}_0(\mathbf{X}) = \det \mathbf{F}(\mathbf{X}) \mathbf{T}(\mathbf{x}) \mathbf{F}(\mathbf{X})^{-T} = \mathbf{T}(\mathbf{x}) \text{Cof } \mathbf{F}(\mathbf{X}). \quad (1.27)$$

Proof (This development follows that of Ciarlet [2]). We will use the Green's formulas (divergence theorems)

$$\int_{G_0} \text{Div } \mathbf{T}_0 dX = \int_{\partial G_0} \mathbf{T}_0 \mathbf{n}_0 dA_0$$

and

$$\int_G \text{div } \mathbf{T} dx = \int_{\partial G} \mathbf{T} \mathbf{n} dA,$$

where

$$\text{Div } \mathbf{T}_0 = \nabla \cdot \mathbf{T}_0 = \frac{\partial (\mathbf{T}_0)_{ij}}{\partial X_j} \mathbf{e}_i,$$

$$\text{div } \mathbf{T} = \frac{\partial}{\partial x_j} T_{ij} \mathbf{e}_i,$$

$$dx = dx_1 dx_2 dx_3 = \det \mathbf{F} dX = \det \mathbf{F} dX_1 dX_2 dX_3.$$

We will also need to use the fact that

$$\frac{\partial}{\partial X_j} (\text{Cof } \nabla \varphi)_{ij} = 0.$$

To show this, we first verify by direct calculation that

$$(\text{Cof } \mathbf{F})_{ij} = (\text{Cof } \nabla \boldsymbol{\varphi})_{ij} = \left(\frac{\partial}{\partial X_{j+i}} \varphi_{i+1} \right) \left(\frac{\partial}{\partial X_{j+2}} \varphi_{i+2} \right) - \left(\frac{\partial}{\partial X_{j+2}} \varphi_{i+1} \right) \left(\frac{\partial}{\partial X_{j+1}} \varphi_{i+2} \right),$$

where no summation is used. Then a direct computation shows that

$$\frac{\partial}{\partial X_j} (\text{Cof } \mathbf{F})_{ij} = 0.$$

Next, set

$$\mathbf{T}_0(\mathbf{X}) = \mathbf{T}(\mathbf{x}) \text{Cof } \mathbf{F}(\mathbf{X}).$$

Noting that

$$\frac{1}{\det \mathbf{F}} (\text{Cof } \mathbf{F})^T = \mathbf{F}^{-1}$$

and

$$\frac{\partial x_i}{\partial X_m} \cdot \frac{\partial X_m}{\partial x_j} = \frac{\partial x_i}{\partial x_j} = \delta_{ij},$$

we see that

$$(\text{Cof } \mathbf{F})_{ij} = \det \mathbf{F} (\mathbf{F}^{-1})_{ji} = \det \mathbf{F} \frac{\partial X_j}{\partial x_i}.$$

Thus,

$$\begin{aligned} \text{Div } \mathbf{T}_0(\mathbf{X}) &= \mathbf{e}_k \frac{\partial}{\partial X_k} (\mathbf{T}_0)_{ij} \mathbf{e}_i \otimes \mathbf{e}_j = \frac{\partial (\mathbf{T}_0)_{ij}}{\partial X_j} \mathbf{e}_i \\ &= \frac{\partial}{\partial X_j} ((\mathbf{T}(\mathbf{x}) \text{Cof } \mathbf{F}(\mathbf{X}))_{ij} \mathbf{e}_i \\ &= \frac{\partial \mathbf{T}_{im}(\mathbf{x})}{\partial x_r} \cdot \frac{\partial x_r}{\partial X_j} \cdot \text{Cof } \mathbf{F}(\mathbf{X})_{mj} \mathbf{e}_i \\ &\quad + \mathbf{T}_{im} \frac{\partial}{\partial X_j} \text{Cof } \mathbf{F}(\mathbf{X})_{mj} \mathbf{e}_i \xrightarrow{0} \\ &= \frac{\partial \mathbf{T}_{im}}{\partial x_r} \cdot \frac{\partial x_r}{\partial X_j} \det \mathbf{F} \frac{\partial X_j}{\partial x_m} \mathbf{e}_i \\ &= \frac{\partial \mathbf{T}_{ir}}{\partial x_r} \mathbf{e}_i \det \mathbf{F} \\ &= \text{div } \mathbf{T} \det \mathbf{F}, \end{aligned}$$

that is,

$$\boxed{\text{Div } \mathbf{T}_0 = \det \mathbf{F} \text{ div } \mathbf{T}.} \quad (1.28)$$

Thus

$$\begin{aligned} \int_{G_0} \text{Div } \mathbf{T}_0 dX &= \int_{G_0} \det \mathbf{F} \text{ div } \mathbf{T} dX = \int_{\partial G_0} \mathbf{T}_0 \mathbf{n}_0 dA_0, \\ \int_{\partial G_0} \mathbf{T}_0 \mathbf{n}_0 dA_0 &= \int_{G_0} \text{div } \mathbf{T} \det \mathbf{F} dX = \int_G \text{div } \mathbf{T} dx = \int_{\partial G} \mathbf{T} \mathbf{n} dA, \end{aligned}$$

as asserted. \square

Corollaries and Observations The Piola transformation provides a means for characterizing the flux of a field through a material surface in the current configuration in terms of the representation of the surface in the reference configuration. It also provides fundamental relationships between differential surface areas and their orientations in the reference and current configurations. We list a few of these as corollaries and observations.

- Since G_0 is arbitrary (symbolically), we obtain

$$\boxed{\mathbf{T}_0 \mathbf{n}_0 dA_0 = \mathbf{T} \mathbf{n} dA.} \quad (1.29)$$

- Set $\mathbf{T} = \mathbf{I} = \text{identity}$. Then

$$\boxed{\det \mathbf{F} \mathbf{F}^{-T} \mathbf{n}_0 dA_0 = \mathbf{n} dA.} \quad (1.30)$$

- Since $\mathbf{n} = \frac{dA_0}{dA} \cdot (\det \mathbf{F}) \mathbf{F}^{-T} \mathbf{n}_0$ and $\|\mathbf{n}\| = 1$, we have

$$\boxed{dA = \det \mathbf{F} \|\mathbf{F}^{-T} \mathbf{n}_0\| dA_0} \quad (\text{Nanson's Formula}), \quad (1.31)$$

where $\|\cdot\|$ denotes the Eulerian norm. Thus

$$\boxed{\mathbf{n} = \frac{\text{Cof } \mathbf{F} \mathbf{n}_0}{\|\text{Cof } \mathbf{F} \mathbf{n}_0\|}.} \quad (1.32)$$

1.6 The Polar Decomposition Theorem

Theorem 1.A (Polar Decomposition) *A real invertible matrix F can be factored in a unique way as*

$$F = RU = VR, \quad (1.33)$$

where R is an orthogonal matrix and U and V are symmetric positive definite matrices.

Proof (We will use as a fact the following lemma: For every symmetric positive definite matrix A , there exists a unique symmetric positive definite matrix B such that $B^2 = A$.) Let us first show the existence of the matrices U and V . Define U by

$$U^2 = F^T F = C$$

(which is possible by virtue of the lemma stated above). Then let

$$R = FU^{-1}.$$

Then

$$R^T R = U^{-1} F^T F U^{-1} = U^{-1} C U^{-1} = I.$$

Thus R is a rotation. We have thus shown that there exists a U such that $F = RU$.

Next, define

$$V = RUR^T.$$

Then

$$VR = RUR^T R = RU = F,$$

as asserted.

To show that U and V are unique, let $F = RU$, R being the rotation matrix. Then $F^T F = UR^T R U = U^2$, which means U is unique by the lemma stated. Since $R = FU^{-1}$, R is also uniquely defined. Finally, if $F = VR$, then $FF^T = B = V^2$, so by the same lemma, V is unique. \square

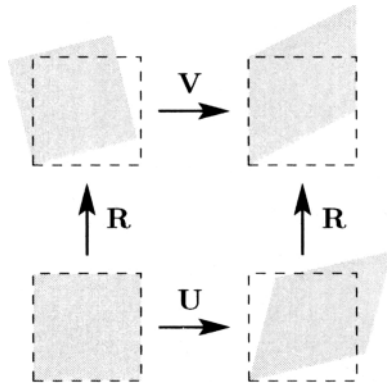


Figure 1.7: The Polar Decomposition Theorem: $\mathbf{F} = \mathbf{R}\mathbf{U} = \mathbf{V}\mathbf{R}$.

Summing up, if $\mathbf{C} = \mathbf{F}^T\mathbf{F}$ and $\mathbf{B} = \mathbf{F}\mathbf{F}^T$ are the right and left Cauchy–Green deformation tensors and

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

then

$$\begin{aligned} \mathbf{C} &= \mathbf{U}^T\mathbf{R}^T\mathbf{R}\mathbf{U} = \mathbf{U}^T\mathbf{U} = \mathbf{U}^2, \\ \mathbf{B} &= \mathbf{V}\mathbf{R}\mathbf{R}^T\mathbf{V} = \mathbf{V}\mathbf{V}^T = \mathbf{V}^2, \end{aligned} \tag{1.34}$$

where \mathbf{U} and \mathbf{V} are the right and left stretch tensors, respectively.

Clearly, the Polar Decomposition Theorem establishes that the deformation gradient \mathbf{F} can be obtained (or can be viewed) as the result of a distortion followed by a rotation or vice versa (see Fig. 1.7).

1.7 Principal Directions and Invariants of Deformation and Strain

For a given deformation tensor field $\mathbf{C}(\mathbf{X})$ and strain field $\mathbf{E}(\mathbf{X})$ (at point \mathbf{X}), recall that $d\mathbf{X}^T\mathbf{C}d\mathbf{X} = 2d\mathbf{X}^T\mathbf{E}d\mathbf{X} - d\mathbf{X}^T d\mathbf{X}$ is the square dS^2 of a material line segment in the current configuration. Suppose the

material line in question is oriented in the direction of a unit vector \mathbf{m} in the reference configuration so that $d\mathbf{X} = \mathbf{m} dS_0$. Then a measure of the stretch or compression of a unit material element originally oriented along a unit vector \mathbf{m} is given by

$$\begin{aligned} \Delta(\mathbf{m}) &= dS^2/dS_0^2 = \mathbf{m}^T \mathbf{C} \mathbf{m}, \\ \mathbf{m} \cdot \mathbf{m} &= \mathbf{m}^T \mathbf{m} = 1. \end{aligned} \quad (1.35)$$

One may ask: Of all possible directions \mathbf{m} at \mathbf{X} , which choice results in the largest (or smallest) value of $\Delta(\mathbf{m})$?

This is a constrained maximization/minimization problem: Find $\mathbf{m} = \mathbf{m}_{\max}$ (or \mathbf{m}_{\min}) that makes $\Delta(\mathbf{m})$ as large (or small) as possible, subject to the constraint $\mathbf{m}^T \mathbf{m} = 1$. To resolve this problem, we use the method of Lagrange multipliers. Denote by $L(\mathbf{m}, \lambda) = \Delta(\mathbf{m}) - \lambda(\mathbf{m}^T \mathbf{m} - 1)$, λ being the Lagrange multiplier. The maxima (on minimize and maximize points) of L satisfy,

$$\frac{\partial L(\mathbf{m}, \lambda)}{\partial \mathbf{m}} = 0 = 2(\mathbf{C} \mathbf{m} - \lambda \mathbf{m}).$$

Thus, unit vectors \mathbf{m} that maximize or minimize $\Delta(\mathbf{m})$ are associated with multipliers λ and satisfy

$$\mathbf{C} \mathbf{m} = \lambda \mathbf{m}, \quad \mathbf{m}^T \mathbf{m} = 1. \quad (1.36)$$

That is, (\mathbf{m}, λ) are eigenvector/eigenvalue pairs of the deformation tensor \mathbf{C} , and \mathbf{m} is normalized so that $\mathbf{m}^T \mathbf{m} = 1$ (or $\|\mathbf{m}\| = 1$).

The following fundamental properties of the above eigenvalue problem can be listed.

1. There are three real eigenvalues and three eigenvectors of \mathbf{C} (at \mathbf{X}); we adopt the ordering $\lambda_1 \geq \lambda_2 \geq \lambda_3$.
2. For $\lambda_i \neq \lambda_j$, the corresponding eigenvectors are orthogonal (for pairs $(\mathbf{m}_i, \lambda_i)$ and $(\mathbf{m}_j, \lambda_j)$), $\mathbf{m}_i^T \mathbf{m}_j = \delta_{ij}$, as can be seen as follows:

$$\mathbf{m}_i^T (\lambda_j \mathbf{m}_j) = \mathbf{m}_i^T \mathbf{C} \mathbf{m}_j = \mathbf{m}_j^T \mathbf{C} \mathbf{m}_i = \mathbf{m}_j^T (\lambda_i \mathbf{m}_i),$$

so

$$(\lambda_i - \lambda_j) \mathbf{m}_i^T \mathbf{m}_j = 0,$$

so

$$\text{if } \lambda_i \neq \lambda_j, \quad \mathbf{m}_i^T \mathbf{m}_j = \delta_{ij}, \quad 1 \leq i, j \leq 3$$

(if $\lambda_i = \lambda_j$, we can always construct \mathbf{m}_j so that it is orthogonal to \mathbf{m}_i).

3. Equation (1.36) can be written as

$$(\mathbf{C} - \lambda \mathbf{I}) \mathbf{m} = \mathbf{0}. \quad (1.37)$$

This equation can have nontrivial solutions only if the determinant of $\mathbf{C} - \lambda \mathbf{I}$ is zero. This is precisely the *characteristic polynomial* of \mathbf{C} :

$$\det(\mathbf{C} - \lambda \mathbf{I}) = -\lambda^3 + I(\mathbf{C})\lambda^2 - \mathbb{I}(\mathbf{C})\lambda + \mathbb{III}(\mathbf{C}), \quad (1.38)$$

where $I, \mathbb{I}, \mathbb{III}$ are the *principal invariants* of \mathbf{C} :

$$\begin{aligned} I(\mathbf{C}) &= \text{trace } \mathbf{C} \equiv \text{tr } \mathbf{C} = C_{ii} = C_{11} + C_{22} + C_{33}, \\ \mathbb{I}(\mathbf{C}) &= \frac{1}{2}(\text{tr } \mathbf{C})^2 - \frac{1}{2} \text{tr } \mathbf{C}^2 = \text{tr } \text{Cof } \mathbf{C}, \\ \mathbb{III}(\mathbf{C}) &= \det \mathbf{C} = \frac{1}{6} \left((\text{tr } \mathbf{C})^3 - 3 \text{tr } \mathbf{C} \text{tr } \mathbf{C}^2 + 2 \text{tr } \mathbf{C}^3 \right). \end{aligned} \quad (1.39)$$

(An invariant of a real matrix \mathbf{C} is any real-valued function $\mu(\mathbf{C})$ with the property $\mu(\mathbf{C}) = \mu(\mathbf{A}^{-1} \mathbf{C} \mathbf{A})$ for all invertible matrices \mathbf{A} .)

4. Because the eigenvectors are all positive, it is customary to write λ_i^2 for the eigenvalues instead of λ_i . Then $(\mathbf{C} - \lambda_i^2 \mathbf{I}) \mathbf{m}_i = \mathbf{0}$. Let \mathbf{N} be the matrix with the mutually orthogonal eigenvectors as rows. Then

$$\mathbf{N}^T \mathbf{C} \mathbf{N} = \begin{bmatrix} \lambda_1^2 & 0 & 0 \\ 0 & \lambda_2^2 & 0 \\ 0 & 0 & \lambda_3^2 \end{bmatrix} = \text{diag}\{\lambda_i^2, i = 1, 2, 3\}. \quad (1.40)$$

The coordinate system defined by the mutually orthogonal triad of eigenvectors define the principal directions and values of \mathbf{C} at \mathbf{X} . For this choice of a basis, we obtain

$$\mathbf{C} = \sum_{i=1}^3 \lambda_i^2 \mathbf{m}_i \otimes \mathbf{m}_i. \quad (1.41)$$

If $\lambda_1^2 \geq \lambda_2^2 \geq \lambda_3^2$, λ_1^2 corresponds to the maximum, λ_3^2 to the minimum, and λ_2^2 to a “mini-max” principal value of \mathbf{C} (or of $\Delta(\mathbf{n})$).

Notice that the stretch along, say, \mathbf{m}_1 is $(\mathbf{m}_1^T \mathbf{C} \mathbf{m}_1)^{1/2} = \lambda_1$, etc. Also,

$$\mathbf{C} = \mathbf{U}^2 = \begin{bmatrix} \lambda_1^2 & 0 & 0 \\ 0 & \lambda_2^2 & 0 \\ 0 & 0 & \lambda_3^2 \end{bmatrix}. \quad (1.42)$$

The principal invariants are thus

$$\begin{aligned} I(\mathbf{C}) &= \lambda_1^2 + \lambda_2^2 + \lambda_3^2, \\ II(\mathbf{C}) &= \lambda_1^2 \lambda_2^2 + \lambda_1^2 \lambda_3^2 + \lambda_2^2 \lambda_3^2, \\ III(\mathbf{C}) &= \lambda_1^2 \lambda_2^2 \lambda_3^2. \end{aligned} \quad (1.43)$$

1.8 The Reynolds' Transport Theorem

We frequently encounter the need to evaluate the total time rate of change of a field, either densities or measures of concentrations per unit volume, defined over a volume $\omega \subset \Omega_t$. For instance, if $\Psi = \Psi(\mathbf{x}, t)$ is a spatial field, either scalar- or vector-valued, suppose we wish to compute $d(\int_{\omega} \Psi dx)/dt$. The following change of integration variables facilitates such a calculation. Let ω_0 be the region of Ω_0 occupied by the material,

while in the reference configuration, that occupies ω in Ω_t . Then,

$$\begin{aligned}
 \frac{d}{dt} \int_{\omega} \Psi \, dx &= \frac{d}{dt} \int_{\omega_0} \Psi_m \det \mathbf{F} \, dX \\
 &= \int_{\omega_0} \frac{d}{dt} (\Psi_m \det \mathbf{F}) \, dX \\
 &= \int_{\omega_0} \left(\frac{\partial \Psi_m}{\partial t} + \mathbf{v} \cdot \text{grad } \Psi_m \right) \det \mathbf{F} \, dX \\
 &\quad + \int_{\omega_0} \Psi_m \overline{\det \mathbf{F}} \det \mathbf{F} \, dX \\
 &= \int_{\omega} \frac{\partial \Psi}{\partial t} \, dx + \int_{\omega} \text{div}(\Psi \mathbf{v}) \, dx.
 \end{aligned}$$

Thus,

$$\boxed{\frac{d}{dt} \int_{\omega} \Psi \, dx = \int_{\omega} \frac{\partial \Psi}{\partial t} \, dx + \int_{\partial \omega} \Psi \mathbf{v} \cdot \mathbf{n} \, dA.} \quad (1.44)$$

This last result is known as the *Reynolds' Transport Theorem*.