

INTRODUCTION TO CONTINUUM MECHANICS

I-SHIH LIU

Instituto de Matemática
Universidade Federal do Rio de Janeiro

2018

Contents

1	Notations and tensor algebra	1
1.1	Vector space, inner product	1
1.2	Linear transformation	2
1.3	Differentiation, gradient	5
1.4	Divergence	8
1.5	Remarks on functions and operators in other coordinate systems	8
2	Kinematics of finite deformation	11
2.1	Configuration and deformation	11
2.2	Strain and rotation	13
2.3	Linear strain tensors	14
2.4	Motions	16
2.5	Relative deformation	19
3	Balance laws	21
3.1	General balance equation	21
3.2	Local balance equation	22
3.3	Balance equations in reference coordinates	22
3.4	Conservation of mass	23
3.5	Equation of motion	23
3.6	Conservation of energy	25
3.7	Basic equations in material coordinates	26
3.8	Boundary value problem	27
4	Euclidean objectivity	29
4.1	Frame of reference, observer	29
4.2	Objective tensors	32
4.3	Transformation properties of motion	33
4.4	Inertial frames	35
4.5	Galilean invariance of balance laws	36
5	Principle of material frame-indifference	39
5.1	Constitutive equations in material description	39
5.2	Principle of material frame-indifference	41
5.3	Constitutive equations in referential description	41
5.4	Simple materials	43
6	Material symmetry	47
6.1	Material symmetry group	47
6.2	Classification of material bodies	48
6.3	Summary on constitutive models of simple materials	49
6.4	Remark on incompressibility	50

7	Elastic solids	53
7.1	Isotropic elastic solid	53
7.2	Representations of isotropic functions	53
7.3	Incompressible isotropic elastic solids	55
7.4	Elastic solid materials	56
7.5	Hooke's law	57
8	Viscoelastic materials	59
8.1	Isotropic viscoelastic solids	59
8.2	Viscoelastic solids	60
8.3	Viscous fluids	62
8.4	Navier-Stokes fluids	62
8.5	Viscous heat-conducting fluids	63
9	Second law of thermodynamics	65
9.1	Entropy principle	66
9.2	Thermodynamics of heat-conducting elastic fluids	67
9.3	Exploitation of entropy principle	68
9.4	Thermodynamic stability	72
10	Some problems in finite elasticity	77
10.1	Universal solutions in elasticity	77
10.2	Simple Shear	79
10.3	Pure Shear	80
10.4	bending of a rectangular block	83
10.5	Deformation of a cylindrical annulus	85
10.6	Appendix: Divergence of a tensor field	87
11	Wave propagation in elastic bodies	89
11.1	Small deformations on a deformed body	89
11.2	The equation of motion in relative description	93
11.3	Plane harmonic waves in a deformed elastic body	95
11.4	Mooney–Rivlin elastic materials	97
11.5	Principal acceleration waves of finite amplitude	99
12	Mixture theory of porous media	103
12.1	Theories of mixtures	103
12.2	Mixture of elastic materials	110
12.3	Saturated porous media	113
12.4	Equations of motion	115
12.5	Linear theory	116
12.6	Problems in poroelasticity	118
12.7	Boundary conditions	120

1 Notations and tensor algebra

The reader is assumed to have a reasonable knowledge of the basic notions of vector spaces and calculus on Euclidean spaces.

1.1 Vector space, inner product

Let V be a finite dimensional vector space, $\dim V = n$, and $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ be a basis of V . Then for any vector $\mathbf{v} \in V$, it can be represented as

$$\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \dots + v_n\mathbf{e}_n,$$

where (v_1, v_2, \dots, v_n) are called the *components* of the vector \mathbf{v} relative to the basis $\{\mathbf{e}_i\}$.

An *inner product* (or scalar product) is defined as a *symmetric, positive definite, bilinear* map such that for $\mathbf{u}, \mathbf{v} \in V$ their inner product, denoted by $\mathbf{u} \cdot \mathbf{v}$, is a scalar.

The *norm* (or length) of the vector \mathbf{v} is defined as

$$|\mathbf{v}| = \sqrt{\mathbf{v} \cdot \mathbf{v}}.$$

We can show that $|\mathbf{u} \cdot \mathbf{v}| \leq |\mathbf{u}||\mathbf{v}|$, so that we can define the *angle* between two vectors as

$$\cos \theta(\mathbf{u}, \mathbf{v}) = \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{u}||\mathbf{v}|}, \quad 0 \leq \theta(\mathbf{u}, \mathbf{v}) \leq \pi,$$

and say that they are *orthogonal* (or perpendicular) if $\mathbf{u} \cdot \mathbf{v} = 0$, so that $\theta(\mathbf{u}, \mathbf{v}) = \pi/2$.

A basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is called *orthonormal* if

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij},$$

where the *Kronecker delta* is defined as

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

For simplicity, we shall assume, from now on, that all bases are orthonormal.

Example. For $\mathbf{u}, \mathbf{v} \in V$,

$$\begin{aligned} \mathbf{v} &= v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \dots + v_n\mathbf{e}_n = \sum_{i=1}^n v_i\mathbf{e}_i = v_i\mathbf{e}_i, \\ \mathbf{u} &= u_1\mathbf{e}_1 + u_2\mathbf{e}_2 + \dots + u_n\mathbf{e}_n = \sum_{j=1}^n u_j\mathbf{e}_j = u_j\mathbf{e}_j, \end{aligned}$$

we have the inner product

$$\mathbf{u} \cdot \mathbf{v} = \left(\sum_{i=1}^n u_i \mathbf{e}_i \right) \cdot \left(\sum_{j=1}^n v_j \mathbf{e}_j \right) = \sum_{i=1}^n \sum_{j=1}^n u_i v_j (\mathbf{e}_i \cdot \mathbf{e}_j) = \sum_{i=1}^n \sum_{j=1}^n u_i v_j \delta_{ij} = \sum_{i=1}^n u_i v_i,$$

or

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n = u_i v_i.$$

In these expressions, we can neglect the summation signs for simplicity. This is called the *summation convention* (due to Einstein), for which every pair of repeated index is summed over its range as understood from the context.

Taking the inner product with the vector, we obtain the component,

$$\mathbf{e}_i \cdot \mathbf{v} = \mathbf{e}_i \cdot (v_j \mathbf{e}_j) = v_j (\mathbf{e}_i \cdot \mathbf{e}_j) = v_j \delta_{ij} = v_i, \quad v_i = \mathbf{e}_i \cdot \mathbf{v}.$$

□

1.2 Linear transformation

Let V be a finite dimensional vector space with an inner product. We call $A : V \rightarrow V$ a *linear transformation* if for any vectors $\mathbf{u}, \mathbf{v} \in V$ and any scalar $a \in \mathbb{R}$,

$$A(a\mathbf{u} + \mathbf{v}) = aA(\mathbf{u}) + A(\mathbf{v}).$$

Let $\mathcal{L}(V)$ be the space of linear transformations on V . The elements of $\mathcal{L}(V)$ are also called (second order) tensors.

For $\mathbf{u}, \mathbf{v} \in V$ we can define their *tensor product*, denoted by $\mathbf{u} \otimes \mathbf{v} \in \mathcal{L}(V)$, defined as a tensor so that for any $\mathbf{w} \in V$,

$$(\mathbf{u} \otimes \mathbf{v})\mathbf{w} = (\mathbf{v} \cdot \mathbf{w}) \mathbf{u}.$$

Let $\{\mathbf{e}_i, i = 1, \dots, n\}$ be a basis of V , then $\{\mathbf{e}_i \otimes \mathbf{e}_j, i, j = 1, \dots, n\}$ is a basis for $\mathcal{L}(V)$, and for any $A \in \mathcal{L}(V)$, the component form can be expressed as

$$A = \sum_{i=1}^n \sum_{j=1}^n A_{ij} \mathbf{e}_i \otimes \mathbf{e}_j = A_{ij} \mathbf{e}_i \otimes \mathbf{e}_j, \quad A_{ij} = \mathbf{e}_i \cdot A \mathbf{e}_j.$$

Here, we have used the summation convention for the two pairs of repeated indices i and j . The components A_{ij} can be represented as the i -th row and j -th column of a matrix.

For $\mathbf{v} = v_i \mathbf{e}_i$, then $A\mathbf{v}$ is a vector, and

$$\begin{aligned} A\mathbf{v} &= (A\mathbf{v})_i \mathbf{e}_i = (A_{ij} \mathbf{e}_i \otimes \mathbf{e}_j)(v_k \mathbf{e}_k) = A_{ij} v_k (\mathbf{e}_i \otimes \mathbf{e}_j) \mathbf{e}_k \\ &= A_{ij} v_k (\mathbf{e}_j \cdot \mathbf{e}_k) \mathbf{e}_i = A_{ij} v_k \delta_{jk} \mathbf{e}_i = A_{ik} v_k \mathbf{e}_i, \end{aligned}$$

which can be written in components or in matrix notation,

$$(A\mathbf{v})_i = A_{ik}v_k, \quad [A\mathbf{v}] = [A][\mathbf{v}],$$

where $[\mathbf{v}]$ is regarded as a column vector. Similarly, for any $A, B \in \mathcal{L}(V)$, the product $AB \in \mathcal{L}(V)$ can be represented as a matrix product,

$$(AB)_{ij} = A_{ik}B_{kj}, \quad [AB] = [A][B].$$

Example. Let $V = \mathbb{R}^2$, and let $\{\mathbf{e}_x = (1, 0), \mathbf{e}_y = (0, 1)\}$ be the standard basis. Any $\mathbf{v} = (x, y) = x\mathbf{e}_x + y\mathbf{e}_y \in \mathbb{R}^2$ can be represented as a column vector,

$$[\mathbf{v}] = \begin{bmatrix} x \\ y \end{bmatrix}.$$

If $\mathbf{u} = (u_1, u_2)$ and $\mathbf{w} = (w_1, w_2)$, the component of their tensor product is

$$(\mathbf{u} \otimes \mathbf{v})_{ij} = \mathbf{e}_i \cdot (\mathbf{u} \otimes \mathbf{v})\mathbf{e}_j = (\mathbf{e}_i \cdot \mathbf{u})(\mathbf{v} \cdot \mathbf{e}_j) = (\mathbf{e}_i \cdot \mathbf{u})(\mathbf{v} \cdot \mathbf{e}_j) = u_i v_j,$$

and can be represented by

$$[\mathbf{u} \otimes \mathbf{w}] = \begin{bmatrix} u_1 w_1 & u_1 w_2 \\ u_2 w_1 & u_2 w_2 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \begin{bmatrix} w_1 & w_2 \end{bmatrix}.$$

Therefore, the standard basis for $\mathcal{L}(\mathbb{R}^2)$ are given by

$$\begin{aligned} [\mathbf{e}_x \otimes \mathbf{e}_x] &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, & [\mathbf{e}_x \otimes \mathbf{e}_y] &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \\ [\mathbf{e}_y \otimes \mathbf{e}_x] &= \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, & [\mathbf{e}_y \otimes \mathbf{e}_y] &= \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \end{aligned}$$

Given a linear transformation $A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by

$$A(x, y) = (2x - 3y, x + 5y), \quad [A] = \begin{bmatrix} 2 & -3 \\ 1 & 5 \end{bmatrix}.$$

It can be written as

$$\begin{bmatrix} 2 & -3 \\ 1 & 5 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x - 3y \\ x + 5y \end{bmatrix}.$$

□

The transpose of $A \in \mathcal{L}(V)$ is defined for any $\mathbf{u}, \mathbf{v} \in V$, such that

$$A^T \mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot A\mathbf{v}.$$

In components, $(A^T)_{ij} = A_{ji}$.

$Q \in \mathcal{L}(V)$ is an *orthogonal* transformation, if it preserves the inner product,

$$Q\mathbf{u} \cdot Q\mathbf{v} = \mathbf{u} \cdot \mathbf{v}.$$

Therefore, an orthogonal transformation preserves both the angle and the norm of vectors. From the definition, it follows that

$$Q^T Q = I, \quad \text{or} \quad Q^{-1} = Q^T,$$

where I is the identity tensor and Q^{-1} is the inverse of Q .

The *trace* of a linear transformation can be defined as a linear scalar function, $\text{tr} : \mathcal{L}(V) \rightarrow \mathbb{R}$, such that for any $\mathbf{u}, \mathbf{v} \in V$,

$$\text{tr} \mathbf{u} \otimes \mathbf{v} = \mathbf{u} \cdot \mathbf{v}.$$

Therefore, for any $A \in \mathcal{L}(V)$,

$$\text{tr} A = \text{tr}(A_{ij} \mathbf{e}_i \otimes \mathbf{e}_j) = A_{ij} \text{tr}(\mathbf{e}_i \otimes \mathbf{e}_j) = A_{ij} \delta_{ij} = A_{jj},$$

which equals the sum of the diagonal elements of the matrix in Cartesian components. One can easily verify that

$$\text{tr} A^T = \text{tr} A, \quad \text{tr} AB = \text{tr} BA.$$

We can define the inner product of two tensors A and B by

$$A : B = \text{tr} AB^T = A_{ij} B_{ij},$$

and the norm $|A|$ can be defined as

$$|A|^2 = A : A = A_{ij} A_{ij},$$

which is the sum of square of all the elements of A by the summation convention.

We are particularly interest in the three-dimensional space, which is the physical space of classical mechanics. Let $\dim V = 3$, we can define the *vector product* of two vectors, $\mathbf{u} \times \mathbf{v} \in V$, in components,

$$(\mathbf{u} \times \mathbf{v})_i = \varepsilon_{ijk} u_j v_k,$$

where ε_{ijk} is the permutation symbol,

$$\varepsilon_{ijk} = \begin{cases} 1, & \text{if } \{i, j, k\} \text{ is an even permutation of } \{1, 2, 3\}, \\ -1, & \text{if } \{i, j, k\} \text{ is an odd permutation of } \{1, 2, 3\}, \\ 0, & \text{if otherwise.} \end{cases}$$

One can easily check the following identity:

$$\varepsilon_{ijk} \varepsilon_{imn} = \delta_{jm} \delta_{kn} - \delta_{jn} \delta_{km}.$$

We can easily show that $|\mathbf{u} \times \mathbf{v}| = |\mathbf{u}| |\mathbf{v}| |\sin \theta(\mathbf{u}, \mathbf{v})|$, which is geometrically the area of the parallelogram formed by the two vectors.

We can also define the *triple product* $\mathbf{u} \cdot \mathbf{v} \times \mathbf{w}$ the *triple product*, which is the volume of the parallelepiped formed by the three vectors. If they are linearly independent then the triple product is different from zero.

For a linear transformation, we can define the *determinant* as the ratio between the deformed volume and the original one for any three linearly independent vectors,

$$\det A = \frac{A\mathbf{u} \cdot A\mathbf{v} \times A\mathbf{w}}{\mathbf{u} \cdot \mathbf{v} \times \mathbf{w}}.$$

We have

$$\det(AB) = \frac{AB\mathbf{u} \cdot AB\mathbf{v} \times AB\mathbf{w}}{\mathbf{u} \cdot \mathbf{v} \times \mathbf{w}} = \frac{A(B\mathbf{u}) \cdot A(B\mathbf{v}) \times A(B\mathbf{w})}{B\mathbf{u} \cdot B\mathbf{v} \times B\mathbf{w}} \cdot \frac{B\mathbf{u} \cdot B\mathbf{v} \times B\mathbf{w}}{\mathbf{u} \cdot \mathbf{v} \times \mathbf{w}},$$

which implies that

$$\det(AB) = (\det A)(\det B).$$

1.3 Differentiation, gradient

Let \mathbb{E} be a three-dimensional Euclidean space and the vector space V be its *translation space*. For any two points $\mathbf{x}, \mathbf{y} \in \mathbb{E}$ there is a unique vector $\mathbf{v} \in V$ associated with their difference,

$$\mathbf{v} = \mathbf{y} - \mathbf{x}, \quad \text{or} \quad \mathbf{y} = \mathbf{x} + \mathbf{v}.$$

We may think of \mathbf{v} as the geometric vector that starts at the point \mathbf{x} and ends at the point \mathbf{y} . The *distance* between \mathbf{x} and \mathbf{y} is then given by

$$d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}| = |\mathbf{v}|.$$

Let \mathcal{D} be an open region in \mathbb{E} and W be any vector space or an Euclidean space. A function $f : \mathcal{D} \rightarrow W$ is said to be *differentiable* at $\mathbf{x} \in \mathcal{D}$ if there exists a linear transformation $\nabla f(\mathbf{x}) : V \rightarrow W$, such that for any $\mathbf{v} \in V$,

$$f(\mathbf{x} + \mathbf{v}) - f(\mathbf{x}) = \nabla f(\mathbf{x})[\mathbf{v}] + o(2),$$

where $o(2)$ denotes the second and higher order terms in $|\mathbf{v}|$. We call ∇f the *gradient* of f with respect to \mathbf{x} , and will also denote it by $\nabla_{\mathbf{x}} f$, or more frequently by $\text{grad } f$. The above definition of gradient can also be written as

$$\nabla f(\mathbf{x})[\mathbf{v}] = \left. \frac{d}{dt} f(\mathbf{x} + t\mathbf{v}) \right|_{t=0}.$$

If $f(\mathbf{x}) \in \mathbb{R}$ is a scalar field for $\mathbf{x} \in \mathcal{D}$, then $\nabla f(\mathbf{x}) \in V$ is a vector field, and if $\mathbf{h}(\mathbf{x}) \in V$ is a vector field, then $\nabla \mathbf{h}(\mathbf{x}) \in \mathcal{L}(V)$ is a tensor field. The above notation has the following meaning:

$$\nabla f(\mathbf{x})[\mathbf{v}] = \nabla f(\mathbf{x}) \cdot \mathbf{v}, \quad \nabla \mathbf{h}(\mathbf{x})[\mathbf{v}] = \nabla \mathbf{h}(\mathbf{x})\mathbf{v}.$$

For functions defined on tensor space, $\mathcal{F} : W_1 \rightarrow W_2$, where W_1, W_2 are some tensor spaces, the differentiation can similarly be defined.

Example. Let $\mathbb{E} = \mathbb{R}^2$, and $f(x, y)$ be a scalar field in Cartesian coordinate system, then

$$\mathbf{x} = (x, y), \quad \mathbf{v} = (v_x, v_y),$$

and

$$\begin{aligned} \nabla f(\mathbf{x}) \cdot \mathbf{v} &= \left. \frac{d}{dt} f(x + tv_x, y + tv_y) \right|_{t=0} \\ &= \frac{\partial f}{\partial x} v_x + \frac{\partial f}{\partial y} v_y = \frac{\partial f}{\partial x} (\mathbf{e}_x \cdot \mathbf{v}) + \frac{\partial f}{\partial y} (\mathbf{e}_y \cdot \mathbf{v}) = \frac{\partial f}{\partial x_j} (\mathbf{e}_j \cdot \mathbf{v}), \end{aligned}$$

which gives

$$\nabla f(x, y) = \frac{\partial f}{\partial x} \mathbf{e}_x + \frac{\partial f}{\partial y} \mathbf{e}_y = \frac{\partial f}{\partial x_i} \mathbf{e}_i.$$

Let

$$\mathbf{h}(x, y) = h_x(x, y) \mathbf{e}_x + h_y(x, y) \mathbf{e}_y = h_i(x, y) \mathbf{e}_i$$

be a vector field, then by the product rule, for any vector \mathbf{v} ,

$$\nabla \mathbf{h}(x, y)[\mathbf{v}] = (\nabla h_i(x, y)[\mathbf{v}]) \mathbf{e}_i + h_i(x, y) (\nabla \mathbf{e}_i[\mathbf{v}]).$$

Since the standard basis is a constant vector field, $\nabla \mathbf{e}_i = 0$, therefore, we have

$$\nabla \mathbf{h}(x, y)[\mathbf{v}] = (\nabla h_i(x, y)[\mathbf{v}]) \mathbf{e}_i = \frac{\partial h_i}{\partial x_j} (\mathbf{e}_j \cdot \mathbf{v}) \mathbf{e}_i = \frac{\partial h_i}{\partial x_j} (\mathbf{e}_i \otimes \mathbf{e}_j) \mathbf{v}.$$

Therefore, we obtain

$$\nabla \mathbf{h}(x, y) = \frac{\partial h_x}{\partial x} \mathbf{e}_x \otimes \mathbf{e}_x + \frac{\partial h_x}{\partial y} \mathbf{e}_x \otimes \mathbf{e}_y + \frac{\partial h_y}{\partial x} \mathbf{e}_y \otimes \mathbf{e}_x + \frac{\partial h_y}{\partial y} \mathbf{e}_y \otimes \mathbf{e}_y = \frac{\partial h_i}{\partial x_j} \mathbf{e}_i \otimes \mathbf{e}_j.$$

In matrix notations,

$$[\nabla f(x, y)] = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}, \quad [\nabla \mathbf{h}(x, y)] = \begin{bmatrix} \frac{\partial h_x}{\partial x} & \frac{\partial h_x}{\partial y} \\ \frac{\partial h_y}{\partial x} & \frac{\partial h_y}{\partial y} \end{bmatrix}.$$

In components,

$$(\nabla f)_i = \frac{\partial f}{\partial x_i} = f_{,i}, \quad (\nabla \mathbf{h})_{ij} = \frac{\partial h_i}{\partial x_j} = h_{i,j},$$

where $i = 1, 2$ refers to coordinate x and y respectively and we have used comma to indicate partial differentiation. \square

Example. Let $\mathcal{F}(A, \mathbf{v}) = \mathbf{v} \cdot A^2 \mathbf{v}$ be a scalar function of a tensor and a vector variables, we have for any $\mathbf{w} \in V$,

$$\begin{aligned} \nabla_{\mathbf{v}} \mathcal{F}(A, \mathbf{v}) \cdot \mathbf{w} &= \left. \frac{d}{dt} (\mathbf{v} + t\mathbf{w}) \cdot A^2 (\mathbf{v} + t\mathbf{w}) \right|_{t=0} \\ &= \mathbf{w} \cdot A^2 \mathbf{v} + \mathbf{v} \cdot A^2 \mathbf{w} = (A^2 \mathbf{v} + (A^2)^T \mathbf{v}) \cdot \mathbf{w}, \end{aligned}$$

and for any $W \in \mathcal{L}(V)$,

$$\begin{aligned} \mathcal{F}(A + W, \mathbf{v}) - \mathcal{F}(A, \mathbf{v}) &= \mathbf{v} \cdot (A + W)(A + W)\mathbf{v} - \mathbf{v} \cdot A^2 \mathbf{v} \\ &= \mathbf{v} \cdot (AW + WA + W^2)\mathbf{v} = \mathbf{v} \cdot (AW + WA)\mathbf{v} + o(2) \\ &= \nabla_A \mathcal{F}(A, \mathbf{v}) : W + o(2), \end{aligned}$$

or

$$\begin{aligned} \nabla_A \mathcal{F}(A, \mathbf{v}) : W &= \left. \frac{d}{dt} (\mathbf{v} \cdot (A + tW)(A + tW)\mathbf{v}) \right|_{t=0} \\ &= \mathbf{v} \cdot (WA)\mathbf{v} + \mathbf{v} \cdot (AW)\mathbf{v} = (\mathbf{v} \otimes A\mathbf{v} + A^T \mathbf{v} \otimes \mathbf{v}) : W, \end{aligned}$$

since

$$\begin{aligned} \mathbf{v} \cdot (WA)\mathbf{v} &= W^T \mathbf{v} \cdot A\mathbf{v} = \text{tr}(W^T \mathbf{v} \otimes A\mathbf{v}) = (\mathbf{v} \otimes A\mathbf{v}) : W, \\ \mathbf{v} \cdot (AW)\mathbf{v} &= A^T \mathbf{v} \cdot W\mathbf{v} = \text{tr}(W \mathbf{v} \otimes A^T \mathbf{v}) = (A^T \mathbf{v} \otimes \mathbf{v}) : W. \end{aligned}$$

So we obtain

$$\begin{aligned} \nabla_{\mathbf{v}} \mathcal{F}(A, \mathbf{v}) &= A^2 \mathbf{v} + (A^2)^T \mathbf{v}, \\ \nabla_A \mathcal{F}(A, \mathbf{v}) &= \mathbf{v} \otimes A\mathbf{v} + A^T \mathbf{v} \otimes \mathbf{v}. \end{aligned}$$

In components,

$$\begin{aligned} (\nabla_{\mathbf{v}} \mathcal{F})_i &= A_{ik} A_{kl} v_l + A_{lk} A_{ki} v_l, \\ (\nabla_A \mathcal{F})_{ij} &= v_i A_{jk} v_k + A_{ki} v_k v_j. \end{aligned}$$

The above differentiations can also be carried out entirely in index notations. Since components are scalar quantities, the usual product rule can easily be applied. For example

$$\begin{aligned} (\nabla_A \mathcal{F})_{ij} &= \frac{\partial (A_{mk} A_{kn} v_m v_n)}{\partial A_{ij}} \\ &= \delta_{mi} \delta_{kj} A_{kn} v_m v_n + A_{mk} \delta_{ki} \delta_{nj} v_m v_n \\ &= A_{jn} v_i v_n + A_{mi} v_m v_j. \end{aligned}$$

In fact, doing tensor calculus entirely in index notation is the simplest way if one is accustomed to the summation convention. The results can easily be converted into the direct notation or matrix notation. \square

1.4 Divergence

For a vector field $\mathbf{v}(\mathbf{x}) \in V$, $\mathbf{x} \in \mathbb{E}$, the gradient $\nabla \mathbf{v}(\mathbf{x}) \in \mathcal{L}(V)$ is a tensor field, then the *divergence* of a vector field is defined as a scalar field by

$$\operatorname{div} \mathbf{v}(\mathbf{x}) = \operatorname{tr}(\nabla \mathbf{v}(\mathbf{x})) \in \mathbb{R}.$$

In components,

$$\operatorname{div} \mathbf{v} = \operatorname{tr}(\nabla \mathbf{v}) = \operatorname{tr}\left(\frac{\partial v_i}{\partial x_j} \mathbf{e}_i \otimes \mathbf{e}_j\right) = \frac{\partial v_i}{\partial x_j} \delta_{ij} = v_{j,j}.$$

Similarly, we can define the *divergence* of a tensor field $A(\mathbf{x}) \in \mathcal{L}(V)$ as a vector field in terms of its components by

$$\operatorname{div} A = (\operatorname{div} A)_i \mathbf{e}_i = \frac{\partial A_{ij}}{\partial x_j} \mathbf{e}_i = A_{i,j,j} \mathbf{e}_i.$$

Example. For $\mathbb{E} = \mathbb{R}^2$, and $\mathbf{v}(x, y) = v_x(x, y) \mathbf{e}_x + v_y(x, y) \mathbf{e}_y$, we have

$$\operatorname{div} \mathbf{v} = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y}.$$

For a tensor field $A(x, y) = A_{ij}(x, y) \mathbf{e}_i \otimes \mathbf{e}_j$, we have

$$[\operatorname{div} A] = \begin{bmatrix} \frac{\partial A_{xx}}{\partial x} + \frac{\partial A_{xy}}{\partial y} \\ \frac{\partial A_{yx}}{\partial x} + \frac{\partial A_{yy}}{\partial y} \end{bmatrix}.$$

□

1.5 Remarks on functions and operators in other coordinate systems

Functions, scalar-, vector- or tensor-valued, are most commonly expressed in Cartesian coordinate system and standard basis in their simplest forms. Of course, they can also be expressed relative to any other coordinate system and its related basis. This is usually done in the theory of tensor analysis in general with the introduction of Christoffel symbols and covariant derivatives.

In fact, to express functions and operators in coordinate system other than the Cartesian coordinate system, as we have done so far, is intrinsically simple. It involves only a change of variables and a change of bases.

Example. Let us consider the gradient of a scalar field from Cartesian to polar coordinate system.

Let (x, y) be the Cartesian coordinate system with the standard basis $(\mathbf{e}_x, \mathbf{e}_y)$ and (r, θ) be the polar coordinate system with orthonormal basis $(\mathbf{e}_r, \mathbf{e}_\theta)$. We have the following change of variables $(x, y) \leftrightarrow (r, \theta)$:

$$\begin{aligned} x &= r \cos \theta, & r &= \sqrt{x^2 + y^2}, \\ y &= r \sin \theta, & \theta &= \arctan \frac{y}{x}, \end{aligned} \quad (1.1)$$

and change of bases $(\mathbf{e}_x, \mathbf{e}_y) \leftrightarrow (\mathbf{e}_r, \mathbf{e}_\theta)$:

$$\begin{aligned} \mathbf{e}_x &= \cos \theta \mathbf{e}_r - \sin \theta \mathbf{e}_\theta, & \mathbf{e}_r &= \cos \theta \mathbf{e}_x + \sin \theta \mathbf{e}_y, \\ \mathbf{e}_y &= \sin \theta \mathbf{e}_r + \cos \theta \mathbf{e}_\theta, & \mathbf{e}_\theta &= -\sin \theta \mathbf{e}_x + \cos \theta \mathbf{e}_y. \end{aligned} \quad (1.2)$$

Consider a scalar-valued function,

$$f(x, y) = f(r, \theta)$$

expressed in Cartesian and polar coordinate systems. The gradient of $f(x, y)$ is given by

$$\nabla f = \frac{\partial f}{\partial x} \mathbf{e}_x + \frac{\partial f}{\partial y} \mathbf{e}_y. \quad (1.3)$$

To change it into polar coordinate, we can first change variables, so that

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial r} \frac{\partial r}{\partial x} + \frac{\partial f}{\partial \theta} \frac{\partial \theta}{\partial x}, \quad \frac{\partial f}{\partial y} = \frac{\partial f}{\partial r} \frac{\partial r}{\partial y} + \frac{\partial f}{\partial \theta} \frac{\partial \theta}{\partial y},$$

and from (1.1)₂,

$$\begin{aligned} \frac{\partial r}{\partial x} &= \frac{x}{r}, & \frac{\partial \theta}{\partial x} &= \frac{-y}{r^2}, \\ \frac{\partial r}{\partial y} &= \frac{y}{r}, & \frac{\partial \theta}{\partial y} &= \frac{x}{r^2}. \end{aligned}$$

Putting together into (1.3) and using (1.2)₁ to change basis, it gives

$$\nabla f = \left(\frac{\partial f}{\partial r} \frac{x}{r} - \frac{\partial f}{\partial \theta} \frac{y}{r^2} \right) (\cos \theta \mathbf{e}_r - \sin \theta \mathbf{e}_\theta) + \left(\frac{\partial f}{\partial r} \frac{y}{r} + \frac{\partial f}{\partial \theta} \frac{x}{r^2} \right) (\sin \theta \mathbf{e}_r + \cos \theta \mathbf{e}_\theta),$$

which after simplification with (1.1)₁ leads to the gradient in the polar coordinate,

$$\text{grad } f = \frac{\partial f}{\partial r} \mathbf{e}_r + \frac{1}{r} \frac{\partial f}{\partial \theta} \mathbf{e}_\theta.$$

□

Example. The divergence of a vector field in polar coordinate: Let

$$\mathbf{u} = u_x(x, y)\mathbf{e}_x + u_y(x, y)\mathbf{e}_y = u_r(r, \theta)\mathbf{e}_r + u_\theta(r, \theta)\mathbf{e}_\theta$$

be a vector field on \mathbb{R}^2 , then with the change $(\mathbf{e}_r, \mathbf{e}_\theta) \rightarrow (\mathbf{e}_x, \mathbf{e}_y)$,

$$\begin{aligned}\mathbf{u} &= u_r(\cos \theta \mathbf{e}_x + \sin \theta \mathbf{e}_y) + u_\theta(-\sin \theta \mathbf{e}_x + \cos \theta \mathbf{e}_y) \\ &= (u_r \cos \theta - u_\theta \sin \theta)\mathbf{e}_x + (u_r \sin \theta + u_\theta \cos \theta)\mathbf{e}_y,\end{aligned}$$

so that

$$u_x = u_r \cos \theta - u_\theta \sin \theta, \quad u_y = u_r \sin \theta + u_\theta \cos \theta,$$

from which we have

$$\begin{aligned}\frac{\partial u_x}{\partial x} &= \left(\frac{\partial u_r}{\partial r} \frac{\partial r}{\partial x} + \frac{\partial u_r}{\partial \theta} \frac{\partial \theta}{\partial x} \right) \cos \theta - u_r \sin \theta \frac{\partial \theta}{\partial x} - \left(\frac{\partial u_\theta}{\partial r} \frac{\partial r}{\partial x} + \frac{\partial u_\theta}{\partial \theta} \frac{\partial \theta}{\partial x} \right) \sin \theta - u_\theta \cos \theta \frac{\partial \theta}{\partial x}, \\ \frac{\partial u_y}{\partial y} &= \left(\frac{\partial u_r}{\partial r} \frac{\partial r}{\partial y} + \frac{\partial u_r}{\partial \theta} \frac{\partial \theta}{\partial y} \right) \sin \theta + u_r \cos \theta \frac{\partial \theta}{\partial y} + \left(\frac{\partial u_\theta}{\partial r} \frac{\partial r}{\partial y} + \frac{\partial u_\theta}{\partial \theta} \frac{\partial \theta}{\partial y} \right) \cos \theta - u_\theta \sin \theta \frac{\partial \theta}{\partial y}.\end{aligned}$$

Therefore, it follows after some calculation that

$$\operatorname{div} \mathbf{u} = \frac{\partial u_x}{\partial x} + \frac{\partial u_y}{\partial y} = \frac{\partial u_r}{\partial r} + \frac{1}{r} \frac{\partial u_\theta}{\partial \theta} + \frac{u_r}{r}.$$

□

2 Kinematics of finite deformation

2.1 Configuration and deformation

A body \mathcal{B} can be identified mathematically with a region in a three-dimensional Euclidean space \mathbb{E} . Such an identification is called a configuration of the body. In other words, a one-to-one mapping from \mathcal{B} into \mathbb{E} is called a *configuration* of \mathcal{B} .

It is more convenient to single out a particular configuration of \mathcal{B} , say κ , as a reference,

$$\kappa : \mathcal{B} \rightarrow \mathbb{E}, \quad \kappa(\mathbf{p}) = \mathbf{X}. \quad (2.1)$$

We call κ a *reference configuration* of \mathcal{B} . The coordinates of \mathbf{X} , $(X^\alpha, \alpha = 1, 2, 3)$ are called the *referential coordinates*, or sometimes called the *material coordinates* since the point \mathbf{X} in the reference configuration κ is often identified with the material point \mathbf{p} of the body when κ is given and fixed. The body \mathcal{B} in the configuration κ will be denoted by \mathcal{B}_κ .

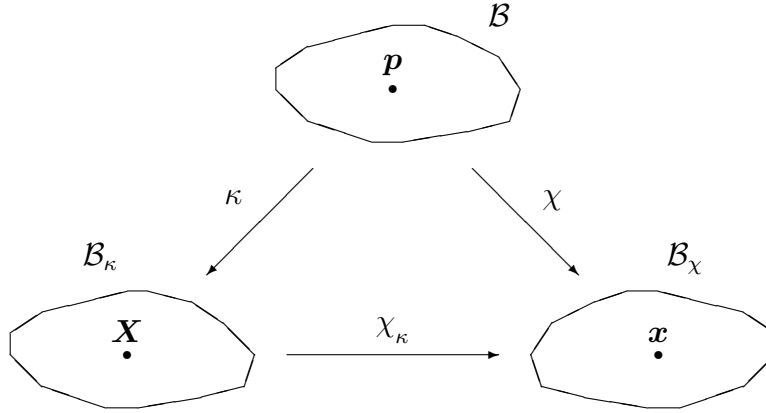


Figure 1: Deformation

Let κ be a reference configuration and χ be an arbitrary configuration of \mathcal{B} . Then the mapping

$$\chi_\kappa = \chi \circ \kappa^{-1} : \mathcal{B}_\kappa \rightarrow \mathcal{B}_\chi, \quad \mathbf{x} = \chi_\kappa(\mathbf{X}) = \chi(\kappa^{-1}(\mathbf{X})), \quad (2.2)$$

is called the *deformation* of \mathcal{B} from κ to χ (Fig. 1). In terms of coordinate systems $(x^i, i = 1, 2, 3)$ and $(X^\alpha, \alpha = 1, 2, 3)$ in the deformed and the reference configurations respectively, the deformation χ_κ is given by

$$x^i = \chi^i(X^\alpha), \quad (2.3)$$

where χ^i are called the *deformation functions*.

The *deformation gradient* of χ relative to κ , denoted by F_κ is defined by

$$F_\kappa = \nabla_{\mathbf{X}} \chi_\kappa. \quad (2.4)$$

By definition, the deformation gradient is the linear approximation of the deformation. Physically, it is a measure of deformation at a point in a small neighborhood. We shall assume that the inverse mapping χ_κ^{-1} exists and the determinant of F_κ is different from zero,

$$J = \det F_\kappa \neq 0. \quad (2.5)$$

When the reference configuration κ is chosen and understood in the context, F_κ will be denoted simply by F .

Relative to the natural bases $\mathbf{e}^\alpha(\mathbf{X})$ and $\mathbf{e}_i(\mathbf{x})$ of the coordinate systems (X^α) and (x^i) respectively, the deformation gradient F can be expressed in the following component form,

$$F = F^i{}_\alpha \mathbf{e}_i(\mathbf{x}) \otimes \mathbf{e}^\alpha(\mathbf{X}), \quad F^i{}_\alpha = \frac{\partial \chi^i}{\partial X^\alpha}. \quad (2.6)$$

In matrix notation,

$$[F] = \begin{bmatrix} \frac{\partial x_1}{\partial X_1} & \frac{\partial x_1}{\partial X_2} & \frac{\partial x_1}{\partial X_3} \\ \frac{\partial x_2}{\partial X_1} & \frac{\partial x_2}{\partial X_2} & \frac{\partial x_2}{\partial X_3} \\ \frac{\partial x_3}{\partial X_1} & \frac{\partial x_3}{\partial X_2} & \frac{\partial x_3}{\partial X_3} \end{bmatrix},$$

where $\frac{\partial \chi^i}{\partial X^\alpha}$ is commonly written as $\frac{\partial x^i}{\partial X^\alpha}$.

Let $d\mathbf{X} = \mathbf{X} - \mathbf{X}_0$ be a small (infinitesimal) material line element in the reference configuration, and $d\mathbf{x} = \chi_\kappa(\mathbf{X}) - \chi_\kappa(\mathbf{X}_0)$ be its image in the deformed configuration, then it follows from the definition that

$$\chi_\kappa(\mathbf{X}) - \chi_\kappa(\mathbf{X}_0) = \nabla_{\mathbf{X}} \chi_\kappa(\mathbf{X}_0)(\mathbf{X} - \mathbf{X}_0) + o(2),$$

or

$$d\mathbf{x} = F d\mathbf{X}, \quad (2.7)$$

since $d\mathbf{X}$ is infinitesimal the higher order term $o(2)$ tends to zero.

Let da_κ and \mathbf{n}_κ be a small material surface element and its unit normal in the reference configuration and da and \mathbf{n} be the corresponding ones in the deformed configuration. And let dv_κ and dv be small material volume elements in the reference and the deformed configurations respectively. Then we have

$$\mathbf{n} da = J F^{-T} \mathbf{n}_\kappa da_\kappa, \quad dv = |J| dv_\kappa. \quad (2.8)$$

To prove these, let the small surface element da_κ be formed from the two line elements $d\mathbf{X}_1$ and $d\mathbf{X}_2$. Then we have from (2.7)

$$\begin{aligned}\mathbf{n}_\kappa &= \frac{d\mathbf{X}_1 \times d\mathbf{X}_2}{|d\mathbf{X}_1 \times d\mathbf{X}_2|}, & \mathbf{n} &= \frac{Fd\mathbf{X}_1 \times Fd\mathbf{X}_2}{|Fd\mathbf{X}_1 \times Fd\mathbf{X}_2|}, \\ da_\kappa &= |d\mathbf{X}_1 \times d\mathbf{X}_2|, & da &= |Fd\mathbf{X}_1 \times Fd\mathbf{X}_2|.\end{aligned}$$

Therefore, for any vector \mathbf{v} , we have

$$\begin{aligned}\mathbf{v} \cdot \mathbf{n} da &= \mathbf{v} \cdot Fd\mathbf{X}_1 \times Fd\mathbf{X}_2 = F(F^{-1}\mathbf{v}) \cdot Fd\mathbf{X}_1 \times Fd\mathbf{X}_2 \\ &= J F^{-1}\mathbf{v} \cdot d\mathbf{X}_1 \times d\mathbf{X}_2 \\ &= J F^{-1}\mathbf{v} \cdot \mathbf{n}_\kappa da_\kappa = \mathbf{v} \cdot JF^{-T}\mathbf{n}_\kappa da_\kappa.\end{aligned}$$

Similarly, let the small volume element dv_κ be formed from three line elements $d\mathbf{X}_1$, $d\mathbf{X}_2$, and $d\mathbf{X}_3$. Then we have

$$\begin{aligned}dv &= |d\mathbf{x}_1 \cdot d\mathbf{x}_2 \times d\mathbf{x}_3| = |Fd\mathbf{X}_1 \cdot Fd\mathbf{X}_2 \times Fd\mathbf{X}_3| \\ &= |\det F| |d\mathbf{X}_1 \cdot d\mathbf{X}_2 \times d\mathbf{X}_3| = |\det F| dv_\kappa.\end{aligned}$$

Note that if $\det F = 1$, the deformation is volume-preserving.

2.2 Strain and rotation

The deformation gradient is a measure of local deformation of the body. We shall introduce other measures of deformation which have more suggestive physical meanings, such as change of shape and orientation. First we shall recall the following theorem from linear algebra:

Theorem (polar decomposition). *For any non-singular tensor F , there exist unique symmetric positive definite tensors V and U and a unique orthogonal tensor R such that*

$$F = RU = VR. \quad (2.9)$$

Since the deformation gradient F is non-singular, the above decomposition holds. We observe that a positive definite symmetric tensor represents a state of pure stretches along three mutually orthogonal axes and an orthogonal tensor a rotation. Therefore, (2.9) assures that any local deformation is a combination of a pure stretch and a rotation.

We call R the *rotation tensor*, while U and V are called the *right* and the *left stretch tensors* respectively. Both stretch tensors measure the local strain, a change of shape, while the tensor R describes the local rotation, a change of orientation, experienced by material elements of the body.

Clearly we have

$$\begin{aligned}U^2 &= F^T F, & V^2 &= F F^T, \\ \det U &= \det V = |\det F|.\end{aligned} \quad (2.10)$$

Since $V = RUR^T$, V and U have the same eigenvalues and their eigenvectors differ only by the rotation R . Their eigenvalues are called the *principal stretches*, and the corresponding eigenvectors are called the *principal directions*.

We shall also introduce the *right* and the *left Cauchy-Green strain tensors* defined by

$$C = F^T F, \quad B = FF^T, \quad (2.11)$$

respectively, which are easier to be calculated than the strain measures U and V from a given F in practice. Note that C and U share the same eigenvectors, while the eigenvalues of U are the positive square root of those of C ; the same is true for B and V .

2.3 Linear strain tensors

The strain tensors introduced in the previous section are valid for finite deformations in general. In the classical linear theory, only small deformations are considered.

We introduce the *displacement vector* from the reference configuration (see Fig. 2),

$$\mathbf{u} = \chi_\kappa(\mathbf{X}) - \mathbf{X},$$

and its gradient,

$$H = \nabla_{\mathbf{X}} \mathbf{u}, \quad [H] = \begin{bmatrix} \frac{\partial u_1}{\partial X_1} & \frac{\partial u_1}{\partial X_2} & \frac{\partial u_1}{\partial X_3} \\ \frac{\partial u_2}{\partial X_1} & \frac{\partial u_2}{\partial X_2} & \frac{\partial u_2}{\partial X_3} \\ \frac{\partial u_3}{\partial X_1} & \frac{\partial u_3}{\partial X_2} & \frac{\partial u_3}{\partial X_3} \end{bmatrix}.$$

Obviously, we have $F = I + H$.

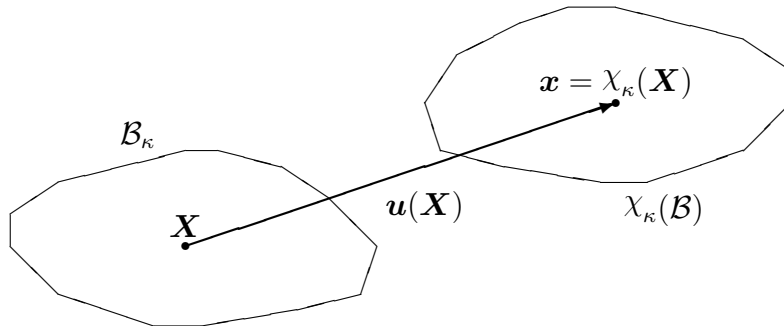


Figure 2: Displacement vector

For small deformations, the displacement gradient H is assumed to be a small quantity, $|H| \ll 1$, and say H is of order $o(1)$. The right stretch tensor U and the rotation tensor

R can then be approximated by

$$\begin{aligned} U &= \sqrt{F^T F} = I + \frac{1}{2}(H + H^T) + o(2) = I + E + o(2), \\ R &= F U^{-1} = I + \frac{1}{2}(H - H^T) + o(2) = I + \tilde{R} + o(2), \end{aligned} \quad (2.12)$$

where

$$E = \frac{1}{2}(H + H^T), \quad \tilde{R} = \frac{1}{2}(H - H^T), \quad (2.13)$$

in components,

$$E_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial X_j} + \frac{\partial u_j}{\partial X_i} \right), \quad \tilde{R}_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial X_j} - \frac{\partial u_j}{\partial X_i} \right),$$

are called the *infinitesimal strain tensor* and the *infinitesimal rotation tensor*, respectively. Note that infinitesimal strain and rotation are the symmetric and skew-symmetric parts of the displacement gradient.

We can give geometrical meanings to the components of the infinitesimal strain tensor E_{ij} relative to a Cartesian coordinate system. Consider two infinitesimal material line segments $d\mathbf{X}_1$ and $d\mathbf{X}_2$ in the reference configuration and their corresponding ones $d\mathbf{x}_1$ and $d\mathbf{x}_2$ in the current configuration. By (2.7), we have

$$d\mathbf{x}_1 \cdot d\mathbf{x}_2 - d\mathbf{X}_1 \cdot d\mathbf{X}_2 = (F^T F - I) d\mathbf{X}_1 \cdot d\mathbf{X}_2 = 2E d\mathbf{X}_1 \cdot d\mathbf{X}_2, \quad (2.14)$$

for small deformations. Now let $d\mathbf{X}_1 = d\mathbf{X}_2 = s_o \mathbf{e}_1$ be a small material line segment in the direction of the unit base vector \mathbf{e}_1 and s be the deformed length. Then we have

$$s^2 - s_o^2 = 2s_o^2 (E \mathbf{e}_1 \cdot \mathbf{e}_1) = 2s_o^2 E_{11},$$

which implies that

$$E_{11} = \frac{s^2 - s_o^2}{2s_o^2} = \frac{(s - s_o)(s + s_o)}{2s_o^2} \simeq \frac{s - s_o}{s_o}.$$

In other words, E_{11} is the change of length per unit original length of a small line segment in the \mathbf{e}_1 -direction. The other diagonal components, E_{22} and E_{33} have similar interpretations as elongation per unit original length in their respective directions.

Similarly, let $d\mathbf{X}_1 = s_o \mathbf{e}_1$ and $d\mathbf{X}_2 = s_o \mathbf{e}_2$ and denote the angle between the two line segments after deformation by θ . Then we have

$$s_o^2 |F \mathbf{e}_1| |F \mathbf{e}_2| \cos \theta - s_o^2 \cos \frac{\pi}{2} = 2s_o^2 (E \mathbf{e}_1 \cdot \mathbf{e}_2),$$

from which, if we write $\gamma = \pi/2 - \theta$, the change from its original right angle, then

$$\frac{\sin \gamma}{2} = \frac{E_{12}}{|F \mathbf{e}_1| |F \mathbf{e}_2|}.$$

Since $|E_{12}| \ll 1$ and $|F\mathbf{e}_i| \simeq 1$, it follows that $\sin \gamma \simeq \gamma$ and we conclude that

$$E_{12} \simeq \frac{\gamma}{2}.$$

Therefore, the component E_{12} is equal to one-half the change of angle between the two line segments originally along the \mathbf{e}_1 - and \mathbf{e}_2 -directions. Other off-diagonal components, E_{23} and E_{13} have similar interpretations as change of angle indicated by their numerical subscripts.

Moreover, since $\det F = \det(1 + H) \simeq 1 + \operatorname{tr} H$ for small deformations, by (2.8)₂ for a small material volume we have

$$\operatorname{tr} E = \operatorname{tr} H \simeq \frac{dv - dv_\kappa}{dv_\kappa}.$$

Thus the sum $E_{11} + E_{22} + E_{33}$ measures the infinitesimal change of volume per unit original volume. Therefore, in the linear theory, if the deformation is incompressible, it follows that

$$\operatorname{tr} E = \operatorname{Div} \mathbf{u} = 0. \quad (2.15)$$

In terms of Cartesian coordinates, the displacement gradient

$$\frac{\partial u_i}{\partial X_j} = \frac{\partial u_i}{\partial x_k} \frac{\partial x_k}{\partial X_j} = \frac{\partial u_i}{\partial x_k} \left(\delta_{kj} + \frac{\partial u_k}{\partial X_j} \right) = \frac{\partial u_i}{\partial x_j} + o(2).$$

In other words, the two displacement gradients

$$\frac{\partial u_i}{\partial X_j} \quad \text{and} \quad \frac{\partial u_i}{\partial x_j}$$

differ in second order terms only. Therefore, since in classical linear theory, the higher order terms are insignificant, it is usually not necessary to introduce the reference configuration in the linear theory. The classical infinitesimal strain and rotation, in the Cartesian coordinate system, are usually defined as

$$E_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \quad \tilde{R}_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} - \frac{\partial u_j}{\partial x_i} \right), \quad (2.16)$$

in the current configuration.

2.4 Motions

A *motion* of the body \mathcal{B} can be regarded as a continuous sequence of deformations in time, *i.e.*, a *motion* χ of \mathcal{B} is regarded as a map,

$$\chi : \mathcal{B}_\kappa \times \mathbb{R} \rightarrow \mathbb{E}, \quad \mathbf{x} = \chi(\mathbf{X}, t). \quad (2.17)$$

We denote the configuration of \mathcal{B} at time t in the motion χ by \mathcal{B}_t .

In practice, the reference configuration κ is often chosen as the configuration in the motion at some instant t_0 , $\kappa = \chi(\cdot, t_0)$, say for example, $t_0 = 0$, so that $\mathbf{X} = \chi(\mathbf{X}, 0)$.

For a fixed material point \mathbf{X} ,

$$\chi(\mathbf{X}, \cdot) : \mathbb{R} \rightarrow \mathbb{E}$$

is a curve called the *path* of the material point \mathbf{X} . The *velocity* \mathbf{v} and the *acceleration* \mathbf{a} are defined as the first and the second time derivatives of position as \mathbf{X} moves along its path,

$$\mathbf{v} = \frac{\partial \chi(\mathbf{X}, t)}{\partial t}, \quad \mathbf{a} = \frac{\partial^2 \chi(\mathbf{X}, t)}{\partial t^2}. \quad (2.18)$$

Lagrangian and Eulerian descriptions

A material body is endowed with some physical properties whose values may change along with the deformation of the body in a motion. A quantity defined on a motion can be described in essentially two different ways: either by the evolution of its value along the path of a material point or by the change of its value at a fixed location in the deformed body. The former is called the material (or a referential description if a reference configuration is used) and the later a spatial description. We shall make them more precise below.

For a given motion χ and a fixed reference configuration κ , consider a quantity, with its value in some space W , defined on the motion of \mathcal{B} by a function

$$f : \mathcal{B} \times \mathbb{R} \rightarrow W. \quad (2.19)$$

Then it can be defined on the reference configuration,

$$\widehat{f} : \mathcal{B}_\kappa \times \mathbb{R} \rightarrow W, \quad (2.20)$$

by

$$\widehat{f}(\mathbf{X}, t) = f(\kappa^{-1}(\mathbf{X}), t) = f(\mathbf{p}, t), \quad \mathbf{X} \in \mathcal{B}_\kappa,$$

and also defined on the position occupied by the body at time t ,

$$\widetilde{f} : \mathcal{B}_t \times \{t\} \rightarrow W, \quad (2.21)$$

by

$$\widetilde{f}(\mathbf{x}, t) = \widehat{f}(\chi^{-1}(\mathbf{x}, t), t) = \widehat{f}(\mathbf{X}, t), \quad \mathbf{x} \in \mathcal{B}_t.$$

As a custom in continuum mechanics, one usually denotes these functions f , \widehat{f} , and \widetilde{f} by the same symbol since they have the same value at the corresponding point, and write, by an abuse of notations,

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

and called them respectively the *material description*, the *referential description* and the *spatial description* of the function f . Sometimes the referential description is referred to as the *Lagrangian description* and the spatial description as the *Eulerian description*.

When a reference configuration is chosen and fixed, one can usually identify the material point \mathbf{p} with its reference position \mathbf{X} . In fact, the material description in (\mathbf{p}, t) is rarely used and the referential description in (\mathbf{X}, t) is often regarded as *the* material description instead.

Possible confusions may arise in such an abuse of notations, especially when differentiations are involved. To avoid such confusions, one may use different notations for differentiation in these situations.

In the referential description, the time derivative is denoted by a dot while the differential operators such as gradient and divergence are denoted by Grad and Div respectively, beginning with capital letters:

$$\dot{f} = \frac{\partial f(\mathbf{X}, t)}{\partial t}, \quad \text{Grad } f = \nabla_{\mathbf{X}} f(\mathbf{X}, t), \quad \text{Div } f(\mathbf{X}, t).$$

In the spatial description, the time derivative is the usual ∂_t and the differential operators beginning with lower-case letters, grad and div:

$$\frac{\partial f}{\partial t} = \frac{\partial f(\mathbf{x}, t)}{\partial t}, \quad \text{grad } f = \nabla_{\mathbf{x}} f(\mathbf{x}, t), \quad \text{div } f(\mathbf{x}, t).$$

The relations between these notations can easily be obtained from the chain rule. Indeed, let f be a scalar field and \mathbf{u} be a vector field. We have

$$\dot{f} = \frac{\partial f}{\partial t} + (\text{grad } f) \cdot \mathbf{v}, \quad \dot{\mathbf{u}} = \frac{\partial \mathbf{u}}{\partial t} + (\text{grad } \mathbf{u})\mathbf{v}, \quad (2.22)$$

and

$$\text{Grad } f = F^T \text{grad } f, \quad \text{Grad } \mathbf{u} = (\text{grad } \mathbf{u})F. \quad (2.23)$$

In particular, taking the velocity \mathbf{v} for \mathbf{u} , it follows that

$$\text{grad } \mathbf{v} = \dot{F}F^{-1}, \quad (2.24)$$

since $\text{Grad } \mathbf{v} = \text{Grad } \dot{\mathbf{x}} = \dot{F}$.

We call \dot{f} the *material time derivative* of f , which is the time derivative of f following the path of the material point. Therefore, by the definition (2.18), we can write the velocity \mathbf{v} and the acceleration \mathbf{a} as

$$\mathbf{v} = \dot{\mathbf{x}}, \quad \mathbf{a} = \ddot{\mathbf{x}},$$

and hence by (2.22)₂,

$$\mathbf{a} = \dot{\mathbf{v}} = \frac{\partial \mathbf{v}}{\partial t} + (\text{grad } \mathbf{v})\mathbf{v}. \quad (2.25)$$

2.5 Relative deformation

Since the reference configuration can be conveniently chosen, we can also choose the current configuration $\chi(\cdot, t)$ as the reference configuration so that past and future deformations can be described relative to the present configuration.

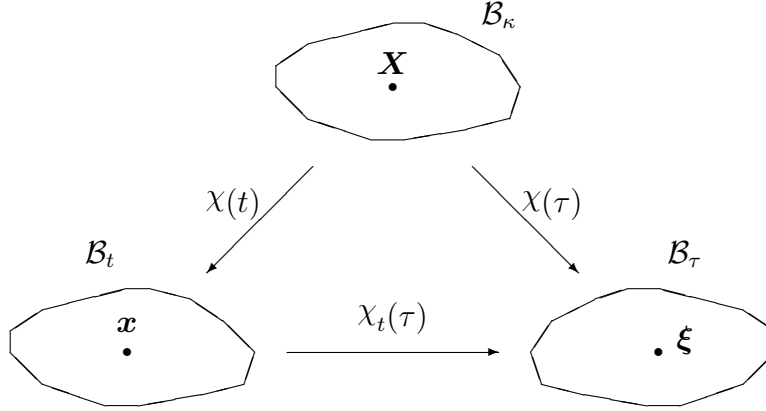


Figure 3: Relative deformation

We denote the position of the material point $\mathbf{X} \in \mathcal{B}_\kappa$ at time τ by $\boldsymbol{\xi}$,

$$\boldsymbol{\xi} = \chi(\mathbf{X}, \tau).$$

Then

$$\mathbf{x} = \chi(\mathbf{X}, t), \quad \boldsymbol{\xi} = \chi_t(\mathbf{x}, \tau) = \chi(\chi^{-1}(\mathbf{x}, t), \tau), \quad (2.26)$$

where $\chi_t(\cdot, \tau) : \mathcal{B}_t \rightarrow \mathcal{B}_\tau$ is the deformation at time τ relative to the configuration at time t or simply called the *relative deformation* (Fig. 3). The *relative deformation gradient* F_t is defined by

$$F_t(\mathbf{x}, \tau) = \nabla_{\mathbf{x}} \chi_t(\mathbf{x}, \tau), \quad (2.27)$$

that is, the deformation gradient at time τ with respect to the configuration at time t . Of course, if $\tau = t$,

$$F_t(\mathbf{x}, t) = I,$$

and we can easily show that

$$F(\mathbf{X}, \tau) = F_t(\mathbf{x}, \tau)F(\mathbf{X}, t). \quad (2.28)$$

Similarly, we can also defined the relative displacement,

$$\mathbf{u}_t(\mathbf{x}, \tau) = \boldsymbol{\xi} - \mathbf{x} = \chi_t(\mathbf{x}, \tau) - \mathbf{x},$$

and the relative displacement gradient,

$$H_t(\mathbf{x}, \tau) = \nabla_{\mathbf{x}} \mathbf{u}_t(\mathbf{x}, \tau).$$

We have

$$F_t(\mathbf{x}, \tau) = I + H_t(\mathbf{x}, \tau),$$

and by the use of (2.28),

$$F(\mathbf{X}, \tau) = (I + H_t(\mathbf{x}, \tau))F(\mathbf{X}, t).$$

Furthermore, from the definition, we have

$$\boldsymbol{\chi}(X, \tau) - \boldsymbol{\chi}(X, t) = \mathbf{u}_t(\boldsymbol{\chi}(X, t), \tau). \quad (2.29)$$

By taking the derivatives with respect to τ , we obtain the velocity and the acceleration of the motion at time τ ,

$$\dot{\boldsymbol{\chi}}(X, \tau) = \frac{\partial \mathbf{u}_t(\mathbf{x}, \tau)}{\partial \tau} = \dot{\mathbf{u}}_t(\mathbf{x}, \tau), \quad \ddot{\boldsymbol{\chi}}(X, \tau) = \ddot{\mathbf{u}}_t(\mathbf{x}, \tau).$$

Note that since $\mathbf{x} = \boldsymbol{\chi}(X, t)$ is independent of τ , the partial derivative with respect to τ keeping \mathbf{x} fixed is nothing but the material time derivative.

Relative description

Recall the material description of a function given by (2.20),

$$f : \mathcal{B}_\kappa \times \mathbb{R} \rightarrow W.$$

By the use of relative motion of the body, we can introduce another description of the function,

$$f_t : \mathcal{B}_t \times \mathbb{R} \rightarrow W,$$

by

$$f_t(\mathbf{x}, \tau) = f(\boldsymbol{\chi}^{-1}(\mathbf{x}, t), \tau) = f(\mathbf{X}, \tau).$$

In fact, we have already used this description above, such as, $F_t(\mathbf{x}, \tau)$, $\mathbf{u}_t(\mathbf{x}, \tau)$, and $H_t(\mathbf{x}, \tau)$. We shall call such description as the *relative description*, in contrast to the frequently used Lagrangian and Eulerian descriptions.

It is interesting to note that for $\tau = t$, the relative description reduces to the Eulerian description, and for $t = t_0$ where t_0 is the time at the reference configuration, then the relative description reduces to the Lagrangian description.

3 Balance laws

3.1 General balance equation

Basic laws of mechanics can all be expressed in general in the following form,

$$\frac{d}{dt} \int_{\mathcal{P}_t} \psi \, dv = \int_{\partial \mathcal{P}_t} \Phi_\psi \mathbf{n} \, da + \int_{\mathcal{P}_t} \sigma_\psi \, dv, \quad (3.1)$$

for any bounded regular subregion of the body, called a part $\mathcal{P} \subset \mathcal{B}$ and the vector field \mathbf{n} , the outward unit normal to the boundary of the region $\mathcal{P}_t \subset \mathcal{B}_t$ in the current configuration. The quantities ψ and σ_ψ are tensor fields of certain order m , and Φ_ψ is a tensor field of order $m + 1$, say $m = 0$ or $m = 1$ so that ψ is a scalar or vector quantity, and respectively Φ_ψ is a vector or second order tensor quantity.

The relation (3.1), called the *general balance* of ψ in integral form, is interpreted as asserting that the rate of change of the quantity ψ in a part \mathcal{P} of a body is affected by the flow of ψ through the boundary of \mathcal{P} and the growth of ψ within \mathcal{P} . We call Φ_ψ the *flux* of ψ and σ_ψ the *supply* of ψ .

We are interested in the local forms of the integral balance (3.1) at a point in the region \mathcal{P}_t . The derivation of local forms rest upon smoothness assumption of the tensor fields ψ , Φ_ψ , and σ_ψ .

First of all, we need the following theorem, which is a three-dimensional version of the formula in calculus for differentiation under the integral sign on a moving interval (Leibniz's rule), namely

$$\frac{\partial}{\partial t} \int_{g(t)}^{f(t)} \psi(x, t) \, dx = \int_{g(t)}^{f(t)} \frac{\partial \psi}{\partial t} \, dx + \psi(f(t), t) \dot{f}(t) - \psi(g(t), t) \dot{g}(t).$$

Theorem (transport theorem). *Let $V(t)$ be a regular region and $u_n(\mathbf{x}, t)$ be the outward normal speed of a surface point $\mathbf{x} \in \partial V(t)$. Then for any smooth tensor field $\psi(\mathbf{x}, t)$, we have*

$$\frac{d}{dt} \int_V \psi \, dv = \int_V \frac{\partial \psi}{\partial t} \, dv + \int_{\partial V} \psi u_n \, da. \quad (3.2)$$

In this theorem, if $V(t)$ is a material region \mathcal{P}_t , *i.e.*, it always consists of the same material points of a part $\mathcal{P} \subset \mathcal{B}$, then $u_n = \dot{\mathbf{x}} \cdot \mathbf{n}$ and (3.2) becomes

$$\frac{d}{dt} \int_{\mathcal{P}_t} \psi \, dv = \int_{\mathcal{P}_t} \frac{\partial \psi}{\partial t} \, dv + \int_{\partial \mathcal{P}_t} \psi \dot{\mathbf{x}} \cdot \mathbf{n} \, da. \quad (3.3)$$

3.2 Local balance equation

For a material region \mathcal{V} , the equation of general balance in integral form (3.1) becomes

$$\int_{\mathcal{V}} \frac{\partial \psi}{\partial t} dv + \int_{\partial \mathcal{V}} \psi \dot{\mathbf{x}} \cdot \mathbf{n} da = \int_{\partial \mathcal{V}} \Phi_{\psi} \mathbf{n} da + \int_{\mathcal{V}} \sigma_{\psi} dv. \quad (3.4)$$

We can obtain the local balance equation at a regular point from the above integral equation. We consider a small material region \mathcal{V} containing \mathbf{x} . By the use of the divergence theorem, (3.4) becomes

$$\int_{\mathcal{V}} \left\{ \frac{\partial \psi}{\partial t} + \operatorname{div}(\psi \dot{\mathbf{x}} - \Phi_{\psi}) - \sigma_{\psi} \right\} dv = 0. \quad (3.5)$$

Since the integrand is smooth and the equation (3.5) holds for any \mathcal{V} , such that $\mathbf{x} \in \mathcal{V}$, the integrand must vanish at \mathbf{x} . Therefore we have

Theorem (local balance equation). *At a regular point \mathbf{x} , the general balance equation reduces to*

$$\frac{\partial \psi}{\partial t} + \operatorname{div}(\psi \dot{\mathbf{x}} - \Phi_{\psi}) - \sigma_{\psi} = 0. \quad (3.6)$$

3.3 Balance equations in reference coordinates

Sometimes, for solid bodies, it is more convenient to use the referential description. The corresponding relations for the balance equation (3.6) can be derived in a similar manner. We begin with the integral form (3.1) now written in the reference configuration κ ,

$$\frac{d}{dt} \int_{\mathcal{P}_{\kappa}} \psi_{\kappa} dv_{\kappa} = \int_{\partial \mathcal{P}_{\kappa}} \Phi_{\kappa}^{\psi} \mathbf{n}_{\kappa} da_{\kappa} + \int_{\mathcal{P}_{\kappa}} \sigma_{\kappa}^{\psi} dv_{\kappa}. \quad (3.7)$$

In view of the relations for volume elements and surface elements (2.8), $dv = J dv_{\kappa}$, and $\mathbf{n} da = J F^{-T} \mathbf{n}_{\kappa} da_{\kappa}$, the corresponding quantities are defined as

$$\psi_{\kappa} = J \psi, \quad \Phi_{\kappa}^{\psi} = J \Phi_{\psi} F^{-T}, \quad \sigma_{\kappa}^{\psi} = J \sigma_{\psi}. \quad (3.8)$$

The transport theorem (3.2) remains valid for $\psi_{\kappa}(\mathbf{X}, t)$ in a movable region $V(t)$,

$$\frac{d}{dt} \int_V \psi_{\kappa} dv_{\kappa} = \int_V \dot{\psi}_{\kappa} dv_{\kappa} + \int_{\partial V} \psi_{\kappa} U_{\kappa} da_{\kappa}, \quad (3.9)$$

where $U_{\kappa}(\mathbf{X}, t)$ is the outward normal speed of a surface point $\mathbf{X} \in \partial V(t)$.

However, for a material region in the reference configuration, the normal speed of the surface points on $\partial \mathcal{V}_{\kappa}$ is zero since a material region is a fixed region in the reference configuration. Therefore, from (3.7), we obtain

$$\int_{\mathcal{V}_{\kappa}} \dot{\psi}_{\kappa} dv_{\kappa} = \int_{\partial \mathcal{V}_{\kappa}} \Phi_{\kappa}^{\psi} \mathbf{n}_{\kappa} da_{\kappa} + \int_{\mathcal{V}_{\kappa}} \sigma_{\kappa}^{\psi} dv_{\kappa}, \quad (3.10)$$

from which we obtain the local balance equation in the reference configuration,

$$\dot{\psi}_{\kappa} - \operatorname{Div} \Phi_{\kappa}^{\psi} - \sigma_{\kappa}^{\psi} = 0. \quad (3.11)$$

3.4 Conservation of mass

Let $\rho(\mathbf{x}, t)$ denote the mass density of \mathcal{B}_t in the current configuration. Since the material is neither destroyed nor created in any motion in the absence of chemical reactions, we have

Conservation of mass. *The total mass of any part $\mathcal{P} \subset \mathcal{B}$ does not change in any motion,*

$$\frac{d}{dt} \int_{\mathcal{P}_t} \rho \, dv = 0. \quad (3.12)$$

By comparison, it is a special case of the general balance equation (3.1) with no flux and no supply,

$$\psi = \rho, \quad \Phi_\psi = 0, \quad \sigma_\psi = 0,$$

and hence from (3.6) we obtain the equation of mass conservation,

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \dot{\mathbf{x}}) = 0, \quad (3.13)$$

which can also be written as

$$\dot{\rho} + \rho \operatorname{div} \dot{\mathbf{x}} = 0.$$

The equation (3.12) states that the total mass of any part is constant in time. In particular, if $\rho_\kappa(\mathbf{X})$ denote the mass density of \mathcal{B}_κ in the reference configuration, then

$$\int_{\mathcal{P}_\kappa} \rho_\kappa \, dv_\kappa = \int_{\mathcal{P}_t} \rho \, dv, \quad (3.14)$$

which implies that

$$\rho_\kappa = \rho J, \quad \text{or} \quad \rho = \frac{\rho_\kappa}{\det F}. \quad (3.15)$$

This is another form of the conservation of mass in the referential description, which also follows from the general expression (3.11) and (3.8).

3.5 Equation of motion

For a deformable body, the linear momentum and the angular momentum with respect to a point $\mathbf{x}_o \in \mathcal{E}$ of a part $\mathcal{P} \subset \mathcal{B}$ in the motion can be defined respectively as

$$\int_{\mathcal{P}_t} \rho \dot{\mathbf{x}} \, dv, \quad \text{and} \quad \int_{\mathcal{P}_t} \rho (\mathbf{x} - \mathbf{x}_o) \times \dot{\mathbf{x}} \, dv.$$

In laying down the laws of motion, we follow the classical approach developed by Newton and Euler, according to which the change of momentum is produced by the action of forces. There are two type of forces, namely, one acts throughout the volume, called the *body force*, and one acts on the surface of the body, called the *surface traction*.

Euler's laws of motion. *Relative to an inertial frame, the motion of any part $\mathcal{P} \subset \mathcal{B}$ satisfies*

$$\begin{aligned}\frac{d}{dt} \int_{\mathcal{P}_t} \rho \dot{\mathbf{x}} dv &= \int_{\mathcal{P}_t} \rho \mathbf{b} dv + \int_{\partial\mathcal{P}_t} \mathbf{t} da, \\ \frac{d}{dt} \int_{\mathcal{P}_t} \rho (\mathbf{x} - \mathbf{x}_o) \times \dot{\mathbf{x}} dv &= \int_{\mathcal{P}_t} \rho (\mathbf{x} - \mathbf{x}_o) \times \mathbf{b} dv + \int_{\partial\mathcal{P}_t} (\mathbf{x} - \mathbf{x}_o) \times \mathbf{t} da.\end{aligned}$$

We remark that the existence of inertial frames (an equivalent of Newton's first law) is essential to establish the Euler's laws (equivalent of Newton's second law) in the above forms. Roughly speaking, a coordinate system at rest for \mathcal{E} is usually regarded as an inertial frame.

We call \mathbf{b} the *body force density* (per unit mass), and \mathbf{t} the *surface traction* (per unit surface area). Unlike the body force $\mathbf{b} = \mathbf{b}(\mathbf{x}, t)$, such as the gravitational force, the traction \mathbf{t} at \mathbf{x} depends, in general, upon the surface $\partial\mathcal{P}_t$ on which \mathbf{x} lies. It is obvious that there are infinite many parts $\mathcal{P} \subset \mathcal{B}$, such that $\partial\mathcal{P}_t$ may also contain \mathbf{x} . However, following Cauchy, it is assumed in classical continuum mechanics that the tractions on all like-oriented surfaces with a common tangent plane at \mathbf{x} are the same.

Postulate (Cauchy). *Let \mathbf{n} be the unit normal to the surface $\partial\mathcal{P}_t$ at \mathbf{x} , then*

$$\mathbf{t} = \mathbf{t}(\mathbf{x}, t, \mathbf{n}). \quad (3.16)$$

An immediate consequence of this postulate is the well-known theorem which ensures the existence of stress tensor. The proof of the theorem can be found in most books of mechanics.

Theorem (Cauchy). *Suppose that $\mathbf{t}(\cdot, \mathbf{n})$ is a continuous function of \mathbf{x} , and $\ddot{\mathbf{x}}$, \mathbf{b} are bounded in \mathcal{B}_t . Then Cauchy's postulate and Euler's first law implies the existence of a second order tensor T , such that*

$$\mathbf{t}(\mathbf{x}, t, \mathbf{n}) = T(\mathbf{x}, t) \mathbf{n}. \quad (3.17)$$

The tensor field $T(\mathbf{x}, t)$ in (3.17) is called the *Cauchy stress tensor*. In components, the traction force (3.17) can be written as

$$t_i = T_{ij} n_j.$$

Therefore, the stress tensor T_{ij} represents the i -th component of the traction force on the surface point with normal in the direction of j -th coordinate.

With (3.17) Euler's first law becomes

$$\frac{d}{dt} \int_{\mathcal{P}_t} \rho \dot{\mathbf{x}} dv = \int_{\mathcal{P}_t} \rho \mathbf{b} dv + \int_{\partial\mathcal{P}_t} T \mathbf{n} da. \quad (3.18)$$

Comparison with the general balance equation (3.1) leads to

$$\psi = \rho \dot{\mathbf{x}}, \quad \Phi_\psi = T, \quad \sigma_\psi = \rho \mathbf{b},$$

in this case, and hence from (3.6) we obtain the balance equation of linear momentum,

$$\frac{\partial}{\partial t}(\rho \dot{\mathbf{x}}) + \operatorname{div}(\rho \dot{\mathbf{x}} \otimes \dot{\mathbf{x}} - T) - \rho \mathbf{b} = 0. \quad (3.19)$$

This equation, also known as the *equation of motion*, can be rewritten in the following more familiar form by the use of (3.13),

$$\rho \ddot{\mathbf{x}} - \operatorname{div} T = \rho \mathbf{b}. \quad (3.20)$$

A similar argument for Euler's second law as a special case of (3.1) with

$$\psi = (\mathbf{x} - \mathbf{x}_o) \times \rho \dot{\mathbf{x}}, \quad \Phi_\psi \mathbf{n} = (\mathbf{x} - \mathbf{x}_o) \times T \mathbf{n}, \quad \sigma_\psi = (\mathbf{x} - \mathbf{x}_o) \times \rho \mathbf{b},$$

leads to

$$T = T^T, \quad (3.21)$$

after some simplification from (3.6) by the use of (3.19). In other words, the symmetry of the stress tensor is a consequence of the conservation of angular momentum.

3.6 Conservation of energy

Besides the kinetic energy, the total energy of a deformable body consists of another part called the internal energy,

$$\int_{\mathcal{P}_t} \left(\rho \varepsilon + \frac{\rho}{2} \dot{\mathbf{x}} \cdot \dot{\mathbf{x}} \right) dv,$$

where $\varepsilon(\mathbf{x}, t)$ is called the specific *internal energy density*. The rate of change of the total energy is partly due to the mechanical power from the forces acting on the body and partly due to the energy inflow over the surface and the external energy supply.

Conservation of energy. *Relative to an inertial frame, the change of energy for any part $\mathcal{P} \subset \mathcal{B}$ is given by*

$$\frac{d}{dt} \int_{\mathcal{P}_t} \left(\rho \varepsilon + \frac{\rho}{2} \dot{\mathbf{x}} \cdot \dot{\mathbf{x}} \right) dv = \int_{\partial \mathcal{P}_t} (\dot{\mathbf{x}} \cdot T \mathbf{n} - \mathbf{q} \cdot \mathbf{n}) da + \int_{\mathcal{P}_t} (\rho \dot{\mathbf{x}} \cdot \mathbf{b} + \rho r) dv. \quad (3.22)$$

We call $\mathbf{q}(\mathbf{x}, t)$ the *heat flux* vector (or energy flux), and $r(\mathbf{x}, t)$ the *energy supply density* due to external sources, such as radiation. Comparison with the general balance equation (3.1), we have

$$\psi = \left(\rho \varepsilon + \frac{\rho}{2} \dot{\mathbf{x}} \cdot \dot{\mathbf{x}} \right), \quad \Phi_\psi = T \dot{\mathbf{x}} - \mathbf{q}, \quad \sigma_\psi = \rho (\dot{\mathbf{x}} \cdot \mathbf{b} + r),$$

and hence we have the following local balance equation of total energy,

$$\frac{\partial}{\partial t} \left(\rho \varepsilon + \frac{\rho}{2} \dot{\mathbf{x}} \cdot \dot{\mathbf{x}} \right) + \operatorname{div} \left((\rho \varepsilon + \frac{\rho}{2} \dot{\mathbf{x}} \cdot \dot{\mathbf{x}}) \dot{\mathbf{x}} + \mathbf{q} - T \dot{\mathbf{x}} \right) = \rho (r + \dot{\mathbf{x}} \cdot \mathbf{b}), \quad (3.23)$$

The energy equation (3.23) can be simplified by subtracting the inner product of the equation of motion (3.20) with the velocity $\dot{\mathbf{x}}$,

$$\rho \dot{\varepsilon} + \operatorname{div} \mathbf{q} = T \cdot \operatorname{grad} \dot{\mathbf{x}} + \rho r. \quad (3.24)$$

This is called the balance equation of internal energy. Note that the internal energy is not conserved and the term $T \cdot \operatorname{grad} \dot{\mathbf{x}}$ is the rate of work due to deformation.

Summary of basic equations

By the use of material time derivative (2.22), the field equations can be written as follows:

$$\begin{aligned} \dot{\rho} + \rho \operatorname{div} \mathbf{v} &= 0, \\ \rho \dot{\mathbf{v}} - \operatorname{div} T &= \rho \mathbf{b}, \\ T &= T^T, \\ \rho \dot{\varepsilon} + \operatorname{div} \mathbf{q} - T \cdot \operatorname{grad} \mathbf{v} &= \rho r, \end{aligned} \quad (3.25)$$

In components:

$$\begin{aligned} \frac{\partial \rho}{\partial t} + v_i \frac{\partial \rho}{\partial x_i} + \rho \frac{\partial v_i}{\partial x_i} &= 0, \\ \rho \left(\frac{\partial v_i}{\partial t} + v_j \frac{\partial v_i}{\partial x_j} \right) - \frac{\partial T_{ij}}{\partial x_j} &= \rho b_i, \\ T_{ij} &= T_{ji}, \\ \rho \left(\frac{\partial \varepsilon}{\partial t} + v_j \frac{\partial \varepsilon}{\partial x_j} \right) + \frac{\partial q_j}{\partial x_j} - T_{ij} \frac{\partial v_i}{\partial x_j} &= \rho r, \end{aligned} \quad (3.26)$$

where (x_i) is the Cartesian coordinate system at the present state. This is the balance equations in Eulerian description (in variables (\mathbf{x}, t)).

3.7 Basic equations in material coordinates

It is sometimes more convenient to rewrite the basic equations in material description relative to a reference configuration κ . They can easily be obtained from (3.11),

$$\begin{aligned} \rho &= \frac{\rho_\kappa}{\det F}, \\ \rho_\kappa \ddot{\mathbf{x}} &= \operatorname{Div} T_\kappa + \rho_\kappa \mathbf{b}, \\ T_\kappa F^T &= F T_\kappa^T, \\ \rho_\kappa \dot{\varepsilon} + \operatorname{Div} \mathbf{q}_\kappa &= T_\kappa \cdot \dot{F} + \rho_\kappa r, \end{aligned} \quad (3.27)$$

where the following definitions have been introduced according to (3.8):

$$T_\kappa = JTF^{-T}, \quad \mathbf{q}_\kappa = JF^{-1}\mathbf{q}. \quad (3.28)$$

In components,

$$(T_\kappa)_{i\alpha} = JT_{ij} \frac{\partial X_\alpha}{\partial x_j}, \quad (q_\kappa)_\alpha = Jq_j \frac{\partial X_\alpha}{\partial x_j}.$$

$J = \det F$ is the determinant of the Jacobian matrix $\left[\begin{smallmatrix} (x_1, x_2, x_3) \\ (X_1, X_2, X_3) \end{smallmatrix} \right]$, where (x_i) and (X_α) are the Cartesian coordinate systems at the present and the reference configurations respectively.

T_κ is called the (First) *Piola–Kirchhoff stress tensor* and \mathbf{q}_κ is called the *material heat flux*. Note that unlike the Cauchy stress tensor T , the Piola–Kirchhoff stress tensor T_κ is not symmetric. The definition has been introduced according to the relation (2.8), which gives the relation,

$$\int_S T\mathbf{n} da = \int_{S_\kappa} T_\kappa \mathbf{n}_\kappa da_\kappa. \quad (3.29)$$

In other words, $T\mathbf{n}$ is the surface traction per unit area in the current configuration, while $T_\kappa \mathbf{n}_\kappa$ is the surface traction measured per unit area in the reference configuration. Note that the magnitude of two traction forces are generally different, however they are parallel vectors.

In components, the equation of motion in material coordinate becomes

$$\begin{aligned} \rho_\kappa \ddot{x}_i &= \frac{\partial (T_\kappa)_{i\alpha}}{\partial X_\alpha} + \rho_\kappa b_i, \\ (T_\kappa)_{i\alpha} \frac{\partial x_j}{\partial X_\alpha} &= \frac{\partial x_i}{\partial X_\alpha} (T_\kappa)_{j\alpha}. \end{aligned} \quad (3.30)$$

This is the balance equations in Lagrangian description (in variables (\mathbf{X}, t)).

3.8 Boundary value problem

Let $\Omega = \mathcal{B}_t$ be the open region occupied by a solid body at the present time t , $\partial\Omega = \Gamma_1 \cup \Gamma_2$ be its boundary, and \mathbf{n} be the exterior unit normal to the boundary.

The balance laws (3.26) in Eulerian description,

$$\begin{aligned} \frac{\partial \rho}{\partial t} + v_i \frac{\partial \rho}{\partial x_i} + \rho \frac{\partial v_i}{\partial x_i} &= 0, \\ \rho \left(\frac{\partial v_i}{\partial t} + v_j \frac{\partial v_i}{\partial x_j} \right) - \frac{\partial T_{ij}}{\partial x_j} &= \rho b_i, \\ T_{ij} &= T_{ji}, \end{aligned}$$

are the governing equations for the initial boundary value problem to determine the fields of density $\rho(\mathbf{x}, t)$ and velocity $\mathbf{v}(\mathbf{x}, t)$ with the following conditions:

Initial condition:

$$\begin{aligned}\rho(\mathbf{x}, 0) &= \rho_0(\mathbf{x}) & \forall \mathbf{x} \in \Omega, \\ \mathbf{v}(\mathbf{x}, 0) &= \mathbf{v}_0(\mathbf{x}) & \forall \mathbf{x} \in \Omega,\end{aligned}$$

Boundary condition:

$$\begin{aligned}\mathbf{v}(\mathbf{x}, t) &= 0 & \forall \mathbf{x} \in \Gamma_1, \\ T(\mathbf{x}, t)\mathbf{n} &= \mathbf{f}(\mathbf{x}, t) & \forall \mathbf{x} \in \Gamma_2,\end{aligned}$$

To solve this initial boundary value problem, we need the constitutive equation for the Cauchy stress $T(\mathbf{x}, t)$ as a function in terms of the fields of density $\rho(\mathbf{x}, t)$ and the velocity $\mathbf{v}(\mathbf{x}, t)$. Constitutive equations of this type characterize the behavior of general fluids, such as elastic fluid, Navier-Stokes fluid, and non-Newtonian fluid.

For solid bodies, it is more convenient to use the Lagrangian description of the equation of motion (3.30),

$$\rho_\kappa \ddot{u}_i = \frac{\partial (T_\kappa)_{i\alpha}}{\partial X_\alpha} + \rho_\kappa b_i,$$

for the determination of the displacement vector $\mathbf{u}(\mathbf{X}, t) = \mathbf{x}(\mathbf{X}, t) - \mathbf{X}$, with the following conditions:

Initial condition:

$$\begin{aligned}\mathbf{u}(\mathbf{X}, 0) &= \mathbf{u}_0(\mathbf{X}) & \forall \mathbf{X} \in \Omega_\kappa, \\ \dot{\mathbf{u}}(\mathbf{X}, 0) &= \mathbf{u}_1(\mathbf{X}) & \forall \mathbf{X} \in \Omega_\kappa,\end{aligned}$$

Boundary condition:

$$\begin{aligned}\mathbf{u}(\mathbf{X}, t) &= \mathbf{u}_\kappa(\mathbf{X}, t) & \forall \mathbf{X} \in \Gamma_{\kappa_1}, \\ T_\kappa(\mathbf{X}, t)\mathbf{n}_\kappa &= \mathbf{f}_\kappa(\mathbf{X}, t) & \forall \mathbf{X} \in \Gamma_{\kappa_2},\end{aligned}$$

where $\Omega_\kappa = \mathcal{B}_\kappa$ is the open region occupied by the body at the reference configuration, and $\partial\Omega_\kappa = \Gamma_{\kappa_1} \cup \Gamma_{\kappa_2}$ its boundary.

To complete the formulation of the boundary value problems, we need the constitutive equation for the stress tensor in terms of the displacement field $\mathbf{u}(\mathbf{X}, t)$. General constitutive theory of material bodies will be discussed in the following chapters.

4 Euclidean objectivity

Properties of material bodies are described mathematically by constitutive equations. Intuitively, there is a simple idea that material properties must be independent of observers, which is fundamental in the formulation of constitutive equations. In order to explain this, one has to know what an *observer* is, so as to define what *independence of observer* means.

4.1 Frame of reference, observer

The event world \mathcal{W} is a four-dimensional space-time in which physical events occur at some places and certain instants. Let \mathcal{T} be the collection of instants and \mathcal{W}_s be the placement space of simultaneous events at the instant s , then the classical space-time can be expressed as the disjoint union of placement spaces of simultaneous events at each instant,

$$\mathcal{W} = \bigcup_{s \in \mathcal{T}} \mathcal{W}_s.$$

A point $p_s \in \mathcal{W}$ is called an event, which occurs at the instant s and the place $p \in \mathcal{W}_s$. At different instants s and \bar{s} , the spaces \mathcal{W}_s and $\mathcal{W}_{\bar{s}}$ are two disjoint spaces. Thus it is impossible to determine the distance between two non-simultaneous events at p_s and $p_{\bar{s}}$ if $s \neq \bar{s}$, and hence \mathcal{W} is not a product space of space and time. However, it can be set into correspondence with a product space through a frame of reference on \mathcal{W} .

Definition. (*Frame of reference*): A frame of reference is a one-to-one mapping

$$\phi : \mathcal{W} \rightarrow \mathbb{E} \times \mathbb{R}, \quad \text{taking } p_s \mapsto (\mathbf{x}, t),$$

i.e., taking $p \mapsto \mathbf{x}$, $s \mapsto t$, where \mathbb{R} is the space of real numbers and \mathbb{E} is a three-dimensional Euclidean space. We shall denote the map taking $p \mapsto \mathbf{x}$ as the map $\phi_s : \mathcal{W}_s \rightarrow \mathbb{E}$.

Of course, there are infinite many frames of reference. Each one of them may be regarded as an *observer*, since it can be depicted as a person taking a snapshot so that the image of ϕ_s is a picture (three-dimensional at least conceptually) of the placements of the events at some instant s , from which the distance between two simultaneous events can be measured. A sequence of events can also be recorded as video clips depicting the change of events in time by an observer.

Now, suppose that two observers are recording the same events with video cameras. In order to compare their video clips regarding the locations and time, they must have a mutual agreement that the clock of their cameras must be synchronized so that simultaneous events can be recognized and since during the recording two observers may move

independently while taking pictures with their cameras from different angles, there will be a relative motion and a relative orientation between them. We shall make such a consensus among observers explicit mathematically.

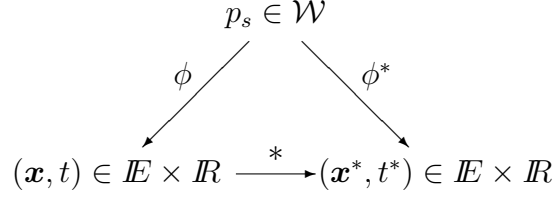


Figure 4: A change of frame

Let ϕ and ϕ^* be two frames of reference. They are related by the composite map $\phi^* \circ \phi^{-1}$,

$$\phi^* \circ \phi^{-1} : \mathbb{E} \times \mathbb{R} \rightarrow \mathbb{E} \times \mathbb{R}, \quad \text{taking } (\mathbf{x}, t) \mapsto (\mathbf{x}^*, t^*),$$

where (\mathbf{x}, t) and (\mathbf{x}^*, t^*) are the position and time of the same event observed by ϕ and ϕ^* simultaneously. Physically, an arbitrary map would be irrelevant as long as we are interested in establishing a consensus among observers, which requires preservation of distance between simultaneous events and time interval as well as the sense of time.

Definition. (*Euclidean change of frame*): A change of frame (observer) from ϕ to ϕ^* taking $(\mathbf{x}, t) \mapsto (\mathbf{x}^*, t^*)$, is an isometry of space and time given by

$$\mathbf{x}^* = Q(t)(\mathbf{x} - \mathbf{x}_0) + \mathbf{c}(t), \quad t^* = t + a, \quad (4.1)$$

for some constant time difference $a \in \mathbb{R}$, some relative translation $\mathbf{c} : \mathbb{R} \rightarrow \mathbb{E}$ with respect to the reference point $\mathbf{x}_0 \in \mathbb{E}$ and some orthogonal transformation $Q : \mathbb{R} \rightarrow \mathcal{O}(V)$.

Such a transformation will be called a *Euclidean transformation*. In particular, $* := \phi_t^* \circ \phi_t^{-1} : \mathbb{E} \rightarrow \mathbb{E}$ is given by

$$*(\mathbf{x}) = \mathbf{x}^* = Q(t)(\mathbf{x} - \mathbf{x}_o) + \mathbf{c}(t), \quad (4.2)$$

which is a time-dependent rigid transformation consisting of an orthogonal transformation and a translation. We shall often call $Q(t)$ the *orthogonal part* of the change of frame from ϕ to ϕ^* .

Euclidean changes of frame will often be called *changes of frame* for simplicity, since they are the only changes of frame among consenting observers of our concern for the purpose of discussing frame-indifference in continuum mechanics.

All consenting observers form an equivalent class, denoted by \mathfrak{E} , among the set of all observers, i.e., for any $\phi, \phi^* \in \mathfrak{E}$, there exists a Euclidean change of frame from $\phi \rightarrow \phi^*$. From now on, only classes of consenting observers will be considered. Therefore, any observer, would mean any observer in some \mathfrak{E} , and a change of frame, would mean a Euclidean change of frame.

Motion and deformation of a body

In Chapter 2, concerning the deformation and the motion, we have tacitly assumed that they are observed by an observer in a frame of reference. Since the later discussions involve different observers, we need to explicitly indicate the frame of reference in the kinematic quantities. Therefore, the placement of a body \mathcal{B} in \mathcal{W}_t is a mapping

$$\chi_t : \mathcal{B} \rightarrow \mathcal{W}_t.$$

for an observer ϕ with $\phi_t : \mathcal{W}_t \rightarrow \mathbb{E}$. The motion can be viewed as a composite mapping $\chi_{\phi_t} := \phi_t \circ \chi_t$,

$$\chi_{\phi_t} : \mathcal{B} \rightarrow \mathbb{E}, \quad \mathbf{x} = \chi_{\phi_t}(p) = \phi_t(\chi_t(p)), \quad p \in \mathcal{B}.$$

This mapping identifies the body with a region in the Euclidean space, $\mathcal{B}_{\chi_t} := \chi_{\phi_t}(\mathcal{B}) \subset \mathbb{E}$ (see the right part of Figure 5). We call χ_{ϕ_t} a *configuration* of \mathcal{B} at the instant t in the frame ϕ , and a motion of \mathcal{B} is a sequence of configurations of \mathcal{B} in time, $\chi_\phi = \{\chi_{\phi_t}, t \in \mathbb{R} \mid \chi_{\phi_t} : \mathcal{B} \rightarrow \mathbb{E}\}$. We can also express a motion as

$$\chi_\phi : \mathcal{B} \times \mathbb{R} \rightarrow \mathbb{E}, \quad \mathbf{x} = \chi_\phi(p, t) = \chi_{\phi_t}(p), \quad p \in \mathcal{B}.$$

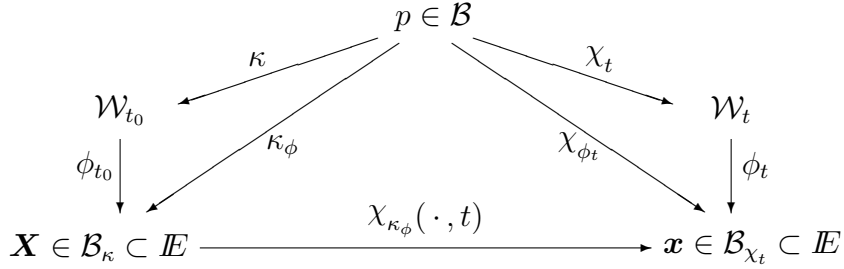


Figure 5: Motion χ_{ϕ_t} , reference configuration κ_ϕ and deformation $\chi_{\kappa_\phi}(\cdot, t)$

Reference configuration

We regard a body \mathcal{B} as a set of material points. Although it is possible to endow the body as a manifold with a differentiable structure and topology for doing mathematics on the body, to avoid such mathematical subtleties, usually a particular configuration is chosen as reference (see the left part of Figure 5),

$$\kappa_\phi : \mathcal{B} \rightarrow \mathbb{E}, \quad \mathbf{X} = \kappa_\phi(p), \quad \mathcal{B}_\kappa := \kappa_\phi(\mathcal{B}) \subset \mathbb{E},$$

so that the motion at an instant t is a one-to-one mapping

$$\chi_{\kappa_\phi}(\cdot, t) : \mathcal{B}_\kappa \rightarrow \mathcal{B}_{\chi_t}, \quad \mathbf{x} = \chi_{\kappa_\phi}(\mathbf{X}, t) = \chi_{\phi_t}(\kappa_\phi^{-1}(\mathbf{X})), \quad \mathbf{X} \in \mathcal{B}_\kappa,$$

defined on a domain in the Euclidean space \mathbb{E} for which topology and differentiability are well defined. This mapping is called a *deformation* from κ to χ_t in the frame ϕ and a motion is then a sequence of deformations in time.

Remember that a configuration is a placement of a body relative to an observer. Therefore, for the reference configuration κ_ϕ , there is some instant, say t_0 , at which the reference placement κ of the body is chosen (see Figure 5).

4.2 Objective tensors

The change of frame (4.1) on the Euclidean space \mathbb{E} gives rise to a linear mapping on the translation space V , in the following way: Let $\mathbf{u}(\phi) = \mathbf{x}_2 - \mathbf{x}_1 \in V$ be the difference vector of $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{E}$ in the frame ϕ , and $\mathbf{u}(\phi^*) = \mathbf{x}_2^* - \mathbf{x}_1^* \in V$ be the corresponding difference vector in the frame ϕ^* , then from (4.1), it follows immediately that

$$\mathbf{u}(\phi^*) = Q(t)\mathbf{u}(\phi),$$

where $Q(t) \in \mathcal{O}(V)$ is the orthogonal part of the change of frame $\phi \rightarrow \phi^*$.

Any vector quantity in V , which has this transformation property, is said to be objective with respect to Euclidean transformations, *objective* in the sense that it pertains to a quantity of its real nature rather than its values as affected by different observers. This concept of objectivity can be generalized to any tensor spaces of V . Let

$$s : \mathfrak{E} \rightarrow \mathbb{R}, \quad \mathbf{u} : \mathfrak{E} \rightarrow V, \quad T : \mathfrak{E} \rightarrow V \otimes V,$$

where \mathfrak{E} is the Euclidean class of frames of reference. They are scalar, vector and (second order) tensor *observable* quantities respectively. We call $f(\phi)$ the value of the quantity f observed in the frame ϕ .

Definition. Let s , \mathbf{u} , and T be scalar-, vector-, (second order) tensor-valued functions respectively. If relative to a change of frame from ϕ to ϕ^* ,

$$\begin{aligned} s(\phi^*) &= s(\phi), \\ \mathbf{u}(\phi^*) &= Q(t)\mathbf{u}(\phi), \\ T(\phi^*) &= Q(t)T(\phi)Q(t)^T, \end{aligned}$$

where $Q(t)$ is the orthogonal part of the change of frame from ϕ to ϕ^* , then s , \mathbf{u} and T are called objective scalar, vector and tensor quantities respectively.

More precisely, they are also said to be frame-indifferent with respect to Euclidean transformations or simply Euclidean objective. For simplicity, we often write $f = f(\phi)$ and $f^* = f(\phi^*)$.

One can easily deduce the transformation properties of functions defined on the position and time under a change of frame. Consider an objective scalar field $\psi(\mathbf{x}, t) =$

$\psi^*(\mathbf{x}^*, t^*)$. Taking the gradient with respect to \mathbf{x} , from (4.2) we obtain

$$\nabla_{\mathbf{x}}\psi(\mathbf{x}, t) = Q(t)^T \nabla_{\mathbf{x}^*}\psi^*(\mathbf{x}^*, t^*) \quad \text{or} \quad (\text{grad } \psi)(\phi^*) = Q(t) (\text{grad } \psi)(\phi),$$

which proves that $(\text{grad } \psi)$ is an objective vector field. Similarly, we can show that if \mathbf{u} is an objective vector field then $(\text{grad } \mathbf{u})$ is an objective tensor field and $(\text{div } \mathbf{u})$ is an objective scalar field. However, one can easily show that the partial derivative $\partial\psi/\partial t$ is not an objective scalar field and neither is $\partial\mathbf{u}/\partial t$ an objective vector field.

4.3 Transformation properties of motion

Let χ_ϕ be a motion of the body in the frame ϕ , and χ_{ϕ^*} be the corresponding motion in ϕ^* ,

$$\mathbf{x} = \chi_\phi(p, t), \quad \mathbf{x}^* = \chi_{\phi^*}(p, t^*), \quad p \in \mathcal{B}.$$

Then from (4.2), we have

$$\chi_{\phi^*}(p, t^*) = Q(t)(\chi_\phi(p, t) - \mathbf{x}_o) + \mathbf{c}(t), \quad p \in \mathcal{B},$$

from which, one can easily show that the velocity and the acceleration are not objective quantities,

$$\begin{aligned} \dot{\mathbf{x}}^* &= Q\dot{\mathbf{x}} + \dot{Q}(\mathbf{x} - \mathbf{x}_o) + \dot{\mathbf{c}}, \\ \ddot{\mathbf{x}}^* &= Q\ddot{\mathbf{x}} + 2\dot{Q}\dot{\mathbf{x}} + \ddot{Q}(\mathbf{x} - \mathbf{x}_o) + \ddot{\mathbf{c}}. \end{aligned} \tag{4.3}$$

A change of frame (4.1) with constant $Q(t)$ and $\mathbf{c}(t) = \mathbf{c}_0 + \mathbf{c}_1 t$, for constant \mathbf{c}_0 and \mathbf{c}_1 , is called a *Galilean transformation*. Therefore, from (4.3) we conclude that the acceleration is not Euclidean objective but it is frame-indifferent with respect to Galilean transformation. Moreover, it also shows that the velocity is neither a Euclidean nor a Galilean objective vector quantity.

Transformation properties of deformation gradient

Let $\kappa : \mathcal{B} \rightarrow \mathcal{W}_{t_0}$ be a reference placement of the body at some instant t_0 (see Figure 6), then

$$\kappa_\phi = \phi_{t_0} \circ \kappa \quad \text{and} \quad \kappa_{\phi^*} = \phi_{t_0}^* \circ \kappa \tag{4.4}$$

are the corresponding reference configurations of \mathcal{B} in the frames ϕ and ϕ^* at the same instant, and

$$\mathbf{X} = \kappa_\phi(p), \quad \mathbf{X}^* = \kappa_{\phi^*}(p), \quad p \in \mathcal{B}.$$

Let us denote by $\gamma = \kappa_{\phi^*} \circ \kappa_\phi^{-1}$ the change of reference configuration from κ_ϕ to κ_{ϕ^*} in the change of frame, then it follows from (4.4) that $\gamma = \phi_{t_0}^* \circ \phi_{t_0}^{-1}$ and by (4.2), we have

$$\mathbf{X}^* = \gamma(\mathbf{X}) = Q(t_0)(\mathbf{X} - \mathbf{x}_o) + \mathbf{c}(t_0). \tag{4.5}$$

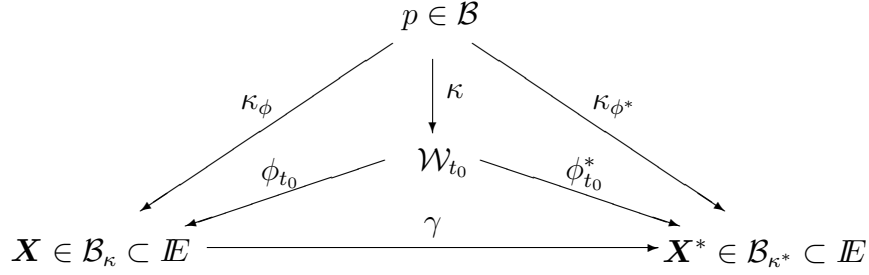


Figure 6: Reference configurations κ_ϕ and κ_{ϕ^*} in the change of frame from $\phi \rightarrow \phi^*$

On the other hand, the motion in referential description relative to the change of frame is given by $\mathbf{x} = \chi_\kappa(\mathbf{X}, t)$ and $\mathbf{x}^* = \chi_{\kappa^*}(\mathbf{X}^*, t^*)$. Hence from (4.2), we have

$$\chi_{\kappa^*}(\mathbf{X}^*, t^*) = Q(t)(\chi_\kappa(\mathbf{X}, t) - \mathbf{x}_o) + \mathbf{c}(t). \quad (4.6)$$

Therefore we obtain for the deformation gradient in the frame ϕ^* , i.e., $F^* = \nabla_{\mathbf{X}^*} \chi_{\kappa^*}$, by taking the gradient with respect to \mathbf{X} and the use of the chain rule and (4.5),

$$F^*(\mathbf{X}^*, t^*)Q(t_0) = Q(t)F(\mathbf{X}, t), \quad \text{or simply} \quad F^* = QFK^T, \quad (4.7)$$

where $K = Q(t_0)$ is a constant orthogonal tensor due to the change of frame for the reference configuration.

Remark. The transformation property (4.7) stands in contrast to $F^* = QF$, the widely used formula which is obtained “provided that the reference configuration is unaffected by the change of frame” as usually implicitly assumed, so that K reduces to the identity transformation. \square

The deformation gradient F is not a Euclidean objective tensor. However, the property (4.7) also shows that it is frame-indifferent with respect to Galilean transformations, since in this case, $K = Q$ is a constant orthogonal transformation.

From (4.7), we can easily obtain the transformation properties of other kinematic quantities associated with the deformation gradient. In particular, let us consider the velocity gradient defined in (2.24). We have

$$L^* = \dot{F}^*(F^*)^{-1} = (Q\dot{F} + \dot{Q}F)K^T(QFK^T)^{-1} = (Q\dot{F} + \dot{Q}F)F^{-1}Q^T,$$

which gives

$$L^* = QLQ^T + \dot{Q}Q^T. \quad (4.8)$$

Moreover, with the decomposition $L = D + W$ into symmetric and skew-symmetric parts, it becomes

$$D^* + W^* = Q(D + W)Q^T + \dot{Q}Q^T.$$

By separating symmetric and skew-symmetric parts, we obtain

$$D^* = QDQ^T, \quad W^* = QWQ^T + \dot{Q}Q^T,$$

since $\dot{Q}Q^T$ is skew-symmetric. Therefore, while the velocity gradient L and the *spin tensor* W are not objective, the *rate of strain tensor* D is an objective tensor.

4.4 Inertial frames

In classical mechanics, Newton's first law, often known as the law of inertia, is essentially a definition of inertial frame.

Definition. (*Inertial frame*): A frame of reference is called an inertial frame if, relative to it, the velocity of a body remains constant unless the body is acted upon by an external force.

We present the first law in this manner in order to emphasize that the existence of inertial frames is essential for the formulation of Newton's second law, which asserts that relative to an inertial frame, the equation of motion takes the simple form:

$$\mathbf{m} \ddot{\mathbf{x}} = \mathbf{f}. \quad (4.9)$$

Now, we shall assume that there is an inertial frame $\phi_0 \in \mathfrak{E}$, for which the equation of motion of a particle is given by (4.9), and we are interested in how the equation is transformed under a change of frame.

Unlike the acceleration, transformation properties of non-kinematic quantities cannot be deduced theoretically. Instead, for the mass and the force, it is conventionally *postulated* that they are Euclidean objective scalar and vector quantities respectively, so that for any change from ϕ_0 to $\phi^* \in \mathfrak{E}$ given by (4.1), we have

$$\mathbf{m}^* = \mathbf{m}, \quad \mathbf{f}^* = Q \mathbf{f},$$

which together with (4.3), by multiplying (4.9) with Q , we obtain the equation of motion in the (non-inertial) frame ϕ^* ,

$$\mathbf{m}^* \ddot{\mathbf{x}}^* = \mathbf{f}^* + \mathbf{m}^* \mathbf{i}^*, \quad (4.10)$$

where \mathbf{i}^* is called the inertial force given by

$$\mathbf{i}^* = \ddot{\mathbf{c}} + 2\Omega(\dot{\mathbf{x}}^* - \dot{\mathbf{c}}) + (\dot{\Omega} - \Omega^2)(\mathbf{x}^* - \mathbf{c}),$$

where $\Omega = \dot{Q}Q^T : \mathbb{R} \rightarrow \mathcal{L}(V)$ is called the spin tensor of the frame ϕ^* relative to the inertial frame ϕ_0 .

Note that the inertial force vanishes if the change of frame $\phi_0 \rightarrow \phi^*$ is a Galilean transformation, i.e., $\dot{Q} = 0$ and $\ddot{\mathbf{c}} = 0$, and hence the equation of motion in the frame ϕ^* also takes the simple form,

$$m^* \ddot{\mathbf{x}}^* = \mathbf{f}^*,$$

which implies that the frame ϕ^* is also an inertial frame.

Therefore, any frame of reference obtained by a Galilean change of frame from an inertial frame is also an inertial frame and thus, all inertial frames form an equivalent class \mathfrak{G} , such that for any $\phi, \phi^* \in \mathfrak{G}$, the change of frame $\phi \rightarrow \phi^*$ is a Galilean transformation. The Galilean class \mathfrak{G} is a subclass of the Euclidean class \mathfrak{E} .

Remark. Since Euclidean change of frame is an equivalence relation, it decomposes all frames of reference into disjoint equivalence classes, i.e., Euclidean classes as we previously called. However, the existence of an inertial frame which is essential in establishing dynamic laws in mechanics, leads to a special choice of Euclidean class of interest.

Let \mathfrak{E} be the Euclidean class which contains an inertial frame. Since different Euclidean classes are not related by any Euclidean transformation, hence, nor by any Galilean transformation, it is obvious that the Euclidean class \mathfrak{E} is the only class containing the subclass \mathfrak{G} of all inertial frames. Consequently, from now on, the only Euclidean class of interest for further discussions, is the one, denoted by \mathfrak{E} , containing Galilean class of all inertial frames. \square

In short, we can assert that physical laws, like the equation of motion, are in general not (Euclidean) frame-indifferent. Nevertheless, the equation of motion is Galilean frame-indifferent, under the assumption that mass and force are frame-indifferent quantities. This is usually referred to as *Galilean invariance* of the equation of motion.

4.5 Galilean invariance of balance laws

In Section 3, the balance laws of mass, linear momentum, and energy for deformable bodies,

$$\begin{aligned}\dot{\rho} + \rho \operatorname{div} \dot{\mathbf{x}} &= 0, \\ \rho \ddot{\mathbf{x}} - \operatorname{div} T &= \rho \mathbf{b}, \\ \rho \dot{\varepsilon} + \operatorname{div} \mathbf{q} - T \cdot \operatorname{grad} \dot{\mathbf{x}} &= \rho r,\end{aligned}\tag{4.11}$$

are tacitly formulated relative to an inertial frame. Consequently, motivated by classical mechanics, they are required to be invariant under Galilean transformation.

Since two inertial frames are related by a Galilean transformation, it means that the equations (4.11) should hold in the same form in any inertial frame. In particular, the balance of linear momentum takes the forms in the inertial frames $\phi, \phi^* \in \mathfrak{G}$,

$$\rho \ddot{\mathbf{x}} - \operatorname{div} T = \rho \mathbf{b}, \quad \rho^* \ddot{\mathbf{x}}^* - (\operatorname{div} T)^* = \rho^* \mathbf{b}^*.$$

Since the acceleration $\ddot{\mathbf{x}}$ is Galilean objective, in order this to hold, it is usually assumed that the mass density ρ , the Cauchy stress tensor T and the body force \mathbf{b} are objective scalar, tensor, and vector quantities respectively. Similarly, for the energy equation, it is also assumed that the internal energy ε and the energy supply r are objective scalars, and the heat flux \mathbf{q} is an objective vector. These assumptions concern the non-kinematic quantities, including external supplies (\mathbf{b}, r) , and the constitutive quantities $(T, \mathbf{q}, \varepsilon)$.

In fact, for Galilean invariance of the balance laws, only frame-indifference with respect to Galilean transformation for all those non-kinematic quantities would be sufficient. However, similar to classical mechanics, it is *postulated* that they are not only Galilean objective but also Euclidean objective. Therefore, with the known transformation properties of the kinematic variables, the balance laws in any arbitrary frame can be deduced.

To emphasize the importance of the objectivity postulate for constitutive theories, it will be referred to as Euclidean objectivity for constitutive quantities:

Euclidean objectivity. *The constitutive quantities: the Cauchy stress T , the heat flux \mathbf{q} and the internal energy density ε , are Euclidean objective (Euclidean frame-indifferent),*

$$T(\phi^*) = Q(t) T(\phi) Q(t)^T, \quad \mathbf{q}(\phi^*) = Q(t) \mathbf{q}(\phi), \quad \varepsilon(\phi^*) = \varepsilon(\phi), \quad (4.12)$$

where $Q(t) \in \mathcal{O}(V)$ is the orthogonal part of the change of frame from ϕ to ϕ^* .

Note that this postulate concerns only frame-indifference properties of balance laws, so that it is a universal property for any deformable bodies, and therefore, do not concern any aspects of material properties of the body.

5 Principle of material frame-indifference

Properties of material bodies are described mathematically by constitutive equations. Classical models, such as Hooke's law of elastic solids, Navier-Stokes law of viscous fluids, and Fourier law of heat conduction, are mostly proposed based on physical experiences and experimental observations. However, even these linear experimental laws did not come without some understanding of theoretical concepts of material behavior. Without it one would neither know what experiments to run nor be able to interpret their results.

For a general and rational formulation of constitutive theories, besides from physical experiences, one should rely on some basic requirements that a mathematical model should obey lest its consequences be contradictory to physical nature. The most fundamental ones are

- principle of material frame-indifference,
- material symmetry,
- second law of thermodynamics.

These requirements impose severe restrictions on material models and hence lead to great simplifications for general constitutive equations. From theoretical viewpoints, the aim of constitutive theories in continuum mechanics is to construct material models consistent with such universal requirements so as to enable us, by formulating and analyzing mathematical problems, to predict the outcomes in material behavior verifiable by experimental observations.

5.1 Constitutive equations in material description

Physically a state of the thermomechanical behavior of a body is characterized by a description of the fields of density $\rho(p, t)$, motion $\chi(p, t)$ and temperature $\theta(p, t)$. The material properties of a body generally depend on the past history of its thermomechanical behavior.

Let us introduce the notion of the past history of a function. Let $h(\bullet, t)$ be a function of time defined on a set \mathcal{X} in some space \mathbb{W} , $h : \mathcal{X} \times \mathbb{R} \rightarrow \mathbb{W}$. The history of h up to time t is defined by

$$h^t(\bullet, s) = h(\bullet, t - s),$$

where $s \in [0, \infty)$ denotes the time-coordinate pointed into the past from the present time t . Clearly $s = 0$ corresponds to the present time, therefore $h^t(\bullet, 0) = h(\bullet, t)$.

Let the set of history functions on a set \mathcal{X} in some space \mathbb{W} be denoted by

$$\mathfrak{H}(\mathcal{X}, \mathbb{W}) = \{h^t : \mathcal{X} \times [0, \infty) \rightarrow \mathbb{W}\}.$$

Mathematical descriptions of material properties are called constitutive equations. We postulate that the history of thermomechanical behavior up to the present time determines the properties of the material body.

Principle of determinism. *Let ϕ be a frame of reference, and \mathcal{C} be a constitutive quantity, then the constitutive equation for \mathcal{C} is given by a functional of the form,*

$$\mathcal{C}(\phi, p, t) = \mathcal{F}_\phi(\rho^t, \chi^t, \theta^t; p), \quad p \in \mathcal{B}, t \in \mathbb{R}, \quad (5.1)$$

where the first three arguments are history functions:

$$\rho^t \in \mathfrak{H}(\mathcal{B}, \mathbb{R}), \quad \chi^t \in \mathfrak{H}(\mathcal{B}, \mathbb{E}), \quad \theta^t \in \mathfrak{H}(\mathcal{B}, \mathbb{R}).$$

We call \mathcal{F}_ϕ the constitutive function of \mathcal{C} in the frame ϕ . Such a functional allows the description of arbitrary non-local effect of an inhomogeneous body with a perfect memory of the past thermomechanical history. With the notation \mathcal{F}_ϕ , we emphasize that the value of a constitutive function may depend on the frame of reference ϕ .

For simplicity, for further discussions on constitutive equations, we shall restrict our attention to material models for mechanical theory only, and only constitutive equations for the stress tensor, $T(\phi, p, t) \in V \otimes V$, will be considered. It can be written as

$$T(\phi, p, t) = \mathcal{F}_\phi(\chi^t; p), \quad \phi \in \mathfrak{E}, \quad p \in \mathcal{B}, \quad \chi^t \in \mathfrak{H}(\mathcal{B}, \mathbb{E}). \quad (5.2)$$

Let $\phi^* \in \mathfrak{E}$ be another observer, then the constitutive equation, $T(\phi^*, p, t^*) \in V \otimes V$, can be written as

$$T(\phi^*, p, t^*) = \mathcal{F}_{\phi^*}((\chi^t)^*; p), \quad p \in \mathcal{B}, \quad (\chi^t)^* \in \mathfrak{H}(\mathcal{B}, \mathbb{E}), \quad (5.3)$$

where the corresponding histories of motion are related by (4.2),

$$(\chi^t)^*(\bar{p}, s) = *(\chi^t(\bar{p}, s)) = Q^t(s)(\chi^t(\bar{p}, s) - \mathbf{x}_o) + \mathbf{c}^t(s),$$

for any $s \in [0, \infty)$ and any $\bar{p} \in \mathcal{B}$ in the change of frame $\phi \rightarrow \phi^*$.

Condition of Euclidean objectivity

We need to bear in mind that according to the assumption referred to as the Euclidean objectivity (4.12), the stress is a frame-indifferent quantity under a change of observer,

$$T(\phi^*, p, t^*) = Q(t)T(\phi, p, t)Q(t)^T.$$

Therefore, it follows immediately that

$$\mathcal{F}_{\phi^*}(*(\chi^t); p) = Q(t)\mathcal{F}_\phi(\chi^t; p)Q(t)^T, \quad (5.4)$$

where $Q(t) \in \mathcal{O}(V)$ is the orthogonal part of the change of frame $\phi \rightarrow \phi^*$.

The relation (5.4) will be referred to as the *condition of Euclidean objectivity*. It is a relation between the constitutive functions relative to two different observers. In other words, different observers cannot independently propose their own constitutive equations. Instead, the condition of Euclidean objectivity (5.4) determines the constitutive function \mathcal{F}_{ϕ^*} once the constitutive function \mathcal{F}_{ϕ} is given or vice-versa. They determine one from the other in a frame-dependent manner because they are bounded by the consensus requirement.

5.2 Principle of material frame-indifference

It is obvious that *not* any proposed constitutive equations can be used as material models. First of all, they may be frame-dependent in general. However, since the constitutive functions must characterize the intrinsic properties of the material body itself, it should be independent of observer. Consequently, there must be some restrictions imposed on the constitutive functions so that they would be indifferent to the change of frame. This is the essential idea of the principle of material frame-indifference.

Principle of material frame-indifference (in material description). *The constitutive function of an objective constitutive quantity must be independent of frame, i.e., for any frames of reference ϕ and ϕ^* , the functionals \mathcal{F}_{ϕ} and \mathcal{F}_{ϕ^*} , defined by (5.2) and (5.3), must have the same form,*

$$\mathcal{F}_{\phi}(\bullet; p) = \mathcal{F}_{\phi^*}(\bullet; p), \quad p \in \mathcal{B}, \quad \bullet \in \mathfrak{H}(\mathcal{B}, \mathbb{E}). \quad (5.5)$$

where \bullet represents the same arguments in both functionals.

Thus, from the condition of Euclidean objectivity (5.4) and the principle of material frame-indifference (5.5), we obtain the following condition:

Condition of material objectivity. *The constitutive function of the stress tensor, in material description, satisfies the condition,*

$$\mathcal{F}_{\phi}(*(\chi^t); p) = Q(t) \mathcal{F}_{\phi}(\chi^t; p) Q(t)^T, \quad p \in \mathcal{B}, \quad \chi^t \in \mathfrak{H}(\mathcal{B}, \mathbb{E}), \quad (5.6)$$

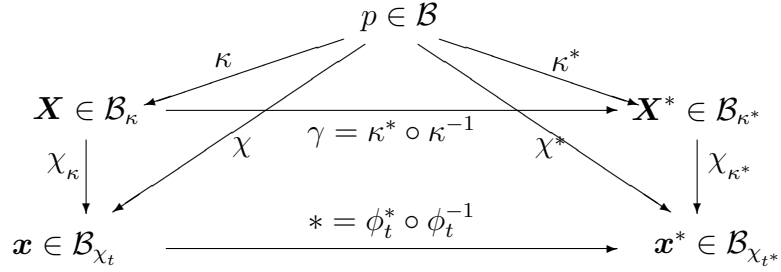
where $Q(t) \in \mathcal{O}(V)$ is the orthogonal part of an arbitrary change of frame $*$.

Since the condition (5.6) involves only the constitutive function in the frame ϕ , it becomes a restriction imposed on the constitutive function \mathcal{F}_{ϕ} . We call it *material objectivity* to emphasize its observer-indifference concerning material properties.

5.3 Constitutive equations in referential description

For mathematical analysis, it is more convenient to use referential description so that motions can be defined on the Euclidean space \mathbb{E} instead of the set of material points

in \mathcal{B} . Therefore, for further discussions, we need to reinterpret the principle of material frame-indifference for constitutive equations relative to a reference configuration.



Let $\kappa : \mathcal{B} \rightarrow \mathbb{E}$ and $\kappa^* : \mathcal{B} \rightarrow \mathbb{E}$ be the two corresponding reference configurations of \mathcal{B} in the frames ϕ and ϕ^* at the same instant t_0 , and

$$\mathbf{X} = \kappa(p) \in \mathbb{E}, \quad \mathbf{X}^* = \kappa^*(p) \in \mathbb{E}, \quad p \in \mathcal{B}.$$

Let us denote by $\gamma = \kappa^* \circ \kappa^{-1}$ the change of reference configuration from κ to κ^* in the change of frame, then from (4.5) we have

$$\mathbf{X}^* = \gamma(\mathbf{X}) = K(\mathbf{X} - \mathbf{x}_o) + \mathbf{c}(t_0), \quad (5.7)$$

where $K = \nabla_{\mathbf{X}} \gamma = Q(t_0)$ is a constant orthogonal tensor.

The motion in referential description relative to the change of frame is given by

$$\begin{aligned}
\mathbf{x} = \chi(p, t) &= \chi(\kappa^{-1}(\mathbf{X}), t) = \chi_\kappa(\mathbf{X}, t), & \chi &= \chi_\kappa \circ \kappa, \\
\mathbf{x}^* = \chi^*(p, t^*) &= \chi^*(\kappa^{*-1}(\mathbf{X}^*), t^*) = \chi_{\kappa^*}(\mathbf{X}^*, t^*), & \chi^* &= \chi_{\kappa^*} \circ \kappa^*.
\end{aligned}$$

From (5.2) and (5.3), we can define the corresponding constitutive functions with respect to the reference configuration,

$$\begin{aligned}
\mathcal{T}_\phi(\chi^t; p) &= \mathcal{T}_\phi(\chi_\kappa^t \circ \kappa; p) := \mathcal{H}_\kappa(\chi_\kappa^t; \mathbf{X}), & \chi_\kappa^t &\in \mathfrak{H}(\mathcal{B}_\kappa, \mathbb{E}), \\
\mathcal{T}_{\phi^*}((\chi^t)^*; p) &= \mathcal{T}_{\phi^*}((\chi_\kappa^t)^* \circ \kappa^*; p) := \mathcal{H}_{\kappa^*}((\chi_\kappa^t)^*; \mathbf{X}^*), & (\chi_\kappa^t)^* &\in \mathfrak{H}(\mathcal{B}_{\kappa^*}, \mathbb{E}).
\end{aligned}$$

From the above definitions, we can obtain the relation between the constitutive functions \mathcal{H}_κ and \mathcal{H}_{κ^*} in the referential description,

$$\begin{aligned}
\mathcal{H}_{\kappa^*}((\chi^t)^*; \mathbf{X}^*) &= \mathcal{T}_{\phi^*}((\chi_\kappa^t)^* \circ \kappa^*; p) = \mathcal{T}_\phi((\chi_\kappa^t)^* \circ \kappa^*; p) \\
&= \mathcal{T}_\phi((\chi_\kappa^t)^* \circ \gamma \circ \kappa; p) = \mathcal{H}_\kappa((\chi^t)^* \circ \gamma; \mathbf{X}),
\end{aligned}$$

which begins with the definition and in the second passage the principle of material frame invariance (5.5), $\mathcal{T}_\phi = \mathcal{T}_{\phi^*}$ has been used, and then κ^* is replaced by $\gamma \circ \kappa$, and finally the definition again.

Unlike $\mathcal{T}_\phi = \mathcal{T}_{\phi^*}$ with the same domain $\mathfrak{H}(\mathcal{B}, \mathbb{E}) \times \mathcal{B}$ in the material description, the constitutive functions \mathcal{H}_κ and \mathcal{H}_{κ^*} have different domains, namely, $\mathfrak{H}(\mathcal{B}_\kappa, \mathbb{E}) \times \mathcal{B}_\kappa$ and

$\mathfrak{H}(\mathcal{B}_{\kappa^*}, \mathbb{E}) \times \mathcal{B}_{\kappa^*}$ in referential description. Therefore, $\mathcal{H}_{\kappa^*} \neq \mathcal{H}_{\kappa}$, but rather they are related by

$$\mathcal{H}_{\kappa^*}(\bullet; \mathbf{X}^*) = \mathcal{H}_{\kappa}(\bullet \circ \gamma; \mathbf{X}), \quad \bullet \in \mathfrak{H}(\mathcal{B}_{\kappa^*}, \mathbb{E}), \quad (5.8)$$

where $\gamma = \kappa^* \circ \kappa^{-1}$ is the change of reference configuration from κ to κ^* in the change of frame $*$ and $\mathbf{X} = \gamma^{-1}(\mathbf{X}^*)$. This is the reinterpretation of the principle of material frame-indifference, for which in expressing observer independence of material properties, one must also take into account the domains of constitutive functions affected by the change of frame on the reference configuration.

The Euclidean objectivity relation (5.4) in referential description can be written in the form,

$$\mathcal{H}_{\kappa^*}(*(\chi_{\kappa}^t); \mathbf{X}^*) = Q(t) \mathcal{H}_{\kappa}(\chi_{\kappa}^t; \mathbf{X}) Q(t)^T, \quad (5.9)$$

where $Q(t)$ is the orthogonal part of the change of frame $*$.

Finally, by combining (5.8) and (5.9), we obtain the condition of material objectivity in referential description,

$$\mathcal{H}_{\kappa}(*(\chi_{\kappa}^t) \circ \gamma; \mathbf{X}) = Q(t) \mathcal{H}_{\kappa}(\chi_{\kappa}^t; \mathbf{X}) Q(t)^T, \quad (5.10)$$

valid for any $\chi_{\kappa}^t \in \mathfrak{H}(\mathcal{B}_{\kappa}, \mathbb{E})$ and for any change of frame $*$. In particular, it is valid for any $Q(t) \in \mathcal{O}(V)$. Note that from (4.6)

$$\begin{aligned} *(\chi_{\kappa}) \circ \gamma(\mathbf{X}, t) &= \chi_{\kappa^*}(\gamma(\mathbf{X}), t^*) = \chi_{\kappa^*}(\mathbf{X}^*, t^*) \\ &= Q(t)(\chi_{\kappa}(\mathbf{X}, t) - \mathbf{x}_0) + c(t), \end{aligned}$$

the condition of material objectivity (5.10) becomes

$$\mathcal{H}_{\kappa}(Q^t(\chi_{\kappa}^t - \mathbf{x}_0) + c^t; \mathbf{X}) = Q(t) \mathcal{H}_{\kappa}(\chi_{\kappa}^t; \mathbf{X}) Q(t)^T, \quad (5.11)$$

for any $Q(t) \in \mathcal{O}(V)$, $\mathbf{x}_0, c(t) \in \mathbb{E}$ and for any $\chi_{\kappa}^t \in \mathfrak{H}(\mathcal{B}_{\kappa}, \mathbb{E})$.

Note that since the conditions (5.8) and (5.9) are valid for any change of frame $\phi \rightarrow \phi^*$, while the condition (5.11) concerns only the frame ϕ , the star frame ϕ^* can then be chosen arbitrarily. Therefore, unlike the conditions (5.8) and (5.9) concerning two frames in the change, the condition (5.11) is valid *for any choice* of $Q(t) \in \mathcal{O}(V)$, $\mathbf{x}_0, c(t) \in \mathbb{E}$, which define an arbitrary change of frame. In other words, the condition (5.11) becomes a restriction on the constitutive function \mathcal{H}_{κ} so that it is valid for any such choices.

5.4 Simple materials

According to the principle of determinism (5.1), thermomechanical histories of any part of the body can affect the response at any point of the body. In most applications, such a non-local property is irrelevant. Therefore it is usually assumed that only thermomechanical histories in an arbitrary small neighborhood of \mathbf{X} affects the material response at the point

\mathbf{X} , and hence if only linear approximation by Taylor expansion in a small neighborhood of \mathbf{X} is concerned, we have

$$\chi_\kappa(\mathbf{Y}, t) = \chi_\kappa(\mathbf{X}, t) + F(\mathbf{X}, t)(\mathbf{Y} - \mathbf{X}) + o(2),$$

and the constitutive equation can be written as

$$\mathcal{T}(\mathbf{X}, t) = \mathcal{H}_\kappa(\chi_\kappa^t(\mathbf{X}), F^t(\mathbf{X}); \mathbf{X}), \quad \mathbf{X} \in \mathcal{B}_\kappa.$$

An immediate consequence of the condition of material objectivity (5.11) can be obtained by the following choice of change of frame such that

$$Q(t) = I, \quad c(t) = c_0,$$

the condition (5.11) implies that

$$\mathcal{H}_\kappa(\chi_\kappa^t + \mathbf{c}_0 - \mathbf{x}_0, F^t; \mathbf{X}) = \mathcal{H}_\kappa(\chi_\kappa^t, F^t; \mathbf{X}).$$

Since $(\mathbf{c}_0 - \mathbf{x}_0) \in V$ is arbitrary, we conclude that \mathcal{H}_κ can not depend on the history of position $\chi_\kappa^t(\mathbf{X}, s)$.

Therefore the constitutive equation can be written as

$$\mathcal{T}(\mathbf{X}, t) = \mathcal{H}_\kappa(F^t(\mathbf{X}); \mathbf{X}), \quad \mathbf{X} \in \mathcal{B}_\kappa, \quad F^t \in \mathfrak{H}(\{\mathbf{X}\}, \mathcal{L}(V)). \quad (5.12)$$

where $F^t = \nabla_{\mathbf{X}} \chi_\kappa^t$ is the deformation gradient and the domain of the history is a single point $\{\mathbf{X}\}$. In other words, the constitutive function depends only on local values at the position \mathbf{X} .

A material with constitutive equation (5.12) is called a *simple material* (due to Noll). The class of simple materials is general enough to include most of the materials of practical interests, such as: elastic solids, viscoelastic solids, as well as elastic fluids, Navier-Stokes fluids and non-Newtonian fluids.

For simple materials, by the use of (5.7), the consequence of the principle of material frame indifference (5.8) takes the form,

$$\mathcal{H}_{\kappa^*}((F^t)^*; \mathbf{X}^*) = \mathcal{H}_\kappa((F^t)^* K; \mathbf{X}),$$

and the Euclidean objectivity condition (5.9) becomes

$$\mathcal{H}_{\kappa^*}((F^t)^*; \mathbf{X}^*) = Q(t) \mathcal{H}_\kappa(F^t; \mathbf{X}) Q(t)^T.$$

Combining the above two conditions and knowing the relation, by the use of (4.7),

$$(F^t)^* K = (Q^t F^t K^T) K = Q^t F^t,$$

we obtain the following

Condition of material objectivity. *Constitutive equation of a simple material must satisfy*

$$\mathcal{H}_\kappa(Q^t F^t; \mathbf{X}) = Q(t) \mathcal{H}_\kappa(F^t; \mathbf{X}) Q(t)^T, \quad (5.13)$$

for any orthogonal transformation $Q(t) \in \mathcal{O}(V)$ and any local history of deformation gradient $F^t \in \mathfrak{H}(\{\mathbf{X}\}, \mathcal{L}(V))$.

Obviously we can also deduce the above condition from (5.11) directly for a simple material.

Remarks. The condition (5.13) is the well-known condition of material objectivity, obtained with the assumption that “reference configuration be unaffected by the change of frame” in the fundamental treatise, *The Non-Linear Field Theories of Mechanics* by Truesdell and Noll (1965). This condition remains valid without such an assumption.

Note that in condition (5.13), no mention of change of frame is involved, and $Q(t)$ can be interpreted as a superimposed orthogonal transformation on the deformation. This interpretation is sometimes viewed as an alternative version of the principle of material objectivity and is called the “principle of invariance under superimposed rigid body motions”. \square

6 Material symmetry

We shall consider homogeneous simple material bodies from now on for simplicity. A body is called *homogeneous* in the configuration κ if the constitutive function does not depend on the argument \mathbf{X} explicitly,

$$T(\mathbf{X}, t) = \mathcal{H}_\kappa(F^t(\mathbf{X})).$$

As indicated, constitutive functions may depend on the reference configuration. Now suppose that $\hat{\kappa}$ is another reference configuration, so that the motion can be written as

$$\mathbf{x} = \chi_\kappa(\mathbf{X}, t) = \chi_{\hat{\kappa}}(\widehat{\mathbf{X}}, t) \quad \text{and} \quad \widehat{\mathbf{X}} = \xi(\mathbf{X}).$$

Let $G = \nabla_{\mathbf{X}}\xi \in \mathcal{L}(V)$, then

$$\nabla_{\mathbf{X}}\chi_\kappa = (\nabla_{\widehat{\mathbf{X}}}\chi_{\hat{\kappa}})(\nabla_{\mathbf{X}}\xi) \quad \text{or} \quad F = \widehat{F}G.$$

Therefore, from the function \mathcal{H}_κ , the constitutive function $\mathcal{H}_{\hat{\kappa}}$ relative to the configuration $\hat{\kappa}$ can be defined as

$$\mathcal{H}_\kappa(F^t) = \mathcal{H}_{\hat{\kappa}}(\widehat{F}^t G) := \mathcal{H}_{\hat{\kappa}}(\widehat{F}^t). \quad (6.14)$$

The two functions \mathcal{H}_κ and $\mathcal{H}_{\hat{\kappa}}$ are in general different. Consequently, a material body subjected to the same experiment (i.e., the same mechanical histories) at two different configurations may have different results.

6.1 Material symmetry group

However, a material body may possess a certain symmetry so that one can not distinguish the outcomes of the same experiments performed at two different configurations. For example, a material body with a cubic crystal structure before and after a rotation of 90° about one of its crystallographic axes is physically indistinguishable.

Definition. *Two reference configurations κ and $\hat{\kappa}$ are said to be materially indistinguishable if their corresponding constitutive functions are the same,*

$$\mathcal{H}_\kappa(\bullet) = \mathcal{H}_{\hat{\kappa}}(\bullet).$$

By the second relation of (6.14), the above condition is equivalent to

$$\mathcal{H}_\kappa(F^t) = \mathcal{H}_\kappa(F^t G), \quad \forall F^t. \quad (6.15)$$

We call a transformation $G \in \mathcal{L}(V)$ which satisfies (6.15) a material symmetry transformation with respect to κ .

We assume that a material symmetry transformation is volume-preserving, since, otherwise, if G is a material symmetry transformation, so is G^n for any $n = 1, 2, \dots$, and since $|\det G^n| = |\det G|^n$, the material could suffer arbitrarily large change of volume with no change in material response – a conclusion that seems physically unacceptable. Therefore, we must require that $G \in \mathcal{U}(V)$, where $\mathcal{U}(V) = \{G \in \mathcal{L}(V) : |\det G| = 1\}$ is called the unimodular group on the vector space V .

It is easy to verify that the set of all material symmetry transformations

$$\mathcal{G}_\kappa = \{G \in \mathcal{U}(V) : \mathcal{H}_\kappa(F^t) = \mathcal{H}_\kappa(F^t G), \forall F^t\}$$

is a subgroup of the unimodular group. We call \mathcal{G}_κ the *material symmetry group* of the material body in the reference configuration κ .

Condition of material symmetry. *Constitutive function of a simple material must satisfy*

$$\mathcal{H}_\kappa(F^t G) = \mathcal{H}_\kappa(F^t), \quad \forall G \in \mathcal{G}_\kappa, \quad \forall F^t. \quad (6.16)$$

Like the condition of material objectivity (5.13), the condition of material symmetry is also a restriction imposed on the constitutive function \mathcal{H}_κ .

6.2 Classification of material bodies

Physical concepts of real materials such as solids and fluids, can be characterized by their symmetry properties. One of such concepts can be interpreted as saying that a solid has a preferred configuration such that any non-rigid deformation from it alters its material response, while for a fluid any deformation that preserves the density should not affect the material response. The following definitions are based on this concept.

Definition. *A simple material body is called a solid body if the symmetry group \mathcal{G}_κ is a subgroup of the orthogonal group, $\mathcal{G}_\kappa \subseteq \mathcal{O}(V)$.*

Definition. *A simple material is called a fluid if the symmetry group is the unimodular group, $\mathcal{G}_\kappa = \mathcal{U}(V)$.*

A simple material which is neither a fluid nor a solid will be called a *fluid crystal*. Another concept concerning material symmetry is the material response due to change of orientation.

Definition. *A simple material body is called isotropic if there exists a configuration κ , such that the symmetry group contains the orthogonal group, $\mathcal{G}_\kappa \supseteq \mathcal{O}(V)$.*

Physically, we can interpret the above definition as saying that any rotation does not alter material response of an isotropic material. The following theorem characterizes isotropic materials (for the proof, see Noll (1965)).

Theorem. *The orthogonal group is maximal in the unimodular group, i.e., if \mathcal{G} is a group such that*

$$\mathcal{O}(V) \subseteq \mathcal{G} \subseteq \mathcal{U}(V),$$

then either $\mathcal{G} = \mathcal{O}(V)$ or $\mathcal{G} = \mathcal{U}(V)$.

Therefore, an isotropic material is either a fluid, $\mathcal{G}_\kappa = \mathcal{U}(V)$ for any κ , or an isotropic solid at some configuration κ , $\mathcal{G}_\kappa = \mathcal{O}(V)$. Any other materials are *anisotropic*. Transversely isotropic solids, crystalline solids and fluid crystals are all anisotropic materials. A solid is isotropic at some configuration may not be isotropic at other configurations.

6.3 Summary on constitutive models of simple materials

So far, we have derived the conditions of material objectivity and material symmetry for the stress tensor only for simplicity. For other vector or scalar constitutive quantities, similar results can be easily obtained. We shall summarize the results for simple materials as follows:

Let \mathcal{G} be the symmetry group of the material and the constitutive equations for the stress T , the heat flux vector \mathbf{q} , and the internal energy density ε be given by

$$T = \mathcal{T}(F^t, \theta^t, \mathbf{g}^t), \quad \mathbf{q} = \mathbf{q}(F^t, \theta^t, \mathbf{g}^t), \quad \varepsilon = \varepsilon(F^t, \theta^t, \mathbf{g}^t),$$

where $\mathbf{g} = \text{grad } \theta$ is the spatial gradient of the temperature.

- **Condition of material objectivity**

$$\begin{aligned} \mathcal{T}(Q^t F^t, \theta^t, Q^t \mathbf{g}^t) &= Q(t) \mathcal{T}(F^t, \theta^t, \mathbf{g}^t) Q(t)^T, \\ \mathbf{q}(Q^t F^t, \theta^t, Q^t \mathbf{g}^t) &= Q(t) \mathbf{q}(F^t, \theta^t, \mathbf{g}^t), \\ \varepsilon(Q^t F^t, \theta^t, Q^t \mathbf{g}^t) &= \varepsilon(F^t, \theta^t, \mathbf{g}^t), \end{aligned} \tag{6.17}$$

for any $Q^t \in \mathcal{O}(V)$ and any thermomechanical histories $(F^t, \theta^t, \mathbf{g}^t)$.

- **Condition of material symmetry**

$$\begin{aligned} \mathcal{T}(F^t G, \theta^t, \mathbf{g}^t) &= \mathcal{T}(F^t, \theta^t, \mathbf{g}^t), \\ \mathbf{q}(F^t G, \theta^t, \mathbf{g}^t) &= \mathbf{q}(F^t, \theta^t, \mathbf{g}^t), \\ \varepsilon(F^t G, \theta^t, \mathbf{g}^t) &= \varepsilon(F^t, \theta^t, \mathbf{g}^t), \end{aligned} \tag{6.18}$$

for any $G \in \mathcal{G}$ and any thermomechanical histories $(F^t, \theta^t, \mathbf{g}^t)$.

These conditions are the most fundamental restrictions imposed on any constitutive functions. Since the constitutive functions of simple materials with memory are in general functionals, i.e., functions of history functions, the analysis of the above conditions requires much more mathematical hardware and is beyond the context of this chapter.

Therefore, in order to analyze these conditions, we shall restrict ourselves to much simpler material models, namely, simple materials without long range memory. In this case, a history function, say $h^t(s) = h(t - s)$ for small s , can be expressed in the Taylor series approximation,

$$h^t(s) = h(t) - \dot{h}(t) s + \frac{1}{2} \ddot{h}(t) s^2 + \dots$$

Therefore, the dependence on the history function can be approximated by the dependence on the values of the function and its derivatives up to a certain order at the present time. With this approximation, constitutive functions become ordinary functions instead of functionals. Constitutive theories of such material models can then be analyzed with linear algebra and differential calculus, no theory of functional analysis will be needed.

We shall consider some simple material models:

- Elastic materials
 $\mathcal{C} = \mathcal{C}(F)$.
- Thermoelastic materials
 $\mathcal{C} = \mathcal{C}(F, \theta, \mathbf{g})$.
- Viscoelastic materials
 $\mathcal{C} = \mathcal{C}(F, \dot{F})$.
- Thermo-viscoelastic materials
 $\mathcal{C} = \mathcal{C}(F, \dot{F}, \theta, \mathbf{g})$.

In the following chapters, we shall analyze the restrictions imposed on the constitutive functions, $\mathcal{C} = \{\mathcal{T}, \mathbf{q}, \varepsilon\}$, by the conditions of material objectivity and material symmetry on these models for solids and fluids.

6.4 Remark on incompressibility

A motion is called *incompressible* if it is volume-preserving, which can be characterized by the condition, $|\det F| = 1$. We call a body an incompressible material body if it is capable of undergoing *only* incompressible motions.

In the discussions of constitutive equations so far, it is assumed that a material body is capable of undergoing any motions. Obviously, for incompressible bodies, some constitutive assumptions must be modified. Indeed, in order to maintain the constant volume in the motion some internal stress is needed to counter the tendency of volume change due to applied forces on the body. This is called the reaction stress which maintains constant volume and hence it should not do any real works in the motion.

Since the rate of work in the motion (see the third term of the equation (3.24)) due to the reaction stress N can be expressed as $(N \cdot \text{grad } \dot{\mathbf{x}})$, we shall require that

$$N \cdot \text{grad } \dot{\mathbf{x}} = 0.$$

Taking the material time derivative of the equation $\det F = \pm 1$ and by the use of $\text{grad } \dot{\mathbf{x}} = \dot{F}F^{-1}$ from (2.24), we obtain

$$\frac{d}{dt}(\det F) = (\det F)F^{-T} \cdot \dot{F} = (\det F)I \cdot \dot{F}F^{-1} = 0 \quad \text{or} \quad I \cdot \text{grad } \dot{\mathbf{x}} = 0.$$

By comparison, we conclude that the reaction stress N must be proportional to the identity tensor, so we can write,

$$N = -pI.$$

Therefore for an incompressible body, the stress tensor can be expressed as sum of the reaction pressure and the extra stress,

$$T = -pI + \tilde{\mathcal{T}}(F^t, \theta^t, \mathbf{g}^t), \quad |\det F^t| = 1.$$

The *principle of determinism* for an incompressible material body now requires that only the extra stress be given by a constitutive function of thermomechanical histories of the body. Consequently, the constitutive function of the extra stress $\tilde{\mathcal{T}}$, instead of the (total) stress, is subject to the conditions of material objectivity and material symmetry as discussed in the previous chapters.

The reaction pressure p is a function depending on the applied forces on the body and can not be determined entirely by the thermomechanical histories of the body. It is often referred to as *indeterminate* pressure.

7 Elastic solids

Elasticity is a quality of a material body of being able to recover its original state independent of any history of deformation. In other words, the elastic behaviors depend solely on the present state of deformation. Therefore, the mathematical model for the class of elastic materials can be characterized by the constitutive equation for the Cauchy stress tensor as a function of the deformation gradient, $T = \mathcal{T}(F)$.

7.1 Isotropic elastic solid

For an isotropic elastic solid, the condition of material symmetry (6.16),

$$\mathcal{T}(F) = \mathcal{T}(FG) \quad \forall F, \quad \forall G \in \mathcal{G} = \mathcal{O}(V).$$

By the use of polar decomposition $F = VR$, and by taking $G = R^T$ which is orthogonal, it follows that

$$\mathcal{T}(F) = \mathcal{T}(V) := \tilde{\mathcal{T}}(B), \quad B = V^2 = FF^T.$$

It implies that the constitutive function $\mathcal{T}(F)$ must reduce to a function of the left stretch tensor V or the left Cauchy-Green tensor B only. It is independent of the rotation part R of the deformation – an expected result for being isotropic, i.e., the same in all directions.

Moreover, the condition of material objectivity (5.13), requires the function $\tilde{\mathcal{T}}(B)$ to satisfy

$$\mathcal{T}(QF) = Q\mathcal{T}(F)Q^T \quad \forall F, \quad \forall Q \in \mathcal{O}(V).$$

Therefore, we have

$$Q\tilde{\mathcal{T}}(B)Q^T = \tilde{\mathcal{T}}((QF)(QF)^T), \quad \forall Q \in \mathcal{O}(V),$$

or

$$Q\tilde{\mathcal{T}}(B)Q^T = \tilde{\mathcal{T}}(QBQ^T), \quad \forall Q \in \mathcal{O}(V). \quad (7.1)$$

Conversely, one can easily show that if $\mathcal{T}(F) = \tilde{\mathcal{T}}(FF^T)$ and if it satisfies the relation (7.1), then both the conditions of material objectivity and material symmetry are satisfied.

A tensor-valued function satisfies the relation (7.1) is called an isotropic tensor function.

7.2 Representations of isotropic functions

Definition. Let $S : \mathcal{L}(V) \rightarrow \mathcal{L}(V)$. We say that S is a tensor-valued isotropic functions, if for any $A \in \mathcal{L}(V)$, it satisfies the following condition:

$$S(QAQ^T) = QS(A)Q^T,$$

for any orthogonal transformation $Q \in \mathcal{O}(V)$.

Before giving the representation theorem for an isotropic function, let us recall a theorem in linear algebra,

Theorem (Cayley-Hamilton). *A linear transformation $A \in \mathcal{L}(V)$ satisfies its characteristic equation,*

$$A^3 - \mathbb{I}_A A^2 + \mathbb{II}_A A - \mathbb{III}_A I = 0, \quad (7.2)$$

where $\{\mathbb{I}_A, \mathbb{II}_A, \mathbb{III}_A\}$ are called the principal invariants of A . They are the coefficients of the characteristic polynomial of A , i.e.,

$$\det(\lambda I - A) = \lambda^3 - \mathbb{I}_A \lambda^2 + \mathbb{II}_A \lambda - \mathbb{III}_A.$$

Since eigenvalues of A are the roots of the characteristic equation, $\det(\lambda I - A) = 0$, if A is symmetric and $\{a_1, a_2, a_3\}$ are three eigenvalues of A , then it follows that

$$\mathbb{I}_A = a_1 + a_2 + a_3, \quad \mathbb{II}_A = a_1 a_2 + a_2 a_3 + a_3 a_1, \quad \mathbb{III}_A = a_1 a_2 a_3.$$

It is obvious that $\mathbb{I}_A = \text{tr } A$ and $\mathbb{III}_A = \det A$ are the trace and the determinant of the tensor A respectively. Moreover, \mathbb{I}_A , \mathbb{II}_A and \mathbb{III}_A are respectively a first order, a second order and a third order quantities of $|A|$.

Let the set of all symmetric linear transformation be denoted by $\text{Sym}(V)$.

Theorem. *Let $S : \text{Sym}(V) \rightarrow \text{Sym}(V)$, then it is an isotropic function if and only if it can be represented by*

$$S(A) = s_0 I + s_1 A + s_2 A^2, \quad (7.3)$$

where s_0 , s_1 and s_2 are arbitrary scalar functions of $(\mathbb{I}_A, \mathbb{II}_A, \mathbb{III}_A)$.

Corollary. *If $S(A)$ is an isotropic and linear function of A , then*

$$S(A) = \lambda (\text{tr } A) I + \mu A, \quad (7.4)$$

where λ and μ are independent of A .

This theorem was first proved by Rivlin & Ericksen (1955). Representations for isotropic functions of any number of vector and tensor variables have been extensively studied and the results are usually tabulated in the literature. We shall give here without proof another theorem for isotropic functions of one vector and one symmetric tensor variables.

Theorem. *Let $\mathcal{D} = V \times \text{Sym}(V)$, and $\phi : \mathcal{D} \rightarrow \mathbb{R}$, $\mathbf{h} : \mathcal{D} \rightarrow V$, and $S : \mathcal{D} \rightarrow \text{Sym}(V)$. Then they are isotropic if and only if they can be represented by*

$$\begin{aligned} \phi &= \varphi(\mathbb{I}_A, \mathbb{II}_A, \mathbb{III}_A, \mathbf{v} \cdot \mathbf{v}, \mathbf{v} \cdot A\mathbf{v}, \mathbf{v} \cdot A^2\mathbf{v}), \\ \mathbf{h} &= h_0 \mathbf{v} + h_1 A\mathbf{v} + h_2 A^2\mathbf{v}, \\ S &= s_0 I + s_1 A + s_2 A^2 + s_3 \mathbf{v} \otimes \mathbf{v} + s_4 (A\mathbf{v} \otimes \mathbf{v} + \mathbf{v} \otimes A\mathbf{v}) + s_5 A\mathbf{v} \otimes A\mathbf{v}, \end{aligned} \quad (7.5)$$

where the coefficients h_0 through h_2 and s_0 through s_5 are arbitrary functions of the variables indicated in the scalar function φ .

Therefore, from the relation (7.1) and the representation theorem, we have

Theorem. *The most general constitutive equation for an Isotropic elastic solid is given by*

$$T = t_0 I + t_1 B + t_2 B^2,$$

where t_0, t_1, t_2 are functions of the principal invariants (I_B, II_B, III_B) .

7.3 Incompressible isotropic elastic solids

Incompressible elastic bodies can be similarly formulated. It is known that the reaction stress for incompressibility is an indeterminate pressure and therefore, the constitutive equation for the stress tensor is given by

$$T = -pI + t_1 B + t_2 B^2, \quad \det B = III_B = 1,$$

where t_1 and t_2 are functions of (I_B, II_B) and p is the indeterminate pressure. By the use of Cayley-Hamilton theorem, $B^2 = I_B B - II_B I + B^{-1}$, it can also be expressed by

$$T = -pI + s_1 B + s_2 B^{-1}, \quad \det B = 1,$$

where the parameters s_1 and s_2 are functions of (I_B, II_B) . Two special cases are of practical interest for finite elasticity, namely, the simple models for which the parameters s_1 and s_2 are constants.

- Neo-Hookean material: $T = -pI + s_1 B$.
- Mooney-Rivlin material: $T = -pI + s_1 B + s_2 B^{-1}$.

Thermodynamic stability analysis¹ requires that

$$s_1 > 0, \quad s_2 < s_1.$$

These incompressible material models are often adopted for rubber-like materials. It provides a reasonable theory of natural rubber at finite strains.

¹ Liu, I-Shih, A note on the Mooney-Rivlin material model, *Continuum Mechanics and Thermodynamics*, 24, 583-590 (2012)

7.4 Elastic solid materials

For elastic solids in general, the constitutive function must satisfy the conditions of material objectivity,

$$\mathcal{T}(QF) = Q\mathcal{T}(F)Q^T, \quad \forall F, \quad \forall Q \in \mathcal{O}(V), \quad (7.6)$$

and the condition of material symmetry,

$$\mathcal{T}(FG) = \mathcal{T}(F), \quad \forall F, \quad \forall G \in \mathcal{G}, \quad (7.7)$$

where $\mathcal{G} \subset \mathcal{O}(V)$ is the symmetry group of the solid material body.

By the use of polar decomposition $F = RU$, and since the condition of material objectivity (7.6) is valid for any orthogonal tensor Q , by taking $Q = R^T$, it follows that

$$\mathcal{T}(F) = Q^T\mathcal{T}(QF)Q = R\mathcal{T}(R^TRU)R^T, \quad \text{or} \quad \mathcal{T}(F) = R\mathcal{T}(U)R^T. \quad (7.8)$$

Therefore, the constitutive function $\mathcal{T}(F)$ satisfies the condition of material objectivity (7.6) if and only if it can be represented in the form (7.8). The representation (7.8) requires that the dependence on F must reduce to a specific form of dependence on the stretch part U and the rotation part R . It can not depend on the deformation gradient F in an arbitrary manner.

The representation (7.8) for the Cauchy stress tensor takes a simpler form in terms of the *second Piola-Kirchhoff stress tensor* defined as

$$S = |\det F| F^{-1}TF^{-T}.$$

It follows that

$$S = |\det(RU)| (RU)^{-1} (R\mathcal{T}(U)R^T) (RU)^{-T} = (\det U) U^{-1}\mathcal{T}(U)U^{-T},$$

where $|\det R| = 1$ has been used since $R \in \mathcal{O}(V)$.

Therefore, the material objectivity condition implies that the second Piola-Kirchhoff stress tensor for elastic materials must reduce to a function of the right stretch tensor U or equivalently of the right Cauchy-Green strain tensor $C = U^2$ only,

$$S = \mathcal{S}(C), \quad C = F^TF.$$

This representation is more convenient in practical calculations in terms of the deformation gradient, because no calculation of U from polar decomposition is necessary.

Moreover, from this representation and the condition of material symmetry (7.7), we have

$$\mathcal{S}(G^TCG) = G^{-1}\mathcal{S}(C)G^{-T}, \quad \forall G \in \mathcal{G} \subset \mathcal{O}(V),$$

for any elastic material with symmetry group \mathcal{G} .

Elastic solid. The second Piola-Kirchhoff stress tensor of an elastic solid with symmetry group $\mathcal{G} \subseteq \mathcal{O}(V)$ is given by

$$S = \mathcal{S}(C), \quad \forall C = F^T F, \quad (7.9)$$

for some function $\mathcal{S} : \text{Sym}(V) \rightarrow \text{Sym}(V)$ satisfying the following condition:

$$\mathcal{S}(Q C Q^T) = Q \mathcal{S}(C) Q^T, \quad \forall Q \in \mathcal{G}, \quad \forall C. \quad (7.10)$$

A function satisfying the relation (7.10) is called invariant relative to the group \mathcal{G} . Explicit representations for constitutive functions of elastic solids invariant relative to transversely isotropic groups and some symmetry groups of crystalline solids can be found in the literature.

7.5 Hooke's law

In the classical theory of linear elasticity, only small deformations are considered. We introduce the displacement vector from the reference configuration and its gradient,

$$\mathbf{u} = \chi_\kappa(\mathbf{X}) - \mathbf{X}, \quad H = \nabla_{\mathbf{X}} \mathbf{u}, \quad \mathbf{X} \in \mathcal{B}_\kappa.$$

We have $H = F - I$. For small deformations, the displacement gradient H is assumed to be a small quantity of order $o(1)$. The Cauchy-Green tensor C can then be approximated by

$$C = F^T F = (I + H)^T (I + H) = I + H + H^T + H^T H = I + 2E + o(2),$$

where the infinitesimal strain tensor E is defined as the symmetric part of the displacement gradient,

$$E = \frac{1}{2}(H + H^T).$$

The function \mathcal{S} of the equation (7.9) can now be approximated by

$$\mathcal{S}(C) = \mathcal{S}(I) + \mathbf{L}[E] + o(2), \quad \mathbf{L}[E] := \left. \frac{d}{dt} \mathcal{S}(I + 2Et) \right|_{t=0}.$$

Here we have defined a fourth order tensor \mathbf{L} as a linear transformation of $\text{Sym}(V)$ into itself. If we further assume that the reference configuration is a natural state, *i.e.*, zero stress at the undeformed state, $\mathcal{T}(I) = 0$, then so is $\mathcal{S}(I) = 0$. Since $\mathbf{L}[E]$ is of order $o(1)$ and $F = I + o(1)$, by neglecting the second order terms in (7.9) we obtain

$$T = \mathbf{L}[E], \quad (7.11)$$

This linear stress-strain relation is known as the *Hooke's law* and \mathbf{L} is called the *elasticity tensor*. Since both the stress and the strain tensors are symmetric, by definition, the elasticity tensor has the following symmetry properties in terms of Cartesian components:

$$T_{ij} = \sum_{k,l=1}^3 L_{ijkl} E_{kl}, \quad L_{ijkl} = L_{jikl} = L_{ijlk}. \quad (7.12)$$

Moreover, the conditions of material objectivity (7.6) and material symmetry (7.7) imply that

$$\mathcal{T}(QFQ^T) = Q\mathcal{T}(F)Q^T, \quad \forall Q \in \mathcal{G} \subseteq \mathcal{O}(V).$$

Since $\mathcal{T}(F) = \mathbf{L}[E(F)]$ and $E(F) = \frac{1}{2}(H + H^T) = \frac{1}{2}(F + F^T) - I$, it follows immediately that

$$\mathbf{L}[QEQ^T] = Q\mathbf{L}[E]Q^T, \quad \forall Q \in \mathcal{G} \subseteq \mathcal{O}(V).$$

In other words, the elasticity tensor $\mathbf{L} : \text{Sym}(V) \rightarrow \text{Sym}(V)$ is linear and invariant relative to the symmetry group $\mathcal{G} \subset \mathcal{O}(V)$ for anisotropic linear elastic solids in general. In particular, for $\mathcal{G} = \mathcal{O}(V)$, by the use of the linear representation (7.4), the Hooke's law for isotropic linear elastic solid body becomes

$$T = \lambda(\text{tr } E)I + \mu E,$$

where the material parameters λ and μ are called Lamé elastic coefficients.

Remark. We should point out that the linear law (7.11) does not satisfy the condition of material objectivity (7.6) for arbitrary orthogonal tensor Q . Therefore, the theory of linear elasticity is meaningless for large deformations.

Indeed, if we choose $F = 1$, which is a natural state by assumption, then the condition (7.6) implies that

$$T(Q) = QT(1)Q^T = 0,$$

for any orthogonal tensor Q . On the other hand, since

$$T(F) = \mathbf{L}[E(F)], \quad E(F) = \frac{1}{2}(H + H^T) = \frac{1}{2}(F + F^T) - 1,$$

if we choose Q as a rotation about z -axis,

$$Q = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

we have

$$E(Q) = \begin{bmatrix} \cos \theta - 1 & 0 & 0 \\ 0 & \cos \theta - 1 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

and for $\theta \neq 0$

$$T(Q) = \mathbf{L}[E(Q)] \neq 0.$$

Hence, the condition of material objectivity is not satisfied in general. Nevertheless, one can see that it is approximately satisfied when both the displacement and the rotation are small (in this case, $\cos \theta \approx 1$). \square

8 Viscoelastic materials

From physical experiences, viscosity is a phenomenon associated with the rate of deformation – the greater the deformation rate, the greater the resistance to motion. Therefore, we shall consider a simple model for viscous materials given by the constitutive equation for the Cauchy stress, $T = \mathcal{T}(F, \dot{F})$.

From (6.17), the constitutive function \mathcal{T} must satisfy the condition of material objectivity,

$$\mathcal{T}(QF, (QF)\dot{\cdot}) = Q\mathcal{T}(F, \dot{F})Q^T, \quad \forall Q \in \mathcal{O}(V), \quad \forall F, \quad (8.13)$$

and from (6.18), the condition of material symmetry,

$$\mathcal{T}(FG, \dot{F}G) = \mathcal{T}(F, \dot{F}), \quad \forall G \in \mathcal{G}, \quad \forall F. \quad (8.14)$$

8.1 Isotropic viscoelastic solids

For isotropic solids, the symmetry group is the orthogonal group $\mathcal{G} = \mathcal{Q}(V)$. To obtain the restrictions imposed on the constitutive function \mathcal{T} , consider the polar decomposition $F = VR$ and take $G = R^T \in \mathcal{G}$ in the condition (8.14),

$$\mathcal{T}(F, \dot{F}) = \mathcal{T}(FR^T, \dot{F}R^T) = \mathcal{T}(VRR^T, \dot{F}F^{-1}FR^T) = \mathcal{T}(V, LV),$$

where we have used the relation $L = \dot{F}F^{-1}$ for velocity gradient. In other words, the dependence on (F, \dot{F}) reduces to the dependence on (V, L) , or equivalently on (B, L) where $B = FF^T$ is the left Cauchy–Green strain tensor. Therefore, we can write

$$\mathcal{T}(F, \dot{F}) = \widehat{\mathcal{T}}(B, L).$$

On the other hand, the function $\widehat{\mathcal{T}}$ must satisfy the condition of material objectivity (8.13) which becomes

$$\widehat{\mathcal{T}}((QF)(QF)^T, (QF)\dot{\cdot}(QF)^{-1}) = Q\widehat{\mathcal{T}}(B, L)Q^T.$$

Simplifying the left-hand side and decomposing $L = D + W$ into symmetric and skew-symmetric parts, we obtain

$$\widehat{\mathcal{T}}(QBQ^T, QDQ^T + QWQ^T + \dot{Q}Q^T) = Q\widehat{\mathcal{T}}(B, L)Q^T. \quad (8.15)$$

This relation must hold for any orthogonal tensor Q . In particular, we can choose the orthogonal tensor (check it!), $Q^t(s) = \exp((t-s)W)$, where W is the skew-symmetric part of the velocity gradient, so that $Q(t) = I$ and $\dot{Q}(t) = -W$ and the above relation reduces to

$$\widehat{\mathcal{T}}(B, L) = \widehat{\mathcal{T}}(B, D),$$

where D is the symmetric part of L . Hence, the constitutive function can not depend on the skew-symmetric part W of the velocity gradient. This, in turns, implies from the above relation (8.15) that $\widehat{\mathcal{T}}(B, D)$ is an isotropic tensor function,

$$\widehat{\mathcal{T}}(QBQ^T, QDQ^T) = Q\widehat{\mathcal{T}}(B, D)Q^T, \quad \forall Q \in \mathcal{O}(V).$$

This isotropic function depends on two symmetric tensor variables. The general representation for such a function can be expressed in the following form²:

$$\begin{aligned} T &= t_1 I + t_2 B + t_3 B^2 + t_4 D + t_5 D^2 \\ &\quad + t_6 (BD + DB) + t_7 (B^2 D + DB^2) + t_8 (BD^2 + D^2 B), \\ t_i &= t_i(\mathbb{I}_B, \mathbb{II}_B, \mathbb{III}_B, \mathbb{I}_D, \mathbb{II}_D, \mathbb{III}_D, \text{tr}(BD), \text{tr}(B^2 D), \text{tr}(BD^2), \text{tr}(B^2 D^2)). \end{aligned} \quad (8.16)$$

For some practical problems of small deformation rate, one can consider a representation which contains up to linear terms in D . It can be written as

$$\begin{aligned} T &= s_0 I + s_1 B + s_2 B^{-1} + \mu_1 D + \mu_2 (BD + DB) + \mu_3 (B^{-1} D + DB^{-1}), \\ s_i &= a_i + b_i \text{tr} D + c_i \text{tr}(BD) + d_i \text{tr}(B^{-1} D), \\ \mu_i &= \mu_i(\mathbb{I}_B, \mathbb{II}_B, \mathbb{III}_B), \end{aligned}$$

and the coefficients a_i , b_i , c_i , and d_i , are also functions of $(\mathbb{I}_B, \mathbb{II}_B, \mathbb{III}_B)$. In these expressions, obtained from (8.16), we have replaced B^2 with B^{-1} as we did in Mooney–Rivlin materials.

8.2 Viscoelastic solids

For viscoelastic solids in general, the constitutive function, $T = \mathcal{T}(F, \dot{F})$, must satisfy the conditions of material objectivity,

$$\mathcal{T}(QF, (QF)\dot{}) = Q\mathcal{T}(F, \dot{F})Q^T, \quad F, \quad \forall Q \in \mathcal{O}(V), \quad (8.17)$$

and the condition of material symmetry,

$$\mathcal{T}(FG, \dot{F}G) = \mathcal{T}(F, \dot{F}), \quad \forall G \in \mathcal{G}, \quad (8.18)$$

where $\mathcal{G} \subset \mathcal{O}(V)$ is the symmetry group of the solid material body.

By the use of polar decomposition $F = RU$, and since the condition of material objectivity (8.17) is valid for any orthogonal tensor Q , by taking $Q = R^T$, it follows that

$$QF = R^T(RU) = U, \quad (QF)\dot{} = (R^T(RU))\dot{} = \dot{U},$$

and hence,

$$\mathcal{T}(F, \dot{F}) = R\mathcal{T}(U, \dot{U})R^T. \quad (8.19)$$

²I-Shih Liu, *Continuum Mechanics*, Springer 2002, (Chapter 4).

Therefore, for any viscoelastic solid, the constitutive equation for the Cauchy stress tensor must reduce to the above representation (8.19).

As before, it follows that the second Piola-Kirchhoff stress tensor must reduce to

$$S(F, \dot{F}) = \mathcal{S}(C, \dot{C}), \quad C = F^T F = U^2.$$

Moreover, from this representation and the condition of material symmetry (8.18), we have

$$\mathcal{S}(G^T C G, G^T \dot{C} G) = G^{-1} \mathcal{S}(C, \dot{C}) G^{-T}, \quad \forall G \in \mathcal{G} \subset \mathcal{O}(V).$$

Viscoelastic solid. *The second Piola-Kirchhoff stress tensor of a viscoelastic solid with symmetry group $\mathcal{G} \subseteq \mathcal{O}(V)$ is given by*

$$S = \mathcal{S}(C, \dot{C}), \quad C = F^T F, \quad (8.20)$$

for some function $\mathcal{S} : \text{Sym}(V) \rightarrow \text{Sym}(V)$ satisfying the following condition:

$$\mathcal{S}(Q C Q^T, Q \dot{C} Q) = Q \mathcal{S}(C, \dot{C}) Q^T, \quad \forall Q \in \mathcal{G}, \quad \forall C, \quad (8.21)$$

that is, it is invariant relative to the group \mathcal{G} .

Remark: Since the general representation of viscoelastic solid is given by (8.19),

$$T(F, \dot{F}) = R \mathcal{T}(U, \dot{U}) R^T. \quad (8.22)$$

in particular, this must hold for isotropic viscoelastic solid which has been proved that it must reduce to

$$T(F, \dot{F}) = \widehat{\mathcal{T}}(B, D). \quad (8.23)$$

Therefore, (8.23) must be a special case of (8.22).

To see this, note that $F = V R = R U$, so that

$$B = V^2 = R U^2 R^T,$$

and, with $L = \dot{F} F^{-1}$ and $D = \frac{1}{2}(L + L^T)$,

$$\dot{C} = (F^T F) \dot{} = F^T (\dot{F} F^{-1} + F^{-T} \dot{F}^T) F = 2 F^T D F,$$

which, with $F = R U$ and $C = U^2$, implies that

$$D = \frac{1}{2} F^{-T} \dot{C} F^{-1} = R U^{-1} (U \dot{U}) U^{-1} R^T = R (\dot{U} U^{-1}) R^T.$$

Therefore,

$$\widehat{\mathcal{T}}(B, D) = \widehat{\mathcal{T}}(R U^2 R^T, R (\dot{U} U^{-1}) R^T),$$

and since the function $\widehat{\mathcal{T}}$ is isotropic and R is orthogonal, it follows that

$$\widehat{\mathcal{T}}(B, D) = R \widehat{\mathcal{T}}(U^2, (\dot{U} U^{-1})) R^T,$$

which is a particular case of (8.22). \square

8.3 Viscous fluids

For fluids, the symmetry group is the unimodular group $\mathcal{G} = \mathcal{U}(V)$. To obtain the restrictions of these conditions imposed on the constitutive function \mathcal{T} , we shall take $G = |\det F|^{1/3} F^{-1}$, obviously $|\det G| = 1$ so that G belongs to the symmetry group, and by (8.14), it follows that

$$\begin{aligned} \mathcal{T}(F, \dot{F}) &= \mathcal{T}(|\det F|^{1/3} F F^{-1}, |\det F|^{1/3} \dot{F} F^{-1}) \\ &= \mathcal{T}(|\det F|^{1/3} I, |\det F|^{1/3} L) := \widehat{\mathcal{T}}(|\det F|, L). \end{aligned}$$

Therefore, for fluids, the material symmetry requires that the dependence of \mathcal{T} on (F, \dot{F}) be reduced to the dependence on the determinant of F and the velocity gradient $L = \dot{F} F^{-1}$ as defined by the constitutive function $\widehat{\mathcal{T}}(|\det F|, L)$.

Furthermore, the function $\widehat{\mathcal{T}}$ must satisfy the condition of material objectivity (8.13) which becomes

$$\widehat{\mathcal{T}}(|\det(QF)|, (QF) \cdot (QF)^{-1}) = Q \widehat{\mathcal{T}}(|\det F|, L) Q^T.$$

By a similar argument from isotropic viscoelastic solids, we arrive the conclusion that the constitutive function can not depend on the skew-symmetric part W of the velocity gradient, and that $\widehat{\mathcal{T}}(|\det F|, D)$ is an isotropic tensor function,

$$\widehat{\mathcal{T}}(|\det F|, QDQ^T) = Q \widehat{\mathcal{T}}(|\det F|, D) Q^T, \quad \forall Q \in \mathcal{O}(V).$$

Moreover, from the conservation of mass, we have $|\det F| = \rho_\kappa / \rho$, where the mass density ρ_κ in the reference configuration is constant. Consequently, by replacing the dependence on $|\det F|$ with the mass density ρ , and by the use of the representation theorem for isotropic functions (7.3), we obtain the constitutive equation,

$$T = \widetilde{\mathcal{T}}(\rho, D) = d_0 I + d_1 D + d_2 D^2, \quad (8.24)$$

where the material parameters d_0, d_1, d_2 are functions of mass density and three principal invariants of the rate of strain tensor, $(\rho, \mathbb{I}_D, \mathbb{II}_D, \mathbb{III}_D)$.

This is the most general constitutive equation for the viscous fluid of the simple model $\mathcal{T}(F, \dot{F})$. It was first derived by Reiner (1945) and by Rivlin (1947) and it is usually known as *Reiner-Rivlin fluid*. However, we should point out that this is by no means the most general constitutive equation for simple viscous fluids. Indeed, one may consider other viscous fluid models which depend also on deformation rates of higher order, for example, a simple fluid of grade-two $\mathcal{T}(F, \dot{F}, \ddot{F})$. In this note, we shall restrict our attention to simple models only.

8.4 Navier-Stokes fluids

The most well-known viscous fluid models is the Navier-Stokes fluids. It is a mathematically simpler model of the general one (8.24) in which only linear dependence on the rate

of strain is relevant and hence by the linear representation (7.4), the constitutive equation for Navier-Stokes fluids is given by

$$T = -p(\rho)I + \lambda(\rho)(\text{tr } D)I + 2\mu(\rho)D. \quad (8.25)$$

The coefficients λ and μ are called the coefficients of viscosity, while μ and $(\lambda + \frac{2}{3}\mu)$ are also known as the *shear* and the *bulk viscosities* respectively. The pressure p and the viscosities λ and μ are functions of ρ .

A Navier-Stokes fluid is also known as a *Newtonian fluid* in fluid mechanics. It is usually assumed that

$$\mu \geq 0, \quad 3\lambda + 2\mu \geq 0.$$

The non-negativeness of the shear and bulk viscosities can be proved from thermodynamic considerations.

It should be pointed out that unlike the Hooke's law in linear elasticity which is an approximate model for small deformations only, the Navier-Stokes fluids defines a class of material models which satisfies both the conditions of material objectivity and material symmetry. It need not be regarded as the linear approximation of a Reiner-Rivlin fluid. Thus it is conceivable that there are some fluids which obey the constitutive equation (8.25) for arbitrary rate of deformation. Indeed, water and air are usually treated as Navier-Stokes fluids in most practical applications with very satisfactory results even under rapid flow conditions

8.5 Viscous heat-conducting fluids

We now consider a simple fluid with heat conduction and viscosity given by the following constitutive equation for $\mathcal{C} = \{\mathcal{T}, \mathbf{q}, \varepsilon\}$,

$$\mathcal{C} = \mathcal{C}(F, \dot{F}, \theta, \mathbf{g}).$$

With the same arguments as before, from the conditions of material objectivity and material symmetry, the constitutive variables $(F, \dot{F}, \theta, \mathbf{g})$ must reduce to $(\rho, D, \theta, \mathbf{g})$ and the constitutive functions are isotropic functions,

$$\begin{aligned} \mathcal{T}(\rho, QDQ^T, \theta, Q\mathbf{g}) &= Q\mathcal{T}(\rho, D, \theta, \mathbf{g})Q^T, \\ \mathbf{q}(\rho, QDQ^T, \theta, Q\mathbf{g}) &= Q\mathbf{q}(\rho, D, \theta, \mathbf{g}), \quad \forall Q \in \mathcal{O}(V), \quad \forall (\rho, D, \theta, \mathbf{g}). \\ \varepsilon(\rho, QDQ^T, \theta, Q\mathbf{g}) &= \varepsilon(\rho, D, \theta, \mathbf{g}), \end{aligned}$$

Therefore, from the representation (7.5), one can immediately write down the most general constitutive equations of a viscous heat-conducting fluid for the stress, the heat flux, and the internal energy,

$$\begin{aligned} T &= \alpha_0 I + \alpha_1 D + \alpha_2 D^2 + \alpha_3 \mathbf{g} \otimes \mathbf{g} + \alpha_4 (D\mathbf{g} \otimes \mathbf{g} + \mathbf{g} \otimes D\mathbf{g}) + \alpha_5 D\mathbf{g} \otimes D\mathbf{g}, \\ \mathbf{q} &= \beta_1 \mathbf{g} + \beta_2 D\mathbf{g} + \beta_3 D^2\mathbf{g}, \\ \varepsilon &= \varepsilon(\rho, \theta, \mathbb{I}_D, \mathbb{II}_D, \mathbb{III}_D, \mathbf{g} \cdot \mathbf{g}, \mathbf{g} \cdot D\mathbf{g}, \mathbf{g} \cdot D^2\mathbf{g}), \end{aligned}$$

where the coefficient α_i and β_j as well as ε are scalar functions of eight variables indicated in the arguments of ε .

The special case, when only up to linear terms in both the strain rate D and the temperature gradient \mathbf{g} are relevant, gives the most widely-used model for viscosity and heat conduction.

Navier-Stokes-Fourier fluids

$$\begin{aligned} T &= -p(\rho, \theta) I + \lambda(\rho, \theta) (\text{tr } D) I + 2\mu(\rho, \theta) D, \\ \mathbf{q} &= -\kappa(\rho, \theta) \mathbf{g}, \\ \varepsilon &= \varepsilon(\rho, \theta) + \varepsilon_1(\rho, \theta) \text{tr } D. \end{aligned}$$

These are the classical *Navier-Stokes* theory and the *Fourier law* of heat conduction. The material parameters λ , μ are the viscosity coefficients and κ is called the thermal conductivity. From thermodynamic considerations, it is possible to prove that the thermal conductivity is non-negative, and the internal energy is independent of the strain rate, so that $\varepsilon = \varepsilon(\rho, \theta)$.

For incompressible fluids, the mass density ρ is a constant field and the equation of mass balance in (4.11) implies that $\text{div } \dot{\mathbf{x}} = 0$, or $\text{tr } D = 0$. Therefore, we have

Incompressible Navier-Stokes fluids

$$\begin{aligned} T &= -p I + 2\mu(\theta) D, \quad \text{tr } D = 0, \\ \mathbf{q} &= -\kappa(\theta) \mathbf{g}, \\ \varepsilon &= \varepsilon(\theta), \end{aligned}$$

where the pressure is no longer a constitutive parameter but rather must be determined from the condition $\text{div } \dot{\mathbf{x}} = 0$ and suitable boundary conditions.

Another important special case is the simplest model in Continuum Mechanics, given by constitutive equations depending on the density and the temperature only. Hence, it is also a special case of elastic materials.

Elastic fluids

$$T = -p(\rho, \theta) I, \quad \mathbf{q} = 0, \quad \varepsilon = \varepsilon(\rho, \theta).$$

This defines an inviscid compressible fluid without heat conduction, also known as Euler fluid or ideal fluid in Fluid Mechanics. Similarly, in the case of incompressible Euler fluids, the mass density ρ is a constant field and the indeterminate pressure p depends on boundary conditions and the condition $\text{div } \dot{\mathbf{x}} = 0$.

9 Second law of thermodynamics

We shall give a brief consideration of thermodynamic restrictions imposed on constitutive equations. We have already mentioned the first law of thermodynamics, i.e., the energy balance. Now we are going to consider the second law for which the essential quantity is the *entropy*,

$$\int_{\mathcal{P}_t} \rho \eta \, dv, \quad (9.1)$$

where $\eta(\mathbf{x}, t)$ is called the specific *entropy density*. Unlike the total energy, the rate of change of total entropy of a body can not be given completely in the form of a balance equation (3.1). There are internal entropy productions in “non-equilibrium” processes.

Entropy production. For any part $\mathcal{P} \subset \mathcal{B}$, the entropy production $\sigma(\mathcal{P}, t)$ is given by

$$\sigma(\mathcal{P}, t) = \frac{d}{dt} \int_{\mathcal{P}_t} \rho \eta \, dv + \int_{\partial \mathcal{P}_t} \boldsymbol{\Phi} \cdot \mathbf{n} \, da - \int_{\mathcal{P}_t} \rho s \, dv.$$

We call $\boldsymbol{\Phi}(\mathbf{x}, t)$ the *entropy flux* and s the external *entropy supply density*. Although entropy is not a quantity associated with some easily measurable physical quantities, its existence is usually inferred from some more fundamental hypotheses concerning thermal behaviors of material bodies, usually known as the second law of thermodynamics. We choose to accept the existence of entropy and state the consequence of such hypotheses directly by saying that the entropy production is a non-negative quantity.

Second law of thermodynamics. The following entropy inequality must hold for any part $\mathcal{P} \subset \mathcal{B}$:

$$\frac{d}{dt} \int_{\mathcal{P}_t} \rho \eta \, dv + \int_{\partial \mathcal{P}_t} \boldsymbol{\Phi} \cdot \mathbf{n} \, da - \int_{\mathcal{P}_t} \rho s \, dv \geq 0. \quad (9.2)$$

Comparing with the general balance equation (3.1) by setting

$$\psi = \rho \eta, \quad \boldsymbol{\Phi}_\psi = \boldsymbol{\Phi}, \quad \sigma_\psi = \rho s,$$

we have the following local form of the entropy inequality,

$$\rho \dot{\eta} + \operatorname{div} \boldsymbol{\Phi} - \rho s \geq 0. \quad (9.3)$$

Another quantity essential in thermodynamics introduced previously is the temperature. Intuitively it is a measurable quantity by contact thermometers. In fact, this property is based on the assumption that on the interface between two bodies, the temperature

is continuous. We shall call such a wall an ideal wall. The continuity of temperature across an ideal wall is often referred as the *zeroth law of thermodynamics*.

In addition to the continuity of temperature, we shall also postulate the absence of energy and entropy productions at an ideal wall so that the normal components of energy flux and entropy flux are continuous across the wall. This may be regarded as a supplement to the laws of thermodynamics. That is, for a wall between bodies I and II,

$$\mathbf{q}_I \cdot \mathbf{n} = \mathbf{q}_{II} \cdot \mathbf{n}, \quad \Phi_I \cdot \mathbf{n} = \Phi_{II} \cdot \mathbf{n}, \quad \text{if} \quad \theta_I = \theta_{II}. \quad (9.4)$$

9.1 Entropy principle

One of the principal objectives of continuum mechanics is to determine or predict the behavior of a body once the external causes are specified. Mathematically, this amounts to solve initial boundary value problems governed by the balance laws of mass, linear momentum and energy,

$$\begin{aligned} \dot{\rho} + \rho \operatorname{div} \dot{\mathbf{x}} &= 0, \\ \rho \ddot{\mathbf{x}} - \operatorname{div} T &= \rho \mathbf{b}, \\ \rho \dot{\varepsilon} + \operatorname{div} \mathbf{q} - T \cdot \operatorname{grad} \dot{\mathbf{x}} &= \rho r, \end{aligned} \quad (9.5)$$

when the external supplies \mathbf{b} and r are given.

The governing field equations are obtained, for the determination of the fields of the density $\rho(\mathbf{X}, t)$, the motion $\chi(\mathbf{X}, t)$, and the temperature $\theta(\mathbf{X}, t)$, after introducing the constitutive relations for T , ε , and \mathbf{q} , into the balance laws (9.5). Any solution $\{\rho(\mathbf{X}, t), \chi(\mathbf{X}, t), \theta(\mathbf{X}, t)\}$ of the field equations is called a *thermodynamic process*.

On the other hand, the behavior of a body must also obey the second law of thermodynamics, *i.e.*, a thermodynamic process must also satisfy the entropy inequality (9.3). Following the idea set forth by Coleman and Noll, the second law of thermodynamics plays an essential role in constitutive theories of continuum mechanics.

Entropy principle. *It is required that constitutive relations be such that the entropy inequality is satisfied identically for any thermodynamic process.*

From this point of view, like the principle of material objectivity and material symmetry, the entropy principle also imposes restrictions on constitutive functions. To find such restrictions is one of the major task in modern continuum thermodynamics. We shall illustrate the procedures of exploiting the entropy principle in this section.

We shall first make some general remarks. Motivated by the results of classical thermostatics, it is often assumed that the entropy flux and the entropy supply are proportional to the heat flux and the heat supply respectively. Moreover, both proportional constants

are assumed to be the reciprocal of the *absolute temperature* θ ,

$$\boldsymbol{\Phi} = \frac{1}{\theta} \mathbf{q}, \quad s = \frac{1}{\theta} r. \quad (9.6)$$

The resulting entropy inequality is called the *Clausius-Duhem inequality*,

$$\rho \dot{\eta} + \operatorname{div} \frac{\mathbf{q}}{\theta} - \rho \frac{r}{\theta} \geq 0. \quad (9.7)$$

Exploitation of entropy principle based on the Clausius-Duhem inequality has been widely adopted in the development of modern continuum thermodynamics following the simple Coleman-Noll procedure.

However, the main assumptions (9.6) while seem to be plausible in all classical theories of continuum mechanics, are not particularly well motivated for materials in general. In fact, the relation (9.6) is known to be inconsistent with the results from the kinetic theory of ideal gases and is also found to be inappropriate to account for thermodynamics of diffusion.

Müller, on the other hand, proposed that if the body is free of external supplies (i.e., $\mathbf{b} = 0$, $r = 0$) the entropy supply must also vanish ($s = 0$), which is certainly much weaker than the assumption (9.6). Since constitutive functions do not depend on the external supplies, in exploring thermodynamic restrictions, it suffices to consider only supply-free bodies.

Exploitation of the entropy principle based on the entropy inequality in its general form (9.3) has been proposed by Müller and the method of Lagrange multipliers proposed by Liu greatly facilitates its procedure (sometimes referred to as Müller-Liu procedure).

9.2 Thermodynamics of heat-conducting elastic fluids

We shall now exploit the entropy principle based on the general entropy inequality for a simple case of heat-conducting elastic fluids following Müller-Liu procedure.

The constitutive relations for heat-conducting elastic fluids are given by

$$\begin{aligned} T &= T(\rho, \theta, \mathbf{g}), & \eta &= \eta(\rho, \theta, \mathbf{g}), \\ \mathbf{q} &= \mathbf{q}(\rho, \theta, \mathbf{g}), & \boldsymbol{\Phi} &= \boldsymbol{\Phi}(\rho, \theta, \mathbf{g}), \\ \varepsilon &= \varepsilon(\rho, \theta, \mathbf{g}), \end{aligned} \quad (9.8)$$

where θ will be regarded as an empirical temperature, and $\mathbf{g} = \nabla \theta$ is its spatial gradient.

9.3 Exploitation of entropy principle

For a supply-free body, we have the following balance laws,

$$\begin{aligned}\dot{\rho} + \rho \operatorname{div} \dot{\mathbf{x}} &= 0, \\ \rho \ddot{\mathbf{x}} - \operatorname{div} T &= 0, \\ \rho \dot{\varepsilon} + \operatorname{div} \mathbf{q} - T \cdot \operatorname{grad} \dot{\mathbf{x}} &= 0,\end{aligned}$$

and the entropy inequality,

$$\rho \dot{\eta} + \operatorname{div} \Phi \geq 0.$$

The entropy principle requires that the above inequality must hold for any thermodynamic process $\{\rho, \chi, \theta\}$. This requirement can be stated in a different way, namely, the fields that satisfy the entropy inequality are constrained by the requirement that they must be solutions of the field equations. Following Liu, we can take care of this requirement by the use of Lagrange multipliers much like that in the classical problems of finding the extremum with constraints.

Method of Lagrange multipliers. *There exist Lagrange multipliers Λ^ρ , Λ^v , and Λ^ε such that the inequality*

$$\rho \dot{\eta} + \operatorname{div} \Phi - \Lambda^\rho(\dot{\rho} + \rho \operatorname{div} \dot{\mathbf{x}}) - \Lambda^v(\rho \ddot{\mathbf{x}} - \operatorname{div} T) - \Lambda^\varepsilon(\rho \dot{\varepsilon} + \operatorname{div} \mathbf{q} - T \cdot \operatorname{grad} \dot{\mathbf{x}}) \geq 0 \quad (9.9)$$

must hold for any fields $\{\rho(\mathbf{x}, t), \chi(\mathbf{x}, t), \theta(\mathbf{x}, t)\}$. Moreover, the Lagrange multipliers are functions of $(\rho, \theta, \mathbf{g})$.

Note that after introducing the constitutive relations (9.8) into (9.9), the inequality assumed the following form:

$$\sum_{b=1}^{26} S_b(X_a) \cdot Y_b + \sigma(X_a) \geq 0, \quad a = 1, \dots, 5, \quad (9.10)$$

where $X_a = (\rho, \theta, \mathbf{g})$ and $Y_b = (\dot{\rho}, \dot{\theta}, \dot{\mathbf{g}}, \nabla \rho, \nabla(\nabla \theta), \nabla \dot{\mathbf{x}}, \ddot{\mathbf{x}})$. Here, we have used the notation, $\nabla \mathbf{g} = \nabla(\nabla \theta)$, to emphasize the symmetry of the second gradient.

Since the inequality (9.9) must hold for any functions $\rho(\mathbf{x}, t)$, $\chi(\mathbf{x}, t)$, and $\theta(\mathbf{x}, t)$, for arbitrary values of X_a and Y_b , one can define such functions in the neighbourhood of a given point and instant, say (\mathbf{x}_0, t_0) which take those given values at (\mathbf{x}_0, t_0) (e.g., by truncated Taylor series). In other words, the inequality (9.10) must hold for any given values of X_a and Y_b .

Note that the inequality (9.10) is linear in Y_b , and the values of Y_b can be given independently of the values of S_b and σ . This implies that S_b (respecting the part involved with the symmetry of the second gradient $\nabla(\nabla \theta)$) must vanish, otherwise, it is possible to choose some values of Y_b such that the inequality is violated.

First of all, from (9.9), vanishing of coefficients of $\ddot{\mathbf{x}}$ and $\nabla \dot{\mathbf{x}}$ leads to

$$\Lambda^v = 0, \quad (9.11)$$

and

$$\Lambda^\varepsilon T - \Lambda^\rho \rho I = 0,$$

which implies that the stress tensor is a pressure only,

$$T = -p I, \quad (9.12)$$

where

$$\Lambda^\rho = -\frac{p}{\rho} \Lambda^\varepsilon. \quad (9.13)$$

With the above results, the inequality (9.9) becomes

$$\rho(\dot{\eta} - \Lambda^\varepsilon \dot{\varepsilon} + \Lambda^\varepsilon \frac{p}{\rho^2} \dot{\rho}) + (\operatorname{div} \Phi - \Lambda^\varepsilon \operatorname{Div} \mathbf{q}) \geq 0. \quad (9.14)$$

and the rest of vanishing coefficients in Y_b can now be listed below:

$$\begin{aligned} \frac{\partial \eta}{\partial \rho} - \Lambda^\varepsilon \frac{\partial \varepsilon}{\partial \rho} + \Lambda^\varepsilon \frac{p}{\rho^2} &= 0, \\ \frac{\partial \eta}{\partial \theta} - \Lambda^\varepsilon \frac{\partial \varepsilon}{\partial \theta} &= 0, \\ \frac{\partial \eta}{\partial \mathbf{g}} - \Lambda^\varepsilon \frac{\partial \varepsilon}{\partial \mathbf{g}} &= 0, \end{aligned} \quad (9.15)$$

and

$$\begin{aligned} \frac{\partial \Phi}{\partial \rho} - \Lambda^\varepsilon \frac{\partial \mathbf{q}}{\partial \rho} &= 0, \\ \left(\frac{\partial \Phi}{\partial \mathbf{g}} - \Lambda^\varepsilon \frac{\partial \mathbf{q}}{\partial \mathbf{g}} \right)_{\operatorname{sym}} &= 0. \end{aligned} \quad (9.16)$$

Finally, the inequality (9.9) reduces to the remaining one which is the entropy production density $\sigma(\rho, \theta, \mathbf{g})$ given by

$$\sigma = \left(\frac{\partial \Phi}{\partial \theta} - \Lambda^\varepsilon \frac{\partial \mathbf{q}}{\partial \theta} \right) \cdot \mathbf{g} \geq 0. \quad (9.17)$$

The relations (9.12), (9.15), (9.16), (9.17) summarize the consequence of the requirement of the entropy principle. These relations contains one essential quantity, the Lagrange multiplier Λ^ε , whose physical significance must be further investigated.

The Lagrange multiplier

The characteristic property of the Lagrange multiplier Λ^ε lies in the relation between the entropy flux and the energy flux. For a heat-conducting elastic fluid, its constitutive functions (9.8) must be isotropic functions. For energy and entropy fluxes, we have

$$\mathbf{q} = -\kappa(\rho, \theta, I_g) \mathbf{g}, \quad \Phi = -\phi(\rho, \theta, I_g) \mathbf{g},$$

where κ and ϕ are scalar functions and $I_g = \mathbf{g} \cdot \mathbf{g}$.

It follows from the relation (9.16)₂ that

$$(\phi - \Lambda^\varepsilon \kappa)I - \left(\frac{\partial \phi}{\partial I_g} - \Lambda^\varepsilon \frac{\partial \kappa}{\partial I_g} \right) \mathbf{g} \otimes \mathbf{g} = 0.$$

This is a tensor equation, and since I and $\mathbf{g} \otimes \mathbf{g}$ are independent tensor elements, their coefficients must vanish,

$$\phi = \Lambda^\varepsilon \kappa, \quad \frac{\partial \phi}{\partial I_g} - \Lambda^\varepsilon \frac{\partial \kappa}{\partial I_g} = 0,$$

which implies that

$$\Phi = \Lambda^\varepsilon \mathbf{q}. \quad (9.18)$$

Taking the gradient of (9.18) with respect to \mathbf{g} and using again the relation (9.16)₂, we obtain

$$\frac{\partial \Lambda^\varepsilon}{\partial \mathbf{g}} \cdot \mathbf{q} = \text{tr} \left(\frac{\partial \Phi}{\partial \mathbf{g}} - \Lambda^\varepsilon \frac{\partial \mathbf{q}}{\partial \mathbf{g}} \right) = 0,$$

from which it implies that Λ^ε must be independent of \mathbf{g} , since \mathbf{q} does not vanish in general. Similarly, by taking the gradient with respect to ρ and using the relation (9.16)₁, it follows immediately that Λ^ε must be independent of ρ . Therefore, we may write

$$\Lambda^\varepsilon = \Lambda(\theta). \quad (9.19)$$

Moreover, the entropy production density (9.17) becomes

$$\sigma = - \left(\frac{\partial \Lambda}{\partial \theta} \right) (\mathbf{g} \cdot \kappa \mathbf{g}) \geq 0. \quad (9.20)$$

Since the entropy production does not vanish identically in heat conducting bodies, we require that $\partial \Lambda / \partial \theta \neq 0$. Consequently, $\Lambda(\theta)$ depends monotonically on θ and hence $\Lambda(\theta)$ can also be taken as a temperature measure referred to as the *coldness*.

Universal coldness

Recall the continuity of normal components of energy and entropy fluxes at an ideal wall where the temperature is continuous. Taking two different fluid bodies in contact of the same class considered here, say, I and II, from (9.18) and the conditions (9.4) we have

$$\mathbf{q}_I \cdot \mathbf{n} = \mathbf{q}_{II} \cdot \mathbf{n}, \quad \Lambda_I \mathbf{q}_I \cdot \mathbf{n} = \Lambda_{II} \mathbf{q}_{II} \cdot \mathbf{n},$$

which implies that

$$\Lambda_I(\theta) = \Lambda_{II}(\theta).$$

In other words, the coldness is the same function of the empirical temperature θ for any fluid of the class under consideration, therefore, it is referred to as the *universal coldness* by Müller.

Absolute temperature

The universal property enables us to determine the coldness function if we can choose it for only one fluid of the kind, say, the ideal gases for simplicity, for which the constitutive relations for the pressure and the internal energy are given by

$$p = R\rho\theta, \quad \varepsilon = c(\theta - \theta_0) + \varepsilon_0,$$

where R is the gas constant, c is the specific heat, and ε_0 is the reference value at the reference temperature θ_0 . In these well-known results, we have chosen the empirical temperature as the *absolute temperature*.

To begin with, we consider the integrability condition for the entropy, by calculating the second derivatives with respect to ρ and θ from (9.15)_{1,2} to give

$$\frac{d\Lambda}{d\theta} \left(\frac{\partial \varepsilon}{\partial \rho} - \frac{p}{\rho^2} \right) = \Lambda \frac{1}{\rho^2} \frac{\partial p}{\partial \theta}.$$

For ideal gases, it reduces to

$$\frac{1}{\Lambda} \frac{d\Lambda}{d\theta} = -\frac{1}{p} \frac{dp}{d\theta} = -\frac{1}{\theta},$$

which integrates to

$$\ln \Lambda = -\ln \theta + k.$$

By conveniently taking constant k to be zero, we arrive at

$$\Lambda(\theta) = \frac{1}{\theta}. \quad (9.21)$$

In other words, we conclude that the Lagrange multiplier Λ^ε can be identified with the reciprocal of the absolute temperature for heat-conducting elastic fluids in general.

Another integrability condition, by the second derivative of the entropy with respect to θ and \mathbf{g} from (9.15) implies the independence of the internal energy and hence also the entropy on \mathbf{g} ,

$$\varepsilon = \varepsilon(\rho, \theta), \quad \eta = \eta(\rho, \theta),$$

and the relations (9.15) can be written as

$$d\eta = \frac{1}{\theta} \left(d\varepsilon - \frac{p}{\rho^2} d\rho \right). \quad (9.22)$$

This is known as the Gibbs's relation. By introducing the free energy function

$$\psi = \varepsilon - \theta\eta,$$

it becomes

$$d\psi = -\eta d\theta + \frac{p}{\rho^2} d\rho,$$

from which we obtain

$$\eta = -\frac{\partial \psi}{\partial \theta}, \quad p = \rho^2 \frac{\partial \psi}{\partial \rho}.$$

Summary of thermodynamic restrictions

Proposition. *For heat-conducting elastic fluids, general constitutive relations (9.8) reduce to*

$$\begin{aligned}
 T &= -p I, & p &= \rho^2 \frac{\partial \psi}{\partial \rho}, \\
 \varepsilon &= \psi - \theta \frac{\partial \psi}{\partial \theta}, \\
 \eta &= -\frac{\partial \psi}{\partial \theta}, \\
 \mathbf{q} &= -\kappa \mathbf{g}, \\
 \Phi &= \frac{1}{\theta} \mathbf{q},
 \end{aligned} \tag{9.23}$$

where θ is the absolute temperature. These relations depend only on two scalar functions, the free energy density

$$\psi = \psi(\rho, \theta)$$

and the heat conductivity coefficient

$$\kappa = \kappa(\rho, \theta, \mathbf{g} \cdot \mathbf{g}) \geq 0,$$

which is non-negative, as an immediate consequence of the remaining inequality (9.20).

From this example, the exploitation of the entropy principle yields, great restrictions on the constitutive functions from (9.8) to (9.23) depending only on two scalar functions, as well as the identification of the absolute temperature, from Müller-Liu procedure. Similar results have been investigated for many other classes of material bodies by either Müller-Liu or Coleman-Noll procedure. Since Coleman-Noll procedure involves more specific assumptions the formulation following it may not be the same as the one following Müller-Liu procedure. However, it turns out that in most classical theories, such as isotropic elastic solids and viscoelastic fluids, they do lead to the same results. In the advance of continuum thermodynamics, the exploitation of entropy principle has been the essential task to achieve the principal objective in the formulation of any constitutive theory.

9.4 Thermodynamic stability

Another important concept associated with the entropy inequality is the thermodynamic stability of a material body.

Stable equilibrium state. *We say that an equilibrium state is stable if any small disturbance away from it will eventually disappear and thus the original state will be restored.*

To establish a stability criterion for a material system, one may try to find a *decreasing function of time* $\mathcal{A}(t)$ from the balance laws and the entropy inequality in integral forms.

Such a function is called an *availability function* of the system, since it is the quantity available to the system for its expense in the course toward equilibrium. Such a function is usually known as a *Liapounov function* in the stability theory of dynamic systems.

To illustrate the basic ideas let us consider a supply-free body occupying a region \mathcal{V} with a fixed adiabatic boundary,

$$\mathbf{v} = \mathbf{0}, \quad \mathbf{q} \cdot \mathbf{n} = 0, \quad \Phi \cdot \mathbf{n} = 0 \quad \text{on } \partial\mathcal{V},$$

and hence the entropy inequality (9.2) and the energy balance (3.22) become

$$\frac{d}{dt} \int_{\mathcal{V}} \rho \eta \, dv \geq 0, \quad \frac{d}{dt} \int_{\mathcal{V}} \rho \left(\varepsilon + \frac{1}{2} \mathbf{v} \cdot \mathbf{v} \right) \, dv = 0. \quad (9.24)$$

In other words, the total entropy must increase in time while the total energy remains constant for a body with fixed adiabatic boundary. Statements of this kind are usually called a *stability criteria*. In the present example, from (9.24)₁ one may define the availability function \mathcal{A} of the system as

$$\mathcal{A}(t) = - \int_{\mathcal{V}} \rho \eta \, dv, \quad \frac{d\mathcal{A}}{dt} \leq 0.$$

Suppose that the region \mathcal{V} is occupied by an elastic fluid in an equilibrium state at rest with constant mass density ρ_0 and internal energy density ε_0 . Now let us consider a small disturbance of internal energy only, from the equilibrium state at the initial time,

$$\varepsilon(\mathbf{x}, 0) = \hat{\varepsilon}(\mathbf{x}),$$

for a small quantity $|\hat{\varepsilon} - \varepsilon_0|$. If we assume that the original state is stable then the perturbed state will eventually return to the original state at later time. Therefore since the total entropy must increase we conclude that

$$\int_{\mathcal{V}} \rho_0 \eta_0 \, dv \geq \int_{\mathcal{V}} \rho_0 \hat{\eta} \, dv, \quad (9.25)$$

where $\eta_0 = \eta(\varepsilon_0, \rho_0)$ and $\hat{\eta} = \eta(\hat{\varepsilon}, \rho_0)$ are the final equilibrium entropy and the perturbed initial entropy. Expanding $\hat{\eta}$ in Taylor series around the equilibrium state, we obtain from (9.25),

$$\int_{\mathcal{V}} \left\{ \frac{\partial \eta}{\partial \varepsilon} \Big|_0 \rho_0 (\hat{\varepsilon} - \varepsilon_0) + \frac{1}{2} \frac{\partial^2 \eta}{\partial \varepsilon^2} \Big|_0 \rho_0 (\hat{\varepsilon} - \varepsilon_0)^2 \right\} \, dv + o(3) \leq 0.$$

Since the total energy remains constant, we have

$$\int_{\mathcal{V}} (\rho_0 \hat{\varepsilon} - \rho_0 \varepsilon_0) \, dv = 0,$$

and hence

$$\int_{\mathcal{V}} \frac{1}{2} \frac{\partial^2 \eta}{\partial \varepsilon^2} \Big|_0 \rho_0 (\hat{\varepsilon} - \varepsilon_0)^2 \, dv = \frac{\partial^2 \eta}{\partial \varepsilon^2} \Big|_0 \int_{\mathcal{V}} \frac{1}{2} \rho_0 (\hat{\varepsilon} - \varepsilon_0)^2 \, dv \leq 0.$$

Since the integral is positive, it implies that

$$\frac{\partial^2 \eta}{\partial \varepsilon^2} \leq 0. \quad (9.26)$$

at any equilibrium state (ε_0, ρ_0) .

In order to give more suggestive meaning to the above condition for stability, note that from the Gibbs relation (9.22), we have

$$\left. \frac{\partial \eta}{\partial \varepsilon} \right|_{\rho} = \frac{1}{\theta},$$

for the entropy as a function of (ρ, ε) . Hence

$$\left. \frac{\partial^2 \eta}{\partial \varepsilon^2} \right|_{\rho} = -\frac{1}{\theta^2} \left. \frac{\partial \varepsilon}{\partial \theta} \right|_{\rho}^{-1}.$$

Therefore, in terms of the variables (ρ, θ) , the stability condition (9.26) reduces to

$$\frac{\partial \varepsilon}{\partial \theta} > 0, \quad (9.27)$$

since we have already assumed that $\varepsilon(\rho, \theta)$ is invertible with respect to θ .

In other words, the positiveness of the specific heat $c_v = \partial \varepsilon / \partial \theta$ is a consequence of thermodynamic stability.

As a second example, we consider a supply-free body with a fixed isothermal boundary,

$$\mathbf{v} = \mathbf{0}, \quad \theta = \theta_o \quad \text{on } \partial \mathcal{V},$$

and assume that the relation $\Phi = \mathbf{q}/\theta$ holds. Then the energy balance (3.22) and the entropy inequality (9.2) lead to

$$\frac{d}{dt} \int_{\mathcal{V}} \rho \left(\varepsilon + \frac{1}{2} \mathbf{v} \cdot \mathbf{v} \right) dv + \int_{\partial \mathcal{V}} \mathbf{q} \cdot \mathbf{n} da = 0,$$

$$\frac{d}{dt} \int_{\mathcal{V}} \rho \eta dv + \frac{1}{\theta_o} \int_{\partial \mathcal{V}} \mathbf{q} \cdot \mathbf{n} da \geq 0.$$

Elimination of the terms containing surface integrals from above, gives

$$\frac{d\mathcal{A}}{dt} \leq 0, \quad \mathcal{A}(t) = \int_{\mathcal{V}} \rho \left(\varepsilon - \theta_o \eta + \frac{1}{2} \mathbf{v} \cdot \mathbf{v} \right) dv. \quad (9.28)$$

In this manner we have found a decreasing function of time, the availability $\mathcal{A}(t)$, which characterizes the stability for this system. Note that

$$\int_{\mathcal{V}} \rho (\varepsilon - \theta_o \eta) dv$$

is the total free energy if $\theta = \theta_o$ throughout the body. Therefore, it follows that for a body with constant uniform temperature in a fixed region the availability \mathcal{A} reduces to the sum of the free energy and the kinetic energy.

Now, suppose that the region \mathcal{V} is occupied by an elastic fluid in an equilibrium state at rest with constant mass density ρ_0 and temperature θ_0 , and let us consider a small disturbance of mass density at the initial time,

$$\rho(\mathbf{x}, 0) = \hat{\rho}(\mathbf{x}),$$

for a small quantity $|\hat{\rho} - \rho_0|$. If we assume that the original state is stable then the perturbed state will eventually return to the original state at later time. Therefore since the free energy must decrease we conclude that

$$\int_{\mathcal{V}} \rho_0 \psi_0 \, dv \leq \int_{\mathcal{V}} \hat{\rho} \hat{\psi} \, dv,$$

where $\psi_0 = \psi(\rho_0, \theta_0)$ and $\hat{\psi} = \psi(\hat{\rho}, \theta_0)$ are the final equilibrium free energy and the perturbed initial free energy. Taking Taylor series expansion around the equilibrium state, we obtain

$$\int_{\mathcal{V}} \left\{ \left. \frac{\partial \psi}{\partial \rho} \right|_0 \hat{\rho} (\hat{\rho} - \rho_0) + \frac{1}{2} \left. \frac{\partial^2 \psi}{\partial \rho^2} \right|_0 \hat{\rho} (\hat{\rho} - \rho_0)^2 \right\} dv + o(3) \geq 0. \quad (9.29)$$

Since total mass remains constant, we have

$$\int_{\mathcal{V}} (\hat{\rho} - \rho_0) \, dv = 0,$$

and hence

$$\int_{\mathcal{V}} \hat{\rho} (\hat{\rho} - \rho_0) \, dv = \int_{\mathcal{V}} (\hat{\rho} - \rho_0)^2 \, dv = \int_{\mathcal{V}} \frac{\hat{\rho}}{\rho_0} (\hat{\rho} - \rho_0)^2 \, dv + o(3).$$

Therefore, up to the second order terms, (9.29) becomes

$$\int_{\mathcal{V}} \left\{ \frac{1}{\rho_0} \left. \frac{\partial \psi}{\partial \rho} \right|_0 + \frac{1}{2} \left. \frac{\partial^2 \psi}{\partial \rho^2} \right|_0 \right\} \hat{\rho} (\hat{\rho} - \rho_0)^2 \, dv \geq 0,$$

which as before, implies that

$$\frac{\partial \psi}{\partial \rho} + \frac{\rho}{2} \frac{\partial^2 \psi}{\partial \rho^2} \geq 0.$$

By the use of the relation (9.23)₂,

$$p = \rho^2 \frac{\partial \psi}{\partial \rho},$$

it implies

$$\frac{\partial p}{\partial \rho} \geq 0.$$

Therefore, the non-negativeness of the *isothermal compressibility* is also a consequence of thermodynamic stability.

As a conclusion, we can Summarize the above two criteria for the stability of equilibrium states as follow:

Criteria of thermodynamic stability.

- 1) *For a body with fixed adiabatic boundary and constant energy, the entropy tends to a maximum in equilibrium.*
- 2) *For a body with fixed boundary and constant uniform temperature, the sum of the free energy and the kinetic energy tends to a minimum in equilibrium.*

We have seen in this section that thermodynamic stability criteria, like the entropy principle, impose further restrictions on properties of the constitutive functions, namely, specific heat and compressibility must be positive. On the other hand, such criteria, besides being used in analyzing stability of solutions, they are the basic principles for the formulation of equilibrium solutions in terms of minimization (or maximization) problems.

10 Some problems in finite elasticity

We shall make an interesting remark on isotropic elastic bodies. From the representations

$$T = s_0 I + s_1 B + s_2 B^{-1} \quad (10.1)$$

for compressible bodies, where the material parameters s_0, s_1, s_2 are scalar functions of $(I_B, \mathbb{I}_B, \mathbb{III}_B)$, and

$$T = -pI + s_1 B + s_2 B^{-1} \quad (10.2)$$

for incompressible bodies, where s_1, s_2 are functions for (I_B, \mathbb{I}_B) and p is the indeterminate pressure, it is obvious that the stress tensor and the left Cauchy–Green tensor commute,

$$TB = BT. \quad (10.3)$$

In particular, for a deformation such that the physical components $B_{\langle 13 \rangle} = B_{\langle 23 \rangle} = 0$, so that $T_{\langle 13 \rangle}$ and $T_{\langle 23 \rangle}$ also vanish, the only non-trivial relation of (10.3) is the expression for the $\langle 12 \rangle$ -component, which reads

$$\frac{T_{\langle 11 \rangle} - T_{\langle 22 \rangle}}{T_{\langle 12 \rangle}} = \frac{B_{\langle 11 \rangle} - B_{\langle 22 \rangle}}{B_{\langle 12 \rangle}}. \quad (10.4)$$

The relations (10.3) and (10.4) between stress and deformation do not depend on any particular constitutive function, and thus they are called *universal relations* of isotropic elastic materials, compressible or incompressible. Relations of this kind are very important for experimental verification of material models, since they reflect a direct consequence from the material symmetry without having to know the constitutive function itself. In other words, if the deformation of a material model does not satisfy the relation (10.3) in its experimental observation it cannot be a material model for an isotropic elastic body.

10.1 Universal solutions in elasticity

For elastic solid bodies, the equation of motion is given by

$$\operatorname{div} T(F) + \rho \mathbf{b} = \rho \ddot{\mathbf{x}}, \quad \rho = \frac{\rho_\kappa}{|\det F|}, \quad (10.5)$$

or, in terms of referential description,

$$\operatorname{Div} T_\kappa(F) + \rho_\kappa \mathbf{b} = \rho_\kappa \ddot{\mathbf{x}}, \quad (10.6)$$

where $T_\kappa = JTF^{-T}$ is the Piola–Kirchhoff stress tensor and ρ_κ is the mass density in the reference configuration κ . The external body force \mathbf{b} is usually given in a specific problem.

An initial boundary value problem in elastic bodies is a problem of finding solutions, $\mathbf{x} = \chi(\mathbf{X}, t)$, of (10.5), or (10.6), with given initial conditions for the position $\chi(\mathbf{X}, t_0)$ and the velocity $\dot{\chi}(\mathbf{X}, t_0)$, as well as certain boundary conditions. The following three types of boundary conditions are often considered.

1) *Traction boundary condition*: The forces acting on the boundary are prescribed,

$$T_\kappa(\mathbf{X})\mathbf{n}_\kappa(\mathbf{X}) = \mathbf{f}_\kappa(\mathbf{X}), \quad \mathbf{X} \in \partial\mathcal{B}_\kappa, \quad (10.7)$$

where \mathbf{f}_κ denotes external surface forces exerted on the boundary and \mathbf{n}_κ denotes the outward unit normal in the reference configuration.

2) *Place boundary condition*: The position of the boundary is prescribed,

$$\chi(\mathbf{X}) = \mathbf{x}_0(\mathbf{X}), \quad \mathbf{X} \in \partial\mathcal{B}_\kappa, \quad (10.8)$$

where $\mathbf{x}_0(\mathbf{X})$ is a given function.

3) *Mixed boundary condition*: The traction is prescribed on a part of the boundary, while on the other part of the boundary the position is prescribed.

Boundary value problems for incompressible elastic bodies can be similarly formulated. From (10.2) the constitutive equation for the stress tensor can be written as

$$T = -p1 + \mathcal{T}(F), \quad \det F = 1,$$

and the equation of motion (10.5) becomes

$$-\text{grad } p + \text{div } \mathcal{T}(F) + \rho \mathbf{b} = \rho \ddot{\mathbf{x}}, \quad (10.9)$$

or, in referential description,

$$-\text{Grad } p + \text{Div } \mathcal{T}_\kappa(F) + \rho_\kappa \mathbf{b} = \rho_\kappa \ddot{\mathbf{x}}, \quad (10.10)$$

where p is the undeterminate hydrostatic pressure.

We shall consider some solutions, called *controllable* deformations. Such a solution is specified by a certain deformation function satisfying the equation of motion such that the body can be maintained in equilibrium by applying suitable surface traction on the boundary alone. In other words, a controllable deformation is a solution of the equilibrium equation,

$$\text{div } T(F) + \rho \mathbf{b} = 0, \quad (10.11)$$

for compressible elastic bodies, or

$$-\text{grad } p + \text{div } \mathcal{T}(F) + \rho \mathbf{b} = 0, \quad (10.12)$$

for incompressible elastic bodies. No additional boundary conditions are prescribed, instead, the boundary tractions are to be determined from (10.7). In the case of incompressible bodies, the pressure field must be suitably chosen so as to satisfy the equilibrium equation.

In general, a controllable deformation for a certain elastic material may not be controllable for a different elastic material, since the equilibrium equation depends on the

constitutive equation. If a deformation function is controllable for a certain type of elastic materials, it will be called a *universal solution* of such materials. It has been shown by Ericksen that homogeneous deformations are the only class of universal solutions for compressible isotropic elastic bodies. However, being allowed to choose a suitable pressure field in order to satisfy the equilibrium equation gives an additional freedom for possible solutions in the case of incompressible bodies. And indeed, there are several other well-known classes of universal solutions for incompressible isotropic elastic bodies. The search for universal solutions is known as “Ericksen’s problem” in the literature. We shall consider some of these solutions in the succeeding sections.

10.2 Simple Shear

We consider an isotropic elastic body subject to a deformation of simple shear given by

$$x = X + \kappa Y, \quad y = Y, \quad z = Z,$$

where the amount of shear κ is a constant, we have the left Cauchy–Green tensor

$$[B_{\langle ij \rangle}] = \begin{bmatrix} 1 + \kappa^2 & \kappa & 0 \\ \kappa & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (10.13)$$

and its inverse

$$[B_{\langle ij \rangle}]^{-1} = \begin{bmatrix} 1 & -\kappa & 0 \\ -\kappa & 1 + \kappa^2 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

so that its principal invariants are given by

$$\mathbb{I}_B = 3 + \kappa^2, \quad \mathbb{II}_B = 3 + \kappa^2, \quad \mathbb{III}_B = 1.$$

It is easy to verify that the simple shear deformation satisfies the equilibrium equation with no body force. The stress tensor is a constant tensor given by

$$\begin{aligned} [T_{\langle ij \rangle}] = & (s_0 + s_1 + s_2) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + (s_1 - s_2)\kappa \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ & + s_1\kappa^2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + s_2\kappa^2 \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \end{aligned} \quad (10.14)$$

where $s_i = s_i(3 + \kappa^2, 3 + \kappa^2, 1) = s_i(\kappa^2)$. Note that simple shear is a volume-preserving deformation and the above results are also valid for an incompressible elastic body, provided that the material parameter s_0 is replaced by $-p$, the undetermined pressure.

The shear stress $T_{\langle xy \rangle}$ on the surface, $Y = Y_0$, has the value

$$T_{\langle xy \rangle} = \hat{\mu}\kappa, \quad (10.15)$$

where

$$\hat{\mu}(\kappa^2) = s_1(\kappa^2) - s_2(\kappa^2)$$

is called the *shear modulus* of the material. For a small κ , then

$$\hat{\mu}(\kappa^2) = \mu + o(\kappa^2),$$

where $\mu = \hat{\mu}(0)$ is the classical shear modulus. Therefore, any discrepancy from the classical result for the shear stress is at least of third-order in the amount of shear κ .

For simple shear, from (10.4) and (10.13) we have the following universal relation,

$$T_{\langle xx \rangle} - T_{\langle yy \rangle} = \kappa T_{\langle xy \rangle}, \quad (10.16)$$

which can be checked immediately from (10.14). If we denote the normal stress on the slanted surface of the block, corresponding to the plane in the reference state $X = X_0$, by N (see Fig. 7), the universal relation (10.16) can be rewritten as

$$N = T_{\langle yy \rangle} - \frac{\kappa}{1 + \kappa^2} T_{\langle xy \rangle}. \quad (10.17)$$

From this expression, it is clear that in order to effect a simple shear on a rectangular block, besides shear stresses, normal stresses must also be applied on the surfaces of the block, since from the above relation the two normal stresses can not vanish simultaneously or even be equal to each other unless there is no shear at all. Moreover, the normal stress difference is a second-order effect in the amount of shear according to (10.15) and (10.17). Therefore, the normal stress difference is more significant than the discrepancy in the shear stress as an indication for the departure from the classical theory. The existence of a normal stress difference is usually known as the *Poynting effect* or simply as the *normal stress effect*.

10.3 Pure Shear

We have noticed that by applying shear stresses alone on the surface of a rectangular block, the body will tend to contract or expand if normal stresses were not supplied properly. To examine such changes quantitatively, we consider a deformation that consists of a homogeneous stretch followed by a simple shear,

$$x = \lambda_1 X + \kappa \lambda_2 Y, \quad y = \lambda_2 Y, \quad z = \lambda_3 Z.$$

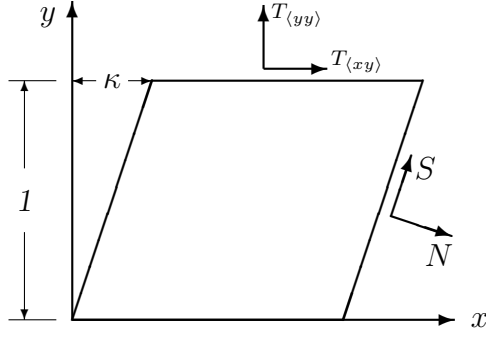


Figure 7: Simple shear

Since this is a homogeneous deformation, it is a controllable universal solution for an elastic body. The deformation gradient relative to the Cartesian coordinate system is given by

$$[F_{\langle i\alpha \rangle}] = \begin{bmatrix} \lambda_1 & \kappa\lambda_2 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}. \quad (10.18)$$

The left Cauchy–Green tensor is given by

$$[B_{\langle ij \rangle}] = \begin{bmatrix} \lambda_1^2 + \kappa^2\lambda_2^2 & \kappa\lambda_2^2 & 0 \\ \kappa\lambda_2^2 & \lambda_2^2 & 0 \\ 0 & 0 & \lambda_3^2 \end{bmatrix}, \quad (10.19)$$

and its inverse

$$[(B^{-1})_{\langle ij \rangle}] = \begin{bmatrix} \frac{1}{\lambda_1^2} & -\frac{\kappa}{\lambda_1^2} & 0 \\ -\frac{\kappa}{\lambda_1^2} & \frac{1}{\lambda_2^2} + \frac{\kappa^2}{\lambda_1^2} & 0 \\ 0 & 0 & \frac{1}{\lambda_3^2} \end{bmatrix}.$$

For isotropic elastic body, the stress tensor $T_{\langle ij \rangle}$ can be calculated from (10.1) for a compressible body or from (10.2) for an incompressible body. In particular, the shear stress on the surface, $Y = Y_0$, is given by

$$T_{\langle xy \rangle} = \kappa \left(s_1 \lambda_2^2 - s_2 \frac{1}{\lambda_1^2} \right), \quad (10.20)$$

where $s_i = s_i(\mathbb{I}_B, \mathbb{II}_B, \mathbb{III}_B)$. This holds for either compressible or incompressible ($\mathbb{III}_B = 1$) bodies. Moreover, from (10.19), the universal relation (10.4) takes the following form,

$$T_{\langle xx \rangle} - T_{\langle yy \rangle} = \frac{\lambda_1^2 - \lambda_2^2 + \kappa^2 \lambda_2^2}{\kappa \lambda_2^2} T_{\langle xy \rangle}. \quad (10.21)$$

Unlike the case of simple shear discussed in the previous section, for a fixed κ , it is now possible to determine the three constants λ_1 , λ_2 , and λ_3 in such a way that three normal stresses vanish on the surface of the block. In the case of an incompressible body, the three conditions for vanishing normal stresses can be used to determine λ_1 , λ_2 and the pressure p , while the condition of incompressibility, $\lambda_1\lambda_2\lambda_3 = 1$, determines λ_3 .

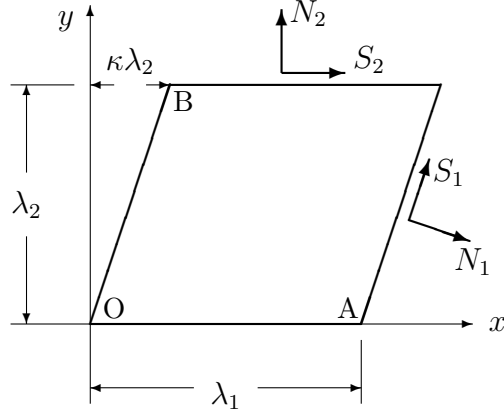


Figure 8: Pure shear

We consider a square block, with sides of unit length. From Fig. 8, we require the block be free of normal stresses,

$$N_1 = 0, \quad N_2 = T_{\langle yy \rangle} = 0, \quad T_{\langle zz \rangle} = 0.$$

The normal stress N_1 is given by

$$N_1 = T_{\langle xx \rangle} - 2\kappa T_{\langle xy \rangle} + \kappa^2 T_{\langle yy \rangle},$$

which together with vanishing of normal stress implies

$$T_{\langle xx \rangle} - 2\kappa T_{\langle xy \rangle} = 0.$$

The last condition combined with the relation (10.21) leads to

$$\lambda_1^2 = (1 + \kappa^2)\lambda_2^2. \quad (10.22)$$

It is interesting to point out that this relation is also a universal relation for elastic bodies and it admits a very simple geometric interpretation, namely, $\overline{OA} = \overline{OB}$, as shown in Fig. 8.

Furthermore, from (10.20) and (10.22), the shear stress τ is given by

$$\tau = T_{\langle xy \rangle} = \kappa(s_1\lambda_2^2 - s_2\lambda_1^{-2}), \quad S_1 = S_2 = \tau. \quad (10.23)$$

Besides the vanishing of normal stresses, the shear stresses on the surfaces, $X = X_0$ and $Y = Y_0$, are equal. Such a state of stress is called a *pure shear*.

Thus we have seen that to effect a state of pure shear on a square block, it is only necessary to apply equal shear stresses on the four surfaces. The amount of shear κ and the stretches λ_1 and λ_2 are adjusted in such a way that the length of the four sides remains the same. A square block becoming a rhombic block is also what one would expect intuitively in a pure shear.

To determine the stretches for the amount of shear τ , explicit constitutive expressions would be needed. As an example, we shall consider the Mooney–Rivlin material with constant material parameters s_1 and s_2 . In the case of two-dimensional deformation, so that the thickness in x_3 -direction remains unchanged, i.e.,

$$\lambda_3 = 1, \quad \text{and} \quad \lambda_1 \lambda_2 = 1,$$

by incompressibility. It follows from (10.22) and (10.23) that

$$\lambda_1 = (1 - \tau^2(s_1 - s_2)^{-2})^{-1/4}. \quad (10.24)$$

Note that according to the requirement, $s_1 > s_2$, the value of λ_1 is greater than 1 so that the side of the rhombus (losango) is greater than the side of original square.

10.4 bending of a rectangular block

We shall consider the case of bending a rectangular block into a circular section for an incompressible isotropic elastic material.

This problem can be described by the following volume-preserving deformation:

$$r = \sqrt{2aX + b}, \quad \theta = cY, \quad z = dZ, \quad \text{for} \quad acd = 1,$$

where (X, Y, Z) are the Cartesian coordinates for the initial configuration and (r, θ, z) are the cylindrical coordinates for the deformed configuration of the body.

For this deformation, the deformation gradient is given by

$$[F^i_{\alpha}] = \begin{bmatrix} \frac{\partial r}{\partial X} & \frac{\partial r}{\partial Y} & \frac{\partial r}{\partial Z} \\ \frac{\partial \theta}{\partial X} & \frac{\partial \theta}{\partial Y} & \frac{\partial \theta}{\partial Z} \\ \frac{\partial z}{\partial X} & \frac{\partial z}{\partial Y} & \frac{\partial z}{\partial Z} \end{bmatrix} = \begin{bmatrix} ar^{-1} & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & d \end{bmatrix}.$$

In terms of physical components, we have

$$[F_{\langle i\alpha \rangle}] = \begin{bmatrix} ar^{-1} & 0 & 0 \\ 0 & cr & 0 \\ 0 & 0 & d \end{bmatrix}, \quad [B_{\langle ij \rangle}] = \begin{bmatrix} a^2r^{-2} & 0 & 0 \\ 0 & c^2r^2 & 0 \\ 0 & 0 & d^2 \end{bmatrix}. \quad (10.25)$$

Obviously, $\det F = \det B = 1$ in physical components, as required by incompressibility.

This deformation belongs to a class of universal solutions for incompressible isotropic elastic bodies given by the constitutive equations (10.2). In other words, this deformation satisfies the equation of equilibrium (10.9) with no body force \mathbf{b} and $\ddot{\mathbf{x}} = 0$,

$$\text{grad } p - \text{div}(s_1 B + s_2 B^{-1}) = 0. \quad (10.26)$$

Since the Cauchy–Green tensor B is a function of r only, the equilibrium equation in the deformed cylindrical coordinate can be written as (see Sec. 10.6)

$$\begin{aligned} \frac{\partial T_{\langle rr \rangle}}{\partial r} + \frac{1}{r}(T_{\langle rr \rangle} - T_{\langle \theta \theta \rangle}) &= 0, \\ \frac{\partial p}{\partial \theta} = 0, \quad \frac{\partial p}{\partial z} &= 0, \end{aligned}$$

which imply that the pressure $p = p(r)$, and

$$T_{\langle rr \rangle} = - \int \frac{1}{r}(T_{\langle rr \rangle} - T_{\langle \theta \theta \rangle}) dr. \quad (10.27)$$

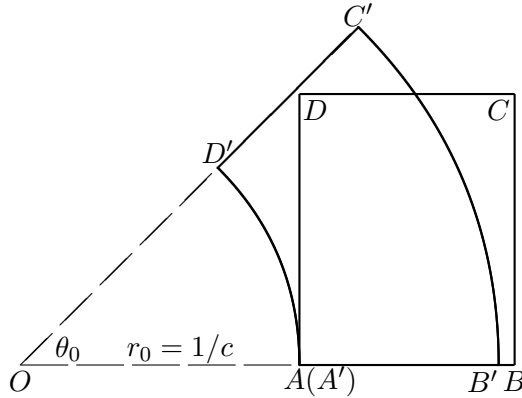


Figure 9: Bending into a circular section

In order to integrate this equation explicitly, we shall consider Mooney–Rivlin material model with constant s_1, s_2 , and for simplicity, taking $d = 1$ so that the deformation

$$r = \sqrt{\frac{2}{c}X + \frac{1}{c^2}}, \quad \theta = cY, \quad z = Z, \quad (10.28)$$

is essentially two-dimensional without change of thickness, as shown in Fig 9. In this case, (10.26) becomes

$$T_{\langle rr \rangle} = -(s_1 - s_2) \int \frac{1}{r} \left(\frac{1}{c^2 r^2} - c^2 r^2 \right) dr.$$

Upon integration, we obtain

$$T_{\langle rr \rangle} = \frac{1}{2}(s_1 - s_2) \left(\frac{1}{c^2 r^2} + c^2 r^2 \right) + K,$$

where K is an integration constant.

Note that from (10.28), we have

$$c^2 r^2 = 1 + 2cX.$$

If we impose a boundary condition that the inner surface of the circular block ($X = 0$) is stress free, i.e.,

$$T_{\langle rr \rangle}|_{X=0} = (s_1 - s_2) + K = 0, \quad (10.29)$$

which implies $K = s_2 - s_1$.

Finally, from (10.2), we obtain the exact solutions for the stresses and the pressure:

$$\begin{aligned} T_{\langle rr \rangle} &= \frac{1}{2}(s_1 - s_2) \left(c^2 r^2 + \frac{1}{c^2 r^2} - 2 \right), \\ T_{\langle \theta\theta \rangle} &= \frac{1}{2}(s_1 - s_2) \left(3c^2 r^2 - \frac{1}{c^2 r^2} - 2 \right), \\ p &= \frac{1}{2} s_1 \left(\frac{1}{c^2 r^2} - c^2 r^2 + 2 \right) + \frac{1}{2} s_2 \left(3c^2 r^2 + \frac{1}{c^2 r^2} - 2 \right). \end{aligned}$$

In order to effect the deformation from a rectangular block into a circular block, besides the condition (10.29), other proper boundary conditions must be applied.

10.5 Deformation of a cylindrical annulus

For inflation, torsion, and extension, the deformation functions, from reference coordinate (R, Θ, Z) to deformed coordinate (r, θ, z) in cylindrical coordinates, given by

$$r = \sqrt{A + BR^2}, \quad \theta = C\Theta + DZ, \quad z = E\Theta + FZ,$$

with the condition $B(CF - DE) = 1$ for incompressibility, is a family of universal solutions for any incompressible elastic bodies.

For the case with no torsion and no extension, we take $B = C = F = 1$ and $D = E = 0$, and the deformation functions become

$$r = \sqrt{R^2 + A}, \quad \theta = \Theta, \quad z = Z. \quad (10.30)$$

In this case, we have the deformation gradient

$$[F^i_\alpha] = \begin{bmatrix} Rr^{-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

In the physical components, it becomes

$$[F_{\langle i\alpha \rangle}] = \begin{bmatrix} Rr^{-1} & 0 & 0 \\ 0 & rR^{-1} & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and $B = FF^T$ is given by

$$[B_{\langle ij \rangle}] = \begin{bmatrix} \frac{R^2}{r^2} & 0 & 0 \\ 0 & \frac{r^2}{R^2} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{r^2 - A}{r^2} & 0 & 0 \\ 0 & \frac{r^2}{r^2 - A} & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (10.31)$$

The equilibrium equation, $-\operatorname{div} T = 0$, in cylindrical coordinates for this case can be written as

$$\begin{aligned} \frac{\partial T_{\langle rr \rangle}}{\partial r} + \frac{1}{r}(T_{\langle rr \rangle} - T_{\langle \theta\theta \rangle}) &= 0, \\ \frac{\partial p}{\partial \theta} = 0, \quad \frac{\partial p}{\partial z} &= 0. \end{aligned} \quad (10.32)$$

Therefore the solutions for the stress T and the pressure p are functions of r only.

For incompressible isotropic elastic materials, $T = -pI + s_1B + s_2B^{-1}$, where s_1 and s_2 are functions of (I_B, \mathbb{I}_B) , from the above equation, we have

$$\begin{aligned} T_{\langle rr \rangle} &= - \int \frac{1}{r}(T_{\langle rr \rangle} - T_{\langle \theta\theta \rangle})dr \\ &= - \int \frac{1}{r} \left(s_1 \left(\frac{r^2 - A}{r^2} - \frac{r^2}{r^2 - A} \right) + s_2 \left(\frac{r^2}{r^2 - A} - \frac{r^2 - A}{r^2} \right) \right) dr. \end{aligned}$$

For Mooney–Rivlin material for which s_1 and s_2 are material constants, we can integrate the above equation and obtain

$$T_{\langle rr \rangle}(r) = \frac{1}{2}(s_1 - s_2) \left(\ln \frac{r^2 - A}{r^2} - \frac{A}{r^2} \right) + K, \quad (10.33)$$

where K is the integration constant.

On the other hand, since

$$T_{\langle rr \rangle} = -p + s_1 \frac{r^2 - A}{r^2} + s_2 \frac{r^2}{r^2 - A},$$

from (10.33), we obtain

$$p(r) = s_1 \frac{r^2 - A}{r^2} + s_2 \frac{r^2}{r^2 - A} - \frac{1}{2}(s_1 - s_2) \left(\ln \frac{r^2 - A}{r^2} - \frac{A}{r^2} \right) - K. \quad (10.34)$$

Similarly, from

$$T_{\langle\theta\theta\rangle} = -p + s_1 \frac{r^2}{r^2 - A} + s_2 \frac{r^2 - A}{r^2},$$

we obtain

$$T_{\langle\theta\theta\rangle}(r) = \frac{1}{2}(s_1 - s_2) \left(\ln \frac{r^2 - A}{r^2} + \frac{2r^2}{r^2 - A} - \frac{2r^2 - A}{r^2} \right) + K. \quad (10.35)$$

Summary of the problem

We consider the deformation function,

$$r = \sqrt{R^2 + A}, \quad \theta = \Theta, \quad z = Z,$$

for reference region $R_0 \leq R \leq R_1$ and deformed region $r_0 \leq r \leq r_1$.

Equilibrium solution for the stress $T = -pI + s_1B + s_2B^{-1}$ of Mooney–Rivlin material is given by

$$\begin{aligned} T_{\langle rr\rangle}(R) &= \frac{1}{2}(s_1 - s_2) \left(\ln \frac{R^2}{R^2 + A} - \frac{A}{R^2 + A} \right) + K, \\ T_{\langle\theta\theta\rangle}(R) &= \frac{1}{2}(s_1 - s_2) \left(\ln \frac{R^2}{R^2 + A} + \frac{2(R^2 + A)}{R^2} - \frac{2R^2 + A}{R^2 + A} \right) + K, \\ p(R) &= s_1 \frac{R^2}{R^2 + A} + s_2 \frac{R^2 + A}{R^2} - \frac{1}{2}(s_1 - s_2) \left(\ln \frac{R^2}{R^2 + A} - \frac{A}{R^2 + A} \right) + K. \end{aligned}$$

The solution contains two free parameters, namely, A and K , to be determined from boundary conditions. One can prescribe the boundary conditions at the inner and outer surfaces of the cylinder, for example,

$$p_0 = -T_{\langle rr\rangle}(R_0), \quad p_1 = -T_{\langle rr\rangle}(R_1).$$

which lead to a system of two equations for the two constants A and K :

$$\begin{aligned} \frac{1}{2}(s_1 - s_2) \left(\ln \frac{R_0^2}{R_0^2 + A} - \frac{A}{R_0^2 + A} \right) + K &= -p_0, \\ \frac{1}{2}(s_1 - s_2) \left(\ln \frac{R_1^2}{R_1^2 + A} - \frac{A}{R_1^2 + A} \right) + K &= -p_1. \end{aligned}$$

The system is difficult to solve analytically, but numerically, it can easily be solved.

10.6 Appendix: Divergence of a tensor field

The divergence in physical components of a symmetric tensor field T :

a) Cartesian coordinate system (x, y, z) :

$$\begin{aligned}(\operatorname{div} T)_{\langle x \rangle} &= \frac{\partial T_{\langle xx \rangle}}{\partial x} + \frac{\partial T_{\langle xy \rangle}}{\partial y} + \frac{\partial T_{\langle xz \rangle}}{\partial z}, \\(\operatorname{div} T)_{\langle y \rangle} &= \frac{\partial T_{\langle xy \rangle}}{\partial x} + \frac{\partial T_{\langle yy \rangle}}{\partial y} + \frac{\partial T_{\langle yz \rangle}}{\partial z}, \\(\operatorname{div} T)_{\langle z \rangle} &= \frac{\partial T_{\langle xz \rangle}}{\partial x} + \frac{\partial T_{\langle yz \rangle}}{\partial y} + \frac{\partial T_{\langle zz \rangle}}{\partial z}.\end{aligned}$$

b) Cylindrical coordinate system (r, θ, z) :

$$\begin{aligned}(\operatorname{div} T)_{\langle r \rangle} &= \frac{\partial T_{\langle rr \rangle}}{\partial r} + \frac{1}{r} \frac{\partial T_{\langle r\theta \rangle}}{\partial \theta} + \frac{\partial T_{\langle rz \rangle}}{\partial z} + \frac{T_{\langle rr \rangle} - T_{\langle \theta\theta \rangle}}{r}, \\(\operatorname{div} T)_{\langle \theta \rangle} &= \frac{\partial T_{\langle r\theta \rangle}}{\partial r} + \frac{1}{r} \frac{\partial T_{\langle \theta\theta \rangle}}{\partial \theta} + \frac{\partial T_{\langle \theta z \rangle}}{\partial z} + \frac{2}{r} T_{\langle r\theta \rangle}, \\(\operatorname{div} T)_{\langle z \rangle} &= \frac{\partial T_{\langle rz \rangle}}{\partial r} + \frac{1}{r} \frac{\partial T_{\langle \theta z \rangle}}{\partial \theta} + \frac{\partial T_{\langle zz \rangle}}{\partial z} + \frac{1}{r} T_{\langle rz \rangle}.\end{aligned}$$

c) Spherical coordinate system (r, θ, ϕ) :

$$\begin{aligned}(\operatorname{div} T)_{\langle r \rangle} &= \frac{\partial T_{\langle rr \rangle}}{\partial r} + \frac{1}{r} \frac{\partial T_{\langle r\theta \rangle}}{\partial \theta} + \frac{1}{r \sin \theta} \frac{\partial T_{\langle r\phi \rangle}}{\partial \phi} \\&\quad + \frac{1}{r} \left(2T_{\langle rr \rangle} - T_{\langle \theta\theta \rangle} - T_{\langle \phi\phi \rangle} + \cot \theta T_{\langle r\theta \rangle} \right), \\(\operatorname{div} T)_{\langle \theta \rangle} &= \frac{\partial T_{\langle r\theta \rangle}}{\partial r} + \frac{1}{r} \frac{\partial T_{\langle \theta\theta \rangle}}{\partial \theta} + \frac{1}{r \sin \theta} \frac{\partial T_{\langle \theta\phi \rangle}}{\partial \phi} \\&\quad + \frac{1}{r} \left(3T_{\langle r\theta \rangle} + \cot \theta (T_{\langle \theta\theta \rangle} - T_{\langle \phi\phi \rangle}) \right), \\(\operatorname{div} T)_{\langle \phi \rangle} &= \frac{\partial T_{\langle r\phi \rangle}}{\partial r} + \frac{1}{r} \frac{\partial T_{\langle \theta\phi \rangle}}{\partial \theta} + \frac{1}{r \sin \theta} \frac{\partial T_{\langle \phi\phi \rangle}}{\partial \phi} \\&\quad + \frac{1}{r} \left(3T_{\langle r\phi \rangle} + 2 \cot \theta T_{\langle \theta\phi \rangle} \right).\end{aligned}$$

11 Wave propagation in elastic bodies

We shall consider small amplitude waves, which are motions of small deformation in a material body. It can be treated in a relatively simple classical approach of harmonic wave propagation. The material body on which the wave propagates is generally deformed so that the body is effectively anisotropic. To study small amplitude waves on a deformed body, we shall first formulate a theory of small deformation on a deformed body.

11.1 Small deformations on a deformed body

To account for propagation of small amplitude waves into a prestressed body, we shall formulate a theory of small deformation on a finitely deformed body. Although this is a well-known problem in finite elasticity³, our approach will not be formulated in Eulerian description nor in Lagrangian description as usual, rather, the problem can be elegantly formulated in *relative description* (see Chapter 2), relative to the pre-deformed state of the body as the reference configuration.

The relative description

Let κ_0 be a reference configuration of the body \mathcal{B} , and κ_t be its deformed configuration at the *present* time t , Let

$$\mathbf{x} = \chi(X, t), \quad X \in \kappa_0(\mathcal{B}),$$

and

$$F(X, t) = \nabla_X(\chi(X, t)),$$

be the deformation and the deformation gradient from κ_0 to κ_t .

Now, at some time τ , consider the deformed configuration κ_τ , and

$$\boldsymbol{\xi} = \chi(X, \tau) := \boldsymbol{\xi}_t(\mathbf{x}, \tau) \in \kappa_\tau(\mathcal{B}), \quad \mathbf{x} = \chi(X, t) \in \kappa_t(\mathcal{B}). \quad (11.1)$$

$\boldsymbol{\xi}_t(\mathbf{x}, \tau)$ is called the *relative motion* with respect to the present configuration at time t , and

$$F_t(\mathbf{x}, \tau) = \nabla_{\mathbf{x}} \boldsymbol{\xi}_t(\mathbf{x}, \tau),$$

is called the relative deformation gradient. We also define the relative displacement vector as

$$\mathbf{u}(\mathbf{x}, \tau) = \boldsymbol{\xi}_t(\mathbf{x}, \tau) - \mathbf{x}, \quad (11.2)$$

³ Green, A.E.; Rivlin R.S.; Shield, R.T.: General theory of small elastic deformations superposed on finite deformations. *Pro. Roy. Soc. London, Ser. A*, **211**, 128-154, (1952)

and

$$H(\mathbf{x}, \tau) = \nabla_{\mathbf{x}} \mathbf{u}(\mathbf{x}, \tau) = F_t(\mathbf{x}, \tau) - I, \quad (11.3)$$

is the relative displacement gradient at time τ with respect to the *present* configuration κ_t (emphasize, not κ_0), and I stands for the identity tensor.

Note that from (11.1),

$$\nabla_{\mathbf{x}} \boldsymbol{\xi}_t(\mathbf{x}, \tau) = \nabla_X(\chi(X, \tau)) \nabla_X(\chi(X, t))^{-1} = F(X, \tau)F(X, t)^{-1},$$

hence, by the use of (11.3) we have

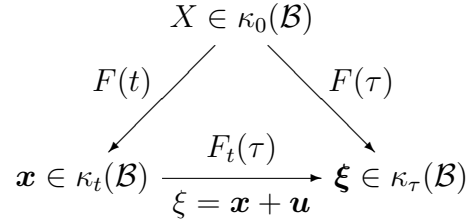
$$H(\mathbf{x}, \tau) = F(X, \tau)F(X, t)^{-1} - I,$$

or simply as

$$F_t(\tau) = I + H(\tau) \quad \text{and} \quad F(\tau) = (I + H(\tau))F(t). \quad (11.4)$$

In these expressions and hereafter, we shall often denote a function F as $F(t)$ to emphasize its value at time t when its spatial variable is self-evident.

We can represent the deformation and deformation gradient schematically in the following diagram:



Furthermore, from (11.1) and (11.2) (or from the above diagram), we have

$$\boldsymbol{\chi}(X, \tau) = \boldsymbol{\chi}(X, t) + \mathbf{u}(\boldsymbol{\chi}(X, t), \tau). \quad (11.5)$$

By taking the derivatives with respect to τ , we obtain the velocity and the acceleration of the motion at time τ ,

$$\dot{\boldsymbol{x}}(X, \tau) = \frac{\partial \mathbf{u}(\mathbf{x}, \tau)}{\partial \tau} = \dot{\mathbf{u}}(\mathbf{x}, \tau), \quad \ddot{\boldsymbol{x}}(X, \tau) = \ddot{\mathbf{u}}(\mathbf{x}, \tau). \quad (11.6)$$

Note that since $\mathbf{x} = \chi(X, t)$ is independent of τ , in relative motion description, the partial derivative with respect to τ keeping \mathbf{x} fixed is nothing but the material time derivative.

The advantage of using relative Lagrangian formulation is that it enable us to linearize constitutive equation relative to the present state and hence with a successive method of Euler's type, we can approximate the nonlinear constitutive functions for large deformations.

Linearized constitutive equations

Let κ_0 be the preferred reference configuration of a viscoelastic body \mathcal{B} , and let the Cauchy stress $T(X, t)$ be given by the constitutive equation in the configuration κ_0 ,

$$T(X, t) = \mathcal{T}(F(X, t)). \quad (11.7)$$

For large deformations, the constitutive function \mathcal{T} is generally a nonlinear function of the deformation gradient F .

We shall regard the present configuration κ_t as an *updated* reference configuration, and consider a small deformation relative to the present state $\kappa_t(\mathcal{B})$ at time $\tau = t + \Delta t$, for small enough time interval Δt . In other words, we shall assume that the relative displacement gradient H is small, $|H| \ll 1$, so that we can linearize the constitutive equation (11.7) at time τ relative to the updated reference configuration at time t , namely,

$$T(\tau) = \mathcal{T}(F(\tau)) = \mathcal{T}(F(t)) + \partial_F \mathcal{T}(F(t))[F(\tau) - F(t)],$$

or by the use of (11.4),

$$T(\tau) = T(t) + \partial_F \mathcal{T}(F(t))[H(\tau)F(t)] + o(2),$$

where $o(2)$ represents higher order terms in the small displacement gradient $|H|$.

The linearized constitutive equation can now be written as

$$T(\tau) = T(t) + L(F(t))[H(\tau)], \quad (11.8)$$

where

$$L(F)[H] := \partial_F \mathcal{T}(F)[HF], \quad (11.9)$$

define the fourth order elasticity tensor $L(F)$ relative to the present configuration κ_t .

The above general definition of the elasticity for any constitutive class of elastic materials $T = \mathcal{T}(F)$, relative to the updated present configuration, will be explicitly determined in the following sections for a particular class, namely a Mooney–Rivlin type materials.

Compressible and nearly incompressible bodies

For an elastic body, the constitutive equation (11.7) relative to the preferred reference configuration κ_0 can be written as

$$T = \mathcal{T}(F) = -p(F)I + \tilde{\mathcal{T}}(F). \quad (11.10)$$

However, for an incompressible body, the pressure p depends also on the boundary conditions, and because it cannot be determined from the deformation of the body alone, it is called an *indeterminate pressure*, which is an independent variable in addition to the displacement vector variable for boundary value problems.

For compressible bodies, we shall assume that the pressure depend on the deformation gradient only through the determinant, or by the use of the mass balance, depend only on the mass density,

$$p = \hat{p}(\det F) = p(\rho), \quad \rho = \frac{\rho_0}{\det F},$$

where ρ_0 is the mass density in the reference configuration κ_0 . We have

$$\begin{aligned} \rho(\tau) - \rho(t) &= \rho_0(\det F(\tau)^{-1} - \det F(t)^{-1}) = \rho(t)(\det(F(\tau)^{-1}F(t)) - 1) \\ &= \rho(t)(\det(I + H(\tau))^{-1} - 1) = -\rho(t) \operatorname{tr} H(\tau) + o(2), \end{aligned}$$

in which the relation (11.4) has been used.

Therefore, it follows that

$$p(\tau) - p(t) = \left(\frac{dp}{d\rho} \right)_t (\rho(\tau) - \rho(t)) + o(2) = - \left(\rho \frac{dp}{d\rho} \right)_t \operatorname{tr} H(\tau) + o(2),$$

or

$$p(\tau) = p(t) - \beta \operatorname{tr} H(\tau) + o(2), \quad (11.11)$$

where $\beta := \left(\rho \frac{dp}{d\rho} \right)_t$ is a material parameter evaluated at the present time t .

From (11.10) and (11.4), let

$$C(F(t))[H(\tau)] := \partial_F \tilde{\mathcal{T}}(F(t))[F(\tau) - F(t)] = \partial_F \tilde{\mathcal{T}}(F(t))[H(\tau)F(t)],$$

then from (11.9)₁, the elasticity tensor becomes

$$L(F)[H] = \beta(\operatorname{tr} H)I + C(F)[H]. \quad (11.12)$$

We call a body nearly incompressible if its density is nearly insensitive to the change of pressure. Therefore, if we regard the density as a function of pressure, $\rho = \rho(p)$, then its derivative with respect to the pressure is nearly zero. In other words, for nearly incompressible bodies, we shall assume that β is a material parameter much greater than 1,

$$\beta \gg 1. \quad (11.13)$$

Note that for compressible or nearly incompressible body, the elasticity tensor does not contain the pressure explicitly. It is only a function of the deformation gradient and the material parameter β at the present time t .

Mooney–Rivlin material

For Mooney–Rivlin material,

$$T = \mathcal{T}(F) = -p(F)I + s_1 B + s_2 B^{-1}, \quad (11.14)$$

The material parameters s_1 and s_2 are assumed to be constants. After taking the gradients of $\tilde{\mathcal{T}}(F)$ with respect to F at $(F, 0)$, from (11.8), we have

$$T(\tau) = T(t) + L(F(t))[H(\tau)], \quad (11.15)$$

where

$$L(F)[H] = \beta(\text{tr } H)I + s_1(HB + BH^T) - s_2(B^{-1}H + H^T B^{-1}). \quad (11.16)$$

11.2 The equation of motion in relative description

Let κ be a reference configuration of the body \mathcal{B} , then in the Lagrangian formulation, we can write the equation of motion at time τ as

$$\rho_\kappa(\mathbf{X})\ddot{\mathbf{x}}(\mathbf{X}, \tau) - \text{Div } T_\kappa(\mathbf{X}, \tau) = \rho_\kappa(\mathbf{X})\mathbf{g}(\mathbf{X}, \tau), \quad \mathbf{X} \in \kappa(\mathcal{B}),$$

where $T_\kappa(\tau)$ is the Piola–Kirchhoff stress tensor at time τ relative to the reference configuration κ , and $\mathbf{g}(\tau)$ is the body force. The operator (Div) stands for the divergence with respect to the coordinate system (\mathbf{X}) in $\kappa(\mathcal{B})$, and the overhead dot ($\dot{\cdot}$) is the material time derivative with respect to time variable τ .

Instead of the fixed reference configuration κ , we can rewrite the equation relative to the configuration κ_t at the present time t as the reference configuration. By a simple substitution, it becomes

$$\rho(\mathbf{x}, t)\ddot{\mathbf{x}}(\mathbf{x}, \tau) - \text{div } T_t(\mathbf{x}, \tau) = \rho(\mathbf{x}, t)\mathbf{g}(\mathbf{x}, \tau), \quad \mathbf{x} \in \kappa_t(\mathcal{B}),$$

where $T_t(\tau) = T_{\kappa_t}(\tau)$ is the Piola–Kirchhoff stress tensor at time τ relative to the present configuration κ_t . The operator (div) stands for the divergence with respect to the coordinate system (\mathbf{x}) in the present configuration.

This is the equation of motion in the relative description. We call this a *relative Lagrangian* formulation in contrast to the usual (total) Lagrangian and Eulerian formulations.

The (first) Piola–Kirchhoff stress tensor at time τ relative to the present configuration at time t , $T_t(\tau)$, is given by

$$\begin{aligned} T_t(\tau) &:= \det F_t(\tau) T(\tau) F_t(\tau)^{-T} = \det(I + H) T(\tau) (I + H)^{-T} \\ &= \det(I + H) (T(t) + L(F)[H]) (I + H)^{-T} \\ &= (I + \text{tr } H) (T(t) + L(F)[H]) (I - H^T) + o(2) \\ &= T(t) + (\text{tr } H) T(t) - T(t) H^T + L(F)[H] + o(2). \end{aligned}$$

We can write the linearized Piola–Kirchhoff stress as

$$T_t(\tau) = T(t) + (\text{tr } H(\tau)) T(t) - T(t) H(\tau)^T + L(F(t))[H(\tau)].$$

Note that when $\tau \rightarrow t$, $H \rightarrow 0$, and hence $T_t(\tau) \rightarrow T(t)$, therefore, the Piola-Kirchhoff stress, becomes the Cauchy stress at the present time t .

We can also write

$$T_t(\tau) = T(t) + K(F(t), T(t))[H(\tau)], \quad (11.17)$$

where the Piola-Kirchhoff elasticity tensor is defined as

$$K(F, T)[H] = (\text{tr } H)T - TH^T + L(F)[H].$$

For Mooney–Rivlin materials, from (11.16), we have

$$\begin{aligned} K(F, T)[H] &= (\text{tr } H)T - TH^T + \beta(\text{tr } H)I \\ &\quad + s_1(HB + BH^T) - s_2(B^{-1}H + H^TB^{-1}). \end{aligned} \quad (11.18)$$

Linearized equation of motion

We shall assume that at the present time t , the deformation gradient F with respect to the preferred reference configuration κ_0 and the Cauchy stress T are known, and that $\tau = t + \Delta t$ with small enough Δt . Then from (11.17), The equation of motion can be written as

$$\rho(t)\ddot{\mathbf{u}}(\tau) - \text{div}\left(K(F(t), T(t))[\nabla\mathbf{u}(\tau)]\right) = \rho(t)\mathbf{g}(\tau) + \text{div } T(t), \quad (11.19)$$

which is a linear partial differential equation for the relative displacement vector $\mathbf{u}(\mathbf{x}, \tau)$.

If we further assume that the body is in equilibrium at time t , i.e.,

$$-\text{div } T(t) = \rho(t)\mathbf{g}(t),$$

then the equation becomes

$$\rho(t)\ddot{\mathbf{u}}(\tau) - \text{div}\left(K(F(t), T(t))[\nabla\mathbf{u}(\tau)]\right) = \rho(t)(\mathbf{g}(\tau) - \mathbf{g}(t)).$$

Note that the right-hand side is the incremental body force, which vanishes if the body force is time-independent. In this case, we have

$$\rho(t)\ddot{\mathbf{u}}(\tau) - \text{div}\left(K(F(t), T(t))[\nabla\mathbf{u}(\tau)]\right) = 0. \quad (11.20)$$

This is the same partial differential equation of linear elasticity, except that the elasticity tensor is not constant, and rather, it depends on the deformed state of the reference body.

Remark: On problem of large deformation

The above formulation of linearized problem in relative description can be successively applied by updating the reference state at each time step, so that the accumulation of successive small deformation can be used as an approximation for a problem of large finite deformation. This is referred to as *Successive Linear Approximation* and has been used for the solutions in numerical schemes in some problems in finite elasticity discussed in the previous chapter⁴.

11.3 Plane harmonic waves in a deformed elastic body

We consider a plane harmonic wave defined as

$$\mathbf{u}(\mathbf{x}, \tau) = \hat{s} \mathbf{a} e^{k_i \mathbf{n} \cdot \mathbf{x}} \cos(\omega \tau - k_r \mathbf{n} \cdot \mathbf{x} + \phi),$$

where \hat{s} is the (scalar) amplitude, \mathbf{a} is the unit amplitude vector, ω is the frequency, the unit vector \mathbf{n} is the wave normal, k_i is called the attenuation factor, k_r is called the wave number, and ϕ is a phase constant. The phase $(\omega \tau - k_r \mathbf{n} \cdot \mathbf{x} + \phi)$ is constant on a plane perpendicular to the wave normal vector \mathbf{n} and ω/k_r is the phase speed.

It is more convenient to express a harmonic wave in the complex form ($i = \sqrt{-1}$),

$$\mathbf{u}(\mathbf{x}, \tau) = \text{Re}(s \mathbf{a} e^{i(\omega \tau - k \mathbf{n} \cdot \mathbf{x})}),$$

where we have defined the complex amplitude s and the complex wave number k by

$$s = \hat{s} e^{i\phi}, \quad k = k_r + i k_i,$$

Note that physically, we are only interested in the real part of the expression, however, for mathematical convenience, we shall write

$$\mathbf{u}(\mathbf{x}, \tau) = s \mathbf{a} e^{i(\omega \tau - k \mathbf{n} \cdot \mathbf{x})}, \quad (11.21)$$

and hence, we have the phase speed U ,

$$U = \frac{\omega}{\text{Re}(k)}.$$

Propagation condition

We consider harmonic wave propagations in a pre-deformed elastic body. In the relative Lagrangian formulation, from (11.20) we can write the equation of motion in component form,

$$\rho \ddot{u}_i = \frac{\partial}{\partial x_j} \left(K_{ijkl} \frac{\partial u_k}{\partial x_l} \right), \quad (11.22)$$

⁴I-Shih Liu, Successive linear approximation for boundary value problems of nonlinear elasticity in relative-descriptive formulation, International Journal of Engineering Science, 49 (2011) 635-645.

where $\mathbf{u}(\mathbf{x}, \tau)$ is the relative displacement vector from the reference state at $\mathbf{x} \in \kappa_t(\mathcal{B})$ and $K(F(\mathbf{x}, t), T(\mathbf{x}, t))$ is the elasticity tensor defined in (11.17). Of course, the values of $(F(\mathbf{x}, t), T(\mathbf{x}, t))$ are assumed to be known at the reference time t , and furthermore, for simplicity, we shall assume that $(F(\mathbf{x}, t), T(\mathbf{x}, t))$ are homogeneous in a small neighborhood of $\mathbf{x} \in \kappa_t(\mathcal{B})$ so that $K(F, T)$ is constant at (\mathbf{x}, t) .

For a solution of harmonic wave, from (11.21), the relative displacement vector can be represented as

$$u_i = s a_i e^{i(\omega\tau - kn_j x_j)},$$

where we have represent the amplitude vector $\tilde{\mathbf{u}} = s \mathbf{a}$, for scalar amplitude s and unit amplitude vector \mathbf{a} .

Upon substitution into (11.22), it leads to

$$\rho \left(\frac{\omega}{k} \right)^2 a_i = K_{ijkl} a_k n_l n_j = Q_{ik} a_k,$$

where the acoustic tensor Q is defined as

$$Q_{ik} = K_{ijkl} n_j n_l. \quad (11.23)$$

The above relation can be written in direct notations as

$$\rho \left(\frac{\omega}{k} \right)^2 \mathbf{a} = K[\mathbf{a} \otimes \mathbf{n}] \mathbf{n} = Q(\mathbf{n}) \mathbf{a}. \quad (11.24)$$

For a given propagation direction \mathbf{n} , the amplitude vector \mathbf{a} must be a real eigenvector of $Q(\mathbf{n})$ and the eigenvalue can be obtained from the relation,

$$\rho \left(\frac{\omega}{k} \right)^2 = \mathbf{a} \cdot Q(\mathbf{n}) \mathbf{a} = \mathbf{a} \cdot K[\mathbf{a} \otimes \mathbf{n}] \mathbf{n}. \quad (11.25)$$

Note that if the right-hand side of (11.25) is non-negative, then the wave number k is real and ω/k is the phase speed, denoted by U_{na} , for the wave propagating in the direction \mathbf{n} with amplitude in the direction \mathbf{a} without attenuation ($\text{Im } k = 0$). If $\mathbf{a} = \mathbf{n}$, the wave is called longitudinal, and if $\mathbf{a} \cdot \mathbf{n} = 0$, the wave is called transversal.

The acoustic tensor $Q(\mathbf{n})$ defined by (11.23) is real, but, in general, it is not symmetric. Therefore, there are no guarantee that there are three eigenvectors, and even there are, they may not be orthogonal. However, in the three-dimensional (real) vector space, there is at least one (real) eigenvalue. Indeed, the propagation direction vector \mathbf{n} is an eigenvector since from (11.24), we have

$$\rho \left(\frac{\omega}{k} \right)^2 \mathbf{n} = K[\mathbf{n} \otimes \mathbf{n}] \mathbf{n},$$

and provided that

$$\rho \left(\frac{\omega}{k} \right)^2 = \mathbf{n} \cdot K[\mathbf{n} \otimes \mathbf{n}] \mathbf{n}, \quad \rho U_{\text{nn}}^2 = K_{ijkl} n_i n_j n_k n_l,$$

is non-negative, it is a longitudinal wave with wave speed U_{nn} .

Note that to guarantee the existence of traveling waves, from (11.25) one can postulate the following requirement:

$$K_{ijkl} v_i u_j v_k u_l > 0, \quad \text{for any } \mathbf{v} \neq 0, \quad \mathbf{u} \neq 0. \quad (11.26)$$

This is usually called the *strong ellipticity* condition. The validity of this requirement depends on constitutive parameters of the material model as well as the reference state of deformation.

11.4 Mooney–Rivlin elastic materials

The propagation condition (11.24) is valid for (Piola–Kirchhoff) elasticity tensor K defined in (11.18) for any elastic material $T = \mathcal{T}(F)$. In particular, for Mooney–Rivlin materials,

$$T = \mathcal{T}(F) = -p(F)I + s_1 B + s_2 B^{-1},$$

from (11.18), we have

$$K_{ijkl} = T_{ij} \delta_{kl} - T_{il} \delta_{jk} + \beta \delta_{ij} \delta_{kl} + s_1 (\delta_{ik} B_{jl} + B_{il} \delta_{jk}) - s_2 (B_{ik}^{-1} \delta_{jl} + \delta_{il} B_{jk}^{-1}),$$

and from $Q_{ik} = K_{ijkl} n_j n_l$, we have

$$Q_{ik} = \beta n_i n_k + s_1 (\delta_{ik} B_{jl} n_j n_l + B_{il} n_l n_k) - s_2 (B_{ik}^{-1} + n_i B_{kj}^{-1} n_j),$$

or in direct notation, the acoustic tensor can be written as

$$Q(\mathbf{n}) = \beta \mathbf{n} \otimes \mathbf{n} + s_1 \left((\mathbf{n} \cdot B \mathbf{n}) I + B \mathbf{n} \otimes \mathbf{n} \right) - s_2 \left(B^{-1} + \mathbf{n} \otimes B^{-1} \mathbf{n} \right). \quad (11.27)$$

Note that from this expression, the acoustic tensor $Q(\mathbf{n})$ is obviously not symmetric unless the direction \mathbf{n} is an eigenvector of the Cauchy–Green deformation tensor B .

We can easily obtain the propagation speed from (11.25) and (11.27), provided that the wave with amplitude \mathbf{a} in the direction \mathbf{n} exists,

$$\begin{aligned} \rho U_{na}^2 = & \beta (\mathbf{a} \cdot \mathbf{n})^2 + s_1 \left(\mathbf{n} \cdot B \mathbf{n} + (\mathbf{a} \cdot B \mathbf{n})(\mathbf{a} \cdot \mathbf{n}) \right) \\ & - s_2 \left(\mathbf{a} \cdot B^{-1} \mathbf{a} + (\mathbf{n} \cdot B^{-1} \mathbf{a})(\mathbf{a} \cdot \mathbf{n}) \right). \end{aligned} \quad (11.28)$$

In particular, for $\mathbf{a} = \mathbf{n}$, the longitudinal wave speed is given by

$$\rho U_m^2 = \beta + 2 \left(s_1 (\mathbf{n} \cdot B \mathbf{n}) - s_2 (\mathbf{n} \cdot B^{-1} \mathbf{n}) \right). \quad (11.29)$$

For $\mathbf{a} \cdot \mathbf{n} = 0$, the transversal wave speed is given by

$$\rho U_{na}^2 = s_1 (\mathbf{n} \cdot B \mathbf{n}) - s_2 (\mathbf{a} \cdot B^{-1} \mathbf{a}). \quad (11.30)$$

In general, for an arbitrary propagation direction \mathbf{n} , the eigenvector of the acoustic tensor $Q(\mathbf{n})$, that is, the amplitude vector \mathbf{a} may not be parallel or perpendicular to \mathbf{n} , therefore, it is an oblique wave with wave velocity given by (11.28).

We shall consider the following special cases:

1. Principal waves

That is, both the propagation direction \mathbf{n} and the amplitude vector \mathbf{a} are in the principal directions of the deformation tensor B , let

$$B = b_1 \mathbf{e}_1 \otimes \mathbf{e}_1 + b_2 \mathbf{e}_2 \otimes \mathbf{e}_2 + b_3 \mathbf{e}_3 \otimes \mathbf{e}_3, \quad \mathbf{n} = \mathbf{e}_1,$$

and

$$\mathbf{a} = \mathbf{e}_1, \quad \mathbf{a} = \mathbf{e}_2, \quad \mathbf{a} = \mathbf{e}_3,$$

respectively, we have one longitudinal and two transversal waves,

$$\begin{aligned} U_{11} &= \sqrt{\frac{1}{\rho} \left(\beta + 2\left(s_1 b_1 - \frac{s_2}{b_1}\right) \right)}, \\ U_{12} &= \sqrt{\frac{1}{\rho} \left(s_1 b_1 - \frac{s_2}{b_2} \right)}, \\ U_{13} &= \sqrt{\frac{1}{\rho} \left(s_1 b_1 - \frac{s_2}{b_3} \right)}. \end{aligned} \tag{11.31}$$

In general, there are nine principle waves, three longitudinal and six transversal, propagating into a deformed body. Therefore, when the field of the deformation is known in the body, one can determined the fields of local wave speed from (11.31) in the principal directions over the body.

2. Small deformation from natural state

For small deformation from the natural state, $B = I + \dots$, $b_1 = b_2 = b_3 = 1$, the results (11.31) reduce to

$$U_{11} \approx \sqrt{\frac{\beta + 2(s_1 - s_2)}{\rho}}, \quad U_{12} = U_{13} \approx \sqrt{\frac{s_1 - s_2}{\rho}}.$$

By comparison with the well-known results of linear elasticity, we can identity the Lamé constants with the material parameters of Mooney–Rivlin materials,

$$\lambda = \beta, \quad \mu = s_1 - s_2, \tag{11.32}$$

in the isotropic natural state.

11.5 Principal acceleration waves of finite amplitude

The results obtained in (11.31) for small amplitude waves can be confirmed from that of finite amplitude acceleration waves treated by Wang & Truesdell⁵ for isotropic elastic bodies. Their results will be cited below.

For an isotropic elastic body, $T = H(B)$, both the Cauchy stress T and the strain tensor B have the same principal directions $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$,

$$\begin{aligned} T &= t_1 \mathbf{e}_1 \otimes \mathbf{e}_1 + t_2 \mathbf{e}_2 \otimes \mathbf{e}_2 + t_3 \mathbf{e}_3 \otimes \mathbf{e}_3, \\ B &= b_1 \mathbf{e}_1 \otimes \mathbf{e}_1 + b_2 \mathbf{e}_2 \otimes \mathbf{e}_2 + b_3 \mathbf{e}_3 \otimes \mathbf{e}_3. \end{aligned}$$

Let the propagation direction $\mathbf{n} = \mathbf{e}_1$, then there are one longitudinal and two transversal principal wave velocities given by the following formulas in terms of principle stresses and strains:

$$\begin{aligned} \mathbf{s} &= s \mathbf{e}_1, & U_{11}^2 &= \frac{2b_1}{\rho} \frac{\partial t_1}{\partial b_1}, \\ \mathbf{s} &= s \mathbf{e}_2, & U_{12}^2 &= \frac{b_1}{\rho} \frac{t_1 - t_2}{b_1 - b_2}, \\ \mathbf{s} &= s \mathbf{e}_3, & U_{13}^2 &= \frac{b_1}{\rho} \frac{t_1 - t_3}{b_1 - b_3}. \end{aligned} \tag{11.33}$$

Note that no explicit constitutive relation is needed in the derivation of the formulas of the principal wave speed (11.33), which is valid for any isotropic elastic material model $T = H(B)$.

Linear elastic materials

For infinitesimal deformation from a natural state, let E be the linear strain tensor, then we have the Cauchy–Green strain tensor

$$B = I + 2E + o(2), \quad E = \nabla \mathbf{u} + \nabla \mathbf{u}^T,$$

where $\mathbf{u} = \mathbf{x} - \mathbf{X}$ is the displacement gradient and $o(2)$ denotes the higher order terms in $|\nabla \mathbf{u}|$. The constitutive equation for small deformation from a natural state is given by the Hook's law,

$$\begin{aligned} T &= H(B) = \lambda(\text{tr } E)I + 2\mu E \\ &= \frac{1}{2}\lambda(\text{tr } B - 3)I + \mu(B - I). \end{aligned}$$

In terms of principal basis,

$$B = \begin{bmatrix} b_1 & & \\ & b_2 & \\ & & b_3 \end{bmatrix} = \begin{bmatrix} 1 + 2E_{11} + \cdots & & \\ & 1 + 2E_{22} + \cdots & \\ & & 1 + 2E_{33} + \cdots \end{bmatrix}.$$

⁵Wang, C.-C., Truesdell, C.; Introduction to Rational Elasticity, Noordhoff International Publishing, 1973

Therefore, we have

$$\begin{aligned} t_1 &= \frac{1}{2}\lambda(b_1 + b_2 + b_3 - 3) + \mu(b_1 - 1) + \cdots, \\ t_2 &= \frac{1}{2}\lambda(b_1 + b_2 + b_3 - 3) + \mu(b_2 - 1) + \cdots, \\ t_3 &= \frac{1}{2}\lambda(b_1 + b_2 + b_3 - 3) + \mu(b_3 - 1) + \cdots, \end{aligned}$$

and from (11.33), we obtain the speed of the principal waves propagating in the direction \mathbf{e}_1 ,

$$\begin{aligned} U_{11}^2 &= \frac{2b_1}{\rho} \frac{\partial t_1}{\partial b_1} = b_1 \frac{\lambda + 2\mu}{\rho} \approx \frac{\lambda + 2\mu}{\rho}, \\ U_{12}^2 &= \frac{b_1}{\rho} \frac{t_1 - t_2}{b_1 - b_2} = b_1 \frac{\mu}{\rho} \approx \frac{\mu}{\rho}, \\ U_{13}^2 &= \frac{b_1}{\rho} \frac{t_1 - t_3}{b_1 - b_3} = b_1 \frac{\mu}{\rho} \approx \frac{\mu}{\rho}. \end{aligned} \tag{11.34}$$

Therefore, we have the well-known results that, there are only two wave speeds: $\sqrt{(\lambda + 2\mu)/\rho}$ for the longitudinal wave and $\sqrt{\mu/\rho}$ for the transversal (shear) wave in isotropic linear elastic bodies.

Mooney–Rivlin elastic material

We consider the Mooney–Rivlin material model of constitutive equation for an isotropic elastic body,

$$T = H(B) = -p(\rho)I + s_1 B + s_2 B^{-1}, \quad \rho = \frac{\rho_\kappa}{\sqrt{\det B}},$$

where s_1 and s_2 are material constants.

In principal basis,

$$\begin{aligned} t_1 &= -p(\rho) + s_1 b_1 + s_2 b_1^{-1}, \\ t_2 &= -p(\rho) + s_1 b_2 + s_2 b_2^{-1}, \\ t_3 &= -p(\rho) + s_1 b_3 + s_2 b_3^{-1}, \end{aligned} \quad \rho = \frac{\rho_\kappa}{\sqrt{b_1 b_2 b_3}},$$

and

$$\frac{\partial p}{\partial b_1} = \frac{dp}{d\rho} \frac{\partial \rho}{\partial b_1} = -\frac{1}{2} \rho \frac{dp}{d\rho} \frac{1}{b_1} = -\frac{1}{2} \frac{\beta}{b_1},$$

where the material parameter $\beta(\rho)$ is defined as

$$\beta = \rho \frac{dp}{d\rho}.$$

Note that if the body is nearly incompressible, the value of $d\rho/dp$ is nearly zero, meaning, even very large pressure can hardly change the mass density. In this case, the value of β is much greater than one, i.e., $\beta \gg 1$.

From (11.33), we obtain the wave speed of the principal waves propagating in the direction \mathbf{e}_1 ,

$$\begin{aligned}
 U_{11}^2 &= \frac{2b_1}{\rho} \frac{\partial t_1}{\partial b_1} = \frac{1}{\rho} \left(\beta + 2(s_1 b_1 - \frac{s_2}{b_1}) \right), \\
 U_{12}^2 &= \frac{b_1}{\rho} \frac{t_1 - t_2}{b_1 - b_2} = \frac{1}{\rho} \left(s_1 b_1 - \frac{s_2}{b_2} \right), \\
 U_{13}^2 &= \frac{b_1}{\rho} \frac{t_1 - t_3}{b_1 - b_3} = \frac{1}{\rho} \left(s_1 b_1 - \frac{s_2}{b_3} \right),
 \end{aligned}
 \tag{11.35}$$

which agree with the results (11.31) for small amplitude waves.

12 Mixture theory of porous media

The theories of mixtures in the framework of continuum mechanics have been developed throughout the sixties and seventies of the last century. Here we shall briefly review some essential features for the governing balance equations and the formulation of constitutive theories of mixtures of different material constituents.

12.1 Theories of mixtures

We consider a mixture of N constituents, all of which are supposed to be able to occupy the same region of space *simultaneously*. Let \mathcal{B}_α denote the α^{th} constituent and κ_α be its reference configuration and denote $B_\alpha = \kappa_\alpha(\mathcal{B}_\alpha)$. The motion of \mathcal{B}_α is a smooth mapping,

$$\chi_\alpha : B_\alpha \times \mathbb{R} \rightarrow \mathbb{E}, \quad \mathbf{x} = \chi_\alpha(\mathbf{X}_\alpha, t), \quad \mathbf{X}_\alpha \in B_\alpha,$$

for each constituent $\alpha = 1, \dots, N$. We denote the mixture body at the instant t as

$$B_t = \chi_\alpha(B_\alpha, t) \subset \mathbb{E},$$

valid for each constituent α . The mapping $\chi_\alpha(t) : B_\alpha \rightarrow B_t$ is smooth and bijective.

Schematically, for $\alpha = \{\text{s}, \text{f}\}$ (stand for solid and fluid constituents), the motion of the constituents can be depicted in the following diagram:

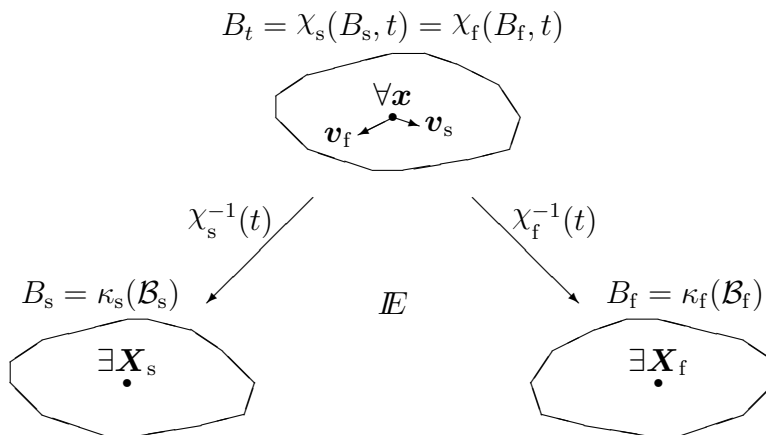


Figure 10: At any instant, there are constituent-points occupied the same position instantaneously with different constituent-velocities individually.

It states that at any instant t , for any spatial position $\mathbf{x} \in B_t$, there is a material point $X_\alpha \in \mathcal{B}_\alpha$ in each constituent, at its reference position $\mathbf{X}_\alpha = \kappa_\alpha(X_\alpha)$, that in their motions

$\mathbf{x} = \chi_\alpha(\mathbf{X}_\alpha, t)$ occupy simultaneously the same spatial position in the Euclidean space \mathbb{E} . — This is the basic assumption that one can regard a mixture body as a continuous medium of multicomponent body.

The velocity and the deformation gradient of each constituent are defined as

$$\mathbf{v}_\alpha = \frac{\partial}{\partial t} \chi_\alpha(\mathbf{X}_\alpha, t), \quad F_\alpha = \nabla_{\mathbf{X}_\alpha} \chi_\alpha(\mathbf{X}_\alpha, t).$$

We introduce the following quantities for the constituent $\alpha \in \{1, \dots, N\}$:

- ρ_α mass density of constituent α .
- \mathbf{v}_α velocity of constituent α .
- T_α stress tensor of constituent α .
- \mathbf{b}_α external body force on constituent α .
- ε_α internal energy density of constituent α .
- \mathbf{q}_α energy flux of constituent α .
- r_α external energy supply of constituent α .
- η_α entropy density of constituent α .
- Φ_α entropy flux of constituent α .
- s_α external entropy supply of constituent α .
- τ_α mass production of constituent α .
- \mathbf{m}_α interaction force on constituent α .
- M_α interaction moment of momentum on constituent α .
- l_α energy production of constituent α .
- σ_α entropy production of constituent α .

Balance laws of each constituent

Following the pioneering work of Truesdell, the basic laws of a mixture are given by the following balance equations for mass, linear momentum, moment of momentum, and energy for each constituent:

$$\begin{aligned} \frac{\partial \rho_\alpha}{\partial t} + \operatorname{div}(\rho_\alpha \mathbf{v}_\alpha) &= \tau_\alpha, \\ \frac{\partial \rho_\alpha \mathbf{v}_\alpha}{\partial t} + \operatorname{div}(\rho_\alpha \mathbf{v}_\alpha \otimes \mathbf{v}_\alpha - T_\alpha) - \rho_\alpha \mathbf{b}_\alpha &= \mathbf{m}_\alpha, \\ T_\alpha - T_\alpha^T &= M_\alpha, \\ \frac{\partial}{\partial t} \left(\rho_\alpha \varepsilon_\alpha + \frac{1}{2} \rho_\alpha \mathbf{v}_\alpha^2 \right) + \operatorname{div} \left(\left(\rho_\alpha \varepsilon_\alpha + \frac{1}{2} \rho_\alpha \mathbf{v}_\alpha^2 \right) \mathbf{v}_\alpha + \mathbf{q}_\alpha - T_\alpha^T \mathbf{v}_\alpha \right) \\ &\quad - \rho_\alpha r_\alpha - \rho_\alpha \mathbf{b}_\alpha \cdot \mathbf{v}_\alpha = l_\alpha, \end{aligned} \tag{12.1}$$

If the mixture reduces to a single constituent, the right hand side of the above equations are zero, and we recover the balance laws for a single body. The terms on the right hand

side thus represent the physical transfers among different constituents, such as chemical reactions, interaction forces and energy transfer. In a mixture as a whole, we assume that such physical transfers are solely due to exchanges among constituents. Therefore, we postulate that the mixture as a whole should behave like a single body. This is expressed by the following relations:

$$\sum_{\alpha} \tau_{\alpha} = 0, \quad \sum_{\alpha} \mathbf{m}_{\alpha} = 0, \quad \sum_{\alpha} M_{\alpha} = 0, \quad \sum_{\alpha} l_{\alpha} = 0. \quad (12.2)$$

Balance laws of the mixture

The motion of a mixture as a whole will be assumed to be governed by the same equations of balance as a single body, by summing up the balance equations (12.1) over all constituents $\alpha \in \{1, \dots, N\}$,

$$\begin{aligned} \sum_{\alpha} \left\{ \frac{\partial \rho_{\alpha}}{\partial t} + \operatorname{div}(\rho_{\alpha} \mathbf{v}_{\alpha}) \right\} &= \sum_{\alpha} \tau_{\alpha}, \\ \sum_{\alpha} \left\{ \frac{\partial \rho_{\alpha} \mathbf{v}_{\alpha}}{\partial t} + \operatorname{div}(\rho_{\alpha} \mathbf{v}_{\alpha} \otimes \mathbf{v}_{\alpha} - T_{\alpha}) - \rho_{\alpha} \mathbf{b}_{\alpha} \right\} &= \sum_{\alpha} \mathbf{m}_{\alpha}, \\ \sum_{\alpha} \left\{ T_{\alpha} - T_{\alpha}^T \right\} &= \sum_{\alpha} M_{\alpha}, \\ \sum_{\alpha} \left\{ \frac{\partial}{\partial t} \left(\rho_{\alpha} \varepsilon_{\alpha} + \frac{1}{2} \rho_{\alpha} \mathbf{v}_{\alpha}^2 \right) + \operatorname{div}(\dots) - \rho_{\alpha} r_{\alpha} - \rho_{\alpha} \mathbf{b}_{\alpha} \cdot \mathbf{v}_{\alpha} \right\} &= \sum_{\alpha} l_{\alpha}. \end{aligned} \quad (12.3)$$

By the use of the relations (12.2) and the following definition of the corresponding quantities for the mixture,

$$\begin{aligned} \rho &= \sum_{\alpha} \rho_{\alpha}, \\ \mathbf{v} &= \sum_{\alpha} \frac{\rho_{\alpha}}{\rho} \mathbf{v}_{\alpha}, \\ T &= \sum_{\alpha} (T_{\alpha} - \rho_{\alpha} \mathbf{u}_{\alpha} \otimes \mathbf{u}_{\alpha}), \\ \mathbf{b} &= \sum_{\alpha} \frac{\rho_{\alpha}}{\rho} \mathbf{b}_{\alpha}, \end{aligned} \quad (12.4)$$

and

$$\begin{aligned} \varepsilon &= \sum_{\alpha} \frac{\rho_{\alpha}}{\rho} \left(\varepsilon_{\alpha} + \frac{1}{2} \mathbf{u}_{\alpha}^2 \right), \\ \mathbf{q} &= \sum_{\alpha} \left(\mathbf{q}_{\alpha} + \rho_{\alpha} \left(\varepsilon_{\alpha} + \frac{1}{2} \mathbf{u}_{\alpha}^2 \right) \mathbf{u}_{\alpha} - T_{\alpha}^T \mathbf{u}_{\alpha} \right), \\ r &= \sum_{\alpha} \frac{\rho_{\alpha}}{\rho} (r_{\alpha} + \mathbf{b}_{\alpha} \cdot \mathbf{u}_{\alpha}), \end{aligned} \quad (12.5)$$

the equations (12.3) reduces to the usual balance equations of mass, linear momentum, and energy for the mixture as a single body,

$$\begin{aligned}\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \mathbf{v}) &= 0, \\ \frac{\partial \rho \mathbf{v}}{\partial t} + \operatorname{div}(\rho \mathbf{v} \otimes \mathbf{v} - T) &= \rho \mathbf{b}, \\ \frac{\partial}{\partial t}(\rho \varepsilon + \frac{1}{2} \rho \mathbf{v}^2) + \operatorname{div}\left((\rho \varepsilon + \frac{1}{2} \rho \mathbf{v}^2) \mathbf{v} + \mathbf{q} - T^T \mathbf{v}\right) &= \rho r + \rho \mathbf{b} \cdot \mathbf{v}.\end{aligned}\tag{12.6}$$

The energy equation can also be written as,

$$\frac{\partial \rho \varepsilon}{\partial t} + \operatorname{div}(\rho \varepsilon \mathbf{v} + \mathbf{q}) - \operatorname{tr}(T \operatorname{grad} \mathbf{v}) = \rho r.\tag{12.7}$$

Note that in the relation (12.4) and (12.5), we have introduced the diffusive velocity of constituent α relative to the mixture,

$$\mathbf{u}_\alpha = \mathbf{v}_\alpha - \mathbf{v},$$

which by the use of (12.4)₂, satisfies the identity,

$$\sum_{\alpha} \rho_{\alpha} \mathbf{u}_{\alpha} = 0.\tag{12.8}$$

Also note that $\sum M_{\alpha} = 0$ implies the symmetry of $\sum T_{\alpha}$ as well as the total stress T , while the partial stress T_{α} is not symmetric in general.

The entropy inequality

The balance of entropy of constituent α can be written as

$$\frac{\partial \rho_{\alpha} \eta_{\alpha}}{\partial t} + \operatorname{div}(\rho_{\alpha} \eta_{\alpha} \mathbf{v}_{\alpha} + \mathbf{\Phi}_{\alpha}) - \rho_{\alpha} s_{\alpha} = \sigma_{\alpha}.$$

Since entropy is not a conservative quantity, we shall allow the partial production of entropy of a constituent be of any amount provided that the total production of entropy for the mixture be non-negative,

$$\sum_{\alpha} \sigma_{\alpha} \geq 0.\tag{12.9}$$

By summing up the equations of all constituents $\alpha \in \{1, \dots, N\}$ and introducing the corresponding quantities for the mixture,

$$\begin{aligned}\eta &= \sum_{\alpha} \frac{\rho_{\alpha}}{\rho} \eta_{\alpha}, \\ \mathbf{\Phi} &= \sum_{\alpha} (\mathbf{\Phi}_{\alpha} + \rho_{\alpha} \eta_{\alpha} \mathbf{u}_{\alpha}), \\ s &= \sum_{\alpha} \frac{\rho_{\alpha}}{\rho} s_{\alpha},\end{aligned}\tag{12.10}$$

we obtain the usual entropy inequality for the mixture as a single body,

$$\frac{\partial \rho \eta}{\partial t} + \operatorname{div}(\rho \eta \mathbf{v} + \boldsymbol{\Phi}) - \rho s \geq 0. \quad (12.11)$$

Note that we do not postulate a similar entropy inequality for individual constituent. The partial entropy production could be negative provided that the total production of the mixture is non-negative, otherwise it will lead to very restrictive physical results.

The use of the general entropy inequality (12.11) for the exploitation of its restrictions on constitutive models is often quite sophisticate and could be rather tedious. In application, assumption for the specific form of the entropy flux and entropy supply for the constituents is usually adopted, namely,

$$\boldsymbol{\Phi}_\alpha = \frac{1}{\theta_\alpha} \mathbf{q}_\alpha, \quad s_\alpha = \frac{1}{\theta_\alpha} r_\alpha, \quad (12.12)$$

where θ_α is the temperature of the constituent α . Under this assumption, from (12.10), we have

$$\begin{aligned} \boldsymbol{\Phi} &= \sum_\alpha \left(\frac{\mathbf{q}_\alpha}{\theta_\alpha} + \rho_\alpha \eta_\alpha \mathbf{u}_\alpha \right), \\ s &= \sum_\alpha \frac{\rho_\alpha}{\rho} \frac{r_\alpha}{\theta_\alpha}. \end{aligned}$$

For further discussions, we shall assume that energy exchange among constituents is quick enough relative to characteristic observation time, so that all the constituents have the same common temperature, then we have

$$\begin{aligned} \boldsymbol{\Phi} &= \frac{1}{\theta} \sum_\alpha \left(\mathbf{q}_\alpha + \theta \rho_\alpha \eta_\alpha \mathbf{u}_\alpha \right), \\ s &= \frac{1}{\theta} \sum_\alpha \frac{\rho_\alpha}{\rho} r_\alpha. \end{aligned} \quad (12.13)$$

Even though we have adopted the Clausius–Duhem assumptions (12.12) for each constituent, for the mixture as a single body, from (12.13) and (12.5) such assumptions are not valid in general. In other words, the usual Clausius–Duhem inequality cannot be adopted in theories of mixtures. Indeed, we can rewrite the relations (12.13) as

$$\begin{aligned} \boldsymbol{\Phi} &= \frac{1}{\theta} \left(\mathbf{q} - \sum_\alpha \rho_\alpha \left(K_\alpha + \frac{1}{2} \mathbf{u}_\alpha^2 I \right) \mathbf{u}_\alpha \right), \\ s &= \frac{1}{\theta} \left(r - \sum_\alpha \frac{\rho_\alpha}{\rho} \mathbf{u}_\alpha \cdot \mathbf{b}_\alpha \right), \end{aligned} \quad (12.14)$$

where

$$K_\alpha = \psi_\alpha I - \frac{T_\alpha^T}{\rho_\alpha}, \quad \psi_\alpha = \varepsilon_\alpha - \theta \eta_\alpha, \quad (12.15)$$

are the chemical potential tensor and the free energy of constituent α respectively. For the free energy of the mixture we define

$$\psi = \sum_{\alpha} \frac{\rho_{\alpha}}{\rho} \psi_{\alpha}. \quad (12.16)$$

Remark. It has been shown that the assumptions (12.12) are appropriate (but not necessarily general enough) to account for the behavior of a mixture within the framework of continuum mechanics. On the other hand, the very expressions in (12.14) show that the Clausius–Duhem assumptions $\Phi = \mathbf{q}/\theta$ and $s = r/\theta$ may not be appropriate for thermodynamic considerations of material bodies in general.

Jump conditions at semi-permeable surface

Let \mathcal{S} be a surface within the mixture and $[A]$ be the jump of a quantity A across \mathcal{S} , defined by

$$[A] = A^+ - A^-.$$

We assume that the one-side limits A^+ and A^- exist and the jump is smooth on \mathcal{S} . Such a surface is called a singular surface. Let \mathbf{u}^* and \mathbf{n} denote the velocity and the unit normal of \mathcal{S} respectively.

In addition to the balance laws (12.6) and (12.11) at a regular point where relevant fields are continuous, at a singular point, i.e., a point on \mathcal{S} , the balance laws take the form of the following jump conditions of mass, momentum, energy and entropy:

$$\begin{aligned} [\rho(\mathbf{v} - \mathbf{u}^*)] \cdot \mathbf{n} &= 0, \\ [\rho\mathbf{v} \otimes (\mathbf{v} - \mathbf{u}^*) - T]\mathbf{n} &= 0, \\ [\mathbf{q} - T\mathbf{v} + \rho(\varepsilon + \frac{1}{2}\mathbf{v}^2)(\mathbf{v} - \mathbf{u}^*)] \cdot \mathbf{n} &= 0, \\ [\Phi + \rho\eta(\mathbf{v} - \mathbf{u}^*)] \cdot \mathbf{n} &\geq 0. \end{aligned} \quad (12.17)$$

A singular surface will be called ideal, if across it the temperature is continuous, $[\theta] = 0$. We assume further that across an ideal singular surface, the entropy jump is null,

$$[\Phi + \rho\eta(\mathbf{v} - \mathbf{u}^*)] \cdot \mathbf{n} = 0, \quad (12.18)$$

and the entropy flux Φ is given by the relation (12.14)₁.

The diffusive velocity of constituent α relative to the surface \mathcal{S} is defined as

$$V_{\alpha} = \mathbf{v}_{\alpha} - \mathbf{u}^*.$$

If $V_{\alpha} = 0$, we say that \mathcal{S} is not permeable to the constituent α , i.e., the constituent does not go across the surface. A singular surface which is not permeable to all constituents

except the constituent ν is called a semi-permeable surface with respect to the constituent ν , or simply a ν -permeable surface. We have the following result: ⁶:

Theorem. *Let \mathcal{S} be an ideal ν -permeable surface, then by the use of the jump conditions (12.17)_{1,2} and (12.18), the energy jump condition (12.17)₃ becomes*

$$[(K_\nu + \frac{1}{2}V_\nu^2 I)\rho_\nu V_\nu] \cdot \mathbf{n} = 0, \quad (12.19)$$

where K_ν is the chemical potential tensor of the constituent ν .

Proof: Note that at a ν -permeable surface, from (12.17), we have

$$\begin{aligned} [\rho_\nu V_\nu] \cdot \mathbf{n} &= 0, \\ [T]\mathbf{n} &= [\mathbf{v} \otimes \rho_\nu V_\nu]\mathbf{n}, \\ [\mathbf{q} - T\mathbf{v}] \cdot \mathbf{n} &= -[(\varepsilon + \frac{1}{2}\mathbf{v}^2)\rho_\nu V_\nu] \cdot \mathbf{n}. \end{aligned} \quad (12.20)$$

We shall work on the left hand side of the energy jump condition (12.20)₃. First, from (12.18) and (12.14)₁, we have

$$[\mathbf{q}] \cdot \mathbf{n} = [-\theta\eta \rho_\nu V_\nu] \cdot \mathbf{n} + \sum [(K_\alpha + \frac{1}{2}\mathbf{u}_\alpha^2 I)\rho_\alpha \mathbf{u}_\alpha] \cdot \mathbf{n}. \quad (12.21)$$

With $\mathbf{u}_\alpha = (\mathbf{v}_\alpha - \mathbf{u}^*) - \sum \frac{\rho_\beta}{\rho}(\mathbf{v}_\beta - \mathbf{u}^*)$ and the relations (12.15) and (12.16), the second term on the right hand side becomes

$$\begin{aligned} \sum [(K_\alpha + \frac{1}{2}\mathbf{u}_\alpha^2 I)\rho_\alpha \mathbf{u}_\alpha] \cdot \mathbf{n} &= [(K_\nu + \frac{1}{2}\mathbf{u}_\nu^2 I)\rho_\nu V_\nu] \cdot \mathbf{n} - \sum [\frac{\rho_\alpha}{\rho} (K_\alpha + \frac{1}{2}\mathbf{u}_\alpha^2 I)\rho_\nu V_\nu] \cdot \mathbf{n} \\ &= [(K_\nu + \frac{1}{2}\mathbf{u}_\nu^2 I)\rho_\nu V_\nu] \cdot \mathbf{n} + [\frac{1}{\rho}(T_I - \rho\psi I + \sum \frac{1}{2}\mathbf{u}_\alpha^2 I)\rho_\nu V_\nu] \cdot \mathbf{n}. \end{aligned}$$

Substituting this into (12.21), by the use of (12.15) and (12.5)₁, we have

$$[\mathbf{q}] \cdot \mathbf{n} = [(K_\nu + \frac{T_I}{\rho} - \varepsilon I + \frac{1}{2}\mathbf{u}_\nu^2 I)\rho_\nu V_\nu] \cdot \mathbf{n}. \quad (12.22)$$

Now, for the second term on the left hand side of (12.20)₃, from (12.4)₃, we have

$$[T\mathbf{v}] \cdot \mathbf{n} = [T_I\mathbf{v}] \cdot \mathbf{n} - \sum [(\rho_\alpha \mathbf{u}_\alpha \otimes \mathbf{u}_\alpha)\mathbf{v}] \cdot \mathbf{n},$$

where $T_I = \sum T_\alpha$ is a symmetric tensor. With $\mathbf{v} = \mathbf{v} - \mathbf{u}^* + \mathbf{u}^* = \frac{1}{\rho} \sum \rho_\beta(\mathbf{v}_\beta - \mathbf{u}^*) + \mathbf{u}^*$, it becomes

$$\begin{aligned} [T\mathbf{v}] \cdot \mathbf{n} &= [\frac{T_I}{\rho}\rho_\nu V_\nu] \cdot \mathbf{n} + [T_I\mathbf{u}^*] \cdot \mathbf{n} - \sum [(\mathbf{u}_\alpha \cdot \mathbf{v})\rho_\alpha \mathbf{u}_\alpha] \cdot \mathbf{n} \\ &= [\frac{T_I}{\rho}\rho_\nu V_\nu] \cdot \mathbf{n} + [T\mathbf{u}^*] \cdot \mathbf{n} + \sum [(\rho_\alpha \mathbf{u}_\alpha \otimes \mathbf{u}_\alpha)\mathbf{u}^*] \cdot \mathbf{n} - \sum [(\mathbf{u}_\alpha \cdot \mathbf{v})\rho_\alpha \mathbf{u}_\alpha] \cdot \mathbf{n}, \end{aligned}$$

⁶I-Shih Liu, On chemical potential and incompressible porous media, *Journal de Mecanique* 19, 327-342 (1980).

and by the use of (12.20)₂ and $\mathbf{u}_\alpha = (\mathbf{v}_\alpha - \mathbf{u}^*) - (\mathbf{v} - \mathbf{u}^*)$, we have

$$[T\mathbf{v}] \cdot \mathbf{n} = \left[\frac{T_I}{\rho} \rho_\nu V_\nu \right] \cdot \mathbf{n} + [((\mathbf{u}^* \cdot \mathbf{v}) + (\mathbf{u}_\nu \cdot \mathbf{u}^*) - (\mathbf{u}_\nu \cdot \mathbf{v})) \rho_\nu V_\nu] \cdot \mathbf{n}. \quad (12.23)$$

Substituting (12.22) and (12.23) into (12.20)₃, we obtain

$$[(K_\nu + (\frac{1}{2}\mathbf{u}_\nu^2 + \frac{1}{2}\mathbf{v}^2 - (\mathbf{u}^* \cdot \mathbf{v}) - (\mathbf{u}_\nu \cdot \mathbf{u}^*) + (\mathbf{u}_\nu \cdot \mathbf{v}))I) \rho_\nu V_\nu] \cdot \mathbf{n} = 0.$$

By summing $[\frac{1}{2}\mathbf{u}^{*2} \rho_\nu V_\nu] \cdot \mathbf{n}$, which is zero by (12.20)₁, to the above equation, we obtain

$$[(K_\nu + \frac{1}{2}(\mathbf{u}_\nu + \mathbf{v} - \mathbf{u}^*)^2 I) \rho_\nu V_\nu] \cdot \mathbf{n} = 0,$$

that proves the theorem. \square

12.2 Mixture of elastic materials

We consider a non-reacting mixture of elastic materials characterized by the constitutive equation of the form:

$$f = \mathcal{F}(\theta, \text{grad } \theta, F_\alpha, \text{grad } F_\alpha, \mathbf{v}_\alpha), \quad (12.24)$$

where θ is the temperature, F_α is the deformation gradient of the constituent α relative to a reference configuration, and

$$f = \{T_\alpha, \varepsilon, \mathbf{q}, \mathbf{m}_\alpha\}$$

are the constitutive quantities for the basic field variables $\{\rho_\alpha, \mathbf{v}_\alpha, \theta\}$ with the following governing equations:

$$\begin{aligned} \frac{\partial \rho_\alpha}{\partial t} + \text{div}(\rho_\alpha \mathbf{v}_\alpha) &= 0, \\ \frac{\partial \rho_\alpha \mathbf{v}_\alpha}{\partial t} + \text{div}(\rho_\alpha \mathbf{v}_\alpha \otimes \mathbf{v}_\alpha - T_\alpha) - \rho_\alpha \mathbf{b}_\alpha &= \mathbf{m}_\alpha, \\ \frac{\partial \rho \varepsilon}{\partial t} + \text{div}(\rho \varepsilon \mathbf{v} + \mathbf{q}) - \text{tr}(T \text{grad } \mathbf{v}) &= \rho r. \end{aligned} \quad (12.25)$$

Thermodynamic considerations of such a mixture theory has been considered by Bowen⁷ in which consequences of the entropy principle have been obtained based on the entropy inequality of the form:

$$\frac{\partial \rho \eta}{\partial t} + \text{div}(\rho \eta \mathbf{v} + \Phi) - \rho s \geq 0, \quad (12.26)$$

⁷ Bowen, R.M.: *The theory of Mixtures in Continuum Mechanics*, Edited by A.C. Eringen, Vol. III, Academic Press (1976). See also Truesdell, C.: *Rational Thermodynamics*, Second Edition, Springer-Verlag, New York Berlin (1984), Appendix 5A.

where for the mixture under consideration, the entropy flux and the entropy supply density are given by the relation (12.14),

$$\begin{aligned}\boldsymbol{\Phi} &= \frac{1}{\theta} \left(\mathbf{q} - \sum_{\alpha} (\rho_{\alpha} (\psi_{\alpha} + \frac{1}{2} \mathbf{u}_{\alpha}^2) I - T_{\alpha}^T) \mathbf{u}_{\alpha} \right), \\ s &= \frac{1}{\theta} \left(r - \frac{1}{\rho} \sum_{\alpha} \rho_{\alpha} \mathbf{u}_{\alpha} \cdot \mathbf{b}_{\alpha} \right),\end{aligned}\tag{12.27}$$

where ψ_{α} is the free energy density of the constituent α .

Note that since only a common temperature is considered for all constituents, only the energy equation of the mixture, instead of the partial energy equations (12.1)₄ of all constituents, is needed in the system of governing equations (12.25). Moreover, since the partial stresses are not assumed to be symmetric, the equation of partial moment of momentum (12.1)₃ is not explicitly needed in (12.25).

Note that the constitutive variables (12.24) for elastic materials not only contain the deformation gradient F_{α} but also the second gradients, $\text{grad } F_{\alpha}$. This makes the mixture under consideration a non-simple mixture (in the sense of Noll's simple materials). The reason to consider such a mixture will become clear later.

Summary of results for elastic solid-fluid mixtures

For a mixture of a solid (with subindex s) and a fluid (with subindex f), to establish field equations of the basic field variables, $\{\rho_s, \rho_f, \chi_s, \chi_f, \theta\}$, constitutive equations for the quantities in the balance equations,

$$f = \{T_s, T_f, \varepsilon, \mathbf{q}, \mathbf{m}_f\},\tag{12.28}$$

must be specified. For an elastic solid-fluid mixture, we consider the constitutive equations of the form:

$$f = \mathcal{F}(\theta, \rho_f, F_s, \text{grad } \theta, \text{grad } \rho_f, \text{grad } F_s, V).\tag{12.29}$$

where θ is the temperature and $V = \mathbf{v}_f - \mathbf{v}_s$ is referred to as the relative velocity.

The results of thermodynamic restrictions of such a mixture obtained by Bowen have also been confirmed from the analysis with the use of Lagrange multipliers by Liu⁸, and

⁸Liu, I-Shih: A solid-fluid mixture theory of porous media, *International Journal of Engineering Sciences* 84, 133-146 (2014).

can be summarized in the following constitutive equations:

$$\begin{aligned}
T_f &= \rho_f \psi_f I - \frac{\partial \rho \psi_I}{\partial \rho_f} \rho_f I + \rho_f \frac{\partial \psi_f}{\partial V} \otimes V, \\
T_s &= \rho_s \psi_s I + \frac{\partial \rho \psi_I}{\partial F_s} F_s^T + \rho_s \frac{\partial \psi_s}{\partial V} \otimes V, \\
\varepsilon &= \psi_I - \theta \frac{\partial \psi_I}{\partial \theta} + \frac{1}{2} \frac{\rho_f \rho_s}{\rho^2} V \cdot V, \\
\mathbf{m}_f^0 &= \frac{\partial \rho_s \psi_s^0}{\partial \rho_f} \text{grad } \rho_f - \frac{\partial \rho_f \psi_f^0}{\partial F_s} \cdot \text{grad } F_s, \quad \mathbf{q}^0 = 0.
\end{aligned} \tag{12.30}$$

where \circ denotes the equilibrium value at the state with $V = 0$ and $\text{grad } \theta = 0$.

These constitutive equations depend solely on the constitutive functions of the free energy,

$$\begin{aligned}
\psi_f &= \psi_f(\theta, \rho_f, F_s, V), & \psi_s &= \psi_s(\theta, \rho_f, F_s, V), \\
\psi_I &= \psi_I(\theta, \rho_f, F_s).
\end{aligned} \tag{12.31}$$

Note that although the partial free energies ψ_f and ψ_s may depend on the relative velocity V , the (inner) free energy ψ_I ,

$$\rho \psi_I = \rho_f \psi_f + \rho_s \psi_s,$$

does not depend on V .

Moreover, from (12.30) and (12.31), the sum of partial stresses becomes

$$\begin{aligned}
T_I &= \rho \psi_I I - \frac{\partial \rho \psi_I}{\partial \rho_f} \rho_f I + \frac{\partial \rho \psi_I}{\partial F_s} F_s^T, \\
T &= T_I(\theta, \rho_f, F_s) - \frac{1}{2} \frac{\rho_s \rho_f}{\rho} V \otimes V.
\end{aligned} \tag{12.32}$$

Similarly, although the partial stresses T_f and T_s may depend on V , the sum of partial stress, $T_I = T_f + T_s$, does not depend on V .

If we define the equilibrium chemical potential of the fluid and the equilibrium partial fluid pressure as

$$\mu_f = \frac{\partial \rho \psi_I}{\partial \rho_f}, \quad p_f = \rho_f (\mu_f - \psi_f^0), \tag{12.33}$$

then the equilibrium fluid stress reduces to the pressure, $T_f^0 = -p_f I$, and

$$T_f = -p_f I + \rho_f (\psi_f - \psi_f^0) I + \rho_f \frac{\partial \psi_f}{\partial V} \otimes V. \tag{12.34}$$

From (12.30)₁, the chemical potential tensor of the fluid becomes

$$K_f = \mu_f I - V \otimes \frac{\partial \psi_f}{\partial V}, \tag{12.35}$$

and from (12.30)₄ and (12.33), the interaction force can be written as

$$\mathbf{m}_f^0 = \frac{p_f}{\rho_f} \text{grad } \rho_f - \rho_f (\text{grad } \psi_f^0) \Big|_0. \tag{12.36}$$

Chemical potential at fluid-permeable surface

At a fluid-permeable surface considered in the previous section, from the relation (12.19), we have the following energy jump condition,

$$[(K_f + \frac{1}{2}V_f^2 I)\rho_f V_f] \cdot \mathbf{n} = 0,$$

At this surface, $\mathbf{v}_s = \mathbf{u}^*$, so that $V_f = \mathbf{v}_f - \mathbf{u}^* = V$. Hence, by the use of (12.35), the condition becomes

$$\left[\mu_f + \frac{1}{2}V^2 - V \cdot \frac{\partial \psi_f}{\partial V} \right] \rho_f V_f \cdot \mathbf{n} = 0,$$

where we have noted that from the mass jump condition (12.17)₁ at the fluid-permeable surface ($\rho_f V_f \cdot \mathbf{n}$) is continuous, and since it is not zero in general, it follows that

$$\left[\mu_f + \frac{1}{2}V^2 - V \cdot \frac{\partial \psi_f}{\partial V} \right] = 0. \quad (12.37)$$

In particular, we have

$$[\mu_f] \rightarrow 0 \quad \text{when} \quad V \rightarrow 0.$$

Thus, we conclude that in equilibrium ($\mathbf{v}_f = \mathbf{u}^*$), the chemical potential of the fluid constituent is continuous across a fluid-permeable surface. This confirms the well-known result of the classical theory.

12.3 Saturated porous media

The solid-fluid mixture considered in the previous section can be regarded as a model for saturated porous media provided that the concept of porosity is introduced. For mixture theory of porous media, a material point is regarded as a representative volume element dV which contains pores through them fluid constituent can flow. Physically, it is assumed that a representative volume element is large enough compare to solid grains (connected or not), yet at the same time small enough compare to the characteristic length of the material body.

Let the volume fraction of pores be denoted by ϕ , then the fractions of representative volume element of the fluid and the solid are

$$dV_f = \phi dV, \quad dV_s = (1 - \phi)dV,$$

if the porous medium is saturated.

Remember that in the mixture theory, the mass densities are defined relative to the mixture volume, so that the fluid and solid mass in the representative volume element are given by

$$\begin{aligned} dM_f &= \rho_f dV = d_f dV_f, \\ dM_s &= \rho_s dV = d_s dV_s, \end{aligned}$$

and hence,

$$\rho_f = \phi d_f, \quad \rho_s = (1 - \phi)d_s, \quad (12.38)$$

where d_f and d_s are the true mass densities of fluid and solid constituents respectively.

We shall also regard the partial fluid pressure p_f in the mixture theory as the outcome of a “microscopic” pressure acting over the area fraction of surface actually occupied by the fluid in the pore, i.e.,

$$p_f dA = P dA_f, \quad \text{hence,} \quad p_f = \phi_a P,$$

where P will be called the *pore fluid pressure* and $\phi_a = dA_f/dA$ is the area fraction of the pores.

In general, the volume fraction ϕ and the area fraction ϕ_a may be different, yet for practical applications, we shall adopt a reasonable assumption that they are the same for simplicity, so that the pore fluid pressure is defined as

$$P = \frac{p_f}{\phi}. \quad (12.39)$$

The pore pressure is an important concept in soil mechanics, we shall see that the definition (12.39) leads to results consistent with the well-known results in soil mechanics.

Pore fluid pressure

From (12.31), we have

$$\psi_f^0 = \widehat{\psi}_f(\theta, \phi d_f, F_s).$$

However, in most applications, it is reasonable to assume that in equilibrium the free energy of the fluid constituent is the same as the free energy of the pure fluid, i.e., we shall assume that

$$\psi_f^0 = \widehat{\psi}_f(\theta, d_f). \quad (12.40)$$

From (12.33), we have

$$\mu = \frac{P}{d_f} + \psi_f^0.$$

Since the true density does not change, $[d_f] = 0$, at an ideal fluid-permeable surface which allows the fluid to go through, it follows that

$$[\mu] = \left[\frac{P}{d_f} \right] + [\widehat{\psi}_f(\theta, d_f)] = \frac{1}{d_f} [P].$$

Therefore the condition (12.37) implies that

$$[P] + d_f \left[\frac{1}{2} (\mathbf{v}_f - \mathbf{u}^*)^2 - (\mathbf{v}_f - \mathbf{u}^*) \cdot \frac{\partial \psi_f}{\partial V} \right] = 0.$$

In particular, if $\mathbf{v}_f = \mathbf{u}^*$, then

$$[P] = 0. \quad (12.41)$$

This result agrees with our physical intuition that if the fluid does not flow from one side to the other side through the pores then the pressures in the pores on both sides must be equal.

A typical fluid pressure measurement is the use of manometric tube attached to the specimen. In the case of a porous body, the junction of the manometric tube and the porous medium can be treated as a fluid-permeable ideal singular surface, where on one side is a solid-fluid mixture and on the other is a pure fluid at rest. Consequently by virtue of the jump condition (12.41) in equilibrium, the fluid pressure measured from the manometric tube is the pore fluid pressure.

Remark. In laying down the jump conditions (12.17) at a singular surface, we have tacitly assumed that there are no surface effects for the mixture as a whole, since the the singular surface is not regarded as a real material surface. Although in the absence of surface effects, it is reasonable to postulate the jump conditions for the mixture, it is not advisable to postulate the similar jump conditions for each constituent. To see this, suppose that for the fluid constituent,

$$[\rho \mathbf{v}_f \otimes (\mathbf{v}_f - \mathbf{u}^*) - T_f] \mathbf{n} = 0, \quad (12.42)$$

at a fluid-permeable surface where the porosity is not continuous, i.e.,

$$[\phi] \neq 0.$$

Now from (12.41) and (12.34), in equilibrium the left-hand side of (12.42) becomes

$$[\rho \mathbf{v}_f \otimes (\mathbf{v}_f - \mathbf{u}^*) - T_f] \mathbf{n} = [p_f] \mathbf{n} = P[\phi] \mathbf{n},$$

which is not zero unless the pore pressure is zero. Therefore, the jump condition (12.42) is not valid in the case of porous media.

12.4 Equations of motion

The equations of motion (12.25)₂ for the fluid and the solid constituents in porous media can be written as

$$\begin{aligned} \phi d_f \dot{\mathbf{v}}_f &= \operatorname{div} T_f + \mathbf{m}_f + \phi d_f \mathbf{g}, \\ (1 - \phi) d_s \dot{\mathbf{v}}_s &= \operatorname{div} T_s - \mathbf{m}_f + (1 - \phi) d_s \mathbf{g}, \end{aligned}$$

where the external body force is the gravitational force \mathbf{g} , and the material derivatives with respect to the constituent have been used,

$$\dot{y}_\alpha = \frac{\partial y_\alpha}{\partial t} + (\operatorname{grad} y_\alpha) \mathbf{v}_\alpha.$$

Let us write the stresses in the following form,

$$\begin{aligned} T_f &= -\phi PI + \bar{T}_f, \\ T_s &= -(1 - \phi)PI + \bar{T}_s. \end{aligned} \quad (12.43)$$

We call \bar{T}_f the *extra fluid stress* and \bar{T}_s the *effective solid stress*, since it reduces to the effective stress widely used in soil mechanics as we shall see later. The equations of motion then become

$$\begin{aligned} \phi d_f \dot{\mathbf{v}}_f &= -\phi \text{grad } P - P \text{grad } \phi + \text{div } \bar{T}_f + \mathbf{m}_f + \phi d_f \mathbf{g}, \\ (1 - \phi) d_s \dot{\mathbf{v}}_s &= -(1 - \phi) \text{grad } P + P \text{grad } \phi + \text{div } \bar{T}_s - \mathbf{m}_f + (1 - \phi) d_s \mathbf{g}. \end{aligned} \quad (12.44)$$

On the other hand, from (12.36), the interactive force \mathbf{m}_f in equilibrium becomes

$$\mathbf{m}_f^0 = P \text{grad } \phi - \phi \mathbf{r}^0, \quad \mathbf{r}^0 = -\frac{P}{d_f} \text{grad } d_f + d_f (\text{grad } \psi_f^0) \Big|_0. \quad (12.45)$$

By canceling out the term $P \text{grad } \phi$ in (12.44) from the interactive force (12.45) leads to the following equations of motion for porous media,

$$\begin{aligned} \phi d_f \dot{\mathbf{v}}_f &= -\phi \text{grad } P + \text{div } \bar{T}_f + (\mathbf{m}_f - \mathbf{m}_f^0) - \phi \mathbf{r}^0 + \phi d_f \mathbf{g}, \\ (1 - \phi) d_s \dot{\mathbf{v}}_s &= -(1 - \phi) \text{grad } P + \text{div } \bar{T}_s - (\mathbf{m}_f - \mathbf{m}_f^0) + \phi \mathbf{r}^0 + (1 - \phi) d_s \mathbf{g}, \end{aligned} \quad (12.46)$$

12.5 Linear theory

Since equilibrium is characterized by the conditions, $\text{grad } \theta = 0$ and $V = 0$, in a linear theory, we shall assume that $|\text{grad } \theta|$ and $|V|$ are small quantities, and that $o(2)$ stands for higher order terms in these quantities.

From (12.34) and (12.43), the extra fluid stress,

$$\bar{T}_f = \phi d_f (\psi_f - \psi_f^0) I + \phi d_f \frac{\partial \psi_f}{\partial V} \otimes V \approx o(2), \quad (12.47)$$

is a second order quantity because the free energy of fluid constituent must be a scalar-valued isotropic function of the vector variable $(V \cdot V)$.

Moreover, we can define the resistive force as

$$\mathbf{r} = \mathbf{r}^0 - \frac{1}{\phi} (\mathbf{m}_f - \mathbf{m}_f^0).$$

It is the force against the flow of the fluid through the medium. Since the non-equilibrium part of the interactive force, $(\mathbf{m}_f - \mathbf{m}_f^0)$, vanishes in equilibrium, we can represent the resistive force as

$$\mathbf{r} = R V + G \text{grad } \theta + \mathbf{r}^0 + o(2). \quad (12.48)$$

The parameter R is called the *resistivity tensor*, and its inverse R^{-1} is called the *permeability tensor*.

The equations of motion (12.46) in the linear theory becomes

$$\begin{aligned} d_f \dot{\mathbf{v}}_f &= -\text{grad } P - \mathbf{r} + d_f \mathbf{g}, \\ (1 - \phi) d_s \dot{\mathbf{v}}_s &= -(1 - \phi) \text{grad } P + \text{div } \bar{T}_s + \phi \mathbf{r} + (1 - \phi) d_s \mathbf{g}. \end{aligned} \quad (12.49)$$

Darcy's law and equilibrium pore fluid pressure

The equation (12.49)₁ for the motion of the fluid is a generalized Darcy's law. Indeed, for stationary case, and only $\mathbf{r} = RV$ is taken into account from (12.48), it reduces to the classical Darcy's law,

$$\mathbf{v}_f - \mathbf{v}_s = -R^{-1}(\text{grad } P - d_f \mathbf{g}).$$

Furthermore, in equilibrium, it becomes

$$\text{grad } P = d_f \mathbf{g}, \quad (12.50)$$

which can be integrated immediately. Suppose that the x -coordinate is in the vertical downward direction, and $\mathbf{g} = g\mathbf{e}_x$. then we have

$$P = d_f g x + P_0, \quad P_0 = P(0). \quad (12.51)$$

This result asserts that the equilibrium pore pressure is the hydrostatic pressure. It agrees with the observation in soil mechanics from experimental measurements that the manometric pressure in the soil is the pressure as if the medium were bulk fluid, unaffected by the presence of the solid constituent in the medium.

We remark that this result is sometimes overlooked in the mixture theory of porous media. It is mainly due to the fact that in the theory of *simple mixture*, which omits the second gradients of deformations as independent constitutive variables, the equilibrium interactive force \mathbf{m}_f^0 is identically zero by constitutive hypothesis (see (12.30)₄). However, from the relation (12.45), \mathbf{m}_f^0 is not a negligible quantity for a body with non-uniform porosity, $\mathbf{m}_f^0 = P \text{grad } \phi$, and it is easy to see that in the absence of this term, the result (12.51) need not follow. This remark, we shall regard as a strong evidence that porous media must be treated as *non-simple mixtures* even for a linear equilibrium theory.

Uplift and effective stress principle

We can obtain an interesting equation for the solid constituent if we eliminate $(\text{grad } P)$ between the two equations in (12.49),

$$(1 - \phi) d_s \dot{\mathbf{v}}_s - \text{div } \bar{T}_s = \mathbf{r} + (1 - \phi)(d_s - d_f) \mathbf{g} + (1 - \phi) d_f \dot{\mathbf{v}}_f. \quad (12.52)$$

From this equation, we notice that the effective stress is not affected by the pore fluid pressure — this is the essential meaning of the *effective stress principle* in soil mechanics.

Note that there are three terms of forces on the right-hand side of the equation (12.52). The first one, $\mathbf{r} = R(\mathbf{v}_f - \mathbf{v}_s) + \dots$, is the usual resistive force of diffusive motion. The second term, $(1 - \phi)(d_s - d_f)\mathbf{g}$, is the weight of solid reduced by the uplift (or buoyancy) from the fluid corresponding to the principle of Archimedes. The importance of uplift in soil structures had been one of the major concern in the development of soil mechanics.

The third term, $(1 - \phi)d_f\dot{\mathbf{v}}_f$, is the inertia force against the displacement of fluid in the motion of the solid through it, which we may think of it as the buoyancy of the inertial force along side with the buoyancy of the gravitational force $(1 - \phi)d_f\mathbf{g}$ in the sense of d'Alembert. The effect of this force seems to be largely unrecognized in the literature.

In the theory of Biot, the relative acceleration was introduced as a part of interactive force between solid and fluid constituents to account for the apparent added mass effect commonly expected. The inertia force considered here seems to correspond to such an effect. However, from the derivation above, it is clear that it is not a part of interactive force, since there is no inertia effect on the motion of fluid (12.49)₁.

12.6 Problems in poroelasticity

Hereafter we shall restrict our attention to mechanical problems (isothermal case) of the theory of elastic porous media, also known as poroelasticity. The governing equations are based on the balance equations of partial mass (12.25)_{1,2} and partial momentum of fluid and solid constituents (12.49), :

$$\begin{cases} (\phi d_f)' + \phi d_f \operatorname{div} \mathbf{v}_f = 0, \\ ((1 - \phi)d_s)' + (1 - \phi)d_s \operatorname{div} \mathbf{v}_s = 0, \\ \phi d_f \dot{\mathbf{v}}_f + \phi \operatorname{grad} P - \operatorname{div} \bar{\mathbf{T}}_f + \phi \mathbf{r} = \phi d_f \mathbf{g}, \\ (1 - \phi)d_s \dot{\mathbf{v}}_s + (1 - \phi) \operatorname{grad} P - \operatorname{div} \bar{\mathbf{T}}_s - \phi \mathbf{r} = (1 - \phi)d_s \mathbf{g}. \end{cases} \quad (12.53)$$

For this system of equations, from (12.43), (12.45), (12.47), and (12.48), we have the following constitutive relations:

$$\begin{aligned} \bar{\mathbf{T}}_f &= \rho_f(\psi_f - \psi_f^0)I + \rho_f \frac{\partial \psi_f}{\partial V} \otimes V \approx o(2), \\ \bar{\mathbf{T}}_s &= T_I + PI - \bar{\mathbf{T}}_f = T_I + PI + o(2), \\ \mathbf{r} &= RV - \frac{P}{d_f} \operatorname{grad} d_f + d_f(\operatorname{grad} \psi_f^0) + o(2). \end{aligned} \quad (12.54)$$

where $o(2)$ stands for higher order terms in $|V|$, and

$$\psi_f = \psi_f(\phi, d_f, F_s, V), \quad P = P(\phi, d_f, F_s), \quad T_I = T_I(\phi, d_f, F_s).$$

Some models of porous media

The governing system (12.53) consists of two scalar and two vector equations, while besides the two vector variables of the motions of fluid and solid constituents, there are

three scalar variables, namely, the two true densities, d_f and d_s , and the porosity ϕ . Therefore, the system is under-determinate, namely, there are less number of equations than the number of independent variables.

Porosity is a microstructural variable of the porous media. To deal with this additional variable, without postulating an additional (evolution or balance) equation for porosity, as proposed in some mixture theories in the literature, there remain some possibilities to formulate deterministic theories from the present theory of porous media. We may consider following models:

1. Firstly, we can regard the porosity as a constitutive quantity, given by a constitutive relation.

Independent variables: $(d_f, d_s, \chi_f, \chi_s)$.

Constitutive variables:

$$\begin{aligned} P &= P(d_f, F_s), & \phi &= \phi(d_f, F_s), & \mathbf{r} &= \mathbf{r}(d_f, F_s, V), \\ \bar{T}_f &= \bar{T}_f(d_f, F_s, V), & \bar{T}_s &= \bar{T}_s(d_f, F_s, V). \end{aligned}$$

Secondly, we can make some assumption of incompressibility of solid or fluid constituent to reduce the number of scalar variables.

2. Incompressible solid constituent: constant d_s .

Independent variables: $(\phi, d_f, \chi_f, \chi_s)$.

Constitutive variables:

$$\begin{aligned} P &= P(\phi, d_f, F_s), & \mathbf{r} &= \mathbf{r}(\phi, d_f, F_s, V), \\ \bar{T}_f &= \bar{T}_f(\phi, d_f, F_s, V), & \bar{T}_s &= \bar{T}_s(\phi, d_f, F_s, V). \end{aligned}$$

3. Incompressible fluid constituent: constant d_f .

Independent variables: $(\phi, d_s, \chi_f, \chi_s)$.

Constitutive variables:

$$\begin{aligned} P &= P(\phi, F_s), & \mathbf{r} &= \mathbf{r}(\phi, F_s, V), \\ \bar{T}_f &= \bar{T}_f(\phi, F_s, V), & \bar{T}_s &= \bar{T}_s(\phi, F_s, V). \end{aligned}$$

4. Incompressible porous medium: constant d_s and d_f .

Independent variables: $(\phi, P, \chi_f, \chi_s)$.

Constitutive variables:

$$\bar{T}_f = \bar{T}_f(\phi, F_s, V), \quad \bar{T}_s = \bar{T}_s(\phi, F_s, V) \quad \mathbf{r} = \mathbf{r}(\phi, F_s, V).$$

Note that even composed with incompressible constituents, the porous body is not necessarily incompressible because the porosity may vary. Moreover, we can regard the pore pressure P as an indeterminate pressure so that the system is deterministic.

12.7 Boundary conditions

Regarding the boundary as a singular surface between the porous body and the external medium, we have the following jump conditions for the mixture as a single body,

$$\begin{aligned} [\rho(\mathbf{v} - \mathbf{u}^*)] \cdot \mathbf{n} &= 0, \\ [\rho\mathbf{v} \otimes (\mathbf{v} - \mathbf{u}^*) - T] \mathbf{n} &= 0, \end{aligned} \quad (12.55)$$

where \mathbf{u}^* is the surface velocity of the boundary.

Therefore, at the boundary of a solid-fluid mixture body, we have $\mathbf{v}_s = \mathbf{u}^*$ and the jump conditions (12.55) becomes,

$$\begin{aligned} [\rho_f V] \cdot \mathbf{n} &= 0, \\ \left[\mathbf{v} \otimes \rho_f V - (T_f + T_s) + \frac{1}{2} \frac{\rho_f \rho_s}{\rho} V \otimes V \right] \mathbf{n} &= 0, \end{aligned} \quad (12.56)$$

Furthermore, the boundary of a porous body can also be regarded as a semipermeable singular surface for the fluid constituent, in other words, the fluid can flow across the boundary and the solid cannot. In a semipermeable boundary, it has been proved that the jump condition of energy is given by (12.37),

$$\left[\mu_f + \frac{1}{2} V^2 - V \cdot \frac{\partial \psi_f}{\partial V} \right] = 0,$$

where $\mu_f = \frac{\partial \rho \psi_f}{\partial \rho_f}$ is the fluid chemical potential. From (12.33), $p_f = \rho_f(\mu_f - \phi_f^0)$, it implies the jump condition for the pore fluid pressure in a porous body,

$$[P] + d_f \left[\psi_f^0 + \frac{1}{2} V^2 - V \cdot \frac{\partial \psi_f}{\partial V} \right] = 0. \quad (12.57)$$

Based on the above jump conditions, we can formulate the boundary condition for the system of partial differential equations. For well-posedness of the problem, two boundary conditions are needed at any point on the boundary, in addition to the proper initial conditions. There are two type of boundary conditions, namely, prescription of the motion of the boundary or the force acting on the free boundary described in the following, where the subindex w denotes the corresponding prescribed value at the exterior side of the boundary.

Dirichlet conditions

These are displacement (velocity) boundary conditions. From (12.56)₁, one can prescribe the solid displacement \mathbf{u}_w and the fluid mass flow m_w ,

$$\mathbf{u}_s = \mathbf{u}_w, \quad \phi d_f(\mathbf{v}_f - \mathbf{v}_s) \cdot \mathbf{n} = m_w,$$

where \mathbf{u}_s is the displacement vector of the solid constituent.

Neumann conditions

Traction boundary conditions must be prescribed according to the relations (12.56)₂ and (12.57). Provided that the fluid mass flux is small enough, one can prescribe the total surface traction \mathbf{t}_w ,

$$T\mathbf{n} = (T_s + T_f)\mathbf{n} = \mathbf{t}_w.$$

If in addition, the equilibrium free energy ψ_f^0 is a function of d_f only, then the second condition implies the continuity of the pore pressure across the boundary,

$$P\mathbf{n} = p_w\mathbf{n},$$

where p_w is the pressure of the adjacent fluid acting on the boundary.

