Published, with minor editorial revisions, as:

J. R. Rice, "Foundations of solid mechanics", in Mechanics and Materials:Fundamentals and Linkages (eds. M. A. Meyers, R. W. Armstrong, and H. Kirchner), Chapter 3, Wiley, 1999, pp. 33-69.

Foundations of Solid Mechanics

James R. Rice

Division of Engineering and Applied Sciences and Department of Earth and Planetary Sciences

Harvard University

January, 1998

1. Introduction

This chapter is devoted to a review of foundations of solid mechanics. It is based on an edited transcript of two lectures by the author on that subject at a 1993 summer school of the NSF Institute for Mechanics and Materials. In university courses, it usually takes a healthy part of the semester to go over the material. So, it is hoped that the student will appreciate that this is indeed a review. It is assumed that the material has to some extent been encountered before, and it is hoped that this quick review will help the reader put things together in his or her own mind. Please see the bibliography for references on the history of the subject, and for further readings.

Figure 1 shows an orthogonal Cartesian reference frame, which we consider to be a Newtonian frame. Coordinates in space are denoted 1, 2, 3; the **e**'s are unit vectors along these coordinate directions. The vector **x**, which denotes a position in space, has components x_1, x_2, x_3. **v** is the velocity vector for this position.

We focus on an arbitrarily chosen region of some body. That region has volume V; its surface is denoted by S. We see in Fig. 1 an element of surface, dS, a unit normal to the surface, $\mathbf{n}$, and an element of volume, dV.

The quantities which are going to enter the equations of motion are, of course, the linear and angular momentum. These are taken as integrals over the region of mass density times velocity integrated over volume, and for the angular momentum, the cross product of mass density times velocity integrated over volume:

$$\textit{Linear momentum} \equiv \mathbf{P} = \int_V \rho \mathbf{v} \, dV \qquad (1)$$

$$\textit{Angular momentum (relative to the coordinate origin)} \equiv \mathbf{H} = \int_V \rho \mathbf{x} \times \mathbf{v} \, dV \qquad (2)$$

The student may reasonably ask: "In microscopic terms, what do these quantities mean?" Of course, mass density is an average of mass per unit volume, and we assume in a theory of the type to be discussed that it is taken over spatial and temporal scales that are large enough to be independent of any fluctuations at the microscopic level. If we ask "What does velocity mean?", especially when we consider the large random fluctuations at the molecular scale, the precise definition of velocity is the vector flux of mass (again, well-defined if taken over large enough scales) divided by the mass density. So, in that sense, a velocity vector is also a well-defined quantity.

2. Newton-Euler Overall Equations of Motion

The equations of motion (called here the Newton-Euler equations of motion, because the independent status of the torque part of these equations, at least in a macroscopic theory not reduced to interacting point masses, was recognized by Euler) state that the total force is the rate of change of linear momentum and the total torque, or moment, is the rate of change of angular momentum. So, if we take the rates of change of the expressions above (Eqns. 1 and 2) for linear and angular momentum,

$$\text{Total } \textit{force} \equiv \mathbf{F} = d\mathbf{P}/dt, \tag{3}$$

$$\text{Total } \textit{torque} \text{ or } \textit{moment} \text{ (relative to the coordinate origin)} \equiv \mathbf{M} = d\mathbf{H}/dt \tag{4}$$

We note that the integral for which we are doing the calculation, let us say of linear momentum of (1), involves mass density times an element of volume. That is an invariant in the motion. So, we do not have to take its time derivative, and we end up just taking the time derivative of velocity; this gives us, for example, mass density times acceleration **a** of material points in the integral over the volume, which explains the first integral term in

$$\mathbf{F} = d\mathbf{P}/dt = \int_V \rho\,\mathbf{a}\,dV + \int_S \mathbf{T}^{mom.flux} dS \tag{5}$$

But in fact, there is really a second term, because there is a microscopic motion, in general, relative to the mass-averaged macroscopic motion. The relative motion causes some momentum flux $\mathbf{T}^{mom.flux}$ per unit area, across surfaces *S*, which themselves move with the mass-averaged motion. In a similar way, the rate of change of angular momentum can be calculated and we have the expression:

$$\mathbf{M} = d\mathbf{H}/dt = \int_V \rho\,\mathbf{x}\times\mathbf{a}\,dV + \int_S \mathbf{x}\times\mathbf{T}^{mom.flux} dS \tag{6}$$

<u>3. Cauchy Stresses and Local Equations of Motion in Terms of Stress</u>

The important hypothesis of Cauchy is the foundation of continuum mechanics and the theory of stress. The idea is that all interactions of an arbitrarily selected region of the body with its surroundings (Fig. 1) can be divided into body forces, **f** per unit volume, and surface forces. $\mathbf{T}^{force}$ is a force vector per unit area acting on the surface of the region. Force and moment derive from those terms:

$$\mathbf{F} = \int_S \mathbf{T}^{force}\, dS + \int_V \mathbf{f}\, dV \tag{7}$$

$$\mathbf{M} = \int_S \mathbf{x} \times \mathbf{T}^{force}\, dS + \int_V \mathbf{x} \times \mathbf{f}\, dV \tag{8}$$

When we equate force and moment of (7) and (8) to the rates of change of linear and angular momentum, given in (5) and (6), we arrive at the final equations of motion.

In those equations the stress (or traction) **T** vector is now defined, as the force vector combined with the term which came from the momentum flux:

$$\mathit{Stress\ vector} \equiv \mathbf{T} = \mathbf{T}^{force} - \mathbf{T}^{mom.\,flux} \tag{9}$$

In a solid at low temperature, virtually the entire stress is made up of the force term; however, in a gaseous system, the momentum flux term is the most important part of the stress. The equations of motion expressing the linear momentum principle are therefore, from (5), (7) and (9),

$$\int_S \mathbf{T}\, dS + \int_V \mathbf{f}\, dV = \int_V \rho\, \mathbf{a}\, dV \tag{10}$$

In a similar way we express the angular momentum principle from (6), (8) and (9),

$$\int_S \mathbf{x} \times \mathbf{T}\, dS + \int_V \mathbf{x} \times \mathbf{f}\, dV = \int_V \rho\, \mathbf{x} \times \mathbf{a}\, dV \tag{11}$$

We will now examine the consequences of requiring that (10) and (11) hold for every conceivable region. First note that we could, upon accepting Cauchy's hypothesis, define a stress vector on any surface orientation at all through a point in the solid. The way that the stress components get introduced is thus shown in Figure 2. If we decide, for example, that we will focus on a little cut in the solid with face that points in the positive 1 direction, then we will have a stress vector **T** acting on that face. Because the face normal points in the 1 direction, the superscript 1 is attached to that vector, and we call it $\mathbf{T}^{(1)}$. But that vector, of course, has components along the three coordinate axes, and those components have the notation σ_{1j} where $j = 1, 2, 3$. More generally, we will let σ_{ij} denote components of the stress vector $\mathbf{T}^{(i)}$ associated with a cut face with normal which points in the positive *i* direction. Thus the first subscript of the σ's denotes the face and the second subscript gives the direction, so that σ_{11} is a normal stress, and σ_{12} and σ_{13} are shear stresses, and they are the components of the stress vector $\mathbf{T}^{(1)}$ (see Fig. 2). Thus

Stress vector associated with surface element pointing in positive i direction:

$$\mathbf{T} \equiv \mathbf{T}^{(i)} = \sigma_{i1}\,\mathbf{e}_1 + \sigma_{i2}\,\mathbf{e}_2 + \sigma_{i3}\,\mathbf{e}_3 \qquad (i = 1,\ 2,\ 3)$$

$$\textit{Stress components:} \quad \sigma_{ij} \qquad (i,\ j = 1,\ 2,\ 3) \tag{12}$$

$$i:\ \text{face} \quad , \quad j:\ \text{direction}$$

How about faces that point in the negative coordinate direction? In particular, suppose we focus on a face that points in the negative 1 direction; such is shown in Figure 3. In order to find out what the traction $\mathbf{T}^{(-1)}$ is on that face, what one does is apply the linear momentum equation (10) to that thin slab, let its thickness go to zero, then let its areal extent go to zero. That leads us to the conclusion that the volume terms go to zero faster than surface terms in (10), and leaves us with the sum of traction vectors adding vectorially to zero.

$$\mathbf{T}^{(-1)} + \mathbf{T}^{(1)} = 0 \,, \quad \mathbf{T}^{(-1)} = -\mathbf{T}^{(1)} \,; \tag{13}$$

So, the traction vector on the negative face is just minus the traction vector on the positive face. Sometimes this is called the principle of action-reaction. Hence the traction vector on any face oriented in the negative i direction is

$$\mathbf{T}^{(-i)} = -\mathbf{T}^{(i)} = -\sigma_{i1}\,\mathbf{e}_1 - \sigma_{i2}\,\mathbf{e}_2 - \sigma_{i3}\,\mathbf{e}_3 \qquad (i = 1,\ 2,\ 3) \tag{14}$$

So far we have thought of stress components merely as components of traction, or stress, vectors on faces that point in the three coordinate directions. Suppose we ask the question that Cauchy asked himself in the 1820s: "What if we now take an arbitrarily oriented face in the solid; what stress vector acts on that face, the face with the unit normal **n**?" We can construct a tetrahedron with faces in the three other coordinate directions, which are negative coordinate directions in the illustration (Fig. 4), and we can draw all the stress vectors. We again use the linear momentum equation, (10), which contains both surface and volume integrals. If we divide the whole equation by the area of that inclined face, and then let the size of the tetrahedron go to zero, we conclude again that the volume integrals all go to zero faster than the surface integrals. The only term that survives in the limit is an integral of the tractions over the faces, which integral must be zero. Rather than writing the tractions multiplied by the elements of area of the faces with which they are associated, we may note that the area of a face like ΔS_1, pointing in the negative 1 direction, is just n_1 times the area the area ΔS of the inclined face, and similarly for other faces. Thus (10) requires that

$$\mathbf{T} + n_1\,\mathbf{T}^{(-1)} + n_2\,\mathbf{T}^{(-2)} + n_3\,\mathbf{T}^{(-3)} = 0 \quad (\text{using } \Delta S^{(i)} / \Delta S = n_i)\,. \tag{15}$$

If we use (14) and solve for the traction vector **T**, and in particular take its component T_j in the *j* direction, this is the sum of components of unit normal times stress components acting in the *j* direction:

$$T_j \equiv \mathbf{e}_j \cdot \mathbf{T} = n_1 \sigma_{1j} + n_2 \sigma_{2j} + n_3 \sigma_{3j} = \sum_{i=1}^{3} n_i \sigma_{ij} \quad (j = 1,\ 2,\ 3) \tag{16}$$

(*Summation convention*: $T_j = n_i \sigma_{ij}$)

Almost always in this subject, when we have a summation, we are summing on an index that happens to be repeated. So, a commonly adopted notation is one in which we rewrite Equation 16 just dispensing with the summation sign. Apparently, it was Einstein who first decided that he was tired of writing summation symbols and such is sometimes called the Einstein summation convention. In this Chapter, summation symbols are almost always kept, but the reader should understand that they could be dispensed with.

There is an important consequence of this equation (16) for the stress vector on an arbitrarily inclined face. We could, if we wanted, have chosen some other coordinate system than the one shown in Figure 4, say a coordinate system — we will call it a primed system x_1', x_2', x_3' — for which one of the axes runs perpendicular to the face of the tetrahedron. Then, the stress vector on that face would just be a stress vector associated with one of the coordinate faces of that primed coordinate system. This consideration lets us at once discover that stress is, in fact, a second rank tensor. If we choose a primed set of axes so that one of these coordinates, say the *k* coordinate of the primed set, is normal to the face of the tetrahedron, then

$$\mathbf{T} = \mathbf{T}^{(k')} = \sigma_{k1}' \mathbf{e}_1' + \sigma_{k2}' \mathbf{e}_2' + \sigma_{k3}' \mathbf{e}_3' \tag{17}$$

We can then extract, say, the *k, l* component of stress in the primed system, σ'_{kl}, by just projecting this stress vector onto the *l* direction of the primed system by the scalar product:
$\sigma'_{kl} = \mathbf{e}'_l \cdot \mathbf{T}^{(k')}$

This leads at once to the transformation law of stress when we realize that **T** is expressible in terms of the unprimed stresses by (16):

$$\textit{Since } \mathbf{e}'_l \cdot \mathbf{T} = \sigma'_{kl} \ , \ \ \sigma'_{kl} = \sum_{i=1}^{3} \sum_{j=1}^{3} \alpha_{ki}\, \alpha_{lj} \sigma_{ij} \ \text{(where } \alpha_{pq} = \mathbf{e}'_p \cdot \mathbf{e}_q \text{)} \tag{18}$$

The stresses in the primed system are related to the stresses in the original system by a transformation law, a double sum on *i* and *j* ; again, these are repeated indices and therefore we could dispense with writing out the summations. The alphas, α_{pq}, are the scalar products of unit vectors along the *p* direction of the primed system and the *q* direction of the unprimed, so α_{pq} just defines the cosine of the angle between those axes.

This expression (18) is indeed the transformation which qualifies a quantity as being a second-rank tensor; so, we see that stress components σ_{ij} form a tensor. For comparison, first-rank tensors, or vectors, and also the Cartesian coordinates, transform as a similar relation with a single term:

$$\textit{Coordinates related by } \ x'_k = \sum_{i=1}^{3} \alpha_{ki}\, x_i \, ; \tag{19}$$

The matrix that defines such a transformation of coordinates is the so-called "orthogonal transformation"; if we assemble the α's into a three-by-three matrix, it satisfies

$$[\alpha]^T[\alpha] = [\alpha][\alpha]^T = [I] \tag{20}$$

which is another way of saying that, for $[\alpha]$, the transpose is in fact the same as the inverse. That is true of an orthogonal transformation matrix: the inverse of the matrix is its transpose.

So far we have applied the linear momentum principle of (10) to a tetrahedron, but now we could also apply it to some arbitrarily selected region of any body. Indeed, when we do that, we extract the differential equations of motion. In the steps that follow, we consider an arbitrary region, with surface S. We make use of Cauchy's relation (16) for the stress vector on an inclined face in terms of stress components at every point along the surface. Remember that the linear momentum equation (Eqn. 10) contains a surface integral of the traction vector plus a volume integral of body force, and on the right hand side a volume integral involving density times acceleration. The surface integral transforms to a volume integral by use of Gauss' divergence theorem, explained in Chapter 1. We simply use the Cauchy relation to rewrite the traction in (10) as a set of products of components of the unit normal vector to the surface times different terms, and then apply the theorem. If we have a surface integral involving, let us say, the first component of the unit normal to that surface, it transforms to a volume integral involving derivative with respect to coordinate in that first direction. Thus:

Using Cauchy tetrahedron relation $T_j = \sum_{i=1}^{3} n_i \sigma_{ij}$ *, and divergence theorem,*

$$\int_S T_j \, dS = \int_S (n_1 \sigma_{1j} + n_2 \sigma_{2j} + n_3 \sigma_{3j}) \, dS = \int_V \left(\frac{\partial \sigma_{1j}}{\partial x_1} + \frac{\partial \sigma_{2j}}{\partial x_2} + \frac{\partial \sigma_{3j}}{\partial x_3} \right) dV . \tag{21}$$

The linear momentum equation (Eqn. 10) already contains two other volume integrals; one involves the body force, the other involves the acceleration. So if we now demand that this linear momentum equation hold for absolutely any choice of volume, we conclude that

$$\frac{\partial\sigma_{1j}}{\partial x_1}+\frac{\partial\sigma_{2j}}{\partial x_2}+\frac{\partial\sigma_{3j}}{\partial x_3}+f_j\equiv\sum_{i=1}^{3}\frac{\partial\sigma_{ij}}{\partial x_i}+f_j=\rho a_j \quad (j=1,\ 2,\ 3) \tag{22}$$

Some common notational short-cuts in writing (22) are as follows:

Summation convention: $$\frac{\partial\sigma_{ij}}{\partial x_i}+f_j=\rho a_j \tag{23}$$

Comma notation: $$\sigma_{ij,i}+f_j=\rho a_j \tag{24}$$

The comma means derivative with respect to the coordinate whose index follows it.

These equations (22), or (23) or (24), are the equations of motion for a continuum. We have three of them, one associated with each coordinate direction.

We also have an angular momentum equation (11), and the details of similarly going through its consequences are omitted here. Its result is that when we use (16) and the divergence theorem to rewrite the surface integral in (11), and use the linear momentum equation to simplify things, the angular momentum equation tells us one piece of information only, and that is that the stress tensor is a symmetric tensor:

$$\sigma_{ij}=\sigma_{ji} \quad (i,\ j=1,\ 2,\ 3) \tag{25}$$

If we interchange indices, we end up with the same value. Put another way, if we interpret Figure 2 as showing a little lump of material, what this is saying is that shear stresses on adjacent faces, say σ_{12} and σ_{21} are equal to one another, and we can see that that indeed is a result that relates to balance of torques which would otherwise generate an angular momentum.

4. Principal Directions and Principal Stresses

Once we know that the stress tensor is symmetric, we can infer additional things about stress, and in particular we can infer that a set of principal axes exist. Any arbitrary stress tensor has the property that there exist three mutually orthogonal directions so that the stress vectors **T** associated with each of those directions are purely normal to the face with which they are associated. In fact, that is because the stress tensor is symmetric; any symmetric three-by-three matrix has the following representation: if we give an appropriate rotation of the coordinate system, the matrix can be made purely diagonal. That is what we do here with stress. The principal stresses are found in an attempt to answer either of the questions which follow. Remember, we have a way of associating a stress vector **T** with any unit vector **n** denoting an orientation through some material point; that is the relation (16) from the Cauchy tetrahedron analysis. So, one question we could ask is: "Do there exist directions **n** such that **T** is parallel to **n**?" If these do exist, then the stress vector is perpendicular to the face; hence, there are no shear stresses on it. Another question we might ask, and it turns out to have an identical answer, is "Are there any directions **n** so that the normal component of the stress vector σ_n, (= $\mathbf{n} \cdot \mathbf{T}$, which is the stress vector projected onto the direction normal to the face) is an extremal (an extremal meaning a local minimum, maximum or saddle point), as we vary the direction **n**?".

The answer to both those questions is given by a simple eigenvalue problem,

$$\sum_{j=1}^{3} \sigma_{ij} n_j - \sigma n_i = 0 \quad (i = 1,\ 2,\ 3) \tag{26}$$

The summation defines the *j* component of the traction vector by (16) and if we temporarily shift the second term to the other side of the equation to get the (vector) form $\mathbf{T} = \sigma\mathbf{n}$, we see this says the traction vector is in the same direction as the unit normal and has magnitude σ. The equations (26) are a set of equations in the components of the unit normal, a homogeneous set of equations. One solution is that we should just set the n's to zero, but that is unacceptable of course because **n** is meant to be a unit vector. We get the other, relevant, solutions when we set the determinant of coefficients of the n's to zero:

$$\textit{Solutions exist for } \sigma = \sigma_n = \sigma_{\mathrm{I}}, \sigma_{\mathrm{II}}, \sigma_{\mathrm{III}}, \textit{ which are roots of}$$

$$\det\begin{bmatrix} \sigma_{11}-\sigma & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22}-\sigma & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33}-\sigma \end{bmatrix} = -\sigma^3 + I_1\sigma^2 + I_2\sigma + I_3 = 0; \qquad (27)$$

$$\sigma_1 \le \sigma_n \le \sigma_{\mathrm{III}}.$$

Here

$$I_1 = \sum_{i=1}^{3} \sigma_{ii}, \quad I_2 = -\frac{1}{2} I_1^2 + \frac{1}{2}\sum_{i=1}^{3}\sum_{j=1}^{3} \sigma_{ij}\sigma_{ji}, \quad I_3 = \det[\sigma], \qquad (28)$$

The three σ's which make the above determinant vanish are the three principal stresses. The associated **n**'s are the three principal directions, and as sketched in Figure 5, they are mutually orthogonal.

In writing out the determinant in (27), one sees that, because it is a three-by-three determinant , that there will be a σ^3 term; there will also be lower-order terms and their coefficients, I_1, I_2, I_3. If we think a moment about it, all of the stress components are going to be involved in determining that set of coefficients, but on the other hand those coefficients cannot possibly depend on what we arbitrarily chose as a set of coordinate axes to which we refer stresses. So, the coefficients, I_1, I_2, I_3, are invariant quantities (invariant to the choice of our reference axes); no matter what coordinate system we choose, the stresses at a point always give the same value for these quantities.

In the theory of isotropic materials, theories of elasticity or elastic-plastic yielding for isotropic materials, invariants play an important role. If, for example, a medium is isotropic and we try to express its strain energy in terms of stress, we would know that the strain energy could at most depend on those three invariant quantities.

5. Formulation of Mechanics Problems

In general, if we want to formulate problems in the mechanics of materials, we need at least three conceptual ingredients. That "at least three" reduces to "exactly three" for problems in which we neglect temperature effects generated by deformation, in which there are no electromagnetic couplings, and in which there is a single medium with no diffusion of one constituent of the medium relative to another (as we would have, for example, for an alloy at high temperatures or for a fluid-infiltrated soil or rock).

The first such ingredient involves the equations of motion, or in simpler problems, the equations of static equilibrium. We have just been through these equations; they are the equations of motion in terms of stress, which we have derived in Section 3 as equations (22) and (25). The second ingredient one has to consider is the geometry of deformation: the way that strains of the material relate to gradients of the displacement field, and considerations of compatibility of strain. In Sections 6 to 9, we examine the geometry of deformation. The third ingredient is the stress-strain relation. That is a large subject, addressed also in other chapters here. We shall consider the general framework and the specific example of linear elasticity in Sections 12 and 13.

6. Geometry of Deformation

For many of the problems we address in solid mechanics, we can be rather casual about geometric matters and address them within the so-called small strain or infinitesimal strain theory. However, if we are considering something like a buckling problem or a finite extrusion of material through a die, or many other problems, we have to understand deformation in a way that is appropriate for deformations of arbitrarily large magnitude. So that is the set-up that we will begin with here.

When we discuss such deformations, a very typical notation is to use upper case **X** to denote the position vectors of material points in a reference configuration, say at time t = zero. We consider that reference configuration as undeformed. That is to say, when we introduce measures of strain, they are going to be such that they will be zero in that reference configuration. The reference configuration is often taken as an unstressed configuration, but that need not be the case and there are many circumstances when it is convenient to take the reference configuration as pre-stressed. Thus

$\mathbf{X} = (X_1, X_2, X_3)$: *Denotes position vectors of material points at time* t = 0

(*in the <u>reference</u> configuration, considered as <u>undeformed</u>*). (29)

$\mathbf{x} = (x_1, x_2, x_3)$: *Denotes position vectors at time* t *(in <u>current</u>, or <u>deformed</u>, configuration).*

A deformation history is simply a mapping of these three components of initial position into the three components of current position at time t, and of course, at time zero, **x** and **X** agree with one another. The displacement is the difference between current and initial positions, and velocities and accelerations can be expressed as below:

Deformation history: $\mathbf{x} = \mathbf{x}(\mathbf{X},t)$; $\mathbf{x}(\mathbf{X},0) = \mathbf{X}$.

(30)

Displacement vector: $\mathbf{u} = \mathbf{x}(\mathbf{X},t) - \mathbf{X}$

$$(also,\ \mathbf{v} = \partial\mathbf{x}(\mathbf{X},t)/\partial t \quad and \quad \mathbf{a} = \partial^2\mathbf{x}(\mathbf{X},t)/\partial t^2)$$

Recall that from a microscopic point of view, **v** is well defined in terms of mass flux and density averages over large enough local space and time scales to be rid of fluctuation effects. Thus we should regard $\partial\mathbf{x}(\mathbf{X},t)/\partial t$ as the physically well-defined quantity, and regard the mapping $\mathbf{x} = \mathbf{x}(\mathbf{X},t)$ as, strictly, being well-defined only by integrating $\partial\mathbf{x}(\mathbf{X},t)/\partial t$ in time.

In discussing strains, an important ingredient is the deformation gradient tensor F_{ij}. Its definition is very simple. Suppose we take some little vector $d\mathbf{X}$, with components dX_j, in the reference configuration. The vector $d\mathbf{X}$, pointing in some particular direction in space, goes over to a vector in another direction in space, and of another length, in the deformation. The vector it goes into is $d\mathbf{x}$, and the components of those two vectors are related by the deformation gradient:

$$\textit{Deformation gradient}\ \ F_{ij}:\ \ dx_i = \sum_{j=1}^{3} F_{ij}\,dX_j \quad (\{dx\} = [F]\{dX\}) \tag{31}$$

It is sometimes convenient to rewrite that as a matrix equation so the columns represent the elements of the material fiber that is being considered. The components of F are just the derivatives of the current spatial coordinates with respect to the reference or material coordinates,

$$F_{ij} = \partial x_i/\partial X_j = \delta_{ij} + \partial u_i/\partial X_j\,. \tag{32}$$

This involves the Kronecker delta, δ_{ij}; it is a quantity which vanishes when the indices differ from one another, and which is equal to unity if the indices agree.

The theory of small strain, that is used in classical elasticity, is of course the theory that results if we have small stretches of all fibers, and also — and this is very important — small rotations. That is the case where all components of displacement derivatives are extremely small compared to one,

$$\left|\partial u_i / \partial X_j\right| << 1 \tag{33}$$

Even if we meet this criterion, there are certain problems for which we still have to be cautious in using the small strain theory; the classic one is Euler's problem of buckling of a thin strut. We really cannot get that right if we think exclusively in terms of small strain theory even though the strains, at least in the early stages of buckling, are generally extremely small strains.

Figure 6 shows a picture of a block of material that, in its reference configuration shown by dashed lines, was a cube with edges along the coordinate directions. A very special kind of strain is considered first: a purely extensional strain which just stretches along the coordinate directions, or perhaps compresses along some directions, but which keeps sides perpendicular to one another. Figure 6b is the same block in the current configuration. A ratio of a length in the deformed state to a length in the undeformed state is called a stretch ratio:

$$\textit{Stretch ratios}:\ \lambda_1 = \Delta x_1 / \Delta X_1,\ \lambda_2 = \Delta x_2 / \Delta X_2,\ \lambda_3 = \Delta x_3 / \Delta X_3 \tag{34}$$

It may seem very special to focus only on extensional strains, but a very important theorem is this: If we consider an absolutely general state of deformation at a point in the material, we could always have chosen an orientation of such a little cube (like shown in Figure 6) at that point, but now not generally lined up with coordinate directions, so that that cube undergoes purely extensional strain with no shear strains. That is called the principal strain orientation. So, it is worthwhile focusing on these extensional strains.

How do we measure strain? We have already introduced the idea of stretch ratios, the λ's as lengths in the deformed configuration divided by lengths in the undeformed. We can write one of these equations for each of the three directions, like in (34). The most common understanding of strain is change in length divided by initial length. So, if we would make that definition of strain, and ask how much did our block stretch in the 1 direction, we would want the change in length in that direction divided by initial length. That is, the Δx_1 minus ΔX_1 gives us change in length, and that also is the difference in displacement from one edge to another of the block as we move in the 1 direction. And we divide that by the initial length to get:

$$\begin{aligned} &\textit{Common definition of strain}: E_{11} = \textit{(change in length)/(initial length)} \\ &\qquad = (\Delta x_1 - \Delta X_1)/\Delta X_1 = \Delta u_1 / \Delta X_1 = \lambda_1 - 1 \end{aligned} \tag{35}$$

All other definitions of strain that we make in this subject agree with this common definition when the stretches are very small (when the λ's are very near 1). But, in fact, we can define an infinite number of strain measures and at least three or four are current in the literature on continuum mechanics.

The idea is this: We define a strain by the function g of the stretch ratio by

$$E_{11} = g(\lambda_1) \;\; \textit{where} \;\; g(1) = 0 \;\; \textit{and} \;\; g'(1) = 1 \tag{36}$$

The restrictions on $g(\lambda)$ make the definition of strain vanish in the reference state and also agree with the simple change in length over initial length definition when the λ is sufficiently close to 1. They make all such strain measures agree with the so-called infinitesimal strain in the right limit. Other than that, $g(\lambda)$ is arbitrary except that we will also want to choose functions so that $g'(\lambda) > 0$ for all λ, $0 < \lambda < \infty$.

One such strain measure, which we call here a strain based on the change of the metric tensor, (it will be explained shortly what justifies that terminology), and is sometimes called the "Green strain", is

$$E_{11}^{M} = (\lambda_1^2 - 1)/2 \tag{37}$$

Another definition of strain is the logarithmic strain. This is very commonly used in discussing plastic flow, and is

$$E_{11}^{L} = \ln(\lambda_1) \tag{38}$$

To review, because we are going to depart for a moment from this topic and then come back to it: What we have looked at are purely extensional strains, and we have found different ways of measuring the strain, so far in a single direction. Ultimately, we have the job of measuring strains in arbitrary coordinate directions.

We need to define another type of strain, a simple shear strain, where the undeformed element is shown by dashes and the deformed by full lines in Figure 7. This looks extremely different from the extensional strains, but as suggested already, there would have been a special orientation of a square element that we could have identified in the undeformed state so that that element would have undergone extensions only (positive in one direction, negative in the other) without shear.

7. Infinitesimal Strains

With that as background, let us spend a moment on the classical infinitesimal strain that was introduced by Cauchy as part of his great work of the 1820s, basically introducing linear elasticity as we know it today. He was preceded in this derivation by Euler, who had already worked out for a fluid a corresponding definition of what we would call a rate of deformation tensor. The idea is given in Figure 8, which shows two line elements labeled as dX_1 and dX_2, line elements or material fibers as they existed in the undeformed or reference configuration. Those line elements have gone over into the two solid line elements in the deformation. If we look at the origin of those elements, it is displaced by u_1 in the direction 1 and by u_2 in the direction 2. But then, if we look at extremities of the elements, they have slightly different displacements. The right hand end of what was, initially, dX_1 has displaced in the 1 direction by u_1, the same amount that the origin displaced, plus a small increase because we have moved the distance dX_1 in the 1 direction so as to accumulate the additional displacement $dX_1(\partial u_1 / \partial X_1)$. In a similar way, the upper end of the element that was dX_2 has moved upwards a different amount than the origin has moved. That amount of movement is identified in Figure 8. We want to make a definition of strain that is appropriate when we have small strains and rotations, that is, when all derivatives of displacements are very much smaller than one. (Incidentally, in that case, we can with impunity just replace derivatives with respect to material coordinates with derivatives with respect to spatial coordinates, above and in other equations, like the equations of motion.) Then, the changes in length of an element, used to define strain, can just be equated to those additional displacements reckoned above.

Dividing those small additional displacements by the initial length of the fibers gives the extensional infinitesimal strains. In the 1 direction, that strain is called ε_{11}. Thus we write as follows:

$$\textit{Extensional strains}: \ \varepsilon_{11} = \partial u_1 / \partial X_1 \ \ \textit{and} \ \ \varepsilon_{22} = \partial u_2 / \partial X_2 \qquad (39)$$

Shear strains are commonly understood as angle changes between initially orthogonal lines. Gamma denotes such an angle change. This turn outs to be twice the so-called tensor shear strain:

$$\textit{Shear strain}: \ \gamma_{12} = 2\,\varepsilon_{12} = \partial u_2 / \partial X_1 + \partial u_1 / \partial X_2 \tag{40}$$

If we carefully look at the diagram (Fig. 8), we will see that this sum adds up to the total angle change between initially orthogonal lines at least within the approximations, valid in the case considered, that $\sin\theta \approx \tan\theta \approx \theta$ and that all line length changes are very small. The general expression which captures these and indeed defines all the strains is:

$$\textit{General expression}: \ \varepsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_j}{\partial X_i} + \frac{\partial u_i}{\partial X_j}\right) \qquad (i, j = 1,\ 2,\ 3) \tag{41}$$

So, Eqn. 41 defines six components of strains — six, not nine, because the definition is symmetric in *i* and *j*. We have three extensional strains and six shear strains, but three of them are identical to the other three; so, effectively, we have three shear strains.

If we work out the strains by measuring displacements and coordinates relative to some rotated coordinate system, we will quickly discover that the strains ε'_{kl} in that rotated system are related to the strains ε_{ij} by the standard law of tensor transformation, the same as in equation (18). So ε is a second rank tensor. Also, because ε is symmetric, there exist principal directions relative to which there are extensional strains only and not shear strains, and the principal values include the least and greatest extensional strains experienced by fibers of any orientation.

8. Compatibility Conditions

Remember that we have six strain components in (41) but they are derived from only three components of displacement, by taking derivatives. So it is not possible that these strain components could vary in a completely arbitrary way from point to point in a medium. We cannot have six functions of position vary in an arbitrary way if they are really defined by only three functions of position. So, there must be equations which restrict the way strains can vary in space. These are called compatibility equations. They are needed when we try to formulate the basic equations of elasticity in terms of stresses as variables, rather than displacements. A typical compatibility equation — one which is the only one needed for plain strain and plain stress analysis — is written out here:

$$\textit{Compatibility}: \; \textit{For example}, \; \partial^2 \varepsilon_{22} / \partial X_1^2 + \partial^2 \varepsilon_{11} / \partial X_2^2 = 2 \partial^2 \varepsilon_{12} / \partial X_1 \partial X_2 \tag{42}$$

No attempt to derive it is made here, but the reader can easily check its validity: If the strains are written in terms of displacements, this will emerge as a simple identity. It is an example of the way these strains must be constrained in their spatial variation to make sure that, in this case, the three strain components which appear in (42) are indeed derivable from the two displacements, u_1 and u_2.

9. Finite Strains

Let us come back to the subject of finite strain and ask how do we characterize a finitely deformed material and how do we write down expressions for strains. The first thing to understand on this topic is the polar decomposition theorem. Remember F (Eqns. 31 and 32), the deformation gradient matrix, the matrix by which we multiply a small vector $d\mathbf{X}$ in the reference configuration to produce a vector $d\mathbf{x}$ in the deformed configuration.

A basic theorem is that any such deformation gradient can be written out as a product of two terms (polar decomposition theorem): an orthogonal transformation matrix *R,* which represents a rigid rotation, and a pure deformation matrix *U* which is its own transpose, that is to say, it is a symmetric matrix. The rotation *R* is just like the coordinate rotation matrix α, which was seen in transformations of stress (Eqn. 18). Thus, in a matrix notation,

$$[F] = [R][U], \; where: \tag{43}$$

$$[R] = rigid \;\; rotation \; ([R]^T [R] = [R][R]^T = [I], \; \det[R] = 1); \tag{44}$$

$$[U] = pure \;\; deformation \; ([U] = [U]^T, \; \det[U] > 0)\,. \tag{45}$$

So, if we think of writing

$$\{dx\} = [F]\{dX\} = [R][U]\{dX\} = [R]([U]\{dX\})\,, \tag{46}$$

the multiplication by *R* means nothing more than take whatever vector is given in the parentheses ( ) and simply rotate it rigidly by *R* without stretching it; and indeed if we consider other vectors in that neighborhood, they all rotate the same. So, we neither stretch nor change angles between vectors by multiplying by *R*.

What are some properties of the pure deformation matrix *U*? It is a symmetric matrix. Let us produce a vector $d\hat{\mathbf{x}}$ by symmetric matrix times $d\mathbf{X}$,

$$\{d\hat{x}\} = [U]\{dX\}\,. \tag{47}$$

Because U is symmetric, there exist three mutually orthogonal directions denoted by unit vectors $\mathbf{N}^{(\mathrm{I})}$, $\mathbf{N}^{(\mathrm{II})}$ and $\mathbf{N}^{(\mathrm{III})}$ which are just the eigenvectors of the matrix $[U]$ and are the principal directions of deformation. They have the property that if the initial material fiber d$\mathbf{X}$ that we consider lies along one of those directions, then so does the $d\hat{\mathbf{x}}$ that is produced. So, there is no shearing relative to those directions. Stretch ratios along $\mathbf{N}^{(\mathrm{I})}$, $\mathbf{N}^{(\mathrm{II})}$, $\mathbf{N}^{(\mathrm{III})}$ are λ_{I}, λ_{II}, λ_{III} (the *eigenvalues* of $[U]$) and are the *extremal* stretch ratios.

Indeed, we can write the pure deformation U in the form

$$U_{ij} = \lambda_{\mathrm{I}} N_i^{(\mathrm{I})} N_j^{(\mathrm{I})} + \lambda_{\mathrm{II}} N_i^{(\mathrm{II})} N_j^{(\mathrm{II})} + \lambda_{\mathrm{III}} N_i^{(\mathrm{III})} N_j^{(\mathrm{III})} \tag{48}$$

Let us go on to a general deformation $[F]$, so that

$$\{dx\} = [F]\{dX\} = [R][U]\{dX\} = [R]\{d\hat{x}\} \tag{49}$$

Fibers having the three mutually orthogonal directions $\mathbf{N}^{(\mathrm{I})}$, $\mathbf{N}^{(\mathrm{II})}$, $\mathbf{N}^{(\mathrm{III})}$ in the reference configuration undergo extensional strain but have no shearing between them. Thus they are rotated by $[R]$ relative to the reference configuration.

A relation that will be used later when we look at a specific definition of finite strain is that

$$[F]^T [F] = [U][U] \equiv [U]^2, \tag{50}$$

which follows from (43) and (44). So, if we write out the ij component of $[F]^T[F]$ in the subscripted notation, we see that it is just the sum of λ^2 times the products of unit vectors:

$$([F]^T[F])_{ij} = \sum_{k=1}^{3} F_{ki}F_{kj} = \lambda_{\text{I}}^2 N_i^{(\text{I})} N_j^{(\text{I})} + \lambda_{\text{II}}^2 N_i^{(\text{II})} N_j^{(\text{II})} + \lambda_{\text{III}}^2 N_i^{(\text{III})} N_j^{(\text{III})} \tag{51}$$

It has already been indicated how we could define a family of finite strain measures. The idea was to choose a function $g(\lambda)$, with $g(1)$ zero and $g'(1)$ unity so that the strain agrees with our small strain definition. To make a definition of a corresponding finite strain tensor, all we do is take $g(\lambda)$ and multiply it by the components of the vector in the principal direction, and sum over directions to get

$$E_{ij} = g(\lambda_{\text{I}}) N_i^{(\text{I})} N_j^{(\text{I})} + g(\lambda_{\text{II}}) N_i^{(\text{II})} N_j^{(\text{II})} + g(\lambda_{\text{III}}) N_i^{(\text{III})} N_j^{(\text{III})} \tag{52}$$

The principal axes of strain $[E]$ are the same as those of the pure deformation $[U]$, and are unaffected by the rotation $[R]$.

Here is a particular example, and an important one: The strain based on change of metric, which is sometimes called the Green strain, is based on

$$g(\lambda) = (\lambda^2 - 1)/2. \tag{53}$$

This strain was introduced in Eqn. 37. If we write out our definition of strain in (52) with this $g(\lambda)$ we see that

$$E_{ij}^M = \frac{1}{2}(\lambda_{\text{I}}^2 - 1) N_i^{(\text{I})} N_j^{(\text{I})} + \frac{1}{2}(\lambda_{\text{II}}^2 - 1) N_i^{(\text{II})} N_j^{(\text{II})} + \frac{1}{2}(\lambda_{\text{III}}^2 - 1) N_i^{(\text{III})} N_j^{(\text{III})} \tag{54}$$

We already know that λ^2 times these vector products adds up to $[F]^T[F]$, like in (51); and unity times those vector products just adds up to the Kronecker delta. Thus

$$E_{ij}^{M} = \frac{1}{2}\left(\sum_{k=1}^{3} F_{ki}F_{kj} - \delta_{ij}\right) = \frac{1}{2}\left(\frac{\partial u_i}{\partial X_j} + \frac{\partial u_j}{\partial X_i} + \sum_{k=1}^{3} \frac{\partial u_k}{\partial X_i}\frac{\partial u_k}{\partial X_j}\right) \tag{55}$$

So, this strain measure has the very nice feature that we can explicitly write it out in terms of the deformation gradient, indeed, in terms of derivatives of displacements, without actually solving for the principal axes. This is what makes it a very convenient finite strain measure and the reason why it is the one most often, although not universally, used.

Why is it called a strain measure based on the change of metric? Imagine that in the reference configuration we draw families of lines in the material along each coordinate direction, and then we let the material deform so that these lines become a curvilinear coordinate system. The metric tensor for that curvilinear coordinate system is just the tensor g_{ij} with the property that the squared length of a line element in that coordinate system is the metric times increments of the coordinates, which in this case are the material coordinates. If we simply work out what that squared length is, we find that

$$d\mathbf{x} \cdot d\mathbf{x} = \sum_{i=1}^{3}\sum_{j=1}^{3}\left(\delta_{ij} + 2E_{ij}^{M}\right) dX_i dX_j \tag{56}$$

so that

$$metric \quad g_{ij} = \delta_{ij} + 2E_{ij}^{M} \tag{57}$$

This means that the change in the metric is just $2E$. Again, this is the reason for sometimes calling the Green strain the strain based on the change of metric.

10. Work Conjugate Stress Tensors

We have a family of finite strain tensors that we can define, one of the most convenient being that based on change of metric, because it is easy to directly calculate in terms of the deformation gradient tensor F or in terms of displacement derivatives. All of these finite strain tensors agree with the infinitesimal strain tensor when we do indeed have small stretches and small rotations. Associated with each finite strain tensor E_{ij}, there is a work conjugate stress tensor S_{ij} that is a symmetric quantity; it is defined by writing out rate by which stress working is done per unit volume of material, where we measure that unit volume in the reference state, and then requiring that

Stress work rate per unit volume of reference state

$$\equiv \det[F] \sum_{i=1}^{3} \sum_{j=1}^{3} \sigma_{ij} \partial \dot{u}_j / \partial x_i = \sum_{k=1}^{3} \sum_{l=1}^{3} S_{kl} \dot{E}_{kl}$$

for arbitrary but related deformation rates $\partial \dot{u}_j / \partial x_i$ *and* $\dot{E}_{kl}$

The dots represent time derivatives. We demand that for every possible combination of derivatives of instantaneous velocity, and corresponding instantaneous strain rates, that this relation holds good. That suffices to define S, the work conjugate stress, in terms of the true (or Cauchy) stress. A different S is defined for each different strain measure E, or for each different $g(\lambda)$ which generates that measure. A relation which is of some interest, and which we will see later, is the one which results for the stress which is conjugate to the strain measure based on change of metric or Green strain. This stress is called the second Piola-Kirchhoff stress, and it is

$$S_{kl} = S_{kl}^{2nd\ P\text{-}K} = \det[F] \sum_{i=1}^{3} \sum_{j=1}^{3} F_{ki}^{-1} F_{lj}^{-1} \sigma_{ij} \qquad (58)$$

It is a linear combination of the actual true stresses but multiplied by certain components of the inverse of the deformation gradient matrix.

11. Principle of Virtual Work

Let us continue with a couple of important diversions before we get to stress-strain relations. The first diversion is to the principle of virtual work. That is a very old principle. It was established in the rapid development of mechanics shortly after Newton by Jean Bernoulli in 1717. He applied the principle of virtual work to objects like pin-connected rigid bodies. A short time later, in the middle-late 1700s, Euler was already applying the principle of virtual work, and related ideas of the calculus of variations, as a way of extracting the equations of statics for objects like beams. At the beginning of the 1800s, this was being done by several French mechanicians for problems of thin plates. This approach to continuum solid mechanics existed, in fact, before the concept of stress was really formalized. Stress was not really understood as a clear concept until the late 1700s. The work of Coulomb stands out, through there are precedents for the idea in the works of James Bernoulli around 1700. The real development of continuum mechanics as we now understand it, in terms of stress tensors and the like, was Cauchy's contribution of the 1820s, but the whole variational approach had actually been developed before. So, the principle of virtual work has an old history and continues to be an important work horse for this subject. The idea is to consider a solid in its deformed configuration at any arbitrary time t and then, taking the configuration at that time, to ask how much work is done on some imagined displacements δu through which we take the loads and stresses of that solid. The δu's are virtual infinitesimal displacements. Associated with them is a variation of strain:

$\delta\mathbf{u} = \delta\mathbf{u}(\mathbf{x}) =$ *virtual, or imagined, infinitesimal displacement field;*

$$\delta\varepsilon_{ij} = \frac{1}{2}\left[\frac{\partial(\delta u_i)}{\partial x_j} + \frac{\partial(\delta u_j)}{\partial x_i}\right] = \textit{associated virtual strain field} \tag{59}$$

Note that these derivatives are based on the current configuration; so $\delta\varepsilon$ is an infinitesimal strain from the current state, a virtual strain field. One can show from the equations of linear and angular momentum, specifically from equations (16), (22) and (25), that the work of the stress vector plus the work of the body force vector (with the d'Alambert procedure of including mass times acceleration of an effective body force) is

$$\int_S \mathbf{T}\cdot\delta\mathbf{u}\,dS + \int_V (\mathbf{f} - \rho\mathbf{a})\cdot\delta\mathbf{u}\,dV = \int_V \sum_{i=1}^{3}\sum_{j=1}^{3}\sigma_{ij}\,\delta\varepsilon_{ij}\,dV \tag{60}$$

This is the principle of virtual work. It is useful for many purposes. It is usually the starting point in proving things like uniqueness or in developing minimum principles and the like. It provides, in fact, an alternative statement of the governing equations. If we assume that the stress is symmetric $(\sigma_{ij} = \sigma_{ji})$, which assumption is typically built in to a stress-strain relation anyway, and if we further assume that the principle of virtual work holds for all possible virtual displacements and compatible virtual strains (i.e., related by (59)), then we are able to obtain the equations of linear momentum:

$$\sigma_{ij,i} + f_j = \rho a_j \ \ in \ \ V, \ \ n_i\sigma_{ij} = T_j \ \ on \ \ S. \tag{61}$$

This use of the principle of virtual work, as an alternative way of stating the equations of motion, is the most common — not the only, but the most common — starting point for the finite element method. In that method, one begins with the statement of the governing equations in the form of the principle of virtual work, and represents the displacement field approximately by displacements of nodes which are then interpolated to define the displacement field within the finite elements. Then one demands that the principle hold good for every virtual displacement field that we could generate by giving arbitrary virtual displacements to the nodes of the finite element system. When stresses are expressed in terms of strains, this provides a discrete system of equations for the nodal displacements which can be solved numerically.

The principle of virtual work has a close connection with a work energy relation, and this is shown below. It is called the mechanical work-energy relation to distinguish it from the first law of thermodynamics, which will be presented in the next section. The linear and angular momentum principles imply, via the principle of virtual work, that the work rate done by the stress vector on the surface, plus the work rate of body forces is the rate of change of the kinetic energy of the body plus the rate at which stresses do work in deformation:

Linear and angular momentum principles imply (via PVW) that

$$\int_S \mathbf{T} \cdot \mathbf{v}\, dS + \int_V \mathbf{f} \cdot \mathbf{v}\, dV = \frac{d}{dt}\int_V \frac{1}{2}\rho\, \mathbf{v} \cdot \mathbf{v}\, dV + \int_V \sum_{i=1}^{3}\sum_{j=1}^{3} \sigma_{ij} D_{ij}\, dV. \tag{62}$$

Here D is the rate of deformation tensor,

$$D_{ij} = (1/2)(\partial v_i / \partial x_j + \partial v_j / \partial x_i) \tag{63}$$

Equation (62) follows at once from (60) if we choose $\delta\mathbf{u} = \mathbf{v}\ \delta t$ and then cancel the common δt. So it holds good simply on the basis of the linear and angular momentum principles. There is no thermodynamics in it, although it looks somewhat similar to the first law of thermodynamics, which is seen in the next section.

12. First Law of Thermodynamics

A question to be asked is: "When do we need to deal with the first law of thermodynamics, as part of the system of equations needed to address problems in solid mechanics?" That is needed if we have any problem in which we consider temperature not as a passive variable (which we somehow prescribe), but if we consider it as a quantity that is generated by and coupled to the deformation field. For example, if we stretch a bar of metal only a little, so that it is still elastic, it actually gets a little bit colder. If we stretch it a lot, into the plastic range, it gets hot. Traditional, simplified, approaches to solid mechanics ignore the effects of any such deformation-induced temperature changes on the mechanical response. If instead we want to deal rigorously with such coupled phenomena, and in the process solve for the temperature field, then we need an additional equation. That is provided by the first law of thermodynamics. The first law is written out here for the continuum as:

$$\int_S (\mathbf{T}\cdot\mathbf{v} - q_n)\,dS + \int_V (\mathbf{f}\cdot\mathbf{v} + r)\,dV = \frac{d}{dt}\int_V \rho\left(e + \frac{1}{2}\mathbf{v}\cdot\mathbf{v}\right)dV$$

$$\begin{aligned} &e = \textit{internal energy per unit mass,} \\ &q_n = \textit{surface heat flux,} \\ &r = \textit{rate of radiant heat supply} \end{aligned} \qquad (64)$$

This balance law, like the balance law of linear momentum, can be used to extract, as consequences, certain equations which hold good at surfaces and certain partial differential equations that must be satisfied throughout a region, which are the local forms of the balance law. These consequences are

Heat flux vector $\mathbf{q}$ *exists such that:* $q_n = \mathbf{n}\cdot\mathbf{q}$ *on* S,

Local form of first law of thermodynamics:

$$-\sum_{i=1}^{3} \partial q_i / \partial x_i + r + \sum_{i=1}^{3}\sum_{j=1}^{3} \sigma_{ij} D_{ij} = \rho\ de/dt \quad \text{in } V. \qquad (65)$$

As regards the first statement, if we apply (64) to a Cauchy tetrahedron, and use (16) to simplify, we conclude that there has to exist a heat flux vector **q** so that the flow out is just the normal component of that heat flux vector. In the second statement of (65), the ordinary *d* is meant to denote a time rate of change following a material point, or following the motion. This is written as an ordinary partial derivative of *e* with respect to time if we write *e* in terms of material coordinates **X** and time. It has an additional transport term that comes in if we write *e* as a function of spatial coordinates **x** and time:

$$\begin{aligned} &de / dt = \textit{derivative following the motion} \\ &= \partial e(\mathbf{X},t) / \partial t = \partial e(\mathbf{x},t) / \partial t + \Sigma\, v_i\, \partial e(\mathbf{x},t) / \partial x_i \end{aligned} \qquad (66)$$

13. Second Law of Thermodynamics and Constitutive Equations

In the modern-day literature on continuum mechanics, the second law seems to be treated in one of two different ways. One approach is to introduce it as a new kind of balance law, then generally called the Clausius-Duhen inequality, where rather than having zero on the right hand side, we have greater than or equal to zero. The quantity being greater or equal to zero is what we think of as an entropy production rate. A second approach is simply to incorporate the second law into writing admissible local variations of state variables. The two approaches turn out to be equivalent in the constraints they deliver on how, for example, stresses can vary with strains, temperature, and possible other local state variables (introduced in certain viscoelastic or viscoplastic theories). The former approach is more general because, while it does not contradict the latter approach, it delivers more goods. Specifically, it delivers further transport constraints, such as for a heat-conducting body that $\mathbf{q} \cdot \nabla\theta \le 0$ (where θ is the thermodynamic temperature), which requires that the thermal conductivity matrix $[K]$, of $\{q\} = -[K]\{\nabla\theta\}$, be positive definite.

The latter approach is, nevertheless, adopted here because it is simpler. Let us think of purely elastic solids. In that case, what are the requirements of the first and second laws? Here we think of applying these laws to local processes within the solid, which are reversible. Understand, if we work in the context of thermoelastic behavior, we can have locally reversible relations between stress, strain, and temperature in a body which undergoes an irreversible process because heat conduction, for example, goes on. That is precisely what happens if we think of thermoelastic dissipation. If we have a vibrating rod, as it vibrates the upper surface is for a moment in extension, another moment in compression. This sets up an unequal temperature field in the body in each cycle of oscillation, which heat conduction tries to even out, and ultimately results in an energy loss mechanism. But locally, at each point of the material, if the solid is elastic, we have a reversible behavior. As we know from elementary thermodynamics, that means an increment of internal energy per unit volume of reference configuration, which is *de* multiplied by ρ_o (the mass density in the reference configuration), is equal to the thermodynamic temperature θ multiplied by an increment of entropy (here s is the entropy per unit mass) plus an increment of work of the relevant force quantities. We have seen that we can always introduce stress conjugates to whatever strain tensor we use, so as to give a work per unit volume. Thus

$$\rho_o \theta\, ds + \sum_{i=1}^{3} \sum_{j=1}^{3} S_{ij}\, dE_{ij} = \rho_o\, de \tag{67}$$

This is the thermodynamic requirement for elastic behavior. Analogs for simple fluid systems will be clear. A consequence is that there exists a potential, or strain energy-like term, so that

$$\begin{aligned} &S_{kl} = \rho_o\, \partial f([E],\theta) / \partial E_{kl} \;, \quad s = -\partial f([E],\theta) / d\theta \\ &\text{where} \quad f([E],\theta) \equiv e - \theta s = \text{Helmholtz free energy.} \end{aligned} \tag{68}$$

Here f is the Helmholtz free energy per unit mass, so that the stresses are given by its derivatives with respect to strains. Since the strain is a symmetric matrix, we understand when we write such equations that we rig things so that the dependence on E_{kl} and E_{lk} is identical.

The entropy is given by the derivative of f with respect to temperature. So this is the structure of stress-strain relations for elastic materials. To generate particular material models, we have to feed in particular definitions of the free energy. If, for example, we wanted to describe an isotopic material, say a block of rubber, then we would choose a free energy which depended only on the invariants of the strain tensor rather than on all components. On the other hand, if we had something like a fiber-reinforced composite, or a single crystal where there are preferred directions, then of course there is a more complex dependence on the strain components.

The Cauchy stress can be expressed in terms of the stress that is work-conjugate to the strain, and in particular, if we choose that stress as the second Piola-Kirchhoff stress, conjugate to Green strain, then the relation is obtainable from (58) as

$$\begin{aligned} &\textit{Cauchy stress } \ \sigma_{ij} = (1/\det[F]) \sum_{k=1}^{3}\sum_{l=1}^{3} F_{ik}F_{jl}S_{kl} \ \ \textit{when} \\ &S_{kl} = 2nd \ \ P\text{-}K \textit{ stress}, \ E_{kl} = \textit{strain based on change of metric (Green strain)} \end{aligned} \tag{69}$$

where the E_{kl} is given by (55) and the S_{kl} is given in terms of E, and θ, by the first line of (68).

If we look at materials which are not elastic but which nevertheless have instantaneous elasticity, we have a very similar structure which applies for constitutive equations. This includes a variety of non-elastic material models, for example, viscoelastic and viscoplastic models. In models commonly in use for polymers, the stress is a unique function of the instantaneous change of strain and temperature, given the prior history of both, so they are included in the class now considered. In fact, the general structure of a constitutive equation is derived by focusing on invariance of a general relation between stress and deformation to superposed rigid rotation. This leads rather directly to the form

$$\sigma_{ij}(t) = (1/\det[F(t)]) \sum_{k=1}^{3}\sum_{l=1}^{3} F_{ik}(t)\ F_{jl}(t) S_{kl}([E(t)],\theta(t);\ [E(t')],\ \theta(t'),\ t_o < t' < t) \tag{70}$$

where the quantity S_{kl} is just the second Piola-Kirchhoff stress. For solids with instantaneous elasticity, it is a direct function of the instantaneous strain and temperature, i.e., of the strain and temperature at time t, as indicated in (70), and is a functional of the prior history of these quantities from the formation time t_o of the material up to the present time; that is what the latter arguments of S are meant to denote, a functional dependence. Further, in view of the instantaneous elastic property, the dependence of S on the instantaneous E and θ must be compatible with the existence of a potential like in the first of equations (68), although f may have a very complex dependence on the prior history. That dependence is only weakly constrained by thermodynamics, which provides only the inequality $\dot{f} \le \dot{E}_{ij}\, \partial f / \partial E_{ij} + \dot{\theta}\, \partial f / \partial \theta$, where summation convention has been used and the derivatives with respect to E and θ represent derivatives relative to instantaneous variations of those quantities. Thus, at fixed E and θ one must have $\dot{f} \le 0$.

In writing specific constitutive models, that dependence on prior history is often replaced by a dependence on a set of state variables, where these state variables satisfy supplementary evolution relations, sometimes called kinetic relations. Those kinetic relations are then required to be consistent with the constraint just cited.

In this constitutive equation description, the instantaneous strain rate does not appear explicitly. This is what is meant by instantaneous elasticity. In terms of classical rheological models described by arrays of springs and dash pots, this sort of thing is o.k. if there is not a viscous element which constrains the entire array against instantaneous (elastic) deformation in response to an instantaneous change in stress. Classical viscous liquids are regarded as being constrained in that way. A solid which is thus constrained is called a Kelvin solid, and such Kelvin viscous response is incompatible with instantaneous elasticity. Note, however, that classical models for viscoplastic flow in elastic-plastic materials do include instantaneous elasticity, specifically because of that elastic part of the response. In such models, it is the expression for the plastic part of the strain rate, not the total strain rate, which is reminiscent of a viscous (but now non-linear, and evolving-state-dependent) liquid. The classical rate-independent elastic-plastic model does not allow instantaneous elastic response for all changes in stress (e.g., not for those directed into the plastic domain), but it does for those which point into the elastic interior of the yield domain, and that is enough to assure that the general structure of (70), with the generalized version of (68), applies.

14. Linear Elasticity and Thermoelasticity

We will now look at linear elasticity and thermoelasticity. The most general stress-strain relation is

$$\sigma_{ij} = \sum_{k=1}^{3}\sum_{l=1}^{3} C_{ijkl}\left(\varepsilon_{kl} - \alpha_{kl}\left(\theta - \theta_o\right)\right) \tag{71}$$

where θ_o is the temperature of the reference state, here unstressed, from which we measure strain. We will say very little about anisotropy, because David Barnett describes the topic in Chapter 4.

Let us focus on what is inside the outer parentheses of Eq. 71. If there was no stress acting and we changed the temperature from, say, the temperature of the reference state, then the strains would just be a set of expansion coefficients multiplied by the change in temperature, at least within a linearized theory. If we have stress acting, we have an additional source of strain, and indeed the difference between the total strain and the thermal strain, which is sometimes called the mechanical strain, enters the relation for stress. The C_{ijkl} are a set of isothermal elastic moduli. If we count the four indices on this set of elastic moduli, then at first the situation looks dreadful, because 3 times 3 times 3 times 3 is 81 different constants. But there are all sorts of symmetries that greatly reduce the number. First, because the strain is symmetric, the alphas have to be symmetric, so there can be at most six of them, and it is meaningless to talk about unsymmetric dependence of C on the last two indices:

$$\varepsilon_{kl} \;\; symmetric \Rightarrow \alpha_{kl} = \alpha_{lk} \;\; and \;\; C_{ijkl} = C_{ijlk} \tag{72}$$

Because the σ_{ij} are symmetric, the C_{ijkl} would have to be symmetric in the first two indices:

$$\sigma_{ij} \;\; symmetric \Rightarrow C_{ijkl} = C_{jikl} \tag{73}$$

This gets us down to not 81 but 36 C's. Then, finally, the equations of thermodynamics imply, as we have seen already, that stresses are derivable from a Helmholtz free energy, essentially a strain-energy like function. So, elastic moduli are just second derivatives of such a functions, giving the further symmetry:

$$Thermodynamics \Rightarrow C_{ijkl} = C_{klij} \;\; (C_{ijkl} = \rho_o \, \partial^2 f([\varepsilon],\theta_o) / \partial\varepsilon_{ij}\partial\varepsilon_{kl}) \tag{74}$$

because the order of differentiation is immaterial. That leaves us in fact with 21 independent elastic constants for the most general material.

Material symmetries reduce the number of constants further. There are two independent constants for the isotropic material, three for a cubic crystal (like aluminum, copper, or iron at low temperature), and five for a hexagonal crystal (for example, zinc). Indeed, five applies for any transversely isotropic material. So, a fiber-reinforced composite with fibers running in a single direction, and rather randomly arranged as they pierce a cross-section perpendicular to the fibers, is a transversely isotropic medium when examined macroscopically at deformation scale lengths that are large compared to fiber spacing or diameter.

There is, in the literature on crystal elasticity, a different notation that one will often encounter. This is also a common notation in many finite element formulations. Recognizing that we really only have six independent stresses, we make a column vector of stresses. We also have a column of six strains, where we use the three extensional strains but then twice the tensor shear strains, to make these the shears coincide with those usually denoted by γ and referring to an angle change. They are indicated below in rows, with the symbols T to transpose them from rows to columns.

$$\begin{aligned}
\{\sigma\} &= (\sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{12}, \sigma_{23}, \sigma_{31})^T \\
\{\varepsilon\} &= (\varepsilon_{11}, \varepsilon_{22}, \varepsilon_{33}, 2\varepsilon_{12}, 2\varepsilon_{23}, 2\varepsilon_{31})^T.
\end{aligned} \tag{75}$$

Then the stress-strain relation is written in a matrix form, where the elastic constants C are represented by a six by six matrix $[c]$ (36 elements). Thermodynamics requires that the matrix be symmetric, and that gets us down to 21 elements, since the diagonal and half of the off-diagonal of elements in the matrix add up to 21. Thus

$$\{\sigma\} = [c](\{\varepsilon\} - \{\alpha\}(\theta - \theta_o)); \; \textit{Thermodynamics} \Rightarrow [c] = [c]^T \tag{76}$$

The free energy is written as

$$\rho_o f([\varepsilon],\theta) = (1/2)\{\varepsilon\}^T [c]\{\varepsilon\} - \{\varepsilon\}^T [c]\{\alpha\}(\theta - \theta_o) + \textit{function of } \; \theta \tag{77}$$

If we neglect change in temperature altogether, it would just be a simple quadratic function of strain. The last term, that is not made very specific, the function of θ, describes the specific heat of the material.

Let us now take up the simplest case, of an isotropic elastic material. To recall some elastic parameters for that case, consider a straight bar aligned in the 1 direction and subject it to uniaxial tensile stress σ_{11}. The strain in that direction is $\varepsilon_{11} = \sigma_{11}/E$, where E is Young's modulus, and in the transverse directions is $\varepsilon_{22} = \varepsilon_{33} = -\nu\, \sigma_{11}/E$ where ν is the Poisson ratio.

Now let us address the question of how to write stress-strain relations for a general stress state like in Figure 2, and in particular, let us see why no additional elastic constants other than E and ν are needed to describe the response. There are different, but equivalent, ways of approaching this. One which draws on the power of the tensor concept is now described. We start by imagining that we have fortuitously chosen coordinate axes that line up with principal directions at a point of interest, so that there are extensional stresses there but no shear stresses (like in Figure 5). Because the medium is isotropic, there must be zero shear strains in that orientation. To write a typical non-zero strain, ε_{11} would be σ_{11}/E plus the Poisson effect due to the stresses in the transverse directions. If there is a change in temperature, we would have an expansion coefficient times the change in temperature, where, for the isotropic material, α is a simple scalar. Thus

$$\varepsilon_{11} = \sigma_{11}/E - \nu(\sigma_{22} + \sigma_{33})/E + \alpha(\theta - \theta_o)$$
$$E = \text{ Young's (tensile) modulus, } \nu = \text{ Poisson ratio.} \tag{78}$$

With a modest rearrangement, we could rewrite Equation 78 as

$$\varepsilon_{11} = (1+\nu)\sigma_{11}/E - \nu(\sigma_{11} + \sigma_{22} + \sigma_{33})/E + \alpha(\theta - \theta_o) \tag{79}$$

Finally, knowing that the Kronecker delta has a "11" component of unity, we could sneak that in in a couple of places without changing anything, to write

$$\varepsilon_{11} = (1+\nu)\sigma_{11}/E - \nu\delta_{11}(\sigma_{11} + \sigma_{22} + \sigma_{33})/E + \alpha\delta_{11}(\theta - \theta_o) \tag{80}$$

A similar equation applies for directions 2 and 3, so we get all the stress-strain relations right for this orientation if we write

$$\varepsilon_{ij} = (1+\nu)\sigma_{ij}/E - \nu\delta_{ij}(\sigma_{11} + \sigma_{22} + \sigma_{33})/E + \alpha\delta_{ij}(\theta - \theta_o), \tag{81}$$

which is correct for shears as well, because we do not have any shear stresses or strains on principal axes and the 12, 23, 31 components of the δ are zero.

This equation is certainly correct relative to principal axes. Now we make use of the powerful result that stress and strain, and also the Kronecker delta, are tensor quantities and that the sum of stresses is an invariant, so it is independent of whatever system of axes is used. If we would write this equation referred to general axes rather than to principal axes then, because we are dealing with tensor quantities, the equation would have exactly the same form. Hence, this stress-strain relation (81) is valid not just on principal axes, but is valid in general for all choices of coordinate systems.

This, of course, leads to a relation between elastic moduli, expressing the shear modulus G in terms of E and ν. That is, if we would evaluate let us say ε_{12}, a shear strain, we would get $(1+\nu)\sigma_{12}/E$. The shear strain based on the change of angle between initially orthogonal lines, γ_{12}, is twice the tensor shear and that γ type of shear strain is used to define the shear modulus by $\gamma_{12} = \sigma_{12}/G$. This gives the relation between the constants:

$$\text{Since } \gamma_{12} = \sigma_{12}/G = 2\varepsilon_{12} = 2(1+\nu)\sigma_{12}/E, \quad G = E/2(1+\nu) \tag{82}$$

Most readers will have encountered the derivation of this relation, done in an elementary way using Mohr's circles, for states of pure shear stress and strain, re-expressed as states of tension and compression at ±45°. But that is an equivalent derivation, because Mohr's circle just describes the tensor transformation in a plane.

In the literature, there are other ways of identifying elastic constants for the isotropic solid. One way is to ask: "Given that the material is isotropic, and that stress and strain are symmetric tensors, what is the most general linear relation that we could have between them?" Throw temperature in too. The most general linear relation is

$$\sigma_{ij} = \lambda\delta_{ij}(\varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33} - 3\alpha(\theta - \theta_o)) + 2\mu\varepsilon_{ij} \tag{83}$$

It is written in a way which identifies constants which are called the Lamé constants. μ is nothing more than a new label for the elastic shear modulus G. λ is a new modulus, and can be written in terms of Poisson's ratio and the other moduli:

$$\begin{aligned} \textit{Lamé constants:} \quad & \mu \equiv G = E/2(1+\nu), \\ & \lambda = 2\nu G/(1-2\nu) = \nu E/(1+\nu)(1-2\nu) \end{aligned} \tag{84}$$

Yet another modulus that we are sometimes concerned with is the bulk modulus. If all the normal stresses are negative and identical to pressure, then the fractional change in volume is the sum of the strains and will be minus the pressure divided by a coefficient that defines the bulk modulus:

$$\begin{aligned} &\textit{Bulk modulus } K: \quad \textit{For } \sigma_{11} = \sigma_{22} = \sigma_{33} = -p \ \ (\textit{pressure}), \\ &\qquad \textit{fractional decrease in volume} = (\varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33}) = -p / K; \\ &\qquad\qquad K = E/3(1-2\nu) = \lambda + 2\mu/3 \end{aligned} \tag{85}$$

We are accustomed to thinking of uniaxial stress because the standard laboratory tensile test is one which approximates to uniaxial stress. However, often we are also concerned, especially in dynamic phenomena (such as wave propagation), with uniaxial strain, where the medium is strained in a single direction, allowing no transverse strains to develop. The modulus for that kind of strain is:

$$\textit{Uniaxial strain modulus} = \lambda + 2\mu = (1-\nu)E/(1+\nu)(1-2\nu) \tag{86}$$

It is this modulus, in fact, which enters an expression for the speed of dilational waves in materials.

We will now distinguish between isothermal vs. isentropic moduli. In thinking of what we know about the deformation of gases, we are quite accustomed of thinking of deformation and temperature changes as being seriously coupled. If we compress a gas isothermically vs. isentropically (isentropic denoting a reversible adiabatic process), there is quite a difference in the stiffness of the gas. For solids, it is quite a different story.

The isothermal moduli are

$$\underline{Isothermal}\ moduli:\ \ C_{ijkl} = \rho_o \partial^2 f([\varepsilon],\theta) / \partial\varepsilon_{ij}\partial\varepsilon_{kl} \tag{87}$$

We define moduli for reversible adiabatic (or isentropic) processes in a similar fashion. The generating potential is not the Helmholtz free energy, but rather the internal energy written in what a thermodynamicist would call a fundamental form, which is to say, written as a function of strain and entropy:

$$\underline{Isentropic}\ moduli:\ \ \overline{C}_{ijkl} = \rho_o \partial^2 e([\varepsilon],s) / \partial\varepsilon_{ij}\partial\varepsilon_{kl} \tag{88}$$

In thermodynamics, we can derive all equilibrium properties of the medium if we write an appropriate thermodynamic potential in terms of the right variables. For example, if we write internal energy *e* in terms of strain and temperature, that will not generate a fundamental equation because there would be certain properties of the medium that we could not extract from it. But, if we write *e* as a function of strain and entropy, that is a fundamental form. In the same way, the Helmholtz function *f*, when written in terms of strain and temperature, is in fundamental form. Every single property, specific heat, stress-strain relations, thermal expansion, etc., is derivable from that kind of a form.

The next question is: how different are the isotropic and isothermal moduli? We can write out the isentropic moduli terms of the isothermic moduli as follows:

$$\begin{gathered}\overline{C}_{ijkl} = C_{ijkl} + (\theta_o / \rho_o c_\varepsilon)\beta_{ij}\beta_{kl} \quad \textit{where} \\ \beta_{pq} = \sum_{r=1}^{3}\sum_{s=1}^{3} C_{pqrs}\,\alpha_{rs} \ \ , \ \ c_\varepsilon = \textit{specific heat per unit mass at constant (zero) strain.}\end{gathered} \tag{89}$$

For the isotropic case, the shear modulus is identical isothermally and isentropically,

$$\overline{G} = G \tag{90}$$

but the bulk moduli are different,

$$\overline{K} = K\,(1 + 9\theta_o K\alpha^2 / \rho_o c_\varepsilon) \tag{91}$$

However, it is important to understand that numerically, in almost all cases the difference is a small one. Let us look at the term inside the parentheses:

$$9\theta_o K\alpha^2 / \rho_o c_\varepsilon \;\; \textit{is typically of the order of 1\% or less for metals and ceramics} \tag{92}$$

We can also work out a related quantity, the fractional change in thermodynamic or absolute temperature when we deform a solid at constant entropy. This is

$$\begin{aligned} &[(\theta - \theta_o)/\theta_o]_{s=const} = -(9\theta_o K\alpha^2 / \rho_o c_\varepsilon)[(\varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33})/3\alpha\theta_o]; \\ &3\alpha\theta_o \;\; \textit{is typically in the range } 10^{-3} \textit{ to } 4\times 10^{-2} \textit{ at room temperature} \end{aligned} \tag{93}$$

So, even if we have a large volume change of order 10^{-2}, we nevertheless typically generate 0.01 or less as a fractional change in temperature. In that sense the coupling effects in the elastic range are small. That ceases to be true, as we know from common experience, when we take materials well into the plastic range. Still, over a wide range of conditions, the coupling of temperature change to deformation is very small indeed. That justifies the (most typically followed) purely mechanical approach, in which temperature is not considered as a dependent variable of the theory (but rather as a given quantity), and no use is made of the first law of thermodynamics in formulating the governing equations.

15. Equations of Linear Elasticity; Waves

We use the linear elastic stress-strain relations of (71); temperature changes, as justified above, are neglected. We use, in these stress-strain relations, the strains written in terms of displacement gradients like in (41). The equations of motion (22), in which the spatial gradient can be replaced by a gradient in material coordinates for the present infinitesimal strain case, then gives the set of three equations of motion,

$$\sum_{i=1}^{3} \frac{\partial}{\partial X_i}\left(\sum_{k=1}^{3} \sum_{l=1}^{3} C_{ijkl} \frac{\partial u_k}{\partial X_l} \right) + f_j = \rho \frac{\partial^2 u_j}{\partial t^2} \quad (j = 1, 2, 3)\,, \tag{94}$$

for the three components of displacement. In the isotropic case, these equations reduce to the so-called Navier equations, the set of three equations

$$(\lambda + \mu) \frac{\partial}{\partial X_j}\left(\sum_{k=1}^{3} \frac{\partial u_k}{\partial X_k} \right) + \mu \sum_{k=1}^{3} \frac{\partial^2 u_j}{\partial X_k^2} + f_j = \rho \frac{\partial^2 u_j}{\partial t^2} \quad (j = 1, 2, 3), \tag{95}$$

restated in vector notation as

$$(\lambda + \mu)\nabla(\nabla \cdot \mathbf{u}) + \mu \nabla^2 \mathbf{u} + \mathbf{f} = \rho \; \partial^2 \mathbf{u} / \partial t^2 \tag{96}$$

I close this brief review by discussing simple solutions of the equations that describe what are called body waves. These describe plane disturbances of an arbitrary pulse shape that propagate steadily in the direction of some arbitrarily chosen unit vector **n**. What is unknown is the speed c of propagation and the polarization **p**, which is the direction of displacement in the wave (each component of displacement being assumed to follow the same shape of pulse). Thus we seek solutions in the form

$$\begin{gathered} \textit{Assume} \quad \mathbf{u}(\mathbf{X}, t) = \mathbf{p}\, f(\mathbf{n}\cdot\mathbf{X} - ct), \\ \textit{unit vector } \mathbf{n} = \textit{ propagation direction, } \mathbf{p} = \textit{ polarization, } c = \textit{wave speed.} \end{gathered} \tag{97}$$

Here **n** is given, the function *f*(…) is chosen arbitrarily, and we seek solutions for **p** and *c*. By inserting the assumed form of solution into (94), we obtain the following eigenvalue problem:

General anisotropic solid: *Solutions exist for arbitrary functions f(…) if*

$$\sum_{k=1}^{3}\left(\sum_{i=1}^{3}\sum_{l=1}^{3} n_i\, C_{ijkl}\, n_l\right) p_k = \rho c^2 p_j \quad (j = 1, 2, 3) \tag{98}$$

The result is that, for each direction **n** in the material, there exist three different wave speeds (two may coincide in degenerate cases, like the isotropic), given as eigenvalues of the above system. Because the quantity in parentheses, called the acoustic tensor when we divide it by ρ, is symmetric and positive definite, the three values of c^2 are real and positive, and further, the corresponding polarization directions are mutually orthogonal to one another (or can be so chosen in the degenerate cases).

In the isotropic case, when the governing equation reduces to (95) or (96), the result is:

Isotropic solid: *Solutions exist for arbitrary functions f(…) if either*

$$c = c_d \equiv \sqrt{(\lambda + 2\mu)/\rho} \quad \textit{and} \quad \mathbf{p} = \mathbf{n} \quad (\underline{\textit{longitudinal}} \textit{ or } \underline{\textit{dilatational}} \textit{ waves}) \tag{99}$$

$$\text{or} \quad c = c_s \equiv \sqrt{\mu/\rho} \quad \textit{and} \quad \mathbf{p}\cdot\mathbf{n} = 0 \quad (\underline{\textit{transverse}} \textit{ or } \underline{\textit{shear}} \textit{ waves}) \tag{100}$$

Acknowledgment: I am grateful to Professors Ronald W. Armstrong (U. of Md.), Helmut Kirchner (Vienna) and Marc A. Meyers (U. of Calif.) for organizing the NSF Institute of Mechanics and Materials (IMM) summer school, at which the material of this chapter was given as a pair of lectures, to Professor Meyers and Dr. Aldo Boccaccini (U. of Calif.) for transcribing the recorded lectures to a readable text, and to Mr. Douglas Galpin (Harvard) for help with graphics and preparation of the final version. Partial support of time has been provided by the ONR Solid Mechanics Program grant N00014-96-10777. The chapter is dedicated to the memory of a friend and respected colleague, Professor Richard Skalak, who served as the founding director of IMM.

Bibliography:

The presentation in *this chapter* follows but expands upon that in

- Rice, J. R., "Mechanics of Solids", section of the article "Mechanics", in *Encyclopaedia Britannica* (1993 printing of the 15th edition), volume 23, pp. 734-747 and 773, 1993.

For the *history* of the subject, a brief account is given in the reference above, and

- Love, A. E. H., *A Treatise on the Mathematical Theory of Elasticity*, 4th ed., 1927, reprinted by Dover Publications, New York, 1947

has a well researched chapter on the origin of elasticity up to the early 1900's. His article on "Elasticity" in the 11th Edition of Encyclopaedia Britannica is also a good source, as are articles on "Mechanics" and "Elasticity" in later editions. The 11th Edition has biographies of the major early figures. Also,

- Timoshenko, S. P., *History of Strength of Materials: With a Brief Account of the Theory of Elasticity and Theory of Structures*, 1953, reprinted by Dover Publications, New York, 1983

has good coverage of most sub-fields of solid mechanics up to the period around 1940, including in some cases detailed but quite readable accounts of specific developments and capsule biographies of major figures. The book

- Truesdell, C., *Essays in the History of Mechanics*, Springer-Verlag, New York, 1968

summarizes the author's studies of original source materials on James Bernoulli, Euler, Leonardo and others, and connects those contributions to some of the developments in what he calls rational mechanics as of the middle 1900's. The articles by

- Truesdell, C. and R. A. Toupin, in *Handbuch Der Physik*, Volume III/1, 1960

and by

- Truesdell, C. and W. Noll, *ibid.*, Volume III/3, 1965

also provide historical background.

For *beginners*, there are many good books, especially among those intended for the education of engineers. One which stands out for its coverage of inelastic solid mechanics as well as the more conventional topics on elementary elasticity and structures is

- Crandall, S. H., N. C. Dahl and T. J. Lardner, eds., *An Introduction to the Mechanics of Solids*, 2nd ed., McGraw-Hill, New York, 1978.

Those with an interest in the physics of materials might want to begin with

- Cottrell, A. H., *Mechanical Properties of Matter*, Wiley, New York, 1964, reprinted 1981.

Some books for beginners aim for a more general introduction to continuum mechanics, including solids and fluids, and one such is

- Fung, Y. C., *A First Course in Continuum Mechanics*, 2nd ed., Prentice-Hall, Englewood Cliffs, NJ, 1977.

A readable introduction to continuum mechanics at a more advanced level, as might be used by scientists and engineers from other fields or by first-year graduate students, is

- Malvern, L. E., *Introduction to the Mechanics of a Continuous Medium*, Prentice-Hall, Englewood Cliffs, NJ, 1969.

The article by Truesdell and Toupin, mentioned above, provides a comprehensive, perhaps overwhelming, treatment of continuum mechanics fundamentals.

For the *advanced student* seeking more specialized treatment of elasticity, the classics are the book by Love, mentioned above, and also

- Timoshenko, S. and J. N. Goodier, *Theory of Elasticity*, 2nd ed., McGraw-Hill, New York, 1951;
- Muskelishvili, N. I. (transl. J. M. Radok), *Some Basic Problems in the Mathematical Theory of Elasticity*, 2nd ed., Noordhoff 1963, reissued 1977;
- Green, A. E. and W. Zerna: *Theoretical Elasticity*, Clarendon/Oxford, 1954;
- Landau, L. D. and E. M. Lifschitz (trans. J. B. Sykes and W. S. Reid): *Theory of Elasticity*, Pergamon/Addison Wesley, 1959.

The article by Truesdell and Noll noted above is a good source on finite elasticity and also on viscoelastic fluids; a standard reference on the latter is

- Bird, R. B., R. C. Armstrong, and O. Hassager, *Dynamics of Polymeric Liquids, Volume 1, Fluid Mechanics*, 2nd ed., Wiley, New York, 1987.

Other books generally regarded as classics in their sub-fields are:

- Hill, R., *The Mathematical Theory of Plasticity*, Clarendon Press, Oxford, 1950;
- Jaeger, J. C. and N. G. W. Cook, *Fundamentals of Rock Mechanics*, 3rd ed., Wiley, New York, 1979;
- Hirth, J. P. and J. Lothe, *Theory of Dislocations*, 2nd ed., Wiley, New York, 1982;
- Aki, K. and P. G. Richards, *Methods of Quantitative Seismology*, W. H. Freeman, 1980.

Other aspects of stress waves in solids are covered by

- Achenbach, J. D., *Wave Propagation in Elastic Solids*, American Elsevier Pub. Co., New York, 1973.

Also, the scope of finite element analysis in solid mechanics and many other areas can be gleaned from

- Zienkiewicz, O. C., *The Finite Element Method*, 3rd ed., McGraw-Hill, New York, 1977,

and that of fracture mechanics from

- Kanninen, M. F. and C. H. Popelar, *Advanced Fracture Mechanics*, Oxford University Press, New York, 1985.

Structural mechanics and issues relating to stability, plasticity, and elastic-plastic stress-strain relations in a way that updates Hill, are presented by

- Bazant, Z. P. and L. Cedolin, *Stability of Structures: Elastic, Inelastic, Fracture and Damage Theories*, Oxford University Press, New York, 1991;
- Lubliner, J., *Plasticity Theory*, Macmillan, New York, 1990.

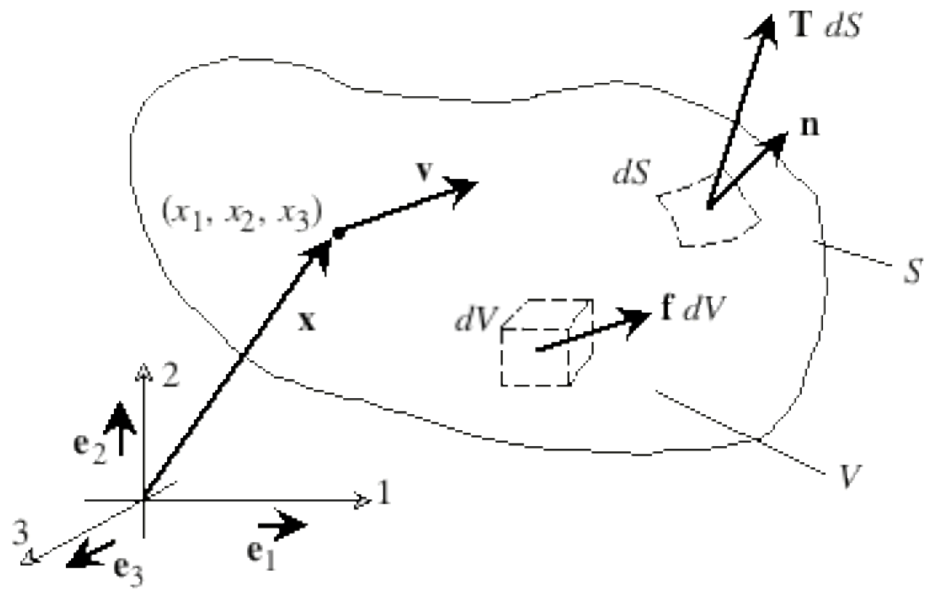

Figure 1: General body with volume and surface elements indicated, with Cartesian reference frame, axes labeled 1, 2 and 3.

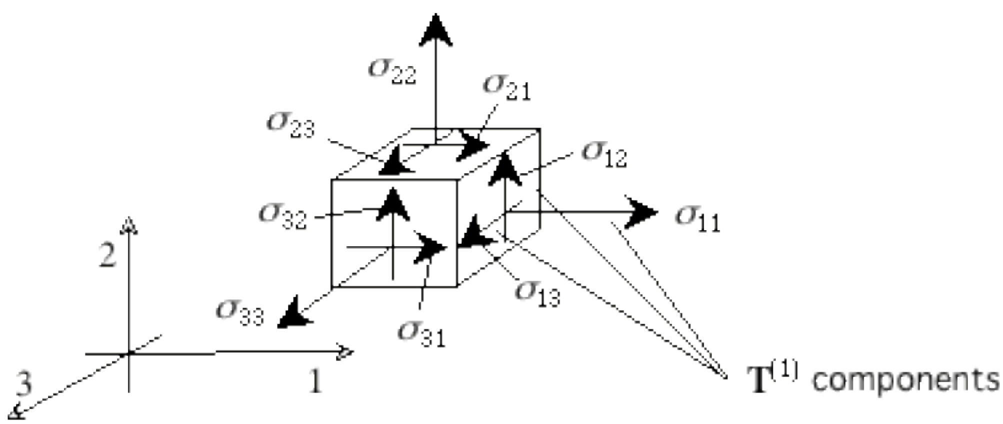

Figure 2: Three orientations (orthogonal to coordinate directions) of cut faces through a generic point of a solid. The nine components of stress σ_{ij} shown are introduced as the components of traction vectors, or stress vectors, $\mathbf{T}^{(i)}$ associated with each of the three face orientations.

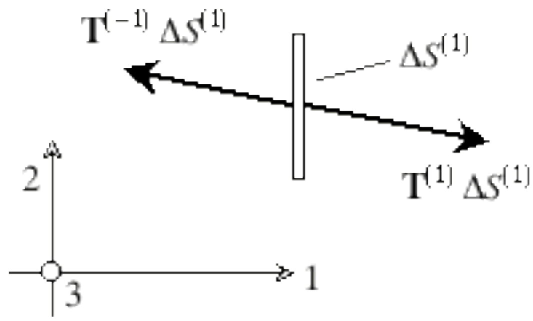

Figure 3: Tractions acting on the two sides of a surface.

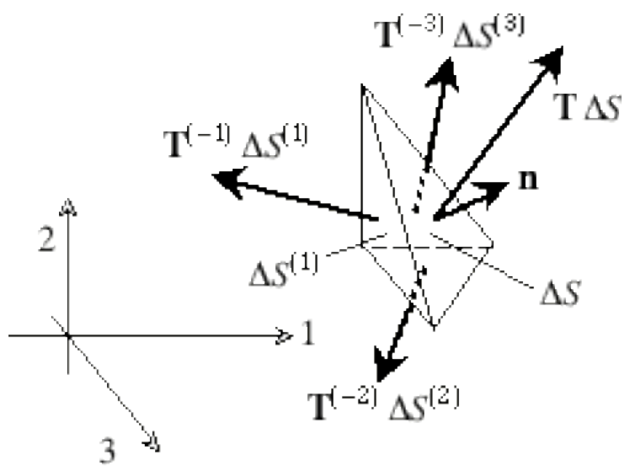

Figure 4: Cauchy tetrahedron with inclined face having some arbitrary orientation **n**; constructed about some material point, and to be shrunk onto that point in the limit to be taken.

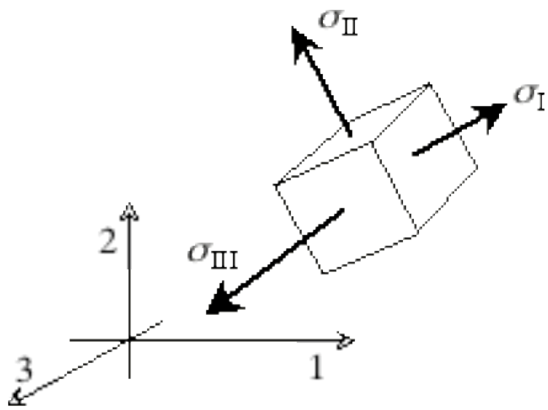

Figure 5: Principal stresses associated with three mutually orthogonal face orientations.

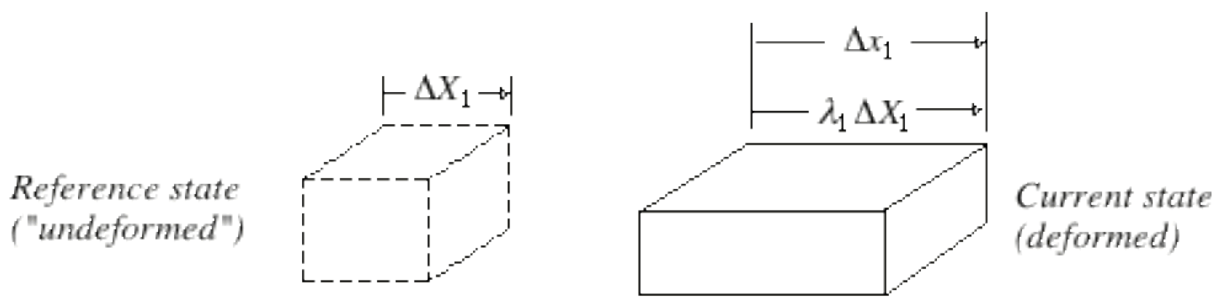

Figure 6: Illustration of extensional strains (λ = stretch ratio)

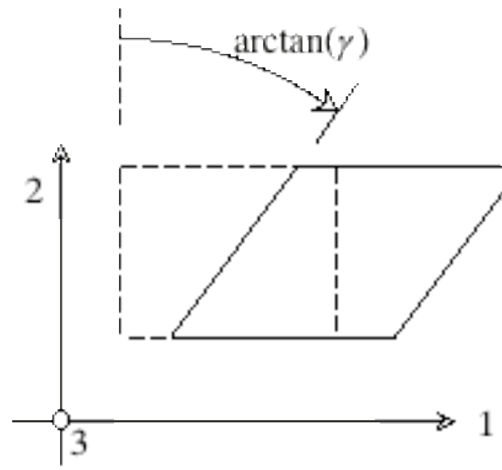

Figure 7: Illustration of "simple" shear strain

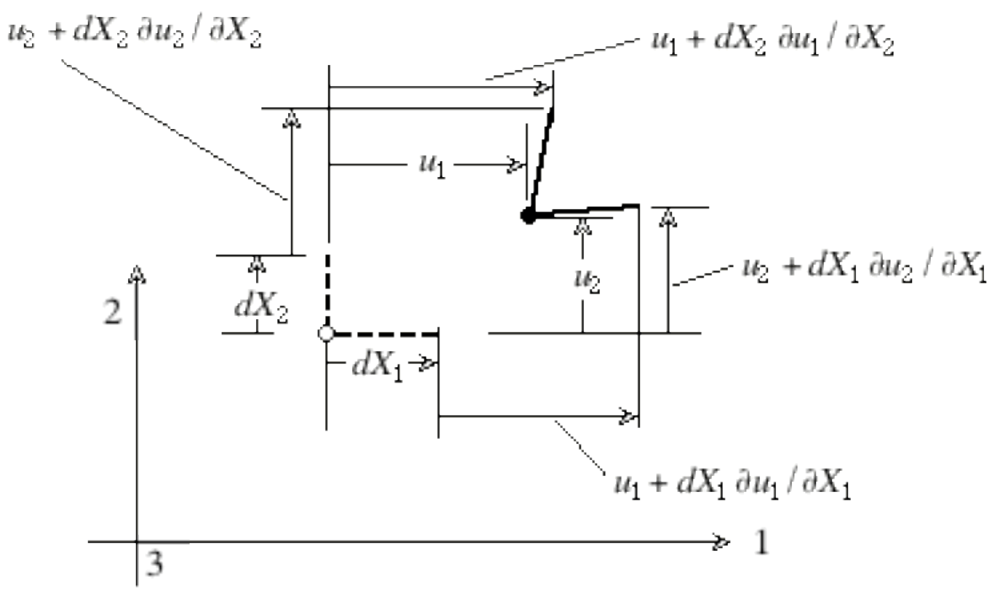

Figure 8: Deformation of line elements dX_1 and dX_2; notice that origin of line elements has displaced by u_1 and u_2, and that the extremities of the elements have displaced by slightly different amounts, due to the displacement gradients $\partial u_i / \partial X_j$ (presumed small for purposes of this diagram, which is used to introduce infinitesimal strain ε_{ij}).