

Lecture 13: Constitutive theory

David Al-Attar, Michaelmas term 2023

Outline and motivation

Within the previous lecture, we obtained the equations of motion for a simple hyperelastic body. The mechanical properties of such bodies are defined by a referential density, ρ , and a strain energy function, W . Using the principle of material frame indifference, we obtained conditions that must be satisfied by the strain energy function. Within the first part of this lecture, we will explore these restrictions in more detail to determine concretely the form the strain energy function must take. We then consider **linearised** motions about a stress-free equilibrium configuration. The resulting equations of motion are appropriate for studying most situations within seismology. To conclude we discuss **material symmetries** which occur in many real materials due to their underlying microscopic structure. By taking account of such symmetries the form of the equations of motion can be significantly simplified.

The polar decomposition theorem

In the previous lecture, we showed that the principle of material frame indifference requires the strain energy function of a simple hyperelastic body to satisfy

$$W(\mathbf{x}, \mathbf{Q}\mathbf{F}) = W(\mathbf{x}, \mathbf{F}), \quad (1)$$

for any deformation gradient \mathbf{F} and any rotation matrix \mathbf{Q} . Physically, this states that the potential energy of the body is invariant under superimposed rigid body rotations. To understand the implications of this condition, we need a result from linear algebra known as the **polar decomposition theorem**. This result states that an invertible matrix \mathbf{F} with $\det \mathbf{F} > 0$ can be written in the form

$$\mathbf{F} = \mathbf{R}\mathbf{U} = \mathbf{V}\mathbf{R} \quad (2)$$

where \mathbf{R} is a rotation matrix, while \mathbf{U} and \mathbf{V} are symmetric and positive-definite matrices with the same eigenvalues. The geometric interpretation of this theorem is shown in fig.1, where we see that the action of a deformation gradient can be viewed as either (i) a stretching about a set of orthogonal axes followed by a rotation, or (ii) the same rotation followed by stretching about co-rotated axes.

Proof of the polar decomposition theorem – NON-EXAMINABLE

You do not need to know the proof of this theorem, but it is simple enough to sketch for those interested. First we define a symmetric matrix

$$\mathbf{C} = \mathbf{F}^T \mathbf{F}, \quad (3)$$

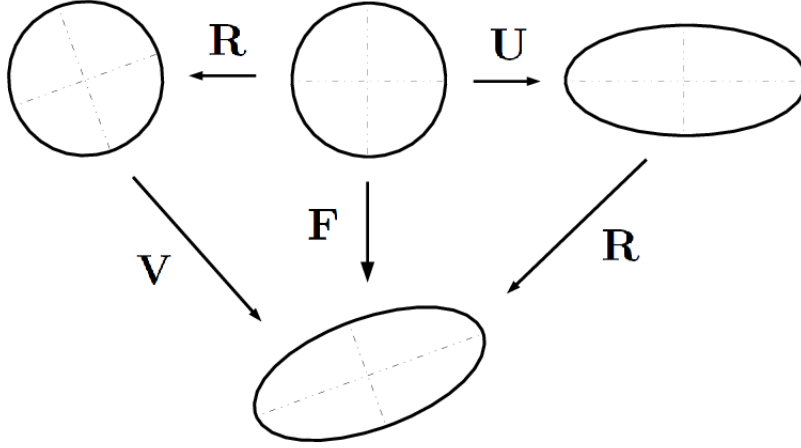


Fig. 1: Under an invertible matrix \mathbf{F} points a sphere is mapped to ellipsoid. Using this fact, the geometric meaning of the polar decomposition theorem is shown. An undeformed circle is shown in the centre of the upper row, with the dashed lines along the eigenvectors of the symmetric matrix \mathbf{U} .

which is positive-definite because \mathbf{F} is invertible. It follows that the square root of \mathbf{C} can be defined¹, and we set $\mathbf{U} = \sqrt{\mathbf{C}}$ which is symmetric and positive-definite by construction. Our claim is that $\mathbf{R} = \mathbf{F}\mathbf{U}^{-1}$ is a rotation matrix. This requires that $\mathbf{R}^T\mathbf{R} = \mathbf{1}$ and also $\det \mathbf{R} = 1$. For the first identity we have

$$\mathbf{R}^T\mathbf{R} = \mathbf{U}^{-1}\mathbf{F}^T\mathbf{F}\mathbf{U}^{-1}, \quad (4)$$

but $\mathbf{F}^T\mathbf{F} = \mathbf{C} = \mathbf{U}^2$, and hence $\mathbf{R}^T\mathbf{R} = \mathbf{1}$ as required. For the determinant condition we note that $\det \mathbf{R} = \frac{\det \mathbf{F}}{\det \mathbf{U}}$, while $\mathbf{U}^2 = \mathbf{F}^T\mathbf{F}$ implies $\det \mathbf{U} = \det \mathbf{F}$. For the second identity, we merely note that

$$\mathbf{F} = \mathbf{R}\mathbf{U} = \mathbf{R}\mathbf{U}\mathbf{R}^T\mathbf{R}, \quad (5)$$

and define $\mathbf{V} = \mathbf{R}\mathbf{U}\mathbf{R}^T$ which is just \mathbf{U} rotated and so has the same eigenvalues.

Implications for the strain energy function

Using the polar decomposition theorem, the deformation gradient can be written

$$\mathbf{F} = \mathbf{R}\mathbf{U}, \quad (6)$$

with \mathbf{R} a rotation matrix, while $\mathbf{U} = \sqrt{\mathbf{C}}$ with $\mathbf{C} = \mathbf{F}^T\mathbf{F}$. It follows from eq.(1) that

$$W(\mathbf{x}, \mathbf{F}) = W(\mathbf{x}, \mathbf{R}\mathbf{U}) = W(\mathbf{x}, \mathbf{U}), \quad (7)$$

and we conclude that the strain energy depends on the deformation gradient only through the symmetric matrix \mathbf{U} , or equivalently on \mathbf{C} . Within continuum mechanics, the symmetric matrix

$$\mathbf{C} = \mathbf{F}^T\mathbf{F}, \quad (8)$$

¹ This is done by first diagonalising $\mathbf{C} = \mathbf{Q} \text{diag}(\lambda_1, \lambda_2, \lambda_3) \mathbf{Q}^T$ with \mathbf{Q} orthogonal, and then setting $\mathbf{U} = \mathbf{Q} \text{diag}(\sqrt{\lambda_1}, \sqrt{\lambda_2}, \sqrt{\lambda_3}) \mathbf{Q}^T$ which is well-defined as the eigenvalues are all positive.

is known as the **right Cauchy-Green deformation tensor**². To summarise, we have shown that a strain energy function satisfies the principle of material frame indifference if and only if it takes the form

$$W(\mathbf{x}, \mathbf{F}) = U(\mathbf{x}, \mathbf{C}), \quad (9)$$

where U is some auxiliary strain energy function dependent on the right Cauchy-Green deformation tensor. Note that we use \mathbf{C} instead of \mathbf{U} in this formula because it simplifies later calculations.

Combining eq.(9) with the definition of the first Piola-Kirchhoff stress tensor

$$T_{ij} = \frac{\partial W}{\partial F_{ij}}, \quad (10)$$

we use the chain rule to obtain

$$T_{ij} = \frac{\partial U}{\partial C_{kl}} \frac{\partial C_{kl}}{\partial F_{ij}}, \quad (11)$$

where we have suppressed the various arguments for clarity. From the definition of the right Cauchy-Green deformation tensor, a routine calculation shows that

$$\frac{\partial C_{kl}}{\partial F_{ij}} = \frac{\partial}{\partial F_{ij}}(F_{mk}F_{ml}) = \delta_{im}\delta_{jk}F_{ml} + F_{mk}\delta_{im}\delta_{jl} = \delta_{jk}F_{il} + \delta_{jl}F_{ik}, \quad (12)$$

and using the symmetry of \mathbf{C} we arrive at the identity

$$T_{ij} = 2F_{ik}\frac{\partial U}{\partial C_{kj}}. \quad (13)$$

If we define the **second Piola-Kirchhoff stress tensor** to be

$$S_{ij} = 2\frac{\partial U}{\partial C_{ij}}, \quad (14)$$

the first Piola-Kirchhoff stress tensor can be decomposed in the form

$$T_{ij} = F_{ik}S_{kj}, \quad (15)$$

or in matrix form as $\mathbf{T} = \mathbf{F}\mathbf{S}$. It is clear that \mathbf{S} is a symmetric matrix, and so we see that the principle of material frame indifference implies that the following combination

$$\mathbf{F}^{-1}\mathbf{T}, \quad (16)$$

of the deformation gradient and the first Piola-Kirchhoff stress tensor is symmetric. In the first problem set you will show that this condition is equivalent to the conservation of angular momentum. We see, therefore, that angular momentum is automatically conserved during the motion of a simple hyperelastic body³. Within other forms of continuum mechanics, such for viscous fluids, angular momentum conservation needs to be explicitly imposed within the governing equations.

² $\mathbf{F}\mathbf{F}^T$ is the *left* Cauchy-Green deformation tensor which is useful in some circumstance.

³ For those that know about such things, this is a consequence of Noether's theorem.

The equilibrium equations

We have seen that the equations of motion for a simple hyperelastic body can be written

$$\rho \frac{\partial v_i}{\partial t} - \frac{\partial T_{ij}}{\partial x_j} = 0, \quad (17)$$

and are subject to the traction-free boundary conditions

$$T_{ij} \hat{n}_j = 0. \quad (18)$$

An **equilibrium configuration** φ^0 is a time-independent solution of these equations, and therefore satisfies

$$-\frac{\partial T_{ij}^0}{\partial x_j} = 0, \quad (19)$$

along with the boundary condition

$$T_{ij}^0 \hat{n}_j = 0. \quad (20)$$

Here we have defined the equilibrium first Piola-Kirchhoff stress tensor

$$T_{ij}^0 = \frac{\partial W}{\partial F_{ij}}[\mathbf{x}, \mathbf{F}^0(\mathbf{x})], \quad (21)$$

with \mathbf{F}^0 the deformation gradient of φ^0 .

Let φ^0 be an equilibrium configuration of a body. As the choice reference body M is arbitrary, we are free to require that for all $\mathbf{x} \in M$ we have

$$\varphi^0(\mathbf{x}) = \mathbf{x}, \quad (22)$$

It follows that $\mathbf{F}^0(\mathbf{x}) = \mathbf{1}$, and recalling eq.(15) we see that at equilibrium the first and second Piola-Kirchhoff stresses are equal, and hence *both* are symmetric. We assume for the first part of this course that $\mathbf{T}^0(\mathbf{x}) = \mathbf{0}$ for every $\mathbf{x} \in M$, which is a non-unique solution of the equations of motion in eq.(19) and the associated boundary conditions in eq.(20). Such an equilibrium configuration is said to be **stress-free**, and corresponds physically to the equilibrium state being one of minimum elastic potential energy. Prior to an earthquake, the Earth is certainly not stress-free, and we will return to this issue later. For many applications, however, it turns out that the neglect equilibrium stress is an acceptable approximation.

Linearised equations of motion

We can now rapidly obtain the equations of motion about a stress-free equilibrium configuration. Suppose that at $t = 0$ the body is displaced from equilibrium, with the initial motion taking the form

$$\varphi_i(\mathbf{x}, 0) = x_i + s U_i(\mathbf{x}) \quad (23)$$

for some given vector \mathbf{U} , while for some \mathbf{V} the initial velocity is given by

$$v_i(\mathbf{x}, 0) = s V_i(\mathbf{x}). \quad (24)$$

Here s is a dimensionless number that indicates the order of terms in a perturbation series. We shall assume that the subsequent motion can be expanded in powers of s , and hence seek a solution of the form

$$\varphi_i(\mathbf{x}, t) = x_i + s u_i(\mathbf{x}, t) + O(s^2), \quad (25)$$

where \mathbf{u} will be known as the **displacement vector** of the linearised motion. Working always to first-order in s , we find that the referential velocity of the motion is given by

$$v_i = s \frac{\partial u_i}{\partial t} + O(s^2), \quad (26)$$

while the deformation gradient can be written

$$F_{ij} = \delta_{ij} + s \frac{\partial u_i}{\partial x_j} + O(s^2). \quad (27)$$

Similarly, the first Piola-Kirchhoff stress tensor can be expanded as

$$T_{ij}(\mathbf{x}, t) = s \frac{\partial^2 W}{\partial F_{ij} \partial F_{kl}}(\mathbf{x}, \mathbf{1}) \frac{\partial u_k}{\partial x_l} + O(s^2), \quad (28)$$

where we have made use of our assumption that the equilibrium stress vanishes. It will be useful to define the **elastic tensor** by

$$A_{ijkl}(\mathbf{x}) = \frac{\partial^2 W}{\partial F_{ij} \partial F_{kl}}(\mathbf{x}, \mathbf{1}), \quad (29)$$

so that we can write the stress tensor more concisely as

$$T_{ij} = s A_{ijkl} \frac{\partial u_k}{\partial x_l} + O(s^2). \quad (30)$$

This linearised relationship is **Hooke's law**. But please note that this is not a “law” of nature nor empirical result. It arises solely as an *approximation* to finite elasticity which is valid for sufficiently small displacements about equilibrium. Due to the equality of mixed partial derivatives, the the elastic tensor A_{ijkl} possesses the so-called **hyperelastic symmetry**

$$A_{ijkl} = A_{klij}. \quad (31)$$

We shall see through this course that this innocuous looking symmetry plays a central role within the theory of elastic wave propagation. In the first problem set you will show that this tensor has additional symmetries

$$A_{ijkl} = A_{jikl} = A_{ijlk} \quad (32)$$

that reduce the number of independent components from 45 to 21. In particular, the symmetry of the elastic tensor in its final two indices means that

$$A_{ijkl} \frac{\partial u_k}{\partial x_l} = A_{ijkl} \frac{1}{2} \left(\frac{\partial u_k}{\partial x_l} + \frac{\partial u_l}{\partial x_k} \right), \quad (33)$$

where on the the right hand side we identify what is commonly called the **strain tensor**. More properly, this quantity the **linearised strain tensor**, which is obtained through

linearisation of the finite-strain $\mathbf{E} = \frac{1}{2}[\mathbf{C} - \mathbf{1}]$. It is possible to develop the general theory in terms of \mathbf{E} and not \mathbf{C} but there is little gained by doing this.

The exact equations of motion for the body are

$$\rho \frac{\partial v_i}{\partial t} - \frac{\partial T_{ij}}{\partial x_j} = 0, \quad (34)$$

and are subject to the traction-free boundary conditions

$$T_{ij} \hat{n}_j = 0. \quad (35)$$

Using the preceding expansions for \mathbf{v} and \mathbf{T} and equating first-order terms, we obtain the linearised equations of motion

$$\rho \frac{\partial^2 u_i}{\partial t^2} - \frac{\partial}{\partial x_j} \left(A_{ijkl} \frac{\partial u_k}{\partial x_l} \right) = 0, \quad (36)$$

along with the linearised boundary conditions

$$n_j A_{ijkl} \frac{\partial u_k}{\partial x_l} = 0. \quad (37)$$

Eq.(36) and (37) are the basic equations governing elastic wave propagation, and in later lectures we will look at their properties. Within these equations there are two physical parameters; the density ρ , and the elastic tensor \mathbf{A} . Both are functions that can vary smoothly within the body. There might also be discontinuities in these parameters across internal surfaces such as the core-mantle boundary, across which continuity of displacement and traction must be imposed.

Material symmetries

In general the elastic tensor has 21 independent components. In certain materials, however, this number is reduced due to the presence of material symmetries. Let us fix a point \mathbf{x} , and suppose there is a rotation matrix \mathbf{Q} such that

$$W(\mathbf{x}, \mathbf{FQ}) = W(\mathbf{x}, \mathbf{F}), \quad (38)$$

for any deformation gradient \mathbf{F} . Using the second part of the polar decomposition theorem we have $\mathbf{F} = \mathbf{VR}$, with \mathbf{V} symmetric and positive-definite, while \mathbf{R} is a rotation matrix. It follows that \mathbf{F} and \mathbf{FQ} represent the same magnitude of stretching, but oriented about different axes. Thus we can say that in eq.(38) there is a symmetry of the strain energy with respect to orientation of stretching. The collection of all such rotation matrices forms the **material symmetry group**⁴ at the point \mathbf{x} .

If the material symmetry group at \mathbf{x} comprises all rotation matrices, we say that the material is **isotropic** at this point. If the material is not isotropic, then it is **anisotropic**. The crystalline materials that make up the Earth's interior are almost always quite strongly anisotropic. However, seismic waves have wavelengths substantially larger than individual crystals, and so they "see" an average over many crystals with many different orientations. The end result is that for most seismological applications the Earth appears to be either isotropic or only slightly anisotropic.

⁴ If you care to, check that this really is a group.

Isotropic materials

Given their practical relevance, let us consider isotropic materials in further detail. To simplify notations, we work at a fixed point \mathbf{x} , but suppress this within the formulae. By assumption the strain energy of an isotropic material satisfies

$$W(\mathbf{F}\mathbf{Q}) = W(\mathbf{F}), \quad (39)$$

for all \mathbf{F} and *any* rotation matrix \mathbf{Q} . From material frame indifference, however, we know that we can always write

$$W(\mathbf{F}) = U(\mathbf{C}), \quad (40)$$

with $\mathbf{C} = \mathbf{F}^T \mathbf{F}$. And hence for an isotropic material we must have

$$U(\mathbf{C}) = U(\mathbf{Q}^T \mathbf{C} \mathbf{Q}). \quad (41)$$

Because \mathbf{C} is symmetric, the matrix $\mathbf{Q}^T \mathbf{C} \mathbf{Q}$ can be made diagonal for an appropriate choice of rotation matrix \mathbf{Q} . We conclude that in an isotropic material the strain energy depends *only on \mathbf{C} through its eigenvalues*. Using this result it is only a matter of tedious algebra (which we will skip) to obtain the general form of the elastic tensor in an isotropic material. The end result is that

$$A_{ijkl} = \lambda \delta_{ij} \delta_{kl} + \mu (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{kj}), \quad (42)$$

where λ and μ are collectively known as the **Lamé parameters**, and μ is called the **shear modulus**. If it is also assumed that the Lamé parameters are constant, it can readily be shown that the linearised equations of motion reduce to

$$\rho \frac{\partial^2 \mathbf{u}}{\partial t^2} - (\lambda + \mu) \nabla (\nabla \cdot \mathbf{u}) - \mu \nabla^2 \mathbf{u} = \mathbf{0}. \quad (43)$$

This is the form of the equations of elasticity covered in the 1B dynamics course. For reference, note that the isotropic elastic tensor can also be expressed using the shear modulus along with the **bulk modulus** κ which is defined in terms of λ and μ through

$$\kappa = \lambda + \frac{2}{3} \mu. \quad (44)$$

As a comment, you might ask why I have taken almost two lectures to get to an equation most of you will already know! There are two reasons. First, eq.(43) is not the equation of motion for a general elastic body, but for the linearised motion subject to a whole host of simplifying assumptions (e.g. isotropy, homogeneity, non initial stress, no gravitation). Such assumptions are generally invalid within seismological applications, and hence we do practically need to consider more general formulations. But beyond this, I think it is always better to have stated clearly and precisely what has and has not been assumed within a physical theory.

Transversely isotropic materials

We describe briefly an important class of anisotropic materials. Let $\hat{\mathbf{v}}$ be a unit vector, and suppose that eq.(39) holds for all rotation matrices \mathbf{Q} that leave $\hat{\mathbf{v}}$ fixed. Such a strain energy function is said to be **transversely isotropic** with symmetry axis $\hat{\mathbf{v}}$. The

most general strain energy function consistent with this condition has been determined, and it can be shown that the resulting elastic tensor takes the following form

$$A_{ijkl} = \lambda \delta_{ij} \delta_{kl} + \mu (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{kj}) + 8\gamma \hat{\nu}_i \hat{\nu}_j \hat{\nu}_k \hat{\nu}_l + 4\xi (\hat{\nu}_i \hat{\nu}_j \delta_{kl} + \delta_{ij} \hat{\nu}_k \hat{\nu}_l) - \zeta (\hat{\nu}_i \hat{\nu}_k \delta_{jl} + \hat{\nu}_j \hat{\nu}_k \delta_{il} + \hat{\nu}_j \hat{\nu}_l \delta_{ik} + \hat{\nu}_i \hat{\nu}_l \delta_{jk}), \quad (45)$$

where λ and μ are as before, while γ , ξ , and ζ are additional elastic moduli. Do not fear, you do not need to remember any such complicated tensor expressions!

Elastic fluids

The discussion of material symmetries above actually applies only to solids. There is, however, an extension of these ideas to so-called **elastic fluids**. This means that the fluid is assumed to be compressible, but to have no viscosity. Within seismological applications this is a good approximation when considering wave propagation within the oceans or in the Earth's outer core. All our basic equations already apply to elastic fluids, and the only difference is the form of the strain energy function. We stated in the previous lecture that an elastic body has a preferred shape and volume. In an elastic fluid, however, there is only a preferred volume. This means that deformations that leave the volume fixed have no effect on the strain energy. Mathematically this is expressed by requiring that

$$W(\mathbf{x}, \mathbf{F}\mathbf{Q}) = W(\mathbf{x}, \mathbf{F}), \quad (46)$$

for all \mathbf{F} and all matrices \mathbf{Q} with determinant equal to one⁵. It follows readily that the strain energy depends on \mathbf{F} only through its determinant, and hence we can write

$$W(\mathbf{x}, \mathbf{F}) = V(\mathbf{x}, J), \quad (47)$$

for some function V . In terms of the linearised theory, it is then a matter of algebra to show that the elastic tensor takes the form

$$A_{ijkl} = \kappa \delta_{ij} \delta_{kl}, \quad (48)$$

which is that of an isotropic solid but with vanishing shear modulus.

What you need to know and be able to do

- (i) The statement of the polar decomposition theorem, and how this can be used to restrict the allowable form of the strain energy function. The proof of the polar decomposition theorem is non-examinable.
- (ii) What a stress-free equilibrium configuration is, and how to derive the linearised equations of motion for small disturbances about such a configuration. Within this course, you will become very familiar with these linearised equations of motion, and it *will* be necessary to memorise their form (in the non-gravitating case discussed in this lecture) for the exam.

⁵ If you know some group theory, we can say that an elastic fluid has symmetry group equal to $SL(3)$, while for a solid the symmetry group is a subgroup of $SO(3)$. It can be shown that there is no proper subgroup intermediate between $SO(3)$ and $SL(3)$, and hence there is no “intermediate” state between solids and fluids. This approach to classifying elastic materials is due to Walter Noll (1925–2017).

- (iii) The definition of the material symmetry group, and its physical interpretation.
- (iv) The meaning of the terms **elastic solid**, **elastic fluid**, **isotropic**, **anisotropic**, and **transversely isotropic**. You do not need to remember the form of the elastic tensor in any particular case, but you should memorise the hyperelastic symmetry $A_{ijkl} = A_{klij}$ due to its great physical importance.