Helmholtz free energy of a rod under uniaxial tension.

Consider a rod, initial length L_0 and cross-sectional area A_0. When a loading machine applies a force f to the rod, the length of the rod becomes L. We record the function $f(L)$ experimentally, which need not be linear. When the length changes from L to $L+dL$, the loading machine does work fdL to the rod.

We model an elastic solid with a Helmholtz free energy, $F(L,T)$. When the length increases by dL, and the temperature increases by dT, the free energy increases by

$$dF = fdL - SdT .$$

Here S is the entropy, which can also be measured experimentally by measuring the heat absorbed by the rod in a slow loading process, $dS = dQ/T$.

Define the stress and strain as

$$\sigma = \frac{f}{A_0},\ \varepsilon = \frac{L-L_0}{L_0}.$$

Define the **free energy density**, w, as the free energy per unit volume, namely

$$w = \frac{F}{A_0 L_0}.$$

We can similarly define the entropy density

$$s = \frac{S}{A_0 L_0}.$$

Here we have used the initial area and initial length to define the stress, the strain, and the free energy density. With these definitions, we can rewrite $dF = fdL - SdT$ as

$$dw = \sigma d\varepsilon - sdT .$$

The free energy density is a function of the strain and the temperature:

$$w = w(\varepsilon, T).$$

For a given solid, the function $w(\varepsilon,T)$ is determined experimentally. Once we know this function, we can obtain the stress-strain relation by taking the differentiation:

$$\sigma = \left.\frac{\partial w(\varepsilon,T)}{\partial \varepsilon}\right|_T .$$

Similarly, we can also calculate the entropy density by

$$s = -\left.\frac{\partial w(\varepsilon,T)}{\partial T}\right|_\varepsilon .$$

We next restrict ourselves to small strains, so that we can expand the function $w(\varepsilon,T)$ into a Taylor series in the strain:

$$w(\varepsilon,T) = w_0 + w_1\varepsilon + \frac{1}{2}w_2\varepsilon^2 .$$

The coefficients w_0, w_1, w_2 are functions of the temperature. We will only go up to the quadratic term in strain.

The stress is obtained by taking partial differentiation:

$$\sigma = w_1 + w_2\varepsilon .$$

Thus, we identify w_2 as Young's modulus *E*. We also identify w_1 as the residual stress σ_o. Recalling the experimental observation of thermal expansion, we may wish to write this residual stress as

$$w_1 = -E\alpha\left(T - T_{ref}\right),$$

where α is the coefficient of thermal expansion, and T_{ref} is a reference temperature at which the strain is set to be zero when the stress vanishes. In most circumstances, E and α only depend on the temperature weakly, and are regarded as material constants. The energy density then becomes

$$w(\varepsilon,T) = w_0 + \sigma_o\varepsilon + \frac{1}{2}E\varepsilon^2 .$$

The entropy density is then given by

$$s = -\frac{\partial w(\varepsilon, T)}{\partial T} = \frac{dw_0}{dT} + \frac{d\sigma_o}{dT}\varepsilon + \frac{1}{2}\frac{dE}{dT}\varepsilon^2 .$$

Experimentally, the entropy density may be determined by measuring heat capacity *c*, namely, the heat absorbed by a unit volume of the solid per unit change in temperature, when the strain is held constant. Thus,

$$ds = \frac{cdT}{T} .$$

Using the experimentally measured value of *c*, we can determine w_0, w_1, w_2.

The above analysis can be generalized in many directions to account for diverse experimental observations. For the time being, we will focus on the stress strain relation of a linearly elastic solid, and make one important generalization: we will consider an anisotropic solid under a multiaxial stress state. We will drop the temperature dependence, and throw in thermal expansion strain at the end of the analysis. When we are not particularly interested in the temperature dependence of the free energy, following a large body of literature, we will call the Helmholtz free energy the elastic energy.

Example: Elastic energy density of a block under a simple shear.

Consider a block of height H_0 and cross-sectional area A_0. Subject to a shear force *F*, the block shear by an angle θ. When the shear angle changes from θ to $\theta + d\theta$, the loading machine does the work $FH_0 d\theta$ to the block. For an elastic solid the work is stored as the elastic energy in the block. Define the shear stress and the (engineering) shear strain as

$$\tau = \frac{F}{A_0}, \quad \gamma = \theta$$

The energy per unit volume is a function of the shear strain, $w(\gamma)$. From the above, we have

$$dw = \tau d\gamma .$$

The shear stress is the differential coefficient of the energy density function.

When the block is made of a linearly elastic solid, under shear load, the stress-strain relation is $\tau = G\gamma$. Consequently, the energy density function is

$$w(\gamma)=\frac{1}{2}G\gamma^2$$

This result holds only for linear elastic solid in pure shear condition.

Multi-axial stress states

The advantage of the energy density function becomes clear under the multi-axial stress state. The energy density is scalar, and the strain is a tensor. The energy density is a function of all components of the stain tensor:

$$w = w(\varepsilon_{11}, \varepsilon_{12}, \ldots).$$

(For people familiar with thermodynamics, we deal energy at a constant temperature, so that the energy is the Helmholz free energy.)

The components of stress tensor are differential coefficients:

$$dw = \sigma_{pq} d\varepsilon_{pq} .$$

That is

$$\sigma_{pq} = \frac{\partial w(\varepsilon_{11}, \varepsilon_{12}, \ldots)}{\partial \varepsilon_{pq}} .$$

In linear elasticity, we assume that stress is linear in strain. **Thus, the energy density is a quadratic form of the stain tensor**, written as

$$w = \frac{1}{2} C_{ijkl} \varepsilon_{ij} \varepsilon_{kl} .$$

Here the C_{ijkl} coefficients are the components of a fourth-rank tensor called the stiffness tensor. Without losing any generality, we can assume the following symmetries:

$$C_{ijkl} = C_{jikl} = C_{ijlk} = C_{klij} .$$

If we count carefully, we should have 21 independent components for a generally anisotropic elastic solid.

The components of the stress tensor are linear in the components of the strain tensor:

$$\sigma_{pq} = C_{pqij}\varepsilon_{ij}\,.$$

We can also invert this relation to express the strain in terms of the stress:

$$\varepsilon_{pq} = S_{pqij}\sigma_{ij}$$

Here S_{pqij} are the components of a fourth-rank tensor called the compliance tensor. They have the same symmetry properties.

Stress-strain relation in a matrix form

We can also write the above equations in another form. The state of strain is specified by the six components:

$$\varepsilon_x, \varepsilon_y, \varepsilon_z, \gamma_{yz}, \gamma_{zx}, \gamma_{xy}$$

In this order we will label them as $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_5, \varepsilon_6$. The six strain components can vary independently. The elastic energy per unit volume is a function of all the six strain components, $w(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_5, \varepsilon_6)$. This is the **energy density function**. When each strain component changes by a small amount, $d\varepsilon_i$, the energy density changes by

$$dw = \sigma_1 d\varepsilon_1 + \sigma_2 d\varepsilon_2 + \sigma_3 d\varepsilon_3 + \sigma_4 d\varepsilon_4 + \sigma_5 d\varepsilon_5 + \sigma_6 d\varepsilon_6\,.$$

Here we use the engineering strains for the shear, rather than the tensor components. We do so to avoid the factor 2 in the above expression. Each stress component is the differential coefficient of the energy density function:

$$\sigma_i = \frac{\partial w}{\partial \varepsilon_i}\,.$$

If the function $w(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4, \varepsilon_5, \varepsilon_6)$ is known, we can determine the six stress-strain relations by the differentiations. Consequently, by introducing the energy density function, we only need to specify one function, rather than six functions, to determine the stress-strain relations.

Linear elastic solids: The above considerations apply to solids with linear or nonlinear stress-strain relations. We now examine linear elastic solids. For the stress components to be linear in the strain components, the energy density function must be a *quadratic form* of the strain components:

$$w = \frac{1}{2}\sum_{i,j} c_{ij}\varepsilon_i\varepsilon_j = \frac{1}{2}\left(c_{11}\varepsilon_1\varepsilon_1 + c_{12}\varepsilon_1\varepsilon_2 + c_{21}\varepsilon_2\varepsilon_1 ...\right).$$

Here c_{ij} are 36 constants. The cross terms come in pairs, e.g.,$(c_{12}+c_{21})\varepsilon_1\varepsilon_2$. Only the combination $c_{12}+c_{21}$ will enter into the stress-strain relation, not c_{12} and c_{21} individually. We can call $c_{12}+c_{21}$ by a different name. A convenient way to say that there is only one independent constant is to just let $c_{12}=c_{21}$. We can do the same for other pairs, namely,

$$c_{ij} = c_{ji}.$$

The matrix c_{ij} is symmetric, with 21 independent elements. Consequently, 21 constants are needed to specify the elasticity of a linear anisotropic elastic solid. Because the elastic energy is positive for any nonzero strain state, the matrix c_{ij} is positive-definite.

Recall that each stress component is the differential coefficient of the energy density function, $\sigma_i = \partial w / \partial \varepsilon_i$. The stress relation becomes

$$\sigma_i = \sum_j c_{ij}\varepsilon_j .$$

In the matrix notation, we write

$$\begin{bmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \\ \sigma_4 \\ \sigma_5 \\ \sigma_6 \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} & c_{15} & c_{16} \\ c_{21} & c_{22} & c_{23} & c_{24} & c_{25} & c_{26} \\ c_{31} & c_{32} & c_{33} & c_{34} & c_{35} & c_{36} \\ c_{41} & c_{42} & c_{43} & c_{44} & c_{45} & c_{46} \\ c_{51} & c_{52} & c_{53} & c_{54} & c_{55} & c_{56} \\ c_{61} & c_{62} & c_{63} & c_{64} & c_{65} & c_{66} \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{bmatrix}$$

The physical significance of the constants c_{ij} is now evident. For example, when the solid is under a *uniaxial strain state,* $\varepsilon_1 \neq 0,\ \varepsilon_2 = \varepsilon_3 = \varepsilon_4 = \varepsilon_5 = \varepsilon_6 = 0$, the six stress components on the solid are $\sigma_1 = c_{11}\varepsilon_1, \sigma_2 = c_{21}\varepsilon_1, ...$ The matrix c_{ij} is known as the **stiffness matrix**.

Inverting the matrix, we express the strain components in terms of the stress components:

$$\begin{bmatrix}\varepsilon_1\\ \varepsilon_2\\ \varepsilon_3\\ \varepsilon_4\\ \varepsilon_5\\ \varepsilon_6\end{bmatrix}=\begin{bmatrix}s_{11} & s_{12} & s_{13} & s_{14} & s_{15} & s_{16}\\ s_{21} & s_{22} & s_{23} & s_{24} & s_{25} & s_{26}\\ s_{31} & s_{32} & s_{33} & s_{34} & s_{35} & s_{36}\\ s_{41} & s_{42} & s_{43} & s_{44} & s_{45} & s_{46}\\ s_{51} & s_{52} & s_{53} & s_{54} & s_{55} & s_{56}\\ s_{61} & s_{62} & s_{63} & s_{64} & s_{65} & s_{66}\end{bmatrix}\begin{bmatrix}\sigma_1\\ \sigma_2\\ \sigma_3\\ \sigma_4\\ \sigma_5\\ \sigma_6\end{bmatrix}$$

The matrix s_{ij} is known as the **compliance matrix**. The compliance matrix is also symmetric and positive definite.

The components of the stiffness tensor relate to the corresponding components of the stiffness matrix as

$$c_{11}=C_{1111},\ \ c_{14}=C_{1123},\ \ c_{44}=C_{2323}.$$

However, the corresponding relations for compliance are

$$s_{11}=S_{1111},\ \ s_{14}=2S_{1123},\ \ s_{44}=4S_{2323}.$$

Anisotropic elasticity. An isotropic, linear elastic solid is characterized by two constants (e.g., Young's modulus and Poisson's ratio) to fully specify the stress-strain relation. Some solids are anisotropic, e.g., fiber reinforced composites, single crystals. Each stress component is a function of all six strain components. Consequently, 21 constants are needed to specify the elasticity of a linear anisotropic elastic solid.

For a crystal of cubic symmetry, such as silicon and germanium, when the coordinates are along the cube edges, the stress-strain relations are

$$\begin{bmatrix}\sigma_1\\ \sigma_2\\ \sigma_3\\ \sigma_4\\ \sigma_5\\ \sigma_6\end{bmatrix}=\begin{bmatrix}c_{11} & c_{12} & c_{12} & 0 & 0 & 0\\ c_{12} & c_{11} & c_{12} & 0 & 0 & 0\\ c_{12} & c_{12} & c_{11} & 0 & 0 & 0\\ 0 & 0 & 0 & c_{44} & 0 & 0\\ 0 & 0 & 0 & 0 & c_{44} & 0\\ 0 & 0 & 0 & 0 & 0 & c_{44}\end{bmatrix}\begin{bmatrix}\varepsilon_1\\ \varepsilon_2\\ \varepsilon_3\\ \varepsilon_4\\ \varepsilon_5\\ \varepsilon_6\end{bmatrix}$$

The three constants c_{11}, c_{12} and c_{44} are independent for a cubic crystal. Isotropic solid is a special case, in which the three constants are related, $c_{44}=(c_{11}-c_{12})/2$.

For a fiber reinforced composite, with fibers in the x_3 direction, the material is isotropic in the x_2- and x_3-directions. The material is said to be transversely isotropic. Five independent elastic constants are needed. The stress-strain relations are

$$\begin{bmatrix}\sigma_1\\ \sigma_2\\ \sigma_3\\ \sigma_4\\ \sigma_5\\ \sigma_6\end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & 0 & 0 & 0\\ c_{12} & c_{11} & c_{13} & 0 & 0 & 0\\ c_{13} & c_{13} & c_{33} & 0 & 0 & 0\\ 0 & 0 & 0 & c_{44} & 0 & 0\\ 0 & 0 & 0 & 0 & c_{44} & 0\\ 0 & 0 & 0 & 0 & 0 & (c_{11}-c_{12})/2 \end{bmatrix} \begin{bmatrix}\varepsilon_1\\ \varepsilon_2\\ \varepsilon_3\\ \varepsilon_4\\ \varepsilon_5\\ \varepsilon_6\end{bmatrix}$$

Invariants and isotropic materials. If a material is isotropic, the elastic energy density w should be a function of the invariants of the strain tensor. For a linear material, w is quadratic in strain. The strain tensor can form only two invariants quadratic in its components: $\varepsilon_{ij}\varepsilon_{ij}$ and $(\varepsilon_{kk})^2$. Consequently, for an isotropic, linearly elastic solid, the elastic energy density takes the form

$$w = \mu\varepsilon_{ij}\varepsilon_{ij} + \frac{1}{2}\lambda(\varepsilon_{kk})^2,$$

where μ and λ are known as the Lame constants. Taking differentiation, we obtain the stress-strain relations:

$$\sigma_{ij} = 2\mu\varepsilon_{ij} + \lambda\varepsilon_{kk}\delta_{ij}.$$

A comparison with the stress-strain relations in the uniaxial stress state shows that

$$\mu = \frac{E}{2(1+\nu)}, \quad \lambda = \frac{\nu E}{(1+\nu)(1-2\nu)}.$$

Exercise. It is reasonable to require that the elastic energy density be a positive-definite quadratic form of the strain tensor. That is, $w > 0$ for any state of strain, except that $w = 0$ when the strain tensor vanishes. For an isotropic, linearly elastic solid, confirm that this positive-definite requirement is equivalent to require that $E > 0$ and $-1 < \nu < 1/2$.

Example: stress in an epitaxial film. Both silicon (Si) and germanium (Ge) are crystals of cubic unit cell. The edge length of the unit cell of Si is $a_{\mathrm{Si}} = 5.428\text{Å}$, and that of Ge is

$a_{\text{Ge}} = 5.658\text{Å}$. A Ge film 10 nm thick is grown epitaxially (i.e. with matching atomic positions) on the [100] surface of a 100 μm thick Si substrate. Calculate the stress and strain components in the Ge film. The respective elastic constants are (in GPa)

Si: $c_{11} = 165.8$, $c_{12} = 63.9$, $c_{44} = 79.6$.

Ge: $c_{11} = 128.5$, $c_{12} = 48.2$, $c_{44} = 66.7$.

Solution. Because the Si substrate is much thicker than the Ge film, the strains in the substrate are much smaller than those in the film. We will neglect these small strains, and assume that the substrate is undeformed. Let axis 3 be normal to the film surface, and axes 1 and 2 be in the plane of the film, parallel to the cube edges of the crystal cell. To register one atom on another, Ge must be compressed in directions 1 and 2 to conform to the undeformed atomic unit cell size of Si. The two in-plane strains in the Ge film are

$$\varepsilon_{11} = \varepsilon_{22} = \frac{a_{\text{Si}} - a_{\text{Ge}}}{a_{\text{Ge}}} = -4\% \ .$$

There will be an elongation normal to the film, $\varepsilon_{33} > 0$. All shear strains vanish. According to the generalized Hooke's law, the stress normal to the film surface relates to the strains as

$$\sigma_{33} = c_{11}\varepsilon_{33} + c_{12}\varepsilon_{11} + c_{12}\varepsilon_{22} .$$

Physically it is evident that there is no stress normal to the surface of the film, $\sigma_{33} = 0$. Inserting into the above expression, we obtain that

$$\varepsilon_{33} = -\frac{2c_{12}}{c_{11}}\varepsilon_{11} = +3\% .$$

The two in-plane stress components are equal, given by

$$\sigma_{11} = \sigma_{22} = c_{11}\varepsilon_{11} + c_{12}\varepsilon_{22} + c_{12}\varepsilon_{33} ,$$

or

$$\sigma_{11} = \sigma_{22} = \left[c_{11} + c_{12} - \frac{2c_{12}^2}{c_{11}} \right]\varepsilon_{11} .$$

Inserting the numerical values, we obtain that $\sigma_{11} = \sigma_{22} = -5.6\text{GPa}$. This is a huge stress that may generate dislocations in the film.

Ferroelastic phase transition. This part goes beyond linear elasticity. Suppose we have the following experimental observations. A crystal has a rectangular symmetry at a high temperature. When the temperature drops below a critical value, T_c, the crystal undergoes a phase transition. The crystal at a low temperature acquires a spontaneous strain in shear. Because of the symmetry, the shear strain can go both directions. (Sketch the geometry.)

We model this crystal with a free energy density

$$w(\gamma, T) = \frac{1}{2}A(T - T_c)\gamma^2 + \frac{1}{4}B\gamma^4,$$

where *A* and *B* are positive constants. Due to symmetry, the crystal is equally likely to shear in two directions, so that we keep the even powers in the strain γ. (Sketch *w* as a function of γ at two temperatures, one above T_c and the other below T_c.)

When $T > T_c$, the coefficient of the γ^2 term is positive, so that the crystal behaves like usual elastic solid, with the shear modulus $A(T - T_c)$. The γ^4 term is unnecessary to describe the behavior of the crystal.

When $T < T_c$, the coefficient of the γ^2 term is negative, and the energy is no longer minimal at $\gamma = 0$. Instead, the energy is minimal at two nonezero strains, known as the spontaneous strains, $\pm\gamma_s$. In this case, the γ^4 term will ensure that energy goes up again when the strain is large enough.

The stress-strain relation is

$$\tau = \frac{\partial w(\gamma, T)}{\partial \gamma} = A(T - T_c)\gamma + B\gamma^3.$$

(Sketch this function. Mark the spontaneous strains and the hysteresis loop.) Setting $\tau = 0$, we find the spontaneous strains:

$$\gamma_s = \pm\sqrt{A(T_c - T)/B}.$$

Because the material is nonlinear, the shear modulus is no longer a constant, and is given by

$$\mu = \frac{\partial \tau(\gamma, T)}{\partial \gamma} = A(T - T_c) + 3B\gamma^2 .$$

At the spontaneous strain, the shear modulus is given by

$$\mu = 2A(T_c - T).$$

Sketch the shear modulus as a function of the temperature, both below and above the critical temperature.