Mechanics of Composites

J R Willis
University of Cambridge

Contents

1 Basic notions

1.1 Introductory remarks

This course is concerned with the response to mechanical load of composite materials. Such materials are nowadays in frequent use. "Low-tech" applications (yacht hulls, car body parts etc.) most usually employ glass fibre-reinforced epoxy. This is light, sufficiently rigid, and has low cost. Structures that require higher performance include yacht masts, tennis raquets, aircraft panels, etc. Here, the material of preference is likely to be carbon fibre-reinforced plastic.

It is fairly easy to visualise what is meant by a composite material, by considering the examples just mentioned: they are particular examples of materials that are strongly heterogeneous on the microscopic scale and yet can be regarded as homogeneous, for the purpose of application. For instance, the flexure of a yacht mast under wind loading of the sails that it supports would be approached by assuming that the material was homogeneous. Furthermore, beam theory (of sufficient generality: the beam would have to be tapered, and allowance for torsion as well as bending would be needed) would most likely be employed. Beam theory would not suffice for the more detailed analysis of the stresses in the region where the mast is secured to the keel, or where it passes through the deck, but even here the mast material would be treated as homogeneous. This is not to imply that the stress and strain distributions agree exactly with those delivered by the calculation assuming homogeneous material: in fact, they will display large fluctuations, on the scale of the microstructure of the material, about "local average" values which will agree, quite closely, with those obtained from the "homogeneous" calculation. This statement in fact has a precise meaning in terms of asymptotic analysis: the mathematical theory of "homogenization" considers systems of partial differential equations whose coefficients oscillate, on a scale ε, while the domain Ω over which the equation is defined has a size of order 1. The solution, $u^\varepsilon(\boldsymbol{x})$, say, depends on ε. A "homogenization theorem" (which certainly applies to problems of stress analysis of the type under present consideration) states that, as $\varepsilon \to 0$, $u^\varepsilon(\boldsymbol{x})$ tends strongly to a limit, $u^0(\boldsymbol{x})$ say, while its gradient tends weakly to the gradient of $u^0(\boldsymbol{x})$[1]. The field $u^0(\boldsymbol{x})$ satisfies a "homogenized" system of partial differential equations, corresponding in our context to the equations of equilibrium for the homogeneous material. In the "engineering" context, it is usual to discuss "effective properties" rather than coefficients of homogenized equations, but the concept is the same. The basic problem treated in these lectures is the determination of the "effective" properties, in terms of the properties of the constituent

[1] In the context of linear equations, the topology relative to which convergence is defined would be that of $L^2(\Omega)$. Details would be out of place here.

materials and the microgeometry.

A great variety of materials display the same character as composites, in the sense that they may be strongly heterogeneous relative to a microscale while appearing homogeneous for the purpose of many applications. With the exception of exotic applications such as some aero engine turbine blades, made from single-crystal nickel alloy, metal objects and structures are composed of metal in polycrystalline form: the individual crystal grains are homogeneous, anisotropic single crystals but the polycrystal is made of many grains, at different orientations; thus, the polycrystal may be isotropic, though it could be anisotropic, if the polycrystal displays "texture", for instance as a result of rolling. Concrete provides another example of a heterogeneous material. Depending on the mix, it may contain stones whose diameter may be up to the order of centimetres. Such a material would appear homogeneous on a sufficiently large scale, for a structure such as a dam, for which a "representative volume element" might have dimensions on the order of tens of centimetres. Testing laboratory samples of such material presents a significant challenge, because any specimen whose size is on the order of 5cm, say, would contain only a few heterogeneities, and different specimens would have different responses. At the other extreme, carbon fibres typically have diameter on the order of 10^{-2}mm, and correspondingly a volume with dimension as small as a few millimetres could be regarded as "representative". The grain size in polycrystalline metal may be of the order of 10^{-3}mm. Thus, a "structure" as small as a paperclip can be regarded as homogeneous.

Instead of pursuing this rather vague, qualitative discussion further, we now proceed to some more precise concepts.

1.2 Definition of effective properties

Consider first material that is *homogeneous*. A way to determine the constitutive response of homogeneous material is to perform a mechanical test, in which a specimen is subjected to loading that produces a homogeneous field of deformation, and a corresponding homogeneous field of stress. Since these fields are both homogeneous, they can be determined from measurements of their values on the surface of the specimen. The functional relation between the two is the desired constitutive relation. Explicitly, in the case of linear elastic response, the Cauchy stress tensor $\boldsymbol{\sigma}$, with components σ_{ij}, is related to the strain tensor $\boldsymbol{\varepsilon}$, with components ε_{ij}, so that

$$\boldsymbol{\sigma} = \boldsymbol{C}\boldsymbol{\varepsilon}, \quad \text{or in suffix notation, } \sigma_{ij} = C_{ijkl}\varepsilon_{kl}. \tag{1.1}$$

A set of experiments in which the six independent components of strain are prescribed, while the stress components are measured, fixes the elastic constant tensor $\boldsymbol{C}$ (whose components

are C_{ijkl}). Alternatively, experiments could be performed in which the six independent components of stress were prescribed, while the strain components were measured. This would yield, directly, a relation

$$\boldsymbol{\varepsilon} = \boldsymbol{S}\boldsymbol{\sigma}, \text{ or in suffix notation, } \varepsilon_{ij} = S_{ijkl}\sigma_{kl}. \tag{1.2}$$

Here, $\boldsymbol{S}$ denotes the tensor of compliances, inverse to $\boldsymbol{C}$ in the sense that

$$\boldsymbol{S}\boldsymbol{C} = \boldsymbol{C}\boldsymbol{S} = \boldsymbol{I}, \text{ or in suffix notation, } S_{ijkl}C_{klmn} = C_{ijkl}S_{klmn} = \tfrac{1}{2}(\delta_{im}\delta_{jn} + \delta_{in}\delta_{jm}). \tag{1.3}$$

Conceptually, one way to realise a uniform strain field in a *homogeneous* body is to impose on the boundary of the specimen displacements that are consistent with uniform strains: if the domain occupied by the body is Ω and its boundary is $\partial\Omega$, impose the boundary condition

$$\boldsymbol{u} = \overline{\boldsymbol{\varepsilon}}\boldsymbol{x}, \text{ or in suffix notation, } u_i = \overline{\varepsilon}_{ij}x_j, \quad \boldsymbol{x} \in \partial\Omega \tag{1.4}$$

(there is no gain in allowing also a rotation). The uniform strain generated in this way is $\overline{\boldsymbol{\varepsilon}}$.

Conversely, a uniform stress field is generated in a homogeneous body by imposing the boundary condition

$$\boldsymbol{t} = \boldsymbol{\sigma}\boldsymbol{n} = \overline{\boldsymbol{\sigma}}\boldsymbol{n}, \text{ or in suffix notation, } t_i = \sigma_{ij}n_j = \overline{\sigma}_{ij}n_j, \quad \boldsymbol{x} \in \partial\Omega. \tag{1.5}$$

The uniform stress that is generated is $\overline{\boldsymbol{\sigma}}$.

In reality, there is no such thing as a homogeneous material: even a perfect crystal is composed of atoms, and so is not even a continuum! The continuum approximation is nevertheless a good one for virtually all engineering applications: the notional experiments just described really could be carried out, and the sort of apparatus that is invariably used would lack the sensitivity to register any deviation from the assumption that the specimen is a continuum. The reason for making this rather trivial remark is to make more acceptable the next comment. This is that "effective" properties can be assigned even to a specimen that is made of material that is *inhomogeneous*.

Suppose now, that the body is inhomogeneous, but that the boundary condition (1.4) is applied. It generates a displacement field $\boldsymbol{u}(\boldsymbol{x})$ and a corresponding strain field $\boldsymbol{\varepsilon}(\boldsymbol{x})$ that is not uniform. Its mean value over Ω is nevertheless equal to $\overline{\boldsymbol{\varepsilon}}$. Conversely, if the body is inhomogeneous and the traction boundary condition (1.5) is applied, the stress field $\boldsymbol{\sigma}(\boldsymbol{x})$ is not uniform but nevertheless has mean value $\overline{\boldsymbol{\sigma}}$. These results are simple consequences of the following

Mean Value Theorems:

(a) The mean value $\overline{\boldsymbol{\varepsilon}}$ over Ω of the strain field $\boldsymbol{\varepsilon}(\boldsymbol{x})$ is expressible in terms of the boundary displacements as

$$\overline{\varepsilon}_{ij} := \frac{1}{|\Omega|}\int_\Omega \varepsilon_{ij}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = \frac{1}{|\Omega|}\int_{\partial\Omega} \tfrac{1}{2}(u_i n_j + u_j n_i)\mathrm{d}S. \tag{1.6}$$

(b) The mean value $\overline{\boldsymbol{\sigma}}$ over Ω of the equilibrium stress field $\boldsymbol{\sigma}(\boldsymbol{x})$ is expressible in terms of the boundary tractions, in the absence of body force, as

$$\overline{\sigma}_{ij} := \frac{1}{|\Omega|}\int_\Omega \sigma_{ij}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} = \frac{1}{|\Omega|}\int_{\partial\Omega} \tfrac{1}{2}(t_i x_j + t_j x_i)\mathrm{d}S. \tag{1.7}$$

Proof:
(a) follows directly from the divergence theorem.
(b) Substitute $t_i = \sigma_{ik}n_k$ into the surface integral. Then by the divergence theorem,

$$\int_{\partial\Omega} t_i x_j \mathrm{d}S = \int_\Omega (\sigma_{ik}x_j)_{,k}\mathrm{d}\boldsymbol{x} = \int_\Omega (\sigma_{ik,k}x_j + \sigma_{ik}\delta_{jk})\mathrm{d}\boldsymbol{x} = \int_\Omega \sigma_{ij}\mathrm{d}\boldsymbol{x},$$

since for equilibrium, $\sigma_{ik,k} = 0$. The result now follows immediately. [Strictly, the symmetric form given in (1.7) is not necessary, but it is desirable at least for aesthetics.]

It should be noted that, if the boundary condition (1.4) is applied, then part (b) of the theorem shows that the mean stress in the body can be obtained from measurements of the surface traction $\boldsymbol{t}$, even in the case that the body is heterogeneous. Conversely, if the boundary condition (1.5) is applied, then part (a) of the theorem shows that the mean strain in the body can be obtained from measurement of the surface displacements.

We are now in a position to define the effective response of a body, or a specimen.
(a) *Linear boundary displacements*. Apply the boundary condition (1.4) and measure the associated mean stress. The effective tensor of elastic moduli $\boldsymbol{C}^{\mathrm{eff,u}}$ is defined so that

$$\overline{\boldsymbol{\sigma}} = \boldsymbol{C}^{\mathrm{eff,u}}\overline{\boldsymbol{\varepsilon}}. \tag{1.8}$$

(b) *Uniform boundary tractions*. Apply the boundary condition (1.5) and measure the associated mean strain. The effective tensor of compliances $\boldsymbol{S}^{\mathrm{eff,t}}$ is defined so that

$$\overline{\boldsymbol{\varepsilon}} = \boldsymbol{S}^{\mathrm{eff,t}}\overline{\boldsymbol{\sigma}}. \tag{1.9}$$

It should be noted that, in general, $\boldsymbol{C}^{\mathrm{eff,u}}$ and $\boldsymbol{S}^{\mathrm{eff,t}}$ are not inverse to one another: the prescription (a) generates an effective compliance $\boldsymbol{S}^{\mathrm{eff,u}} = (\boldsymbol{C}^{\mathrm{eff,u}})^{-1}$, and prescription (b) generates a tensor of effective moduli $\boldsymbol{C}^{\mathrm{eff,t}} = (\boldsymbol{S}^{\mathrm{eff,t}})^{-1}$. If, however, the body or specimen

is made of composite material that appears, "on average"[2], as uniform, and the specimen is large enough relative to the microstructure, then it is to be expected that the effective properties resulting from either boundary condition will coincide. In this case, the superscripts 'u' or 't' become irrelevant, and will later be omitted.

A further interesting property follows rigorously, if either linear boundary displacements or uniform boundary tractions are applied. Whereas, in general, the average of a product is different from the product of the averages, either of these conditions gives the result

$$\overline{\sigma_{ij}\varepsilon_{ij}} := \frac{1}{|\Omega|}\int_\Omega \sigma_{ij}\varepsilon_{ij}\mathrm{d}\boldsymbol{x} = \overline{\sigma}_{ij}\overline{\varepsilon}_{ij}. \tag{1.10}$$

This result is commonly known as the Hill relation, because it was first discussed by Rodney Hill, around 1951. The proof for the linear boundary displacement condition is as follows.

$$\begin{aligned}
\int_\Omega \sigma_{ij}\varepsilon_{ij}\mathrm{d}\boldsymbol{x} &= \int_\Omega \tfrac{1}{2}\sigma_{ij}(u_{i,j}+u_{j,i})\mathrm{d}\boldsymbol{x} \\
&= \int_\Omega \sigma_{ij}u_{i,j}\mathrm{d}\boldsymbol{x} \ \text{(by the symmetry of the stress tensor)} \\
&= \int_\Omega (\sigma_{ij}u_i)_{,j}\mathrm{d}\boldsymbol{x} \ \text{(from the equilibrium equations)} \\
&= \int_{\partial\Omega} \sigma_{ij}n_j u_i \mathrm{d}S = \int_{\partial\Omega} \sigma_{ij}\overline{\varepsilon}_{ik}x_k n_j \mathrm{d}S,
\end{aligned} \tag{1.11}$$

the last equality following from the boundary condition. The result follows by applying the divergence theorem to express the last surface integral as a volume integral, remembering now that $\overline{\boldsymbol{\varepsilon}}$ is constant. The proof for the uniform traction condition is similar.

1.3 Representative volume element

It was noted in the introductory subsection that there is a rigorous asymptotic result, that for a body of fixed size, and subject to fixed boundary conditions (and body forces), the displacement field $\boldsymbol{u}^\varepsilon$ approaches, asymptotically, a field $\boldsymbol{u}^0$ as $\varepsilon \to 0$, and that $\boldsymbol{u}^0$ satisfies a set of equations which include a "homogenized" constitutive relation, which we now write as $\boldsymbol{\sigma} = \boldsymbol{C}^{\mathrm{eff}}\boldsymbol{\varepsilon}$. If the microscopic length is small enough – or, equivalently, the dimensions of the body are large enough relative to the microstructure – then the *same* "effective modulus" tensor will apply to all boundary value problems and so, in particular, the tensors which we defined as $\boldsymbol{C}^{\mathrm{eff,u}}$ and $\boldsymbol{C}^{\mathrm{eff,t}}$ will coincide. A "representative volume element" must be at least large enough for $\boldsymbol{C}^{\mathrm{eff}}$ to be independent of boundary conditions, to some suitable level

[2]This will be discussed more fully later.

of accuracy. If this is the case for all boundary conditions, then it has to be the case for the linear displacement and uniform traction conditions. Thus, a minimal requirement for a representative volume element is that $\boldsymbol{C}^{\mathrm{eff,u}} = \boldsymbol{C}^{\mathrm{eff,t}}$. Another, rather similar, requirement is that the Hill relation (1.10) should hold for *all* (sufficiently smooth) boundary conditions. It was in this context that Rodney Hill introduced the relation in 1951; the proof that it held rigorously for certain boundary conditions was not given until 1963.

1.4 Other properties

The ideas presented above are applicable to properties other than elasticity. Any form of conduction (thermal, electrical), and also magnetism, follows a similar pattern. A flux (now a vector rather than a tensor) $\boldsymbol{\sigma}$ is related to a potential gradient $\boldsymbol{\varepsilon}$, which in turn is derived from a scalar function u which is the negative of the usual scalar potential. Then,

$$\boldsymbol{\sigma} = \boldsymbol{C}\boldsymbol{\varepsilon}, \quad \text{or} \quad \sigma_i = C_{ij}\varepsilon_j, \quad \text{with} \quad \boldsymbol{\varepsilon} = \nabla u \quad \text{or} \quad \varepsilon_i = u_{,i}. \tag{1.12}$$

The second order tensor $\boldsymbol{C}$ is the conductivity tensor (or dielectric tensor, or magnetic permeability depending on context), and the equation of equilibrium is

$$\mathrm{div}\boldsymbol{\sigma} + f = 0, \quad \text{or} \quad \sigma_{i,i} + f = 0, \quad \boldsymbol{x} \in \Omega, \tag{1.13}$$

where f is the source term.

It is true also that the mean value theorems, and the Hill relation, do not rely on any constitutive relation at all.[3] Hence, they are equally available for exploitation for nonlinear response. One such response, still for small deformations, is physically-nonlinear elasticity (or deformation theory of plasticity), in which stress is related to strain via a potential function $W(\boldsymbol{\varepsilon})$ so that

$$\boldsymbol{\sigma} = W'(\boldsymbol{\varepsilon}), \quad \text{or} \quad \sigma_{ij} = \frac{\partial W(\boldsymbol{\varepsilon})}{\partial \varepsilon_{ij}}. \tag{1.14}$$

It is usual to assume that the function $W(\boldsymbol{\varepsilon})$ is convex, the linear case being recovered when W is a quadratic function.

The case of incremental plasticity (flow theory) can also be studied with the help of the formulae developed above. Since the constitutive response depends on the path in strain or stress space, it is necessary to treat such problems incrementally or, equivalently, in terms

[3] In particular, the Hill relation remains true when $\boldsymbol{\sigma}$ and $\boldsymbol{\varepsilon}$ have no relation to one another, even being associated with different boundary value problems. This observation will be exploited in a succeeding section.

of rates of stress and strain. These also obey the relations given, with the formal addition of a superposed dot to signify a rate of change. The complications presented by incremental plasticity are sufficiently severe that these notes will contain very little discussion of the subject.

The basic results given so far also generalise to large deformations. The theory goes through, virtually as presented above, if a Lagrangian description is adopted, with the strain $\boldsymbol{\varepsilon}$ replaced by the non-symmetric deformation gradient (usually called $\boldsymbol{F}$), and the Cauchy stress $\boldsymbol{\sigma}$ replaced by the conjugate of $\boldsymbol{F}$, the non-symmetric Boussinesq stress tensor $\boldsymbol{B}$, also called the Piola–Lagrange, or the first Piola–Kirchhoff stress tensor.

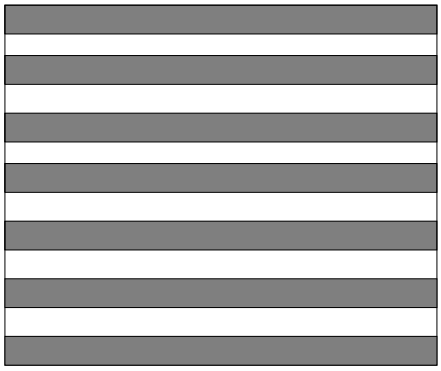

Figure 1: A simple laminate, with interfaces $x_3 = const.$

2 Theory of Laminates

2.1 An example

Figure 1 illustrates a simple laminated medium. It consists of a set of uniform laminae, with alternating properties, bonded together across interfaces, all of which are planes $x_3 = const.$ Suppose, for simplicity, that the two materials from which the laminae are composed are isotropic, with Young's moduli E^r and Poisson's ratios ν^r $(r = 1, 2)$. The volume fractions are c_r $(r = 1, 2)$, so that $c_1 + c_2 = 1$.

Our objective is to calculate the effective moduli of this laminate. Thus, we shall consider a domain, filled with the laminated material, which is large enough relative to the scale of the lamination to be regarded as a "representative volume element". The effective moduli are found by imposing boundary conditions that generate, "on average", uniform stresses and strains throughout the domain. Since by hypothesis we are dealing with a representative volume element, the exact choice of boundary conditions is irrelevant. In fact, the construction given below provides uniform mean stresses and strains, but the implied boundary conditions are neither of (1.4), (1.5): imposition of either of these conditions would generate fields that would differ from those that we shall construct, but only in a "boundary layer" whose thickness would be on the order of the scale of lamination, so that volume averages would be essentially unaffected.

Evidently, the effective response of this composite medium will display transverse isotropy, with symmetry axis parallel to the x_3-axis. The stress-strain relations for isotropic material have the form

$$E\varepsilon_{ij} = (1 + \nu)\sigma_{ij} - \nu\delta_{ij}\sigma_{kk}. \tag{2.1}$$

The stress-strain relations for transverse isotropy can be given in the form

$$\begin{aligned} E_1\varepsilon_{11} &= \sigma_{11} - \nu_{12}\sigma_{22} - \nu_{13}\sigma_{33}, \\ E_1\varepsilon_{22} &= \sigma_{22} - \nu_{12}\sigma_{11} - \nu_{13}\sigma_{33}, \\ E_1\varepsilon_{12} &= (1+\nu_{12})\sigma_{12}, \\ E_3\varepsilon_{33} &= \sigma_{33} - \nu_{31}(\sigma_{11}+\sigma_{22}), \\ 2G_{13}\varepsilon_{13} &= \sigma_{13}, \quad 2G_{13}\varepsilon_{23} = \sigma_{23}, \end{aligned} \tag{2.2}$$

with the interrelation $\nu_{13}/E_1 = \nu_{31}/E_3$.

Consider first the effective constant E_3^{eff}. This is obtained by imposing upon the composite a mean stress whose only non-zero component is $\overline{\sigma}_{33}$. In fact, for equilibrium, it has to follow that $\sigma_{33} = \overline{\sigma}_{33}$ throughout the composite. Poisson effects will generate non-zero values for ε_{11} and ε_{22} which are equal, from symmetry, in each lamina. It is possible to find a field such that these strains are constant in each lamina. In this case, necessarily, $\varepsilon_{11} = \varepsilon_{22} = \overline{\varepsilon}_{11}$, since otherwise the displacement would not be continuous across interfaces. The value of $\overline{\varepsilon}_{11}$ is fixed by the requirement that the mean value of the stress component σ_{11} must be zero.

Now in any one of the laminae,

$$E\varepsilon_{33} = \overline{\sigma}_{33} - 2\nu\sigma_{11}, \tag{2.3}$$

$$E\overline{\varepsilon}_{11} = (1-\nu)\sigma_{11} - \nu\overline{\sigma}_{33}. \tag{2.4}$$

The second of these equations gives

$$\sigma_{11} = \frac{E}{(1-\nu)}\overline{\varepsilon}_{11} + \frac{\nu}{(1-\nu)}\overline{\sigma}_{33}. \tag{2.5}$$

The requirement that the mean value of σ_{11} must be zero gives

$$\overline{\varepsilon}_{11} = -\left\langle \frac{E}{(1-\nu)} \right\rangle^{-1} \left\langle \frac{\nu}{(1-\nu)} \right\rangle \overline{\sigma}_{33}. \tag{2.6}$$

Here and in what follows, the angled bracket is employed as an alternative notation for the mean value: $\langle\phi\rangle = c_1\phi^1 + c_2\phi^2$ in the case of a two-component laminate.

Comparison of the first of equations (2.2) with (2.6) gives, immediately,

$$\frac{\nu_{13}^{\text{eff}}}{E_1^{\text{eff}}} = \left\langle \frac{E}{(1-\nu)} \right\rangle^{-1} \left\langle \frac{\nu}{(1-\nu)} \right\rangle = \frac{\nu_{31}^{\text{eff}}}{E_3^{\text{eff}}}. \tag{2.7}$$

Substitution of (2.6) into (2.5) gives, with (2.3),

$$\varepsilon_{33} = \left\{ \frac{(1-2\nu)(1+\nu)}{E(1-\nu)} + \frac{2\nu}{(1-\nu)} \left\langle \frac{E}{(1-\nu)} \right\rangle^{-1} \left\langle \frac{\nu}{(1-\nu)} \right\rangle \right\} \overline{\sigma}_{33}. \tag{2.8}$$

The average of this gives, by definition, $\overline{\sigma}_{33}/E_3^{\text{eff}}$. Thus,

$$E_3^{\text{eff}} = \left\langle \frac{(1-2\nu)(1+\nu)}{E(1-\nu)} + \frac{2\nu}{(1-\nu)} \left\langle \frac{E}{(1-\nu)} \right\rangle^{-1} \left\langle \frac{\nu}{(1-\nu)} \right\rangle \right\rangle^{-1}. \tag{2.9}$$

Next, choose a mean stress with $\overline{\sigma}_{11} = \overline{\sigma}_{22}$, all other components being zero. Correspondingly, $\varepsilon_{11} = \varepsilon_{22} = \overline{\varepsilon}_{11}$ (this being so far unknown). Also, $\sigma_{33} = 0$. Thus,

$$E\overline{\varepsilon}_{11} = (1-\nu)\sigma_{11}, \tag{2.10}$$

$$E\varepsilon_{33} = -2\nu\sigma_{11}. \tag{2.11}$$

It follows that

$$\sigma_{11} = \frac{E}{(1-\nu)}\overline{\varepsilon}_{11} \tag{2.12}$$

and therefore by averaging,

$$\frac{E_1^{\text{eff}}}{(1-\nu_{12}^{\text{eff}})} = \left\langle \frac{E}{(1-\nu)} \right\rangle. \tag{2.13}$$

Also, from the equations above,

$$\varepsilon_{33} = -\frac{2\nu}{(1-\nu)}\overline{\varepsilon}_{11} \tag{2.14}$$

and hence, by averaging and comparing with the corresponding transversely isotropic effective relation,

$$\frac{\nu_{13}^{\text{eff}}}{(1-\nu_{12}^{\text{eff}})} = \left\langle \frac{\nu}{(1-\nu)} \right\rangle. \tag{2.15}$$

A further independent relation is obtained by imposing on the laminate a mean strain whose only non-zero component is $\overline{\varepsilon}_{12}$. Then, continuity of displacements across interfaces requires that $\varepsilon_{12} = \overline{\varepsilon}_{12}$ in each lamina. The stress component σ_{12} is therefore $E/(1+\nu)$ times $\overline{\varepsilon}_{12}$. Hence, by averaging,

$$\frac{E_1^{\text{eff}}}{(1+\nu_{12}^{\text{eff}})} = \left\langle \frac{E}{(1+\nu)} \right\rangle. \tag{2.16}$$

Finally, prescribing $\overline{\sigma}_{13}$, which implies that σ_{13} takes that value throughout, with other components equal to zero, gives $\varepsilon_{13} = \frac{(1+\nu)}{E}\overline{\sigma}_{13}$ and therefore, by averaging,

$$2G_{13}^{\text{eff}} = \left\langle \frac{(1+\nu)}{E} \right\rangle^{-1} . \tag{2.17}$$

In concluding this subsection, it is remarked that, although the formulae were introduced through considering a two-component composite, the reasoning applies unchanged to a laminate made of any number of isotropic materials.

In practical applications, it is usual that the individual laminae will be anisotropic: each lamina is often a fibre-reinforced composite, for example. The next subsection shows how the algebra can be completed in a concise way, even in this case.

2.2 A more general discussion of simple laminates

Consider now a general two-component laminate, for which the direction of lamination is defined by the common normal $\boldsymbol{n}$ of all of the interfaces. The elastic constant tensors $\boldsymbol{C}^r$ $(r = 1, 2)$ are allowed to be anisotropic. As in the previous subsection, solutions which deliver $\boldsymbol{C}^{\text{eff}}$ can be constructed in which the stress and strain fields are piecewise constant. It is not necessary that the lamination has to display periodicity: the formulae to be given remain valid even if the laminate has a random structure. It is actually convenient to work in terms of the displacement gradient[4] $\boldsymbol{d} = \nabla \otimes \boldsymbol{u}$ rather than its symmetric part, $\boldsymbol{\varepsilon}$. It is consistent to assume that the displacement gradient in component r takes the constant value $\boldsymbol{d}^r$; the stress in component r then takes the constant value $\boldsymbol{\sigma}^r = \boldsymbol{C}^r\boldsymbol{d}^r \equiv \boldsymbol{C}^r\boldsymbol{\varepsilon}^r$.[5] The requirement that the displacement should be continuous as well as piecewise-linear means that it must comprise the sum of a linear function of $\boldsymbol{x}$, $\boldsymbol{u}^{\text{l}}$ say, and a continuous but piecewise-linear function, $\boldsymbol{u}^{\text{p-l}}$ say, of the "normal" coordinate $\boldsymbol{x}\cdot\boldsymbol{n}$. Correspondingly,

$$\boldsymbol{d} = \boldsymbol{d}^{\text{l}} + \boldsymbol{n} \otimes (\boldsymbol{u}^{\text{p-l}})', \tag{2.18}$$

where the prime signifies differentiation with respect to $\boldsymbol{x}\cdot\boldsymbol{n}$; the function $(\boldsymbol{u}^{\text{p-l}})'$ is piecewise-constant. The requirement that the mean displacement gradient should have the value $\overline{\boldsymbol{d}}$ (with symmetric part $\overline{\boldsymbol{\varepsilon}}$) allows (2.18) to be reduced to the form

$$\boldsymbol{d}^1 = \overline{\boldsymbol{d}} - c_2\boldsymbol{n} \otimes \boldsymbol{\alpha}, \quad \boldsymbol{d}^2 = \overline{\boldsymbol{d}} + c_1\boldsymbol{n} \otimes \boldsymbol{\alpha}, \tag{2.19}$$

[4]The definition used here gives the transpose of the usual "$\nabla\boldsymbol{u}$", *i.e.* $\boldsymbol{d}$ has components $d_{ij} = (\partial/\partial x_i)u_j \equiv u_{j,i}$.

[5]The symmetry of the elastic constant tensor ensures that this expression only depends on the value of the strain.

where $\boldsymbol{\alpha}$ is so far unknown.

The stresses $\boldsymbol{\sigma}^r$ $(r = 1, 2)$ can now be given in the form

$$\sigma^1_{ij} = C^1_{ijkl}(\overline{\varepsilon}_{kl} - c_2\alpha_k n_l), \quad \sigma^2_{ij} = C^2_{ijkl}(\overline{\varepsilon}_{kl} + c_1\alpha_k n_l). \tag{2.20}$$

Finally, equilibrium requires that $\sigma^r_{ij}n_j$ is the same for all r, and so is equal to $\overline{\sigma}_{ij}n_j$, in which

$$\overline{\boldsymbol{\sigma}} = c_1\boldsymbol{\sigma}^1 + c_2\boldsymbol{\sigma}^2. \tag{2.21}$$

Thus, with the notation

$$K^r_{ik} = C^r_{ijkl}n_j n_l, \tag{2.22}$$

considering the component 1,

$$n_j C^1_{ijkl}\overline{\varepsilon}_{kl} - c_2 K^1_{ik}\alpha_k = n_j\overline{\sigma}_{ij}. \tag{2.23}$$

A similar relation applies to the component 2 but it carries the same information. Then, in an obvious symbolic notation,

$$\boldsymbol{\alpha} = (c_2)^{-1}(\boldsymbol{K}^1)^{-1}[\boldsymbol{n}(\boldsymbol{C}^1\overline{\boldsymbol{\varepsilon}} - \overline{\boldsymbol{\sigma}})]. \tag{2.24}$$

Therefore, substituting this into equations (2.19),

$$\begin{aligned} \boldsymbol{d}^1 &= \overline{\boldsymbol{d}} - \boldsymbol{\Gamma}^1(\boldsymbol{n})\boldsymbol{C}^1\overline{\boldsymbol{\varepsilon}} + \boldsymbol{\Gamma}^1(\boldsymbol{n})\overline{\boldsymbol{\sigma}}, \\ \boldsymbol{d}^2 &= \overline{\boldsymbol{d}} + (c_2)^{-1}c_1\boldsymbol{\Gamma}^1(\boldsymbol{n})\boldsymbol{C}^1\overline{\boldsymbol{\varepsilon}} - (c_2)^{-1}c_1\boldsymbol{\Gamma}^1(\boldsymbol{n})\overline{\boldsymbol{\sigma}}, \end{aligned} \tag{2.25}$$

where

$$\boldsymbol{\Gamma}^r(\boldsymbol{n}) = \boldsymbol{n} \otimes (\boldsymbol{K}^r)^{-1} \otimes \boldsymbol{n}. \tag{2.26}$$

It follows now that

$$\overline{\boldsymbol{\sigma}} = \overline{\boldsymbol{C}}\overline{\boldsymbol{\varepsilon}} - c_1\boldsymbol{C}^1\boldsymbol{\Gamma}^1(\boldsymbol{n})\boldsymbol{C}^1\overline{\boldsymbol{\varepsilon}} + c_1\boldsymbol{C}^2\boldsymbol{\Gamma}^1(\boldsymbol{n})\boldsymbol{C}^1\overline{\boldsymbol{\varepsilon}} + c_1(\boldsymbol{C}^1 - \boldsymbol{C}^2)\boldsymbol{\Gamma}^1(\boldsymbol{n})\overline{\boldsymbol{\sigma}}. \tag{2.27}$$

In fact, because the elastic constant tensors, and the mean stress, display the usual symmetries, only the symmetrized part of the operator $\boldsymbol{\Gamma}^1$ participates in equation (2.27), and hence $\boldsymbol{\Gamma}^1$ can be replaced by its symmetrized form, $\tilde{\boldsymbol{\Gamma}}^1$ say[6], which has components

$$\tilde{\Gamma}^1_{ijkl} = \tfrac{1}{4}\{\Gamma^1_{ijkl} + \Gamma^1_{jikl} + \Gamma^1_{ijlk} + \Gamma^1_{jilk}\}. \tag{2.28}$$

[6]This notation is chosen because the operator $\tilde{\boldsymbol{\Gamma}}^1$ will emerge, from entirely different considerations, in Section 6.

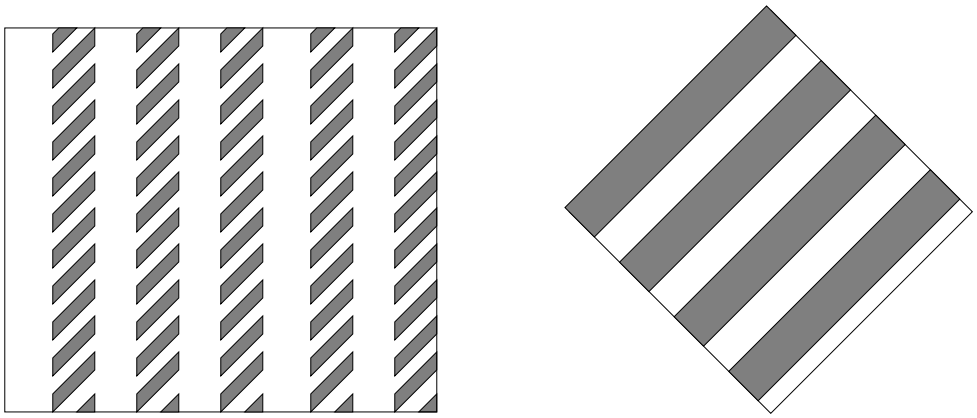

Figure 2: (a) A rank-2 laminate, made by laminating "material 1" with material which itself is a laminate of "material 1" with "material 2", (b) a magnified picture of the simple laminate.

Solving equation (2.27) for $\overline{\boldsymbol{\sigma}}$ gives

$$\overline{\boldsymbol{\sigma}} = \boldsymbol{C}^{\text{eff}}\overline{\boldsymbol{\varepsilon}}, \tag{2.29}$$

where

$$\begin{aligned}
\boldsymbol{C}^{\text{eff}} &= [\boldsymbol{I} + c_1(\boldsymbol{C}^2 - \boldsymbol{C}^1)\tilde{\boldsymbol{\Gamma}}^1(\boldsymbol{n})]^{-1}\{c_1\boldsymbol{C}^1 + c_2\boldsymbol{C}^2 + c_1(\boldsymbol{C}^2 - \boldsymbol{C}^1)\tilde{\boldsymbol{\Gamma}}^1(\boldsymbol{n})\boldsymbol{C}^1\} \\
&= [\boldsymbol{I} + c_1(\boldsymbol{C}^2 - \boldsymbol{C}^1)\tilde{\boldsymbol{\Gamma}}^1(\boldsymbol{n})]^{-1}\{[\boldsymbol{I} + c_1(\boldsymbol{C}^2 - \boldsymbol{C}^1)\tilde{\boldsymbol{\Gamma}}^1(\boldsymbol{n})]\boldsymbol{C}^1 + c_2(\boldsymbol{C}^2 - \boldsymbol{C}^1)\} \\
&= \boldsymbol{C}^1 + c_2[\boldsymbol{I} + c_1(\boldsymbol{C}^2 - \boldsymbol{C}^1)\tilde{\boldsymbol{\Gamma}}^1(\boldsymbol{n})]^{-1}(\boldsymbol{C}^2 - \boldsymbol{C}^1).
\end{aligned} \tag{2.30}$$

The expression for $\boldsymbol{C}^{\text{eff}}$ can be given in a variety of forms. For instance, the reasoning could be repeated with the indices 1 and 2 interchanged. It is quite difficult to confirm, algebraically, that the resulting $\boldsymbol{C}^{\text{eff}}$ is the same! It relies on the easily-verified identity

$$\boldsymbol{\Gamma}^1(\boldsymbol{n})(\boldsymbol{C}^1 - \boldsymbol{C}^2)\boldsymbol{\Gamma}^2(\boldsymbol{n}) = \boldsymbol{\Gamma}^2 - \boldsymbol{\Gamma}^1. \tag{2.31}$$

The same type of reasoning leads to an explicit expression for $\boldsymbol{C}^{\text{eff}}$ in the case of a laminate with n components. This is not pursued because a very convenient formula will be developed by use of more general methodology, in a later section.

2.3 Hierarchical laminates

Figure 2a shows a two-component laminate which is constructed by laminating one of the media – medium 1 say – with another medium which itself is a laminate, a magnified version

of which is shown in Figure 2b. Assuming that there is a separation of scales, the response of the medium shown in Figure 2a can be calculated using equation (2.30), with the new "medium 2", now itself a laminate but on a much finer scale, treated as a homogeneous medium with elastic constant tensor given by the appropriate realisation of (2.30). The calculation is facilitated by placing the relation (2.30) in the form

$$(\boldsymbol{C}^{\text{eff}} - \boldsymbol{C}^1)^{-1} = (c_2)^{-1}(\boldsymbol{C}^2 - \boldsymbol{C}^1)^{-1} + (c_2)^{-1}c_1\tilde{\boldsymbol{\Gamma}}^1(\boldsymbol{n}). \tag{2.32}$$

Now we introduce the parameters defining the rank-2 laminate of Figure 2. First, the simple laminate is composed by laminating medium 1, whose elastic constant tensor is $\boldsymbol{C}^1$, with medium 2, whose elastic constant tensor is $\boldsymbol{C}^2$, at volume fraction $c_2^{(1)}$; the volume fraction of medium 1 is then $c_1^{(1)} = 1 - c_2^{(1)}$. The normal defining the direction of lamination is $\boldsymbol{n}^1$. Then, the relation (2.32) shows that its effective elastic constant tensor $\boldsymbol{C}^{\text{eff},1}$ satisfies

$$(\boldsymbol{C}^{\text{eff},1} - \boldsymbol{C}^1)^{-1} = (c_2^{(1)})^{-1}(\boldsymbol{C}^2 - \boldsymbol{C}^1)^{-1} + (c_2^{(1)})^{-1}c_1^{(1)}\tilde{\boldsymbol{\Gamma}}^1(\boldsymbol{n}^1). \tag{2.33}$$

Now create the rank-2 laminate by laminating medium 1 with a medium with elastic constant tensor $\boldsymbol{C}^{\text{eff},1}$, in direction $\boldsymbol{n}^2$, at volume fraction $c_2^{(2)}$. The volume fraction of the original medium 2 is now $c_2 = c_2^{(1)}c_2^{(2)}$, and that of medium 1 is $c_1 = 1 - c_2$. Application of the formula (2.32) shows that the effective elastic constant tensor $\boldsymbol{C}^{\text{eff},2}$ of the rank-2 laminate satisfies

$$\begin{aligned}
(\boldsymbol{C}^{\text{eff},2} - \boldsymbol{C}^1)^{-1} &= (c_2^{(2)})^{-1}(\boldsymbol{C}^{\text{eff},1} - \boldsymbol{C}^1)^{-1} + (c_2^{(2)})^{-1}c_1^{(2)}\tilde{\boldsymbol{\Gamma}}^1(\boldsymbol{n}^2) \\
&= (c_2^{(1)}c_2^{(2)})^{-1}(\boldsymbol{C}^2 - \boldsymbol{C}^1)^{-1} + (c_2^{(1)}c_2^{(2)})^{-1}c_1^{(1)}\tilde{\boldsymbol{\Gamma}}^1(\boldsymbol{n}^1) + (c_2^{(2)})^{-1}c_1^{(2)}\tilde{\boldsymbol{\Gamma}}^1(\boldsymbol{n}^2) \\
&= (c_2)^{-1}(\boldsymbol{C}^2 - \boldsymbol{C}^1)^{-1} + (c_2)^{-1}c_1^{(1)}\tilde{\boldsymbol{\Gamma}}^1(\boldsymbol{n}^1) + (c_2^{(2)})^{-1}c_1^{(2)}\tilde{\boldsymbol{\Gamma}}^1(\boldsymbol{n}^2).
\end{aligned} \tag{2.34}$$

The interesting thing to note about this formula is that the sum of the "weights" of the terms involving $\tilde{\boldsymbol{\Gamma}}^1$ is

$$(c_2)^{-1}(c_1^{(1)} + c_2^{(1)}c_1^{(2)}) = (c_2)^{-1}(1 - c_2^{(1)}c_2^{(2)}) = (c_2)^{-1}(1 - c_2) = (c_2)^{-1}c_1,$$

just as in the simple lamination formula (2.32). Since $\boldsymbol{n}^1$ and $\boldsymbol{n}^2$ are unit vectors, the function $\tilde{\boldsymbol{\Gamma}}^1$ can be considered to be evaluated at points on the unit sphere. The lamination formula (2.34) requires the value of the function $\tilde{\boldsymbol{\Gamma}}^1$ to be shared between two points, with the same total weight, c_1/c_2, as for the simple laminate with the same volume fractions. This pattern repeats, in fact, for any hierarchical laminate. In the limit of an infinite hierarchy,

the weighted sum of the values of $\tilde{\mathbf{\Gamma}}^1$ over the unit sphere becomes an integral. The weights define a measure over the unit sphere, called the H-measure (H for homogenization). A finite-rank laminate can be viewed in the same framework, by recognising the measure as a sum of Dirac masses. The H-measure will appear in a different context later.

3 Energy relations

The stress-strain relation for a linearly elastic body is (1.1) where, in general, the tensor of moduli $\boldsymbol{C}$ varies with position, $\boldsymbol{x}$. It is assumed throughout these lectures that $\boldsymbol{C}$ has the symmetries

$$C_{ijkl} = C_{klij} \tag{3.1}$$

as well as $C_{ijkl} = C_{jikl} = C_{ijlk}$ which follow from the symmetry of the stress and strain tensors. The important symmetry (3.1) ensures the existence of an energy density $W(\boldsymbol{\varepsilon})$ such that

$$W(\boldsymbol{\varepsilon}) = \tfrac{1}{2} C_{ijkl}\varepsilon_{ij}\varepsilon_{kl} \;\text{ or, symbolically, }\; W(\boldsymbol{\varepsilon}) = \tfrac{1}{2}\boldsymbol{\varepsilon}\boldsymbol{C}\boldsymbol{\varepsilon}. \tag{3.2}$$

The energy is assumed to be a positive-definite function of $\boldsymbol{\varepsilon}$, and therefore convex.

3.1 The principle of virtual work

In this section and elsewhere, use will be made of a simple consequence of the divergence theorem, known as the principle of virtual work:
If $\boldsymbol{\sigma}$ is a stress field that satisfies

$$\sigma_{ij,j} + f_i = 0, \quad \boldsymbol{x} \in \Omega, \tag{3.3}$$

then

$$\int_\Omega \sigma_{ij}\varepsilon'_{ij}\,\mathrm{d}\boldsymbol{x} = \int_\Omega f_i u'_i\,\mathrm{d}\boldsymbol{x} + \int_{\partial\Omega} \sigma_{ij}n_j u'_i\,\mathrm{d}S, \tag{3.4}$$

where $\boldsymbol{u}'$ is *any* displacement field and $\boldsymbol{\varepsilon}'$ is the associated strain field.
Proof: The symmetry ($\sigma_{ij} = \sigma_{ji}$) of the stress tensor allows ε'_{ij} on the left side of (3.4) to be replaced by the displacement gradient $u'_{i,j}$. Thus,

$$\int_\Omega \sigma_{ij}\varepsilon'_{ij}\,\mathrm{d}\boldsymbol{x} = \int_\Omega \sigma_{ij}u'_{i,j}\,\mathrm{d}\boldsymbol{x} = \int_\Omega [(\sigma_{ij}u'_i)_{,j} - \sigma_{ij,j}u'_i]\,\mathrm{d}\boldsymbol{x}. \tag{3.5}$$

The result (3.4) now follows from the divergence theorem, coupled with the fact that the stress field satisfies equations (3.3).

It should be noted that the identity (3.4) continues to hold, even if the stress field $\boldsymbol{\sigma}$ is discontinuous across a set of internal surfaces, provided that $\sigma_{ij}n_j$ is continuous across any such surface: this follows by applying the basic identity to each sub-domain within which $\sigma_{ij,j} + f_i = 0$, and recognising that the integrals over the internal surfaces cancel out. *Throughout these notes, the equilibrium equation (3.3) will be given this "generalized function" interpretation without explicit comment.*

3.2 The classical energy principles

Subject to the assumption that the energy function is positive-definite, equilibrium of the body is governed by the

Minimum Energy Principle: if the body is subjected to body-force $\boldsymbol{f}$ per unit volume, and a part S_u of its boundary is subjected to prescribed displacements $\boldsymbol{u}^0$ while the complementary part S_t is subjected to prescribed tractions $\boldsymbol{t}^0$, the equilibrium displacement minimises the energy functional

$$\mathcal{F}(\boldsymbol{u}) := \int_\Omega (\tfrac{1}{2} C_{ijkl}\varepsilon_{ij}\varepsilon_{kl} - f_i u_i)\mathrm{d}\boldsymbol{x} - \int_{S_t} t_i^0 u_i \mathrm{d}S, \tag{3.6}$$

amongst displacement fields which take the prescribed boundary values over S_u.[7]

Proof: Let $\boldsymbol{u}$ be the displacement field that minimises $\mathcal{F}$ subject to the given conditions, and let $\boldsymbol{u}'$ be any trial field (so that also $\boldsymbol{u}' = \boldsymbol{u}^0$ on S_u). Then

$$\begin{aligned}
\mathcal{F}(\boldsymbol{u}') - \mathcal{F}(\boldsymbol{u}) &= \int_\Omega [\tfrac{1}{2} C_{ijkl}(\varepsilon'_{ij} - \varepsilon_{ij})(\varepsilon'_{kl} - \varepsilon_{kl}) + C_{ijkl}(\varepsilon'_{ij} - \varepsilon_{ij})\varepsilon_{kl} - f_i(u'_i - u_i)]\,\mathrm{d}\boldsymbol{x} \\
&\qquad - \int_{S_t} t_i^0 (u'_i - u_i)\,\mathrm{d}S \\
&= \int_\Omega [\tfrac{1}{2} C_{ijkl}(\varepsilon'_{ij} - \varepsilon_{ij})(\varepsilon'_{kl} - \varepsilon_{kl}) + \sigma_{ij}(\varepsilon'_{ij} - \varepsilon_{ij}) - f_i(u'_i - u_i)]\,\mathrm{d}\boldsymbol{x} \\
&\qquad - \int_{S_t} t_i^0 (u'_i - u_i)\,\mathrm{d}S.
\end{aligned} \tag{3.7}$$

Here, we have written $\sigma_{ij} = C_{ijkl}\varepsilon_{kl}$. Use of the divergence theorem as in the proof of the principle of virtual work now gives

$$\begin{aligned}
\mathcal{F}(\boldsymbol{u}') - \mathcal{F}(\boldsymbol{u}) &= \int_\Omega [\tfrac{1}{2} C_{ijkl}(\varepsilon'_{ij} - \varepsilon_{ij})(\varepsilon'_{kl} - \varepsilon_{kl}) - (\sigma_{ij,j} + f_i)(u'_i - u_i)]\,\mathrm{d}\boldsymbol{x} \\
&\qquad + \int_{S_t} (\sigma_{ij} n_j - t_i^0)(u'_i - u_i)\,\mathrm{d}S,
\end{aligned} \tag{3.8}$$

since $u'_i - u_i = 0$ on S_u. It follows that

$$\mathcal{F}(\boldsymbol{u}') - \mathcal{F}(\boldsymbol{u}) = \int_\Omega \tfrac{1}{2} C_{ijkl}(\varepsilon'_{ij} - \varepsilon_{ij})(\varepsilon'_{kl} - \varepsilon_{kl})\,\mathrm{d}\boldsymbol{x} \tag{3.9}$$

for all trial displacements $\boldsymbol{u}'$, if and only if

$$\sigma_{ij,j} + f_i = 0, \quad \text{with } \sigma_{ij} = C_{ijkl}\varepsilon_{kl}, \quad \boldsymbol{x} \in \Omega, \tag{3.10}$$

[7] A more precise discussion would specify that $\boldsymbol{u}$ must belong to the Sobolev space $H^1(\Omega)$ of functions whose gradients are square-integrable over Ω, and place a suitable restriction on the prescribed boundary values.

and

$$\sigma_{ij}n_j = t_i^0, \quad \boldsymbol{x} \in S_t. \tag{3.11}$$

These are the conditions for stationarity of the functional $\mathcal{F}$. Now assume that the energy density function is positive-definite: it follows immediately that $\boldsymbol{u}$ provides the minimum value for $\mathcal{F}$. The usual uniqueness theorem for elastostatics also follows immediately: if $\boldsymbol{u}'$ were also a solution, it would follow that

$$\int_\Omega \tfrac{1}{2}C_{ijkl}(\varepsilon'_{ij} - \varepsilon_{ij})(\varepsilon'_{kl} - \varepsilon_{kl})\,\mathrm{d}\boldsymbol{x} = 0, \tag{3.12}$$

and hence that $\boldsymbol{\varepsilon}' = \boldsymbol{\varepsilon}$, almost everywhere in Ω. This shows that stress and strain are unique, and that displacement gradients can differ at most by a field of pure rotation. This is possible only if the rotation corresponds to that of a rigid body, and this degree of non-uniqueness is allowed if S_t is the whole of the boundary $\partial\Omega$. Otherwise, if displacements are prescribed over some part of $\partial\Omega$, the rigid rotation must be zero and the displacement field is unique.

It is possible also to define a complementary energy density

$$W^*(\boldsymbol{\sigma}) = \tfrac{1}{2}S_{ijkl}\sigma_{ij}\sigma_{kl} \equiv \tfrac{1}{2}\boldsymbol{\sigma}\boldsymbol{S}\boldsymbol{\sigma}. \tag{3.13}$$

The fact that this is numerically equal to $W(\boldsymbol{\varepsilon})$ when $\boldsymbol{\sigma} = \boldsymbol{C}\boldsymbol{\varepsilon}$ is not so important as the fact that W and W^* are Legendre duals:

$$W^*(\boldsymbol{\sigma}) = \sigma_{ij}\varepsilon_{ij} - W(\boldsymbol{\varepsilon}); \quad \boldsymbol{\varepsilon} = \boldsymbol{S}\boldsymbol{\sigma}, \quad \boldsymbol{\sigma} = \boldsymbol{C}\boldsymbol{\varepsilon}. \tag{3.14}$$

(more will be said about this later). Equilibrium is equally defined by the
Complementary Energy Principle: For the boundary value problem described above, the actual stress field minimises the functional

$$\mathcal{G}(\boldsymbol{\sigma}) := \int_\Omega \tfrac{1}{2}S_{ijkl}\sigma_{ij}\sigma_{kl}\mathrm{d}\boldsymbol{x} - \int_{S_u} \sigma_{ij}n_j u_i^0 \mathrm{d}S, \tag{3.15}$$

amongst stress fields that satisfy the equations of equilibrium $\sigma_{ij,j} + f_i = 0$ in Ω, and the given traction conditions on S_t.

The proof of the complementary energy principle is very similar to that for the minimum energy principle and is left as an exercise.

3.3 Implications for effective properties

To save notation in performing volume averages, we adopt the convention of taking the unit of length to be such that the domain Ω has unit volume.

(a) Linear displacement boundary condition. Under this boundary condition, and with no body force, the minimum energy principle gives

$$W^{\rm eff}(\overline{\boldsymbol{\varepsilon}}) := \int_\Omega W(\boldsymbol{\varepsilon}, \boldsymbol{x})\,{\rm d}\boldsymbol{x} \le \int_\Omega W(\boldsymbol{\varepsilon}', \boldsymbol{x})\,{\rm d}\boldsymbol{x}, \tag{3.16}$$

where $\boldsymbol{u}$ is the actual displacement and $\boldsymbol{u}'$ is *any* displacement that satisfies the linear displacement boundary condition (1.4).

Furthermore, it can be proved that

$$\overline{\sigma}_{ij} \equiv C^{\rm eff}_{ijkl}\overline{\varepsilon}_{kl} = \frac{\partial W^{\rm eff}(\overline{\boldsymbol{\varepsilon}})}{\partial \overline{\varepsilon}_{ij}}, \tag{3.17}$$

as follows.

Let $\boldsymbol{u} + \delta\boldsymbol{u}$ be the actual displacement field associated with the imposed mean strain $\overline{\boldsymbol{\varepsilon}} + \delta\overline{\boldsymbol{\varepsilon}}$. Then, to first order,

$$\begin{aligned}
\delta W^{\rm eff} &= W^{\rm eff}(\overline{\boldsymbol{\varepsilon}} + \delta\overline{\boldsymbol{\varepsilon}}) - W^{\rm eff}(\overline{\boldsymbol{\varepsilon}}) = \int_\Omega \{W(\boldsymbol{\varepsilon} + \delta\boldsymbol{\varepsilon}, \boldsymbol{x}) - W(\boldsymbol{\varepsilon}, \boldsymbol{x})\}\,{\rm d}\boldsymbol{x} \\
&= \int_\Omega \frac{\partial W(\boldsymbol{\varepsilon}, \boldsymbol{x})}{\partial \varepsilon_{ij}}\delta\varepsilon_{ij}\,{\rm d}\boldsymbol{x} = \int_\Omega \sigma_{ij}\delta\varepsilon_{ij}\,{\rm d}\boldsymbol{x} \\
&= \int_{\partial\Omega} \sigma_{ij}n_j\delta u_i\,{\rm d}S \ \text{ (by the principle of virtual work)} \\
&= \int_{\partial\Omega} \sigma_{ij}n_j\delta\overline{\varepsilon}_{ik}x_k\,{\rm d}S \\
&= \overline{\sigma}_{ik}\delta\overline{\varepsilon}_{ik} \ \text{ (by the relation (1.7)).}
\end{aligned} \tag{3.18}$$

This implies the relation (3.17) which demonstrates, in turn, that the tensor of effective moduli $\boldsymbol{C}^{\rm eff}$ has the symmetry

$$C^{\rm eff}_{ijkl} = C^{\rm eff}_{klij}, \tag{3.19}$$

and that

$$W^{\rm eff}(\overline{\boldsymbol{\varepsilon}}) = \tfrac{1}{2}\overline{\varepsilon}_{ij}C^{\rm eff}_{ijkl}\overline{\varepsilon}_{kl}. \tag{3.20}$$

Consider now the application of the complementary energy principle to the problem with the linear displacement boundary condition imposed. This gives

$$\int_\Omega \tfrac{1}{2}\sigma_{ij}S_{ijkl}(\boldsymbol{x})\sigma_{kl}\,{\rm d}\boldsymbol{x} - \int_{\partial\Omega} \sigma_{ij}n_j\overline{\varepsilon}_{ik}x_k\,{\rm d}S \le \int_\Omega \tfrac{1}{2}\sigma'_{ij}S_{ijkl}(\boldsymbol{x})\sigma'_{kl}\,{\rm d}\boldsymbol{x} - \int_{\partial\Omega} \sigma'_{ij}n_j\overline{\varepsilon}_{ik}x_k\,{\rm d}S, \tag{3.21}$$

where $\boldsymbol{\sigma}$ is the actual stress field and $\boldsymbol{\sigma}'$ is any stress field, in equilibrium with zero body force. Now the left side of (3.21) can equivalently be written as $-W^{\rm eff}(\overline{\boldsymbol{\varepsilon}})$, since the actual

stress is related to the actual displacement via (1.1), (1.2). Hence, the complementary energy principle and the minimum energy principle together give

$$\int_{\partial\Omega} \sigma'_{ij} n_j \overline{\varepsilon}_{ik} x_k \, \mathrm{d}S - \int_\Omega \tfrac{1}{2}\sigma'_{ij} S_{ijkl}(\boldsymbol{x}) \sigma'_{kl} \, \mathrm{d}\boldsymbol{x} \le W^{\mathrm{eff}}(\overline{\boldsymbol{\varepsilon}}) \le \int_\Omega \tfrac{1}{2}\varepsilon'_{ij} C_{ijkl}(\boldsymbol{x}) \varepsilon'_{kl} \, \mathrm{d}\boldsymbol{x}, \tag{3.22}$$

where $\boldsymbol{\sigma}'$ is any self-equilibrated stress field and $\boldsymbol{u}'$ is any displacement field that takes the prescribed values on $\partial\Omega$.

(b) Uniform boundary traction condition. Similar arguments can be advanced in relation to the uniform traction boundary condition, with the roles of the minimum energy and complementary energy principles interchanged. The results are

$$\overline{\varepsilon}_{ij} \equiv S^{\mathrm{eff}}_{ijkl} \overline{\sigma}_{kl} = \frac{\partial W^{*\,\mathrm{eff}}(\overline{\boldsymbol{\sigma}})}{\partial \overline{\sigma}_{ij}}, \tag{3.23}$$

$$S^{\mathrm{eff}}_{ijkl} = S^{\mathrm{eff}}_{klij}, \quad W^{*\,\mathrm{eff}}(\overline{\boldsymbol{\sigma}}) = \tfrac{1}{2}\overline{\sigma}_{ij} S^{\mathrm{eff}}_{ijkl} \overline{\sigma}_{kl}, \tag{3.24}$$

and

$$\int_{\partial\Omega} \overline{\sigma}_{ij} n_j u'_i \, \mathrm{d}S - \int_\Omega \tfrac{1}{2}\varepsilon'_{ij} C_{ijkl}(\boldsymbol{x}) \varepsilon'_{kl} \, \mathrm{d}\boldsymbol{x} \le W^{*\mathrm{eff}}(\overline{\boldsymbol{\sigma}}) \le \int_\Omega \tfrac{1}{2}\sigma'_{ij} S_{ijkl}(\boldsymbol{x}) \sigma'_{kl} \, \mathrm{d}\boldsymbol{x}, \tag{3.25}$$

where $\boldsymbol{u}'$ is any displacement field and $\boldsymbol{\sigma}'$ is any self-equilibrated stress field that satisfies the traction boundary conditions.

3.4 Elementary bounds for overall properties

The simplest "trial" fields for substitution into the bound formulae (3.22) are the displacement field $u'_i = \overline{\varepsilon}_{ij} x_j$, so that $\boldsymbol{\varepsilon}' = \overline{\boldsymbol{\varepsilon}}$, and the uniform stress field $\boldsymbol{\sigma}' = \boldsymbol{\sigma}^*$, constant, so that the restrictions are satisfied trivially. These fields give the inequalities

$$\sigma^*_{ij}\overline{\varepsilon}_{ij} - \tfrac{1}{2}\sigma^*_{ij}\overline{S}_{ijkl}\sigma^*_{kl} \le W^{\mathrm{eff}}(\overline{\boldsymbol{\varepsilon}}) \equiv \tfrac{1}{2}\overline{\varepsilon}_{ij} C^{\mathrm{eff}}_{ijkl} \overline{\varepsilon}_{kl} \le \tfrac{1}{2}\overline{\varepsilon}_{ij}\overline{C}_{ijkl}\overline{\varepsilon}_{kl}. \tag{3.26}$$

The lower bound given by the left inequality is true for any constant stress $\boldsymbol{\sigma}^*$. It is maximised by choosing the constants σ^*_{ij} so that

$$\overline{\varepsilon}_{ij} = \overline{S}_{ijkl}\sigma^*_{kl}, \quad \text{or} \quad \boldsymbol{\sigma}^* = (\overline{\boldsymbol{S}})^{-1}\overline{\boldsymbol{\varepsilon}}. \tag{3.27}$$

Thus, employing symbolic notation,

$$\tfrac{1}{2}\overline{\boldsymbol{\varepsilon}}(\overline{\boldsymbol{S}})^{-1}\overline{\boldsymbol{\varepsilon}} \le \tfrac{1}{2}\overline{\boldsymbol{\varepsilon}}\boldsymbol{C}^{\mathrm{eff}}\overline{\boldsymbol{\varepsilon}} \le \tfrac{1}{2}\overline{\boldsymbol{\varepsilon}}\overline{\boldsymbol{C}}\overline{\boldsymbol{\varepsilon}}. \tag{3.28}$$

Essentially the same bounds follow from the inequalities (3.25). They give

$$\tfrac{1}{2}\overline{\boldsymbol{\sigma}}(\overline{\boldsymbol{C}})^{-1}\overline{\boldsymbol{\sigma}} \leq \tfrac{1}{2}\overline{\boldsymbol{\sigma}}\boldsymbol{S}^{\text{eff}}\overline{\boldsymbol{\sigma}} \leq \tfrac{1}{2}\overline{\boldsymbol{\sigma}}\overline{\boldsymbol{S}}\overline{\boldsymbol{\sigma}}. \tag{3.29}$$

The elementary approximation $\boldsymbol{C}^{\text{eff}} \approx \overline{\boldsymbol{C}}$ (so that $\boldsymbol{S}^{\text{eff}} \approx (\overline{\boldsymbol{C}})^{-1}$) can be obtained directly by *assuming* that the strain field generated in the composite really is the constant field $\overline{\boldsymbol{\varepsilon}}$. Then, the stress at any point $\boldsymbol{x}$ is obtained as $\boldsymbol{\sigma}(\boldsymbol{x}) \approx \boldsymbol{C}(\boldsymbol{x})\overline{\boldsymbol{\varepsilon}}(\boldsymbol{x})$, and so $\overline{\boldsymbol{\sigma}} \approx \overline{\boldsymbol{C}}\overline{\boldsymbol{\varepsilon}}$. This is called the Voigt approximation, after its originator. Similarly, the elementary approximation $\boldsymbol{S}^{\text{eff}} \approx \overline{\boldsymbol{S}}$ (so that $\boldsymbol{C}^{\text{eff}} \approx (\overline{\boldsymbol{S}})^{-1}$) follows by *assuming* that the stress field that is generated in the composite really is a constant field. This is called the Reuss approximation. They were demonstrated to deliver bounds, subject to the assumed validity of the Hill condition, in about 1951; their rigorous status as bounds followed in 1963, when the Hill condition was deduced for the particular two types of boundary conditions that have been adopted here.

Since the bounds (3.28) involve only volume averages, the only information that they require about the composite consists of the volume fractions. That is both their strength and their weakness: they are true, universally (which is good), but they contain no information about the geometrical arrangement of the composite (it could be a laminate, or fibre-reinforced, or contain a dispersion of spherical inclusions, etc.) and so give no indication about the sensitivity of the effective properties to the microgeometry. The subject of bounds will receive further attention later.

4 Some General Relations for Composites

This section considers a general n-component (also called n-phase) composite. Phase r has elastic constant tensor $\boldsymbol{C}^r$ and volume fraction c_r. Here and throughout these notes, an overbar will imply a volume average. Thus, for example, for the elastic constant tensor,

$$\overline{\boldsymbol{C}} = \sum_{r=1}^{n} c_r \boldsymbol{C}^r. \tag{4.1}$$

4.1 Concentration tensors

Suppose that the composite is subjected to an overall mean strain $\overline{\boldsymbol{\varepsilon}}$, through imposition of the linear displacement boundary condition (1.4). The actual strain generated in the composite is $\boldsymbol{\varepsilon}(\boldsymbol{x})$. Let its average, over the total region occupied by phase r, be denoted $\boldsymbol{\varepsilon}^r$. Since the entire problem is linear, the strain at any point must be a linear function of the parameters $\overline{\boldsymbol{\varepsilon}}$ which define the boundary data, and its average over phase r must be likewise a linear function:

$$\boldsymbol{\varepsilon}^r = \boldsymbol{A}^r \overline{\boldsymbol{\varepsilon}}. \tag{4.2}$$

The fourth-order tensor $\boldsymbol{A}^r$ is the strain concentration tensor for phase r. The identity

$$\overline{\boldsymbol{\varepsilon}} = \sum_{r=1}^{n} c_r \boldsymbol{\varepsilon}^r \tag{4.3}$$

induces the identity

$$\sum_{r=1}^{n} c_r \boldsymbol{A}^r = \boldsymbol{I} \tag{4.4}$$

between the concentration tensors $\boldsymbol{A}^r$ $(r = 1, 2 \cdots n)$.

If the average of the stress over phase r is denoted $\boldsymbol{\sigma}^r$, it follows from the stress-strain relation that

$$\boldsymbol{\sigma}^r = \boldsymbol{C}^r \boldsymbol{\varepsilon}^r = \boldsymbol{C}^r \boldsymbol{A}^r \overline{\boldsymbol{\varepsilon}} \tag{4.5}$$

and hence that the overall mean stress is

$$\overline{\boldsymbol{\sigma}} = \boldsymbol{C}^{\text{eff}} \overline{\boldsymbol{\varepsilon}}, \tag{4.6}$$

where

$$\boldsymbol{C}^{\text{eff}} = \sum_{r=1}^{n} c_r \boldsymbol{C}^r \boldsymbol{A}^r. \tag{4.7}$$

If the composite consists of a matrix, phase n, say, containing different types of inclusions, it is natural to eliminate $\boldsymbol{A}^n$, using the identity (4.4), to give

$$\boldsymbol{C}^{\text{eff}} = \boldsymbol{C}^n + \sum_{r=1}^{n-1} c_r(\boldsymbol{C}^r - \boldsymbol{C}^n)\boldsymbol{A}^r. \tag{4.8}$$

4.2 A dilute suspension

In general, estimation of the concentration tensors $\boldsymbol{A}^r$ requires the solution – or at least approximate solution – of a complicated problem with interactions between inhomogeneities. If the suspension is dilute, however, then by definition interactions between different inhomogeneities are weak, and to lowest order, can be neglected. The problem then becomes that of estimating the strain within a single inclusion of type r, in a matrix with elastic constant tensor $\boldsymbol{C}^n$, subjected to the mean strain $\overline{\boldsymbol{\varepsilon}}$. It suffices, in fact, to take the matrix as infinite in extent, and to impose the condition that $\boldsymbol{\varepsilon} \to \overline{\boldsymbol{\varepsilon}}$ as $|\boldsymbol{x}| \to \infty$.

4.3 Isotropic matrix

This subsection presents, without derivation, some explicit formulae, in the case that the matrix and inclusions are isotropic. The derivation will be supplied, also for anisotropic media, in Section 6. It is first desirable to introduce notation that facilitates the algebra, now and also in later sections.

Notation for isotropic fourth-order tensors

When the matrix is isotropic, with Lamé moduli λ, μ and bulk modulus $\kappa = \lambda + \frac{2}{3}\mu$, the elastic constant tensor $\boldsymbol{C}$ can be given in the form

$$C_{ijkl} = \kappa\delta_{ij}\delta_{kl} + \mu(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk} - \tfrac{2}{3}\delta_{ij}\delta_{kl}). \tag{4.9}$$

It is convenient to employ the symbolic notation

$$\boldsymbol{C} = (3\kappa, 2\mu). \tag{4.10}$$

Then, for two isotropic fourth-order tensors with the symmetries associated with tensors of moduli, $\boldsymbol{C}_1 = (3\kappa_1, 2\mu_1)$ and $\boldsymbol{C}_2 = (3\kappa_2, 2\mu_2)$, their product becomes

$$\boldsymbol{C}_1\boldsymbol{C}_2 = ((3\kappa_1)(3\kappa_2), (2\mu_1)(2\mu_2)). \tag{4.11}$$

The identity tensor has the representation

$$\boldsymbol{I} = (1, 1) \tag{4.12}$$

and

$$\boldsymbol{C}^{-1} = (1/(3\kappa), 1/(2\mu)). \tag{4.13}$$

Isotropic spherical inclusion

If the inclusion is a sphere and is composed of isotropic material with Lamé moduli λ', μ' and bulk modulus $\kappa' = \lambda' + \frac{2}{3}\mu'$, then the associated strain concentration tensor $\boldsymbol{A}$ is isotropic and is given by the formula

$$\boldsymbol{A} = (3\kappa_A, 2\mu_A), \quad \text{or} \quad A_{ijkl} = \kappa_A \delta_{ij}\delta_{kl} + \mu_A(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk} - \tfrac{2}{3}\delta_{ij}\delta_{kl}), \tag{4.14}$$

where

$$3\kappa_A = \frac{3\kappa + 4\mu}{3\kappa' + 4\mu}, \quad 2\mu_A = \frac{5\mu(3\kappa + 4\mu)}{\mu(9\kappa + 8\mu) + 6\mu'(\kappa + 2\mu)}. \tag{4.15}$$

Thus, for a dilute suspension of isotropic spherical inclusions, at volume fraction c, in an isotropic matrix, the formula (4.8) gives the effective response as isotropic, with bulk and shear moduli κ^{eff}, μ^{eff} given by

$$\begin{aligned} \kappa^{\text{eff}} &= \kappa + c\frac{(\kappa' - \kappa)(3\kappa + 4\mu)}{3\kappa' + 4\mu}, \\ \mu^{\text{eff}} &= \mu + c\frac{5(\mu' - \mu)\mu(3\kappa + 4\mu)}{\mu(9\kappa + 8\mu) + 6\mu'(\kappa + 2\mu)}. \end{aligned} \tag{4.16}$$

Two particular cases are noted explicitly.

A dilute array of rigid spherical inclusions: If κ' and μ' tend to infinity, the formulae (4.16) reduce to

$$\kappa^{\text{eff, rigid}} = \kappa + c(\kappa + 4\mu/3), \quad \mu^{\text{eff, rigid}} = \mu + c\frac{5\mu(3\kappa + 4\mu)}{6(\kappa + 2\mu)}. \tag{4.17}$$

If, in addition, the matrix is incompressible, the effective bulk modulus becomes infinite, while the effective shear modulus reduces to

$$\mu^{\text{eff, rigid, incomp}} = \mu(1 + 5c/2). \tag{4.18}$$

The formula was first derived by Einstein, in 1906, in the mathematically identical context of incompressible viscous flow.

A matrix containing spherical voids: The case $\kappa' = \mu' = 0$ corresponds to a matrix containing voids; the formulae (4.16) reduce to

$$\kappa^{\text{eff, voids}} = \kappa\left(1 - c\frac{(3\kappa + 4\mu)}{4\mu}\right), \quad \mu^{\text{eff, voids}} = \mu\left(1 - c\frac{5(3\kappa + 4\mu)}{(9\kappa + 8\mu)}\right). \tag{4.19}$$

4.4 Approximation for a general composite

If the composite is non-dilute, the problem of estimating its properties becomes very complicated unless resort is made to some simplifying approximation. One such approximation is considered now.

Suppose first that the composite is an aggregate of different materials, without any clear matrix-inclusion structure. A simple approximation for the concentration tensor of material of type r is obtained by assuming that each piece of material of type r is spherical, and that the strain within it can be estimated by embedding a single sphere, with tensor of moduli $\boldsymbol{C}^r$, in a uniform matrix with tensor of moduli $\boldsymbol{C}^0$ (to be chosen judiciously later!). This assumption means that the effect of the surrounding heterogeneous material on this piece of type r is approximately the same as the effect of surrounding it with material with tensor of moduli $\boldsymbol{C}^0$. The mean strain in the composite is $\overline{\boldsymbol{\varepsilon}}$. This, however, is "screened" by the heterogeneities, and so we assume that the strain in material r is reproduced by subjecting the uniform matrix with moduli $\boldsymbol{C}^0$ to a remote strain $\boldsymbol{\varepsilon}^0$ (also to be chosen later). It follows now, in an obvious notation, that the mean strain $\boldsymbol{\varepsilon}^r$ in material r is approximated as

$$\boldsymbol{\varepsilon}^r = \boldsymbol{A}^{r,0}\boldsymbol{\varepsilon}^0. \tag{4.20}$$

The requirement (4.3) gives

$$\boldsymbol{\varepsilon}^0 = \left(\sum c_r \boldsymbol{A}^{r,0}\right)^{-1}\overline{\boldsymbol{\varepsilon}} =: \langle \boldsymbol{A}^0\rangle^{-1}\overline{\boldsymbol{\varepsilon}}, \tag{4.21}$$

and hence

$$\boldsymbol{A}^r \approx \boldsymbol{A}^{r,0}\langle \boldsymbol{A}^0\rangle^{-1}. \tag{4.22}$$

Formula (4.7) now provides the approximation

$$\boldsymbol{C}^{\text{eff}} \approx \sum_{r=1}^{n} c_r \boldsymbol{C}^r \boldsymbol{A}^{r,0}\left(\sum_{s=1}^{n} c_s \boldsymbol{A}^{s,0}\right)^{-1}. \tag{4.23}$$

This approximation can be used, formally, even when the composite has a clearly-defined matrix. It is not at all obvious that it is appropriate to treat the actual matrix as though it is a spherical inclusion in material with tensor of moduli $\boldsymbol{C}^0$, but at least this assumption provides a formula, and later theoretical developments will show why it works quite well!

Explicit formulae

For an isotropic composite, with isotropic phases, the formula for $\boldsymbol{A}^{r,0}$ becomes, using the notation already established for isotropic tensors,

$$\boldsymbol{A}^{r,0} = (3\kappa_A^{r,0}, 2\mu_A^{r,0}) \tag{4.24}$$

where, from (4.15),

$$3\kappa_A^{r,0} = \frac{3\kappa^0 + 4\mu^0}{3\kappa^r + 4\mu^0}, \quad 2\mu_A^{r,0} = \frac{5\mu^0(3\kappa^0 + 4\mu^0)}{\mu^0(9\kappa^0 + 8\mu^0) + 6\mu^r(\kappa^0 + 2\mu^0)}. \tag{4.25}$$

Hence, (4.23) gives

$$\kappa^{\text{eff}} \approx \frac{\sum_{r=1}^n c_r\kappa^r/(3\kappa^r + 4\mu^0)}{\sum_{s=1}^n c_s/(3\kappa^s + 4\mu^0)}, \quad \mu^{\text{eff}} \approx \frac{\sum_{r=1}^n c_r\mu^r/[\mu^0(9\kappa^0 + 8\mu^0) + 6\mu^r(\kappa^0 + 2\mu^0)]}{\sum_{s=1}^n c_s/[\mu^0(9\kappa^0 + 8\mu^0) + 6\mu^s(\kappa^0 + 2\mu^0)]}. \tag{4.26}$$

It is noted, for future use, that the expressions (4.26) can be rearranged to give

$$\begin{aligned} \frac{1}{3\kappa^{\text{eff}} + 4\mu^0} &= \sum_{r=1}^n \frac{c_r}{3\kappa_r + 4\mu^0}, \\ \frac{1}{\mu^0(9\kappa^0 + 8\mu^0) + 6\mu^{\text{eff}}(\kappa^0 + 2\mu^0)} &= \sum_{r=1}^n \frac{c_r}{\mu^0(9\kappa^0 + 8\mu^0) + 6\mu^r(\kappa^0 + 2\mu^0)}. \end{aligned} \tag{4.27}$$

Simple special cases

Suppose we allow the moduli $\boldsymbol{C}^0$ to tend to infinity, corresponding to rigid material. Physically, it follows that the strain in the inclusion must be $\boldsymbol{\varepsilon}^0$, for each r, and hence that

$$\boldsymbol{A}^r \approx \boldsymbol{I}. \tag{4.28}$$

Correspondingly, from (4.23),

$$\boldsymbol{C}^{\text{eff}} \approx \overline{\boldsymbol{C}}, \tag{4.29}$$

the Voigt approximation. This is confirmed by taking the limit directly in equations (4.26).

Now, conversely, let $\boldsymbol{C}^0 \to 0$. Equation (4.26) gives

$$\boldsymbol{C}^{\text{eff}} \approx (\overline{\boldsymbol{S}})^{-1}, \tag{4.30}$$

the Reuss approximation. An interpretation of this result is as follows. $\boldsymbol{A}^{r,0} \to 0$ and the formula (4.22) becomes degenerate. What happens in the limit is that $\boldsymbol{\varepsilon}^0 \to \infty$ in such a

way that the corresponding stress $\boldsymbol{\sigma}^0$ remains finite. The stress in the inclusion also takes the value $\boldsymbol{\sigma}^0$ and hence the strain becomes $\boldsymbol{S}^r\boldsymbol{\sigma}^0$. Equation (4.3) now requires that

$$\boldsymbol{\sigma}^0 = (\overline{\boldsymbol{S}})^{-1}\overline{\boldsymbol{\varepsilon}}. \tag{4.31}$$

Since the stress in each piece of material is estimated to take the same constant value $\boldsymbol{\sigma}^0$, this represents the estimated mean stress in the composite, and the Reuss approximation follows.

Evidently, choosing intermediate values for $\boldsymbol{C}^0$ will yield estimates for $\boldsymbol{C}^{\text{eff}}$ between the Reuss and Voigt approximations.

4.5 Self-consistent approximation

Now here is a proposal for a particular choice of $\boldsymbol{C}^0$. Suppose that the best choice for $\boldsymbol{C}^0$ is the actual $\boldsymbol{C}^{\text{eff}}$. The latter of course is not known, but an approximation is given in terms of $\boldsymbol{C}^0$ by equation (4.23). The so-called "self-consistent" approximation is given by taking $\boldsymbol{C}^{\text{eff}} = \boldsymbol{C}^0$, where $\boldsymbol{C}^0$ is the solution of the equation

$$\boldsymbol{C}^0 = \sum_{r=1}^{n} c_r \boldsymbol{C}^r \boldsymbol{A}^{r,0} \left(\sum_{s=1}^{n} c_s \boldsymbol{A}^{s,0} \right)^{-1}. \tag{4.32}$$

Explicit equations, for a composite with isotropic phases, are provided by (4.26) or (4.27), with κ^{eff}, μ^{eff} replaced by κ^0, μ^0.

It is worth noting that the form of the equations obtained from (4.27) gives, immediately, that

$$\sum_{s=1}^{n} c_s \boldsymbol{A}^{s,0} = \boldsymbol{I}, \tag{4.33}$$

and hence that

$$\boldsymbol{A}^r = \boldsymbol{A}^{r,0}, \tag{4.34}$$

when $\boldsymbol{C}^0$ is chosen self-consistently. This observation provides a different perspective for the self-consistent approximation (though not for the more general formulae (4.26)): estimate $\boldsymbol{A}^r$ by embedding an inclusion of type r directly into a matrix whose elastic constant tensor $\boldsymbol{C}^0$ is chosen to be $\boldsymbol{C}^{\text{eff}}$. The matrix is subjected to remote strain $\overline{\boldsymbol{\varepsilon}}$.

It was remarked in subsection 4.1 that, for a matrix-inclusion composite (the matrix being material n), it was natural to estimate the tensors $\boldsymbol{A}^r$ for $1 \leq r \leq n-1$ by embedding an inclusion in a matrix, but then to obtain $\boldsymbol{A}^n$ by use of the identity (4.4). Then, the expression for $\boldsymbol{C}^{\text{eff}}$ takes the form (4.8). The observation that (4.33) holds, for the self-consistent choice

of $\boldsymbol{C}^0$, shows that use of this second prescription, in conjunction with the self-consistent approximation, gives exactly the same result as that obtained from estimating $\boldsymbol{A}^n$ by embedding an inclusion of matrix material in material with moduli $\boldsymbol{C}^0$! These comments have been made in relation to the explicit formulae (4.26), (4.27). They do, in fact, have validity also for anisotropic media, for which formulae will be presented later.

In the case of a 2-component composite, in which an isotropic matrix with elastic constant tensor $\boldsymbol{C}^2$ has embedded in it a single population of isotropic spherical inclusions, with elastic constant tensor $\boldsymbol{C}^1$, at volume fraction c_1, the self-consistent approximation to $\boldsymbol{C}^{\rm eff}$ is defined by the equations

$$\begin{aligned}\kappa^0 &= \kappa^2 + c_1\frac{(\kappa^1-\kappa^2)(3\kappa^0+4\mu^0)}{3\kappa^1+4\mu^0},\\ \mu^0 &= \mu^2 + c_1\frac{5(\mu^1-\mu^2)\mu^0(3\kappa^0+4\mu^0)}{\mu^0(9\kappa^0+8\mu^0)+6\mu^1(\kappa^0+2\mu^0)}.\end{aligned} \tag{4.35}$$

When $c_1 \ll 1$, the development of the self-consistent solution in a power series in c_1, truncated at first order in c_1, agrees with the dilute approximation (4.16).

In the special case of rigid inclusions ($\boldsymbol{C}^1 \to \infty$), the formulae (4.35) reduce to

$$\begin{aligned}\kappa^{\rm eff} &= \kappa^2 + c_1(\kappa^0+4\mu^0/3),\\ \mu^{\rm eff} &= \mu^2 + c_1\frac{5\mu^0(3\kappa^0+4\mu^0)}{6(\kappa^0+2\mu^0)}.\end{aligned} \tag{4.36}$$

In the special case of cavities ($\boldsymbol{C}^1 = 0$), they reduce to

$$\begin{aligned}\kappa^{\rm eff} &= \kappa^2\left(1 - c_1\frac{(3\kappa_0+4\mu^0)}{4\mu^0}\right),\\ \mu^{\rm eff} &= \mu^2\left(1 - c_1\frac{5(3\kappa^0+4\mu^0)}{(9\kappa^0+8\mu^0)}\right).\end{aligned} \tag{4.37}$$

5 Thermoelastic Response of a Composite

5.1 General relations

The Helmholtz free energy density F of a thermoelastic solid is a function of the strain, $\boldsymbol{\varepsilon}$, and the temperature, θ. If strain and temperature are measured relative to a stress-free state, then, to lowest order, the free energy density becomes a quadratic function of $\boldsymbol{\varepsilon}$ and θ:

$$F(\boldsymbol{\varepsilon},\theta) = \tfrac{1}{2}\boldsymbol{\varepsilon}\boldsymbol{C}\boldsymbol{\varepsilon} - (\boldsymbol{\varepsilon}\boldsymbol{\beta})\theta - \tfrac{1}{2}f\theta^2 \equiv \tfrac{1}{2}\varepsilon_{ij}C_{ijkl}\varepsilon_{kl} - \varepsilon_{ij}\beta_{ij}\theta - \tfrac{1}{2}f\theta^2, \tag{5.1}$$

having set F to zero in the reference state. The constitutive equations associated with (5.1) are

$$\sigma_{ij} = \frac{\partial F}{\partial \varepsilon_{ij}} = C_{ijkl}\varepsilon_{kl} - \beta_{ij}\theta, \tag{5.2}$$

$$\eta = -\frac{\partial F}{\partial \theta} = \beta_{ij}\varepsilon_{ij} + f\theta. \tag{5.3}$$

Here, η is the entropy density. The elastic constant tensor $\boldsymbol{C}$ applies to isothermal deformations. With θ_R denoting the temperature in the reference state, $\theta_R f$ is the specific heat capacity $(\partial\eta/\partial\theta)$ at constant strain. The second-order tensor $\boldsymbol{\beta}$ is the thermal stress tensor, giving (minus) the change in stress associated with change in temperature, at constant strain. The thermal strain tensor, $\boldsymbol{\alpha}$, gives the change in strain with change of temperature, at constant stress. It follows by setting $\boldsymbol{\sigma} = 0$ in (5.2) and solving for $\boldsymbol{\varepsilon}$. Thus,

$$\boldsymbol{\alpha} = \boldsymbol{S}\boldsymbol{\beta}, \quad \boldsymbol{\beta} = \boldsymbol{C}\boldsymbol{\alpha}, \tag{5.4}$$

where $\boldsymbol{S}$ is the isothermal tensor of compliances.

For an n-phase composite, the constants take the values $\boldsymbol{C}^r$, $\boldsymbol{\beta}^r$, f^r in phase r. Effective constants for the composite are established by finding the mean values of stress and entropy, when the composite is subjected to conditions that would generate within it, uniform strain and temperature fields if it were homogeneous. Thus, it is subjected, at its boundary, to linear displacement or uniform traction conditions, and to a uniform temperature. Since the concern here is only for equilibrium conditions, uniform temperature on the boundary implies uniform temperature throughout the composite.

In analogy with the purely mechanical case, concentration tensors are now defined so that

$$\boldsymbol{\varepsilon}^r = \boldsymbol{A}^r\overline{\boldsymbol{\varepsilon}} + \boldsymbol{a}^r\theta. \tag{5.5}$$

These are subject to the restrictions

$$\sum_{r=1}^{n} c_r \boldsymbol{A}^r = \boldsymbol{I}, \quad \sum_{r=1}^{n} c_r \boldsymbol{a}^r = 0. \tag{5.6}$$

Now define F^{eff} to be the mean value over Ω of F. F^{eff} must be a quadratic function of the parameters $\overline{\boldsymbol{\varepsilon}}$, θ which define the boundary value problem from which it is constructed. Taking, as previously, Ω to have unit volume,

$$F^{\text{eff}} := \int_{\Omega} F(\boldsymbol{\varepsilon}, \theta) \mathrm{d}\boldsymbol{x}. \tag{5.7}$$

Now change the mean strain to $\overline{\boldsymbol{\varepsilon}} + \delta\overline{\boldsymbol{\varepsilon}}$ and the temperature to $\theta + \delta\theta$. The strain within the composite changes to $\boldsymbol{\varepsilon} + \delta\boldsymbol{\varepsilon}$. The change in F^{eff} is then, to first order,

$$\begin{aligned}
\delta F^{\text{eff}} &\equiv F^{\text{eff}}_{,\overline{\boldsymbol{\varepsilon}}} \delta\overline{\boldsymbol{\varepsilon}} + F^{\text{eff}}_{,\theta} \delta\theta = \int_{\Omega} \{F_{,\boldsymbol{\varepsilon}} \delta\boldsymbol{\varepsilon} + F_{,\theta} \delta\theta\} \mathrm{d}\boldsymbol{x} \\
&= \int_{\Omega} \{\boldsymbol{\sigma} \delta\boldsymbol{\varepsilon} - \eta \delta\theta\} \mathrm{d}\boldsymbol{x} \\
&= \int_{\Omega} \{\boldsymbol{\sigma} \delta\overline{\boldsymbol{\varepsilon}} - \eta \delta\theta\} \mathrm{d}\boldsymbol{x} \ \text{ (by the principle of virtual work)} \\
&= \overline{\boldsymbol{\sigma}} \delta\overline{\boldsymbol{\varepsilon}} - \overline{\eta} \delta\theta.
\end{aligned} \tag{5.8}$$

Hence,

$$\overline{\boldsymbol{\sigma}} \equiv \boldsymbol{C}^{\text{eff}} \overline{\boldsymbol{\varepsilon}} - \boldsymbol{\beta}^{\text{eff}} \theta = F^{\text{eff}}_{,\overline{\boldsymbol{\varepsilon}}}. \tag{5.9}$$

Also,

$$\overline{\eta} = -F^{\text{eff}}_{,\theta}. \tag{5.10}$$

It follows, using (5.5), that

$$\overline{\boldsymbol{\sigma}} = \sum_{r=1}^{n} c_r \{\boldsymbol{C}^r (\boldsymbol{A}^r \overline{\boldsymbol{\varepsilon}} + \boldsymbol{a}^r \theta) - \boldsymbol{\beta}^r \theta\}, \tag{5.11}$$

$$\overline{\eta} = \sum_{r=1}^{n} c_r \{\boldsymbol{\beta}^r (\boldsymbol{A}^r \overline{\boldsymbol{\varepsilon}} + \boldsymbol{a}^r \theta) + f^r \theta\}. \tag{5.12}$$

Therefore, considering also (5.9) and (5.10),

$$\boldsymbol{C}^{\text{eff}} = \sum_{r=1}^{n} c_r \boldsymbol{C}^r \boldsymbol{A}^r \tag{5.13}$$

is symmetric (as in the purely mechanical case), and

$$\boldsymbol{\beta}^{\rm eff} = \sum_{r=1}^{n} c_r\{\boldsymbol{\beta}^r - \boldsymbol{C}^r\boldsymbol{a}^r\} \equiv \sum_{r=1}^{n} c_r(\boldsymbol{A}^r)^T\boldsymbol{\beta}^r. \tag{5.14}$$

Here, the superscript T denotes transposition, in the sense that the $ijkl$ component of $(\boldsymbol{A}^r)^T$ is A^r_{klij}. [This is about the only point at which the symbolic notation becomes potentially unclear: if in doubt, revert to suffix notation!] The equivalence recorded here provides an additional relation between the concentration tensors. It demonstrates, remarkably, that the effective thermal stress tensor can be deduced, once the purely mechanical strain concentration tensors have been found: there is no need to solve any thermoelastic problem, for this purpose.

The fact that $F^{\rm eff}$ is a homogeneous function of degree 2 in its arguments gives the relation

$$F^{\rm eff}(\overline{\boldsymbol{\varepsilon}}, \theta) = \tfrac{1}{2}(\overline{\boldsymbol{\sigma}}\,\overline{\boldsymbol{\varepsilon}} - \overline{\eta}\theta). \tag{5.15}$$

Thus,

$$F^{\rm eff}(\overline{\boldsymbol{\varepsilon}}, \theta) = \tfrac{1}{2}\overline{\boldsymbol{\varepsilon}}\boldsymbol{C}^{\rm eff}\overline{\boldsymbol{\varepsilon}} - (\boldsymbol{\beta}^{\rm eff}\overline{\boldsymbol{\varepsilon}})\theta - \tfrac{1}{2}f^{\rm eff}\theta^2, \tag{5.16}$$

where $f^{\rm eff}$ is given by

$$f^{\rm eff} = \sum_{r=1}^{n} c_r(f^r + \boldsymbol{\beta}^r\boldsymbol{a}^r). \tag{5.17}$$

If the composite consists of a matrix (phase n say) containing inclusions, elimination of $\boldsymbol{A}^n$ and $\boldsymbol{a}^n$ using equations (5.6) gives

$$\begin{aligned}
\boldsymbol{C}^{\rm eff} &= \boldsymbol{C}^n + \sum_{r=1}^{n-1} c_r(\boldsymbol{C}^r - \boldsymbol{C}^n)\boldsymbol{A}^r, \\
\boldsymbol{\beta}^{\rm eff} &= \boldsymbol{\beta}^n + \sum_{r=1}^{n-1} c_r(\boldsymbol{A}^r)^T(\boldsymbol{\beta}^r - \boldsymbol{\beta}^n), \\
f^{\rm eff} &= \overline{f} + \sum_{r=1}^{n-1} c_r(\boldsymbol{\beta}^r - \boldsymbol{\beta}^n)\boldsymbol{a}^r.
\end{aligned} \tag{5.18}$$

The alternative form,

$$\begin{aligned}
\boldsymbol{C}^{\rm eff} &= \overline{\boldsymbol{C}} + \sum_{r=1}^{n-1} c_r(\boldsymbol{C}^r - \boldsymbol{C}^n)(\boldsymbol{A}^r - \boldsymbol{I}), \\
\boldsymbol{\beta}^{\rm eff} &= \overline{\boldsymbol{\beta}} + \sum_{r=1}^{n-1} c_r(\boldsymbol{A}^r - \boldsymbol{I})^T(\boldsymbol{\beta}^r - \boldsymbol{\beta}^n),
\end{aligned}$$

$$f^{\rm eff} \;=\; \overline{f} + \sum_{r=1}^{n-1} c_r(\boldsymbol{\beta}^r - \boldsymbol{\beta}^n)\boldsymbol{a}^r. \tag{5.19}$$

will be convenient in what follows.

5.2 The Levin relations

Further remarkable relations (due to V M Levin, about 1967) follow if the composite has only two phases. When combined with (5.6), the equivalence given in (5.14) provides

$$\boldsymbol{a}^1 = -(\boldsymbol{C}^1 - \boldsymbol{C}^2)^{-1}(\boldsymbol{A}^1 - \boldsymbol{I})^T(\boldsymbol{\beta}^1 - \boldsymbol{\beta}^2). \tag{5.20}$$

Also, from the first of equations (5.19), with $n = 2$,

$$c_1(\boldsymbol{A}^1 - \boldsymbol{I})^T = (\boldsymbol{C}^{\rm eff} - \overline{\boldsymbol{C}})(\boldsymbol{C}^1 - \boldsymbol{C}^2)^{-1}, \tag{5.21}$$

and hence

$$c_1\boldsymbol{a}^1 = -(\boldsymbol{C}^1 - \boldsymbol{C}^2)^{-1}(\boldsymbol{C}^{\rm eff} - \overline{\boldsymbol{C}})(\boldsymbol{C}^1 - \boldsymbol{C}^2)^{-1}(\boldsymbol{\beta}^1 - \boldsymbol{\beta}^2). \tag{5.22}$$

It follows now, from the second and third of equations (5.19) that

$$\begin{aligned} \boldsymbol{\beta}^{\rm eff} &= \overline{\boldsymbol{\beta}} + (\boldsymbol{C}^{\rm eff} - \overline{\boldsymbol{C}})(\boldsymbol{C}^1 - \boldsymbol{C}^2)^{-1}(\boldsymbol{\beta}^1 - \boldsymbol{\beta}^2), \\ f^{\rm eff} &= \overline{f} - (\boldsymbol{\beta}^1 - \boldsymbol{\beta}^2)(\boldsymbol{C}^1 - \boldsymbol{C}^2)^{-1}(\boldsymbol{C}^{\rm eff} - \overline{\boldsymbol{C}})(\boldsymbol{C}^1 - \boldsymbol{C}^2)^{-1}(\boldsymbol{\beta}^1 - \boldsymbol{\beta}^2). \end{aligned} \tag{5.23}$$

These are the Levin relations: they give the effective thermoelastic properties explicitly, in terms of the purely mechanical effective modulus tensor.

5.3 The isotropic spherical inclusion

Just for the record, it is easy to solve the problem of the thermal expansion of an isotropic spherical inclusion in an isotropic matrix, because the problem displays spherical symmetry. If the bulk and shear moduli are, respectively, κ, μ for the matrix and κ', μ' for the inclusion, and if the thermal stress tensors have components $\beta\delta_{ij}$ for the matrix and $\beta'\delta_{ij}$ for the inclusion, then the tensor $\boldsymbol{a}'$ for the inclusion has components $a'\delta_{ij}$, where

$$a' = \frac{(\beta' - \beta)}{(3\kappa' + 4\mu)}. \tag{5.24}$$

This also follows from the Levin relations, applied to a composite consisting of a dilute suspension of spheres in a matrix.

6 A General Formulation for Heterogeneous Media

6.1 A fundamental integral equation

This sub-section considers the solution of problems for a medium which occupies a domain Ω, with general tensor of elastic moduli $\boldsymbol{C}(\boldsymbol{x})$. To avoid unnecessary complication, only the displacement boundary value problem will be addressed. Thus, the stress, strain and displacement must conform to the equations

$$\begin{aligned} \operatorname{div}\boldsymbol{\sigma} + \boldsymbol{f} = 0, \quad \text{or} \quad \sigma_{ij,j} + f_i = 0, \;\; \boldsymbol{x} \in \Omega, \\ \boldsymbol{\sigma} = \boldsymbol{C}\boldsymbol{\varepsilon}, \;\; \boldsymbol{\varepsilon} = \tfrac{1}{2}[\nabla\boldsymbol{u} + (\nabla\boldsymbol{u})^T], \quad \text{or} \quad \sigma_{ij} = C_{ijkl}\varepsilon_{kl}, \;\; \varepsilon_{ij} = \tfrac{1}{2}[u_{i,j} + u_{j,i}], \;\; \boldsymbol{x} \in \Omega, \\ \boldsymbol{u}(\boldsymbol{x}) = \boldsymbol{u}^0(\boldsymbol{x}), \;\; \boldsymbol{x} \in \partial\Omega. \end{aligned} \tag{6.1}$$

Introduce a "comparison medium", with elastic constant tensor $\boldsymbol{C}^0$: in general, this could vary with position $\boldsymbol{x}$ but the formulation will be most useful (at least for the present) if the comparison medium is uniform. A "stress polarisation tensor" $\boldsymbol{\tau}(\boldsymbol{x})$ is defined so that the stress $\boldsymbol{\sigma}$ and the strain $\boldsymbol{\varepsilon}$ in the medium satisfy

$$\boldsymbol{\sigma} = \boldsymbol{C}\boldsymbol{\varepsilon} \equiv \boldsymbol{C}^0\boldsymbol{\varepsilon} + \boldsymbol{\tau}. \tag{6.2}$$

Thus,

$$\boldsymbol{\tau} = (\boldsymbol{C} - \boldsymbol{C}^0)\boldsymbol{\varepsilon}. \tag{6.3}$$

Substituting (6.2) into the equilibrium equations, and also expressing strain in terms of displacement, gives the system of equations

$$(C^0_{ijkl}u_{k,l})_{,j} + \tau_{ij,j} + f_i = 0, \;\; \boldsymbol{x} \in \Omega. \tag{6.4}$$

Thus, the displacement and strain fields in the actual medium are generated in the comparison medium, if this is subjected to "body force" $\boldsymbol{f} + \operatorname{div}\boldsymbol{\tau}$. The stress in the actual body is given by (6.2).

Since equations (6.4) are linear, their solution can be broken down as follows. First, solve the equations, with the real body-force $\boldsymbol{f}$ and the boundary condition but without the additional body-force $\operatorname{div}\boldsymbol{\tau}$: call this solution $\boldsymbol{u}^0$ (now defined over all of Ω). Now, regarding $\boldsymbol{\tau}$ for the moment as known, solve the equations

$$(C^0_{ijkl}u_{k,l})_{,j} + \tau_{ij,j} = 0, \;\; \boldsymbol{x} \in \Omega, \tag{6.5}$$

subject to the boundary condition $\boldsymbol{u} = 0$ on $\partial\Omega$: call this solution $\boldsymbol{u}^1$. The required solution is then

$$\boldsymbol{u} = \boldsymbol{u}^0 + \boldsymbol{u}^1. \tag{6.6}$$

The field $\boldsymbol{u}^1$ is hard to find explicitly, but it can be represented in terms of Green's function $\boldsymbol{G}$ for the comparison body. This has to satisfy the equations

$$\begin{aligned}[C^0_{ijkl}(\boldsymbol{x})G_{kp,l}(\boldsymbol{x},\boldsymbol{x}')]_{,j} + \delta_{ip}\delta(\boldsymbol{x}-\boldsymbol{x}') &= 0, \quad \boldsymbol{x}\in\Omega,\\ G_{ip}(\boldsymbol{x},\boldsymbol{x}') &= 0, \quad \boldsymbol{x}\in\partial\Omega.\end{aligned} \tag{6.7}$$

Then, by superposition,

$$u^1_i(\boldsymbol{x}) = \int_\Omega G_{ip}(\boldsymbol{x},\boldsymbol{x}')\tau_{pq,q}(\boldsymbol{x}')\,\mathrm{d}\boldsymbol{x}'. \tag{6.8}$$

Hence, by integrating by parts and combining with (6.6),

$$u_i(\boldsymbol{x}) = u^0_i(\boldsymbol{x}) - \int_\Omega \frac{\partial G_{ip}(\boldsymbol{x},\boldsymbol{x}')}{\partial x'_q}\tau_{pq}(\boldsymbol{x}')\,\mathrm{d}\boldsymbol{x}'. \tag{6.9}$$

Differentiating and symmetrising now gives

$$\varepsilon_{ij}(\boldsymbol{x}) = \varepsilon^0_{ij}(\boldsymbol{x}) - \int_\Omega \Gamma_{ijpq}(\boldsymbol{x},\boldsymbol{x}')\tau_{pq}(\boldsymbol{x}')\,\mathrm{d}\boldsymbol{x}', \tag{6.10}$$

where

$$\Gamma_{ijpq}(\boldsymbol{x},\boldsymbol{x}') = \left.\frac{\partial^2 G_{ip}(\boldsymbol{x},\boldsymbol{x}')}{\partial x_j \partial x'_q}\right|_{(ij),(pq)}, \tag{6.11}$$

the bracketed subscripts implying symmetrisation[8]. The representation (6.10) will be written in symbolic notation

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}^0 - \boldsymbol{\Gamma}\boldsymbol{\tau}. \tag{6.12}$$

An integral equation for $\boldsymbol{\tau}$ now follows by noting, from (6.3), that $\boldsymbol{\varepsilon} = (\boldsymbol{C}-\boldsymbol{C}^0)^{-1}\boldsymbol{\tau}$. Thus,

$$(\boldsymbol{C}-\boldsymbol{C}^0)^{-1}\boldsymbol{\tau} + \boldsymbol{\Gamma}\boldsymbol{\tau} = \boldsymbol{\varepsilon}^0. \tag{6.13}$$

6.2 Green's function for an infinite body

Further progress requires knowledge of the Green's function. This can be found, fairly easily, if the body is infinite and uniform. In this case, translation invariance shows that $\boldsymbol{G}$ depends on $\boldsymbol{x}$, $\boldsymbol{x}'$ only in the combination $(\boldsymbol{x}-\boldsymbol{x}')$, and so $\boldsymbol{x}'$ can be set to zero, without loss of generality.

[8]The integral in (6.10) will be singular, and must be interpreted in the sense of generalised functions. An alternative procedure, which follows the derivation, is first to evaluate the integral in (6.9) and then to perform the differentiation with respect to $\boldsymbol{x}$.

Note first the basic results

$$\begin{aligned} \Delta\left(\frac{1}{4\pi r}\right) + \delta(\boldsymbol{x}) &= 0 \ \text{(in 3 dimensions)}, \\ \Delta\left(\frac{\log(1/r)}{2\pi}\right) + \delta(\boldsymbol{x}) &= 0 \ \text{(in 2 dimensions)}. \end{aligned} \tag{6.14}$$

where $r = |\boldsymbol{x}|$ and Δ denotes the Laplacian. Note also that

$$\Delta r = 2/r \ \text{ (in 3 dimensions)}, \ \ \Delta r^2 \log(1/r) = 4(\log(1/r) - 1) \ \text{ (in 2 dimensions)}. \tag{6.15}$$

Isotropic medium

For an isotropic body, with Lamé moduli λ, μ which are constants, Green's function must satisfy the equations

$$(\lambda + \mu) G_{jp,ji} + \mu G_{ip,jj} + \delta_{ip}\delta(\boldsymbol{x}) = 0. \tag{6.16}$$

Motivated by the results (6.14) for the Laplacian operator, try

$$G_{ip}(\boldsymbol{x}) = \begin{cases} \frac{1}{\mu}\left[\frac{\delta_{ip}}{4\pi r} + \alpha r_{,ip}\right] & \text{(in 3 dimensions)} \\ \frac{1}{\mu}\left[\frac{\log(1/r)}{2\pi} + \alpha\{r^2 \log(1/r)\}_{,ip}\right] & \text{(in 2 dimensions).} \end{cases} \tag{6.17}$$

Substituting into (6.16) shows that, in either two or three dimensions,

$$\alpha = -\frac{\lambda + \mu}{8\pi(\lambda + 2\mu)}. \tag{6.18}$$

Generally-anisotropic medium

Now consider a uniform infinite medium with generally-anisotropic tensor of elastic moduli $\boldsymbol{C}$. Green's function now satisfies the equation

$$C_{ijkl} G_{kp,jl} + \delta_{ip}\delta(\boldsymbol{x}) = 0. \tag{6.19}$$

An attractive representation can be given for $\boldsymbol{G}(\boldsymbol{x})$ in three dimensions, by noting the elementary relation

$$\int_{|\boldsymbol{\xi}|=1} \delta(\boldsymbol{\xi}.\boldsymbol{x}) \, \mathrm{d}S = 2\pi/r. \tag{6.20}$$

The integral here is over the surface of the unit sphere $|\boldsymbol{\xi}| = 1$. One way to see this result is to transform to polar coordinates (ρ, θ, ϕ) (where $\rho \equiv |\boldsymbol{\xi}| = 1$ on the surface of the sphere), with the axis $\theta = 0$ aligned with $\boldsymbol{x}$. The integral becomes

$$\int_0^{2\pi} \mathrm{d}\phi \int_0^{\pi} \sin\theta \, \mathrm{d}\theta \, \delta(r\cos\theta),$$

which can be evaluated by elementary means, after employing the further variable transformation $s = r\cos\theta$.

Notice now that, for any function f of the scalar variable $s = \boldsymbol{\xi}.\boldsymbol{x}$, $\partial f(\boldsymbol{\xi}.\boldsymbol{x})/\partial x_i = \xi_i f'(\boldsymbol{\xi}.\boldsymbol{x})$. Therefore, by applying the Laplacian operator to both sides of (6.20), it is obtained that

$$\delta(\boldsymbol{x}) = \frac{-1}{8\pi^2} \int_{|\boldsymbol{\xi}|=1} \delta''(\boldsymbol{\xi}.\boldsymbol{x}) \, \mathrm{d}S, \tag{6.21}$$

since $\xi_j \xi_j = 1$ on the surface of the unit sphere. The solution of (6.19) can now be built up by superposition. First, solve the problem

$$C_{ijkl} G^{\boldsymbol{\xi}}_{kp,jl} + \delta_{ip} \delta''(\boldsymbol{\xi}.\boldsymbol{x}) = 0. \tag{6.22}$$

Evidently, it is only necessary to take $\boldsymbol{G}^{\boldsymbol{\xi}}$ to be a function of $s = \boldsymbol{\xi}.\boldsymbol{x}$. Then, (6.22) reduces to

$$\boldsymbol{K}(\boldsymbol{\xi}) \boldsymbol{G}^{\boldsymbol{\xi}\prime\prime}(s) + \boldsymbol{I}\delta''(s) = 0, \tag{6.23}$$

where

$$K_{ik}(\boldsymbol{\xi}) = C_{ijkl} \xi_j \xi_l. \tag{6.24}$$

Hence, we may take

$$\boldsymbol{G}^{\boldsymbol{\xi}}(s) = -[\boldsymbol{K}(\boldsymbol{\xi})]^{-1} \delta(s). \tag{6.25}$$

It follows now, by the superposition shown in equation (6.21), that

$$\boldsymbol{G}(\boldsymbol{x}) = \frac{1}{8\pi^2} \int_{|\boldsymbol{\xi}|=1} [\boldsymbol{K}(\boldsymbol{\xi})]^{-1} \delta(\boldsymbol{\xi}.\boldsymbol{x}) \, \mathrm{d}S. \tag{6.26}$$

The kernel of the operator $\boldsymbol{\Gamma}$ is obtained by performing the differentiations indicated in (6.11). Thus,

$$\boldsymbol{\Gamma} = \frac{-1}{8\pi^2} \int_{|\boldsymbol{\xi}|=1} \tilde{\boldsymbol{\Gamma}}(\xi) \delta''(\boldsymbol{\xi}.\boldsymbol{x}) \, \mathrm{d}S, \tag{6.27}$$

where

$$\tilde{\Gamma}_{ijpq} = \xi_{(i} \{[\boldsymbol{K}(\boldsymbol{\xi})]^{-1}\}_{j)(p} \xi_{q)} \tag{6.28}$$

(brackets on subscripts again implying symmetrisation).

There are corresponding expressions for Green's function and the associated Γ-operator in two dimensions but they are more complicated and are not considered here.

6.3 Inclusion problems

The preceding formulation finds immediate use for solving the problem of an isolated inclusion occupying a finite region V, embedded in an infinite uniform matrix which is in a state of uniform stress and strain, remote from V. Let the matrix have elastic constant tensor $\boldsymbol{C}^0$, while the inclusion has elastic constant tensor $\boldsymbol{C}$. Then since $\boldsymbol{C} = \boldsymbol{C}^0$ in the matrix, equation (6.3) shows that $\boldsymbol{\tau}$ is non-zero only over the region V; correspondingly, equation (6.13) applies only for $\boldsymbol{x} \in V$. Since the strain is uniform far from V it would be identically uniform if the medium contained no inclusion, and hence $\boldsymbol{\varepsilon}^0$ is constant, taking the value of the strain remote from V.

Ellipsoidal inclusion

In the special case that V is an ellipsoid, the solution of this problem has the remarkable property that the stress and strain, and hence also the polarisation $\boldsymbol{\tau}$, are constant over the inclusion V. This can be verified by the semi-inverse method of assuming that $\boldsymbol{\tau}$ is constant, and confirming that the equation is satisfied if the constant is chosen appropriately. To simplify the calculation, suppose first that the inclusion in fact occupies the sphere $V = \{\boldsymbol{x} : |\boldsymbol{x}| \leq a\}$. Since $\boldsymbol{\tau}$ is to be taken constant, (6.13) requires the evaluation of the integral

$$\boldsymbol{P} := \frac{-1}{8\pi^2} \int_{|\boldsymbol{\xi}|=1} \mathrm{d}S \tilde{\boldsymbol{\Gamma}}(\boldsymbol{\xi}) \int_{\boldsymbol{x}' \in V} \mathrm{d}\boldsymbol{x}' \delta''(\boldsymbol{\xi}.(\boldsymbol{x} - \boldsymbol{x}')), \tag{6.29}$$

when $\boldsymbol{x} \in V$. Thus, it is necessary to evaluate the integral

$$J(p) := \int_{\boldsymbol{x}' \in V} \delta(\boldsymbol{\xi}.\boldsymbol{x}' - p) \mathrm{d}\boldsymbol{x}', \tag{6.30}$$

and then to differentiate the result twice with respect to p, and finally setting $p = \boldsymbol{\xi}.\boldsymbol{x}$. Now when $\boldsymbol{x} \in V$, we have $|p| < a$, since $\boldsymbol{\xi}$ is a unit vector. The value of $J(p)$ is the area of the disc defined by the intersection of the plane $\boldsymbol{\xi}.\boldsymbol{x}' = p$ with the sphere V. Thus, since $|p| < a$,

$$J(p) = \pi(a^2 - p^2) \tag{6.31}$$

and consequently

$$J''(p) = -2\pi \tag{6.32}$$

for all p such that $|p| < a$. Hence,

$$\boldsymbol{P} = \frac{1}{4\pi} \int_{|\boldsymbol{\xi}|=1} \tilde{\boldsymbol{\Gamma}}(\boldsymbol{\xi}) \, \mathrm{d}S, \tag{6.33}$$

which is constant, as asserted. Substituting back into equation (6.13), therefore,

$$[(\boldsymbol{C}-\boldsymbol{C}^0)^{-1}+\boldsymbol{P}]\boldsymbol{\tau}=\boldsymbol{\varepsilon}^0. \tag{6.34}$$

Thus, within the inclusion,

$$\begin{aligned}\boldsymbol{\tau} &= [(\boldsymbol{C}-\boldsymbol{C}^0)^{-1}+\boldsymbol{P}]^{-1}\boldsymbol{\varepsilon}^0, &(6.35)\\ \boldsymbol{\varepsilon} &= (\boldsymbol{C}-\boldsymbol{C}^0)^{-1}\boldsymbol{\tau}=[\boldsymbol{I}+\boldsymbol{P}(\boldsymbol{C}-\boldsymbol{C}^0)]^{-1}\boldsymbol{\varepsilon}^0, &(6.36)\\ \boldsymbol{\sigma} &= \boldsymbol{C}\boldsymbol{\varepsilon}=\boldsymbol{C}[\boldsymbol{I}+\boldsymbol{P}(\boldsymbol{C}-\boldsymbol{C}^0)]^{-1}\boldsymbol{\varepsilon}^0, &(6.37)\end{aligned}$$

all constants, as asserted.

The corresponding result when V is the ellipsoid $V=\{\boldsymbol{x}:\boldsymbol{x}^T(A^TA)^{-1}\boldsymbol{x}<1\}$ can be deduced from the same working, by first defining $\boldsymbol{y}=A^{-T}\boldsymbol{x}$ and $\boldsymbol{\zeta}=A\boldsymbol{\xi}$, so that $|\boldsymbol{y}|<1$ when $\boldsymbol{x}\in V$, and $\boldsymbol{\xi}.\boldsymbol{x}=\boldsymbol{\zeta}.\boldsymbol{y}$. The only differences are that $\boldsymbol{\zeta}$ is not a unit vector: $|\boldsymbol{\zeta}|=(\boldsymbol{\xi}^TA^TA\boldsymbol{\xi})^{1/2}$, and $\mathrm{d}\boldsymbol{x}'=\det(A)\mathrm{d}\boldsymbol{y}$. Following the calculation through gives

$$\boldsymbol{P}=\frac{\det(A)}{4\pi}\int_{|\boldsymbol{\xi}|=1}\frac{\tilde{\boldsymbol{\Gamma}}(\boldsymbol{\xi})}{(\boldsymbol{\xi}^TA^TA\,\boldsymbol{\xi})^{3/2}}\,\mathrm{d}S. \tag{6.38}$$

6.4 Implications for a dilute suspension of ellipsoids

It follows immediately from the second of equations (6.37) that the strain concentration tensor corresponding to a dilute suspension of ellipsoidal particles, at volume fraction c, is

$$\boldsymbol{A}=[\boldsymbol{I}+\boldsymbol{P}(\boldsymbol{C}-\boldsymbol{C}^0)]^{-1}. \tag{6.39}$$

Then, by suitable specialisation of (4.8), it is obtained that

$$\boldsymbol{C}^{\mathrm{eff}}=\boldsymbol{C}^0+c(\boldsymbol{C}-\boldsymbol{C}^0)[\boldsymbol{I}+\boldsymbol{P}(\boldsymbol{C}-\boldsymbol{C}^0)]^{-1}=\boldsymbol{C}^0+c[(\boldsymbol{C}-\boldsymbol{C}^0)^{-1}+\boldsymbol{P}]^{-1}. \tag{6.40}$$

(The second form also follows by averaging equation (6.2).) Note that this formula applies, for any ellipsoidal inclusion, of arbitrary anisotropy, in an anisotropic matrix. It generalises immediately to the case of a suspension of several different types of ellipsoidal inclusion (different shapes, or orientations, or elastic constants, or all of these), subject to the dilute approximation.

6.5 Isotropic matrix

For an isotropic medium with Lamé moduli λ, μ, the matrix $\boldsymbol{K}(\boldsymbol{\xi})$ takes the form

$$K_{ik}(\boldsymbol{\xi}) = (\lambda+\mu)\xi_i\xi_k + \mu|\boldsymbol{\xi}|^2\delta_{ik} \equiv (\lambda+2\mu)\xi_i\xi_k + \mu(|\boldsymbol{\xi}|^2\delta_{ik} - \xi_i\xi_k). \tag{6.41}$$

The second form permits its immediate inversion:

$$\{[\boldsymbol{K}(\boldsymbol{\xi})]^{-1}\}_{ik} = \frac{1}{|\boldsymbol{\xi}|^4}\left\{\frac{\xi_i\xi_k}{\lambda+2\mu} + \frac{|\boldsymbol{\xi}|^2\delta_{ik}-\xi_i\xi_k}{\mu}\right\} = \frac{1}{\mu|\boldsymbol{\xi}|^2}\left\{\delta_{ik} - \frac{\lambda+\mu}{\lambda+2\mu}\,\frac{\xi_i\xi_k}{|\boldsymbol{\xi}|^2}\right\}. \tag{6.42}$$

The tensor $\tilde{\boldsymbol{\Gamma}}(\boldsymbol{\xi})$ now becomes, when $|\boldsymbol{\xi}| = 1$,

$$\tilde{\Gamma}_{ijpq}(\boldsymbol{\xi}) = \frac{1}{4\mu}\left(\xi_i\delta_{jp}\xi_q + \xi_j\delta_{ip}\xi_q + \xi_i\delta_{jq}\xi_p + \xi_j\delta_{iq}\xi_p\right) - \frac{\lambda+\mu}{\mu(\lambda+2\mu)}\xi_i\xi_j\xi_p\xi_q. \tag{6.43}$$

It is now easy to evaluate the tensor $\boldsymbol{P}$ in the case of a sphere. First, formula (6.33) shows that it is necessary to average $\xi_i\xi_q$ over the unit sphere. The result must be an isotropic second-order tensor, and hence a multiple of δ_{iq}:

$$\frac{1}{4\pi}\int_{|\boldsymbol{\xi}|=1}\xi_i\xi_q\,\mathrm{d}S = \alpha\delta_{iq}. \tag{6.44}$$

The constant α follows by setting $i = q$ and summing. Thus, $\alpha = 1/3$. Similarly, the average of $\xi_i\xi_j\xi_p\xi_q$ over the unit sphere is an isotropic fourth-order tensor, which is also completely symmetric in its indices. The constant $1/15$ in the formula

$$\frac{1}{4\pi}\int_{|\boldsymbol{\xi}|=1}\xi_i\xi_j\xi_p\xi_q\,\mathrm{d}S = \tfrac{1}{15}(\delta_{ij}\delta_{pq} + \delta_{ip}\delta_{jq} + \delta_{iq}\delta_{jp}) \tag{6.45}$$

is obtained by setting $i = j$, and $p = q$, and summing over i and p. Putting these results together now gives

$$P_{ijpq} = \frac{1}{6\mu}(\delta_{iq}\delta_{jp} + \delta_{jq}\delta_{ip}) - \frac{(\lambda+\mu)}{15\mu(\lambda+2\mu)}(\delta_{ij}\delta_{pq} + \delta_{ip}\delta_{jq} + \delta_{iq}\delta_{jp}). \tag{6.46}$$

This result can be expressed

$$\boldsymbol{P} = (3\kappa_P, 2\mu_P), \quad \text{or} \quad P_{ijpq} = \kappa_P\delta_{ij}\delta_{pq} + \mu_P(\delta_{ip}\delta_{jq} + \delta_{iq}\delta_{jp} - \tfrac{2}{3}\delta_{ij}\delta_{pq}), \tag{6.47}$$

where

$$3\kappa_P = \frac{1}{3\kappa+4\mu}, \quad 2\mu_P = \frac{3(\kappa+2\mu)}{5\mu(3\kappa+4\mu)}. \tag{6.48}$$

The results of subsection 4.3 are now easily confirmed.

7 The Hashin–Shtrikman Variational Principle and its Implications

The integral equation (6.13) has an associated variational principle. The operator $\boldsymbol{\Gamma}$ is symmetric, following from the symmetry $\boldsymbol{G}(\boldsymbol{x},\boldsymbol{x}') = \boldsymbol{G}^T(\boldsymbol{x}',\boldsymbol{x})$, and hence, immediately, the stationary principle

$$\delta\mathcal{H}(\boldsymbol{\tau}) = 0, \tag{7.1}$$

where

$$\mathcal{H}(\boldsymbol{\tau}) = \int_\Omega \{\boldsymbol{\varepsilon}^0\boldsymbol{\tau} - \tfrac{1}{2}\boldsymbol{\tau}(\boldsymbol{C}-\boldsymbol{C}^0)^{-1}\boldsymbol{\tau} - \tfrac{1}{2}\boldsymbol{\tau}\boldsymbol{\Gamma}\boldsymbol{\tau}\}\,\mathrm{d}\boldsymbol{x}, \tag{7.2}$$

generates (6.13). This is, basically, the Hashin–Shtrikman variational principle. Much more insight, as well as generality, will follow from the derivation given below.

7.1 Derivation of the Hashin–Shtrikman principle

Begin with the minimum energy principle in the form

$$\mathcal{F}(\boldsymbol{u}) = \inf_{\boldsymbol{u}^*}\int_\Omega \{W(\boldsymbol{\varepsilon}^*) - \boldsymbol{f}\boldsymbol{u}^*\}\,\mathrm{d}\boldsymbol{x};\ \ \boldsymbol{u}^* = \boldsymbol{u}^0,\ \ \boldsymbol{x}\in\partial\Omega. \tag{7.3}$$

The energy function $W(\boldsymbol{\varepsilon})$ will be assumed to be convex in $\boldsymbol{\varepsilon}$; it may be quadratic, as in (3.2), but it does not have to be. The energy density function may vary with position, $\boldsymbol{x}$, even though this is not shown explicitly. The adoption of the displacement boundary condition $\boldsymbol{u} = \boldsymbol{u}^0$ on $\partial\Omega$ is made for consistency with subsection 6.1, some of whose results will be invoked in the sequel.

Introduce a "comparison potential" $W^0(\boldsymbol{\varepsilon})$: this is likely, in practice, to be chosen to be a quadratic function, $\frac{1}{2}\boldsymbol{\varepsilon}\boldsymbol{C}^0\boldsymbol{\varepsilon}$, but the reasoning to follow does not require this. Now define

$$(W-W^0)^*(\boldsymbol{\tau}) := \sup_{\boldsymbol{\varepsilon}}\{\boldsymbol{\tau}\boldsymbol{\varepsilon} - W(\boldsymbol{\varepsilon}) + W^0(\boldsymbol{\varepsilon})\}. \tag{7.4}$$

Except in pathological cases, the supremum will be attained for some $\boldsymbol{\varepsilon}$, which will satisfy the condition of stationarity

$$\boldsymbol{\tau} = W'(\boldsymbol{\varepsilon}) - W^{0\prime}(\boldsymbol{\varepsilon}). \tag{7.5}$$

Thus, at least locally, $(W-W^0)^*$ will be a classical Legendre transform. In the case that $(W-W^0)$ is strictly convex, there will be only one stationary point, and this will deliver the supremum. If there is more than one stationary point, the classical Legendre transform will be multi-valued; the definition (7.4) selects the maximum.

Since, by definition, $(W-W^0)^*$ is the greatest possible value of the function in braces in (7.4), the Fenchel inequality

$$(W-W^0)^*(\boldsymbol{\tau}) \geq \boldsymbol{\tau}\boldsymbol{\varepsilon} - W(\boldsymbol{\varepsilon}) + W^0(\boldsymbol{\varepsilon}) \tag{7.6}$$

must hold for all $\boldsymbol{\tau}$ and $\boldsymbol{\varepsilon}$. The minimum energy principle (7.3) and (7.6) together imply

$$\mathcal{F}(\boldsymbol{u}) \geq \inf_{\boldsymbol{u}^*} \int_\Omega \{\boldsymbol{\tau}\boldsymbol{\varepsilon}^* + W^0(\boldsymbol{\varepsilon}^*) - (W-W^0)^*(\boldsymbol{\tau}) - \boldsymbol{f}\boldsymbol{u}^*\}\,\mathrm{d}\boldsymbol{x}, \quad \boldsymbol{u}^* = \boldsymbol{u}^0, \quad \boldsymbol{x} \in \partial\Omega, \tag{7.7}$$

for any $\boldsymbol{\tau}(\boldsymbol{x})$.

Now, starting from a different function W^0, define

$$(W-W^0)_*(\boldsymbol{\tau}) := \inf_{\boldsymbol{\varepsilon}} \{\boldsymbol{\tau}\boldsymbol{\varepsilon} - W(\boldsymbol{\varepsilon}) + W^0(\boldsymbol{\varepsilon})\}. \tag{7.8}$$

Since $(W-W^0)_*(\boldsymbol{\tau})$ is the smallest possible value of the bracketed function,

$$(W-W^0)_*(\boldsymbol{\tau}) \leq \boldsymbol{\varepsilon}\boldsymbol{\tau} - W(\boldsymbol{\varepsilon}) + W^0(\boldsymbol{\varepsilon}) \tag{7.9}$$

holds for all $\boldsymbol{\tau}$ and $\boldsymbol{\varepsilon}$, and the minimum energy principle implies

$$\mathcal{F}(\boldsymbol{u}) \leq \int_\Omega \{\boldsymbol{\tau}\boldsymbol{\varepsilon}^* + W^0(\boldsymbol{\varepsilon}^*) - (W-W^0)_*(\boldsymbol{\tau}) - \boldsymbol{f}\boldsymbol{u}^*\}\,\mathrm{d}\boldsymbol{x}, \quad \boldsymbol{u}^* = \boldsymbol{u}^0, \quad \boldsymbol{x} \in \partial\Omega, \tag{7.10}$$

for any $\boldsymbol{\tau}(\boldsymbol{x})$ and any $\boldsymbol{u}^*$ that satisfies the boundary conditions; the infimum over $\boldsymbol{u}^*$ is not needed because all the inequalities run the same way.

The inequalities (7.7) and (7.10) provide a rather general statement of the Hashin–Shtrikman variational principle(s). To see their relation to (7.1), specialise to a linear comparison medium, so that $W^0(\boldsymbol{\varepsilon}) = \frac{1}{2}\boldsymbol{\varepsilon}\boldsymbol{C}^0\boldsymbol{\varepsilon}$, and suppose that this is positive-definite. It is necessary to evaluate the infimum over $\boldsymbol{u}^*$ in (7.7), and it is desirable, to get the best possible result, in (7.10). The infimum is attained when

$$(C^0_{ijkl}u^*_{k,l})_{,j} + \tau_{ij,j} + f_i = 0, \quad \boldsymbol{x} \in \Omega, \tag{7.11}$$

with the boundary condition $\boldsymbol{u}^* = \boldsymbol{u}^0$ on $\partial\Omega$. Therefore, $\boldsymbol{u}^*$ is given by (6.9) and $\boldsymbol{\varepsilon}^*$ is given by (6.10). Write these, just for the moment, $\boldsymbol{u}^* = \boldsymbol{u}^0 + \boldsymbol{u}^1$, $\boldsymbol{\varepsilon}^* = \boldsymbol{\varepsilon}^0 + \boldsymbol{\varepsilon}^1$. The relevant part of the energy function (7.7) or (7.10) is then

$$\int_\Omega \{\boldsymbol{\tau}\boldsymbol{\varepsilon}^0 + \boldsymbol{\tau}\boldsymbol{\varepsilon}^1 + \tfrac{1}{2}\boldsymbol{\varepsilon}^1\boldsymbol{C}^0\boldsymbol{\varepsilon}^1 + \tfrac{1}{2}\boldsymbol{\varepsilon}^0\boldsymbol{C}^0\boldsymbol{\varepsilon}^0 - \boldsymbol{f}\boldsymbol{u}^0 + (\boldsymbol{\varepsilon}^1\boldsymbol{C}^0\boldsymbol{\varepsilon}^0 - \boldsymbol{u}^1\boldsymbol{f})\}\mathrm{d}\boldsymbol{x}.$$

Now

$$\int_\Omega (\boldsymbol{\varepsilon}^1 \boldsymbol{C}^0 \boldsymbol{\varepsilon}^0 - \boldsymbol{u}^1 \boldsymbol{f})\,\mathrm{d}\boldsymbol{x} = 0,$$

by the principle of virtual work, since the field $\boldsymbol{u}^0$ satisfies the equations of equilibrium with the body-force $\boldsymbol{f}$, while $\boldsymbol{u}^1 = 0$ on $\partial\Omega$. Similarly,

$$\int_\Omega \boldsymbol{\varepsilon}^1 (\boldsymbol{C}^0 \boldsymbol{\varepsilon}^1 + \boldsymbol{\tau})\,\mathrm{d}\boldsymbol{x} = 0,$$

because $\boldsymbol{C}^0\boldsymbol{\varepsilon}^1 + \boldsymbol{\tau}$ has zero divergence. Hence, keeping the terms that remain, (7.7) gives

$$\mathcal{F}(\boldsymbol{u}) \geq \int_\Omega \{\boldsymbol{\tau}\boldsymbol{\varepsilon}^0 - \tfrac{1}{2}\boldsymbol{\tau}\boldsymbol{\Gamma}\boldsymbol{\tau} - (W - W^0)^*(\boldsymbol{\tau}) + W^0(\boldsymbol{\varepsilon}^0) - \boldsymbol{f}\boldsymbol{u}^0\}\mathrm{d}\boldsymbol{x}, \tag{7.12}$$

having written $\boldsymbol{\varepsilon}^1 = -\boldsymbol{\Gamma}\boldsymbol{\tau}$, in line with (6.10). Similarly, (7.10) gives

$$\mathcal{F}(\boldsymbol{u}) \leq \int_\Omega \{\boldsymbol{\tau}\boldsymbol{\varepsilon}^0 - \tfrac{1}{2}\boldsymbol{\tau}\boldsymbol{\Gamma}\boldsymbol{\tau} - (W - W^0)_*(\boldsymbol{\tau}) + W^0(\boldsymbol{\varepsilon}^0) - \boldsymbol{f}\boldsymbol{u}^0\}\mathrm{d}\boldsymbol{x}. \tag{7.13}$$

In the case that $W(\boldsymbol{\varepsilon}) = \frac{1}{2}\boldsymbol{\varepsilon}\boldsymbol{C}\boldsymbol{\varepsilon}$, corresponding to linear elasticity, elementary calculation gives

$$(W - W^0)^*(\boldsymbol{\tau}) = \tfrac{1}{2}\boldsymbol{\tau}(\boldsymbol{C} - \boldsymbol{C}^0)^{-1}\boldsymbol{\tau}. \tag{7.14}$$

The expression on the right side of (7.14) in fact is the stationary value of the function $\boldsymbol{\tau}\boldsymbol{\varepsilon} - (W - W^0)(\boldsymbol{\varepsilon})$. It is the desired supremum so long as the quadratic form $\boldsymbol{\varepsilon}(\boldsymbol{C} - \boldsymbol{C}^0)\boldsymbol{\varepsilon}$ is positive-definite. Otherwise, the supremum is $+\infty$: the inequality (7.7) remains true, but becomes trivial! A similar conclusion can be reached in the case of the upper bound, given by (7.13); for this, the upper bound is non-trivial so long as $\boldsymbol{\varepsilon}(\boldsymbol{C} - \boldsymbol{C}^0)\boldsymbol{\varepsilon}$ is negative-definite. The result can be summarised thus:

$$\mathcal{F}(\boldsymbol{u}) \geq (\leq) \int_\Omega \{\boldsymbol{\tau}\boldsymbol{\varepsilon}^0 - \tfrac{1}{2}\boldsymbol{\tau}\boldsymbol{\Gamma}\boldsymbol{\tau} - \tfrac{1}{2}\boldsymbol{\tau}(\boldsymbol{C} - \boldsymbol{C}^0)^{-1}\boldsymbol{\tau} + \tfrac{1}{2}\boldsymbol{\varepsilon}^0\boldsymbol{C}^0\boldsymbol{\varepsilon}^0 - \boldsymbol{f}\boldsymbol{u}^0\}\,\mathrm{d}\boldsymbol{x} \tag{7.15}$$

for any $\boldsymbol{\tau}$, so long as the quadratic form $\boldsymbol{\varepsilon}(\boldsymbol{C} - \boldsymbol{C}^0)\boldsymbol{\varepsilon}$ is positive- (negative-) definite. The inequalities (7.12) and (7.13) are both true for any $\boldsymbol{\tau}$. The condition that the lower bound is maximised with respect to $\boldsymbol{\tau}$ is formally the same as the condition that the upper bound is minimised. It is that the integrand in (7.15) should be stationary. Thus,

$$(\boldsymbol{C} - \boldsymbol{C}^0)^{-1}\boldsymbol{\tau} + \boldsymbol{\Gamma}\boldsymbol{\tau} = \boldsymbol{\varepsilon}^0, \tag{7.16}$$

which is a repeat of equation (6.13). When this condition is met, the associated strain field $\boldsymbol{\varepsilon}$ is given by (6.12), and it follows that (6.3) is also satisfied. Thus, the exact solution to the problem is generated. Only conditions for stationarity have been imposed here; thus, stationarity of the Hashin–Shtrikman functional $\mathcal{H}$ is equivalent to the original formulation of the problem, for any choice of comparison medium $\boldsymbol{C}^0$. However, it provides a maximum or a minimum principle only when $\boldsymbol{C}^0$ is restricted as indicated.

7.2 Random media

It is very often appropriate to treat a composite as a random medium: even in the unusual event that the composite has a perfectly periodic structure, with period cell Q, say, it would be quite exceptional for the exact position of one chosen cell to be known exactly, relative to the boundary of the specimen or structure. In this sense, even a medium with periodic microstructure becomes random: a natural assumption is that a specified corner of a specified cell can occupy any position within a cell, Q_0, say, whose corner lies at the origin of coordinates. Of course, once the position of that one corner is fixed, the whole of the geometry of the composite is defined.

Now more generally, let a realisation of the composite be defined by a parameter (possibly infinite-dimensional) $\alpha \in \mathcal{A}$, where $\mathcal{A}$ is the sample space, over which a probability measure p is defined. For the purpose of this work, it suffices to describe the composite in terms of the set of characteristic functions χ_r of the region occupied by material of type r ($r = 1, 2, \cdots n$). Since the composite is to be treated as random, the functions χ_r depend on position $\boldsymbol{x}$ and on the parameter α:

$$\chi_r(\boldsymbol{x}, \alpha) = \begin{cases} 1 & \text{if } \boldsymbol{x} \in \text{phase } r \\ 0 & \text{otherwise.} \end{cases} \tag{7.17}$$

In the case of a periodic medium, as discussed above, the parameter α could be identified with the coordinates of the corner of the chosen cell. Then, the corresponding sample space $\mathcal{A}$ would be Q_0, and p would be the uniform measure on Q_0.

The ensemble mean of any quantity $\phi(\alpha)$ is now defined as

$$\langle \phi \rangle := \int_{\mathcal{A}} \phi(\alpha) \mathrm{p}(\mathrm{d}\alpha). \tag{7.18}$$

The ensemble mean of χ_r gives the probability $p_r(\boldsymbol{x})$ of finding material of type r at position $\boldsymbol{x}$:

$$p_r(\boldsymbol{x}) := \langle \chi_r \rangle(\boldsymbol{x}) = \int_{\mathcal{A}} \chi_r(\boldsymbol{x}, \alpha) \, \mathrm{p}(\mathrm{d}\alpha). \tag{7.19}$$

Similarly,

$$p_{rs}(\boldsymbol{x}, \boldsymbol{x}') := \int_{\mathcal{A}} \chi_r(\boldsymbol{x}, \alpha) \chi_s(\boldsymbol{x}', \alpha) \, \mathrm{p}(\mathrm{d}\alpha) \tag{7.20}$$

is the probability of finding simultaneously phase r at $\boldsymbol{x}$ and phase s at $\boldsymbol{x}'$. Probabilities involving more points can be defined similarly.

In the case that the composite is "statistically uniform", multipoint probabilities are insensitive to translations. Thus, the probability p_r becomes independent of $\boldsymbol{x}$, and so equal to the volume fraction c_r. The two-point function p_{rs} becomes a function of $\boldsymbol{x}$ and $\boldsymbol{x}'$ only in the combination $\boldsymbol{x} - \boldsymbol{x}'$.

7.3 Bounds

The plan now is to substitute trial polarisations into (7.15) to obtain upper and lower bounds for the energy in a composite. Its elastic constant tensor $\boldsymbol{C}(\boldsymbol{x},\alpha)$ takes the form

$$\boldsymbol{C}(\boldsymbol{x},\alpha)=\sum_{r=1}^{n}\boldsymbol{C}^r\chi_r(\boldsymbol{x},\alpha). \tag{7.21}$$

Correspondingly, the simplest choice that can be made for the polarisation, so that it depends on the microstructure, is

$$\boldsymbol{\tau}(\boldsymbol{x},\alpha)=\sum_{r=1}^{n}\boldsymbol{\tau}^r(\boldsymbol{x})\chi_r(\boldsymbol{x},\alpha), \tag{7.22}$$

where the functions $\boldsymbol{\tau}^r$ are sure, *i.e.* they are the same for every realisation α. Substitution into (7.15) now gives

$$\begin{aligned}\mathcal{F}(\boldsymbol{u}) \;\; \geq (\leq) \;\; \int_\Omega \Big\{\sum_{r=1}^{n}\boldsymbol{\tau}^r\left(\boldsymbol{\varepsilon}^0-\tfrac{1}{2}(\boldsymbol{C}^r-\boldsymbol{C}^0)^{-1}\boldsymbol{\tau}^r\right)\chi_r-\tfrac{1}{2}\sum_{r=1}^{n}\sum_{s=1}^{n}\boldsymbol{\tau}^r\boldsymbol{\Gamma}\boldsymbol{\tau}^s\chi_r\chi_s\\ +\tfrac{1}{2}\boldsymbol{\varepsilon}^0\boldsymbol{C}^0\boldsymbol{\varepsilon}^0-\boldsymbol{f}\boldsymbol{u}^0\Big\}\,\mathrm{d}\boldsymbol{x}.\end{aligned} \tag{7.23}$$

Now the intention is to exploit the notion that, if the microstructure is fine enough, the energy will not vary significantly from one realisation to another. An easy way through is to note that, since the inequalities (7.23) are true for all realisations, they survive the operation of taking the ensemble mean. Thus,

$$\begin{aligned}\langle\mathcal{F}(\boldsymbol{u})\rangle \;\; \geq (\leq) \;\; \int_\Omega \Big\{\sum_{r=1}^{n}p_r\boldsymbol{\tau}^r\left(\boldsymbol{\varepsilon}^0-\tfrac{1}{2}(\boldsymbol{C}^r-\boldsymbol{C}^0)^{-1}\boldsymbol{\tau}^r\right)-\tfrac{1}{2}\sum_{r=1}^{n}\sum_{s=1}^{n}\boldsymbol{\tau}^r\{\boldsymbol{\Gamma}p_{rs}\}\boldsymbol{\tau}^s\\ +\tfrac{1}{2}\boldsymbol{\varepsilon}^0\boldsymbol{C}^0\boldsymbol{\varepsilon}^0-\boldsymbol{f}\boldsymbol{u}^0\Big\}\,\mathrm{d}\boldsymbol{x}.\end{aligned} \tag{7.24}$$

The term in the double sum requires explanation. The operation of $\boldsymbol{\Gamma}$ on a function ϕ produces

$$(\boldsymbol{\Gamma}\phi)(\boldsymbol{x})=\int_\Omega\boldsymbol{\Gamma}(\boldsymbol{x},\boldsymbol{x}')\phi(\boldsymbol{x}')\,\mathrm{d}\boldsymbol{x}'.$$

The term $\{\boldsymbol{\Gamma}p_{rs}\}$ represents the operator which, applied to ϕ, gives

$$(\{\boldsymbol{\Gamma}p_{rs}\}\phi)(\boldsymbol{x})=\int_\Omega\boldsymbol{\Gamma}(\boldsymbol{x},\boldsymbol{x}')p_{rs}(\boldsymbol{x},\boldsymbol{x}')\phi(\boldsymbol{x}')\,\mathrm{d}\boldsymbol{x}'.$$

The right side of (7.23) is now optimised (to produce either a lower or an upper bound), by requiring it to be stationary with respect to variations in $\boldsymbol{\tau}^r$. Thus,

$$p_r(\boldsymbol{C}^r - \boldsymbol{C}^0)^{-1}\boldsymbol{\tau}^r + \sum_{s=1}^{n} \{\boldsymbol{\Gamma} p_{rs}\}\boldsymbol{\tau}^s = p_r\boldsymbol{\varepsilon}^0. \tag{7.25}$$

This represents the "best approximation" to the exact equation (6.13) amongst the set of polarisations of the form (7.22). It requires only knowledge of the one- and two-point probabilities p_r and p_{rs}. Any more complicated polarisation would inevitably require more statistical information for its optimal evaluation: the resulting bounds would be better, because they would contain that information, but availability of such information would be the exception rather than the rule.

It should be noted that, once an optimal set of polarisations $\boldsymbol{\tau}^r$ has been obtained, it provides, via (6.10), the approximation

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}^0 - \sum_{r=1}^{n} \boldsymbol{\Gamma}(\boldsymbol{\tau}^r\chi_r) \tag{7.26}$$

for the strain field, and the corresponding approximation for the stress field,

$$\boldsymbol{\sigma} = \boldsymbol{C}^0\boldsymbol{\varepsilon} + \sum_{r=1}^{n} \boldsymbol{\tau}^r\chi_r. \tag{7.27}$$

The mean strain and stress are obtained by ensemble averaging these quantities:

$$\langle\boldsymbol{\varepsilon}\rangle = \boldsymbol{\varepsilon}^0 - \boldsymbol{\Gamma}\langle\boldsymbol{\tau}\rangle, \quad \langle\boldsymbol{\sigma}\rangle = \boldsymbol{C}^0\langle\boldsymbol{\varepsilon}\rangle + \langle\boldsymbol{\tau}\rangle, \tag{7.28}$$

where

$$\langle\boldsymbol{\tau}\rangle = \sum_{r=1}^{n} p_r\boldsymbol{\tau}^r. \tag{7.29}$$

Taken by themselves, equations (7.28) are exact; the approximation is contained in (7.29).

In conclusion of this subsection, it is noted that substituting the optimal polarisations into (7.24) gives

$$\langle\mathcal{F}(\boldsymbol{u})\rangle \geq (\leq) \int_\Omega \{\tfrac{1}{2}\langle\boldsymbol{\tau}\rangle\boldsymbol{\varepsilon}^0 + \tfrac{1}{2}\boldsymbol{\varepsilon}^0\boldsymbol{C}^0\boldsymbol{\varepsilon}^0 - \boldsymbol{f}\boldsymbol{u}^0\}\,\mathrm{d}\boldsymbol{x}. \tag{7.30}$$

7.4 The classical Hashin–Shtrikman bounds

The classical Hashin–Shtrikman bounds – generalised rather beyond their original exposition – are obtained by specialising the inequalities (7.30) to the case in which

$$\boldsymbol{\varepsilon}^0 = \overline{\boldsymbol{\varepsilon}}, \text{ constant, and } \boldsymbol{f} = 0.$$

It is necessary to solve the optimal equations (7.25) in this case. This task is reduced by noting that, even in the general case, $\boldsymbol{\varepsilon}^0$ can be eliminated in favour of $\langle\boldsymbol{\varepsilon}\rangle$ by use of the first of equations (7.28). Thus,

$$p_r(\boldsymbol{C}^r - \boldsymbol{C}^0)^{-1}\boldsymbol{\tau}^r + \sum_{s=1}^{n}\{\boldsymbol{\Gamma}(p_{rs} - p_r p_s)\}\boldsymbol{\tau}^s = p_r\langle\boldsymbol{\varepsilon}\rangle. \tag{7.31}$$

Although $\langle\boldsymbol{\varepsilon}\rangle$ is not known, this equation has particular significance because it directly relates the polarisation to the mean strain.

Suppose now that the dimensions of the body are large in comparison with the typical dimension of the microstructure. If the composite has no long-range order, the term $(p_{rs} - p_r p_s)$ will tend to zero, typically exponentially, as $|\boldsymbol{x} - \boldsymbol{x}'|$ becomes large, relative to the microscale. This means, essentially, that the operator $\{\boldsymbol{\Gamma}(p_{rs} - p_r p_s)\}$ is sensitive only to values of the function upon which it acts, within a small neighbourhood of the point $\boldsymbol{x}$. Therefore, except in a boundary layer close to $\partial\Omega$, the effect of the operator is dominated by its singularity at $\boldsymbol{x}' = \boldsymbol{x}$. This means that $\boldsymbol{\Gamma}$ can be approximated by its infinite-body form, for which it is known exactly: see equation (6.26).

Suppose further that the material from which the specimen has been produced is "statistically uniform", so that $p_r = c_r$, constant, and p_{rs} is a function of $\boldsymbol{x} - \boldsymbol{x}'$ only. Except in a boundary layer, the ensemble mean of $\boldsymbol{\varepsilon}$ will coincide with its spatial mean, $\overline{\boldsymbol{\varepsilon}}$, which is constant, also equal to $\boldsymbol{\varepsilon}^0$. Equations (7.25) become translation-invariant, and hence will give $\boldsymbol{\tau}^r =$ constant, where the constant values satisfy the algebraic equations

$$c_r(\boldsymbol{C}^r - \boldsymbol{C}^0)^{-1}\boldsymbol{\tau}^r + \sum_{s=1}^{n}\boldsymbol{A}_{rs}\boldsymbol{\tau}^s = c_r\overline{\boldsymbol{\varepsilon}}, \tag{7.32}$$

with

$$\boldsymbol{A}_{rs} = \int \boldsymbol{\Gamma}(\boldsymbol{x} - \boldsymbol{x}')(p_{rs} - p_r p_s)(\boldsymbol{x} - \boldsymbol{x}')\,\mathrm{d}\boldsymbol{x}' \equiv \int \boldsymbol{\Gamma}(\boldsymbol{x})(p_{rs} - p_r p_s)(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}. \tag{7.33}$$

Isotropic statistics

The constant tensors can be made more explicit in the case that the two-point probabilities p_{rs} are isotropic, and so functions only of $r = |\boldsymbol{x}|$. Employing the representation (6.26) shows the need to evaluate the integral

$$J(p) := \int \delta(\boldsymbol{\xi}.\boldsymbol{x} - p)\psi_{rs}(r)\,\mathrm{d}\boldsymbol{x} \tag{7.34}$$

and then to calculate $J''(0)$. Here, the notation

$$\psi_{rs} := p_{rs} - p_r p_s$$

has been introduced.

Notice now that $J(p)$ requires the integration of ψ_{rs} over the plane $\boldsymbol{\xi}.\boldsymbol{x} = p$. Thus, employing cylindrical polar coordinates (ρ, θ, z), with the z-axis aligned with $\boldsymbol{x}$,

$$J(p) = \int_0^{2\pi} \mathrm{d}\theta \int_0^{\infty} \rho\,\mathrm{d}\rho\,\psi_{rs}(\sqrt{\rho^2 + p^2}). \tag{7.35}$$

Integrating with respect to θ and transforming the variable ρ to $r = \sqrt{\rho^2 + p^2}$ gives

$$J(p) = 2\pi \int_p^{\infty} r\mathrm{d}r\,\psi_{rs}(r). \tag{7.36}$$

Hence,

$$J''(p) = -2\pi(p\psi'_{rs}(p) + \psi_{rs}(p)), \tag{7.37}$$

so

$$J''(0) = -2\pi\psi_{rs}(0) = -2\pi(p_{rs}(0) - c_r c_s) \equiv -2\pi c_r(\delta_{rs} - c_s) \ \text{ (no sum on } \ r). \tag{7.38}$$

The value of $p_{rs}(0)$ has to be c_r if $r = s$ or 0 otherwise, because only one material can occupy a single point.

It follows that

$$\boldsymbol{A}_{rs} = \boldsymbol{P}c_r(\delta_{rs} - c_s) \ \text{ (no sum on } \ r), \tag{7.39}$$

where the tensor $\boldsymbol{P}$ is given by (6.33). Equations (7.25) now reduce to

$$(\boldsymbol{C}^r - \boldsymbol{C}^0)^{-1}\boldsymbol{\tau}^r + \boldsymbol{P}(\boldsymbol{\tau}^r - \langle\boldsymbol{\tau}\rangle) = \overline{\boldsymbol{\varepsilon}}. \tag{7.40}$$

These equations can be solved, as follows. First,

$$\boldsymbol{\tau}^r = [(\boldsymbol{C}^r - \boldsymbol{C}^0)^{-1} + \boldsymbol{P}]^{-1}[\overline{\boldsymbol{\varepsilon}} + \boldsymbol{P}\langle\boldsymbol{\tau}\rangle].$$

Therefore, by averaging,

$$\langle \boldsymbol{\tau} \rangle = \left\langle [(\boldsymbol{C} - \boldsymbol{C}^0)^{-1} + \boldsymbol{P}]^{-1} \right\rangle [\overline{\boldsymbol{\varepsilon}} + \boldsymbol{P} \langle \boldsymbol{\tau} \rangle].$$

This equation can be solved for $\langle \boldsymbol{\tau} \rangle$, to give

$$\langle \boldsymbol{\tau} \rangle = \left\langle [\boldsymbol{I} + (\boldsymbol{C} - \boldsymbol{C}^0)\boldsymbol{P}]^{-1} \right\rangle^{-1} \left\langle [(\boldsymbol{C} - \boldsymbol{C}^0)^{-1} + \boldsymbol{P}]^{-1} \right\rangle \overline{\boldsymbol{\varepsilon}}. \tag{7.41}$$

It is possible to continue, to obtain each $\boldsymbol{\tau}^r$, but this is not necessary for evaluation of the bound. Substituting into (7.30), specialised to the present case, gives the result

$$\tfrac{1}{2} \overline{\boldsymbol{\varepsilon}} \boldsymbol{C}^{\text{eff}} \overline{\boldsymbol{\varepsilon}} \geq (\leq) \, \tfrac{1}{2} \overline{\boldsymbol{\varepsilon}} \boldsymbol{C}^{HS} \overline{\boldsymbol{\varepsilon}}, \tag{7.42}$$

where

$$\begin{aligned} \boldsymbol{C}^{HS} &:= \boldsymbol{C}^0 + \left\langle [\boldsymbol{I} + (\boldsymbol{C} - \boldsymbol{C}^0)\boldsymbol{P}]^{-1} \right\rangle^{-1} \left\langle [(\boldsymbol{C} - \boldsymbol{C}^0)^{-1} + \boldsymbol{P}]^{-1} \right\rangle \\ &\equiv \left\langle [\boldsymbol{I} + (\boldsymbol{C} - \boldsymbol{C}^0)\boldsymbol{P}]^{-1} \right\rangle^{-1} \left\langle [\boldsymbol{I} + (\boldsymbol{C} - \boldsymbol{C}^0)\boldsymbol{P}]^{-1} \boldsymbol{C} \right\rangle. \end{aligned} \tag{7.43}$$

The result may be summarised concisely as follows:

$$(\boldsymbol{C}^{\text{eff}} - \boldsymbol{C}^{HS}) \geq (\leq) \, 0, \quad \text{for all } \boldsymbol{C}^0 \text{ such that } (\boldsymbol{C}^r - \boldsymbol{C}^0) \geq (\leq) \, 0 \text{ for all } r, \tag{7.44}$$

the inequalities here implying order relations in the sense of the associated quadratic form.

Isotropic phases

If the elastic constant tensors are also isotropic, the above result can be made completely explicit, by employing the expression (6.47) in conjunction with (6.48). Noting, too, that an isotropic elastic constant tensor $\boldsymbol{C} = (3\kappa, 2\mu)$ is positive-definite if and only if $\kappa > 0$ and $\mu > 0$, the result is

$$\kappa^{\text{eff}} \geq (\leq) \, \kappa^{HS}, \quad \mu^{\text{eff}} \geq (\leq) \, \mu^{HS} \tag{7.45}$$

whenever $\kappa^r \geq (\leq) \, \kappa^0$ and $\mu^r \geq (\leq) \, \mu^0$ for all r, where

$$\begin{aligned} \kappa^{HS} &= \left\{ \sum_r c_r \frac{(3\kappa^0 + 4\mu^0)}{(3\kappa^r + 4\mu^0)} \right\}^{-1} \sum_s c_s \frac{(3\kappa^0 + 4\mu^0)\kappa^s}{(3\kappa^s + 4\mu^0)}, \\ \mu^{HS} &= \left\{ \sum_r c_r \frac{5\mu^0(3\kappa^0 + 4\mu^0)}{6\mu^r(\kappa^0 + 2\mu^0) + \mu^0(9\kappa^0 + 8\mu^0)} \right\}^{-1} \\ &\qquad \sum_s c_s \frac{5\mu^0(3\kappa^0 + 4\mu^0)\mu^s}{6\mu^s(\kappa^0 + 2\mu^0) + \mu^0(9\kappa^0 + 8\mu^0)}. \end{aligned} \tag{7.46}$$

It should be noted that these formulae are *identical* to (4.26), which were developed as an *ad hoc* approximation!

7.5 More general two-point statistics

Ellipsoidal statistics

The general formulae (7.43) survive unchanged, if all of the two-point probabilities are functions of $\boldsymbol{x}$ in the combination $\rho = (\boldsymbol{x}^T(A^TA)^{-1}\boldsymbol{x})^{1/2}$; this could be achieved, for example, if a composite with intially isotropic statistics were subjected to an affine transformation $\boldsymbol{x} \to A^{-1}\boldsymbol{x}$. The only difference is that the tensor $\boldsymbol{P}$ is given by (6.38) instead of (6.47).

A general two-phase composite

If the composite has only two phases, the identity $\chi_1 + \chi_2 = 1$ induces the relations

$$\begin{aligned} p_1(\boldsymbol{x}) + p_2(\boldsymbol{x}) &= 1, \\ p_{11}(\boldsymbol{x},\boldsymbol{x}') + p_{21}(\boldsymbol{x},\boldsymbol{x}') &= p_1(\boldsymbol{x}'), \quad p_{11}(\boldsymbol{x},\boldsymbol{x}') + p_{12}(\boldsymbol{x},\boldsymbol{x}') = p_1(\boldsymbol{x}), \\ p_{12}(\boldsymbol{x},\boldsymbol{x}') + p_{22}(\boldsymbol{x},\boldsymbol{x}') &= p_2(\boldsymbol{x}'), \quad p_{21}(\boldsymbol{x},\boldsymbol{x}') + p_{22}(\boldsymbol{x},\boldsymbol{x}') = p_2(\boldsymbol{x}). \end{aligned}$$

It follows, when the composite is statistically uniform, that

$$\psi_{11} = \psi_{22} = -\psi_{12} = -\psi_{21} = c_1c_2h, \tag{7.47}$$

where h is an even function of $\boldsymbol{x}-\boldsymbol{x}'$, with $h(0) = 1$. Also, if the composite has no long-range order, $h \to 0$ as $|\boldsymbol{x} - \boldsymbol{x}'| \to \infty$. Hence,

$$\boldsymbol{A}_{11} = \boldsymbol{A}_{22} = -\boldsymbol{A}_{12} = -\boldsymbol{A}_{21} = c_1c_2\boldsymbol{P}, \tag{7.48}$$

where now

$$\boldsymbol{P} = \int \boldsymbol{\Gamma}(\boldsymbol{x})h(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}. \tag{7.49}$$

The interesting aspect of these results is that they give $\boldsymbol{A}_{rs}$ in exactly the form (7.39). The reasoning follows through unchanged, to give the Hashin–Shtrikman bounds (7.43), now for any two-point statistics.

7.6 A general formula for a laminate

The construction in Section 2 showed that, in a laminate subjected to uniform mean strain $\boldsymbol{\varepsilon}^0 = \overline{\boldsymbol{\varepsilon}}$, the fields were uniform in each phase. Hence, relative to any uniform comparison medium, the polarisations $\boldsymbol{\tau}^r$ will be exactly constant, and the Hashin–Shtrikman variational approximation will correspondingly lead to the exact effective tensor $\boldsymbol{C}^{\mathrm{eff}}$ for the laminate.

In fact, since the fields vary only in the direction $\boldsymbol{n}$, equation (6.4) reduces to the form

$$\boldsymbol{K}(\boldsymbol{n})\boldsymbol{u}''(s) + \boldsymbol{n}\boldsymbol{\tau}'(s) + \boldsymbol{f}(s) = 0, \tag{7.50}$$

where $s = \boldsymbol{n}.\boldsymbol{x}$. The corresponding Green's function satisfies

$$\boldsymbol{K}(\boldsymbol{n})\boldsymbol{G}''(s) + \boldsymbol{I}\delta(s) = 0, \tag{7.51}$$

while the operator $\boldsymbol{\Gamma}$ has kernel

$$\boldsymbol{\Gamma}(s) = \boldsymbol{n} \otimes [\boldsymbol{K}(\boldsymbol{n})]^{-1} \otimes \boldsymbol{n}\Big|_{\text{symmetrised}} \delta(s) \equiv \tilde{\boldsymbol{\Gamma}}(\boldsymbol{n})\delta(s). \tag{7.52}$$

The integral equation (7.25) therefore reduces to the algebraic equation

$$(\boldsymbol{C} - \boldsymbol{C}^0)^{-1}\boldsymbol{\tau} + \tilde{\boldsymbol{\Gamma}}(\boldsymbol{n})(\boldsymbol{\tau} - \langle\boldsymbol{\tau}\rangle) = \overline{\boldsymbol{\varepsilon}}, \tag{7.53}$$

which is identical to (7.40), with $\tilde{\boldsymbol{\Gamma}}(\boldsymbol{n})$ now written in place of $\boldsymbol{P}$. It follows that $\boldsymbol{C}^{\text{eff}}$ for a laminate is given exactly by (7.43), with $\boldsymbol{P} = \tilde{\boldsymbol{\Gamma}}(\boldsymbol{n})$. That this formula is consistent with (2.30) can be confirmed by specialising to a two-phase laminate, and identifying $\boldsymbol{C}^0$ with $\boldsymbol{C}^1$.

7.7 A remark on optimality

Consider the (generalised) Hashin–Shtrikman bounds, for a general two-phase medium, so that the tensor $\boldsymbol{P}$ is given by (7.49). Use of the representation (6.27) generates the expression

$$\boldsymbol{P} = \int_{|\boldsymbol{\xi}|=1} \tilde{\boldsymbol{\Gamma}}(\boldsymbol{\xi})\hat{h}(\boldsymbol{\xi})\,\mathrm{d}S, \tag{7.54}$$

where

$$\hat{h}(\boldsymbol{\xi}) := \frac{-1}{8\pi^2}\int \delta''(\boldsymbol{\xi}.\boldsymbol{x})h(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = \frac{-1}{8\pi^2}\frac{\partial^2}{\partial p^2}\int \delta(\boldsymbol{\xi}.\boldsymbol{x} - p)h(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}\Big|_{(p=0)}. \tag{7.55}$$

The function $\hat{h}(\boldsymbol{\xi})$ defines a measure on the surface of the unit sphere. Its total mass, M say, is obtained by integrating it over the unit sphere. In view of the result (6.21),

$$M = \int \delta(\boldsymbol{x})h(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = h(0) = 1. \tag{7.56}$$

Thus, $\hat{h}(\boldsymbol{\xi})$ defines an H-measure, of total mass 1.

Now suppose that the two phases are well-ordered, in the sense that $\boldsymbol{C}^1 - \boldsymbol{C}^2$ defines a quadratic form which is definite – say positive-definite, without loss. Choosing $\boldsymbol{C}^0 = \boldsymbol{C}^1$ in the generalised Hashin–Shtrikman bound therefore generates an upper bound which, by simple algebra, can be placed in the form

$$(\boldsymbol{C}^{HS} - \boldsymbol{C}^1)^{-1} = (c_2)^{-1}(\boldsymbol{C}^2 - \boldsymbol{C}^1)^{-1} + (c_2)^{-1}c_1\boldsymbol{P}. \tag{7.57}$$

Section 2 indicated the construction of a hierarchical laminate whose effective elastic constant tensor would approach $\boldsymbol{C}^{HS}$ arbitrarily closely. The Hashin–Shrikman bound in this case is therefore optimal. Interchanging materials 1 and 2 shows that the lower bound is also optimal.

7.8 A remark on the self-consistent approximation

The Hashin–Shtrikman approximation yields a bound for the energy if $\boldsymbol{C}^0$ is suitably restricted; when $\boldsymbol{C}^0$ is not so restricted, it still provides a stationary approximation for the energy. In this latter context, one possible choice for the comparison medium is to take $\boldsymbol{C}^0$ equal to the $\boldsymbol{C}^{\text{eff}}$ that it predicts. This is an approximation of "self-consistent" type.

In the case of a composite with isotropic (or, more generally, ellipsoidal) two-point statistics, this self-consistent prescription gives $\boldsymbol{C}^{\text{eff}} = \boldsymbol{C}^0$, where

$$\boldsymbol{C}^0 = \left\langle [\boldsymbol{I} + (\boldsymbol{C} - \boldsymbol{C}^0)\boldsymbol{P}^0]^{-1} \right\rangle^{-1} \left\langle [\boldsymbol{I} + (\boldsymbol{C} - \boldsymbol{C}^0)\boldsymbol{P}^0]^{-1}\boldsymbol{C} \right\rangle. \tag{7.58}$$

The superscript 0 has been placed on $\boldsymbol{P}$ to emphasise that it is calculated from $\boldsymbol{C}^0$. Equation (7.58) is well suited to solution by iteration. Its content is that the energy that is predicted in the composite is equal to the energy that the homogeneous comparison medium would have, when subjected to the same mean strain $\overline{\boldsymbol{\varepsilon}}$.

An equivalent form is

$$\left\langle [\boldsymbol{I} + (\boldsymbol{C} - \boldsymbol{C}^0)\boldsymbol{P}^0]^{-1} \right\rangle^{-1} \left\langle [(\boldsymbol{C} - \boldsymbol{C}^0)^{-1} + \boldsymbol{P}^0]^{-1} \right\rangle = 0, \tag{7.59}$$

and hence

$$\left\langle [(\boldsymbol{C} - \boldsymbol{C}^0)^{-1} + \boldsymbol{P}^0]^{-1} \right\rangle = 0. \tag{7.60}$$

These equations state that the comparison medium is chosen so that the mean polarisation is zero.

Exactly the same formulae apply to a general two-phase composite, with $\boldsymbol{P}^0$ suitably defined.

The formula (7.60) can be manipulated to give

$$\left\langle [\boldsymbol{I} + (\boldsymbol{C} - \boldsymbol{C}^0)\boldsymbol{P}^0]^{-1} \right\rangle = \boldsymbol{I}. \tag{7.61}$$

Thus, also,

$$\boldsymbol{C}^0 = \left\langle [\boldsymbol{I} + (\boldsymbol{C} - \boldsymbol{C}^0)\boldsymbol{P}^0]^{-1}\boldsymbol{C} \right\rangle, \tag{7.62}$$

and this can be expressed in the form

$$\boldsymbol{C}^0 = \boldsymbol{C}^n + \sum_{r=1}^{n-1} c_r [\boldsymbol{I} + (\boldsymbol{C}^r - \boldsymbol{C}^0)\boldsymbol{P}^0]^{-1}(\boldsymbol{C}^r - \boldsymbol{C}^n). \tag{7.63}$$

This is in precise agreement with the self-consistent approximation of subsection 4.5, which was derived from the solution of an inclusion problem for a matrix containing spheres. The formula thus has a rational foundation much more generally.

8 Nonlinear Response of Composites

This section presents just a brief outline of work on composites whose constitutive response is nonlinear. Almost all of this is of very recent origin, and the subject remains under active development. There are two main strands. One is the development of rigorous bounds, generalising Section 7. The other is the exploitation of results, such as those of Section 7, for linear composites, now employed as "comparison media" for the nonlinear composites of present concern. The theory is most advanced for composites whose behaviour can be described via a convex potential function $W(\boldsymbol{\varepsilon})$ (as in (1.14)). Path-dependent response, as realised by models of plasticity, requires incremental treatment. Each increment can be analysed by the methods given in these notes. However, the state from which any given increment commences is known only to some level of approximation, and the theory consequently suffers some loss of precision. An extended discussion is beyond the scope of these notes.

8.1 Elementary bounds

Suppose that the constitutive response of the composite can be described by equation (1.14), except that the potential W depends on position $\boldsymbol{x}$:

$$\boldsymbol{\sigma} = W'(\boldsymbol{\varepsilon}, \boldsymbol{x}), \quad \text{or} \quad \sigma_{ij} = \frac{\partial W(\boldsymbol{\varepsilon}, \boldsymbol{x})}{\partial \varepsilon_{ij}}, \tag{8.1}$$

where

$$W(\boldsymbol{\varepsilon}, \boldsymbol{x}) = \sum_{r=1}^{n} W^r(\boldsymbol{\varepsilon}) \chi_r(\boldsymbol{x}). \tag{8.2}$$

The potential $W^r(\boldsymbol{\varepsilon})$ associated with phase r is assumed convex, for each r. In the context of small-deformation but physically-nonlinear elasticity (or the deformation theory of plasticity), $\boldsymbol{\varepsilon}$ is the strain, which is derived from displacement $\boldsymbol{u}$. However, (8.1) also models nonlinear creep behaviour, if $\boldsymbol{\varepsilon}$ is interpreted as the strain-rate associated with velocity field $\boldsymbol{u}$. The mathematics is the same in either case. The following discussion will be phrased in the context of physically-nonlinear elasticity.

The classical energy principles given in subsection 3.2 generalise immediately to the response defined by (8.1).

Minimum Energy Principle: if the body which occupies a domain Ω is subjected to body-force $\boldsymbol{f}$ per unit volume, and a part S_u of its surface is subjected to prescribed displacements $\boldsymbol{u}^0$ while the complementary part S_t is subjected to prescribed tractions $\boldsymbol{t}^0$, the equilibrium

displacement minimises the energy functional

$$\mathcal{F}(\boldsymbol{u}) := \int_\Omega (W(\boldsymbol{\varepsilon}) - f_i u_i)\,\mathrm{d}\boldsymbol{x} - \int_{S_t} t_i^0 u_i\,\mathrm{d}S, \tag{8.3}$$

amongst displacement fields which take the prescribed values over S_u.

The complementary energy density is defined as

$$\begin{aligned} W^*(\boldsymbol{\sigma},\boldsymbol{x}) &= \sup_{\boldsymbol{\varepsilon}}\{\boldsymbol{\sigma}\boldsymbol{\varepsilon} - W(\boldsymbol{\varepsilon},\boldsymbol{x})\} \\ &\equiv \boldsymbol{\sigma}\boldsymbol{\varepsilon} - W(\boldsymbol{\varepsilon},\boldsymbol{x});\ \ \boldsymbol{\sigma} = W'(\boldsymbol{\varepsilon},\boldsymbol{x}),\ \ \boldsymbol{\varepsilon} = W^{*\prime}(\boldsymbol{\sigma},\boldsymbol{x}). \end{aligned} \tag{8.4}$$

Complementary energy principle: For the boundary value problem described above, the actual stress field minimises the functional

$$\mathcal{G}(\boldsymbol{\sigma}) := \int_\Omega W^*(\boldsymbol{\sigma},\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} - \int_{S_u} \sigma_{ij} n_j u_i^0\,\mathrm{d}S, \tag{8.5}$$

amongst stress fields that satisfy the equations of equilibrium $\sigma_{ij,j} + f_i = 0$ in Ω, and the given traction conditions on S_t.

(a) Linear displacement boundary condition. The minimum energy principle takes the form (3.16), there being no body force. Substituting the trial field $\boldsymbol{\varepsilon}' = \overline{\boldsymbol{\varepsilon}}$ gives the upper bound

$$W^{\mathrm{eff}}(\overline{\boldsymbol{\varepsilon}}) \le \overline{W}(\overline{\boldsymbol{\varepsilon}}) \equiv \int_\Omega W(\overline{\boldsymbol{\varepsilon}},\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \equiv \sum_{r=1}^{n} c_r W^r(\overline{\boldsymbol{\varepsilon}}). \tag{8.6}$$

In the nonlinear context, this is usually called the Taylor bound. Here, as in subsection 3.3, units are chosen so that the domain Ω has unit volume.

Substituting the constant stress $\overline{\boldsymbol{\sigma}}$ as a trial field in the complementary energy principle gives

$$W^{*\,\mathrm{eff}}(\overline{\boldsymbol{\sigma}}) \le \overline{W^*}(\overline{\boldsymbol{\sigma}}) \equiv \sum_{r=1}^{n} c_r W^{r*}(\overline{\boldsymbol{\sigma}}). \tag{8.7}$$

Hence,[9]

$$W^{\mathrm{eff}}(\overline{\boldsymbol{\varepsilon}}) = \sup_{\overline{\boldsymbol{\sigma}}}\{\overline{\boldsymbol{\varepsilon}}\,\overline{\boldsymbol{\sigma}} - W^{*\,\mathrm{eff}}(\overline{\boldsymbol{\sigma}})\} \ge \sup_{\overline{\boldsymbol{\sigma}}}\{\overline{\boldsymbol{\varepsilon}}\,\overline{\boldsymbol{\sigma}} - \overline{W}^*(\overline{\boldsymbol{\sigma}})\} = \left\{\overline{W^*}\right\}^*(\overline{\boldsymbol{\varepsilon}}). \tag{8.8}$$

This is usually called the Sachs bound.

[9]This manipulation relies on the fact that $W^{*\,\mathrm{eff}}$ is the dual of W^{eff}, which is a convex function, and so equal to $(W^{*\,\mathrm{eff}})^*$. Proofs of these claims are omitted in this brief outline.

In summary, therefore,

$$\left\{\overline{W^*}\right\}^*(\overline{\boldsymbol{\varepsilon}}) \le W^{\text{eff}}(\overline{\boldsymbol{\varepsilon}}) \le \overline{W}(\overline{\boldsymbol{\varepsilon}}). \tag{8.9}$$

(b) Uniform traction boundary condition. Similar arguments produce the bounds

$$\left\{\overline{W}\right\}^*(\overline{\boldsymbol{\sigma}}) \le W^{*\,\text{eff}}(\overline{\boldsymbol{\sigma}}) \le \overline{W^*}(\overline{\boldsymbol{\sigma}}). \tag{8.10}$$

They carry exactly the same information as the bounds (8.9).

8.2 Hashin–Shtrikman formalism

The use of a comparison medium, with potential $W^0(\boldsymbol{\varepsilon})$, was introduced in Section 6 and developed further in Section 7. In the nonlinear context, the stress is represented in the form

$$\boldsymbol{\sigma} = W'(\boldsymbol{\varepsilon}) = W^{0\prime}(\boldsymbol{\varepsilon}) + \boldsymbol{\tau}. \tag{8.11}$$

The potential W depends on $\boldsymbol{x}$, as in (8.2), but this is suppressed for conciseness. The potential W^0 may also depend on $\boldsymbol{x}$, and does not in the most general case need to be a quadratic function. The reasoning at the start of subsection 7.1 applies, and the inequalities (7.7) and (7.10) apply at this level of generality.

Now one natural possibility is to take the potential W^0 to be quadratic:

$$W^0(\boldsymbol{\varepsilon}) = \tfrac{1}{2}\boldsymbol{\varepsilon}\boldsymbol{C}^0\boldsymbol{\varepsilon}, \tag{8.12}$$

(where $\boldsymbol{C}^0$ could still depend on $\boldsymbol{x}$). Considering first the lower bound (7.7), it is not necessary for $W^r - W^0$ to be convex for each r but it is necessary that each W^r should grow at least quadratically as $\|\boldsymbol{\varepsilon}\| \to \infty$, since otherwise $(W - W^0)^*$ will be infinite and the lower bound will be $-\infty$: true but useless! The opposite statement applies to the upper bound (7.10): it is necessary, if $(W - W^0)_*$ is to be finite, that each W^r should grow no faster than quadratically as $\|\boldsymbol{\varepsilon}\| \to \infty$. Thus, except in very special cases, if the composite is nonlinear, it is impossible for the technology so far developed to produce more than one bound: either an upper bound, or a lower bound, but not both. And if the composite were such that the potential of some phase grew faster than quadratically while the potential of another grew slower than quadratically, no bound would be obtained at all. The answer to finding the missing bound(s) has to be to introduce a nonlinear comparison medium. Some research has been done on this topic but it is mathematically too advanced to be described here; it is also rather limited in its practical application.

For the sake of this outline, suppose that each potential of the composite grows faster than quadratically, The lower bound formula (7.7) then gives a non-trivial result.

Consider first the case that $\boldsymbol{C}^0$ is uniform, and adopt the form (7.22) for $\boldsymbol{\tau}$:

$$\boldsymbol{\tau} = \sum_{r=1}^{n} \boldsymbol{\tau}^r \chi_r.$$

The displacement field associated with $\boldsymbol{\tau}$ is calculated exactly as in subsection 7.1, and it leads to the bound (7.12). This is now optimised with respect to the $\boldsymbol{\tau}^r$ by requiring that

$$p_r(W^r - W^0)^{*\prime}(\boldsymbol{\tau}^r) + \sum_{s=1}^{n} \{\boldsymbol{\Gamma}(p_{rs} - p_r p_s)\}\boldsymbol{\tau}^s = \langle \boldsymbol{\varepsilon} \rangle, \tag{8.13}$$

which is the natural extension of equation (7.31) to nonlinear behaviour[10]. The very first bounds which improved on the elementary bounds (8.9) were obtained by precisely this method, around 1985.

It was recognised, later, that results for linear composites could be exploited directly in the nonlinear setting, by taking $\boldsymbol{C}^0$ itself to correspond to a composite, with the same microgeometry as the given one. Thus,

$$\boldsymbol{C}^0(\boldsymbol{x}) = \sum_{r=1}^{n} \hat{\boldsymbol{C}}^r \chi_r(\boldsymbol{x}). \tag{8.14}$$

The intention is now to select the constants $\hat{\boldsymbol{C}}^r$ so that the field in the comparison linear composite mimics as closely as possible the field in the actual nonlinear composite. The mathematics that achieves this is very simple! First, specialising the inequality (7.7) to linear displacement boundary conditions and no body-force,

$$W^{\text{eff}}(\overline{\boldsymbol{\varepsilon}}) \geq \inf_{\boldsymbol{\varepsilon}} \int_{\Omega} \{\boldsymbol{\tau}\overline{\boldsymbol{\varepsilon}} + W^0(\boldsymbol{\varepsilon}) - (W - W^0)^*(\boldsymbol{\tau})\}\, \mathrm{d}\boldsymbol{x}, \tag{8.15}$$

where the infimum is taken over strain fields whose associated displacements satisfy the linear displacement boundary condition. It usually suffices to choose $\boldsymbol{\tau} = 0$. Then,

$$(W - W^0)^*(0) = \sup_{\boldsymbol{\varepsilon}} \{-(W - W^0)(\boldsymbol{\varepsilon})\} = -\min(W - W^0).$$

[10]It could happen that $(W^r - W^0)^*$ has corners and so is not differentable everywhere. The more general procedure is to replace $(W^r - W^0)^{*\prime}(\boldsymbol{\tau}^r)$ by a subgradient at $\boldsymbol{\tau}^r$, $\boldsymbol{\varepsilon}^r$ say. The relevant theory is not presented here.

Also,

$$\inf_{\boldsymbol{\varepsilon}} \int_\Omega W^0(\boldsymbol{\varepsilon})\,\mathrm{d}\boldsymbol{x} = W^{0,\mathrm{eff}}(\overline{\boldsymbol{\varepsilon}}),$$

the effective energy density of the linear comparison composite. Thus,

$$W^{\mathrm{eff}}(\overline{\boldsymbol{\varepsilon}}) \geq W^{0,\mathrm{eff}}(\overline{\boldsymbol{\varepsilon}}) + \sum_{r=1}^{n} c_r \min(W^r - \hat{W}^r), \tag{8.16}$$

where $\hat{W}^r(\boldsymbol{\varepsilon}) = \frac{1}{2}\boldsymbol{\varepsilon}\hat{\boldsymbol{C}}^r\boldsymbol{\varepsilon}$. A bound is obtained by replacing $W^{0,\mathrm{eff}}(\overline{\boldsymbol{\varepsilon}})$ by any bound already known from linear technology. It is optimised by choosing the best elastic constant tensors $\hat{\boldsymbol{C}}^r$. Remarkably, when this simple prescription is used in conjunction with the linear Hashin–Shtrikman bound, it gives the same result as the prescription based on solving equations (8.13), except in certain pathological cases, when it does not perform as well.

Some insight into the structure of this prescription follows from observing that $\min(W^r - \hat{W}^r)$ is achieved, at $\boldsymbol{\varepsilon}^r$ say, when

$$\boldsymbol{\sigma}^r := W^{r\prime}(\boldsymbol{\varepsilon}^r) = \hat{\boldsymbol{C}}^r\boldsymbol{\varepsilon}^r.$$

The stress and strain in phase r of the composite, insofar as they are represented by $\boldsymbol{\sigma}^r$ and $\boldsymbol{\varepsilon}^r$, are also realised in the comparison linear composite. However, the prescription leads to a bound, independently of any physical interpretation.

An alternative direct derivation of the formula (8.16) is as follows.

$$\begin{aligned} W^{\mathrm{eff}}(\overline{\boldsymbol{\tau}}) &= \inf_{\boldsymbol{\varepsilon}} \int_\Omega \{W^0 + (W - W^0)\}\mathrm{d}\boldsymbol{x} \\ &\geq \inf_{\boldsymbol{\varepsilon}} \int_\Omega W^0\,\mathrm{d}\boldsymbol{x} + \inf_{\boldsymbol{\varepsilon}} \int_\Omega (W - W^0)\,\mathrm{d}\boldsymbol{x} \\ &\geq \inf_{\boldsymbol{\varepsilon}} \int_\Omega W^0\,\mathrm{d}\boldsymbol{x} + \int_\Omega \min(W - W^0)\,\mathrm{d}\boldsymbol{x} \\ &= W^{0,\mathrm{eff}}(\overline{\boldsymbol{\varepsilon}}) + \sum_{r=1}^{n} c_r \min(W^r - \hat{W}^r). \end{aligned} \tag{8.17}$$

8.3 Other approximations

Approximations, as opposed to bounds, may follow from a variational structure or may be developed on the basis of some insight, either mathematical or physical. An advantage to approximations derived variationally is that an error of order ϵ, say, in the trial field will generate an error of order ϵ^2 (or perhaps smaller) in the energy. Variational methods are not available for path-dependent problems. One favourable course of action is to attempt

to develop prescriptions of general applicability directly, that have a variational status when they are applied to problems which have that structure.

One successful attempt in this direction employed a comparison linear composite, with pre-stress. Thus, it had a potential which, in region r, took the form

$$\hat{W}^r(\boldsymbol{\varepsilon}) = \tfrac{1}{2}\boldsymbol{\varepsilon}\hat{\boldsymbol{C}}^r\boldsymbol{\varepsilon} + \boldsymbol{\tau}^r\boldsymbol{\varepsilon} + \hat{W}^r(0). \tag{8.18}$$

The idea was to select the constants $\hat{W}^r(0)$, $\hat{\boldsymbol{C}}^r$ and $\boldsymbol{\tau}^r$ so that, in the vicinity of the strain level $\boldsymbol{\varepsilon}^r$, the actual potential W^r and the quadratic potential $\hat{W}^r$ would agree to second order. That is,

$$\tfrac{1}{2}\boldsymbol{\varepsilon}^r\hat{\boldsymbol{C}}^r\boldsymbol{\varepsilon}^r + \boldsymbol{\tau}^r\boldsymbol{\varepsilon}^r + \hat{W}^r(0) = W^r(\boldsymbol{\varepsilon}^r), \quad \hat{\boldsymbol{C}}^r\boldsymbol{\varepsilon}^r + \boldsymbol{\tau}^r = W^{r\prime}(\boldsymbol{\varepsilon}^r) \text{ and } \hat{\boldsymbol{C}}^r = W^{r\prime\prime}(\boldsymbol{\varepsilon}^r). \tag{8.19}$$

The strain $\boldsymbol{\varepsilon}^r$ is intended to be the mean strain achieved in region r of the actual nonlinear composite. This is not known, so instead it is taken as the mean strain over region r that is predicted by solving the problem for the linear composite. This scheme was originally proposed intuitively, essentially as just described, but then it was demonstrated to follow from a variational structure like that of the preceding subsection, except that $(W - W^0)^*$ was taken to be the appropriate branch of the classical Legendre transform rather than the supremum, or convex dual. This allowed the actual material response to be matched as closely as possible, but at the expense of loss of an inequality: the resulting variational principle is a stationary principle only. The second of conditions (8.19) corresponds to the condition of stationarity; the third condition makes the stationary point second-order, rather than just first-order, and the approximation for the energy should be correspondingly more accurate.

This scheme is quite closely related to a method called "the affine method", which can be applied even when there is no potential structure. The affine method starts from the same type of comparison medium, but the constants are fixed by direct matching of stresses, since there is no potential.