NISSUNA UMANA INVESTIGAZIONE SI PUO DIMANDARE
VERA SCIENZIA S'ESSA NON PASSA PER LE
MATEMATICHE DIMOSTRAZIONI
LEONARDO DA VINCI

vol. 5 no. 1 2017

MATHEMATICS AND MECHANICS of Complex Systems

msp

MATHEMATICS AND MECHANICS OF COMPLEX SYSTEMS

msp.org/memocs

MEMOCS (ISSN 2325-3444 electronic, 2326-7186 printed) is a journal of the International Research Center for the Mathematics and Mechanics of Complex Systems at the Università dell’Aquila, Italy.

Cover image: “Tangle” by © John Horigan; produced using the *Context Free* program (contextfreeart.org).

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

MATHEMATICS AND MECHANICS OF COMPLEX SYSTEMS
Vol. 5, No. 1, 2017
dx.doi.org/10.2140/memocs.2017.5.1

REDUCIBLE AND IRREDUCIBLE FORMS OF STABILISED GRADIENT ELASTICITY IN DYNAMICS

HARM ASKES AND INNA M. GITMAN

The continualisation of discrete particle models has been a popular tool to formulate higher-order gradient elasticity models. However, a straightforward continualisation leads to unstable continuum models. Padé approximations can be used to stabilise the model, but the resulting formulation depends on the particular equation that is transformed with the Padé approximation. In this contribution, we study two different stabilised gradient elasticity models; one is an irreducible form with displacement degrees of freedom only, and the other is a reducible form where the primary unknowns are not only displacements but also the Cauchy stresses — this turns out to be Eringen's theory of gradient elasticity. Although they are derived from the same discrete model, there are significant differences in variationally consistent boundary conditions and resulting finite element implementations, with implications for the capability (or otherwise) to suppress crack tip singularities.

1. Introduction

Gradient elasticity is a methodology to enrich the continuum equations of elasticity with additional higher-order spatial (and occasionally temporal) derivatives of certain state variables. There are different versions of gradient elasticity, such as those equipped with strain gradients, stress gradients and acceleration gradients; see for instance [Askes and Aifantis 2011] for a recent (but by no means complete) review.

Certain formats of gradient elasticity bear a close relationship with discrete lattice models of materials with microstructure; indeed, it is often possible to derive gradient elasticity theories by continualising the response of a discrete model, for instance using Taylor series approximations [Chang and Gao 1995; Mühlhaus and Oka 1996; Suiker et al. 2001a; Suiker et al. 2001b; Ioannidou et al. 2001; Askes and Metrikine 2005]. However, such models often suffer from intrinsic deficiencies, such as loss of stability in dynamics and loss of uniqueness in statics [Askes et al. 2002]. This can be amended by applying Padé approximations or similar

Communicated by Francesco dell'Isola.
MSC2010: 74-XX.
Keywords: gradient elasticity, mixed formulation, length scale, nonlocal elasticity.

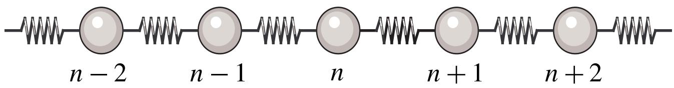

Figure 1. One-dimensional chain of masses connected by springs.

techniques, as has for instance been demonstrated in [Rosenau 1984; Rubin et al. 1995; Chen and Fish 2001; Andrianov 2002; Andrianov et al. 2003; Charlotte and Truskinovsky 2008]. Thus, stabilised gradient elasticity theories can be formulated that maintain their close link with discrete lattice models, thereby facilitating simple identification of the higher-order constitutive parameters (usually known as "intrinsic length scales" or "microstructural length scales").

In this paper, we compare two versions of stabilised gradient elasticity. Both can be derived from the response of a discrete lattice model, which is shown for the one-dimensional case. Variational formulations are presented for the multidimensional extensions. Throughout, a distinction is made between the so-called *irreducible form* where the only unknowns are the displacements and the *reducible form* where the unknowns are the displacements as well as the Cauchy stresses. The difference between these two forms has important consequences for the variationally consistent boundary conditions and finite element implementations. A numerical example will show the ability (or otherwise) of the two formulations to suppress singularities — this has historically been an important motivation for using gradient elasticity theories, and certain formats have been demonstrated to remove singularities even under restrictive conditions such as anisotropic material behaviour and bimaterial interface cracks [Kwong and Gitman 2012]. We also discuss the relation of the reducible form with Eringen's [1983] differential theory of nonlocal elasticity.

2. Continualisation of the response of a discrete chain

To illustrate the concepts of continualisation (this section) and stabilisation via Padé approximations (Section 3), the one-dimensional chain of particles and springs in Figure 1 is studied. All particles have mass M, and all springs have stiffness K. Furthermore, the interparticle distance is denoted by d. The equation of motion of particle n thus reads

$$M\ddot{u}_n = K(u_{n+1} - 2u_n + u_{n-1}) \tag{1}$$

where u_i is the displacement of particle i. A continuum approximation is obtained by replacing u_n with $u(x)$ and $u_{n\pm1}$ with $u(x \pm d)$. Taylor series expansions are applied according to

$$u(x \pm d) = u(x) \pm d\frac{\partial u}{\partial x} + \frac{1}{2}d^2\frac{\partial^2 u}{\partial x^2} \pm \frac{1}{6}d^3\frac{\partial^3 u}{\partial x^3} + \frac{1}{24}d^4\frac{\partial^4 u}{\partial x^4} \pm \cdots \tag{2}$$

so that (1) can be rewritten as

$$\rho \ddot{u} = E\left(\frac{\partial^2 u}{\partial x^2} + \frac{1}{12} d^2 \frac{\partial^4 u}{\partial x^4} + \cdots\right) \tag{3}$$

where the mass density $\rho = M/Ad$ and the Young's modulus $E = Kd/A$, with A the (unit) cross-sectional area of the system. Multidimensional formulations in the spirit of (3) have been derived by Chang and Gao [1995], Mühlhaus and Oka [1996] and Suiker et al. [2001a; Suiker et al. [2001b], among others.

Apart from the lowest-order, standard terms, (3) also contains higher-order terms proportional to d^2, d^4, etc. These additional terms capture the microstructural effects that are present in the discrete model of (1) but that are absent in standard continuum theories as retrieved by taking $d = 0$ in (3). The simplest continuum model that incorporates microstructural effects is obtained by truncating the series in (3) after the term that is proportional to d^2; unfortunately, such a model is unstable and its solutions in a boundary-value problem may lack uniqueness [Askes et al. 2002]. Although stability and uniqueness can be restored by incorporating the next term, i.e., truncating after the d^4 term, the numerical implementation of such a model is complicated [Askes et al. 2002]; thus, alternative solution strategies are explored here.

3. Stabilising the continuum equations

Unstable gradient theories can be turned into stable gradient theories by means of Padé approximations, as has been explored in [Andrianov et al. 2003; Andrianov and Awrejcewicz 2008; Andrianov et al. 2010]. However, there are various ways to do this, and the format of the resulting equations depends on which equations are transformed by the Padé approximation.

3.1. ***Irreducible form.*** Firstly, (3) is truncated after the first nonstandard term. The various spatial derivatives are factorised as

$$\rho \ddot{u} = \left(1 + \frac{1}{12} d^2 \frac{\partial^2}{\partial x^2}\right) E \frac{\partial^2 u}{\partial x^2}. \tag{4}$$

A [0, 1]-Padé approximation is used according to

$$1 + a \approx \frac{1}{1 - a} \quad \text{for } a \ll 1. \tag{5}$$

For a in (5), we will substitute the operator $\frac{1}{12} d^2\, \partial^2/\partial x^2$, which allows us to rewrite (4) as

$$\left(1 - \frac{1}{12} d^2 \frac{\partial^2}{\partial x^2}\right) \rho \ddot{u} = E \frac{\partial^2 u}{\partial x^2}. \tag{6}$$

The higher-order gradient term now appears on the inertia side of the equation, and for this reason, it has been called microinertia, internal inertia or higher-order inertia in the literature [Vardoulakis and Aifantis 1994; Wang and Sun 2002; Bennett et al. 2007]. Equation (6), or slight variations thereof, has also been obtained by various other researchers using asymptotic series equivalence; see for instance the work of Rubin et al. [1995], Chen and Fish [2001] and Pichugin et al. [2008].

Note that the only unknown appearing in (6) is the displacement; for this reason, this format is denoted as *irreducible*. Although at first sight it may appear that the micromechanical background of the higher-order terms is lost through the Padé approximation, an alternative interpretation of the microinertia contribution in terms of long-range interactions has been provided in [Askes and Gitman 2014].

3.2. ***Reducible form.*** It is also possible to extract a (one-dimensional) relation between stress σ and strain ε from (3) such that

$$\rho\ddot{u} = \frac{\partial \sigma}{\partial x} \quad \text{and} \quad \varepsilon = \frac{\partial u}{\partial x}. \tag{7}$$

The stress-strain relation then follows as

$$\sigma = E\left(\varepsilon + \frac{1}{12}d^2\frac{\partial^2\varepsilon}{\partial x^2}\right) = E\left(1 + \frac{1}{12}d^2\frac{\partial^2}{\partial x^2}\right)\varepsilon \tag{8}$$

where series have again been truncated after the first nonstandard term. Applying the [0, 1]-Padé approximation to (8) yields

$$\left(1 - \frac{1}{12}d^2\frac{\partial^2}{\partial x^2}\right)\sigma = E\varepsilon. \tag{9}$$

Equations (7) and (9) can be combined into a system of coupled equations,

$$\rho\ddot{u} = \frac{\partial \sigma}{\partial x} \tag{10a}$$

together with

$$\sigma - \frac{1}{12}d^2\frac{\partial^2\sigma}{\partial x^2} = E\frac{\partial u}{\partial x} \tag{10b}$$

where the unknowns are the displacement u as well as the stress σ. In contrast to the single fourth-order equation (6), (10) is a set of two second-order equations. They are termed *reducible* because it is possible to eliminate one of the unknowns, namely the stress σ. To do this, the second-order spatial derivative of (10a) must be taken and, multiplied with $\frac{1}{12}d^2$, subtracted from the original expression (10a):

$$\rho\left(\ddot{u} - \frac{1}{12}d^2\frac{\partial^2\ddot{u}}{\partial x^2}\right) = \frac{\partial}{\partial x}\left(\sigma - \frac{1}{12}d^2\frac{\partial^2\sigma}{\partial x^2}\right). \tag{11}$$

If (10b) is substituted into the right-hand side of (11), the stress will disappear from the expressions and thus it is possible to retrieve (6). This reduction of the number of unknowns, and its consequences, will be discussed in more depth below in Section 4.2.

4. Energy functionals for the multidimensional case

Above, the governing equations have been derived from simple mechanical and mathematical arguments in a one-dimensional context. Next, we will show how the analogous multidimensional equations can be derived from variational principles. Hamilton's action S is defined as

$$S = \int_{t_0}^{t_1} L \,\mathrm{d}t. \tag{12}$$

The governing equations of the models can be derived by requiring stationarity of S, that is, $\delta S = 0$. The energy functional (or Lagrangian function) L is defined individually for the two different models below, but we will assume that L depends on the displacements u_i and their spatial and temporal derivatives, as well as on the stresses σ_{ij} and their spatial derivatives:

$$L = L(u_i; u_{i,j}; \dot{u}_i; \dot{u}_{i,j}; \sigma_{ij}; \sigma_{ij,k}). \tag{13}$$

Substituting (13) into (12) and requiring $\delta S = 0$ yields

$$\begin{aligned}\int_{t_0}^{t_1} \delta L \,\mathrm{d}t &= \int_{t_0}^{t_1} \left(\delta u_i \frac{\partial L}{\partial u_i} + \delta u_{i,j} \frac{\partial L}{\partial u_{i,j}} + \delta \dot{u}_i \frac{\partial L}{\partial \dot{u}_i} + \delta \dot{u}_{i,j} \frac{\partial L}{\partial \dot{u}_{i,j}} + \delta \sigma_{ij} \frac{\partial L}{\partial \sigma_{ij}} + \delta \sigma_{ij,k} \frac{\partial L}{\partial \sigma_{ij,k}} \right) \mathrm{d}t \\ &= 0, \end{aligned} \tag{14}$$

which, as usual, can be rewritten as

$$\begin{aligned}&\int_{t_0}^{t_1} \delta u_i \left(\frac{\partial L}{\partial u_i} - \frac{\partial}{\partial x_j} \frac{\partial L}{\partial u_{i,j}} - \frac{\partial}{\partial t} \frac{\partial L}{\partial \dot{u}_i} + \frac{\partial^2}{\partial x_j \partial t} \frac{\partial L}{\partial \dot{u}_{i,j}} \right) \mathrm{d}t \\ &\quad + \int_{t_0}^{t_1} \frac{\partial}{\partial x_j} \left(\delta u_i \frac{\partial L}{\partial u_{i,j}} - \delta u_i \frac{\partial}{\partial t} \frac{\partial L}{\partial \dot{u}_{i,j}} \right) \mathrm{d}t + \int_{t_0}^{t_1} \frac{\partial}{\partial t} \left(\delta u_i \frac{\partial L}{\partial \dot{u}_i} + \delta u_{i,j} \frac{\partial L}{\partial \dot{u}_{i,j}} \right) \mathrm{d}t \\ &\qquad + \int_{t_0}^{t_1} \delta \sigma_{ij} \left(\frac{\partial L}{\partial \sigma_{ij}} - \frac{\partial}{\partial x_k} \frac{\partial L}{\partial \sigma_{ij,k}} \right) \mathrm{d}t + \int_{t_0}^{t_1} \frac{\partial}{\partial x_k} \left(\delta \sigma_{ij} \frac{\partial L}{\partial \sigma_{ij,k}} \right) \mathrm{d}t = 0. \end{aligned} \tag{15}$$

The third integral cancels through the requirement that $\delta u_i = 0$ and $\delta u_{i,j} = 0$ for $t = t_0$ and for $t = t_1$. The first and fourth integrals will lead to field equations, whereas the second and fifth will contribute to the natural boundary conditions.

4.1. *Irreducible form.* The Lagrangian function of the irreducible form can be written as

$$L_{\text{irred}} = \int_{\Omega} \tfrac{1}{2}\rho(\dot{u}_i\dot{u}_i + \ell^2\dot{u}_{i,j}\dot{u}_{i,j})\,\mathrm{d}V - \int_{\Omega} \tfrac{1}{2}u_{i,j}C_{ijkl}u_{k,l}\,\mathrm{d}V + \int_{\Omega} u_i b_i\,\mathrm{d}V + \int_{\Gamma_n} u_i t_i\,\mathrm{d}S \quad (16)$$

where the first integral is the kinetic energy, the second integral is the stored strain energy and the last two terms represent the work of the external forces. Thus, for this model, the Lagrangian takes the usual format of "kinetic energy minus potential energy", whereby the nonstandard contributions are included in the kinetic energy only [Lazar and Anastassiadis 2007; Polizzotto 2012]. Note that for the internal length scale we have now used the generic notation ℓ rather than the notation d that was used in the previous section in relation to the discrete model.

Substituting (16) into (15) and noting that $\delta u_i = 0$ on Γ_e leads to

$$\int_{t_0}^{t_1}\int_{\Omega} \delta u_i(b_i + C_{ijkl}u_{k,jl} - \rho\ddot{u}_i + \rho\ell^2\ddot{u}_{i,jj})\,\mathrm{d}V\,\mathrm{d}t + \int_{t_0}^{t_1}\int_{\Gamma_n} \delta u_i(t_i - n_j(C_{ijkl}u_{k,l} + \rho\ell^2\ddot{u}_{i,j}))\,\mathrm{d}S\,\mathrm{d}t = 0, \quad (17)$$

where, as usual, the boundary Γ of the domain Ω is decomposed into parts Γ_n and Γ_e associated with natural and essential boundary conditions: $\Gamma = \Gamma_n \cup \Gamma_e$ and $\varnothing = \Gamma_n \cap \Gamma_e$.

A symmetric Hookean stress $\tau_{ij}^{\mathrm{H}} = C_{ijkl}u_{k,l}$ can be identified in terms of which the field equations and natural boundary conditions can be written as

$$\rho(\ddot{u}_i - \ell^2\ddot{u}_{i,jj}) = \tau_{ij,j}^{\mathrm{H}} + b_i \quad \text{in } \Omega, \quad (18a)$$

$$n_j(\tau_{ij}^{\mathrm{H}} + \rho\ell^2\ddot{u}_{i,j}) = t_i \quad \text{on } \Gamma_n. \quad (18b)$$

In our opinion, Hookean stress is appropriate terminology for τ_{ij}^{H}, not Cauchy stress, since the equations of motion and the natural boundary conditions contain additional gradients of the acceleration that are not included in the definition of τ_{ij}^{H}. In Appendix A this particular terminology is motivated.

Remark. A nonsymmetric stress tensor τ_{ij}^{B} can be identified as (see [Lazar and Anastassiadis 2007])

$$\tau_{ij}^{\mathrm{B}} = C_{ijkl}u_{k,l} + \rho\ell^2\ddot{u}_{i,j} \quad (19)$$

This would enable one to write the equations of motion and natural boundary conditions in terms of a stress tensor that is similar in role to a standard Cauchy stress as explained in Appendix A. However, since τ_{ij}^{B} is nonsymmetric, using the term Cauchy stress for this tensor is not obvious. This issue of nomenclature is left for future debate and discussion.

4.2. ***Reducible form.*** For the reducible form, the Lagrangian function adopts a less common appearance, which, to the authors' best knowledge, is novel:

$$L_{\text{red}} = \int_\Omega \tfrac{1}{2}\rho \dot{u}_i \dot{u}_i \,\mathrm{d}V - \int_\Omega u_{i,j}\sigma_{ij}\,\mathrm{d}V + \int_\Omega \tfrac{1}{2}(\sigma_{ij}S_{ijkl}\sigma_{kl} + \ell^2\sigma_{ij,m}S_{ijkl}\sigma_{kl,m})\,\mathrm{d}V + \int_\Omega u_i b_i\,\mathrm{d}V + \int_{\Gamma_n} u_i t_i\,\mathrm{d}S, \quad (20)$$

where S_{ijkl} is the elastic compliance tensor. The first integral is again the kinetic energy, whilst the last two integrals contain the external work. The third integral contains the stored complementary energy with a positive rather than negative sign, but the effects of the lower-order part are offset by the effects of the second integral, which couples the effects of the two sets of unknowns, namely displacements and stresses. In the reducible form, the displacement derivative $u_{i,j}$ is no longer energy-conjugated to the (symmetric) stress σ_{ij}, unless $\ell = 0$. Therefore, the second integrand does *not* have the meaning of internal work. Expression (20) can also be rewritten as a Hellinger–Reissner functional whereby the displacements act as Lagrange multipliers to enforce balance of momentum in Ω and on Γ [Askes and Gutiérrez 2006; Polizzotto 2015].

Again making use of $\delta u_i = 0$ on Γ_e, substitution of (20) into (15) yields

$$\int_{t_0}^{t_1}\int_\Omega \delta u_i(b_i + \sigma_{ij,j} - \rho\ddot{u}_i)\,\mathrm{d}V\,\mathrm{d}t + \int_{t_0}^{t_1}\int_{\Gamma_n} \delta u_i(t_i - n_j\sigma_{ij})\,\mathrm{d}S\,\mathrm{d}t + \int_{t_0}^{t_1}\int_\Omega \delta\sigma_{ij}(-u_{i,j} + S_{ijkl}\sigma_{kl} - \ell^2 S_{ijkl}\sigma_{kl,mm})\,\mathrm{d}V\,\mathrm{d}t + \int_{t_0}^{t_1}\oint_\Gamma \delta\sigma_{ij}n_m S_{ijkl}\sigma_{kl,m}\,\mathrm{d}S\,\mathrm{d}t = 0 \quad (21)$$

so that the following set of coupled governing equations can be identified:

$$\rho\ddot{u}_i = \sigma_{ij,j} + b_i \quad \text{in } \Omega, \quad (22a)$$

$$n_j\sigma_{ij} = t_i \quad \text{on } \Gamma_n, \quad (22b)$$

$$S_{ijkl}(\sigma_{kl} - \ell^2\sigma_{kl,mm}) = \tfrac{1}{2}(u_{i,j} + u_{j,i}) \quad \text{in } \Omega, \quad (22c)$$

$$n_m\ell^2 S_{ijkl}\sigma_{kl,m} = 0 \quad \text{on } \Gamma. \quad (22d)$$

From the format of (22a) and (22b), it is clear that the meaning of σ_{ij} in the reducible model is that of the Cauchy stress. Equations (22) have also been derived, using different arguments, by Eringen [1983]; see Appendix B for a discussion.

Equations (22) form a set of coupled equations with independent unknowns u_i and σ_{ij}, but they are reducible in the sense that it is possible to eliminate the stresses σ_{ij}. To do so, firstly the Laplacian of (22a) is taken and multiplied with ℓ^2, after which the result is subtracted from the original expression (22a). This gives

$$\rho(\ddot{u}_i - \ell^2\ddot{u}_{i,jj}) = \sigma_{ij,j} - \ell^2\sigma_{ij,kk} + b_i - \ell^2 b_{i,jj}. \quad (23)$$

Next, (22c) is premultiplied with the elastic stiffness tensor C_{ijkl} and substituted into (23), leading to

$$\rho(\ddot{u}_i - \ell^2 \ddot{u}_{i,jj}) = C_{ijkl} u_{k,jl} + b_i - \ell^2 b_{i,jj}, \tag{24}$$

which is equivalent to (18a) except for the presence of the Laplacian of the body forces $b_{i,jj}$ and a mismatch in the associated variationally consistent boundary conditions. Note that the effect of the higher-order gradients disappears altogether in statics in the case $b_{i,jj} = 0$.

Remark. From (22c) it is clear that the gradient enrichment affects the constitutive part of the field equations, and therefore the term "gradient elasticity" seems appropriate for what is here denoted as the reducible form. In contrast, it could be argued that using the term "gradient elasticity" is less suitable for the irreducible format represented in (24), because the gradient enrichment operates on the accelerations, not stresses or strains — i.e., the elasticity part of the irreducible form retains its classical format. However, we still prefer to refer to the irreducible form as a particular variant of gradient elasticity, because of the close relation between the reducible form and the irreducible form. Due to the coupling between the equations of motion and the constitutive equations, the gradient enrichment of the accelerations will affect the stresses and strains, albeit indirectly.

5. Finite element equations

In order to obtain solutions of the relevant partial differential equations for domains of arbitrary geometry, a numerical solution strategy is required. Here, the finite element method will be used for the spatial discretisation, whereas the Newmark time integrator will be adopted to progress the solution in the time domain. The finite element equations of the irreducible form are well established and need not be revisited here — the interested reader is referred to [Fish et al. 2002a; 2002b; Askes and Aifantis 2011].

For the reducible form, we write $\underline{u} = \boldsymbol{N}_u \boldsymbol{d}$ and $\underline{\sigma} = \boldsymbol{N}_\sigma \boldsymbol{s}$ where $\underline{u}$ and $\underline{\sigma}$ are column vectors containing the relevant components of the displacements and Cauchy stresses, respectively. Furthermore, the matrices $\boldsymbol{N}_u$ and $\boldsymbol{N}_\sigma$ contain the shape functions for displacements and Cauchy stresses whereas $\boldsymbol{d}$ and $\boldsymbol{s}$ are the nodal displacements and nodal Cauchy stresses. The spatial discretisation of (20) can thus be written as

$$\begin{aligned} L_{\text{red}}^{\text{FE}} &= \int_\Omega \tfrac{1}{2} \rho \dot{\boldsymbol{d}}^T \boldsymbol{N}_u^T \boldsymbol{N}_u \dot{\boldsymbol{d}} \, \mathrm{d}V - \int_\Omega \boldsymbol{d}^T \boldsymbol{B}_u^T \boldsymbol{N}_\sigma \boldsymbol{s} \, \mathrm{d}V \\ &+ \int_\Omega \tfrac{1}{2} \boldsymbol{s}^T \left(\boldsymbol{N}_\sigma^T \boldsymbol{S} \boldsymbol{N}_\sigma + \sum_{i=1}^{3} \ell^2 \frac{\partial \boldsymbol{N}_\sigma^T}{\partial x_i} \boldsymbol{S} \frac{\partial \boldsymbol{N}_\sigma}{\partial x_i} \right) \boldsymbol{s} \, \mathrm{d}V + \int_\Omega \boldsymbol{d}^T \boldsymbol{N}_u^T \underline{b} \, \mathrm{d}V + \int_{\Gamma_n} \boldsymbol{d}^T \boldsymbol{N}_u^T \underline{t} \, \mathrm{d}S \end{aligned} \tag{25}$$

where $\underline{b}$ and $\underline{t}$ contain the components of the distributed body and surface forces, respectively. Furthermore, $\boldsymbol{B}_u$ is the standard strain-displacement matrix with derivatives of the displacement shape functions $\boldsymbol{N}_u$ and $\boldsymbol{S}$ is the matrix counterpart of the compliance tensor S_{ijkl}.

Requiring $\delta L^{\mathrm{FE}}_{\mathrm{red}} = 0$ leads to a system of finite element equations according to

$$\begin{bmatrix} \boldsymbol{M}_{uu} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \ddot{\boldsymbol{d}} \\ \ddot{\boldsymbol{s}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{0} & \boldsymbol{K}_{u\sigma} \\ \boldsymbol{K}_{\sigma u} & \boldsymbol{K}_{\sigma\sigma} \end{bmatrix} \begin{bmatrix} \boldsymbol{d} \\ \boldsymbol{s} \end{bmatrix} = \begin{bmatrix} \boldsymbol{f} \\ \boldsymbol{0} \end{bmatrix}, \tag{26}$$

where

$$\boldsymbol{M}_{uu} = \int_\Omega \rho \boldsymbol{N}_u^T \boldsymbol{N}_u \,\mathrm{d}V, \tag{27a}$$

$$\boldsymbol{K}_{u\sigma} = \boldsymbol{K}_{\sigma u}^T = \int_\Omega \boldsymbol{B}_u^T \boldsymbol{N}_\sigma \,\mathrm{d}V, \tag{27b}$$

$$\boldsymbol{K}_{\sigma\sigma} = -\int_\Omega \left(\boldsymbol{N}_\sigma^T \boldsymbol{S} \boldsymbol{N}_\sigma + \sum_{i=1}^{3} \ell^2 \frac{\partial \boldsymbol{N}_\sigma^T}{\partial x_i} \boldsymbol{S} \frac{\partial \boldsymbol{N}_\sigma}{\partial x_i} \right) \mathrm{d}V. \tag{27c}$$

Finite-element implementation of (26) was carried out using the recommendations of the statics theory given in [Askes and Gutiérrez 2006], in particular the use of quadratic shape functions for $\boldsymbol{s}$ and linear shape functions for $\boldsymbol{d}$. This particular choice of shape functions avoids oscillations in the displacement field, although a formal investigation of the inf-sup condition may require further refinement of the two sets of interpolations.

6. Numerical example

Although the reducible form can be transformed into the irreducible form as shown in (23) and (24), the associated change in variationally consistent boundary conditions has implications when it comes to the simulation of crack tip stresses. This will be demonstrated by means of the numerical example shown in Figure 2.

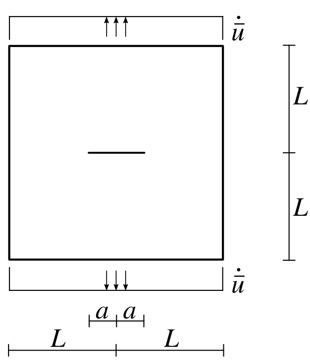

Figure 2. Strip with central crack: geometry and loading conditions.

A square strip with dimension $2L = 2\,\mathrm{m}$ has a central crack of length $2a = 0.5\,\mathrm{m}$. The material properties are mass density $\rho = 1\,\mathrm{kg/m^3}$, Young's modulus $E = 100\,\mathrm{N/m^2}$ and Poisson's ratio $\nu = \frac{1}{4}$, whilst a plane stress assumption has been made. Furthermore, the gradient elasticity length scale $\ell = 0.1\,\mathrm{m}$. The strip is subjected to outward vertical velocities $\dot{\bar{u}} = 10\,\mathrm{m/s}$ imposed on the top and bottom edges, as indicated, which leads to stress waves propagating towards the centre of the strip. Away from the crack, the stress waves will have the shape of a block wave due to the nature of the loading conditions, but the presence of the crack will disturb this pattern, and indeed in a classical elasticity setting, this will lead to singular stresses and strains at the tips of the crack. It is the aim of this example to verify whether these singularities can be avoided in the reducible and irreducible formulations of gradient elasticity discussed above. For reasons of symmetry, only the top quarter of the strip is modelled.

The irreducible format of gradient elasticity is implemented with four-noded quadrilateral elements for the displacements. The reducible format is implemented with eight-noded elements for the stresses and four-noded quadrilateral elements for the displacements — see [Askes and Gutiérrez 2006] for details on this particular choice. Structured finite element meshes consisting of square elements are used, and a sequence of uniformly refined meshes is taken to monitor the behaviour of the stresses at the crack tip. Since in the irreducible format the stresses are postprocessed from linear displacements whereas in the reducible format the stresses are primary unknowns interpolated with quadratic shape functions, there is an obvious mismatch in stress resolution between the two formats. To address this mismatch, the meshes used range from 16×16 to 128×128 elements for the irreducible format, whereas they range from 8×8 to 64×64 for the reducible format.

Regarding the imposition of traction boundary conditions, it must be realised that the stresses are primary variables in the reducible formulation, whereas they are derived quantities in the irreducible formulation. In the reducible formulation, traction boundary conditions are thus essential boundary conditions and are imposed by assigning prescribed values to the relevant stress components (e.g., $\sigma_{yy} = 0$ on the crack face). On the other hand, traction boundary conditions are natural boundary conditions in the irreducible formulation; applying zero tractions on the crack face means that the left-hand-side of (18b) is set equal to zero, which is handled straightforwardly in a finite element context. Finally, and for the sake of completeness, it is noted that displacement (and velocity) boundary conditions have been implemented using Lagrange Multipliers in the reducible formulation.

The Newmark constant average acceleration scheme is used for the time integration. This scheme is unconditionally stable; therefore, the only criterion for selecting the time step is accuracy. Following the recommendations given in [Askes et al. 2008; Bennett and Askes 2009], the time step is chosen such that waves

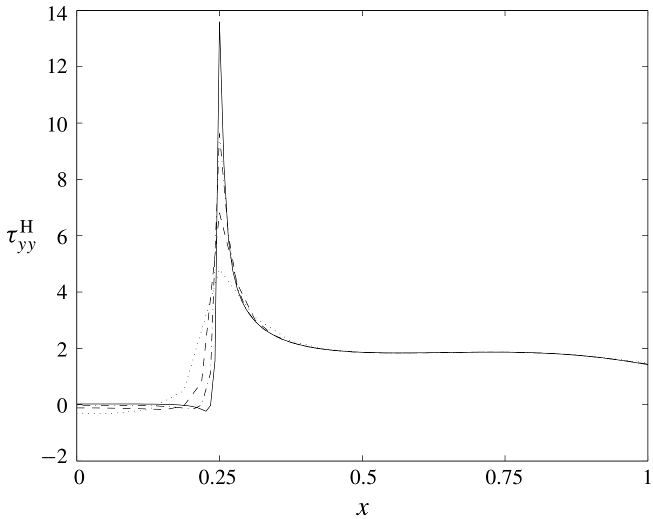

Figure 3. Vertical normal stress τ^{H}_{yy} (N/m^2) versus x (m) for the irreducible format — 16×16 elements (dotted), 32×32 elements (dashed), 64×64 elements (dot-dashed) and 128×128 elements (solid).

propagate approximately half an element per time step. Time domain simulations were carried out from time $t = 0$ s to $t = 0.2$ s.

Figures 3 and 4 show the profiles of the vertical normal stress for both formats and the indicated range of finite element meshes, where the origin of the coordinate system is chosen at the centre of the crack. For the irreducible format (Figure 3), we have plotted the Hookean stress τ^{H}_{yy} (see Section 4.1) whilst for the reducible format the Cauchy stress σ_{yy} is plotted (Figure 4).

The stress profiles for the irreducible formulation appear to converge towards a unique solution, except for the crack tip value. At the crack tip, the stress increases significantly for every refinement of the mesh. This is an indication that a stress singularity is present at the crack tip. To analyse this in more depth, Richardson extrapolations have been carried out for the crack tip stresses. Table 1 reports the

mesh	τ^{H}_{yy}	extrapolation
16×16	4.8077	
32×32	6.8091	8.8105
64×64	9.6223	13.6438
128×128	13.5923	20.0751

Table 1. Crack tip stress and Richardson extrapolation in N/m^2 for irreducible form.

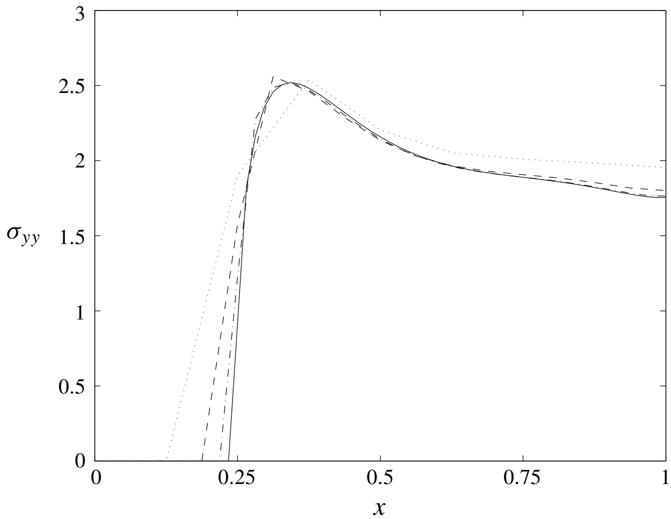

Figure 4. Vertical normal stress σ_{yy} (N/m^2) versus x (m) for the reducible format — 8×8 elements (dotted), 16×16 elements (dashed), 32×32 elements (dot-dashed) and 64×64 elements (solid).

values of the crack tip stress and their extrapolations. (The first extrapolation is a two-point extrapolation based on the coarsest two meshes, the second is a three-point extrapolation based on the coarsest three meshes, and mutatis mutandis for the last extrapolation.) The numerical results confirm that the crack tip stress grows in a seemingly unbounded manner, whereas the difference between numerical stress and extrapolated stress increases with refinement of the mesh. This confirms the suggestion that a singularity is present. Thus, it must be concluded that the irreducible format is not capable of avoiding stress singularities. This is reported for the Hookean stress τ_{yy}^{H} but will carry over to the pseudo Cauchy stress τ_{yy}^{B} since the latter quantity includes the former.

On the other hand, the results of the reducible format clearly converge towards a unique, nonsingular solution, and the singularities that plague classical elasticity formulations are avoided. However, it must be noted that the maximum stress occurs not at the crack tip but further inside the material. This is in line with the analysis and results reported in [Simone et al. 2004].

7. Conclusions

We have reviewed and systematically compared two formats of gradient elasticity. Both formats can be derived by continualising a one-dimensional discrete model and stabilising the resulting equations, but the models differ in respect of which particular equation is stabilised — either the field equation (leading to what is denoted as the "irreducible format") or the constitutive equation (leading to the "reducible format"). The multidimensional case, including the associated boundary conditions, has been derived from a variational principle. It is noted that the field equations of the irreducible format can be retrieved from those of the reducible format (assuming that the Laplacian of the body forces vanishes), but the variationally consistent boundary conditions are different for the two models.

This has implications for the solution of initial-boundary-value problems. We have presented a crack problem, and it was demonstrated that the irreducible format is not capable of avoiding singularities in the stress field. On the other hand, no singularities were found when the reducible format was used. Thus, for the dynamic analysis of stresses around sharp cracks, the reducible format is to be preferred.

Appendix A: Nomenclature in gradient elasticity: Cauchy stress

In the literature, there is a lack of consistency in which quantity is denoted as the Cauchy stress in gradient elasticity theories. Some eminent authors have used this term to indicate the derivative of the strain energy density with respect to the strain — see for instance [Mindlin 1964, p. 57] or [Shu et al. 1999, p. 375]. However, we have followed the arguments set out by Borino and Polizzotto [2003, Remark 3], who state that the term Cauchy stress should be used for the total stress quantity as it appears in the equilibrium equations; conversely, we have used the term Hookean stress for the derivative of the strain energy density with respect to the strain. We believe the former is in line with the conceptualisation of Cauchy himself, who discussed stresses as forming equilibrium (or indeed accelerating) systems by acting on surfaces, rather than as derivatives of energy functionals — see for instance [Cauchy 1823; 1827; 1843].

However, it is also noted that extending the concept of Cauchy stress as "force divided by area" to gradient-enriched continua leads, in general, to much more complicated expressions. This is illustrated by the format of the natural boundary conditions in Mindlin's [1964, pp. 67–68] theory of gradient elasticity. Askes and Metrikine [2005] as well as Froiio et al. [2010] have provided physical interpretations of the nonstandard boundary conditions.

Appendix B: Eringen's 1983 differential theory of nonlocal elasticity

The reducible format presented in Section 4 has been derived earlier in [Eringen 1983] from an integral formulation. Because the coupled nature of the governing equations of Eringen's theory is not always appreciated, it is worthwhile to summarise Eringen's theory. Adopting his notation unless stated otherwise, the equations of motion are given by [Eringen 1983, (2.1)] as

$$t_{kl,k} + \rho(f_l - \ddot{u}_l) = 0 \tag{28}$$

where t_{kl} is the Cauchy stress tensor and f_l is the body force density. With the restriction to isotropic linear elasticity, a Hookean stress σ^0_{kl} is defined via [Eringen 1983, (2.3) and (2.4)] as

$$\sigma^0_{kl} = \lambda \delta_{kl} u_{j,j} + \mu u_{k,l} + \mu u_{l,k} \tag{29}$$

where a superscript 0 is included in σ^0 to avoid confusion with the Cauchy stress of the reducible theory discussed in Section 4.2. Furthermore, λ and μ are the Lamé constants and δ_{kl} is the Kronecker delta.

The field equations are completed by a differential relation between the Cauchy stress t_{kl} and the Hookean stress σ^0_{kl}. The particular relation that seems to have attracted most interest in the literature is given in [Eringen 1983, (3.19)] as

$$t_{kl} - \ell^2 t_{kl,jj} = \sigma^0_{kl} \tag{30}$$

where the higher-order coefficient is simply indicated by ℓ^2 (Eringen uses a more intricate notation with multiple symbols, which are not required in the present discussion).

Eringen [1983, pp. 4704–4705] also discusses the elimination of the stress t_{kl} from the system of equations. Combining (3.13) and (3.18), he arrives at the irreducible form

$$\sigma^0_{kl,k} + (1 - \ell^2 \nabla^2)(\rho f_l - \rho \ddot{u}_l) = 0. \tag{31}$$

Next, he notes that the particular case of statics with vanishing body forces leads to

$$\sigma^0_{kl,k} = 0. \tag{32}$$

However, regarding natural boundary conditions, Eringen [1983, p. 4704] explicitly states that "[b]oundary conditions involving tractions [are] based on the stress tensor t_{kl}, *not* on σ^0_{kl}", while Eringen [2002, p. 100] also emphasises that "the real stress is not σ^0_{kl} but t_{kl}" — in both quotations we have added the superscript 0 to σ as explained above. This means that (32) cannot be used in isolation to solve *general* boundary-value problems involving prescribed tractions.

In summary, in our opinion, a divergence-free Hookean stress σ^0 should not be considered as a fundamental equation of the Eringen theory because, firstly, it can only be retrieved by making the assumptions of zero body force and zero acceleration and, secondly, it cannot be used to solve general equilibrium problems due to a lack of associated traction boundary conditions. In this respect, we disagree with Lazar and Polyzos [2015], who suggest that (32) is an equilibrium equation in its own right — although these authors do confirm that the correct natural boundary conditions are in terms of t_{kl} rather than σ^0_{kl}.

References

[Andrianov 2002] I. V. Andrianov, "The specific features of the limiting transition from a discrete elastic medium to a continuous one", *J. Appl. Math. Mech.* **66**:2 (2002), 261–265.

[Andrianov and Awrejcewicz 2008] I. V. Andrianov and J. Awrejcewicz, "Continuous models for 2D discrete media valid for higher-frequency domain", *Comput. Struct.* **86**:1–2 (2008), 140–144.

[Andrianov et al. 2003] I. V. Andrianov, J. Awrejcewicz, and R. G. Barantsev, "Asymptotic approaches in mechanics: New parameters and procedures", *Appl. Mech. Rev.* **56**:1 (2003), 87–110.

[Andrianov et al. 2010] I. V. Andrianov, J. Awrejcewicz, and D. Weichert, "Improved continuous models for discrete media", *Math. Probl. Eng.* **2010** (2010), 986242.

[Askes and Aifantis 2011] H. Askes and E. C. Aifantis, "Gradient elasticity in statics and dynamics: an overview of formulations, length scale identification procedures, finite element implementations and new results", *Int. J. Solids Struct.* **48**:13 (2011), 1962–1990.

[Askes and Gitman 2014] H. Askes and I. M. Gitman, "A computational mechanics perspective on long-range interactions in gradient elasticity with microinertia", *J. Mech. Behav. Mater.* **23**:1–2 (2014), 37–40.

[Askes and Gutiérrez 2006] H. Askes and M. A. Gutiérrez, "Implicit gradient elasticity", *Internat. J. Numer. Methods Engrg.* **67**:3 (2006), 400–416.

[Askes and Metrikine 2005] H. Askes and A. V. Metrikine, "Higher-order continua derived from discrete media: continualisation aspects and boundary conditions", *Int. J. Solids Struct.* **42**:1 (2005), 187–202.

[Askes et al. 2002] H. Askes, A. S. J. Suiker, and L. J. Sluys, "A classification of higher-order strain-gradient models: linear analysis", *Arch. Appl. Mech.* **72**:2 (2002), 171–188.

[Askes et al. 2008] H. Askes, B. Wang, and T. Bennett, "Element size and time step selection procedures for the numerical analysis of elasticity with higher-order inertia", *J. Sound Vibr.* **314**:3–5 (2008), 650–656.

[Bennett and Askes 2009] T. Bennett and H. Askes, "Finite element modelling of wave dispersion with dynamically consistent gradient elasticity", *Comput. Mech.* **43**:6 (2009), 815–825.

[Bennett et al. 2007] T. Bennett, I. M. Gitman, and H. Askes, "Elasticity theories with higher-order gradients of inertia and stiffness for the modelling of wave dispersion in laminates", *Int. J. Fract.* **148**:2 (2007), 185–193.

[Borino and Polizzotto 2003] G. Borino and C. Polizzotto, "Letter to the editor: 'Higher-order strain/higher-order stress gradient models derived from a discrete microstructure, with application to fracture', by C. S. Chang, H. Askes and L. J. Sluys; Engineering Fracture Mechanics 69 (2002), 1907–1924", *Eng. Fract. Mech.* **70**:9 (2003), 1219–1221.

[Cauchy 1823] A.-L. Cauchy, "Recherches sur l'équilibre et le mouvement intérieur des corps solides ou fluides, élastiques ou non élastiques", *B. Soc. Philomat.* **1823** (1823), 9–13.

[Cauchy 1827] A.-L. Cauchy, "Sur les relations qui existent dans l'état d'équilibre d'un corps solide ou fluide, entre les pressions ou tensions et les forces accélératrices", pp. 108–111 in *Exercices de mathématiques*, vol. 2, Bure Frères, Paris, 1827.

[Cauchy 1843] A.-L. Cauchy, "Note sur les pressions supportées, dans un corps solide ou fluide, par deux portions de surface très voisines, l'une extérieure, l'autre intérieure à ce même corps", *C. R. Hebd. Acad. Sci.* **16** (1843), 151–155.

[Chang and Gao 1995] C. S. Chang and J. Gao, "Second-gradient constitutive theory for granular material with random packing structure", *Int. J. Solids Struct.* **32**:16 (1995), 2279–2293.

[Charlotte and Truskinovsky 2008] M. Charlotte and L. Truskinovsky, "Towards multi-scale continuum elasticity theory", *Contin. Mech. Thermodyn.* **20** (2008), 133–161.

[Chen and Fish 2001] W. Chen and J. Fish, "A dispersive model for wave propagation in periodic heterogeneous media based on homogenization with multiple spatial and temporal scales", *J. Appl. Mech.* **68**:2 (2001), 153–161.

[Eringen 1983] A. C. Eringen, "On differential equations of nonlocal elasticity and solutions of screw dislocation and surface waves", *J. Appl. Phys.* **54**:9 (1983), 4703–4710.

[Eringen 2002] A. C. Eringen, *Nonlocal continuum field theories*, Springer, New York, 2002.

[Fish et al. 2002a] J. Fish, W. Chen, and G. Nagai, "Non-local dispersive model for wave propagation in heterogeneous media: multi-dimensional case", *Internat. J. Numer. Methods Engrg.* **54**:3 (2002), 347–363.

[Fish et al. 2002b] J. Fish, W. Chen, and G. Nagai, "Non-local dispersive model for wave propagation in heterogeneous media: one-dimensional case", *Internat. J. Numer. Methods Engrg.* **54**:3 (2002), 331–346.

[Froiio et al. 2010] F. Froiio, A. Zervos, and I. Vardoulakis, "On natural boundary conditions in linear 2nd-grade elasticity", pp. 211–221 in *Mechanics of generalized continua*, edited by G. A. Maugin and A. V. Metrikine, Adv. Mech. Math. **21**, Springer, New York, 2010.

[Ioannidou et al. 2001] T. Ioannidou, J. Pouget, and E. C. Aifantis, "Kink dynamics in a lattice model with long-range interactions", *J. Phys. A* **34**:20 (2001), 4269–4280.

[Kwong and Gitman 2012] M. T. Kwong and I. M. Gitman, "Gradient elastic stress analysis for anisotropic bimaterial interface with arbitrarily oriented crack", *Int. J. Fract.* **173**:1 (2012), 79–85.

[Lazar and Anastassiadis 2007] M. Lazar and C. Anastassiadis, "Lie point symmetries, conservation and balance laws in linear gradient elastodynamics", *J. Elasticity* **88**:1 (2007), 5–25.

[Lazar and Polyzos 2015] M. Lazar and D. Polyzos, "On non-singular crack fields in Helmholtz type enriched elasticity theories", *Int. J. Solids Struct.* **62** (2015), 1–7.

[Mindlin 1964] R. D. Mindlin, "Micro-structure in linear elasticity", *Arch. Rational Mech. Anal.* **16** (1964), 51–78.

[Mühlhaus and Oka 1996] H. B. Mühlhaus and F. Oka, "Dispersion and wave propagation in discrete and continuous models for granular materials", *Int. J. Solids Struct.* **33**:19 (1996), 2841–2858.

[Pichugin et al. 2008] A. V. Pichugin, H. Askes, and A. Tyas, "Asymptotic equivalence of homogenisation procedures and fine-tuning of continuum theories", *J. Sound Vibr.* **313**:3–5 (2008), 858–874.

[Polizzotto 2012] C. Polizzotto, "A gradient elasticity theory for second-grade materials and higher order inertia", *Int. J. Solids Struct.* **49**:15–16 (2012), 2121–2137.

[Polizzotto 2015] C. Polizzotto, "A unifying variational framework for stress gradient and strain gradient elasticity theories", *Eur. J. Mech. A Solids* **49** (2015), 430–440.

[Rosenau 1984] P. Rosenau, "Dynamics of nonlinear mass-spring chains near the continuum limit", *Phys. Lett. A* **118**:5 (1984), 222–227.

[Rubin et al. 1995] M. B. Rubin, P. Rosenau, and O. Gottlieb, "Continuum model of dispersion caused by an inherent material characteristic length", *J. Appl. Phys.* **77**:8 (1995), 4054–4063.

[Shu et al. 1999] J. Y. Shu, W. E. King, and N. A. Fleck, "Finite elements for materials with strain gradient effects", *Internat. J. Numer. Methods Engrg.* **44**:3 (1999), 373–391.

[Simone et al. 2004] A. Simone, H. Askes, and L. J. Sluys, "Incorrect initiation and propagation of failure in non-local and gradient-enhanced media", *Int. J. Solids Struct.* **41**:2 (2004), 351–363.

[Suiker et al. 2001a] A. S. J. Suiker, R. de Borst, and C. S. Chang, "Micro-mechanical modelling of granular material, I: Derivation of a second-gradient micro-polar constitutive theory", *Acta Mech.* **149**:1 (2001), 161–180.

[Suiker et al. 2001b] A. S. J. Suiker, A. V. Metrikine, and R. de Borst, "Comparison of wave propagation characteristics of the Cosserat continuum and corresponding discrete lattice models", *Int. J. Solids Struct.* **38**:9 (2001), 1563–1583.

[Vardoulakis and Aifantis 1994] I. Vardoulakis and E. C. Aifantis, "On the role of microstructure in the behavior of soils: effects of higher order gradients and internal inertia", *Mech. Mater.* **18**:2 (1994), 151–158.

[Wang and Sun 2002] Z.-P. Wang and C. T. Sun, "Modeling micro-inertia in heterogeneous materials under dynamic loading", *Wave Motion* **36**:4 (2002), 473–485.

Received 9 Mar 2016. Revised 22 Jul 2016. Accepted 26 Sep 2016.

HARM ASKES: h.askes@sheffield.ac.uk
Department of Civil and Structural Engineering, University of Sheffield, Sheffield, S1 3JD, United Kingdom

INNA M. GITMAN: i.gitman@sheffield.ac.uk
Department of Mechanical Engineering, University of Sheffield, Sheffield, S1 3JD, United Kingdom

MATHEMATICS AND MECHANICS OF COMPLEX SYSTEMS
Vol. 5, No. 1, 2017
dx.doi.org/10.2140/memocs.2017.5.19

DATING HYPATIA'S BIRTH: A PROBABILISTIC MODEL

CANIO BENEDETTO, STEFANO ISOLA AND LUCIO RUSSO

We propose a probabilistic approach as a dating methodology for events like the birth of a historical figure. The method is then applied to the controversial birth date of the Alexandrian scientist Hypatia, proving to be surprisingly effective.

1. Introduction

Although in historical investigation it may appear meaningless to do experiments on the basis of a preexisting theory — and in particular, it does not make sense to prove theorems of history — it can make perfect sense to use forms of reasoning typical of the exact sciences as an aid to increase the degree of reliability of a particular statement regarding a historical event. This paper deals with the problem of dating the birth of a historical figure when the only information available about it is indirect — for example, a set of testimonies, or scattered statements, about various aspects of his/her life. The strategy is then based on the construction of a probability distribution for the birth date out of each testimony and subsequently combining the distributions so obtained in a sensible way. One might raise several objections to this program. According to Charles Sanders Peirce [1901], a probability "is the known ratio of frequency of a specific event to a generic event", but a birth is neither a specific event nor a generic event but an "individual event". Nevertheless, probabilistic reasoning is used quite often in situations dealing with events that can be classified as "individual". In probabilistic forecasting, one tries to summarize what is known about future events with the assignment of a probability to each of a number of different outcomes that are often events of this kind. For instance, in sport betting, a summary of bettors' opinions about the likely outcome of a race

Communicated by Francesco dell'Isola.
MSC2010: 01A20, 62C05, 62P99.
Keywords: dating, probabilistic method, historical testimonies, decision theory, Hypatia.

is produced in order to set bookmakers' pay-off rates. By the way, this type of observation lies at the basis of the theoretical formulation of the subjective approach in probability theory [de Finetti 1931]. Although we do not endorse de Finetti's approach in all its implications, we embrace its severe criticism of the exclusive use of the frequentist interpretation in the application of probability theory to concrete problems. In particular, we feel entitled to look at an "individual" event of the historical past with a spirit similar to that with which one bets on a future outcome (this is a well known issue in the philosophy of probability; see, e.g., [Dubucs 1993]). Plainly, as the information about an event like the birth of an historical figure is first extracted by material drawn from various literary sources and then treated with mathematical tools, both our approach and goal are interdisciplinary in their essence.

2. A probabilistic method for combining testimonies

Let $X = [x_-, x_+] \subset \mathbb{Z}$ be the time interval that includes all possible birth dates of a given subject (*terminus ad quem*). X can be regarded as a set of mutually exclusive statements about a singular phenomenon (the birth of a given subject in a given year), only one of which is true, and can be made a probability space $(X, \mathcal{F}, P_0)$, with $\mathcal{F}$ the σ-algebra made of the $2^{|X|}$ events of interest and P_0 the uniform probability measure on $\mathcal{F}$ (reference measure): $P_0(A) = |A|/|X|$ (where $|A|$ denotes the number of elements of A). In the context of decision theory, the assignment of this probability space can be regarded as the expression of a basic state of knowledge, in the absence of any information that can be used to discriminate among the possible statements on the given phenomenon, namely a situation in which Laplace's *principle of indifference* can be legitimately applied.

Now suppose we have k testimonies T_i, $i = 1, \dots, k$, which in first approximation we may assume independent of each other, each providing some kind of information about the life of the subject, and which can be translated into a probability distribution p_i on $\mathcal{F}$ so that $p_i(x)$ is the probability that the subject is born in the year $x \in X$ based on the information given by the testimony T_i, assumed true, along with supplementary information such as, e.g., life tables for the historical period considered. The precise criteria for the construction of these probability distributions depends on the kind of information carried by each testimony and will be discussed case by case in the next section. Of course, we shall also take into account the possibility that some testimonies are false, thereby not producing any additional information. We model this possibility by assuming that the corresponding distributions equal the reference measure P_0.

The problem that we want to discuss in this section is the following: how can one combine the distributions p_i in such a way to get a single probability distribution Q

that somehow optimizes the available information? To address this question, let us observe that from the k testimonies taken together, each one with the possibility to be true or false, one gets $N = 2^k$ combinations, corresponding to as many binary words $\sigma_s = \sigma_s(1) \cdots \sigma_s(k) \in \{0, 1\}^k$, which can be ordered lexicographically according to $s = \sum_{i=1}^k \sigma_s(i) \cdot 2^{i-1} \in \{0, 1, \ldots, N-1\}$, and given by

$$P_s(\cdot) = \frac{\prod_{i=1}^k p_i^{\sigma_s(i)}(\cdot)}{\sum_{x\in X} \prod_{i=1}^k p_i^{\sigma_s(i)}(x)}, \qquad p_i^{\sigma_s(i)} = \begin{cases} p_i, & \sigma_s(i) = 1, \\ P_0, & \sigma_s(i) = 0. \end{cases} \tag{2-1}$$

In particular, one readily verifies that P_0 is but the reference uniform measure.

Now, if Ω denotes the class of probability distributions $Q : X \to [0, 1]$, we look for a *pooling operator* $T : \Omega^N \to \Omega$ that combines the distributions P_s by weighing them in a sensible way. The simplest candidate has the general form of a linear combination

$$T(P_0, \ldots, P_{N-1}) = \sum_{s=0}^{N-1} w_s P_s, \quad w_s \geq 0, \ \sum_{s=0}^{N-1} w_s = 1, \tag{2-2}$$

which, as we shall see, can also be obtained by minimizing some information-theoretic function.

Remark 2.1. The issue we are discussing here has been the object of a vast amount of literature regarding the normative aspects of the formation of aggregate opinions in several contexts (see, e.g., [Genest and Zidek 1986] and references therein). In particular, it has been shown by McConway [1981] that, if one requires the existence of a function $F : [0, 1]^N \to [0, 1]$ such that

$$T(P_0, \ldots, P_{N-1})(A) = F(P_0(A), \ldots, P_{N-1}(A)) \quad \text{for all } A \in \mathcal{F} \tag{2-3}$$

with $P_s(A) = \sum_{x\in A} P_s(x)$, then whenever $|X| \geq 3$, F must necessarily have the form of a linear combination as in (2-2). The above condition implies in particular that the value of the combined distribution on coordinates depends only on the corresponding values on the coordinates of the distributions P_s, namely that the pooling operator commutes with marginalization.

However, some drawbacks of the linear pooling operator have also been highlighted. For example, it does not "preserve independence" in general: if $|X| \geq 5$, it is not true that $P_s(A \cap B) = P_s(A)P_s(B)$, $s = 0, \ldots, N-1$, entails

$$T(P_0, \ldots, P_{N-1})(A \cap B) = T(P_0, \ldots, P_{N-1})(A)T(P_0, \ldots, P_{N-1})(B)$$

unless $w_s = 1$ for some s and 0 for all others [Lehrer and Wagner 1983; Genest and Wagner 1987].

(Another form of the pooling operator considered in the literature to overcome the difficulties associated with the use of (2-2) is the log-linear combination

$$T(P_0, \ldots, P_{N-1}) = C \prod_{s=0}^{N-1} P_s^{w_s}, \quad w_s \geq 0, \sum_{s=0}^{N-1} w_s = 1, \tag{2-4}$$

where C is a normalizing constant [Genest and Zidek 1986; Abbas 2009].)

On the other hand, in our context, the independence preservation property does not seem so desirable: the final distribution $T(P_0, \ldots, P_{N-1})$ relies on a set of information much wider than that associated with the single distributions P_s, and one can easily imagine how the alleged independence between two events can disappear as the information about them increases.

2.1. ***Optimization.*** The linear combination (2-2) can also be viewed as the marginal distribution[1] of $x \in X$ under the hypothesis that one of the distributions $P_0, \ldots, P_{N-1}$ is the "true" one (without knowing which) [Genest and McConway 1990]. In this perspective, (2-2) can be obtained by minimizing the expected loss of information due to the need to compromise, namely a function of the form

$$I(w, Q) = \sum_{s=0}^{N-1} w_s \, D(P_s \parallel Q) \geq 0, \tag{2-5}$$

where

$$D(P \parallel Q) = \sum_{x \in X} P(x) \log\left(\frac{P(x)}{Q(x)}\right) \tag{2-6}$$

is the *Kullback–Leibler divergence* [1951], representing the information loss using the measure Q instead of P. Note that the concavity of the logarithm and the Jensen inequality yield

$$-\sum_x P(x) \log \frac{P(x)}{Q(x)} \leq \log \sum_x P(x) \frac{Q(x)}{P(x)} = 0$$

and therefore

$$D(P \parallel Q) \geq 0 \quad \text{and} \quad D(P \parallel Q) = 0 \iff Q \equiv P. \tag{2-7}$$

We have the following result.

Lemma 2.2. *Given a probability vector $w = (w_0, w_1, \ldots, w_{N-1})$,*

$$\arg\min_{Q \in \Omega} I(w, Q) = Q_w \equiv \sum_s w_s P_s. \tag{2-8}$$

Moreover,

$$I(w, Q_w) = H\left(\sum_s w_s P_s\right) - \sum_s w_s H(P_s), \tag{2-9}$$

where $H(Q) = -\sum_{x \in X} Q(x) \log Q(x)$ is the entropy of $Q \in \Omega$.

[1] In the sense that a marginal probability can be obtained by averaging conditional probabilities.

Proof. Equation (2-8) can be obtained using the method of Lagrange multipliers. An alternative argument makes use of the easily derived "parallelogram rule":

$$\sum_s w_s D(P_s \parallel Q) = \sum_s w_s D(P_s \parallel Q_w) + D(Q_w \parallel Q) \quad \text{for all } Q \in \Omega. \tag{2-10}$$

From (2-7), we thus get $I(w, Q_w) \le I(w, Q)$ for all $Q \in \Omega$. The uniqueness of the minimum follows from the convexity of $D(P \parallel Q)$ with respect to Q. Finally, checking (2-9) is a simple exercise. □

Remark 2.3. It is worth mentioning that, if we took $\sum_s w_s D(Q \parallel P_s)$ (instead of $\sum_s w_s D(P_s \parallel Q)$) as the function to be minimized (still varying Q with w fixed), then instead of the "arithmetic mean" (2-2), the "optimal" distribution would have been the "geometric mean" (2-4) (see also [Abbas 2009]).

2.2. ***Allocating the weights.*** We have seen that for each probability vector w in the N-dimensional simplex $\{w_s \ge 0 : \sum_{s=0}^{N-1} w_s = 1\}$ the distribution $Q_w = \sum_s w_s P_s$ is the "optimal" one. We are now left with the problem of determining a sensible choice for w. This cannot be achieved by using the same criterion, in that by (2-7) $\inf_w I(w, Q_w) = 0$ and the minimum is realized whenever $w_s = 1$ for some s and 0 for all others.

A suitable expression for the weights w_s can be obtained by observing that the term $\sum_{x \in X} \prod_{i=1}^k p_i^{\sigma_s(i)}(x)$ is proportional to the probability of the event (in the product space $X^{[1,k]}$) that the birth dates of k different subjects, with the i-th birth date distributed according to $p_i^{\sigma_s(i)}$, coincide, and thus, it furnishes a measure of the degree of compatibility of the distributions p_i involved in the product associated with the word σ_s.

It thus appears natural to consider the weights

$$w_s = \frac{\sum_{x \in X} \prod_{i=1}^k p_i^{\sigma_s(i)}(x)}{\sum_{s=0}^{N-1} \sum_{x \in X} \prod_{i=1}^k p_i^{\sigma_s(i)}(x)}, \tag{2-11}$$

which, once inserted in (2-2), yield the expression

$$T(P_0, \dots, P_{N-1})(\cdot) = \frac{\sum_{s=0}^{N-1} \prod_{i=1}^k p_i^{\sigma_s(i)}(\cdot)}{\sum_{x \in X} \sum_{s=0}^{N-1} \prod_{i=1}^k p_i^{\sigma_s(i)}(x)}. \tag{2-12}$$

Remark 2.4. There are at least $k+1$ strictly positive coefficients w_s. They correspond to the words $\sigma_s^{(i)}$ with $\sigma_s^{(i)}(i) = 1$ for some $i \in \{1, \dots, k\}$ and $\sigma_s^{(i)}(j) = 0$ for $j \ne i$, plus one to the word 0^k, that is, to the distributions $P_{s^{(i)}} \equiv p_i$, $i \in \{0, 1, \dots, k\}$, where $p_0 \equiv P_0$.

2.3. ***Weights as likelihoods.*** A somewhat complementary argument to justify the choice (2-11) for the coefficients w_s can be formulated in the language of probabilistic inference, showing that they can be interpreted as (normalized) *average* *likelihoods* associated with the various combinations corresponding to the words σ_s. More precisely, with each pair of "hypotheses" of the form

$$D_i^e = \begin{cases} \{T_i \text{ true}\}, & e = 1, \\ \{T_i \text{ false}\}, & e = 0, \end{cases}$$

we associate its *likelihood, given the event that the birth date is* $x \in X$, with the expression[2]

$$V(D_i^e \mid x) = \frac{P(x \mid D_i^e)}{P(x)} = \begin{cases} p_i(x)/p_0(x), & e = 1, \\ 1, & e = 0, \end{cases} \tag{2-13}$$

with $i \in \{1, \dots, k\}$ and $p_0 \equiv P_0$. In this way, the posterior probability $P(D_i^e \mid x)$ (the probability of D_i^e in light of the event that the subject was born in the year $x \in X$) is given by the product of $V(D_i^e \mid x)$ with the prior probability $P(D_i^e)$, according to Bayes's formula.

If we now consider two pairs of "hypotheses" $D_i^{e_i}$ and $D_j^{e_j}$, which we assume conditionally independent (without being necessarily independent), that is,

$$P(D_i^{e_i}, D_j^{e_j} \mid x) = P(D_i^{e_i} \mid x) P(D_j^{e_j} \mid x), \quad e_i, e_j \in \{0, 1\},$$

then we find

$$\begin{aligned} P(D_i^{e_i}, D_j^{e_j} \mid x) &= \frac{P(x \mid D_i^{e_i}, D_j^{e_j})}{P(x)} = \frac{P(D_i^{e_i}, D_j^{e_j} \mid x)}{P(D_i^{e_i}, D_j^{e_j})} = \frac{P(D_i^{e_i} \mid x) P(D_j^{e_j} \mid x)}{P(D_i^{e_i}, D_j^{e_j})} \\ &= \frac{P(D_i^{e_i}) P(D_j^{e_j})}{P(D_i^{e_i}, D_j^{e_j})} \cdot V(D_i^{e_i} \mid x) V(D_j^{e_j} \mid x). \end{aligned}$$

More generally, given k testimonies T_i, to each of which there corresponds the pair of events D_i^e, and given a word $\sigma_s \in \{0, 1\}^k$, if we assume the conditional independence of the events $(D_1^{\sigma_s(1)}, \dots, D_k^{\sigma_s(k)})$, we get

$$V(D_1^{\sigma_s(1)}, \dots, D_k^{\sigma_s(k)} \mid x) = \rho_s \prod_{i=1}^{k} V(D_i^{\sigma_s(i)} \mid x) \tag{2-14}$$

where

$$\rho_s = \frac{\prod_{i=1}^{k} P(D_i^{\sigma_s(i)})}{P(D_1^{\sigma_s(1)}, \dots, D_k^{\sigma_s(k)})}. \tag{2-15}$$

If, in addition, there is grounds to assume unconditional independence, i.e., $\rho_s = 1$, then (2-14) simply reduces to the product rule. Under this assumption, we can

[2]Here the symbol P denotes either the reference measure P_0 or any probability measure on X compatible with it.

evaluate the *average likelihood* of the set of information $(D_1^{\sigma_s(1)}, \dots, D_k^{\sigma_s(k)})$ with the expression

$$V_s = \frac{1}{|X|} \sum_{x \in X} V(D_1^{\sigma_s(1)}, \dots, D_k^{\sigma_s(k)} \mid x) = |X|^{k-1} \sum_{x \in X} \prod_{i=1}^{k} p_i^{\sigma_s(i)}(x). \tag{2-16}$$

Comparing with (2-11), we see that

$$w_s = \frac{V_s}{\sum_{s=0}^{N-1} V_s}. \tag{2-17}$$

In other words, within the hypotheses made so far, the allocation of the coefficients (2-11) corresponds to assigning to each distribution P_s a weight proportional to the average likelihood of the set of information from which it is constructed.

3. Application to Hypatia

This method is now applied to a particular dating process, the one of Hypatia's birth. This choice stems from the desire to study a case both easy to handle and potentially useful in its results. The problem of dating Hypatia's birth is indeed open, in that there are different possible resolutions of the constraints imposed by the available data. According to the reconstruction given by Deakin [2007, p. 51], "Hypatia's birth has been placed as early as 350 and as late as 375. Most authors settle for 'around 370'". There are not many testimonies (historical records) concerning the birth of the Alexandrian scientist (far more are about her infamous death), but they have the desirable feature of being independent of one another, as will be apparent in the sequel, so that the scheme discussed in the previous section can be directly applied. The hope is to obtain something that is qualitatively significant when compared to the preexisting proposals, based on a qualitative discussion of the sources, and quantitatively unambiguous. A probability distribution for the year of Hypatia's birth is extracted from each testimony, the specific reasoning being briefly discussed in each case. Eventually all distributions are combined according to the criteria outlined in the previous section.

3.1. ***Hypatia was at her peak between 395 and 408.*** Under the entry Ὑπατία, the Suda (a Byzantine lexicon) informs us that she flourished under the emperor Arcadius (ἤκμασεν ἐπί τῆς βασιλείας Ἀρκαδίου).[3]

It is well established that Arcadius, the first ruler of the Byzantine Empire, reigned from 395 to 408. Guessing an age or age interval based on the Greek ἤκμασεν, however, is less straightforward. The word is related to ἀκμή, 'peak',

[3] Υ166. See http://www.stoa.org/sol-bin/search.pl?field=adlerhw_gr&searchstr=upsilon,166.

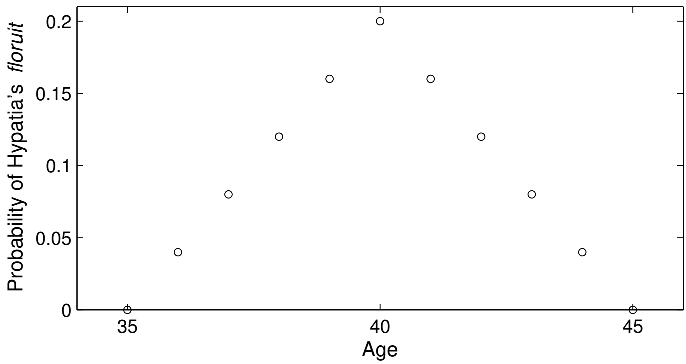

Figure 1. The probability distribution $f(x)$ assumed associated with one's peak years.

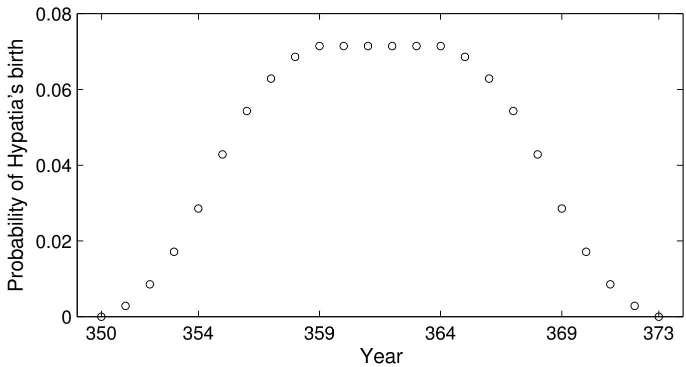

Figure 2. The probability distribution $\Upsilon_f(\xi)$ for Hypatia's birth based on her peak years.

and we follow the rule of thumb, going back to Antiquity, that it refers to the period of one's life around 40 years of age. Specifically, we adopt the probability distribution $f(x)$ in Figure 1 to model how old Hypatia would have been at her "peak" in Arcadius' reign.

Figure 2 shows $\Upsilon_f(\xi)$, the probability distribution for the year of Hypatia's birth deduced from this historical datum; it is obtained by averaging fourteen copies of the triangular $f(x)$, each centered around one of the years from 355 through 368 — the beginning and end points of Arcadius's empire, shifted back by the 40 years corresponding to the peak of $f(x)$.

3.2. ***Hypatia was intellectually active in 415.*** The sources ascribe Hypatia's martyrdom at the hands of a mob of Christian fanatics to the envy that many felt on account of her extraordinary intelligence, freedom of thought, and political influence, being a woman. Her entry in the Suda, already mentioned, states:

> Τοῦτο δὲ πέπονθε διὰ φθόνον καὶ τὴν ὑπερβάλλουσαν σοφίαν, καὶ μάλιστα εἰς τὰ περὶ ἀστρονομίαν.[4]

Socrates Scholasticus, in his Εκκλησιαστική Ιστορία, reports:

> On account of the self-possession and ease of manner, which she had acquired in consequence of the cultivation of her mind, she not infrequently appeared in public in presence of the magistrates. Neither did she feel abashed in coming to an assembly of men. For all men on account of her extraordinary dignity and virtue admired her the more. Yet even she fell a victim to the political jealousy which at that time prevailed. For as she had frequent interviews with Orestes, it was calumniously reported among the Christian populace, that it was she who prevented Orestes from being reconciled to the bishop.[5]

Because of these and similar testimonies, it seems reasonable to mark 415 as a year of intellectual activity in Hypatia's life.

To get from this information a probability distribution for the year of birth, it is necessary to have the probability distribution of being intellectually active at a given age. This can be calculated given the probability of being alive at any given age and of being active at any given age (if alive), by simple multiplication.

To derive the first of these probability distributions we have used data from a 1974 mortality table for Italian males,[6] clipping off ages under 18 since the subject was known to be intellectually active. The resulting probability distribution, $a(x)$, is shown in Figure 3.

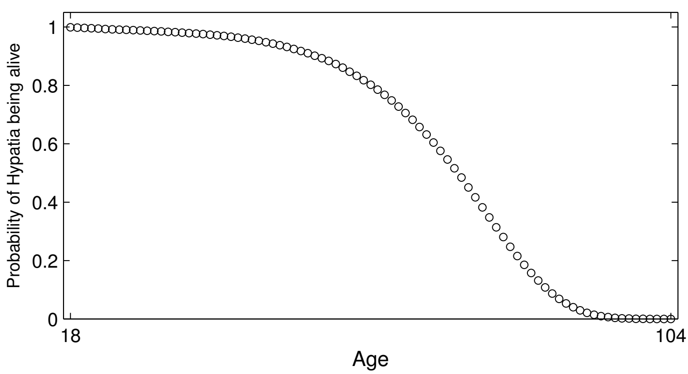

Figure 3. The probability distribution $a(x)$ for an adult to reach a given age. The life expectancy comes to 71.8 years.

[4]She suffered this [violent death] because of the envy for her extraordinary wisdom, especially in the field of astronomy.

[5]Book VII, Chapter 15; translation from [Socrates Scholasticus, p. 160].

[6]All data are taken from http://www.mortality.org.

Name	Dates of birth and death	Lifespan
Accius, Lucius	170–circa 86 BC	~84
Adrianus (Hadrianus) of Tyre	circa AD 113–193	~80
Aelian (Claudius Aelianus)	AD 165/170–230/235	~65
Aeschines	circa 397–circa 322 BC	~65
Aeschylus	524/525–456/455 BC	~70
Agathocles (2) (of Cyzicus)	circa 275/265–circa 200/190 BC	~75
Alexander of Tralles	AD 525–605	80
Alexis	circa 375–circa 275 BC	~100
Ammianus Marcellinus	circa AD 330–395	~65
Anaxagoras	probably 500–428 BC	~72
Anaximenes (2) of Lampsacus	circa 380–320 BC	~60
Andocides	circa 440–circa 390 BC	~50
Androtion	circa 410–340 BC	~70
Antiphon	circa 480–411 BC	~69
Apollonius of Citium	circa 90–15 BC?	~75
Arcesilaus	316/315–242/241 BC	~74
Aristarchus of Samothrace	circa 216–144 BC	~72
Aristophanes of Byzantium	circa 257–180 BC	~77
Aristotle	384–322 BC	62
Arius	circa AD 260–336	~76
Arrian (Lucius Flavius Arrianus)	circa AD 86–160	~74
Aspasius	circa AD 100–150	~50
Athanasius	circa AD 295–373	~78
Atticus	circa AD 150–200	~50
Augustine, Saint	AD 354–430	76
Bacchius of Tanagra	probably 275–200 BC	~75
Bacchylides	circa 520–450 BC	~70
Basil of Caesarea	circa AD 330–379	~49
Bion of Borysthenes	circa 335–circa 245 BC	~90
Carneades	214/213–129/128 BC	~85
Cassius (1)	31 BC–AD 37	68
Cassius Longinus	circa AD 213–273	~60
Cato (Censorius)	234–149 BC	85
Chrysippus of Soli	circa 280–207 BC	~73
Chrysostom, John	circa AD 354–407	~53
Cinesias	circa 450–390 BC	~60
Claudius Atticus Herodes (2) Tiberius	circa AD 101–177	~76
Cleanthes of Assos	331–232 BC	99
Clitomachus	187/186–110/119 BC	~77
Colotes (RE 1) of Lampsacus	circa 325–260 BC	~65
Cornelius (RE 157) Fronto, Marcus	circa AD 95–circa 166	~71
Crantor of Soli in Cilicia	circa 335–275 BC	~60
Crates (2)	circa 368/365–288/285 BC	~80
Demades	circa 380–319 BC	~61
Demochares	circa 360–275 BC	~85
Democritus (of Abdera)	circa 460–370 BC	~90
Demosthenes (2)	384–322 BC	62
Dinarchus	circa 360–circa 290 BC	~70
Dio Cocceianus	circa 40/50–110/120 BC	~70
Diodorus (3) of Agyrium, Sicily	circa 90–30 BC	~60
Diogenes (3) (of Babylon)	circa 240–152 BC	~88

Diogenes (2) the Cynic	circa 412/403–circa 324/321 BC	~85
Duris	circa 340–circa 260 BC	~80
Empedocles	circa 492–432 BC	~60
Ennius, Quintus	239–169 BC	70
Ennodius, Magnus Felix	AD 473/474–521	~48
Ephorus of Cyme	circa 405–330 BC	~75
Epicurus	341–270 BC	71
Epiphanius	circa AD 315–403	~88
Erasistratus	circa 315–240 BC	~75
Eratosthenes of Cyrene	circa 285–194 BC	~91
Eubulus (1)	circa 405–circa 335 BC	~70
Euclides (1) of Megara	circa 450–380 BC	~70
Euripides	probably 480s–407/406 BC	~78
Eusebius of Caesarea	circa AD 260–339	~79
Evagrius Scholasticus	circa AD 535–circa 600	~65
Favorinus	circa AD 85–155	~70
Fenestella	52 BC–AD 19 or 35 BC–AD 36	71
Galen of Pergamum	AD 129–216	87
Gorgias (1) of Leontini	circa 485–circa 380 BC	~105
Gregory (2) of Nazianzus	AD 329–389	60
Gregory (3) of Nyssa	circa AD 330–395	~65
Gregory (4) Thaumaturgus	circa AD 213–circa 275	~62
Hecataeus (2) of Abdera	circa 360–290 BC	~70
Hegesippus (1)	circa 390–circa 325 BC	~65
Hellanicus (1) of Lesbos	circa 480–395 BC	~85
Hellanicus (2)	circa 230/220–160/150 BC	~70
Herophilus of Chalcedon	circa 330–260 BC	~70
Hieronymus (2) of Rhodes	circa 290–230 BC	~60
Himerius	circa AD 310–circa 390	~80
Horace (Quintus Horatius Flaccus)	65–8 BC	57
Idomeneus (2)	circa 325–circa 270 BC	~55
Irenaeus	circa AD 130–circa 202	~72
Isaeus (1)	circa 420–340s BC	~75
Isocrates	436–338 BC	98
Ister	circa 250–200 BC	~50
Jerome (Eusebius Hieronymus)	circa AD 347–420	~73
Laberius, Decimus	circa 106–43 BC	~63
Libanius	AD 314–circa 393	~63
Livius Andronicus, Lucius	circa 280/270–200 BC	~75
Livy (Titus Livius)	59 BC–AD 17 or 64 BC–AD 12	76
Lucilius (1) Gaius	probably 180–102/101 BC	~75
Lucretius (Titus Lucretius Carus)	circa 94–55/51 BC	~41
Lyco	circa 300/298–226/224 BC	~74
Lycurgus (3)	circa 390–circa 325/324 BC	~65
Lydus	AD 490–circa 560	~70
Lysias	459/458–circa 380 BC or circa 445–circa 380 BC	~72
Malalas	circa AD 480–circa 570	~90
Mantias	circa 165–85 BC	~80
Megasthenes	circa 350–290 BC	~60

Table 1. Life spans of the first 100 "ancient intellectuals" in *The Oxford Classical Dictionary*. The average, 71.7 years, is taken as typical.

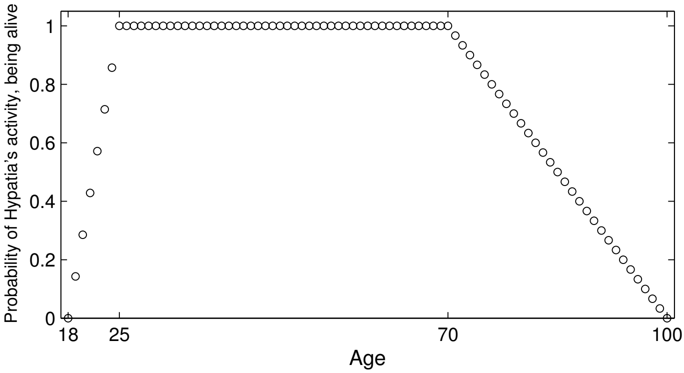

Figure 4. The probability distribution $a_a(x)$ for being active at a given age, if alive.

The choice made for this distribution might appear questionable on two grounds: Is it appropriate to use modern data in studying an Alexandrian scholar of the fourth century AD? And assuming this is so, is the particular mortality table chosen adequate?

Our chief justification for keeping this choice of $a(x)$ is that its most important feature for our purposes, the life expectancy, is in excellent agreement with a control value calculated for this purpose: the average lifespan of the first one hundred (in alphabetical order) "well dated" intellectuals found in *The Oxford Classical Dictionary* [Hornblower et al. 2012][7] (see Table 1). This suggests that using $a(x)$ as an approximation for the mortality distribution of the population of interest is consistent with the available quantitative evidence.

To model the probability $a_a(x)$ of being intellectually active at a given age if alive at that age we make some reasonable, if somewhat arbitrary, assumptions reflected in the graph in Figure 4.

Combining the two distributions $a(x)$ and $a_a(x)$ as explained, the probability of being active at any given age is calculated and — knowing that Hypatia was so in 415 — the probability distribution $\Upsilon_a(\xi)$ for the year of Hypatia's birth deduced from this historical datum is obtained in a straightforward manner (see Figure 5).

3.3. ***Hypatia reached old age.*** In his Χρονογραφία, John Malalas tells us that our subject was an old woman when she died:

> Κατ᾽ ἐκεῖνον δὲ τὸν καιρὸν παρρησίαν λαβόντες ὑπὸ τοῦ ἐπισκόπου οἱ Ἀλεξανδρεῖς ἔκαυσαν φρυγάνοις αὐθεντήσαντες Ὑπατίαν τὴν περι-

[7]The cutoff at 100 gives a convenient sample size large enough to be representative. Using all "ancient intellectuals" as the control population and not only those who lived in the third and fourth centuries AD is necessary in order to obtain a statistically significant sample.

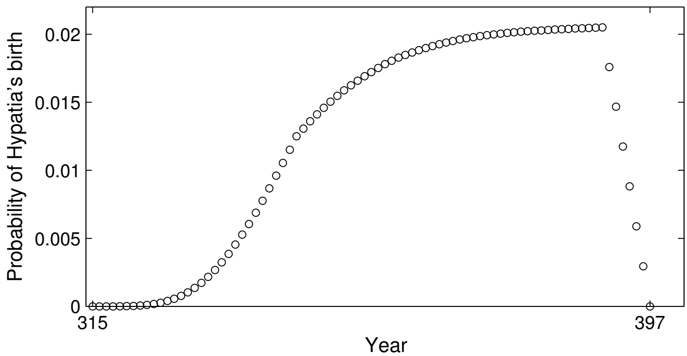

Figure 5. The probability distribution $\Upsilon_a(\xi)$ for Hypatia's birth based on her being active when she died.

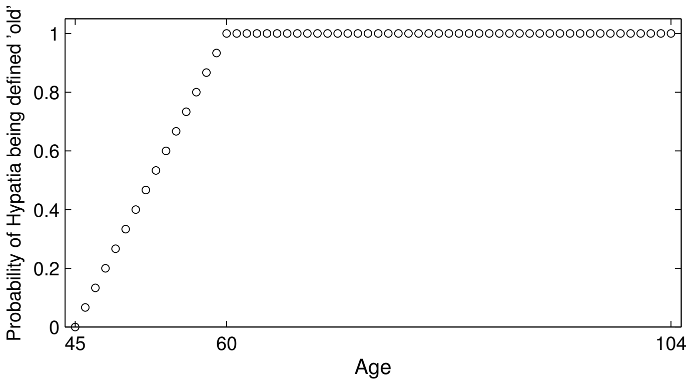

Figure 6. The probability distribution $o(x)$ for being regarded as an old woman.

βόητον φιλόσοφον, περὶ ἧς μεγάλα ἐφέρετο ἦν δὲ παλαιὰ γυνή.[8]

In light of the average lifespan of ancient intellectuals (Table 1), even a conservative interpretation of "old woman" would preclude an age much below 50.[9] Hence we model the probability distribution of someone being "old woman" by the function $o(x)$ shown in Figure 6. The resulting probability distribution, $\Upsilon_o(\xi)$, for the year of Hypatia's birth based on this datum is then easily obtained; see Figure 7.

[8] At that time the Alexandrians, given free rein by their bishop, seized and burnt on a pyre of brushwood Hypatia the famous philosopher, who had a great reputation and who was an old woman [Malalas, XIV.12].

[9] This agrees with the authoritative opinion of many historians; thus Maria Dzielska [1995]: "John Malalas argues persuasively that at the time of her ghastly death Hypatia was an elderly woman — not twenty-five years old (as Kingsley wants), nor even forty-five, as popularly assumed. Following Malalas, some scholars, including Wolf, correctly argue that Hypatia was born around 355 and was about sixty when she died".

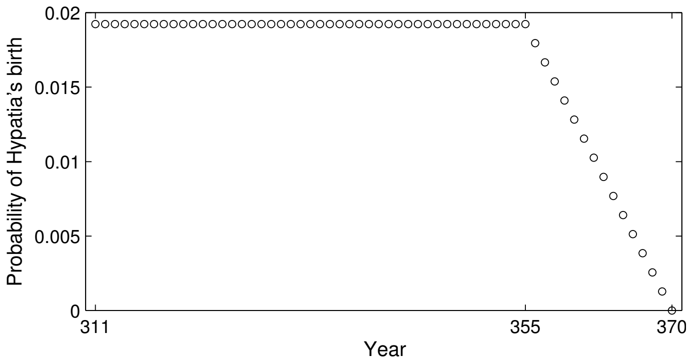

Figure 7. The probability distribution $\Upsilon_o(\xi)$ for Hypatia's birth given that she reached old age.

3.4. ***Hypatia, daughter of Theon.*** Theon of Alexandria, best known for allowing the transmission of Euclid's *Elements* to the present day, was Hypatia's father. By knowing his birth year, one might think of deducing a probability distribution for the year of Hypatia's birth; sadly, this is unknown as well. Therefore, it is necessary to calculate a probability distribution for the year of Theon's birth first. To this end, two recorded facts are useful:

- Theon was intellectually active between 364 and 377.[10]
- Hypatia overhauled the third book of Theon's *Commentary on the Almagest* (Theon refers to this in the *Commentary* itself).

This second datum makes it unlikely that Hypatia was born in Theon's old age; it also make it less probable that he stopped being intellectually active at a young age, since he was still active while his daughter made her contribution to his work. To quantify this reasoning, we define notation for the relevant events:

- F_i, Theon becomes a father at age i.
- $A_i^{T/I}$, Theon/Hypatia is intellectually active at the age of i.
- C, Theon is able to collaborate with Hypatia (both are intellectually active).
- $B_k^{T/I}$, Theon/Hypatia begins being intellectually active at age k.
- $S_k^{T/I}$, Theon/Hypatia stops being intellectually active at age k.

The probability of Theon becoming a father at various ages is described approximately by the model distribution $F(x)$ shown in Figure 8.

[10] In the *Little Commentary on Ptolemy's Handy Tables*, Theon mentioned some astronomical observations that can be dated with certainty: the two solar eclipses of June 15th and November 26th, 364 and an astral conjunction in 377. It is reasonable to assume that he was also active in the interval between those two years.

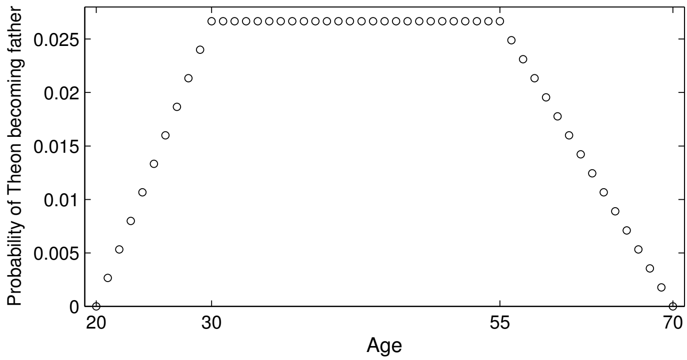

Figure 8. The probability distribution $F(x)$ for Theon's age at the time of Hypatia's birth.

The probability of a subject (Theon or Hypatia) beginning their intellectual activity at a given age is described approximately by the model distribution $B(x)$ shown in Figure 9.

The probability distribution $S(x)$ for the subject ending her intellectual activity at a given age is taken to be, up to age 70, just the probability of dying (derived from the distribution $a(x)$ of Figure 3), while after that it is the probabily of dying conditioned to that of being active, as obtained in Section 3.2. See Figure 10.

The probability of event C is therefore

$$P(C) = \sum_i \sum_k P(A^T_{i+k} \cap F_i \cap B^I_k).$$

By the definition of conditional probability,

$$\sum_i \sum_k P(A^T_{i+k} \cap F_i \cap B^I_k) = \sum_i \sum_k P(A^T_{i+k} \cap I^I_k \mid F_i) \cdot P(F_i),$$

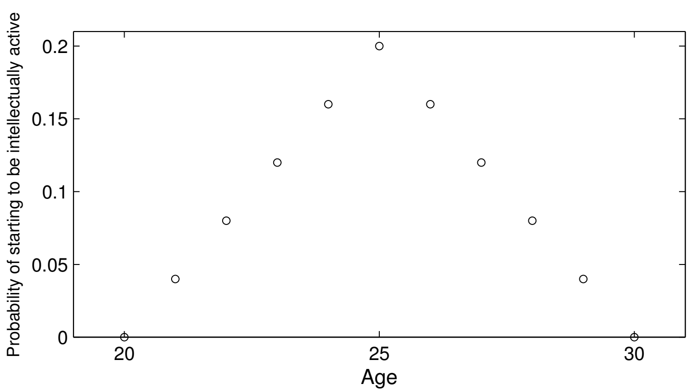

Figure 9. The probability distribution $B(x)$ for the starting point of one's intellectual career.

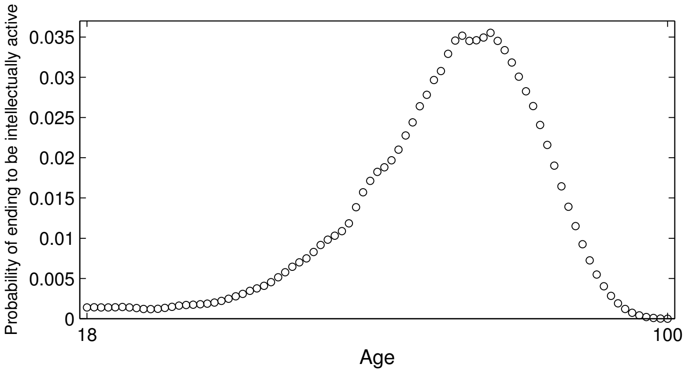

Figure 10. The probability distribution $S(x)$ for the endpoint of one's intellectual career.

and since the beginning of the active life of Hypatia does not depend on her father's activity, the following simplification can be made:

$$\sum_i \sum_k P(A^T_{i+k} \cap B^I_k \mid F_i) \cdot P(F_i) = \sum_i \sum_k P(B^I_k) \cdot P(A^T_{i+k} \mid F_i) \cdot P(F_i).$$

Without committing a large error, it is possible to confuse the probability of being active at age $i+k$ having had a daughter at age i, $P(A^T_{i+k} \mid F_i)$, with the one of being active at age $i+k$ having been alive at age i (V_i),[11] $P(A^T_{i+k} \mid V_i)$:

$$P(A^T_{i+k} \mid F_i) \approx P(A^T_{i+k} \mid V_i) = \frac{P(A^T_{i+k})}{P(V_i)}.$$

In the end, the following equation can be written:

$$P(C) = \sum_i \sum_k P(B^I_k) \cdot \frac{P(A^T_{i+k})}{P(V_i)} \cdot P(F_i).$$

Based on the idea previously introduced, the next step is to calculate $P(F_i \mid C)$ and $P(S^T_k \mid C)$ (and so $P(A^T_i \mid C) = 1 - \sum_k P(S^T_k \mid C)$):

$$P(F_i \mid C) = \frac{P(F_i \cap C)}{P(C)} = \frac{\sum_k P(B^I_k) \cdot (P(A^T_{i+k})/P(V_i)) \cdot P(F_i)}{\sum_i \sum_k P(B^I_k) \cdot (P(A^T_{i+k})/P(V_i)) \cdot P(F_i)},$$

$$P(S^T_k \mid C) = \frac{P(S_k \cap C)}{P(C)} = \frac{\sum_{i,j:i+j \le k} P(S^T_k) \cdot P(F_i) \cdot P(B^I_j)}{\sum_i \sum_j P(B^I_j) \cdot (P(A^T_{i+j})/P(V_i)) \cdot P(F_i)}.$$

$A^T_C(x)$ is the probability distribution of Theon being active at a given age, conditioned to the C event; see Figure 11.

[11] V_i is obtained from the above-mentioned 1974 Italian male mortality data set.

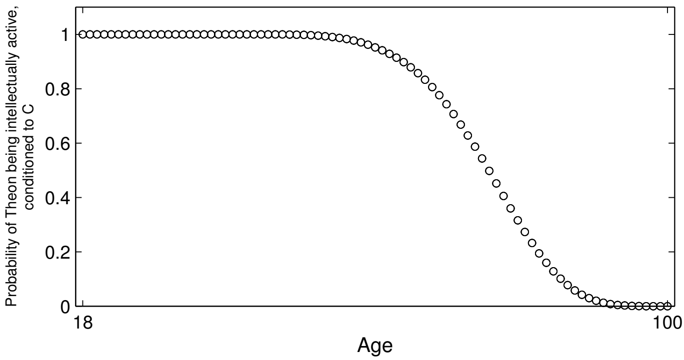

Figure 11. The probability distribution $A_C^T(x)$ for Theon being active at a given age, given that his and Hypatia's periods of activity overlap.

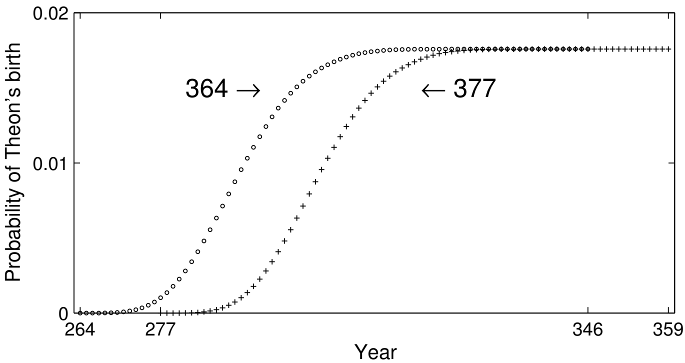

Figure 12. The probability distributions $364(\xi)$ and $377(\xi)$.

Keeping in mind the two years in which Theon was surely active (364 and 377), two distributions $364(\xi)$ and $377(\xi)$ for Theon's year of birth are deduced as previously shown in Section 3.2 (see Figure 12). Then, following the procedure introduced in Section 2, a single distribution $\Theta(\xi)$ is obtained (see Figure 13).

Finally, in order to calculate $\Upsilon_d(\xi)$, the probability distribution for the year of Hypatia's birth based on her being Theon's daughter, the probability of the various events "the age difference between father and daughter is i years" conditioned on event C must be known. This is indeed the above-calculated $P(F_i \mid C)$, now written as the function $F_C(x)$ (see Figure 14) so that $\Upsilon_d(\xi)$ is straightforward to calculate:[12]

$$\Upsilon_d(\xi) = \sum_x \Theta(\xi) \cdot F_C(\xi - \xi).$$

(See Figure 15.)

[12] The sum is taken over the whole domain of $\Theta(\xi)$.

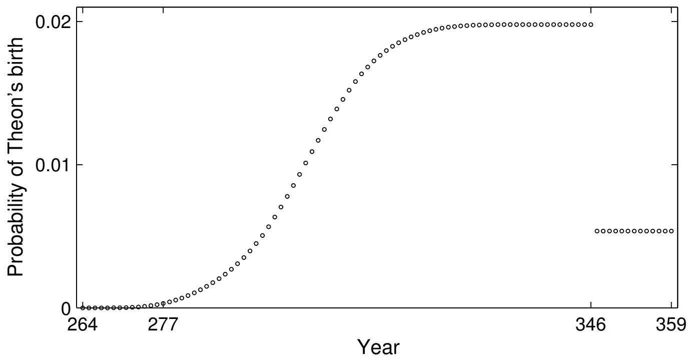

Figure 13. The probability distribution $\Theta(\xi)$ for Theon's birth.

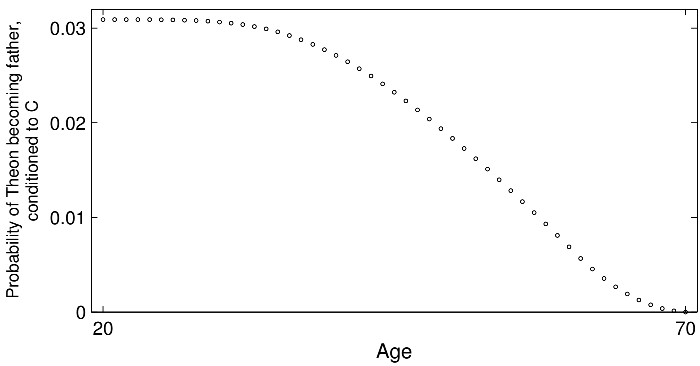

Figure 14. The probability distribution $F_C(x)$ for the difference in age between father and daughter, given that their periods of activity overlap.

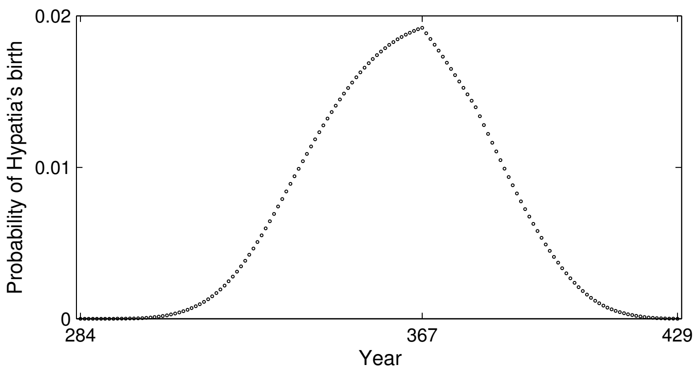

Figure 15. The probability distribution $\Upsilon_d(\xi)$ for Hypatia's birth based on her relationship to Theon.

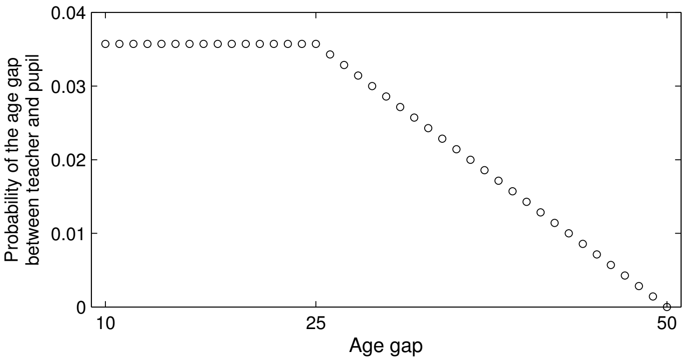

Figure 16. The probability distribution $T(x)$ for the age gap between teacher and disciple.

3.5. *Hypatia, teacher of Synesius.* Synesius of Cyrene, neo-Platonic philosopher and bishop of Ptolemais, was a disciple of Hypatia, as shown by a close correspondence between the two.

For instance, from his deathbed, Synesius wrote:

> Τῇ φιλοσόφῳ.
> Κλινοπετὴς ὑπηγόρευσα τὴν ἐπιστολήν, ἣν ὑγιαίνουσα κομίσαιο, μῆτερ καὶ ἀδελφὴ καὶ διδάσκαλε καὶ διὰ πάντων τούτων εὐεργετικὴ καὶ πᾶν ὅ τι τίμιον καὶ πρᾶγμα καὶ ὄνομα.[13]

The distribution $T(x)$ is introduced as a model to describe the probability of a difference of x years of age between teacher and pupil (see Figure 16).

$\Upsilon_t(\xi)$, the probability distribution for the year of Hypatia's birth deduced from this historical datum, is obtained in a straightforward manner by taking 370 as the year of birth of Synesius[14] (see Figure 17).

3.6. *Combined distribution.* Combining the five probability distributions deduced above for the year of Hypatia's birth, one final distribution, $\Upsilon(\xi)$, can be obtained following the rules introduced in Section 2. This final distribution $\Upsilon(\xi)$ can be compared to the distribution given by the simple arithmetic mean of the various distributions resulting from every possible combination of testimonies being considered true at the same time, $\Upsilon_A(\xi)$ (see Figure 18).

Therefore, the most probable year for the birth of Hypatia is 355 ($\sim$14.5%) with a total probability of the interval [350, 360] of about 90%.

[13]To the Philosopher. I am dictating this letter to you from my bed, but may you receive it in good health, mother, sister, teacher, and withal benefactress, and whatsoever is honored in name and deed [Synesius of Cyrene, Incipit of Letter 16].

[14]See, for example, [Hornblower et al. 2012].

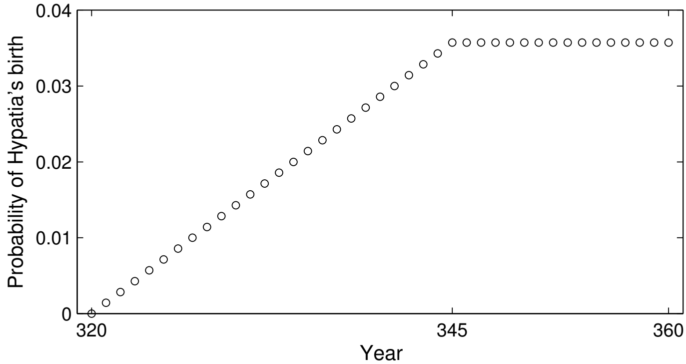

Figure 17. The probability distribution $\Upsilon_t(\xi)$ for Hypatia's age based on her having been a teacher of Synesius.

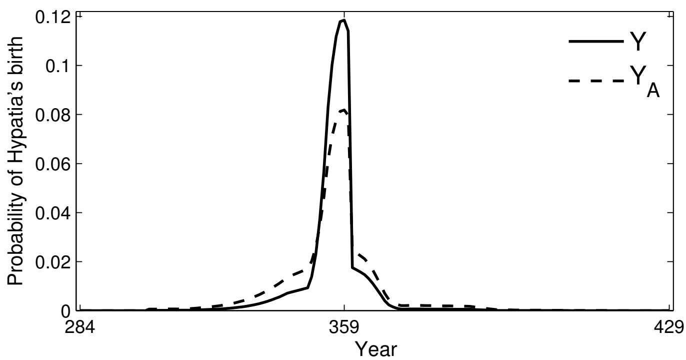

Figure 18. The final probability distribution $\Upsilon(\xi)$ calculated for the birth of Hypatia using the method in Section 2 and an average distribution $\Upsilon_A(\xi)$ based on the same historical data.

4. Conclusions

The probabilistic dating model proposed in this work, structured in three steps, could be summarized by making use of a culinary analogy. The first step is represented by the collection of enough raw ingredients (testimonies) to be refined or "cooked" in the second step (turned into probability distributions) and — finally, in the third step — put together following a recipe (provided in Section 2) so that they blend well (as a single probability distribution).

Its application to the case of Hypatia proved to be satisfactory in that the final probability distribution shows a marked peak, making it possible to give a date with good precision. The result so obtained contradicts the prevalent opinion (cf. page 25) but is in agreement with the minority view held by some highly-regarded scholars working on the issue. We have already mentioned the authoritative opinion of Maria Dzielska, who deems that Hypatia died at about age 60, having been, consequently, born around the year 355. A similar opinion is expressed in [Deakin 2007, p. 52].

Future applications appear to be far-reaching as the method could serve not only in cases strictly analogous to the one presented here but also in dating any event provided with a sufficient number of testimonies able to be turned into probability distributions.

References

[Abbas 2009] A. E. Abbas, "A Kullback–Leibler view of linear and log-linear pools", *Decision Anal.* **6**:1 (2009), 25–37.

[Deakin 2007] M. A. B. Deakin, *Hypatia of Alexandria: mathematician and martyr*, Prometheus Books, Amherst, NY, 2007.

[Dubucs 1993] J.-P. Dubucs (editor), *Philosophy of probability*, Philosophical Studies Series **56**, Springer, Dordrecht, 1993.

[Dzielska 1995] M. Dzielska, *Hypatia of Alexandria*, Revealing Antiquity **8**, Harvard University, Cambridge, MA, 1995.

[de Finetti 1931] B. de Finetti, "Sul significato soggettivo della probabilità", *Fund. Math.* **17** (1931), 298–329.

[Genest and McConway 1990] C. Genest and K. J. McConway, "Allocating the weights in the linear opinion pool", *J. Forecasting* **9**:1 (1990), 53–73.

[Genest and Wagner 1987] C. Genest and C. G. Wagner, "Further evidence against independence preservation in expert judgement synthesis", *Aequationes Math.* **32**:1 (1987), 74–86.

[Genest and Zidek 1986] C. Genest and J. V. Zidek, "Combining probability distributions: a critique and an annotated bibliography", *Statist. Sci.* **1**:1 (1986), 114–148.

[Hornblower et al. 2012] S. Hornblower, A. Spawforth, and E. Eidinow (editors), *The Oxford Classical Dictionary*, 4th ed., Oxford University, 2012.

[Kullback and Leibler 1951] S. Kullback and R. A. Leibler, "On information and sufficiency", *Ann. Math. Statistics* **22** (1951), 79–86.

[Lehrer and Wagner 1983] K. Lehrer and C. Wagner, "Probability amalgamation and the independence issue: a reply to Laddaga", *Synthese* **55**:3 (1983), 339–346.

[Malalas] E. Jeffreys, M. Jeffreys, and R. Scott (translators), *The Chronicle of John Malalas*, Byzantina Australiensia **41**, Melbourne Australian Association for Byzantine Studies, 1986.

[McConway 1981] K. J. McConway, "Marginalization and linear opinion pools", *J. Amer. Statist. Assoc.* **76**:374 (1981), 410–414.

[Peirce 1901] C. S. Peirce, "On the logic of drawing history from ancient documents especially from testimonies", manuscript published in his *Collected Papers*, vol. 7/8, paragraphs 164-231, edited by A. W. Burks, Belknap Press, Cambridge, MA, 1966. Reprinted in *The essential Peirce: selected philosophical writings*, vol. 2, Indiana University, Bloomington, 1998, pp. 75–114.

[Socrates Scholasticus] A. C. Zenos (translation editor), *Socrates Scholasticus: Ecclesiastical history*, Select Library of the Nicene and post-Nicene Fathers of the Christian Church (2) **2**, Christian Literature Company, New York, 1890.

[Synesius of Cyrene] A. Fitzgerald (translator), *The letters of Synesius of Cyrene*, Oxford University, London, 1926.

Received 25 Apr 2016. Revised 3 Jun 2016. Accepted 17 Jul 2016.

CANIO BENEDETTO: canio.benedetto@gmail.com
Via Leonardo da Vinci 20, I-85100 Potenza, Italy

STEFANO ISOLA: stefano.isola@gmail.com
Scuola di Scienze e Tecnologie, Università degli Studi di Camerino, I-62032 Camerino, Italy

LUCIO RUSSO: russo@axp.mat.uniroma2.it
Dipartimento di Matematica, Università degli Studi di Roma Tor Vergata, I-00173 Roma, Italy

MATHEMATICS AND MECHANICS OF COMPLEX SYSTEMS
Vol. 5, No. 1, 2017
dx.doi.org/10.2140/memocs.2017.5.41

ON THE POSSIBLE EFFECTIVE ELASTICITY TENSORS OF 2-DIMENSIONAL AND 3-DIMENSIONAL PRINTED MATERIALS

GRAEME W. MILTON, MARC BRIANE AND DAVIT HARUTYUNYAN

The set GU_f of possible effective elastic tensors of composites built from two materials with elasticity tensors $\boldsymbol{C}_1 > 0$ and $\boldsymbol{C}_2 = 0$ comprising the set $U = \{\boldsymbol{C}_1, \boldsymbol{C}_2\}$ and mixed in proportions f and $1-f$ is partly characterized. The material with tensor $\boldsymbol{C}_2 = 0$ corresponds to a material which is void. (For technical reasons $\boldsymbol{C}_2$ is actually taken to be nonzero and we take the limit $\boldsymbol{C}_2 \to 0$). Specifically, recalling that GU_f is completely characterized through minimums of sums of energies, involving a set of applied strains, and complementary energies, involving a set of applied stresses, we provide descriptions of microgeometries that in appropriate limits achieve the minimums in many cases. In these cases the calculation of the minimum is reduced to a finite-dimensional minimization problem that can be done numerically. Each microgeometry consists of a union of walls in appropriate directions, where the material in the wall is an appropriate p-mode material that is easily compliant to $p \leq 5$ independent applied strains, yet supports any stress in the orthogonal space. Thus the material can easily slip in certain directions along the walls. The region outside the walls contains "complementary Avellaneda material", which is a hierarchical laminate that minimizes the sum of complementary energies.

1. Introduction

Here we consider what effective elasticity tensors can be produced in the limit $\delta \to 0$ if we mix in prescribed proportions two materials with positive definite and bounded elasticity tensors $\boldsymbol{C}_1$ and $\boldsymbol{C}_2 = \delta \boldsymbol{C}_0$. In the limit $\delta \to 0$ this represents a mixture of an elastic phase and an extremely compliant phase. Thus we are given a set $U = \{\boldsymbol{C}_1, \delta \boldsymbol{C}_0\}$ and we are aiming to characterize as best we can the set GU_f of all possible effective tensors of composites having a volume fraction f of phase 1. The elasticity tensor $\boldsymbol{C}_1$ need not be isotropic but if it is anisotropic we require that it has a fixed orientation throughout the composite. Our results are summarized by the theorems in Section 10.

Communicated by Robert P. Lipton.
MSC2010: 74Q20, 35Q74.
Keywords: printed materials, elastic G-closures, metamaterials.

To get an idea of the enormity of the problem one has to recognize that in three dimensions elasticity tensors can be represented by 6×6 matrices and these have 21 independent elements. The set of possible elasticity tensors is thus represented as a set in a 21-dimensional space. Even a distorted multidimensional cube in a 21-dimensional space needs about 44 million real numbers to represent it (specifying the position in 21-dimensional space of each of the 2^{21} vertices). In the case where the two phases are isotropic, one is free to rotate the material to obtain an equivalent structure. Thus the set of possible elasticity tensors is invariant under rotations. As rotations involve three parameters (the Euler angles) this reduces the number of constants needed to describe the elasticity tensor from 21 to $21 - 3 = 18$, and thus the elasticity tensor can be represented in an 18-dimensional space of tensor invariants. For example, in the generic case, one can take these 18 invariants as follows: the six eigenvalues of the elasticity tensor; the two independent elements of the normalized eigenstrain associated with the lowest eigenvalue that can be assumed to be diagonal by an appropriate choice of the coordinate axes (which then fixes these axes); the four independent elements of the normalized eigenstrain associated with the second lowest eigenvalue that is orthogonal to the first eigenstrain; the three independent elements of the normalized eigenstrain associated with the third lowest eigenvalue that is orthogonal to the first two eigenstrains; the two independent elements of the normalized eigenstrain associated with the third lowest eigenvalue that is orthogonal to the first three eigenstrains; and the one independent element of the normalized eigenstrain associated with the third lowest eigenvalue that is orthogonal to the first four eigenstrains. This brings the total to $6+2+4+3+2+1 = 18$. In the same way that it takes two parameters (the bulk and shear moduli) to specify the elastic behavior of an isotropic material, it takes 18 parameters to specify the elastic behavior of a fully anisotropic material.

A distorted cube in this 18-dimensional space still requires about 4.7 million numbers to represent it. This makes exploring the range of possible elasticity tensors a daunting, if not impossible, numerical task. Some numerical exploration of this space has been done by Sigmund [1994; 1995], but we emphasize that this exploration covers only a tiny fraction of the number of possibilities.

Furthermore, the microstructures we found that lie near the boundary of GU_f have quite complicated multiscale architectures and thus would be difficult to find numerically. Also, it is not clear whether there are significantly simpler microstructures that can do the job. The numerical route of Sigmund should provide some simpler alternatives for the strut configurations in the multimode structures in the walls, although even then one needs to make subtle multiscale replacements (such as those appearing later in Figures 9 and 10) to achieve the desired performance. Numerical tests need to be made to see whether one can achieve the same performance with simpler structures. While strut configurations might be suitable at low volume fractions they are unlikely to be ideal at high volume fractions. Work by Allaire and Aubry [1999] shows that sometimes optimal microstructures necessarily have structure on multiple length scales. Even if one could numerically explore the question, it is not clear how one could summarize the results in a concise way.

From the applied side there is growing interest in trying to characterize the effective elasticity tensors of microstructures that can be produced by 3-dimensional or 2-dimensional printing. A dramatic example of such a microstructure is given in Figure 1. Our results have obvious relevance to this problem in the case where the 3-dimensional printed material uses only one isotropic material plus void. Although our microstructures are somewhat extreme, they provide benchmarks that show what is theoretically possible. What is possible in practice will be a subset of this.

The microstructures we consider involve taking three limits. First, as they have structure on multiple length scales, the homogenization limit where the ratio between length scales goes to infinity needs to be taken. Second, the limit $\delta \to 0$ needs to be taken. Third, as the structure involves thin walls of width ϵ, along which the material can "slip", the limit $\epsilon \to 0$ needs to be taken so the contribution to the complementary energy of these walls goes to zero, when the structure supports an applied stress. (Here ϵ should not be confused with the size of the unit cell, as is common in homogenization theory). The limits should be taken in this order, as, for example, standard homogenization theory is justified only if $\delta \neq 0$, so we need to take the homogenization limit before taking the limit $\delta \to 0$. In the walled structures the material may only occupy a small volume fraction, but this is ultimately irrelevant as the thin walled structures themselves occupy only a very small volume fraction in the final material (which goes to zero as $\epsilon \to 0$).

The case, applicable to printed materials, when phase 2 is actually void, rather than almost void, requires special care. To justify the homogenization steps taken here one has to first replace the void phase 2 with a composite foam having a small amount of phase 1 as the matrix phase, so that its effective elasticity tensor is nonzero, but approaches zero as the proportion of phase 1 in it tends to zero. The microgeometry in this composite needs to be much smaller than the scales in the geometries discussed here, which would involve mixtures of it and phase 1.

We emphasize, too, that our analysis is valid only for linear elasticity, and ignores nonlinear effects such as buckling. In reality the structures will easily buckle under compression. This buckling will occur, for example, in the square beam array structure of Figure 10. Additionally, some of the multimode materials are constructed via a superposition of appropriately shifted and deformed pentamode materials, and these substructures will interact under finite deformations. Also, in practice it would be difficult to realize the delicate multiscale materials that come close to attaining the bounds. Thus what is practically realizable will be just a subset, dependent on the current state of technology, of the set GU_f.

While the title refers only to printed materials, the results are also applicable to any periodic, or statistically homogeneous, material containing voids or pores in a homogeneous material. Printed materials are more interesting than typical porous materials as they allow one to explore a wider range of interesting structures.

In a companion paper [Milton et al. 2017] we consider the opposite limit $\delta \to \infty$, corresponding to a mixture of an elastic material and an almost rigid material.

2. Review of some bounds on the elastic moduli of two-phase composites and geometries that attain them

Here we review a selection of results on sharp bounds on the elastic response of two-phase composites and the associated problem of identifying optimal geometries that attain them. The interested reader is encouraged to look at the books of Nemat-Nasser and Hori [1998], Cherkaev [2000], Milton [2002], Allaire [2002], Torquato [2002] and Tartar [2009], which provide a much more comprehensive survey.

The most elementary bounds on the elastic properties of composites are the classical bounds of Hill [1952], who implicitly showed that

$$\langle [\boldsymbol{C}(\boldsymbol{x})]^{-1} \rangle^{-1} \leq \boldsymbol{C}_* \leq \langle \boldsymbol{C}(\boldsymbol{x}) \rangle. \tag{2-1}$$

Here the angular brackets $\langle \cdot \rangle$ denote a volume average, and the inequality holds in the sense of quadratic forms, i.e., for fourth-order tensors $\boldsymbol{A}$ and $\boldsymbol{B}$ satisfying the symmetries of elasticity tensors we say that $\boldsymbol{A} \geq \boldsymbol{B}$ if $\boldsymbol{\epsilon} : \boldsymbol{A}\boldsymbol{\epsilon} \geq \boldsymbol{\epsilon} : \boldsymbol{B}\boldsymbol{\epsilon}$ for all matrices $\boldsymbol{\epsilon}$. While these bounds were not explicitly stated by Hill in his 1952 paper they are an immediate and obvious consequence of his equation (2). If the two phases are isotropic the spectral decomposition of the elasticity tensors $\boldsymbol{C}_1$ and $\boldsymbol{C}_2$ of the two phases is

$$\boldsymbol{C}_1 = 3\kappa_1 \boldsymbol{\Lambda}_h + 2\mu_1 \boldsymbol{\Lambda}_s \quad \text{and} \quad \boldsymbol{C}_2 = 3\kappa_2 \boldsymbol{\Lambda}_h + 2\mu_2 \boldsymbol{\Lambda}_s, \tag{2-2}$$

where κ_1 and κ_2 are the bulk moduli of the two phases, μ_1 and μ_2 are the shear moduli, and

$$\{\boldsymbol{\Lambda}_h\}_{ijk\ell} = \tfrac{1}{3}\delta_{ij}\delta_{k\ell}, \quad \{\boldsymbol{\Lambda}_s\}_{ijk\ell} = \tfrac{1}{2}[\delta_{ik}\delta_{j\ell} + \delta_{i\ell}\delta_{kj}] - \tfrac{1}{3}\delta_{ij}\delta_{k\ell} \tag{2-3}$$

act as projections. The tensor $\boldsymbol{\Lambda}_h$ projects onto the 1-dimensional space of matrices proportional to the second-order identity matrix, while $\boldsymbol{\Lambda}_s$ projects onto the 5-dimensional space of trace-free matrices. Similarly if the effective elasticity tensor $\boldsymbol{C}_*$ is isotropic we have that $\boldsymbol{C}_* = 3\kappa_1 \boldsymbol{\Lambda}_h + 2\mu_1 \boldsymbol{\Lambda}_s$, where κ_* and μ_* are the effective bulk and shear moduli of the composite. In this paper we are interested in the case where the two phases are well-ordered in the sense that

$$\boldsymbol{C}_1 \geq \boldsymbol{C}_2, \tag{2-4}$$

and we will take the limit as $\boldsymbol{C}_2 \to 0$, meaning that all the eigenvalues of $\boldsymbol{C}_2$ approach zero. In the case of isotropic components this well-ordering assumption is satisfied if $\kappa_1 \geq \kappa_2$ and $\mu_1 \geq \mu_2$, and we will take the limit as $\kappa_2, \mu_2 \to 0$.

For isotropic composites of two well-ordered materials Hashin and Shtrikman [1963] and Hill [1963] obtained the celebrated bounds

$$\begin{aligned}
\kappa_* &\geq f\kappa_1+(1-f)\kappa_2-\frac{f(1-f)(\kappa_1-\kappa_2)^2}{(1-f)\kappa_1+f\kappa_2+4\mu_2/3},\\
\kappa_* &\leq f\kappa_1+(1-f)\kappa_2-\frac{f(1-f)(\kappa_1-\kappa_2)^2}{(1-f)\kappa_1+f\kappa_2+4\mu_1/3},\\
\mu_* &\geq f\mu_1+(1-f)\mu_2-\frac{f(1-f)(\mu_1-\mu_2)^2}{(1-f)\mu_1+f\mu_2+\mu_2(9\kappa_2+8\mu_2)/[6(\kappa_2+2\mu_2)]},\\
\mu_* &\leq f\mu_1+(1-f)\mu_2-\frac{f(1-f)(\mu_1-\mu_2)^2}{(1-f)\mu_1+f\mu_2+\mu_1(9\kappa_1+8\mu_1)/[6(\kappa_1+2\mu_1)]}.
\end{aligned} \tag{2-5}$$

In fact these bounds (and the variational principles they derive from) hold even if one component has a negative bulk modulus, so long as the composite is stable [Kochmann and Milton 2014]. For 2-dimensional composites (fiber reinforced materials) analogous bounds on the effective elastic moduli were found by Hill [1964] and Hashin [1965]. Bounds on the complex effective bulk and shear moduli of isotropic two-phase 2-dimensional or 3-dimensional composites were also obtained [Gibiansky and Milton 1993; Milton and Berryman 1997; Gibiansky et al. 1993; 1999; Gibiansky and Lakes 1993; 1997]: these are appropriate to the propagation of fixed frequency elastic waves in composites when one or both of the phases is viscoelastic, and when the wavelength is much larger than the microstructure.

An important "attainability principle" is that bounds obtained by substituting a trial field in a variational principle will be attained when the geometry is such that the actual field matches this trial field. This principle was used, for example, in [Milton 1981c] to find geometries that attain the Hashin–Shtrikman bounds on the effective bulk modulus of composites with three or more phases (see also [Gibiansky and Sigmund 2000]). The Hashin–Shtrikman variational principles involve a minimization over trial polarization fields, and the actual polarization field depends on the choice of the elasticity tensor $\boldsymbol{C}_0$ of a "reference medium" (typically chosen to be positive definite) and is defined by

$$\boldsymbol{P}(\boldsymbol{x}) = (\boldsymbol{C}(\boldsymbol{x}) - \boldsymbol{C}_0)\boldsymbol{\epsilon}(\boldsymbol{x}) = \boldsymbol{\sigma}(\boldsymbol{x}) - \boldsymbol{C}_0\boldsymbol{\epsilon}(\boldsymbol{x}). \tag{2-6}$$

The variational principles require that $\boldsymbol{C}(\boldsymbol{x}) - \boldsymbol{C}_0$ be either positive semidefinite or negative semidefinite, so in the case of a well-ordered material natural choices of $\boldsymbol{C}_0$ are $\boldsymbol{C}_1$ or $\boldsymbol{C}_2$ and correspondingly the field will be zero in phase 1 or phase 2, respectively. The bounds are obtained by assuming it is constant in the other phase (proportional to the identity in case of the bulk modulus bounds, and trace-free for the shear modulus bounds). Hashin and Shtrikman [1963] recognized that the effective bulk modulus would be attained by the Hashin assemblage of coated spheres [Hashin 1962]. A single coated sphere can be a neutral inclusion: if the surrounding "matrix" material has an appropriate bulk modulus (with a specific value between κ_1 and κ_2) one can insert it in the matrix material without disturbing a surrounding hydrostatic field (this is the principle behind the unfeelability cloak of Bückmann, Thiel, Kadic, Schittny and Wegener [Bückmann et al. 2014]). The inclusion is invisible to the surrounding field and one can continue to insert similar inclusions, scaled to sizes ranging to the very small, until one essentially obtains a two-phase composite with effective bulk modulus the same as the original matrix material. Due to radial symmetry the forces acting on the spherical inner core will be equally distributed around the boundary and directed radially: thus the field inside the core material is hydrostatic and constant, and hence by the attainability principle, and due to their neutrality, sphere assemblages must attain the effective bulk modulus bounds in (2-5).

One very important class of microgeometries for which the field is constant in one phase are the sequentially layered laminates (first introduced by Maxwell [1873]) built by layering phase 2 with phase 1 in a direction $\boldsymbol{n}_1$ (by which we mean $\boldsymbol{n}_1$ is perpendicular to the layers), then taking this laminate and layering it again on a much larger length scale with phase 1 in a direction $\boldsymbol{n}_2$ to obtain a "rank 2" laminate, and continuing this process until one obtains a rank m laminate, containing in a sense a "core" of phase 2 surrounded by layers of phase 1. The field is then constant in the core material of phase 2. An explicit formula for the effective elasticity tensor of such sequentially layered laminates was obtained by Francfort and Murat [1986], generalizing the analogous formulas obtained by Tartar [1985] for conductivity. Of course one can switch the roles of the phases in this construction and thus obtain a material where the field is constant in phase 1. It then immediately follows from the attainability principle (without requiring any calculation!) that one can attain the Hashin–Shtrikman shear modulus bounds (2-5) (and simultaneously the bulk modulus bounds) if one can find a sequentially layered laminate that has an isotropic elasticity tensor, and the easiest way to do this is to do the lamination sequentially by adding infinitesimal layers in random directions. This established the attainability of the Hashin–Shtrikman shear modulus bound [Milton 1986], also established independently and at the same time by Norris [1985], using the differential scheme that was known to be realizable [Milton 1985; Avellaneda 1987a] — in fact Roscoe [1973] had earlier realized the differential approximation scheme could produce the desired shear modulus — and at the same time elegantly by Francfort and Murat [1986], using sequentially layered laminates with just five directions of lamination (in the case of 3-dimensional composites).

Hill [1963] proved that the bulk modulus bounds are valid also in the non-well-ordered case where $\mu_1 \geq \mu_2$ but $\kappa_1 \leq \kappa_2$. As far as we know, the tightest bounds on the effective shear modulus of 3-dimensional composites in the non-well-ordered case where $\mu_1 \geq \mu_2$ but $\kappa_1 \leq \kappa_2$ are those of Milton and Phan-Thien [1982]:

$$\begin{aligned}\min_{\substack{\zeta\\0\leq\zeta\leq1}} \frac{8\langle 6/\mu+7/\kappa\rangle_\zeta+15/\mu_2}{2\big(\langle 21/\mu+2/\kappa\rangle_\zeta/\mu_2+40\langle 1/\mu\rangle_\zeta\langle 1/\kappa\rangle_\zeta\big)}&\\ \leq \frac{f(1-f)(\mu_1-\mu_2)^2}{f\mu_1+(1-f)\mu_2-\mu_*}&-(1-f)\mu_1-f\mu_2\\ \leq \max_{\substack{\zeta\\0\leq\zeta\leq1}} \frac{8\mu_1\langle 6\kappa+7\mu\rangle_\zeta+15\langle\mu\rangle_\zeta\langle\kappa\rangle_\zeta}{2\big(\langle 21\kappa+2\mu\rangle_\zeta+40\mu_1\big)}&,\end{aligned}\tag{2-7}$$

where for any quantity a taking values a_1 and a_2 in phase 1 and phase 2, respectively, we define $\langle a\rangle_\zeta \equiv \zeta a_1+(1-\zeta)a_2$. These bounds are obtained by eliminating the geometric parameters from the bounds of Milton and Phan-Thien [1982] and are tighter than the better-known Walpole bounds [1966], and are in fact sharp (as they coincide with the Hashin–Shtrikman formula, which corresponds to particular geometries as we have discussed) when the moduli are slightly non-well-ordered. Specifically, the first bound in (2-7) is sharp when the minimum over ζ is attained at $\zeta=0$, which occurs when

$$\kappa_1-\kappa_2 \geq -\frac{(3\kappa_2+8\mu_2)^2}{42\kappa_2^2}\frac{\kappa_1\kappa_2}{\mu_1\mu_2}(\mu_1-\mu_2),\tag{2-8}$$

and the second bound in (2-7) is sharp when the maximum over ζ is attained at $\zeta=1$, which occurs when

$$\kappa_1-\kappa_2 \geq -\frac{(3\kappa_1+8\mu_1)^2}{42\mu_1^2}(\mu_1-\mu_2).\tag{2-9}$$

The bounds (2-5) and (2-7) constrain the pair (κ_*, μ_*) to lie in a rectangular box. Berryman and Milton [1988] obtained tighter coupled bounds which slice off two opposing corner regions of the box by eliminating the geometric parameters from the bulk modulus bounds of Beran and Molyneux [1966] (as simplified by Milton [1981b]) and from the shear modulus bounds of Milton and Phan-Thien [1982]. There is good reason to believe these bounds can be improved as the analogous 2-dimensional bounds are not as tight as the bounds of Cherkaev and Gibiansky [1993] coupling κ_* and μ_*, which were derived using the translation method.

For anisotropic composites with an effective tensor $\boldsymbol{C}_*$, the microstructure independent bounds that are directly analogous to the Hashin–Shtrikman–Hill bounds, given by (2-5), are the "trace bounds"

$$\begin{aligned} f\,\mathrm{Tr}[\boldsymbol{\Lambda}_h(\boldsymbol{C}_*-\boldsymbol{C}_2)^{-1}] &\le \frac{1}{3(\kappa_1-\kappa_2)}+\frac{1-f}{3\kappa_2+4\mu_2},\\ (1-f)\,\mathrm{Tr}[\boldsymbol{\Lambda}_h(\boldsymbol{C}_1-\boldsymbol{C}_*)^{-1}] &\le \frac{1}{3(\kappa_1-\kappa_2)}-\frac{f}{3\kappa_1+4\mu_1},\\ f\,\mathrm{Tr}[\boldsymbol{\Lambda}_s(\boldsymbol{C}_*-\boldsymbol{C}_2)^{-1}] &\le \frac{5}{2(\mu_1-\mu_2)}+\frac{3(\kappa_2+2\mu_2)(1-f)}{\mu_2(3\kappa_2+4\mu_2)},\\ (1-f)\,\mathrm{Tr}[\boldsymbol{\Lambda}_s(\boldsymbol{C}_1-\boldsymbol{C}_*)^{-1}] &\le \frac{5}{2(\mu_1-\mu_2)}-\frac{3(\kappa_1+2\mu_1)f}{\mu_1(3\kappa_1+4\mu_1)}, \end{aligned} \tag{2-10}$$

obtained independently by Milton and Kohn [1988] and Zhikov [1988; 1991a; 1991b]. In these expressions the fourth-order tensors $\boldsymbol{\Lambda}_h$ multiply the fourth-order tensors on their right, and

$$\mathrm{Tr}[\boldsymbol{A}] = A_{ijij} \tag{2-11}$$

defines the "trace" of a fourth-order tensor (see also [Francfort and Murat 1986] and [Nemat-Nasser and Hori 1993] for related bounds). From the attainability principle it follows that these bounds will be achieved whenever the composite is a sequentially layered laminate, with a core of one phase, surrounded by layers (on widely separated length scales) of the other phase. When $\boldsymbol{C}_*$ is isotropic these bounds (2-10) reduce to the Hashin–Shtrikman–Hill bounds (2-5). In the case where the two phases, and hence the composite, are incompressible we can define the five effective shear moduli μ_1^*, μ_2^*, μ_3^*, μ_4^*, μ_5^* to be the five finite eigenvalues of $\frac{1}{2}\boldsymbol{C}_*$, and the second pair of bounds in (2-10) reduce to

$$\begin{aligned} \sum_{i=1}^{5}\frac{f}{2(\mu_{*i}-\mu_2)} &\le \frac{5}{2(\mu_1-\mu_2)}+\frac{3(\kappa_2+2\mu_2)(1-f)}{\mu_2(3\kappa_2+4\mu_2)},\\ \sum_{i=1}^{5}\frac{1-f}{2(\mu_1-\mu_{*i})} &\le \frac{5}{2(\mu_1-\mu_2)}-\frac{3(\kappa_1+2\mu_1)f}{\mu_1(3\kappa_1+4\mu_1)}. \end{aligned} \tag{2-12}$$

Lipton [1988] established that the analogous bounds for the two effective shear moduli μ_1^* and μ_2^* of 2-dimensional composites of two incompressible isotropic phases completely characterize GU_f.

Earlier, Willis [1977] considered anisotropic composites and used the Hashin–Shtrikman variational principle with a trial polarization that was zero in one phase and constant in the other to obtain bounds on the elastic energy of a two-phase composite. He found that these bounds are not microgeometry independent, but rather involve the two-point correlation function, i.e., the probability that a rod with fixed orientation lands with both ends in phase 1 when thrown randomly in a composite. It follows from the attainability principle that the Willis bounds will be achieved when the composite is a sequentially layered laminate, with a core of one phase, surrounded by layers (on widely separated length scales) of the other phase.

In a major advance, Avellaneda [1987b] recognized that for any composite of two phases with well-ordered tensors not all the information contained in the two-point correlation function was relevant to determining the bounds: what was relevant was a “reduced two-point correlation function” that could be represented as a positive measure $\mu(\boldsymbol{\xi})$ (with unit integral) on the sphere $|\boldsymbol{\xi}|=1$. Roughly speaking one takes the Fourier transform of the two-point correlation function and integrates it over rays $\boldsymbol{k}=k\boldsymbol{\xi}$ in “Fourier space” keeping $\boldsymbol{\xi}$ fixed and integrating over k from 0 to infinity. Most importantly, every such measure could be realized to an arbitrarily high degree of approximation by the measure of a suitable sequentially layered laminate. For example, a measure with weighted delta functions in directions $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ would be realized by a second-rank sequentially layered laminate with layers normal to $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$. (We note in passing that these reduced two-point correlation functions of Avellaneda are a special case of the H-measures introduced at the same time by Tartar [1989; 1990], in terms of which he could calculate second-order corrections to the effective tensor of a nearly homogeneous composite. H-measures were also introduced independently by Gérard [1989; 1994] under the name of microlocal defect measures. For composites of two isotropic phases the Hashin–Shtrikman conductivity bounds, and indeed variational conductivity bounds at any order, can be naturally expressed in terms of the series expansion coefficients of the effective tensor up to a corresponding order for a nearly homogeneous composite, as shown by Milton and McPhedran [1982].)

The fantastic implication was that by summing the Willis bounds [1977], and then minimizing over all positive measures on the sphere, one would get sharp bounds on the sum of elastic complementary energies

$$W_f^0(\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0,\boldsymbol{\sigma}_3^0,\boldsymbol{\sigma}_4^0,\boldsymbol{\sigma}_5^0,\boldsymbol{\sigma}_6^0)=\min_{\boldsymbol{C}_*\in GU_f}\sum_{j=1}^{6}\boldsymbol{\sigma}_j^0:\boldsymbol{C}_*^{-1}\boldsymbol{\sigma}_j^0, \tag{2-13}$$

and similarly one could get sharp bounds on the sum of elastic energies

$$W_f^6(\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0,\boldsymbol{\epsilon}_3^0,\boldsymbol{\epsilon}_4^0,\boldsymbol{\epsilon}_5^0,\boldsymbol{\epsilon}_6^0)=\min_{\boldsymbol{C}_*\in GU_f}\sum_{i=1}^{6}\boldsymbol{\epsilon}_i^0:\boldsymbol{C}_*\boldsymbol{\epsilon}_i^0. \tag{2-14}$$

Here some of the applied stresses $\boldsymbol{\sigma}_j^0$ or the applied strains $\boldsymbol{\epsilon}_i^0$ could be zero. So the evaluation of the functions $W_f^0(\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0,\boldsymbol{\sigma}_3^0,\boldsymbol{\sigma}_4^0,\boldsymbol{\sigma}_5^0,\boldsymbol{\sigma}_6^0)$ and $W_f^6(\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0,\boldsymbol{\epsilon}_3^0,\boldsymbol{\epsilon}_4^0,\boldsymbol{\epsilon}_5^0,\boldsymbol{\epsilon}_6^0)$ reduces to a finite-dimensional minimization problem which can be done numerically. Hence we will treat the functions $W_f^0(\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0,\boldsymbol{\sigma}_3^0,\boldsymbol{\sigma}_4^0,\boldsymbol{\sigma}_5^0,\boldsymbol{\sigma}_6^0)$ as being known, and we will call an “Avellaneda material” an associated sequentially layered laminate material with effective tensor $\boldsymbol{C}_*=\boldsymbol{C}_f^A(\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0,\boldsymbol{\epsilon}_3^0,\boldsymbol{\epsilon}_4^0,\boldsymbol{\epsilon}_5^0,\boldsymbol{\epsilon}_6^0)$ that attains the minimum in (2-14), and similarly we call a "complementary Avellaneda material" an associated sequentially layered laminate material with effective tensor $\boldsymbol{C}_* = \widetilde{\boldsymbol{C}}_f^A(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0, \boldsymbol{\sigma}_5^0, \boldsymbol{\sigma}_6^0) \in GU_f$ that attains the minimum in (2-13). Explicit analytical formulas for the tensors $\boldsymbol{C}_f^A(\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0, \boldsymbol{\epsilon}_5^0, \boldsymbol{\epsilon}_6^0)$ and $\widetilde{\boldsymbol{C}}_f^A(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0, \boldsymbol{\sigma}_5^0, \boldsymbol{\sigma}_6^0)$ are not generally available, but rather have to be found by numerical computation. When $\boldsymbol{C}_1 \geq \boldsymbol{C}_2$ one needs to take the minimum in (2-13) over the $\boldsymbol{C}_*$ of sequentially layered laminates with a core material of phase 2. Similarly, when $\boldsymbol{C}_1^{-1} \geq \boldsymbol{C}_2^{-1}$ the minimum in (2-14) also can be taken over the $\boldsymbol{C}_*$ of sequentially layered laminates with a core material of phase 2. We remark that although Avellaneda assumed the tensors $\boldsymbol{C}_1$ and $\boldsymbol{C}_2$ were isotropic, his analysis easily extends to the case where the tensors are anisotropic but well-ordered (either with $\boldsymbol{C}_1 \geq \boldsymbol{C}_2$ or $\boldsymbol{C}_2 \geq \boldsymbol{C}_1$) and with constant orientation throughout the composite: see, for example, Section 23.3 in [Milton 2002].

These $\boldsymbol{C}_*$ of sequentially layered laminates are given by the formula of Francfort and Murat [1986] and Gibiansky and Cherkaev [1997b]:

$$(1-f)(\boldsymbol{C}_1 - \boldsymbol{C}_*)^{-1} = (\boldsymbol{C}_1 - \boldsymbol{C}_2)^{-1} - f \sum_{j=1}^{r} c_j \boldsymbol{\Gamma}(\boldsymbol{n}_j), \tag{2-15}$$

where r is the rank of the sequential laminate, the positive weights c_j sum to 1, the $\boldsymbol{n}_i$ are the lamination directions, and $\boldsymbol{\Gamma}(\boldsymbol{n})$ is the fourth-order tensor with elements given by

$$\{\boldsymbol{\Gamma}(\boldsymbol{n})\}_{hik\ell} = \tfrac{1}{4}\big(n_h\{\boldsymbol{C}(\boldsymbol{n})^{-1}\}_{ik}n_\ell + n_h\{\boldsymbol{C}(\boldsymbol{n})^{-1}\}_{i\ell}n_k + n_i\{\boldsymbol{C}(\boldsymbol{n})^{-1}\}_{hk}n_\ell + n_i\{\boldsymbol{C}(\boldsymbol{n})^{-1}\}_{h\ell}n_k\big), \tag{2-16}$$

in which $\boldsymbol{C}(\boldsymbol{n}) = \boldsymbol{n} \cdot \boldsymbol{C}_1 \boldsymbol{n}$ is the 3×3 matrix known as the acoustic tensor, with elements

$$\{\boldsymbol{C}(\boldsymbol{n})\}_{ik} = \{\boldsymbol{n} \cdot \boldsymbol{C}_1 \boldsymbol{n}\}_{ik} = n_h \{\boldsymbol{C}_1\}_{hik\ell} n_\ell. \tag{2-17}$$

Thus the minimum needs to be taken over the rank r of the sequential laminate, over the positive weights c_j, which sum to 1, and over the lamination directions $\boldsymbol{n}_j$. In the case where phase 1 is isotropic, with bulk modulus κ_1 and shear modulus μ_1, $\boldsymbol{C}(\boldsymbol{n})$ can be easily calculated and one obtains

$$\{\boldsymbol{\Gamma}(\boldsymbol{n}_j)\}_{hik\ell} = \frac{3n_h n_i n_k n_\ell}{3\kappa_1 + 4\mu_1} + \frac{1}{4\mu_1}(n_h\delta_{ik}n_\ell + n_h\delta_{i\ell}n_k + n_i\delta_{hk}n_\ell + n_i\delta_{h\ell}n_k - 4n_h n_i n_k n_\ell). \tag{2-18}$$

Francfort, Murat, and Tartar [Francfort et al. 1995] proved that when $\boldsymbol{C}_1$ is isotropic it suffices to limit attention to laminates of rank $r \leq 6$. When $\boldsymbol{C}_1$ is anisotropic we extend an argument due to Avellaneda [1987b]. Consider the set $\mathcal{A}$ consisting of all fourth-order tensors $\boldsymbol{A}$ of the form

$$\boldsymbol{A} = \int_{|\boldsymbol{n}|=1} \boldsymbol{\Gamma}(\boldsymbol{n})\, m(d\boldsymbol{n}), \tag{2-19}$$

where $m(d\boldsymbol{n})$ is a nonnegative measure on the unit sphere having an integral of 1 over the sphere. Since $\boldsymbol{A}$ satisfies

$$\{\boldsymbol{A}\}_{hik\ell}\{\boldsymbol{C}_1\}_{hik\ell} = \int_{|\boldsymbol{n}|=1} \{\boldsymbol{C}(\boldsymbol{n})^{-1}\}_{ik}\{\boldsymbol{C}(\boldsymbol{n})\}_{ik}\, m(d\boldsymbol{n}) = 3, \tag{2-20}$$

it follows that $\mathcal{A}$ is a convex set in a space of dimension $\nu = 20$ (with 20 of the 21 independent matrix elements of $\boldsymbol{A}$ as coordinates, and the remaining element being determined by (2-20)). The extreme points correspond to point masses on the unit sphere. Hence any tensor of the form (2-19) is a convex combination of at most $\nu + 1$ extreme points. Thus the sum (2-15) can be limited to $r \leq 21$; i.e., it suffices to consider laminates up to rank 21. Lipton [1991; 1992; 1994] obtained a complete algebraic characterization of the possible sequentially layered laminates having transverse or orthotropic symmetry and derived explicit expressions for many of the associated bounds. The Avellaneda materials are of course difficult to build in practice since they have structure on multiple length scales. However, if f is small and one phase is void, Bourdin and Kohn [2008] showed that it suffices to use a walled structure (similar to the structure on the right in Figure 4, but with walls in many directions, not just two, and with the wall thickness depending on orientation).

As observed by Avellaneda [1987b], the implications of course also apply to 2-dimensional elasticity. Define

$$W_f^0(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0) = \min_{\boldsymbol{C}_* \in GU_f} \sum_{j=1}^{3} \boldsymbol{\sigma}_j^0 : \boldsymbol{C}_*^{-1}\boldsymbol{\sigma}_j^0 \tag{2-21}$$

and

$$W_f^3(\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0) = \min_{\boldsymbol{C}_* \in GU_f} \sum_{i=1}^{3} \boldsymbol{\epsilon}_i^0 : \boldsymbol{C}_*\boldsymbol{\epsilon}_i^0. \tag{2-22}$$

Then there is an Avellaneda material with effective tensor $\boldsymbol{C}_* = \boldsymbol{C}_f^A(\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0)$ that attains the minimum in (2-22), and a complementary Avellaneda material with effective tensor $\boldsymbol{C}_* = \widetilde{\boldsymbol{C}}_f^A(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0) \in GU_f$ that attains the minimum in (2-21). In 2-dimensional elasticity, sequentially layered laminates have elasticity tensors given by (2-15)–(2-17) when the tensor $\boldsymbol{C}_1$ is anisotropic. When the elasticity tensor $\boldsymbol{C}_1$ of phase 1 is isotropic, the sequentially layered laminates of rank r have effective compliance tensors $\boldsymbol{S}_* = (\boldsymbol{C}_*)^{-1}$ given by the Gibiansky–Cherkaev formula

$$(1-f)(\boldsymbol{S}_1 - \boldsymbol{S}_*)^{-1} = (\boldsymbol{S}_1 - \boldsymbol{S}_2)^{-1} - f[(4\kappa_2)^{-1} + (4\mu_2)^{-1}]\boldsymbol{M} \tag{2-23}$$

(see [Gibiansky and Cherkaev 1997b, equations (2.37) and (2.38)] and see also [Lurie et al. 1982], in which Lurie, Cherkaev, and Fedorov derived an equivalent, but less simple, formula), where $\boldsymbol{S}_1 = (\boldsymbol{C}_1)^{-1}$ and $\boldsymbol{S}_2 = (\boldsymbol{C}_2)^{-1}$ are the compliance tensors of the two phases, occupying volume fractions f and $1 - f$, respectively, and $\boldsymbol{M}$ has elements

$$\{\boldsymbol{M}\}_{hik\ell} = \sum_{j=1}^{r} c_j \boldsymbol{t}_{j_h} \boldsymbol{t}_{j_i} \boldsymbol{t}_{j_k} \boldsymbol{t}_{j_\ell}, \tag{2-24}$$

in which the $\boldsymbol{t}_j$ are unit vectors perpendicular to the directions of lamination (i.e., parallel to the layer boundaries), and the c_j are any set of positive weights, summing to 1, giving the proportions of phase 1 laminated in the various directions. The tensor $\boldsymbol{M}$ is clearly positive semidefinite and has the property that

$$\{\boldsymbol{M}\}_{hkhk} = \{\boldsymbol{M}\}_{hhkk} = 1. \tag{2-25}$$

Conversely, Avellaneda and Milton [1989] have shown that given a positive semidefinite fourth-order tensor $\boldsymbol{M}$ satisfying (2-25) there is a sequential layered laminate of rank $r \le 3$ that corresponds to it, i.e., such that (2-23) holds for some choice of unit vectors $\boldsymbol{t}_j$ and weights c_j (see also Theorem 2.2 of [Francfort et al. 1995]). Thus when $\boldsymbol{C}_1$ is isotropic, the computation of the complementary Avellaneda tensor $\widetilde{\boldsymbol{C}}_f^A(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0)$ reduces to a minimization over positive semidefinite fourth-order tensors $\boldsymbol{M}$ satisfying (2-25). When $\boldsymbol{C}_1$ is anisotropic, by the same argument as in the 3-dimensional case, it suffices to consider sequential layered laminates of rank at most 6.

We also remark that aside from hierarchical laminates there are many other structures that have a uniform field in one phase, sometimes only for certain applied fields. These include assemblages of confocal ellipses and ellipsoids [Milton 1980; 1981a; Grabovsky and Kohn 1995a], the periodic Vigdergauz geometries [Vigdergauz 1986; 1994; 1996; 1999; Grabovsky and Kohn 1995b], the Sigmund structures [2000], and the periodic E-inclusions of Liu, James, and Leo [Liu et al. 2007] (see also Section 23.9 of [Milton 2002]). Usually these attain the bounds when the measure $\mu(\boldsymbol{\xi})$ minimizing the sum of Willis bounds is not required to be a discrete measure. Allaire and Aubry [1999] have shown that sometimes the best microstructure necessarily has structure on multiple length scales (like sequentially layered laminates).

For single energies for anisotropic two-phase composites, the Hill bounds (2-1) imply

$$\begin{aligned}\boldsymbol{\epsilon}_0 : [f\boldsymbol{C}_1^{-1} + (1-f)\boldsymbol{C}_2^{-1}]^{-1}\boldsymbol{\epsilon}_0 \le \boldsymbol{\epsilon}_0 : \boldsymbol{C}_*\boldsymbol{\epsilon}_0 \le \boldsymbol{\epsilon}_0 : [f\boldsymbol{C}_1 + (1-f)\boldsymbol{C}_2]\boldsymbol{\epsilon}_0,\\ \boldsymbol{\sigma}_0 : [f\boldsymbol{C}_1 + (1-f)\boldsymbol{C}_2]^{-1}\boldsymbol{\sigma}_0 \le \boldsymbol{\sigma}_0 : \boldsymbol{C}_*^{-1}\boldsymbol{\sigma}_0 \le \boldsymbol{\sigma}_0 : [f\boldsymbol{C}_1^{-1} + (1-f)\boldsymbol{C}_2^{-1}]\boldsymbol{\sigma}_0.\end{aligned} \tag{2-26}$$

Improved, and in fact sharp, upper and lower bounds on the elastic energy $\boldsymbol{\epsilon}_0 : \boldsymbol{C}_*\boldsymbol{\epsilon}_0$ in terms of the given applied strain $\boldsymbol{\epsilon}_0$ and sharp upper and lower bounds on the complementary elastic energy $\boldsymbol{\sigma}_0 : \boldsymbol{C}_*^{-1}\boldsymbol{\sigma}_0$ in terms of the given applied stress $\boldsymbol{\sigma}_0$ were obtained for isotropic component materials by Gibiansky and Cherkaev [1997a], Kohn and Lipton [1988], and Allaire and Kohn [1993a; 1993b; 1994]. The paper of Gibiansky and Cherkaev [1997a] was for the fourth-order plate equation, but this can be mapped to the equivalent 2-dimensional elasticity problem considered by Allaire and Kohn [1993b]. Their lower bounds on $\boldsymbol{\sigma}^0 : \boldsymbol{C}_*^{-1}\boldsymbol{\sigma}^0$ are equivalent to the bounds that for any tensor $\boldsymbol{C}_* \in GU_f$,

$$\boldsymbol{\sigma}^0 : \boldsymbol{C}_*^{-1}\boldsymbol{\sigma}^0 \geq \boldsymbol{\sigma}^0 : [\widetilde{\boldsymbol{C}}_f^A(\boldsymbol{\sigma}^0, 0, 0)]^{-1}\boldsymbol{\sigma}^0, \tag{2-27}$$

and they provided an explicit formula for the right-hand side for any 2×2 symmetric matrix $\boldsymbol{\sigma}^0$ representing the applied stress. This bound can be viewed in two ways: in the way originally interpreted, i.e., as a bound on the possible (elastic energy, average stress, volume fraction) triplets; or as a bound

$$\boldsymbol{\sigma}^0 : \boldsymbol{\epsilon}^0 \geq \boldsymbol{\sigma}^0 : [\widetilde{\boldsymbol{C}}_f^A(\boldsymbol{\sigma}^0, 0, 0)]^{-1}\boldsymbol{\sigma}^0 \tag{2-28}$$

on the possible (average stress, average strain, volume fraction) triplets. Here $\boldsymbol{\epsilon}^0 = \boldsymbol{C}_*^{-1}\boldsymbol{\sigma}^0$ is the strain associated with $\boldsymbol{\sigma}^0$. Significantly, Milton, Serkov, and Movchan [Milton et al. 2003] found that the inequality (2-28) completely characterizes the possible (average stress, average strain, volume fraction) triplets in the limit in which one phase becomes void, when the other phase is isotropic. Specifically, given any triplet $(\boldsymbol{\sigma}^0, \boldsymbol{\epsilon}^0, f)$ satisfying (2-28) as an inequality, they give a recipe for constructing a 2-dimensional microstructure with effective tensor $\boldsymbol{C}_*$ and having phase 1 occupy a volume fraction f such that $\boldsymbol{\sigma}^0 = \boldsymbol{C}_*\boldsymbol{\epsilon}^0$.

For 3-dimensional composites explicit expressions for the optimal upper energy bound were found by Gibiansky and Cherkaev [1997b] and Allaire [1994] for the case of a two-phase composite where one of the phases is void or rigid [Gibiansky and Cherkaev 1997b]. Grabovsky [1996] obtained energy bounds for two-phase composites containing anisotropic phases, each with a constant orientation.

Another major advance was made by Milton and Cherkaev [1995], who showed that any desired positive definite fourth-order tensor which has the symmetries of an elasticity tensor could be realized as the effective elasticity tensor $\boldsymbol{C}_*$ of a composite of a sufficiently stiff isotropic material and a sufficiently compliant isotropic material. One key to this advance was the realization that certain structures called pentamode materials could be (arbitrarily) stiff to one applied stress $\boldsymbol{\sigma}_1^0$ and yet have five mutually orthogonal strains $\boldsymbol{\epsilon}_1^0$, $\boldsymbol{\epsilon}_2^0$, $\boldsymbol{\epsilon}_3^0$, $\boldsymbol{\epsilon}_4^0$, $\boldsymbol{\epsilon}_5^0$, each orthogonal to $\boldsymbol{\sigma}_1^0$ as five (arbitrarily compliant) easy modes of deformations (hence the name pentamode).

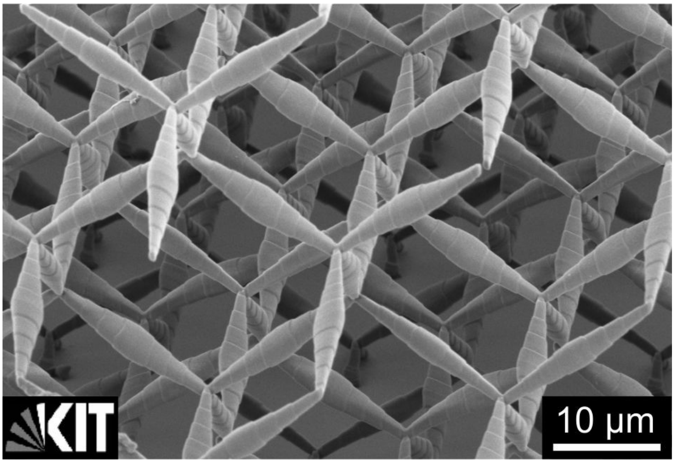

Figure 1. An electron micrograph of the pentamode structure created by Kadic, Bückmann, Stenger, Thiel and Wegener [Kadic et al. 2012] using a 3-dimensional lithography technique. (Used with the kind permission of Martin Wegener.)

For such a pentamode

$$W_f^5(\boldsymbol{\sigma}_1^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0, \boldsymbol{\epsilon}_5^0) = \min_{\boldsymbol{C}_* \in GU_f} \left[\left(\sum_{i=1}^{5} \boldsymbol{\epsilon}_i^0 : \boldsymbol{C}_* \boldsymbol{\epsilon}_i^0 \right) + \boldsymbol{\sigma}_1^0 : \boldsymbol{C}_*^{-1} \boldsymbol{\sigma}_1^0 \right] \tag{2-29}$$

approaches zero as the constituent stiff isotropic material becomes increasingly stiff and the constituent compliant isotropic material becomes increasingly compliant. The lattice structure of a pentamode is similar to that of diamond with a stiff double cone structure replacing each carbon bond. This structure ensures that the tips of four double cone structures meet at each vertex. This is the essential feature: treating the double cone structures as struts, the tension in one determines uniquely the tension in the other three. This is simply the balance of forces. Thus the structure as a whole can essentially support only one stress. Pentamode structures were experimentally realized by Kadic, Bückmann, Stenger, Thiel and Wegener [Kadic et al. 2012] in an incredible feat of precision three-dimensional lithography. One of their electron micrographs of the structure is shown in Figure 1. Pentamode structures were also independently discovered in 1995 by Sigmund, although he did not find the complete span of pentamode structures needed here: one needs pentamodes that can support any chosen stress, not just a hydrostatic one. It is this aspect of pentamodes that makes them more interesting than, for example, a gel. Gels are examples of pentamodes as they are easy to shear, but difficult to compress under a hydrostatic loading $\boldsymbol{\sigma}_1 = \boldsymbol{I}$. By contrast the pentamodes of Milton and Cherkaev could be stiff to any desired stress $\boldsymbol{\sigma}_1^0$: this desired stress may be a mixture of shear and compression, and may have eigenvalues of mixed signs. A simple argument for seeing that these pentamodes can achieve any desired elasticity tensor was given in the foreword of the book edited by Phani and Hussein [2017]. To recapitulate that argument, one expresses the desired $\boldsymbol{C}_*$ in terms of its eigenvectors and eigenvalues,

$$\boldsymbol{C}_* = \sum_{i=1}^{6} \lambda_i \boldsymbol{v}_i \otimes \boldsymbol{v}_i. \tag{2-30}$$

The idea, roughly speaking, is to find six pentamode structures each supporting a stress represented by the vector $\boldsymbol{v}_i$ for $i = 1, 2, \dots, 6$. The stiffness of the material and the necks of the junction regions at the vertices need to be adjusted so each pentamode structure has an effective elasticity tensor close to

$$\boldsymbol{C}_*^{(i)} = \lambda_i \boldsymbol{v}_i \otimes \boldsymbol{v}_i. \tag{2-31}$$

Then one successively superimposes all these six pentamode structures, with their lattice structures being offset to avoid collisions. Additionally one may need to deform the structures appropriately to avoid these collisions as described in [Milton and Cherkaev 1995], and when one does this it is necessary to readjust the stiffness of the material in the structure to maintain the value of λ_i. Then the remaining void in the structure is replaced by an extremely compliant material. (Its presence is needed just for technical reasons, to ensure that the assumptions of homogenization theory are valid so that the elastic properties can be described by an effective tensor.) But it is so compliant that essentially the effective elasticity tensor is just a sum of the effective elasticity tensors of the six pentamodes, i.e., the elastic interaction between the six pentamodes is negligible. In this way we arrive at a material with (approximately) the desired elasticity tensor $\boldsymbol{C}_*$.

It is worth mentioning that with extremely high contrast materials the homogenized equations are not necessarily the usual linear elasticity equations, but can also include nonlocal terms. Nonlocal interactions can be obtained for example with an extremely stiff dumbbell-shaped inclusion with the balls arbitrarily distant. If the bar joining them is not only extremely stiff but also extremely thin, then it does not directly couple with the surrounding elastic material (except in the very near vicinity of the bar, where it is obviously deformed by it), but provides a nonlocal interaction between the balls. In fact, amazingly, Camar-Eddine and Seppecher [2003] have completely characterized all possible linear macroscopic behaviors of any high contrast composite: they showed that any energetically stable behavior can be obtained using materials with such dumbbell-shaped inclusions interacting at many length scales. Some interesting examples of high contrast materials with exotic effective behaviors have been given by Seppecher, Alibert, and dell'Isola [Seppecher et al. 2011].

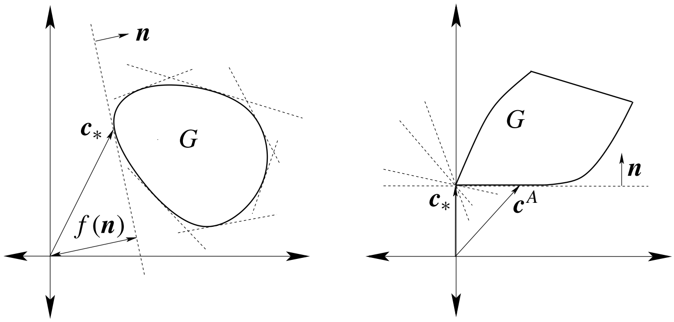

Figure 2. Left: A convex set is the envelope of its tangent planes. The positions of the two tangent planes with normal $\boldsymbol{n}$ are determined by the Legendre transform $f(\boldsymbol{n})$ and $f(-\boldsymbol{n})$ defined by (3-1). Specifically $f(\boldsymbol{n})$ and $f(-\boldsymbol{n})$ give the distances of the tangent planes from the origin. Right: An example highlighting an interesting case discussed in the text that helps give a geometrical interpretation of the results of the paper.

3. Characterizing convex sets and G-closures for elasticity

Let G be a convex set of real d-dimensional vectors, meaning that if $\boldsymbol{c}_1, \boldsymbol{c}_2 \in G$ then $\theta \boldsymbol{c}_1 + (1-\theta)\boldsymbol{c}_2 \in G$ for all $\theta \in [0, 1]$. As shown in Figure 2 (left) for $d = 2$ such a convex set can be completely characterized by its Legendre transform,

$$f(\boldsymbol{n}) = \min_{\boldsymbol{c}\in G} \boldsymbol{n} \cdot \boldsymbol{c}. \tag{3-1}$$

Clearly this function satisfies the homogeneity property that

$$f(\lambda \boldsymbol{n}) = \lambda f(\boldsymbol{n}) \quad \text{for all } \lambda > 0, \tag{3-2}$$

and consequently it suffices to know $f(\boldsymbol{n})$ for all unit vectors $\boldsymbol{n}$ to recover the function $f(\boldsymbol{n})$ for any vector $\boldsymbol{n}$. The values of $f(\boldsymbol{n})$ and $f(-\boldsymbol{n})$ give the positions of the two planes with normals $\pm\boldsymbol{n}$ that are tangent to G: specifically $|f(\boldsymbol{n})|$ and $|f(-\boldsymbol{n})|$ give the distances from these tangent planes to the origin. By varying $\boldsymbol{n}$ and taking the intersection of the regions between the planes one recovers G: the set G is the envelope of its tangent planes as illustrated in Figure 2 (left). Thus the Legendre transform function $f(\boldsymbol{n})$ with $|\boldsymbol{n}| = 1$ completely characterizes G.

The example of Figure 2 (right) is also illuminating for the purposes of this paper. Let $\boldsymbol{n}$ and $\boldsymbol{m}$ be the vectors

$$\boldsymbol{n} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \boldsymbol{m} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{3-3}$$

and consider $f(\boldsymbol{n}+\alpha\boldsymbol{m})$ for $\alpha \geq 0$ in the context of this example. (Of course $\boldsymbol{n}+\alpha\boldsymbol{m}$ is only a unit vector when $\alpha = 0$.) As the boundary of G contains a flat section orthogonal to $\boldsymbol{n}$, the vector $\boldsymbol{c}$ which attains the minimum in (3-1) is not unique. In the diagram both $\boldsymbol{c}^A$ and $\boldsymbol{c}_*$ are minimizers. However, for an infinitesimal value of $\alpha > 0$, $\boldsymbol{c}_*$ is selected as the unique minimizer and remains the minimizer no matter how large $\alpha > 0$ becomes. Furthermore, since $\boldsymbol{c}_*$ is orthogonal to $\boldsymbol{m}$ the value of $f(\boldsymbol{n}+\alpha\boldsymbol{m})$ remains constant for all $\alpha \geq 0$.

If G is a convex set of, say, real $d \times d$ matrices it can be similarly characterized by its Legendre transform,

$$f(\boldsymbol{N}) = \min_{\boldsymbol{C}\in G}(\boldsymbol{N}, \boldsymbol{C}), \tag{3-4}$$

defined for all $d \times d$ matrices $\boldsymbol{N}$, where $(\boldsymbol{N}, \boldsymbol{C})$ is an inner product on the space of matrices which we may take to be

$$(\boldsymbol{N}, \boldsymbol{C}) = N_{ij}C_{ij} \equiv \boldsymbol{N} : \boldsymbol{C}, \tag{3-5}$$

where we have adopted the Einstein summation convention that sums over repeated indices are assumed, and the double dot ":" denotes a double contraction of indices. This is exactly equivalent to (3-1) if we think of the matrix $\boldsymbol{C}$ being represented by the vector $\boldsymbol{c}$ of its matrix elements. Note that if G only contains symmetric matrices, then it suffices to take $\boldsymbol{N}$ as a symmetric matrix since $(\boldsymbol{A}, \boldsymbol{C}) = 0$ if C is symmetric and A is antisymmetric.

Similarly, if G is a convex set of fourth-order elasticity tensors $\boldsymbol{C}$ satisfying the usual symmetries

$$C_{ijk\ell} = C_{jik\ell} = C_{k\ell ij}, \tag{3-6}$$

then it can be characterized by the Legendre transform (3-4) with an inner product

$$(\boldsymbol{N}, \boldsymbol{C}) = N_{ijk\ell}C_{ijk\ell}, \tag{3-7}$$

and again it suffices to assume $\boldsymbol{N}$ has the same symmetries as $\boldsymbol{C}$, i.e., those in (3-6).

However, G-closures (i.e., sets of all possible effective tensors) are not generally convex sets. Nevertheless, they do have some convexity properties as a consequence of their stability under lamination. In the case of the set GU_f where $U = \{\boldsymbol{C}_1, \delta\boldsymbol{C}_2\}$, we can take two materials with effective tensors $\boldsymbol{C}_1^*, \boldsymbol{C}_2^* \in GU_f$ and laminate them together in a direction $\boldsymbol{n}$ (representing the vector perpendicular to the layers) in proportions θ and $1-\theta$ to obtain an effective tensor $\boldsymbol{C}_*(\boldsymbol{n}, \theta)$ which necessarily lies in the set GU_f for all $\theta \in [0, 1]$. While $\boldsymbol{C}_*(\boldsymbol{n}, \theta)$ is not a linear average of $\boldsymbol{C}_1^*$ and $\boldsymbol{C}_2^*$, there exist fractional linear transformations $T_{\boldsymbol{n}}$ of fourth-order tensors such that lamination in direction $\boldsymbol{n}$ reduces to a linear average [Backus 1962; Milton 1990] (see also [Tartar 1979]):

$$T_{\boldsymbol{n}}(\boldsymbol{C}_*(\boldsymbol{n}, \theta)) = \theta T_{\boldsymbol{n}}(\boldsymbol{C}_1^*) + (1-\theta)T_{\boldsymbol{n}}(\boldsymbol{C}_2^*) \quad \text{for all } \theta \in [0, 1]. \tag{3-8}$$

Thus $T_{\boldsymbol{n}}(GU_f)$ must be a convex set of fourth-order tensors. In the particular case where a set of effective tensors has no interior, i.e., is constrained to lie on a manifold of dimension m smaller than the dimension of the space of fourth-order tensors satisfying the symmetries of elasticity tensors (i.e., $m < 21$ for 3-dimensional composites and $m < 6$ for 2-dimensional composites), then as recognized by Grabovsky [1998] (see also [Grabovsky and Sage 1998]) $T_{\boldsymbol{n}}$ must map this manifold to a subset of a hyperplane of dimension m for any value of $\boldsymbol{n}$. This places rather severe constraints on the form of such manifolds. Identifying such manifolds is important as they represent exact relations satisfied by effective tensors, no matter what the geometry of the composite happens to be. Thus these constraints provide necessary conditions for an exact relation. Later, sufficient conditions for an exact relation to hold were obtained [Grabovsky et al. 2000].

Unfortunately, the use of Legendre transforms of the convex set $T_{\boldsymbol{n}}(GU_f)$ is not useful to us as we are unaware of any direct variational principles for $T_{\boldsymbol{n}}(\boldsymbol{C}_*)$. An alternative approach was prompted by work of Cherkaev and Gibiansky [1992; 1993], who found that bounding sums of energies and complementary energies could lead to very useful bounds on G-closures. It was proved by Francfort and Milton [Francfort and Milton 1994; Milton 1994] that minimums over $\boldsymbol{C}_* \in GU_f$ of such sums of energies and complementary energies completely characterize GU_f in much the same way that Legendre transforms characterize convex sets: the stability under lamination of GU_f is what allows one to recover GU_f from the values of these minimums (see also Chapter 30 in [Milton 2002]). Figure 3 captures the idea of this characterization.

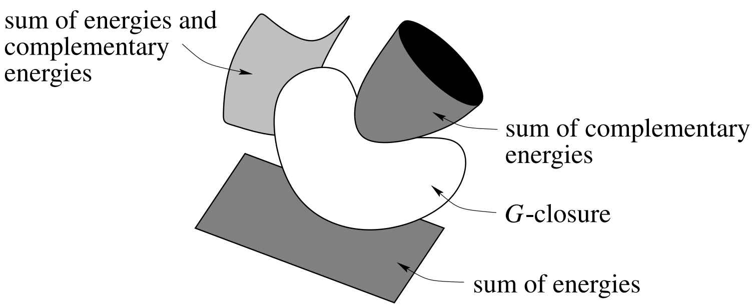

Figure 3. G-closures are characterized by minimums of sums of energies and complementary energies. The coordinates here represent the elements of the effective elasticity tensor $\boldsymbol{C}_*$. Then a plane represents a surface where a sum of energies is constant, and when this sum takes its minimum value the plane is tangent to the G-closure. The convexity properties of the G-closure guarantee that the surfaces corresponding to the minimums of sums of energies and complementary energies wrap around the G-closure and touch each point on its boundary. (Reproduction of Figure 30.1 in [Milton 2002].)

Specifically, in the case of 3-dimensional elasticity, the set GU_f is completely characterized if we know the seven "energy functions",

$$\begin{aligned}
W_f^0(\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0,\boldsymbol{\sigma}_3^0,\boldsymbol{\sigma}_4^0,\boldsymbol{\sigma}_5^0,\boldsymbol{\sigma}_6^0) &= \min_{\boldsymbol{C}_*\in GU_f}\sum_{j=1}^{6}\boldsymbol{\sigma}_j^0:\boldsymbol{C}_*^{-1}\boldsymbol{\sigma}_j^0,\\
W_f^1(\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0,\boldsymbol{\sigma}_3^0,\boldsymbol{\sigma}_4^0,\boldsymbol{\sigma}_5^0,\boldsymbol{\epsilon}_1^0) &= \min_{\boldsymbol{C}_*\in GU_f}\Bigg[\boldsymbol{\epsilon}_1^0:\boldsymbol{C}_*\boldsymbol{\epsilon}_1^0+\sum_{j=1}^{5}\boldsymbol{\sigma}_j^0:\boldsymbol{C}_*^{-1}\boldsymbol{\sigma}_j^0\Bigg],\\
W_f^2(\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0,\boldsymbol{\sigma}_3^0,\boldsymbol{\sigma}_4^0,\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0) &= \min_{\boldsymbol{C}_*\in GU_f}\Bigg[\sum_{i=1}^{2}\boldsymbol{\epsilon}_i^0:\boldsymbol{C}_*\boldsymbol{\epsilon}_i^0+\sum_{j=1}^{4}\boldsymbol{\sigma}_j^0:\boldsymbol{C}_*^{-1}\boldsymbol{\sigma}_j^0\Bigg],\\
W_f^3(\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0,\boldsymbol{\sigma}_3^0,\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0,\boldsymbol{\epsilon}_3^0) &= \min_{\boldsymbol{C}_*\in GU_f}\Bigg[\sum_{i=1}^{3}\boldsymbol{\epsilon}_i^0:\boldsymbol{C}_*\boldsymbol{\epsilon}_i^0+\sum_{j=1}^{3}\boldsymbol{\sigma}_j^0:\boldsymbol{C}_*^{-1}\boldsymbol{\sigma}_j^0\Bigg],\\
W_f^4(\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0,\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0,\boldsymbol{\epsilon}_3^0,\boldsymbol{\epsilon}_4^0) &= \min_{\boldsymbol{C}_*\in GU_f}\Bigg[\sum_{i=1}^{4}\boldsymbol{\epsilon}_i^0:\boldsymbol{C}_*\boldsymbol{\epsilon}_i^0+\sum_{j=1}^{2}\boldsymbol{\sigma}_j^0:\boldsymbol{C}_*^{-1}\boldsymbol{\sigma}_j^0\Bigg],\\
W_f^5(\boldsymbol{\sigma}_1^0,\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0,\boldsymbol{\epsilon}_3^0,\boldsymbol{\epsilon}_4^0,\boldsymbol{\epsilon}_5^0) &= \min_{\boldsymbol{C}_*\in GU_f}\Bigg[\Bigg(\sum_{i=1}^{5}\boldsymbol{\epsilon}_i^0:\boldsymbol{C}_*\boldsymbol{\epsilon}_i^0\Bigg)+\boldsymbol{\sigma}_1^0:\boldsymbol{C}_*^{-1}\boldsymbol{\sigma}_1^0\Bigg],\\
W_f^6(\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0,\boldsymbol{\epsilon}_3^0,\boldsymbol{\epsilon}_4^0,\boldsymbol{\epsilon}_5^0,\boldsymbol{\epsilon}_6^0) &= \min_{\boldsymbol{C}_*\in GU_f}\sum_{i=1}^{6}\boldsymbol{\epsilon}_i^0:\boldsymbol{C}_*\boldsymbol{\epsilon}_i^0.
\end{aligned}\tag{3-9}$$

In fact, it suffices [Milton and Cherkaev 1995] to know these functions for sets of applied strains $\boldsymbol{\epsilon}_i^0$ and applied stresses $\boldsymbol{\sigma}_j^0$ that are mutually orthogonal:

$$(\boldsymbol{\epsilon}_i^0,\boldsymbol{\sigma}_j^0)=0,\quad (\boldsymbol{\epsilon}_i^0,\boldsymbol{\epsilon}_k^0)=0,\quad (\boldsymbol{\sigma}_j^0,\boldsymbol{\sigma}_\ell^0)=0,\qquad \text{for all } i,j,k,\ell \text{ with } i\neq j,\ i\neq k,\ j\neq \ell. \tag{3-10}$$

Each of these terms in the minimums has a physical significance. For example, in the expression for W_f^2,

$$\sum_{i=1}^{2}\boldsymbol{\epsilon}_i^0:\boldsymbol{C}_*\boldsymbol{\epsilon}_i^0+\sum_{j=1}^{4}\boldsymbol{\sigma}_j^0:\boldsymbol{C}_*^{-1}\boldsymbol{\sigma}_j^0 \tag{3-11}$$

has the physical interpretation of being the sum of energies per unit volume stored in the composite with effective elasticity tensor $\boldsymbol{C}_*$ when successively subjected to the two applied strains $\boldsymbol{\epsilon}_1^0$ and $\boldsymbol{\epsilon}_2^0$ and then to the four applied stresses $\boldsymbol{\sigma}_1^0$, $\boldsymbol{\sigma}_2^0$, $\boldsymbol{\sigma}_3^0$ and $\boldsymbol{\sigma}_4^0$. To distinguish the terms $\boldsymbol{\epsilon}_i^0:\boldsymbol{C}_*\boldsymbol{\epsilon}_i^0$ and $\boldsymbol{\sigma}_j^0:\boldsymbol{C}_*^{-1}\boldsymbol{\sigma}_j^0$, the first is called an energy (it is really an energy per unit volume associated with the applied strain $\boldsymbol{\epsilon}_i^0$)

and the second is called a complementary energy, although it too physically represents an energy per unit volume associated with the applied stress $\boldsymbol{\sigma}_j^0$. Note that the quantity (3-11) can be equivalently written as

$$(\boldsymbol{C}_*, \boldsymbol{N}) + (\boldsymbol{C}_*^{-1}, \boldsymbol{N}'), \tag{3-12}$$

where

$$\boldsymbol{N} = \sum_{i=1}^{2} \boldsymbol{\epsilon}_i^0 \otimes \boldsymbol{\epsilon}_i^0, \quad \boldsymbol{N}' = \sum_{j=1}^{4} \boldsymbol{\sigma}_j^0 \otimes \boldsymbol{\sigma}_j^0, \tag{3-13}$$

in which for any $d \times d$ symmetric matrix $\boldsymbol{A}$, the tensor $\boldsymbol{A} \otimes \boldsymbol{A}$ is defined to be the fourth-order tensor with elements

$$\{\boldsymbol{A} \otimes \boldsymbol{A}\}_{ijk\ell} = \{\boldsymbol{A}\}_{ij}\{\boldsymbol{A}\}_{k\ell}. \tag{3-14}$$

If we decompose the positive semidefinite tensors $\boldsymbol{N}$ and $\boldsymbol{N}'$ into their spectral decompositions

$$\boldsymbol{N} = \sum_{i=1}^{2} \lambda_i \boldsymbol{v}_i \otimes \boldsymbol{v}_i, \quad \boldsymbol{N}' = \sum_{j=1}^{4} \lambda_j' \boldsymbol{v}_j' \otimes \boldsymbol{v}_j', \tag{3-15}$$

with eigenmatrices $\boldsymbol{v}_i$ and $\boldsymbol{v}_j'$ and corresponding nonnegative eigenvalues λ_i and λ_j', then, with the orthogonality constraints (3-10), we can make the identifications

$$\boldsymbol{\epsilon}_i^0 = \sqrt{\lambda_i}\boldsymbol{v}_i, \quad \boldsymbol{\sigma}_j^0 = \sqrt{\lambda_j'}\boldsymbol{v}_j. \tag{3-16}$$

Note that due to the orthogonality conditions (3-10) the fourth-order tensors $\boldsymbol{N}$ and $\boldsymbol{N}'$ have the property that the product $\boldsymbol{N}\boldsymbol{N}'$ is zero. Here the product of two fourth-order tensors $\boldsymbol{C}$ and $\boldsymbol{C}'$ is given by

$$\{\boldsymbol{C}\boldsymbol{C}'\}_{ijk\ell} = \{\boldsymbol{C}\}_{ijmn}\{\boldsymbol{C}'\}_{mnk\ell}. \tag{3-17}$$

Thus in the same way that convex sets are the envelope of planes, the G-closure GU_f is the envelope of special surfaces parametrized by positive semidefinite fourth-order tensors $\boldsymbol{N}$ and $\boldsymbol{N}'$ satisfying the symmetries of elasticity tensors, and having zero product $\boldsymbol{N}\boldsymbol{N}' = \boldsymbol{N}'\boldsymbol{N} = 0$ (i.e., the range of $\boldsymbol{N}'$ is in the null space of $\boldsymbol{N}$, and conversely the range of $\boldsymbol{N}$ is in the null space of $\boldsymbol{N}'$). These special surfaces consist of all positive definite fourth-order tensors $\boldsymbol{C}$ satisfying

$$(\boldsymbol{C}, \boldsymbol{N}) + (\boldsymbol{C}^{-1}, \boldsymbol{N}') = c, \tag{3-18}$$

where c is a positive real constant. In the case $\boldsymbol{N}' = 0$ this does represent a hyperplane, but its orientation is restricted by the fact that the outward normal to the surface $\boldsymbol{N}$ is restricted to be a positive definite fourth-order tensor (by outward normal we mean the normal pointing away from the origin). Knowledge of the seven functions W_f^i given by (3-9) is clearly equivalent to knowledge of the function

$$W_f(\boldsymbol{N}, \boldsymbol{N}') = \min_{\boldsymbol{C}_* \in GU_f} (\boldsymbol{C}_*, \boldsymbol{N}) + (\boldsymbol{C}_*^{-1}, \boldsymbol{N}') \tag{3-19}$$

for all positive semidefinite fourth-order tensors $\boldsymbol{N}$ and $\boldsymbol{N}'$ satisfying the symmetries of elasticity tensors and having $\boldsymbol{N}\boldsymbol{N}' = 0$. The formula for recovering GU_f from $W_f(\boldsymbol{N}, \boldsymbol{N}')$ is then

$$\bigcap_{\substack{\boldsymbol{N},\boldsymbol{N}' \geq 0 \\ \boldsymbol{N}\boldsymbol{N}'=0}} \left\{\boldsymbol{C} : (\boldsymbol{C}, \boldsymbol{N}) + (\boldsymbol{C}^{-1}, \boldsymbol{N}') \geq W_f(\boldsymbol{N}, \boldsymbol{N}')\right\} = GU_f. \tag{3-20}$$

More generally if we replace GU_f in (3-19) by another set G of positive definite matrices, and if the left-hand side of (3-20) is again G, then we may say G is "W-convex".

An explicit definition of W-convexity is as follows: a set G of positive definite symmetric matrices is said to be strictly W-convex if G is simply connected and if for every pair of positive semidefinite symmetric matrices $\boldsymbol{N}$ and $\boldsymbol{N}'$, not both zero, the minimum in

$$\min_{\boldsymbol{C} \in G} (\boldsymbol{C}, \boldsymbol{N}) + (\boldsymbol{C}^{-1}, \boldsymbol{N}') \tag{3-21}$$

is uniquely attained by only one $\boldsymbol{C} \in G$. Geometrically, G is strictly W-convex if for all positive semidefinite symmetric matrices $\boldsymbol{N}$ and $\boldsymbol{N}'$, not both zero, the surface that consists of all positive definite matrices $\boldsymbol{C}$ satisfying

$$(\boldsymbol{C}, \boldsymbol{N}) + (\boldsymbol{C}^{-1}, \boldsymbol{N}') = k, \tag{3-22}$$

where k is chosen as the smallest value for which this surface touches G, has the property that it touches G at only one point. A set G is W-convex if it is a limit of strictly W-convex sets. If the set G has a smooth boundary, then the condition for W-convexity can be expressed in terms of the curvature of the boundary of G: when G is a set of matrices, this curvature at each point on the surface of G is a fourth-order tensor; when G is a set of fourth-order elasticity tensors, this curvature is an eighth-order tensor. (See equation (3.51) in [Milton 1994], or equation (30.11) in [Milton 2002], for the explicit inequalities that the curvature must satisfy.)

The stability of GU_f under lamination implies it is W-convex, but W-convexity probably does not imply stability under lamination, as stability under lamination depends on the underlying partial differential equations. Associated with any set G of symmetric positive definite matrices $\boldsymbol{C}$ is its W-transform, defined as

$$W(\boldsymbol{N}, \boldsymbol{N}') = \min_{\boldsymbol{C} \in G} (\boldsymbol{C}, \boldsymbol{N}) + (\boldsymbol{C}^{-1}, \boldsymbol{N}'), \tag{3-23}$$

where $\boldsymbol{N}$ and $\boldsymbol{N}'$ are symmetric positive semidefinite matrices satisfying $\boldsymbol{N}\boldsymbol{N}' = 0$, and the inner product of two symmetric matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ can be taken as $(\boldsymbol{A}, \boldsymbol{B}) =$

$\mathrm{Tr}(\boldsymbol{AB})$, where Tr denotes the trace (sum of diagonal elements) of a matrix. To see some of the properties of W-transforms it is helpful to extend the definition of the transform to allow for matrices $\boldsymbol{N}$ and $\boldsymbol{N}'$ that have a nonzero product, $\boldsymbol{N}\boldsymbol{N}' \neq 0$. The defining equation, (3-23), remains the same. Then consider a weighted average of $(\boldsymbol{N}_1, \boldsymbol{N}_1')$ and $(\boldsymbol{N}_2, \boldsymbol{N}_2')$, with weights θ and $1-\theta$, where the four matrices $\boldsymbol{N}_1, \boldsymbol{N}_1', \boldsymbol{N}_2, \boldsymbol{N}_2'$ are positive semidefinite. Then for any $\theta \in (0, 1)$, we have

$$\begin{aligned} W(\theta\boldsymbol{N}_1+(1-\theta)\boldsymbol{N}_2, \theta\boldsymbol{N}_1'+(1-\theta)\boldsymbol{N}_2') &= \min_{\boldsymbol{C}\in G}\bigl\{\theta[(\boldsymbol{C},\boldsymbol{N}_1)+(\boldsymbol{C}^{-1},\boldsymbol{N}_1')]+(1-\theta)[(\boldsymbol{C},\boldsymbol{N}_2)+(\boldsymbol{C}^{-1},\boldsymbol{N}_2')]\bigr\} \\ &\geq \theta\bigl\{\min_{\boldsymbol{C}\in G}(\boldsymbol{C},\boldsymbol{N}_1)+(\boldsymbol{C}^{-1},\boldsymbol{N}_1')\bigr\}+(1-\theta)\bigl\{\min_{\boldsymbol{C}\in G}(\boldsymbol{C},\boldsymbol{N}_2)+(\boldsymbol{C}^{-1},\boldsymbol{N}_2')\bigr\} \\ &\geq \theta W_f(\boldsymbol{N}_1,\boldsymbol{N}_1')+(1-\theta)W_f(\boldsymbol{N}_2,\boldsymbol{N}_2'), \end{aligned} \tag{3-24}$$

which (by definition) implies $W(\boldsymbol{N}, \boldsymbol{N}')$ is a jointly concave function of $\boldsymbol{N}$ and $\boldsymbol{N}'$. This concavity is a well-known property of Legendre transforms.

4. Variational principles

Upper bounds on the sums of energies and complementary energies can easily be obtained from classic energy minimization variational principles. For example, in the case of the sum (3-11), we have

$$\begin{aligned} &\sum_{i=1}^{2}\boldsymbol{\epsilon}_i^0 : \boldsymbol{C}_*\boldsymbol{\epsilon}_i^0+\sum_{j=1}^{4}\boldsymbol{\sigma}_j^0 : \boldsymbol{C}_*^{-1}\boldsymbol{\sigma}_j^0 \\ &\quad = \min_{\underline{\boldsymbol{\epsilon}}_1,\underline{\boldsymbol{\epsilon}}_2,\underline{\boldsymbol{\sigma}}_1,\underline{\boldsymbol{\sigma}}_2,\underline{\boldsymbol{\sigma}}_3,\underline{\boldsymbol{\sigma}}_4}\left\langle\sum_{i=1}^{2}\underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x}):\boldsymbol{C}(\boldsymbol{x})\underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x})+\sum_{j=1}^{4}\underline{\boldsymbol{\sigma}}_j(\boldsymbol{x}):[\boldsymbol{C}(\boldsymbol{x})]^{-1}\underline{\boldsymbol{\sigma}}_j(\boldsymbol{x})\right\rangle, \end{aligned} \tag{4-1}$$

where the minimum is over a set of two trial strain fields $\underline{\boldsymbol{\epsilon}}_1(\boldsymbol{x})$ and $\underline{\boldsymbol{\epsilon}}_2(\boldsymbol{x})$ and a set of four trial stress fields $\underline{\boldsymbol{\sigma}}_1(\boldsymbol{x})$, $\underline{\boldsymbol{\sigma}}_2(\boldsymbol{x})$, $\underline{\boldsymbol{\sigma}}_3(\boldsymbol{x})$, and $\underline{\boldsymbol{\sigma}}_4(\boldsymbol{x})$ that have the prescribed average values

$$\langle\underline{\boldsymbol{\epsilon}}_i\rangle = \boldsymbol{\epsilon}_i^0 \quad \text{for } i = 1, 2, \qquad \langle\underline{\boldsymbol{\sigma}}_j\rangle = \boldsymbol{\sigma}_j^0 \quad \text{for } j = 1, 2, 3, 4, \tag{4-2}$$

and are subject to the differential constraints that

$$\begin{gathered} \underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x}) = \tfrac{1}{2}\bigl(\nabla\underline{\boldsymbol{u}}_i(\boldsymbol{x})+(\nabla\underline{\boldsymbol{u}}_i(\boldsymbol{x}))^T\bigr) \quad \text{for } i = 1, 2, \\ \nabla\cdot\underline{\boldsymbol{\sigma}}_j(\boldsymbol{x}) = 0 \quad \text{for } j = 1, 2, 3, 4, \end{gathered} \tag{4-3}$$

where T denotes the transpose (reflecting the matrix about its diagonal) and $\underline{\boldsymbol{u}}_i(\boldsymbol{x})$ is the trial displacement field associated with the trial stress field $\underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x})$. The trial strain fields $\underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x})$ and the trial stress fields $\underline{\boldsymbol{\sigma}}_j(\boldsymbol{x})$ (but not the trial displacement fields) should be chosen to be periodic (if the composite is periodic), quasiperiodic (if the composite is quasiperiodic), or statistically homogeneous (if the composite is statistically homogeneous). It may be the case that the material has structure on widely separated length scales. Maybe it can be viewed as a mixture of two composites, one with effective tensor $\boldsymbol{C}_*^1$ and a second with effective tensor $\boldsymbol{C}_*^2$, so that at the mesoscale it has a geometry described by a characteristic function $\chi_*(\boldsymbol{x})$, where $\chi_*(\boldsymbol{x})$ is 1 in the composite with effective tensor $\boldsymbol{C}_*^1$ and 0 in the material with effective tensor $\boldsymbol{C}_*^2$. Naturally the length scale, or length scales, of variations in $\chi_*(\boldsymbol{x})$ should be much larger than the variations in the microstructure of the materials that have the effective tensors $\boldsymbol{C}_*^1$ and $\boldsymbol{C}_*^2$. Then we can treat the material having effective tensor as a composite of the materials $\boldsymbol{C}_*^1$ and $\boldsymbol{C}_*^2$ and we have the variational principle

$$\begin{aligned}\sum_{i=1}^{2}\boldsymbol{\epsilon}_i^0:\boldsymbol{C}_*\boldsymbol{\epsilon}_i^0+\sum_{j=1}^{4}\boldsymbol{\sigma}_j^0:\boldsymbol{C}_*^{-1}\boldsymbol{\sigma}_j^0 \\ =\min_{\underline{\boldsymbol{\epsilon}}_1,\underline{\boldsymbol{\epsilon}}_2,\underline{\boldsymbol{\sigma}}_1,\underline{\boldsymbol{\sigma}}_2,\underline{\boldsymbol{\sigma}}_3,\underline{\boldsymbol{\sigma}}_4}\Big\langle\sum_{i=1}^{2}\underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x}):\big[\chi_*(\boldsymbol{x})\boldsymbol{C}_*^1+(1-\chi_*(\boldsymbol{x}))\boldsymbol{C}_*^2\big]\underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x}) \\ +\sum_{j=1}^{4}\underline{\boldsymbol{\sigma}}_j(\boldsymbol{x}):\big[\chi_*(\boldsymbol{x})\boldsymbol{C}_*^1+(1-\chi_*(\boldsymbol{x}))\boldsymbol{C}_*^2\big]^{-1}\underline{\boldsymbol{\sigma}}_j(\boldsymbol{x})\Big\rangle,\end{aligned} \tag{4-4}$$

where again the minimum is over fields subject to the appropriate average values and differential constraints. Particular choices of trial fields will then lead to an upper bound on this sum of energies and complementary energies. To bound the quantities on the right one may again use variational principles. When $\boldsymbol{x}$ is in the material $\boldsymbol{C}_*^k$ for $k=1$ or 2, one has the variational principles

$$\begin{aligned}\underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x}):\boldsymbol{C}_*^k\underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x})&=\min_{\underline{\underline{\boldsymbol{\epsilon}}}_i}\big\langle\underline{\underline{\boldsymbol{\epsilon}}}_i(\boldsymbol{x},\boldsymbol{y}):\boldsymbol{C}^k(\boldsymbol{y})\underline{\underline{\boldsymbol{\epsilon}}}_i(\boldsymbol{x},\boldsymbol{y})\big\rangle_{\boldsymbol{y}},\\ \underline{\boldsymbol{\sigma}}_j(\boldsymbol{x}):[\boldsymbol{C}_*^k]^{-1}\underline{\boldsymbol{\sigma}}_j(\boldsymbol{x})&=\min_{\underline{\underline{\boldsymbol{\sigma}}}_j}\big\langle\underline{\underline{\boldsymbol{\sigma}}}_j(\boldsymbol{x},\boldsymbol{y}):[\boldsymbol{C}^k(\boldsymbol{y})]^{-1}\underline{\underline{\boldsymbol{\sigma}}}_j(\boldsymbol{x},\boldsymbol{y})\big\rangle_{\boldsymbol{y}},\end{aligned} \tag{4-5}$$

where $\langle\cdot\rangle_{\boldsymbol{y}}$ now denotes an average over the $\boldsymbol{y}$ variable ($\boldsymbol{x}$ is the "slow variable" and $\boldsymbol{y}$ is the "fast variable") and

$$\boldsymbol{C}^k(\boldsymbol{y})=\chi^k(\boldsymbol{y})\boldsymbol{C}_1+(1-\chi^k(\boldsymbol{y}))\boldsymbol{C}_2, \tag{4-6}$$

in which $\chi^k(\boldsymbol{y})$ is the characteristic function representing the geometry associated with the effective tensor $\boldsymbol{C}_*^k$, taking a value 1 in the material with tensor $\boldsymbol{C}_1$ and 0 in the material with tensor $\boldsymbol{C}_2$. Here the trial fields have the prescribed average values

$$\langle\underline{\underline{\boldsymbol{\epsilon}}}_i(\boldsymbol{x},\boldsymbol{y})\rangle_{\boldsymbol{y}}=\underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x})\ \text{ for } i=1,2,\qquad \langle\underline{\underline{\boldsymbol{\sigma}}}_j(\boldsymbol{x},\boldsymbol{y})\rangle_{\boldsymbol{y}}=\underline{\boldsymbol{\sigma}}_j(\boldsymbol{x})\ \text{ for } j=1,2,3,4, \tag{4-7}$$

and are subject to the differential constraints

$$\begin{aligned}\underline{\underline{\boldsymbol{\epsilon}}}_i(\boldsymbol{x},\boldsymbol{y})&=\tfrac{1}{2}\big(\nabla_y\underline{\underline{\boldsymbol{u}}}_i(\boldsymbol{x},\boldsymbol{y})+(\nabla\underline{\underline{\boldsymbol{u}}}_i(\boldsymbol{x},\boldsymbol{y}))^T\big)\quad\text{for } i=1,2,\\ \nabla\cdot_y\underline{\underline{\boldsymbol{\sigma}}}_j(\boldsymbol{x},\boldsymbol{y})&=0\quad\text{for } j=1,2,3,4,\end{aligned} \tag{4-8}$$

where ∇_y and $\nabla\cdot_y$ are the gradient and divergence with respect to the $\boldsymbol{y}$ variables. We call the step of replacing the variational principle (4-1) by the variational principles (4-4) and (4-5) the "homogenization at intermediate scales step".

In this paper we will choose trial fields that satisfy the *local orthogonality condition* that

$$\underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x}):\underline{\boldsymbol{\sigma}}_j(\boldsymbol{x})=0, \quad \text{for all } \boldsymbol{x}. \tag{4-9}$$

Using the differential constraints satisfied by the trial fields, and integration by parts, one sees that the associated average fields are necessarily orthogonal too:

$$\boldsymbol{\epsilon}_i^0:\boldsymbol{\sigma}_j^0=\langle\underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x})\rangle:\langle\underline{\boldsymbol{\sigma}}_j(\boldsymbol{x})\rangle=\langle\underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x}):\underline{\boldsymbol{\sigma}}_j(\boldsymbol{x})\rangle=0. \tag{4-10}$$

5. Finding most of the energy functions

Recall from Section 2 that an complementary Avellaneda material is a sequentially layered laminate material with phase 1 occupying a volume fraction f and with effective tensor

$$\widetilde{\boldsymbol{C}}_f^A(\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0,\boldsymbol{\sigma}_3^0,\boldsymbol{\sigma}_4^0,\boldsymbol{\sigma}_5^0,0)$$

that attains equality in (2-13). It is found by minimizing the right-hand side of (2-13) as $\boldsymbol{C}_*$ varies within the class of tensors given by (2-15)–(2-17) with $\boldsymbol{C}_2=0$, as the rank r, the positive weights c_j which sum to 1, and the unit vectors $\boldsymbol{n}_i$ are varied. Here some of the applied stresses $\boldsymbol{\sigma}_j^0$ could be zero. Since the energy $\boldsymbol{\sigma}_j^0:\boldsymbol{C}_*^{-1}\boldsymbol{\sigma}_j^0$ associated with any applied stress $\boldsymbol{\sigma}_j^0$ is necessarily nonnegative, we obtain from (3-9) the bounds

$$\begin{aligned}
\sum_{j=1}^{5}\boldsymbol{\sigma}_j^0:[\widetilde{\boldsymbol{C}}_f^A(\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0,\boldsymbol{\sigma}_3^0,\boldsymbol{\sigma}_4^0,\boldsymbol{\sigma}_5^0,0)]^{-1}\boldsymbol{\sigma}_j^0&\le W_f^1(\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0,\boldsymbol{\sigma}_3^0,\boldsymbol{\sigma}_4^0,\boldsymbol{\sigma}_5^0,\boldsymbol{\epsilon}_1^0),\\
\sum_{j=1}^{4}\boldsymbol{\sigma}_j^0:[\widetilde{\boldsymbol{C}}_f^A(\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0,\boldsymbol{\sigma}_3^0,\boldsymbol{\sigma}_4^0,0,0)]^{-1}\boldsymbol{\sigma}_j^0&\le W_f^2(\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0,\boldsymbol{\sigma}_3^0,\boldsymbol{\sigma}_4^0,\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0),\\
\sum_{j=1}^{3}\boldsymbol{\sigma}_j^0:[\widetilde{\boldsymbol{C}}_f^A(\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0,\boldsymbol{\sigma}_3^0,0,0,0)]^{-1}\boldsymbol{\sigma}_j^0&\le W_f^3(\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0,\boldsymbol{\sigma}_3^0,\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0,\boldsymbol{\epsilon}_3^0),\\
\sum_{j=1}^{2}\boldsymbol{\sigma}_j^0:[\widetilde{\boldsymbol{C}}_f^A(\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0,0,0,0,0)]^{-1}\boldsymbol{\sigma}_j^0&\le W_f^4(\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0,\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0,\boldsymbol{\epsilon}_3^0,\boldsymbol{\epsilon}_4^0),\\
\boldsymbol{\sigma}_1^0:[\widetilde{\boldsymbol{C}}_f^A(\boldsymbol{\sigma}_1^0,0,0,0,0,0)]^{-1}\boldsymbol{\sigma}_1^0&\le W_f^5(\boldsymbol{\sigma}_1^0,\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0,\boldsymbol{\epsilon}_3^0,\boldsymbol{\epsilon}_4^0,\boldsymbol{\epsilon}_5^0),\\
0&\le W_f^6(\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0,\boldsymbol{\epsilon}_3^0,\boldsymbol{\epsilon}_4^0,\boldsymbol{\epsilon}_5^0,\boldsymbol{\epsilon}_6^0).
\end{aligned} \tag{5-1}$$

The last inequality is clearly sharp, being attained when the composite consists of islands of phase 1 surrounded by a phase 2 (so that $\boldsymbol{C}_*$ approaches 0 as $\delta\to 0$).

The objective of this paper is to show that many of the other inequalities are also sharp in the limit $\delta \to 0$, at least when the spaces spanned by the applied strains $\boldsymbol{\epsilon}_j^0$ for $j = 1, 2, \ldots, p$ satisfy certain properties. This space of applied strains $\mathcal{V}_p$, associated with W_f^p, has dimension p and is spanned by $\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \ldots, \boldsymbol{\epsilon}_p^0$.

The recipe for doing this is to simply insert into a relevant complementary Avellaneda material a microstructure occupying a thin walled region, such that the material can slip along the walls when the applied strain lies in appropriate spaces $\mathcal{V}_p$, yet which is such that the combination of Avellaneda material and walled material can support without slip any applied stress in the subspace orthogonal to $\mathcal{V}_p$. This will be possible only when $\mathcal{V}_p$ is spanned by symmetrized rank 1 matrices, taking the form

$$\boldsymbol{\epsilon}^{(k)} = \tfrac{1}{2}(\boldsymbol{a}_k \boldsymbol{n}_k^T + \boldsymbol{n}_k \boldsymbol{a}_k^T), \quad \text{for } k = 1, \ldots, p. \tag{5-2}$$

The existence of such matrices $\boldsymbol{\epsilon}^{(k)}$ is proved in Section 7. The proof uses small perturbations of the applied stresses and strains. But, due to the continuity of the energy functions W_f^k established in Section 9, the small perturbations do not modify the generic result. The vectors $\boldsymbol{n}_k$ determine the orientation of the walls in the structure. For each $\boldsymbol{n}_k$ there is a set of parallel walls perpendicular to $\boldsymbol{n}_k$ that allow slip given by the strain $\boldsymbol{\epsilon}^{(k)}$. We say slip but it should be recognized that $\boldsymbol{\epsilon}^{(k)}$ is not generally a pure shear, but rather a combination of dilation and shear, since it does not generally have zero trace.

To define the thin walled structure, introduce the periodic function $H_c(x)$ with period 1 which takes the value 1 if $x - [x] \leq c$, where $[x]$ is the greatest integer less than x, and $c \in [0, 1]$ gives the thickness of each wall relative to the spacing between walls (which is unity). Then for the unit vectors $\boldsymbol{n}_1, \boldsymbol{n}_2, \ldots, \boldsymbol{n}_p$ appearing in (5-2), and for a small relative wall thickness $c = \epsilon$, define the characteristic functions

$$\eta_k(\boldsymbol{x}) = H_\epsilon(\boldsymbol{x} \cdot \boldsymbol{n}_k + k/p). \tag{5-3}$$

This characteristic function defines a series of parallel walls, as shown on the left in Figure 4, each perpendicular to the vector $\boldsymbol{n}_j$, where $\eta_j(\boldsymbol{x}) = 1$ in the wall material. The additional shift term k/p in (5-3) ensures the walls associated with k_1 and k_2 do not intersect when it happens that $\boldsymbol{n}_{k_1} = \boldsymbol{n}_{k_2}$, at least when ϵ is small. Note that ϵ is a volume fraction, not a homogenization parameter. We will be taking the limit $\epsilon \to 0$ *after* taking the homogenization limit.

Now define the characteristic function

$$\chi_*(\boldsymbol{x}) = \prod_{k=1}^{p} (1 - \eta_k(\boldsymbol{x})). \tag{5-4}$$

If $p \leq 3$, this is usually a periodic function of $\boldsymbol{x}$, an exception being if $p = 3$ and there are no nonzero integers z_1, z_2, and z_3 such that $z_1\boldsymbol{n}_1 + z_2\boldsymbol{n}_2 + z_3\boldsymbol{n}_3 = 0$. More

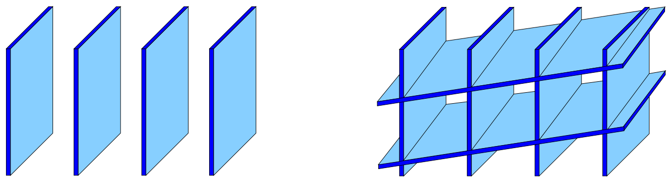

Figure 4. Example of walled structures. On the left we have a "rank 1" walled structure and on the right a "rank 2" walled structure. The generalization to walled structures of any rank is obvious, and precisely defined by the characteristic function (5-4) that is 0 in the walls, and 1 in the remaining material.

generally, $\chi_*(\boldsymbol{x})$ is a quasiperiodic function of $\boldsymbol{x}$. The walled structure is where $\chi_*(\boldsymbol{x})$ takes the value 0. In the case $p=2$ the walled structure is illustrated on the right in Figure 4.

Recall that a p-mode material is a material for which there are p independent strains to which the material is easily compliant, yet the material is much more resistant to any strain in the $(6-p)$-dimensional orthogonal subspace. In this sense the microstructure of Figure 1 is a pentamode material. We consider a subclass of multimode materials which can still support stresses in the limit $\delta \to 0$. We say a composite with effective tensor $\boldsymbol{C}_*$ built from the two materials $\boldsymbol{C}_1$ and $\boldsymbol{C}_2 = \delta \boldsymbol{C}_0$ is easily compliant to a strain $\boldsymbol{\epsilon}_i^0$ if the elastic energy $\boldsymbol{\epsilon}_i^0 : \boldsymbol{C}_* \boldsymbol{\epsilon}_i^0$ goes to zero as $\delta \to 0$, and supports a stress $\boldsymbol{\sigma}_j^0$ if the complementary energy $\boldsymbol{\sigma}_j^0 : \boldsymbol{C}_*^{-1} \boldsymbol{\sigma}_j^0$ has a nonzero limit as $\delta \to 0$. We desire p-mode materials for which there are p independent strains to which the material is easily compliant, yet for which the material supports any stress in the $(6-p)$-dimensional orthogonal subspace. The pentamode structure of Figure 1 needs to be modified as all its elastic moduli go to zero as $\delta \to 0$. The multimode structures we will introduce have structure on multiple length scales and it is important that one takes the limit of an infinite separation of length scales (so one can apply homogenization theory) before taking the limit $\delta \to 0$.

Inside the walled structure, where $\chi_*(\boldsymbol{x}) = 0$, we put a p-mode material with effective tensor $\boldsymbol{C}_*^2 = \boldsymbol{C}_*(\mathcal{V}_p)$ that supports any applied stress $\boldsymbol{\sigma}^0$ in the space orthogonal to $\mathcal{V}_p$ and which is easily compliant to any strain $\boldsymbol{\epsilon}^0$ in the space $\mathcal{V}_p$. When we take the six matrices

$$\boldsymbol{v}_1 = \boldsymbol{\sigma}_1^0/|\boldsymbol{\sigma}_1^0|, \dots, \boldsymbol{v}_{6-p} = \boldsymbol{\sigma}_{6-p}^0/|\boldsymbol{\sigma}_{6-p}^0|, \boldsymbol{v}_{7-p} = \boldsymbol{\epsilon}_1^0/|\boldsymbol{\epsilon}_1^0|, \dots, \boldsymbol{v}_6 = \boldsymbol{\epsilon}_p^0/|\boldsymbol{\epsilon}_p^0| \quad (5\text{-}5)$$

as an orthonormal basis for the space of 6×6 matrices, we need to find a p-mode material for which the elasticity tensor $\boldsymbol{C}_*^2$ in this basis is such that

$$\lim_{\delta \to 0} \boldsymbol{C}_*^2 = \begin{pmatrix} \boldsymbol{A} & 0 \\ 0 & 0 \end{pmatrix}, \quad (5\text{-}6)$$

where $\boldsymbol{A}$ represents a (strictly) positive definite $(6-p)\times(6-p)$ matrix and the 0 on the diagonal represents the $p\times p$ zero matrix.

Outside the walled structure, where $\chi_*(\boldsymbol{x})=1$, we put the complementary Avellaneda material with effective elasticity tensor

$$\boldsymbol{C}^1_* = \widetilde{\boldsymbol{C}}^A_f(\boldsymbol{\sigma}^0_1,\dots,\boldsymbol{\sigma}^0_{6-p},0,\dots,0).$$

In a variational principle similar to (4-4) (i.e., treating the complementary Avellaneda material and the p-mode material both as homogeneous materials with effective tensors $\boldsymbol{C}^1_*$ and $\boldsymbol{C}^2_*$, respectively) we choose trial stress fields that are constant,

$$\underline{\boldsymbol{\sigma}}_j(\boldsymbol{x}) = \boldsymbol{\sigma}^0_j, \tag{5-7}$$

thus trivially fulfilling the differential constraints, and trial strain fields of the form

$$\underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x}) = \sum_{k=1}^{p} \boldsymbol{\epsilon}_{i,k}\eta_k(\boldsymbol{x})/\epsilon, \tag{5-8}$$

which are required to have the average values

$$\boldsymbol{\epsilon}^0_i = \langle\underline{\boldsymbol{\epsilon}}_i\rangle = \sum_{k=1}^{p}\boldsymbol{\epsilon}_{i,k}, \tag{5-9}$$

and the matrices $\boldsymbol{\epsilon}_{i,k}$ have the form

$$\boldsymbol{\epsilon}_{i,k} = a_{i,k}\boldsymbol{\epsilon}^{(k)}, \tag{5-10}$$

for some choice of constants $a_{i,k}$ which ensures they are symmetrized rank 1 matrices lying in the space $\mathcal{V}_p$ (so they cost very little energy), and which ensures that the $\boldsymbol{\epsilon}^0_i$ given by (5-9) are orthogonal. This symmetrized rank 1 form ensures that $\underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x})$ derives from a displacement field. Specifically we have

$$\underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x}) = \tfrac{1}{2}\big(\nabla\underline{\boldsymbol{u}}_i(\boldsymbol{x}) + (\nabla\underline{\boldsymbol{u}}_i(\boldsymbol{x}))^T\big), \tag{5-11}$$

with

$$\underline{\boldsymbol{u}}_i(\boldsymbol{x}) = \sum_{k=1}^{p} a_{i,k}\boldsymbol{a}_k\big\{(\boldsymbol{n}_k\cdot\boldsymbol{x})\eta_k(\boldsymbol{x})/\epsilon + ([\boldsymbol{n}_k\cdot\boldsymbol{x}]+1)(1-\eta_k(\boldsymbol{x}))\big\}, \tag{5-12}$$

where, as before, $[\boldsymbol{n}_j\cdot\boldsymbol{x}]$ is the greatest integer less than $\boldsymbol{n}_j\cdot\boldsymbol{x}$. One can easily check that this displacement field is continuous at the wall interfaces.

To find upper bounds on the energy associated with this trial strain field, first consider those parts of the walled structure that are outside of any junction regions, i.e., where for some k we have $\eta_k(\boldsymbol{x})=1$, while $\eta_s(\boldsymbol{x})=0$ for all $s\neq k$. An upper bound for the volume fraction occupied by the region where $\eta_k(\boldsymbol{x})=1$ while $\eta_s(\boldsymbol{x})=0$ for all $s\neq k$ is of course ϵ, as this represents the volume of the region where $\eta_k(\boldsymbol{x}) = 1$. The associated energy per unit volume of the trial strain field in those parts of the walled structure that are outside of any junction regions is bounded above by

$$\sum_{k=1}^{p} \boldsymbol{\epsilon}_{i,k} : \boldsymbol{C}_*(\mathcal{V}_p)\boldsymbol{\epsilon}_{i,k}/\epsilon. \tag{5-13}$$

We will see in Section 8 that with an appropriate choice of multimode material, $\boldsymbol{\epsilon}_{i,k} : \boldsymbol{C}_*(\mathcal{V}_p)\boldsymbol{\epsilon}_{i,k}$ is bounded above by a quantity proportional to δ, essentially because all the strain is concentrated in phase 2. So we require that the limits $\delta \to 0$ and $\epsilon \to 0$ be taken so that $\delta/\epsilon \to 0$ to ensure that the quantity (5-13) goes to zero in this limit.

Next, consider those junction regions where only two walls meet, i.e., where for some k_1 and $k_2 > k_1$, $\boldsymbol{x}$ is such that $\eta_{k_1}(\boldsymbol{x}) = \eta_{k_2}(\boldsymbol{x}) = 1$ while $\eta_s(\boldsymbol{x}) = 0$ for all s not equal to k_1 or k_2. Provided $\boldsymbol{n}_{k_1} \neq \boldsymbol{n}_{k_2}$, an upper bound for the volume fraction occupied by each such junction region is ϵ^2. Then the associated energy per unit volume of the trial strain field in these junction regions where only two walls meet is bounded above by

$$\sum_{k_1=1}^{p} \sum_{k_2=k_1+1}^{p} (\boldsymbol{\epsilon}_{i,k_1} + \boldsymbol{\epsilon}_{i,k_2}) : \boldsymbol{C}_*(\mathcal{V}_p)(\boldsymbol{\epsilon}_{i,k_1} + \boldsymbol{\epsilon}_{i,k_2}). \tag{5-14}$$

Thus, the powers of ϵ cancel and this energy density goes to zero if the multimode material is easily compliant to the strains $\boldsymbol{\epsilon}_{i,k_1} + \boldsymbol{\epsilon}_{i,k_2}$ for all k_1 and k_2 with $k_2 > k_1$.

Finally, consider those junction regions where three or more walls meet, i.e., for some k_1, $k_2 > k_1$, and $k_3 > k_2$, $\boldsymbol{x}$ is such that $\eta_{k_i}(\boldsymbol{x}) = 1$ for $i = 1, 2, 3$. For a given choice of k_1, $k_2 > k_1$, and $k_3 > k_2$ such that the three vectors $\boldsymbol{n}_{k_1}$, $\boldsymbol{n}_{k_2}$, and $\boldsymbol{n}_{k_3}$ are not coplanar, an upper bound for the volume fraction occupied by this region is ϵ^3. In the case that the three vectors $\boldsymbol{n}_{k_1}$, $\boldsymbol{n}_{k_2}$, and $\boldsymbol{n}_{k_3}$ are coplanar, we can ensure that the volume fraction occupied by this region is ϵ^3 or less by appropriately translating one or two wall structures, i.e., by replacing $\eta_{k_m}(\boldsymbol{x})$ with $\eta_{k_m}(\boldsymbol{x} + \alpha_i \boldsymbol{n}_{k_m})$ for $m = 2, 3$, for an appropriate choice of α_2 and α_3 between 0 and 1. Since the energy density of the trial field in these regions scales as $\epsilon^3/\epsilon^2 = \epsilon$, we can ignore this contribution in the limit $\epsilon \to 0$ as it goes to zero too.

From this analysis of the energy densities associated with the trial fields it follows that one does not necessarily need the pentamode, quadramode, trimode, bimode, and unimode materials as appropriate for the material inside the walled structure. Instead, by modifying the construction, it suffices to use only unimode and bimode materials. In the walled structure we now put unimode materials in those sections where for some k we have $\eta_k(\boldsymbol{x}) = 1$ while $\eta_{k'}(\boldsymbol{x}) = 0$ for all $k' \neq k$. Each unimode material is easily compliant to the single strain $\boldsymbol{\epsilon}^{(k)}$ appropriate to the wall under consideration. A prescription for constructing 3-dimensional unimode materials that are multiple rank laminates, and which are easily compliant under any desired single strain, is given in Section 5.1 of [Milton and Cherkaev 1995]. In each junction region of the walled structure where $\eta_{k_1}(\boldsymbol{x}) = \eta_{k_2}(\boldsymbol{x}) = 1$ for some $k_1 \neq k_2$ while $\eta_k(\boldsymbol{x}) = 0$ for all k not equal to k_1 or k_2, we put a bimode material which is easily compliant to any strain in the subspace spanned by $\boldsymbol{\epsilon}^{(k_1)}$ and $\boldsymbol{\epsilon}^{(k_2)}$ as appropriate to the junction region under consideration. At present we do not know of any recipe in three dimensions for constructing bimode materials that have any desired pair of strains as their easy modes of deformation, other than to superimpose four pentamode structures as described in Section 8. In the remaining junction regions of the walled structure (where three or more walls intersect) we put phase 1. The contribution to the average energy of the fields in these regions vanishes as $\epsilon \to 0$ as discussed above.

By these constructions we effectively obtain materials with elasticity tensors $\boldsymbol{C}_*$ such that

$$\lim_{\delta \to 0} \boldsymbol{C}_* = (\boldsymbol{I} - \Pi_p)\widetilde{\boldsymbol{C}}_f^A(\boldsymbol{I} - \Pi_p), \tag{5-15}$$

where $\boldsymbol{I}$ is the fourth-order identity matrix, Π_p is the fourth-order tensor that is the projection onto the space $\mathcal{V}_p$, $\boldsymbol{I} - \Pi_p$ is the projection onto the orthogonal complement of $\mathcal{V}_p$, and $\widetilde{\boldsymbol{C}}_f^A$ is the relevant complementary Avellaneda material. In the basis (5-5) $\boldsymbol{I} - \Pi_p$ is represented by the 6×6 matrix that has the block form

$$\boldsymbol{I} - \Pi_p = \begin{pmatrix} \boldsymbol{I}_{6-p} & 0 \\ 0 & 0 \end{pmatrix}, \tag{5-16}$$

where $\boldsymbol{I}_{6-p}$ represents the $(6-p) \times (6-p)$ identity matrix and the 0 on the diagonal represents the $p \times p$ zero matrix.

6. Simplifications for 2-dimensional printed materials

For 2-dimensional printed materials, or any 2-dimensional two-phase composite with one phase being void, the analysis simplifies as then the space of 2×2 symmetric matrices has dimension 3, so there are only four energy functions to consider:

$$\begin{aligned} W_f^0(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0) &= \min_{\boldsymbol{C}_* \in GU_f} \sum_{j=1}^{3} \boldsymbol{\sigma}_j^0 : \boldsymbol{C}_*^{-1} \boldsymbol{\sigma}_j^0, \\ W_f^1(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\epsilon}_1^0) &= \min_{\boldsymbol{C}_* \in GU_f} \left[\boldsymbol{\epsilon}_1^0 : \boldsymbol{C}_* \boldsymbol{\epsilon}_1^0 + \sum_{j=1}^{2} \boldsymbol{\sigma}_j^0 : \boldsymbol{C}_*^{-1} \boldsymbol{\sigma}_j^0 \right], \\ W_f^2(\boldsymbol{\sigma}_1^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0) &= \min_{\boldsymbol{C}_* \in GU_f} \left[\left(\sum_{i=1}^{2} \boldsymbol{\epsilon}_i^0 : \boldsymbol{C}_* \boldsymbol{\epsilon}_i^0 \right) + \boldsymbol{\sigma}_1^0 : \boldsymbol{C}_*^{-1} \boldsymbol{\sigma}_1^0 \right], \\ W_f^3(\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0) &= \min_{\boldsymbol{C}_* \in GU_f} \sum_{i=1}^{3} \boldsymbol{\epsilon}_i^0 : \boldsymbol{C}_* \boldsymbol{\epsilon}_i^0. \end{aligned} \tag{6-1}$$

Again $W_f^0(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0)$ is attained for a "complementary Avellaneda material" consisting of a sequentially layered laminate geometry having an effective tensor $\boldsymbol{C}_* = \widetilde{\boldsymbol{C}}_f^A(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0) \in GU_f$, and we have the inequalities

$$\begin{aligned} \sum_{j=1}^{2} \boldsymbol{\sigma}_j^0 : [\widetilde{\boldsymbol{C}}_f^A(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, 0)]^{-1} \boldsymbol{\sigma}_j^0 &\leq W_f^1(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\epsilon}_1^0), \\ \boldsymbol{\sigma}_1^0 : [\widetilde{\boldsymbol{C}}_f^A(\boldsymbol{\sigma}_1^0, 0, 0)]^{-1} \boldsymbol{\sigma}_1^0 &\leq W_f^2(\boldsymbol{\sigma}_1^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0), \\ 0 &\leq W_f^3(\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0), \end{aligned} \tag{6-2}$$

where, as before, the last inequality is sharp in the limit $\delta \to 0$ being attained when the material consists of islands of phase 1 surrounded by a phase 2.

The recipe for showing that the bound (6-1) on $W_f^1(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\epsilon}_1^0)$ is sharp for certain values of $\boldsymbol{\epsilon}_1^0$ and that the bound (6-1) on $W_f^2(\boldsymbol{\sigma}_1^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0)$ is sharp for certain values of $\boldsymbol{\epsilon}_1^0$ and $\boldsymbol{\epsilon}_2^0$ is almost exactly the same as in the 3-dimensional case: insert into the complementary Avellaneda material a thin walled structure of respectively unimode and bimode materials so that slips can occur along these walls, allowing with very little energetic cost the average strain $\boldsymbol{\epsilon}_1^0$ in the case of W_f^1, or any strain in the space spanned by $\boldsymbol{\epsilon}_1^0$ and $\boldsymbol{\epsilon}_2^0$ in the case of W_f^2.

7. The algebraic problem: characterizing those symmetric matrix pencils spanned by symmetrized rank 1 matrices

We are interested in the following question: *Given k linearly independent symmetric $d \times d$ matrices $\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_k$, find necessary and sufficient conditions such that there exist linearly independent matrices $\{\boldsymbol{B}_i\}_{i=1}^k$ spanned by the basis elements $\boldsymbol{A}_i$ so that each matrix $\boldsymbol{B}_i$ is a symmetrized rank 1 matrix, i.e., there exist vectors $\boldsymbol{a}_i$ and $\boldsymbol{b}_i$, with $|\boldsymbol{b}_i| = 1$, such that*

$$\boldsymbol{B}_i = \tfrac{1}{2}(\boldsymbol{b}_i \boldsymbol{a}_i^T + \boldsymbol{a}_i \boldsymbol{b}_i^T).$$

It is assumed that $d = 2$ or 3 and $1 \leq k \leq k_d$, where $k_2 = 2$ and $k_3 = 5$. Here we are working in the generic situation, i.e., we prove the algebraic result for a dense set of matrices. The continuity result of Section 9 will allow us to conclude for the whole set of matrices. Actually, the proof below also shows that the algebraic result holds for the complement of a zero measure set of matrices.

Theorem 7.1. *The above problem is solvable if and only if the matrices $\boldsymbol{A}_i$ for $i = 1, \ldots, k$ satisfy the following conditions*:

(i) $$\det(\boldsymbol{A}_1) \leq 0, \quad \text{if } k = 1,\ d = 2, \tag{7-1}$$

$\boldsymbol{A}_1$ has two eigenvalues of opposite signs and one zero eigenvalue,
or has two zero eigenvalues, $\quad$ *if* $k = 1,\ d = 3.$ (7-2)

(ii) *If* $k = d = 2$,

$$\det(\boldsymbol{A}_1) < 0$$

or

$$f(t) = \det(\boldsymbol{A}_1 + t\boldsymbol{A}_2) \textit{ is quadratic and has two distinct roots for } t, \\ \textit{or is linear in } t \textit{ with a nonzero coefficient of } t. \quad (7\text{-}3)$$

(iii) *If* $k = 2$ *and* $d = 3$, *defining* $\boldsymbol{A}(\eta, \mu) = \eta \boldsymbol{A}_1 + \mu \boldsymbol{A}_2$, *the numbers*

$$\det(\boldsymbol{A}(\eta, \mu)), \quad \{\boldsymbol{A}(\eta, \mu)\}_{11}\{\boldsymbol{A}(\eta, \mu)\}_{22} - \{\boldsymbol{A}(\eta, \mu)\}_{12}^2, \quad \{\boldsymbol{A}(\eta, \mu)\}_{11} \quad (7\text{-}4)$$

are never simultaneously nonnegative for any choice of η *and* μ *not both zero (equivalently* $\boldsymbol{A}(\eta, \mu)$ *is never strictly positive definite for any values of* η *and* μ*), and*

$$\triangle = 18\det(\boldsymbol{A}_1)\det(\boldsymbol{A}_2)S_1S_2 - 4S_1^3\det(\boldsymbol{A}_2) + S_1^2S_2^2 - 4S_2^3\det(\boldsymbol{A}_1) \\ - 27\det(\boldsymbol{A}_1)^2\det(\boldsymbol{A}_2)^2 > 0, \quad (7\text{-}5)$$

where $S_i = \sum_{j=1}^3 s_{ij}$ *for* $i = 1, 2$ *and* s_{ij} *is the determinant of the matrix obtained by replacing the* j*-th row of* $\boldsymbol{A}_i$ *by the* j*-th row of* $\boldsymbol{A}_{i+1}$, *where by convention we have* $\boldsymbol{A}_3 = \boldsymbol{A}_1$ *(equivalently* $\boldsymbol{A}(\eta, \mu)$ *has three distinct roots).*

(iv) *Always solvable if* $k \geq 3,\ d = 3$. (7-6)

Remark. In fact, the condition (7-2) and the last condition in (7-3), that $f(t)$ is linear in t, could be withdrawn since we are considering the generic case. They are inserted because we can treat them explicitly.

Proof. Case (i): $k = 1$, $d = 2$ or 3. In this case $\boldsymbol{A}_1$ must be a multiple of $\boldsymbol{B}_1$ and hence must be a symmetrized rank 1 matrix. To see more clearly the condition for a matrix $\boldsymbol{B}$ to be a symmetrized rank 1 matrix, i.e., have the form $\boldsymbol{B} = \frac{1}{2}(\boldsymbol{b}\boldsymbol{a}^T + \boldsymbol{a}\boldsymbol{b}^T)$, let us, without loss of generality, choose our coordinates so that $\boldsymbol{b} = [1, 0]^T$ when $d = 2$ and $\boldsymbol{b} = [1, 0, 0]^T$ when $d = 3$. Then $\boldsymbol{B}$ has the representation

$$\boldsymbol{B} = \begin{pmatrix} a_1 & \frac{1}{2}a_2 \\ \frac{1}{2}a_2 & 0 \end{pmatrix} \text{ when } d = 2, \quad \boldsymbol{B} = \begin{pmatrix} a_1 & \frac{1}{2}a_2 & \frac{1}{2}a_3 \\ \frac{1}{2}a_2 & 0 & 0 \\ \frac{1}{2}a_3 & 0 & 0 \end{pmatrix} \text{ when } d = 3. \quad (7\text{-}7)$$

These have eigenvalues

$$\begin{aligned} \lambda &= \frac{1}{2}\Big(a_1 \pm \sqrt{a_1^2 + a_2^2}\Big) \quad \text{when } d = 2, \\ \lambda &= \frac{1}{2}\Big(a_1 \pm \sqrt{a_1^2 + a_2^2 + a_3^2}\Big) \text{ and } \lambda = 0 \quad \text{when } d = 3. \end{aligned} \quad (7\text{-}8)$$

So, clearly $\boldsymbol{B}$ is a symmetrized rank 1 matrix in two dimensions if and only if $\det(\boldsymbol{B}) \leq 0$, and is a symmetrized rank 1 matrix in three dimensions if and only if it has two eigenvalues of opposite signs and one zero eigenvalue, or has two zero eigenvalues.

Case (ii): $k=2$, $d=2$. In this case there should be two distinct values of t such that $\det(\boldsymbol{A}_1+t\boldsymbol{A}_2)<0$, which by continuity of this determinant as a function of t is guaranteed if any of the conditions in (7-3) are met. Note that the case where $\det(\boldsymbol{A}_1+t\boldsymbol{A}_2)=0$ for all t can be ruled out from consideration since this can only happen when $\boldsymbol{A}_2$ is proportional to $\boldsymbol{A}_1$, as can be easily seen by working in a basis where $\boldsymbol{A}_2$ is diagonal.

Case (iii): $k=2$, $d=3$. Consider the matrix pencil (over reals η and μ) $\boldsymbol{A}(\eta,\mu)=\eta\boldsymbol{A}_1+\mu\boldsymbol{A}_2$. Assuming that $\det\boldsymbol{A}(\eta,\mu)$ is not zero for all η and μ, there are at least two matrices on the pencil which have nonzero determinant. Let us relabel them as $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$. Then the equation $\det(\boldsymbol{A}(1,\mu))=0$ must have either two or three roots $\mu=z_i$ for $i=1,2$ or $i=1,2,3$, where the z_i are obtained by changing the sign of the generalized eigenvalues. This gives Cardan's condition:

$$\triangle = 18\det(\boldsymbol{A}_1)\det(\boldsymbol{A}_2)S_1S_2-4S_1^3\det(\boldsymbol{A}_2)+S_1^2S_2^2-4S_2^3\det(\boldsymbol{A}_1) \\ -27\det(\boldsymbol{A}_1)^2\det(\boldsymbol{A}_2)^2\geq 0. \quad \text{(7-9)}$$

Suppose that $\boldsymbol{A}_1+\mu\boldsymbol{A}_2$ contains a symmetric matrix with two zero eigenvalues (a rank 1 matrix) as μ is varied. Then by redefining $\boldsymbol{A}_2$ we can assume $\boldsymbol{A}_2$ is this matrix, now with zero determinant, and by using a basis where $\boldsymbol{A}_2$ is diagonal, we see that $\det(\boldsymbol{A}_1+\mu\boldsymbol{A}_2)$ depends linearly on μ and $\det(\boldsymbol{A}_1+\mu\boldsymbol{A}_2)$ can only have one root: (7-9) must be violated. So we can exclude this possibility: $\boldsymbol{A}_1+\mu\boldsymbol{A}_2$ has at most one zero eigenvalue for any value of μ. Now consider the eigenvalues of $\boldsymbol{A}(\theta)\equiv\boldsymbol{A}(\cos\theta,\sin\theta)$ as θ is varied. As $\boldsymbol{A}(-\theta)=-\boldsymbol{A}(\theta)$ it suffices to consider the interval of θ between 0 and π. Some scenarios for the eigenvalue trajectories are plotted in Figure 5. At the values $\theta_i=\arctan^{-1}(z_i)$ at least one of the eigenvalues must be zero, and the favorable situation is when there are two remaining eigenvalues of opposite signs or only one nonzero eigenvalue. Such angles θ_i are marked by the vertical dashed lines in the figure. The unfavorable situation is when there are two nonzero eigenvalues of the same sign, marked by the red vertical lines in Figure 5 (left). First suppose that $\boldsymbol{A}(\theta)$ is positive definite for some $\theta=\theta_0$. By refining θ as the old θ minus θ_0, let us suppose $\boldsymbol{A}(0)$ is positive definite. Then the scenario is that in Figure 5 (left), or some variant of it in which eigenvalues cross, which is unfavorable. The only way to avoid this is for $\boldsymbol{A}(\theta)$ to have two zero eigenvalues at the smallest and largest values of $\theta\in[0,\pi]$ for which $\det\boldsymbol{A}(\theta)=0$, as in Figure 5 (middle), but we have ruled out the possibility that $\boldsymbol{A}(\theta)$ has two zero eigenvalues for any value of θ. We are left with Figure 5 (right) as being the only possible suitable scenario. In conclusion, we require that the matrix $\boldsymbol{A}(\theta)$ not be positive semidefinite for any choice of θ; i.e., the three quantities

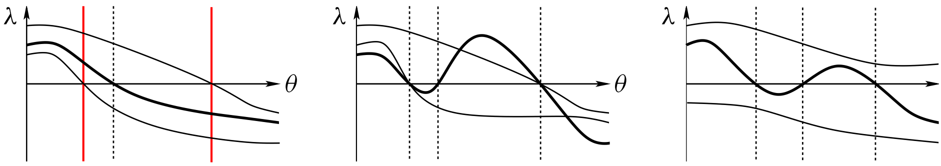

Figure 5. Some scenarios for the eigenvalues λ of $A(\theta) = \cos\theta A_1 + \sin\theta A_2$ as θ is varied.

$$\det(A(\eta,\mu)), \quad \{A(\eta,\mu)\}_{11}\{A(\eta,\mu)\}_{22} - \{A(\eta,\mu)\}_{12}^2, \quad \{A(\eta,\mu)\}_{11} \tag{7-10}$$

are never simultaneously nonnegative for any choice of η and μ not both zero. This condition could be made explicit by using the formula for the roots of a cubic to determine the generalized eigenvalues $-z_i$.

Case (iv): $k \ge 3$, $d = 3$. The case $k = 3$ is a straightforward consequence of Lemma 7.2 below.

It remains to consider $k \ge 4$ and $d = 3$. By the previous step, in the space spanned by A_1, A_2, and A_3 there are three matrices A_1', A_2', and $B_3 = A_3 + \eta_3 A_1' + \mu_3 A_2'$ that are linearly independent, symmetrized and of rank 1. Then, again by the previous step, we can find linearly independent matrices $B_1, \dots, B_k$ that have the form $B_1 = A_1'$, $B_2 = A_2'$, and $B_i = A_i + \eta_i A_1' + \mu_i A_2'$ for $3 \le i \le k$ and that are of rank 1. $\square$

In the sequel we write

$$a \otimes b := ab^T \quad \text{and} \quad a \odot b := \tfrac{1}{2}(a \otimes b + b \otimes a) \qquad \text{for } a, b \in \mathbb{R}^3. \tag{7-11}$$

Lemma 7.2. *Let A, B, C be three symmetric matrices of $\mathbb{R}^{3\times 3}$.*

(i) *Up to small perturbations of A, B, C, there exist a basis (x, y, z) of $\mathbb{R}^3$ and three vectors a, b, c of $\mathbb{R}^3$ satisfying*

$$\begin{cases} a \in \{Ax, Bx, Cx\}^\perp \setminus \{0\}, \\ b \in \{Ay, By, Cy\}^\perp \setminus \{0\}, \\ c \in \{Az, Bz, Cz\}^\perp \setminus \{0\}, \end{cases} \tag{7-12}$$

or equivalently,

$$a \odot x,\ b \odot y,\ c \odot z \in \{A, B, C\}^\perp \setminus \{0\}. \tag{7-13}$$

(ii) *Up to small perturbations of A, B, C, there exist three independent symmetrized rank 1 matrices in the space $\{A, B, C\}^\perp$.*

Proof. (i) Let F be the cubic function defined by

$$F(x) := \det(Ax, Bx, Cx) \quad \text{for } x \in \mathbb{R}^3. \tag{7-14}$$

If $F \equiv 0$ in $\mathbb{R}^3$, then condition (7-12) is immediately satisfied. Otherwise, there exists a basis $(\boldsymbol{x}_0, \boldsymbol{u}_0, \boldsymbol{v}_0)$ of $\mathbb{R}^3$ in the nonempty open set $\{F \neq 0\}$. Since we have

$$F(\boldsymbol{x}_0 + s\boldsymbol{u}_0) \underset{|s|\to\infty}{\sim} s^3 F(\boldsymbol{u}_0) \underset{|s|\to\infty}{\longrightarrow} \pm\infty, \tag{7-15}$$

there exists $s, t \in \mathbb{R} \setminus \{0\}$ such that $\boldsymbol{x} := \boldsymbol{x}_0 + s\boldsymbol{u}_0$ and $\boldsymbol{y} := \boldsymbol{x}_0 + t\boldsymbol{v}_0$ are two independent vectors in the set $\{F = 0\}$.

First, assume that the set $\{F = 0\}$ is not contained in the plane $\mathrm{Span}\{\boldsymbol{x}, \boldsymbol{y}\}$. Then there exists a basis $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ of $\mathbb{R}^3$ in the set $\{F = 0\}$. Therefore, there exist three vectors $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ of $\mathbb{R}^3$ satisfying (7-12), or equivalently (7-13).

Now, assume that $\{F = 0\} \subset \mathrm{Span}\{\boldsymbol{x}, \boldsymbol{y}\}$. First of all, up to small perturbations we can assume that the matrices $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ are invertible. Since $\boldsymbol{B}^{-1}\boldsymbol{C}$ is a 3×3 real matrix, it has at least a real eigenvalue λ. The perturbation procedure is now divided into two cases.

First case: The matrix $\boldsymbol{B}^{-1}\boldsymbol{C}$ has two complex conjugate eigenvalues.

Then the eigenspace $\mathrm{Ker}(\boldsymbol{B}^{-1}\boldsymbol{C} - \lambda \boldsymbol{I}_3)$ is a line of $\mathbb{R}^3$ spanned by $\boldsymbol{e} \in \mathbb{R}^3 \setminus \{\boldsymbol{0}\}$. Consider a basis $(\boldsymbol{x}_0, \boldsymbol{u}_0, \boldsymbol{v}_0)$ of $\mathbb{R}^3$ in the set $\{F \neq 0\}$ such that $(\boldsymbol{e}, \boldsymbol{x}_0, \boldsymbol{u}_0)$ and $(\boldsymbol{e}, \boldsymbol{x}_0, \boldsymbol{v}_0)$ are also two bases of $\mathbb{R}^3$. As previously there exist $s, t \in \mathbb{R} \setminus \{0\}$ such that $\boldsymbol{x} := \boldsymbol{x}_0 + s\boldsymbol{u}_0$ and $\boldsymbol{y} := \boldsymbol{x}_0 + t\boldsymbol{v}_0$ are two independent vectors of the set $\{F = 0\}$. Moreover, since $(\boldsymbol{e}, \boldsymbol{x})$ and $(\boldsymbol{e}, \boldsymbol{y})$ are two families of independent vectors and $\mathbb{R}\boldsymbol{e}$ is the unique real eigenspace of the matrix $\boldsymbol{B}^{-1}\boldsymbol{C}$, we have

$$\boldsymbol{B}\boldsymbol{x} \times \boldsymbol{C}\boldsymbol{x} \neq \boldsymbol{0} \quad \text{and} \quad \boldsymbol{B}\boldsymbol{y} \times \boldsymbol{C}\boldsymbol{y} \neq \boldsymbol{0}. \tag{7-16}$$

Now, consider a vector $\boldsymbol{u} \in \{\boldsymbol{x}, \boldsymbol{y}\}^{\perp} \setminus \{\boldsymbol{0}\}$ and the matrix $\boldsymbol{M} \in \mathbb{R}^{3\times 3}$ defined by

$$\boldsymbol{M}\boldsymbol{x} = \boldsymbol{\xi}, \quad \boldsymbol{M}\boldsymbol{y} = \boldsymbol{\eta}, \quad \boldsymbol{M}\boldsymbol{u} = \boldsymbol{0}, \tag{7-17}$$

where the vectors $\boldsymbol{\xi}, \boldsymbol{\eta}$ will be chosen later. Define for $\tau > 0$ the perturbed function

$$F_\tau(z) := \det(\boldsymbol{A}z + \tau \boldsymbol{M}z, \boldsymbol{B}z, \boldsymbol{C}z) \quad \text{for } z \in \mathbb{R}^3. \tag{7-18}$$

We have

$$\begin{cases} F_\tau(\boldsymbol{x} + \tau\boldsymbol{u}) = \tau\boldsymbol{\xi} \cdot \big(\boldsymbol{B}\boldsymbol{x} \times \boldsymbol{C}\boldsymbol{x} + \boldsymbol{O}(\tau)\big) + O(\tau), \\ F_\tau(\boldsymbol{y} + \tau\boldsymbol{u}) = \tau\boldsymbol{\eta} \cdot \big(\boldsymbol{B}\boldsymbol{y} \times \boldsymbol{C}\boldsymbol{y} + \boldsymbol{O}(\tau)\big) + O(\tau), \end{cases} \tag{7-19}$$

where the $\boldsymbol{O}(\tau)$ denote some first-order vectors in τ and $O(\tau)$ some first-order real numbers in τ which are independent of $\boldsymbol{\xi}, \boldsymbol{\eta}$. Condition (7-16) then allows us to choose $\boldsymbol{\xi} = \boldsymbol{\xi}_\tau$ and $\boldsymbol{\eta} = \boldsymbol{\eta}_\tau$ such that $F_\tau(\boldsymbol{x} + \tau\boldsymbol{u}) = F_\tau(\boldsymbol{y} + \tau\boldsymbol{u}) = 0$. Therefore, since $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u})$ is a basis of $\mathbb{R}^3$, $(\boldsymbol{x}, \boldsymbol{x} + \tau\boldsymbol{u}, \boldsymbol{y} + \tau\boldsymbol{u})$ is also a basis of $\mathbb{R}^3$, which in addition lies in the set $\{F_\tau = 0\}$. This leads us to condition (7-12) with the matrices $\boldsymbol{A} + \tau\boldsymbol{M}, \boldsymbol{B}, \boldsymbol{C}$.

Second case: The matrix $\boldsymbol{B}^{-1}\boldsymbol{C}$ has only real eigenvalues.

Then there exists a small perturbation $\boldsymbol{C}_\tau$ of $\boldsymbol{C}$ such that the perturbed matrix $\boldsymbol{B}^{-1}\boldsymbol{C}_\tau$ has three distinct real eigenvalues. Hence, the matrix $\boldsymbol{B}^{-1}\boldsymbol{C}_\tau$ admits a basis $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ of eigenvectors, which implies that

$$\boldsymbol{C}_\tau \boldsymbol{x} - \lambda \boldsymbol{B}\boldsymbol{x} = \boldsymbol{C}_\tau \boldsymbol{y} - \lambda \boldsymbol{B}\boldsymbol{y} = \boldsymbol{C}_\tau \boldsymbol{z} - \lambda \boldsymbol{B}\boldsymbol{z} = \boldsymbol{0}. \tag{7-20}$$

Therefore, the perturbed function

$$F_\tau(\boldsymbol{u}) := \det(\boldsymbol{A}\boldsymbol{u}, \boldsymbol{B}\boldsymbol{u}, \boldsymbol{C}_\tau \boldsymbol{u}) \quad \text{for } \boldsymbol{u} \in \mathbb{R}^3 \tag{7-21}$$

satisfies $F_\tau(\boldsymbol{x}) = F_\tau(\boldsymbol{y}) = F_\tau(\boldsymbol{z}) = 0$, which again leads us to condition (7-12) with the matrices $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}_\tau$.

(ii) We will distinguish four cases according to whether the following conditions are satisfied by the basis $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ of $\mathbb{R}^3$ and the vectors $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathbb{R}^3 \setminus \{0\}$ obtained in step (i):

$$\begin{cases} \boldsymbol{a} \in \operatorname{Span}\{\boldsymbol{x}, \boldsymbol{y}\} \cap \operatorname{Span}\{\boldsymbol{x}, \boldsymbol{z}\}, \\ \boldsymbol{b} \in \operatorname{Span}\{\boldsymbol{y}, \boldsymbol{x}\} \cap \operatorname{Span}\{\boldsymbol{y}, \boldsymbol{z}\}, \\ \boldsymbol{c} \in \operatorname{Span}\{\boldsymbol{z}, \boldsymbol{x}\} \cap \operatorname{Span}\{\boldsymbol{z}, \boldsymbol{y}\}. \end{cases} \tag{7-22}$$

First case: $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ satisfy conditions (7-22).

Then, since $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ is a basis of $\mathbb{R}^3$, we have necessarily $\boldsymbol{a} \in \mathbb{R}\boldsymbol{x}$, $\boldsymbol{b} \in \mathbb{R}\boldsymbol{y}$, $\boldsymbol{c} \in \mathbb{R}\boldsymbol{z}$. Therefore, $\boldsymbol{x} \odot \boldsymbol{x}$, $\boldsymbol{y} \odot \boldsymbol{y}$, $\boldsymbol{z} \odot \boldsymbol{z}$ are clearly three independent matrices of $\{\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}\}^\perp$.

Second case: $\boldsymbol{b}$ and $\boldsymbol{c}$ satisfy conditions (7-22) but $\boldsymbol{a}$ does not.

Then, for example, $(\boldsymbol{a}, \boldsymbol{x}, \boldsymbol{y})$ is a basis of $\mathbb{R}^3$, and $\boldsymbol{b} \in \mathbb{R}\boldsymbol{y}$, $\boldsymbol{c} \in \mathbb{R}\boldsymbol{z}$. Let $\boldsymbol{u} \in \{\boldsymbol{y}, \boldsymbol{z}\}^\perp \setminus \{\boldsymbol{0}\}$, and let $\alpha, \beta, \gamma \in \mathbb{R}$ be such that $\alpha\, \boldsymbol{a} \odot \boldsymbol{x} + \beta\, \boldsymbol{y} \odot \boldsymbol{y} + \gamma\, \boldsymbol{z} \odot \boldsymbol{z} = \boldsymbol{0}$. Multiplying by $\boldsymbol{u}$ we get that $\alpha(\boldsymbol{x} \cdot \boldsymbol{u})\boldsymbol{a} + \alpha(\boldsymbol{a} \cdot \boldsymbol{u})\boldsymbol{x} = \boldsymbol{0}$; hence $\alpha = 0$ since $\boldsymbol{x} \cdot \boldsymbol{u} \neq 0$. We deduce immediately that $\beta = \gamma = 0$. Therefore, $\boldsymbol{a} \odot \boldsymbol{x}$, $\boldsymbol{y} \odot \boldsymbol{y}$, $\boldsymbol{z} \odot \boldsymbol{z}$ are three independent matrices of $\{\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}\}^\perp$.

Third case: $\boldsymbol{a}$ and $\boldsymbol{b}$ do not satisfy conditions (7-22), with $\boldsymbol{a} \notin \operatorname{Span}\{\boldsymbol{x}, \boldsymbol{y}\}$ and $\boldsymbol{b} \notin \mathbb{R}\boldsymbol{a} \cup \mathbb{R}\boldsymbol{x}$ (respectively $\boldsymbol{a} \notin \operatorname{Span}\{\boldsymbol{x}, \boldsymbol{z}\}$ and $\boldsymbol{c} \notin \mathbb{R}\boldsymbol{a} \cup \mathbb{R}\boldsymbol{x}$).

Then $(\boldsymbol{a}, \boldsymbol{x}, \boldsymbol{y})$ is a basis of $\mathbb{R}^3$. Let $\boldsymbol{u} \in \{\boldsymbol{x}, \boldsymbol{y}\}^\perp \setminus \{\boldsymbol{0}\}$, and let $\alpha, \beta \in \mathbb{R}$ be such that $\alpha\, \boldsymbol{a} \odot \boldsymbol{x} + \beta\, \boldsymbol{b} \odot \boldsymbol{y} = \boldsymbol{0}$. Multiplying by $\boldsymbol{u}$ we get that $\alpha(\boldsymbol{a} \cdot \boldsymbol{u})\boldsymbol{x} + \beta(\boldsymbol{b} \cdot \boldsymbol{u})\boldsymbol{y} = \boldsymbol{0}$; hence $\alpha = 0$ since $\boldsymbol{a} \cdot \boldsymbol{u} \neq 0$, and thus $\beta = 0$. Therefore, $\boldsymbol{a} \odot \boldsymbol{x}$, $\boldsymbol{b} \odot \boldsymbol{y}$ are two independent matrices of $\{\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}\}^\perp$, which have two eigenvalues of opposite sign and one 0 eigenvalue.

Let us prove by contradiction that

$$\exists\, t \in \mathbb{R} \setminus \{0\}, \quad \det(\boldsymbol{a} \odot \boldsymbol{x} + t\, \boldsymbol{b} \odot \boldsymbol{y}) \neq 0. \tag{7-23}$$

Otherwise, for any $t \neq 0$, there exists $z_t \in \operatorname{Ker}(\boldsymbol{a} \odot \boldsymbol{x} + t\,\boldsymbol{b} \odot \boldsymbol{y}) \setminus \{\boldsymbol{0}\}$; hence

$$(\boldsymbol{x} \cdot z_t)\boldsymbol{a} + (\boldsymbol{a} \cdot z_t)\boldsymbol{x} + t(\boldsymbol{y} \cdot z_t)\boldsymbol{b} + t(\boldsymbol{b} \cdot z_t)\boldsymbol{y} = \boldsymbol{0}. \tag{7-24}$$

Since $(\boldsymbol{a}, \boldsymbol{x}, \boldsymbol{y})$ is a basis of $\mathbb{R}^3$ and $z_t \neq \boldsymbol{0}$, we have necessarily $\boldsymbol{y} \cdot z_t \neq 0$, which implies that

$$-\boldsymbol{b} = \frac{\boldsymbol{x} \cdot z_t}{t(\boldsymbol{y} \cdot z_t)}\boldsymbol{a} + \frac{\boldsymbol{a} \cdot z_t}{t(\boldsymbol{y} \cdot z_t)}\boldsymbol{x} + \frac{\boldsymbol{b} \cdot z_t}{\boldsymbol{y} \cdot z_t}\boldsymbol{y} = \alpha\,\boldsymbol{a} + \beta\,\boldsymbol{x} + \gamma\,\boldsymbol{y}, \tag{7-25}$$

where α, β, γ are independent of t, and

$$(\boldsymbol{x} - \alpha t\,\boldsymbol{y}) \cdot z_t = (\boldsymbol{a} - \beta t\,\boldsymbol{y}) \cdot z_t = (\boldsymbol{b} - \gamma\,\boldsymbol{y}) \cdot z_t = 0. \tag{7-26}$$

Since $z_t \neq \boldsymbol{0}$ there exists $(p_t, q_t, r_t) \in \mathbb{R}^3 \setminus \{0\}$ such that

$$\begin{aligned} &p_t(\boldsymbol{x} - \alpha t\,\boldsymbol{y}) + q_t(\boldsymbol{a} - \beta t\,\boldsymbol{y}) + r_t(\boldsymbol{b} - \gamma\,\boldsymbol{y}) \\ &\qquad = (q_t - \alpha r_t)\boldsymbol{a} + (p_t - \beta r_t)\boldsymbol{x} - (\alpha t p_t + \beta t q_t + 2\gamma r_t)\boldsymbol{y} = \boldsymbol{0}, \end{aligned} \tag{7-27}$$

which implies that $q_t = \alpha r_t$, $p_t = \beta r_t$ and $r_t(\alpha\beta t + \gamma) = 0$. Since $(p_t, q_t, r_t) \neq \boldsymbol{0}$, we have $r_t \neq 0$ and $\alpha\beta t + \gamma = 0$ for any $t \neq 0$; hence $\alpha\beta = 0$ and $\gamma = 0$. This yields a contradiction between (7-25) and $\boldsymbol{b} \notin \mathbb{R}\boldsymbol{a} \cup \mathbb{R}\boldsymbol{x}$.

By virtue of (7-23) there exist two nonzero real numbers $\alpha \neq \beta$ such that the matrices

$$\boldsymbol{M} := \boldsymbol{a} \odot \boldsymbol{x} + \alpha\,\boldsymbol{b} \odot \boldsymbol{y} \quad \text{and} \quad \boldsymbol{N} := \boldsymbol{a} \odot \boldsymbol{x} + \beta\,\boldsymbol{b} \odot \boldsymbol{y} \tag{7-28}$$

are invertible. The function $p(t) := \det(\beta\boldsymbol{M} - t\boldsymbol{N})$ is a polynomial of degree 3 whose α, β are two distinct roots. Then the polynomial $p(t)$ must change sign by crossing α, for example (the conclusion is similar for β). Let $\lambda_1(t) \le \lambda_2(t) \le \lambda_3(t)$ be the well-ordered eigenvalues of the symmetric matrix $\beta\boldsymbol{M} - t\boldsymbol{N}$. Since the vectors $\boldsymbol{a}, \boldsymbol{x}$ are independent, $\boldsymbol{a} \odot \boldsymbol{x}$ has two eigenvalues of opposite sign and one 0 eigenvalue; hence $\lambda_1(\alpha) < \lambda_2(\alpha) = 0 < \lambda_3(\alpha)$.

Now, let $\boldsymbol{P}_\tau$ for a small $\tau > 0$ be a symmetric matrix in the space $\{\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}\}^\perp$, such that $|\boldsymbol{P}_\tau - \boldsymbol{a} \odot \boldsymbol{x}| = O(\tau)$, and such that the three matrices $\boldsymbol{a} \odot \boldsymbol{x}, \boldsymbol{b} \odot \boldsymbol{y}, \boldsymbol{P}_\tau$ are independent (note that the dimension of $\{\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}\}^\perp$ is ≥ 3). Define the two perturbed matrices

$$\boldsymbol{M}_\tau := \boldsymbol{P}_\tau + \alpha\,\boldsymbol{b} \odot \boldsymbol{y} \quad \text{and} \quad \boldsymbol{N}_\tau := \boldsymbol{P}_\tau + \beta\,\boldsymbol{b} \odot \boldsymbol{y}. \tag{7-29}$$

Since the well-ordered eigenvalues of a real symmetric matrix $\boldsymbol{S}$ are Lipschitz-continuous with respect to $\boldsymbol{S}$ (see, e.g., [Ciarlet 1989], Theorem 2.3-2), the eigenvalues $\lambda_1^\tau(t) \le \lambda_2^\tau(t) \le \lambda_3^\tau(t)$ of $\beta\boldsymbol{M}_\tau - t\boldsymbol{N}_\tau$ converge uniformly as $\tau \to 0$ to the eigenvalues $\lambda_1(t) \le \lambda_2(t) \le \lambda_3(t)$ of $\beta\boldsymbol{M} - t\boldsymbol{N}$, with respect to t in a neighborhood of α. Hence, for $\tau > 0$ small enough, there exist α_τ close to α such that $\alpha_\tau \neq \beta$ and

$\lambda_1^\tau(\alpha_\tau) < \lambda_2^\tau(\alpha_\tau) = 0 < \lambda_3^\tau(\alpha_\tau)$. Then by (7-29) there exist $\boldsymbol{c}_\tau, \boldsymbol{z}_\tau \in \mathbb{R}^3$ such that

$$\beta \boldsymbol{M}_\tau - \alpha_\tau \boldsymbol{N}_\tau = \boldsymbol{c}_\tau \odot \boldsymbol{z}_\tau = (\beta - \alpha_\tau)\boldsymbol{P}_\tau + \beta(\alpha - \alpha_\tau)\boldsymbol{b} \odot \boldsymbol{y}, \quad \text{with } \beta - \alpha_\tau \neq 0. \tag{7-30}$$

Therefore, $\boldsymbol{a} \odot \boldsymbol{x}$, $\boldsymbol{b} \odot \boldsymbol{y}$, $\boldsymbol{c}_\tau \odot \boldsymbol{z}_\tau$ are three independent symmetrized rank 1 matrices in the space $\{\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}\}^\perp$.

Fourth case: $\boldsymbol{a}$ and $\boldsymbol{b}$ do not satisfy conditions (7-22), with $\boldsymbol{a} \notin \mathrm{Span}\{\boldsymbol{x}, \boldsymbol{y}\}$ and $\boldsymbol{b} \in \mathbb{R}\boldsymbol{a} \cup \mathbb{R}\boldsymbol{x}$ (respectively $\boldsymbol{a} \notin \mathrm{Span}\{\boldsymbol{x}, \boldsymbol{z}\}$ and $\boldsymbol{c} \in \mathbb{R}\boldsymbol{a} \cup \mathbb{R}\boldsymbol{x}$).

For example, we have $\boldsymbol{b} \in \mathbb{R}\boldsymbol{a}$. We thus start from the matrices $\boldsymbol{a} \odot \boldsymbol{x}$ and $\boldsymbol{a} \odot \boldsymbol{y}$ in the space $\{\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}\}^\perp$, where $(\boldsymbol{a}, \boldsymbol{x}, \boldsymbol{y})$ is a basis of $\mathbb{R}^3$. We will consider a perturbation of $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ for leading us to the third case.

Let $\boldsymbol{t} \in \{\boldsymbol{a}, \boldsymbol{x}\}^\perp \setminus \{\boldsymbol{0}\}$, let $\boldsymbol{d} \in \mathbb{R}^3 \setminus (\mathbb{R}\boldsymbol{a} + \mathbb{R}\boldsymbol{x})$, and consider, for a small $\tau > 0$, the perturbed vector $\boldsymbol{b}_\tau := \boldsymbol{a} + \tau \boldsymbol{d} \notin \mathbb{R}\boldsymbol{a} \cup \mathbb{R}\boldsymbol{x}$ and the perturbed matrices

$$\boldsymbol{A}_\tau := \boldsymbol{A} + \tau \boldsymbol{t} \odot \boldsymbol{u}_\tau, \quad \boldsymbol{B}_\tau := \boldsymbol{B} + \tau \boldsymbol{t} \odot \boldsymbol{v}_\tau, \quad \boldsymbol{C}_\tau := \boldsymbol{C} + \tau \boldsymbol{t} \odot \boldsymbol{w}_\tau, \tag{7-31}$$

where the vectors $\boldsymbol{u}_\tau$, $\boldsymbol{v}_\tau$, $\boldsymbol{w}_\tau$ will be chosen later. Clearly, $\boldsymbol{a} \odot \boldsymbol{x} \in \{\boldsymbol{A}_\tau, \boldsymbol{B}_\tau, \boldsymbol{C}_\tau\}^\perp$. On the other hand, we have

$$\boldsymbol{A}_\tau : \boldsymbol{b}_\tau \odot \boldsymbol{y} = \tau(\boldsymbol{A} : \boldsymbol{d} \odot \boldsymbol{y} + \boldsymbol{t} \odot \boldsymbol{u}_\tau : \boldsymbol{a} \odot \boldsymbol{y} + \tau \boldsymbol{t} \odot \boldsymbol{u}_\tau : \boldsymbol{d} \odot \boldsymbol{y}). \tag{7-32}$$

Since $2\boldsymbol{t} \odot \boldsymbol{u}_\tau : \boldsymbol{a} \odot \boldsymbol{y} = (\boldsymbol{t} \cdot \boldsymbol{y})\boldsymbol{a} \cdot \boldsymbol{u}_\tau$ with $\boldsymbol{t} \cdot \boldsymbol{y} \neq 0$, we can choose $\boldsymbol{u}_\tau = \boldsymbol{O}(1)$ with respect to τ such that $\boldsymbol{A}_\tau : \boldsymbol{b}_\tau \odot \boldsymbol{y} = 0$. Hence, choosing $\boldsymbol{v}_\tau$ and $\boldsymbol{w}_\tau$ similarly, we get that $\boldsymbol{b}_\tau \odot \boldsymbol{y} \in \{\boldsymbol{A}_\tau, \boldsymbol{B}_\tau, \boldsymbol{C}_\tau\}^\perp$. Therefore, the vectors $\boldsymbol{a}$, $\boldsymbol{b}_\tau$, $\boldsymbol{x}$, $\boldsymbol{y}$ satisfy the conditions of the third case with the perturbed matrices $\boldsymbol{A}_\tau$, $\boldsymbol{B}_\tau$, $\boldsymbol{C}_\tau$. □

8. Constructing suitable multimode materials for the wall microstructure

Let us specify the construction of the desired multimode materials in two dimensions and then move to three dimensions. We begin by constructing bimode materials that can only support one stress. One could use the fourth-rank laminate structure described in detail in Section 30.7 of [Milton 2002]. The analysis would then be essentially a repeat of that analysis, which builds the appropriate trial stress and strain fields at each length scale. The key feature is that these trial fields need to be chosen so the trial stress associated with the average stress $\boldsymbol{\sigma}^0$ we want to achieve at the macroscopic scale is concentrated entirely in phase 1 (apart from boundary layers that we ignore, whose contribution to the energy vanishes in the homogenization limit), and so the trial strain associated with an average strain that is orthogonal to $\boldsymbol{\sigma}^0$ is concentrated entirely in phase 2.

Rather than doing this, it is more instructive to build trial stress and strain fields that are concentrated in phase 1 and phase 2, respectively, for the honeycomb and inverted honeycomb bimode structures of Figure 6, as the ideas here carry over to pentamode materials. The trial stress is easy. It is taken to be macroscopically

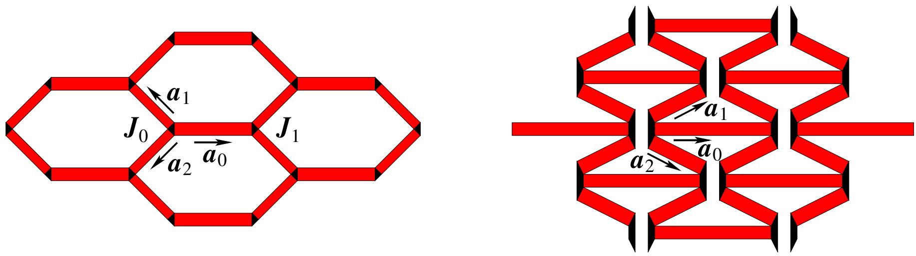

Figure 6. 2-dimensional bimode materials that can only support one average stress field $\boldsymbol{\sigma}^0$, and which are easily compliant to any strain orthogonal to $\boldsymbol{\sigma}^0$. Here the red struts are laminates of the two phases with the interfaces in the laminate parallel to the direction of the struts. The geometry on the left is appropriate if $\det \boldsymbol{\sigma}^0 > 0$, the geometry on the right is appropriate if $\det \boldsymbol{\sigma}^0 < 0$, and if $\det \boldsymbol{\sigma}^0 = 0$ it suffices to use a simple laminate with the layer surfaces perpendicular to the null vector of $\boldsymbol{\sigma}^0$.

constant with a value $\alpha_i \boldsymbol{a}_i \boldsymbol{a}_i^T$ in each strut which is parallel to the unit vector $\boldsymbol{a}_i$ in Figure 7. Let w_i denote the width of the strut parallel to $\boldsymbol{a}_i$, for $i = 0, 1, 2$. Since the net "force" on the black junction regions in the top left and top right of Figure 7 must be zero, we obtain

$$0 = -\sum_{k=0}^{2} w_i(\alpha_i \boldsymbol{a}_i \boldsymbol{a}_i^T)\boldsymbol{a}_i = -\sum_{k=0}^{2} w_i \alpha_i \boldsymbol{a}_i. \tag{8-1}$$

Since $w_1 = w_2$ and $\boldsymbol{a}_0$ points in the horizontal direction, while $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ have the same horizontal component and equal but opposite vertical components, we get

$$\alpha_1 = \alpha_2 = -w_0\alpha_0/[2w_1(\boldsymbol{a}_1 \cdot \boldsymbol{a}_2)]. \tag{8-2}$$

The symmetry of the trial stress field implies there is no associated torque acting on the junction regions. The trial stress in the junction regions is really not that important. One choice is the stress field that satisfies the elasticity equations appropriate to phase 1 filling the junction region when constant tractions act on the three sides. The average value of the trial stress does not depend on the choice of trial stress in the junctions. Indeed, since $\nabla \cdot (\boldsymbol{\sigma}) = 0$ it follows from integration by parts of $\nabla \cdot (\boldsymbol{\sigma x})$ (where $\boldsymbol{\sigma x}$ is a third-order tensor) that

$$\int_{\Omega} \boldsymbol{\sigma}\, d\boldsymbol{x} = \int_{\partial\Omega} \boldsymbol{t}\boldsymbol{x}^T\, dS, \quad \text{where } \boldsymbol{t} = \boldsymbol{\sigma}\boldsymbol{n} \text{ is the surface traction,} \tag{8-3}$$

in which Ω is any region with boundary $\partial\Omega$. For example, the boundary of Ω could be the outermost boundary of the shape in the top left or top right of Figure 7, where we include the dashed lines as part of the boundary.

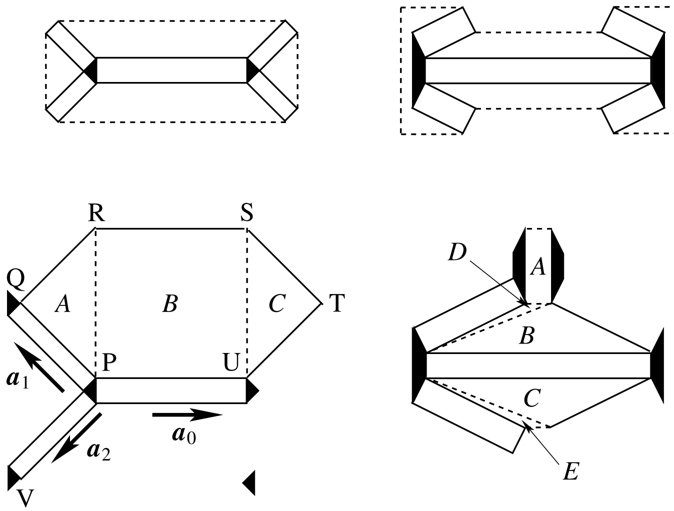

Figure 7. The honeycomb structure of Figure 6 (left) can be taken to have the unit cell shown at top left. Similarly the inverted honeycomb structure of Figure 6 (right) can be taken to have the unit cell shown at top right. The space outside the struts and junction regions (which is occupied by phase 2) has been triangulated with boundaries marked by the dashed lines to make the construction of the trial stress fields easy.

In passing, we remark that if $\boldsymbol{\sigma}^0$ is proportional to the identity matrix, then the microstructure of Figure 7 (top left) resembles a Sigmund microstructure (see the last subfigure in Figure 2 in [Sigmund 2000]). However, we do not require the tuning of layer widths in the struts that makes his structure optimal. Suboptimal structures are perfectly fine in the walls, since the walls ultimately occupy a vanishingly small volume fraction in the final material.

To obtain a trial easy strain it suffices to specify the trial displacement in the unit cell. We only choose motions so the junction regions (triangular in Figure 7 (bottom left) and quadrilateral in Figure 7 (bottom right)) undergo rigid body translations, so there is no strain inside them. Thus associated with Figure 7 (bottom left) one can clearly identify two independent macroscopic modes of motion. The first is where the line RS moves vertically upwards while the line PU remains fixed, and Q and T move in such a way that the lengths QR, QP, TS, and TU remain equal and preserved in length. One can choose the displacement to be linear in each of the three regions A, B, and C so that it matches the displacement on the boundary. The second is where the line RS moves horizontally while the line PU remains fixed, and Q and T move in such a way that the lengths QR, QP, TS, and TU remain equal and preserved in length. In either case inside the horizontal laminate arm there is no

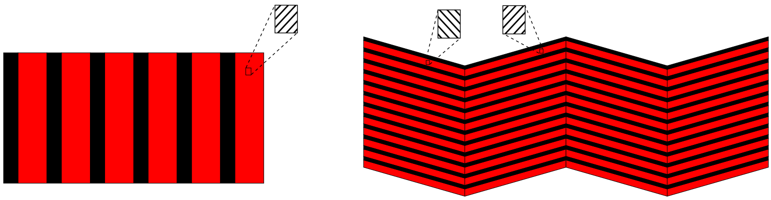

Figure 8. 2-dimensional unimode materials that are easily compliant to one average strain field $\boldsymbol{\epsilon}^0$, and which can support any stress orthogonal to $\boldsymbol{\epsilon}^0$. In both, the red region represents a laminate as indicated by the inserts. The second-rank laminate geometry on the left is appropriate if $\det \boldsymbol{\epsilon}^0 < 0$ and the third-rank geometry on the right is appropriate for any $\boldsymbol{\epsilon}^0$.

strain, while inside the inclined laminate arms there is an infinitesimal shear so the junction at P remains fixed, while the junction at Q moves perpendicular to $\boldsymbol{a}_1$ and the junction at V moves perpendicular to $\boldsymbol{a}_2$. We also note that there is also an easy microscopic motion which results in no macroscopic motion. Define the center of each triangular junction to be the point which is at the junction of the perpendicular bisector of the three faces. Then if all the triangular junctions undergo the same infinitesimal rotation about these centers while the laminate material in the struts shears at the same time, it will cost very little energy. The trial strain field is bounded and nonzero only in phase 2, and therefore the associated upper bound on the elastic energy scales in proportion to δ.

The situation in Figure 7 (bottom right) is basically similar. The two black quadrilateral junction regions at the bottom of the figure can remain fixed. Then one mode is the symmetric one, where the region A undergoes uniaxial compression in the horizontal direction and at the same time moves downwards. The second is where the region A undergoes pure shear, so the junction on the left side of it moves up, while the right side moves down. The strain field can be taken constant in the regions A, B, C, D, and E, and in the inclined laminate strut arms is also constant and corresponds to pure shear. These strains are easily determined from the value of the trial displacement field at the boundaries of each region. Again, the trial strain field is bounded and nonzero only in phase 2, and therefore the associated upper bound on the elastic energy scales in proportion to δ.

The structures of Figure 8 give suitable 2-dimensional unimode materials. We will not specify the appropriate trial stress and strain fields which prove that these structures have the desired elastic behavior, as they are exactly the same as those given in Section 30.6 of [Milton 2002].

We now describe the pentamodes and the trial fields in them. Given four vectors $\boldsymbol{a}_0$, $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, and $\boldsymbol{a}_4$ (no longer required to be unit vectors) we position a point P at

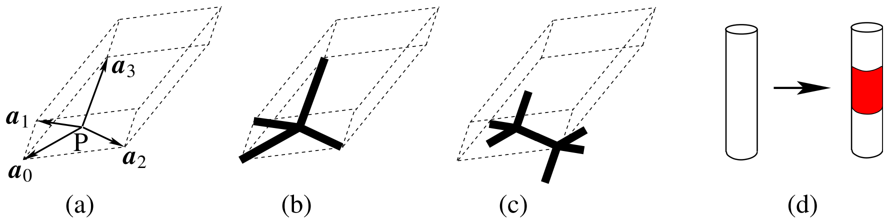

Figure 9. The procedure for constructing the desired pentamodes. In (d) a shearable section is inserted into each strut. This section, shown in red, has the structure of parallel square fibers, as illustrated in Figure 10, with the fibers aligned parallel to the strut.

the origin, and join P to the four points $\boldsymbol{x} = \boldsymbol{a}_i$ for $i = 0, 1, 2, 3$, with four infinitesimally thin rods, as in Figure 9(a). We then take as our unit cell of periodicity the parallelepiped with the eight points $\boldsymbol{x} = \boldsymbol{a}_i$, $\boldsymbol{x} = \boldsymbol{a}_1 + \boldsymbol{a}_2 + \boldsymbol{a}_3 - \boldsymbol{a}_i - \boldsymbol{a}_0$ for $i = 0, 1, 2, 3$ (the three vectors $\boldsymbol{v}_i = \boldsymbol{a}_i - \boldsymbol{a}_0$ for $i = 1, 2, 3$ are the primitive lattice vectors). We require that $\boldsymbol{a}_0, \boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_4$ be chosen so P lies within this parallelepiped. After periodically extending the rod structure (with rods joining $k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_1 + k_3\boldsymbol{v}_1$ with the four points $k_1\boldsymbol{v}_1 + k_2\boldsymbol{v}_1 + k_3\boldsymbol{v}_1 + a_i$ for $i = 0, 1, 2, 3$, for any integers k_1, k_2, and k_3), we then coat this periodic rod structure with phase 1, as illustrated in Figure 9(b), so that any point $\boldsymbol{x}$ is in phase 1 if and only if it is within a distance r of the rod structure. Here r should be chosen appropriately small so that the coatings of each rod contain a cylindrical section that we refer to as a strut. Figure 9(b) is misleading as it suggests that the unit cell only contains one junction region. The true structure which should be periodically repeated (by making copies shifted by vectors $k_1\boldsymbol{v}_1+k_2\boldsymbol{v}_1+k_3\boldsymbol{v}_1$ for all combinations of integers k_1, k_2, and k_3) is shown in Figure 9(c) and contains the junction of Figure 9(b) plus the one obtained by inverting it under the transformation $\boldsymbol{x} \to -\boldsymbol{x}$. The final step, illustrated in Figure 9(d), is to take a cylindrical subsection of each cylindrical section between junctions and replace it with a pentamode material that supports any stress proportional to $\boldsymbol{a}_i\boldsymbol{a}_i^T$. It is convenient to take end faces of the cylindrical subsection to be perpendicular to the cylinder axis, i.e., perpendicular to the vector $\boldsymbol{a}_i$ that is parallel to the cylinder axis. Now we define the junction regions to be those connected regions of phase 1 that are bounded by the cylindrical subsections.

To obtain the trial stress field, we first solve for the tensions in the rods of Figure 9(a) when the rods are completely rigid and supporting a stress. These are found just by balance of forces at the junctions. If the rods parallel to $\boldsymbol{a}_i$ have a tension T_i (which could be negative) then we take in the cylindrical subsection of the corresponding strut of the final pentamode a trial stress field $T_i\boldsymbol{a}_i\boldsymbol{a}_i^T/(|\boldsymbol{a}_i|^2\pi r^2)$ giving rise to a net force T_i pulling (pushing if T_i is negative) on the adjacent junction regions. Inside the junction region we take a stress field that satisfies the

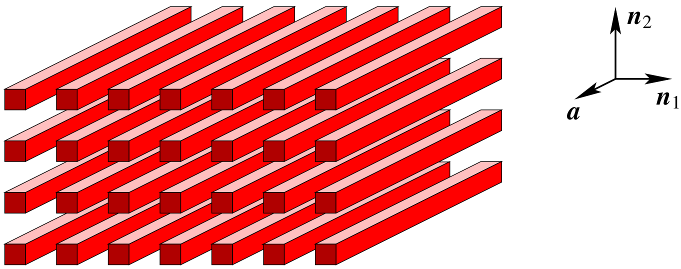

Figure 10. A detailed view of the square beam array microstructure which is used as the easily shearable section in the pentamode cylindrical struts. The vector $\boldsymbol{a}$ is chosen to be one of the four vectors $\boldsymbol{a}_k$ for $k = 0, 1, 2, 3$, as appropriate to each pentamode strut orientated parallel to $\boldsymbol{a}_k$. The square beams can support tension (or compression) in the direction of the beam, and in particular can support a constant macroscopic stress $\boldsymbol{\sigma}_k^B = \alpha_k \boldsymbol{a}_k \boldsymbol{a}_k^T$. As we are working in the framework of linear elasticity, we ignore the very real possibility that the beams will buckle.

elasticity equations appropriate to phase 1 filling the junction region when constant tractions $T_i/(\pi r^2)$ act on the four disks that border the cylindrical subsections, and there are no forces on the remaining surface of the junction regions.

Obtaining appropriate trial strain fields is also not too difficult. We first consider an infinitesimal motion that the rod model with Figure 9(a) as the unit cell can undergo when the rods are rigid but the pin junctions are flexible. Then in the final pentamode the junction regions are taken to undergo a rigid body translation which is the same as that of the corresponding pin junction in the rod model. The cylindrical subsections undergo appropriate shears to ensure continuity of the displacement. The trial displacement in the remaining multiconnected region of phase 2 bordered by the junction regions and the cylindrical subsection can be somewhat arbitrary, and is not really important. One could take it as the solution for the displacement field when phase 2 has some nonzero elastic moduli, and the displacement at the boundary of the junctions and cylindrical subsections matches that of the trial field just specified. The trial strain field is bounded and nonzero only in phase 2, and therefore the associated upper bound on the elastic energy scales in proportion to δ.

It is clear from the choice of these trial stress and strain fields that the macroscopic stress the material supports and the easy motions it permits are exactly the same as those for the ideal model with rods and pin junctions that has the unit cell pictured in Figure 9(a), and which provided the basis for our construction. That this structure can support any desired average stress, and only that average stress, is then a direct consequence of the analysis in Section 5.2 of [Milton and Cherkaev 1995].

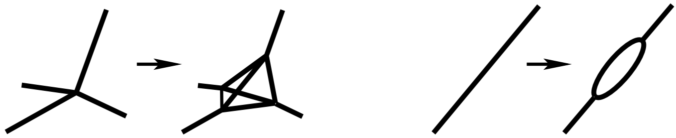

Figure 11. Some of the replacements that are needed to obtain desired unimode, bimode, trimode, or quadramode materials.

To obtain any desired unimode, bimode, trimode, or quadramode material, having respectively $p = 1, 2, 3, 4$ independent easy modes of deformation, and supporting respectively $6-p$ applied stresses $\boldsymbol{\sigma}_j^0$ for $j = 1, \ldots, 6-p$, we follow the prescription given by Milton and Cherkaev [1995]. That is, we superimpose, one at a time, $6-p$ pentamode structures, each supporting one of the stresses $\boldsymbol{\sigma}_j^0$, with struts which are sufficiently thin to ensure that one can (with appropriate modification specified below) superimpose the structures without collision. When doing this superimposition we first remove phase 2 and shift the lattice structures to try to avoid unwanted intersections of phase 1. This may not always be possible, so in the event two vertices clash we make the replacement in Figure 11 (left) in one of the structures (which may of course then cause additional unwanted intersections of the struts). Then if two (or more) struts intersect we make the replacement in Figure 11 (right) in all but one of the struts (which then passes through each hole). The remaining possibility we want to avoid is that two pentamode struts are parallel and intersect when we superimpose the structures. Due to the freedom in the choice of the $\boldsymbol{a}_k$ that give a desired $\boldsymbol{\sigma}_j^0$, we can always choose our $6-p$ pentamode structures to avoid such clashes. Finally, the shearable section in each pentamode strut should be placed in a section that has not been modified, so it still is parallel to one of the $\boldsymbol{a}_k$. At the very end any remaining space that is not filled by phase 1 should be filled by the extremely compliant phase 2.

9. Continuity of the energy functions

It follows from the preceding analysis that we can determine the three energy functions

$$W_f^3(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0),$$
$$W_f^4(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0),$$
$$W_f^5(\boldsymbol{\sigma}_1^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0, \boldsymbol{\epsilon}_5^0)$$

in the limit $\delta \to 0$ for almost all combinations of applied fields. Here we establish that these energy functions are continuous functions of the applied fields in the limit $\delta \to 0$, and therefore we obtain expressions for the energy functions for all combinations of applied fields in this limit.

Recall that the set GU_f is characterized by its W-transform. For example, part of it is described by the function

$$W_f^4(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0) = \min_{\boldsymbol{C}_* \in GU_f} \left[\sum_{i=1}^4 \boldsymbol{\epsilon}_i^0 : \boldsymbol{C}_* \boldsymbol{\epsilon}_i^0 + \sum_{j=1}^2 \boldsymbol{\sigma}_j^0 : \boldsymbol{C}_*^{-1} \boldsymbol{\sigma}_j^0 \right]. \quad (9\text{-}1)$$

Here we want to show that such energy functions are continuous in their arguments. Let the tensor $\boldsymbol{C}_*(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0)$ be a minimizer of (9-1), and suppose we perturb the applied stress fields $\boldsymbol{\sigma}_j^0$ by $\delta\boldsymbol{\sigma}_j^0$ and the applied strain fields $\boldsymbol{\epsilon}_i^0$ by $\delta\boldsymbol{\epsilon}_i^0$. Now consider the walled material with a geometry described by the characteristic function

$$\chi_w(\boldsymbol{x}) = \prod_{k=1}^3 (1 - H_{\epsilon'}(\boldsymbol{x} \cdot \boldsymbol{n}_k)), \quad (9\text{-}2)$$

where $\boldsymbol{n}_1$, $\boldsymbol{n}_2$, and $\boldsymbol{n}_3$ are the three orthogonal unit vectors

$$\boldsymbol{n}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{n}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{n}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad (9\text{-}3)$$

and ϵ' is a small parameter that gives the thickness of the walls. Inside the walls, where $\chi_w(\boldsymbol{x}) = 0$, we put an isotropic composite of phase 1 and phase 2, mixed in the proportions f and $1 - f$ with isotropic effective elasticity tensor $\boldsymbol{C}(\kappa_0, \mu_0)$, where κ_0 is the effective bulk modulus and μ_0 is the effective shear modulus, which are assumed to have nonzero limits as $\delta \to 0$. (The isotropic composite could consist of islands of void surrounded by phase 1.) Outside the walls, where $\chi_w(\boldsymbol{x}) = 1$, we put the material that has an effective tensor

$$\boldsymbol{C}_*^1 = \boldsymbol{C}_*(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0).$$

Let $\boldsymbol{C}'_*$ be the effective tensor of the composite. We have the variational principle

$$\begin{aligned} &\sum_{i=1}^4 (\boldsymbol{\epsilon}_i^0 + \delta\boldsymbol{\epsilon}_i^0) : \boldsymbol{C}'_*(\boldsymbol{\epsilon}_i^0 + \delta\boldsymbol{\epsilon}_i^0) + \sum_{j=1}^2 (\boldsymbol{\sigma}_j^0 + \delta\boldsymbol{\sigma}_j^0) : (\boldsymbol{C}'_*)^{-1} (\boldsymbol{\sigma}_j^0 + \delta\boldsymbol{\sigma}_j^0) \\ &\quad = \min_{\underline{\boldsymbol{\epsilon}}_1, \underline{\boldsymbol{\epsilon}}_2, \underline{\boldsymbol{\epsilon}}_3, \underline{\boldsymbol{\epsilon}}_4, \underline{\boldsymbol{\sigma}}_1, \underline{\boldsymbol{\sigma}}_2} \Big\langle \sum_{i=1}^4 \underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x}) : [\chi_w(\boldsymbol{x}) \boldsymbol{C}_*^1 + (1 - \chi_w(\boldsymbol{x})) \boldsymbol{C}(\kappa_0, \mu_0)] \underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x}) \\ &\qquad\qquad + \sum_{j=1}^2 \underline{\boldsymbol{\sigma}}_j(\boldsymbol{x}) : [\chi_w(\boldsymbol{x}) \boldsymbol{C}_*^1 + (1 - \chi_w(\boldsymbol{x})) \boldsymbol{C}(\kappa_0, \mu_0)]^{-1} \underline{\boldsymbol{\sigma}}_j(\boldsymbol{x}) \Big\rangle, \end{aligned} \quad (9\text{-}4)$$

where the minimum is over fields subject to the appropriate average values and differential constraints. We choose constant trial strain fields

$$\underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x}) = \boldsymbol{\epsilon}_i^0 + \delta\boldsymbol{\epsilon}_i^0, \quad i = 1, 2, 3, 4, \quad (9\text{-}5)$$

and trial stress fields

$$\underline{\boldsymbol{\sigma}}_j(\boldsymbol{x}) = \boldsymbol{\sigma}_j^0 + \delta\underline{\boldsymbol{\sigma}}_j(\boldsymbol{x}), \quad j = 1, 2, \tag{9-6}$$

where $\delta\underline{\boldsymbol{\sigma}}_j(\boldsymbol{x})$ has average value $\delta\boldsymbol{\sigma}_j^0$ and is concentrated in the walls. Specifically, if $\{\delta\boldsymbol{\sigma}_j^0\}_{k\ell}$ denote the matrix elements of $\delta\boldsymbol{\sigma}_j^0$, and letting

$$\delta\boldsymbol{\sigma}_j^1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & \{\delta\boldsymbol{\sigma}_j^0\}_{23} \\ 0 & \{\delta\boldsymbol{\sigma}_j^0\}_{32} & \{\delta\boldsymbol{\sigma}_j^0\}_{33} \end{pmatrix},$$

$$\delta\boldsymbol{\sigma}_j^2 = \begin{pmatrix} \{\delta\boldsymbol{\sigma}_j^0\}_{11} & 0 & \{\delta\boldsymbol{\sigma}_j^0\}_{13} \\ 0 & 0 & 0 \\ \{\delta\boldsymbol{\sigma}_j^0\}_{31} & 0 & 0 \end{pmatrix}, \tag{9-7}$$

$$\delta\boldsymbol{\sigma}_j^3 = \begin{pmatrix} 0 & \{\delta\boldsymbol{\sigma}_j^0\}_{12} & 0 \\ \{\delta\boldsymbol{\sigma}_j^0\}_{21} & \{\delta\boldsymbol{\sigma}_j^0\}_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

then we choose

$$\delta\underline{\boldsymbol{\sigma}}_j(\boldsymbol{x}) = \sum_{k=1}^{3} \delta\boldsymbol{\sigma}_j^k H_{\epsilon'}(\boldsymbol{x}\cdot\boldsymbol{n}_k)/\epsilon', \tag{9-8}$$

which has the required average value $\delta\boldsymbol{\sigma}_j^0$ and satisfies the differential constraints appropriate to a stress field because $\delta\boldsymbol{\sigma}_j^k\boldsymbol{n}_k = 0$.

Hence, there exist positive constants α and β such that for sufficiently small ϵ' and for sufficiently small variations $\delta\boldsymbol{\sigma}_j^0$ and $\delta\boldsymbol{\epsilon}_i^0$ in the applied fields, we have

$$\begin{aligned}\Big\langle \sum_{i=1}^{4} \underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x}) : [\chi_w(\boldsymbol{x})\boldsymbol{C}_*^1 + (1-\chi_w(\boldsymbol{x}))\boldsymbol{C}(\kappa_0,\mu_0)]\underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x}) \\ + \sum_{j=1}^{2} \underline{\boldsymbol{\sigma}}_j(\boldsymbol{x}) : [\chi_w(\boldsymbol{x})\boldsymbol{C}_*^1 + (1-\chi_w(\boldsymbol{x}))\boldsymbol{C}(\kappa_0,\mu_0)]^{-1}\underline{\boldsymbol{\sigma}}_j(\boldsymbol{x})\Big\rangle \\ \le W_f^4(\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0,\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0,\boldsymbol{\epsilon}_3^0,\boldsymbol{\epsilon}_4^0) + \alpha\epsilon' + \beta K/\epsilon',\end{aligned} \tag{9-9}$$

where K represents the norm

$$K = \sqrt{\sum_{i=1}^{4} \delta\boldsymbol{\epsilon}_i^0 : \delta\boldsymbol{\epsilon}_i^0 + \sum_{j=1}^{2} \delta\boldsymbol{\sigma}_j^0 : \delta\boldsymbol{\sigma}_j^0}, \tag{9-10}$$

of the field variations. Choosing $\epsilon' = \sqrt{\beta K/\alpha}$ to minimize the right-hand side of (9-9), we obtain

$$W_f^4(\boldsymbol{\sigma}_1^0+\delta\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0+\delta\boldsymbol{\sigma}_2^0, \boldsymbol{\epsilon}_1^0+\delta\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0+\delta\boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0+\delta\boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0+\delta\boldsymbol{\epsilon}_4^0) \\ \leq W_f^4(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0) + 2\sqrt{\alpha\beta K}. \quad (9\text{-}11)$$

In obtaining the bound (9-9) we have used the fact that K^2 is less than K for sufficiently small K, specifically $K < 1$. Clearly the right-hand side of (9-11) approaches $W_f^4(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0)$ as $K \to 0$. On the other hand, by repeating the same argument with the roles of

$$W_f^4(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0)$$

and

$$W_f^4(\boldsymbol{\sigma}_1^0+\delta\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0+\delta\boldsymbol{\sigma}_2^0, \boldsymbol{\epsilon}_1^0+\delta\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0+\delta\boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0+\delta\boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0+\delta\boldsymbol{\epsilon}_4^0)$$

reversed, and with

$$\boldsymbol{C}_*(\boldsymbol{\sigma}_1^0+\delta\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0+\delta\boldsymbol{\sigma}_2^0, \boldsymbol{\epsilon}_1^0+\delta\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0+\delta\boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0+\delta\boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0+\delta\boldsymbol{\epsilon}_4^0)$$

replacing

$$\boldsymbol{C}_*(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0),$$

we deduce that

$$W_f^4(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0) \\ \leq W_f^4(\boldsymbol{\sigma}_1^0+\delta\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0+\delta\boldsymbol{\sigma}_2^0, \boldsymbol{\epsilon}_1^0+\delta\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0+\delta\boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0+\delta\boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0+\delta\boldsymbol{\epsilon}_4^0) \\ + 2\sqrt{\alpha\beta K}. \quad (9\text{-}12)$$

This, together with (9-11), establishes the continuity of $W_f^4(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0)$. The continuity of the other energy functions follows by the same argument.

10. Conclusion

We have established the following two theorems.

Theorem 10.1. *Consider composites in three dimensions of two materials with positive definite elasticity tensors $\boldsymbol{C}_1$ and $\boldsymbol{C}_2 = \delta\boldsymbol{C}_0$ mixed in proportions f and $1-f$. Let the seven energy functions W_f^k, for $k = 0, 1, \ldots, 6$, that characterize the set GU_f (with $U = (\boldsymbol{C}_1, \delta\boldsymbol{C}_0)$) of possible elastic tensors be defined by (3-9). These energy functions involve a set of applied strains $\boldsymbol{\epsilon}_i^0$ and applied stresses $\boldsymbol{\sigma}_j^0$ meeting the orthogonality condition (3-10). The energy function W_f^0 is given by*

$$W_f^0(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0, \boldsymbol{\sigma}_5^0, \boldsymbol{\sigma}_6^0) = \sum_{j=1}^{6} \boldsymbol{\sigma}_j^0 : \widetilde{\boldsymbol{C}}_f^A(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0, \boldsymbol{\sigma}_5^0, \boldsymbol{\sigma}_6^0)\boldsymbol{\sigma}_j^0 \quad (10\text{-}1)$$

(as proved by Avellaneda [1987b]). *Here* $\widetilde{\boldsymbol{C}}_f^A(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0, \boldsymbol{\sigma}_5^0, \boldsymbol{\sigma}_6^0)$ *is the effective elasticity tensor of a complementary Avellaneda material that is a sequentially layered laminate with the minimum value of the sum of complementary energies*

$$\sum_{j=1}^{6} \boldsymbol{\sigma}_j^0 : \boldsymbol{C}_*^{-1} \boldsymbol{\sigma}_j^0. \tag{10-2}$$

Additionally, we now have

$$\begin{aligned}
&\lim_{\delta\to 0} W_f^3(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0) = \sum_{j=1}^{3} \boldsymbol{\sigma}_j^0 : [\widetilde{\boldsymbol{C}}_f^A(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, 0, 0, 0)]^{-1} \boldsymbol{\sigma}_j^0, \\
&\lim_{\delta\to 0} W_f^4(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0) = \sum_{j=1}^{2} \boldsymbol{\sigma}_j^0 : [\widetilde{\boldsymbol{C}}_f^A(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, 0, 0, 0, 0)]^{-1} \boldsymbol{\sigma}_j^0, \\
&\lim_{\delta\to 0} W_f^5(\boldsymbol{\sigma}_1^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0, \boldsymbol{\epsilon}_5^0) = \boldsymbol{\sigma}_1^0 : [\widetilde{\boldsymbol{C}}_f^A(\boldsymbol{\sigma}_1^0, 0, 0, 0, 0, 0)]^{-1} \boldsymbol{\sigma}_1^0, \\
&\lim_{\delta\to 0} W_f^6(\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0, \boldsymbol{\epsilon}_5^0, \boldsymbol{\epsilon}_6^0) = 0.
\end{aligned} \tag{10-3}$$

for all combinations of applied stresses $\boldsymbol{\sigma}_j^0$ *and applied strains* $\boldsymbol{\epsilon}_i^0$. *When* $\det \boldsymbol{\epsilon}_1^0 = 0$ *but* $\boldsymbol{\epsilon}_1^0$ *is not positive semidefinite or negative semidefinite, we have*

$$\lim_{\delta\to 0} W_f^1(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0, \boldsymbol{\sigma}_5^0, \boldsymbol{\epsilon}_1^0) = \sum_{j=1}^{5} \boldsymbol{\sigma}_j^0 : [\widetilde{\boldsymbol{C}}_f^A(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0, \boldsymbol{\sigma}_5^0, 0)]^{-1} \boldsymbol{\sigma}_j^0, \tag{10-4}$$

while when the equation $\det(\boldsymbol{\epsilon}_1^0 + t\boldsymbol{\epsilon}_2^0)$ *has at least two distinct roots for* t (*the condition for which is given by* (7-5)), *and additionally, the matrix pencil* $\boldsymbol{\epsilon}(t) = \boldsymbol{\epsilon}_1^0 + t\boldsymbol{\epsilon}_2^0$ *does not contain any positive definite or negative definite matrices as* t *is varied* (*which requires that the quantities in* (7-4) *are never all positive, or all negative*), *we have*

$$\lim_{\delta\to 0} W_f^2(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0) = \sum_{j=1}^{4} \boldsymbol{\sigma}_j^0 : [\widetilde{\boldsymbol{C}}_f^A(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0, 0, 0)]^{-1} \boldsymbol{\sigma}_j^0. \tag{10-5}$$

Theorem 10.2. *For* 2*-dimensional composites, the four energy functions* W_f^k, *for* $k = 0, 1, 2, 3$, *are defined by* (6-1), *and these characterize the set* GU_f, *with* $U = (\boldsymbol{C}_1, \delta\boldsymbol{C}_0)$, *of possible elastic tensors* $\boldsymbol{C}_*$ *of composites of two phases with positive definite elasticity tensors* $\boldsymbol{C}_1$ *and* $\boldsymbol{C}_2 = \delta\boldsymbol{C}_0$. *These energy functions involve a set of applied strains* $\boldsymbol{\epsilon}_i^0$ *and applied stresses* $\boldsymbol{\sigma}_j^0$ *meeting the orthogonality condition* (3-10). *The energy function* W_f^0 *is given by*

$$W_f^0(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0) = \sum_{j=1}^{3} \boldsymbol{\sigma}_j^0 : \widetilde{\boldsymbol{C}}_f^A(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0) \boldsymbol{\sigma}_j^0 \tag{10-6}$$

(as proved by Avellaneda [1987b]), *where* $\widetilde{\boldsymbol{C}}^A_f(\boldsymbol{\sigma}^0_1, \boldsymbol{\sigma}^0_2, \boldsymbol{\sigma}^0_3)$ *is the effective elasticity tensor of a complementary Avellaneda material that is a sequentially layered laminate with the minimum value of the sum of complementary energies*

$$\sum_{j=1}^{3} \boldsymbol{\sigma}^0_j : \boldsymbol{C}_*^{-1} \boldsymbol{\sigma}^0_j. \tag{10-7}$$

We also have the trivial result that

$$\lim_{\delta \to 0} W^3_f(\boldsymbol{\epsilon}^0_1, \boldsymbol{\epsilon}^0_2, \boldsymbol{\epsilon}^0_3) = 0. \tag{10-8}$$

When $\det \boldsymbol{\epsilon}^0_1 \le 0$ *we have*

$$\lim_{\delta \to 0} W^1_f(\boldsymbol{\sigma}^0_1, \boldsymbol{\sigma}^0_2, \boldsymbol{\epsilon}^0_1) = \sum_{j=1}^{2} \boldsymbol{\sigma}^0_j : [\widetilde{\boldsymbol{C}}^A_f(\boldsymbol{\sigma}^0_1, \boldsymbol{\sigma}^0_2, 0)]^{-1} \boldsymbol{\sigma}^0_j, \tag{10-9}$$

while when $\det \boldsymbol{\epsilon}^0_1 < 0$ *or when* $f(t) = \det(\boldsymbol{\epsilon}^0_1 + t\boldsymbol{\epsilon}^0_2)$ *is quadratic in* t *with two distinct roots*, *or when* $f(t)$ *is linear in* t *with a nonzero* t *coefficient*, *we have*

$$\lim_{\delta \to 0} W^2_f(\boldsymbol{\sigma}^0_1, \boldsymbol{\epsilon}^0_1, \boldsymbol{\epsilon}^0_2) = \boldsymbol{\sigma}^0_1 : [\widetilde{\boldsymbol{C}}^A_f(\boldsymbol{\sigma}^0_1, 0, 0)]^{-1} \boldsymbol{\sigma}^0_1. \tag{10-10}$$

These theorems, and the accompanying microstructures, help define what sort of elastic behaviors are theoretically possible in 2- and 3-dimensional printed materials. They should serve as benchmarks for the construction of more realistic microstructures that can be manufactured. We have found the minimum over all microstructures of various sums of energies and complementary energies. More realistic designs can be obtained by adding to this sum a term that penalizes the surface area as done for a single energy minimization by Kohn and Wirth [2014; 2016].

It remains an open problem to find expressions for the energy functions in the cases not covered by these theorems. Even for an isotropic composite with a bulk modulus κ_* and a shear modulus μ_*, the set of all possible pairs (κ_*, μ_*) is still not completely characterized either in the limit $\delta \to 0$ or in the limit $\delta \to \infty$. In these limits the bounds of Berryman and Milton [1988] and Cherkaev and Gibiansky [1993] decouple and provide no extra information beyond that provided by the Hashin–Shtrikman–Hill bounds [Hashin and Shtrikman 1963; Hashin 1965; Hill 1963; 1964]. While the results of this paper show that in the limit $\delta \to 0$ one can obtain 2- or 3-dimensional structures attaining the Hashin–Shtrikman–Hill upper bound on κ_*, while having $\mu_* = 0$, it is not clear what the maximum value for μ_* is, given that $\kappa_* = 0$.

One important corollary of this work is that it gives a complete characterization of the possible triplets $(\boldsymbol{\epsilon}^0, \boldsymbol{\sigma}^0, f)$ of average strain $\boldsymbol{\epsilon}^0$, average stress $\boldsymbol{\sigma}^0$, and volume fraction f that can occur in 2-dimensional and 3-dimensional printed materials in the limit $\delta \to 0$. This will be discussed in a separate paper [Milton and Camar-Eddine 2016].

Acknowledgements

The authors thank the National Science Foundation for support through grant DMS-1211359. M. Briane wishes to thank the Department of Mathematics of the University of Utah for his stay during March 25–April 3, 2016. Mohamed Camar-Eddine is thanked for initial collaborations that marked the beginning of this work. The authors are grateful to Martin Wegener and his group for allowing them to use one of their electron micrographs of the 3-dimensional pentamode structure, and to Muamer Kadic for help with processing the micrograph.

References

[Allaire 1994] G. Allaire, "Explicit lamination parameters for three-dimensional shape optimization", *Control Cybernet.* **23**:3 (1994), 309–326.

[Allaire 2002] G. Allaire, *Shape optimization by the homogenization method*, Applied Mathematical Sciences **146**, Springer, New York, 2002.

[Allaire and Aubry 1999] G. Allaire and S. Aubry, "On optimal microstructures for a plane shape optimization problem", *Struct. Optimiz.* **17**:2 (1999), 86–94.

[Allaire and Kohn 1993a] G. Allaire and R. V. Kohn, "Optimal bounds on the effective behavior of a mixture of two well-ordered elastic materials", *Quart. Appl. Math.* **51**:4 (1993), 643–674.

[Allaire and Kohn 1993b] G. Allaire and R. V. Kohn, "Explicit optimal bounds on the elastic energy of a two-phase composite in two space dimensions", *Quart. Appl. Math.* **51**:4 (1993), 675–699.

[Allaire and Kohn 1994] G. Allaire and R. V. Kohn, "Optimal lower bounds on the elastic energy of a composite made from two non-well-ordered isotropic materials", *Quart. Appl. Math.* **52**:2 (1994), 311–333.

[Avellaneda 1987a] M. Avellaneda, "Iterated homogenization, differential effective medium theory and applications", *Comm. Pure Appl. Math.* **40**:5 (1987), 527–554.

[Avellaneda 1987b] M. Avellaneda, "Optimal bounds and microgeometries for elastic two-phase composites", *SIAM J. Appl. Math.* **47**:6 (1987), 1216–1228.

[Avellaneda and Milton 1989] M. Avellaneda and G. W. Milton, "Bounds on the effective elasticity tensor of composites based on two-point correlations", pp. 89–93 in *Composite material technology, 1989: presented at the Twelfth Annual Energy-Sources Technology Conference and Exhibition* (Houston, 1989), American Society of Mechanical Engineers, Petroleum Division **24**, American Society of Mechanical Engineers, New York, 1989.

[Backus 1962] G. E. Backus, "Long-wave elastic anisotropy produced by horizontal layering", *J. Geophys. Res.* **67**:11 (1962), 4427–4440.

[Beran and Molyneux 1966] M. Beran and J. E. Molyneux, "Use of classical variational principles to determine bounds for the effective bulk modulus in heterogeneous media", *Q. Appl. Math.* **24** (1966), 107–118.

[Berryman and Milton 1988] J. G. Berryman and G. W. Milton, "Microgeometry of random composites and porous media", *J. Phys. D* **21**:1 (1988), 87–94.

[Bourdin and Kohn 2008] B. Bourdin and R. V. Kohn, "Optimization of structural topology in the high-porosity regime", *J. Mech. Phys. Solids* **56**:3 (2008), 1043–1064.

[Bückmann et al. 2014] T. Bückmann, M. Thiel, M. Kadic, R. Schittny, and M. Wegener, "An elasto-mechanical unfeelability cloak made of pentamode metamaterials", *Nature Commun.* **5** (2014), art. id. 4130, 6 pp.

[Camar-Eddine and Seppecher 2003] M. Camar-Eddine and P. Seppecher, "Determination of the closure of the set of elasticity functionals", *Arch. Ration. Mech. Anal.* **170**:3 (2003), 211–245.

[Cherkaev 2000] A. Cherkaev, *Variational methods for structural optimization*, Applied Mathematical Sciences **140**, Springer, New York, 2000.

[Cherkaev and Gibiansky 1992] A. V. Cherkaev and L. V. Gibiansky, "The exact coupled bounds for effective tensors of electrical and magnetic properties of two-component two-dimensional composites", *Proc. Roy. Soc. Edinburgh Sect. A* **122**:1-2 (1992), 93–125.

[Cherkaev and Gibiansky 1993] A. V. Cherkaev and L. V. Gibiansky, "Coupled estimates for the bulk and shear moduli of a two-dimensional isotropic elastic composite", *J. Mech. Phys. Solids* **41**:5 (1993), 937–980.

[Ciarlet 1989] P. G. Ciarlet, *Introduction to numerical linear algebra and optimisation*, Cambridge Univ. Press, 1989.

[Francfort and Milton 1994] G. A. Francfort and G. W. Milton, "Sets of conductivity and elasticity tensors stable under lamination", *Comm. Pure Appl. Math.* **47**:3 (1994), 257–279.

[Francfort and Murat 1986] G. A. Francfort and F. Murat, "Homogenization and optimal bounds in linear elasticity", *Arch. Rational Mech. Anal.* **94**:4 (1986), 307–334.

[Francfort et al. 1995] G. Francfort, F. Murat, and L. Tartar, "Fourth-order moments of nonnegative measures on S^2 and applications", *Arch. Rational Mech. Anal.* **131**:4 (1995), 305–333.

[Gérard 1989] P. Gérard, "Compacité par compensation et régularité 2-microlocale", exposé VI, 1–18 in *Séminaire sur les Équations aux Dérivées Partielles, 1988–1989*, École Polytech., Palaiseau, FR, 1989.

[Gérard 1994] P. Gérard, "Microlocal analysis of compactness", pp. 75–86 in *Nonlinear partial differential equations and their applications* (Paris, 1991–1993), edited by H. Brezis and J.-L. Lions, Pitman Res. Notes Math. Ser. **302**, Longman, Essex, 1994.

[Gibiansky and Cherkaev 1997a] L. V. Gibiansky and A. V. Cherkaev, "Design of composite plates of extremal rigidity", pp. 95–137 in *Topics in the mathematical modelling of composite materials*, edited by A. Cherkaev and R. Kohn, Progr. Nonlinear Differential Equations Appl. **31**, Birkhäuser, Boston, 1997.

[Gibiansky and Cherkaev 1997b] L. V. Gibiansky and A. V. Cherkaev, "Microstructures of composites of extremal rigidity and exact bounds on the associated energy density", pp. 273–317 in *Topics in the mathematical modelling of composite materials*, edited by A. Cherkaev and R. Kohn, Progr. Nonlinear Differential Equations Appl. **31**, Birkhäuser, Boston, 1997.

[Gibiansky and Lakes 1993] L. V. Gibiansky and R. Lakes, "Bounds on the complex bulk modulus of a two-phase viscoelastic composite with arbitrary volume fractions of the components", *Mech. Mater.* **16**:3 (1993), 317–331.

[Gibiansky and Lakes 1997] L. V. Gibiansky and R. S. Lakes, "Bounds on the complex bulk and shear moduli of a two-dimensional two-phase viscoelastic composite", *Mech. Mater.* **25**:2 (1997), 79–95.

[Gibiansky and Milton 1993] L. V. Gibiansky and G. W. Milton, "On the effective viscoelastic moduli of two-phase media, I: Rigorous bounds on the complex bulk modulus", *Proc. Roy. Soc. London Ser. A* **440**:1908 (1993), 163–188.

[Gibiansky and Sigmund 2000] L. V. Gibiansky and O. Sigmund, "Multiphase composites with extremal bulk modulus", *J. Mech. Phys. Solids* **48**:3 (2000), 461–498.

[Gibiansky et al. 1993] L. V. Gibiansky, R. S. Lakes, and G. W. Milton, "Viscoelastic composites with extremal properties", pp. 369–376 in *Structural Optimization 93* (Rio de Janeiro, 1993), World Cong. Opt. Design Struct. Syst. Proc. **1**, Fed. Univ. Rio de Janeiro, 1993.

[Gibiansky et al. 1999] L. V. Gibiansky, G. W. Milton, and J. G. Berryman, "On the effective viscoelastic moduli of two-phase media, III: Rigorous bounds on the complex shear modulus in two dimensions", *Proc. Roy. Soc. London Ser. A* **455**:1986 (1999), 2117–2149.

[Grabovsky 1996] Y. Grabovsky, "Bounds and extremal microstructures for two-component composites: a unified treatment based on the translation method", *Proc. Roy. Soc. London Ser. A* **452**:1947 (1996), 919–944.

[Grabovsky 1998] Y. Grabovsky, "Exact relations for effective tensors of polycrystals, I: Necessary conditions", *Arch. Rational Mech. Anal.* **143**:4 (1998), 309–329.

[Grabovsky and Kohn 1995a] Y. Grabovsky and R. V. Kohn, "Microstructures minimizing the energy of a two-phase elastic composite in two space dimensions, I: The confocal ellipse construction", *J. Mech. Phys. Solids* **43**:6 (1995), 933–947.

[Grabovsky and Kohn 1995b] Y. Grabovsky and R. V. Kohn, "Microstructures minimizing the energy of a two-phase elastic composite in two space dimensions, II: The Vigdergauz microstructure", *J. Mech. Phys. Solids* **43**:6 (1995), 949–972.

[Grabovsky and Sage 1998] Y. Grabovsky and D. S. Sage, "Exact relations for effective tensors of polycrystals, II: Applications to elasticity and piezoelectricity", *Arch. Rational Mech. Anal.* **143**:4 (1998), 331–356.

[Grabovsky et al. 2000] Y. Grabovsky, G. W. Milton, and D. S. Sage, "Exact relations for effective tensors of composites: necessary conditions and sufficient conditions", *Comm. Pure Appl. Math.* **53**:3 (2000), 300–353.

[Hashin 1962] Z. Hashin, "The elastic moduli of heterogeneous materials", *J. Appl. Math.* **29**:1 (1962), 143–150.

[Hashin 1965] H. Hashin, "On elastic behavior of fibre reinforced materials of arbitrary transverse phase geometry", *J. Mech. Phys. Solids* **13**:3 (1965), 119–134.

[Hashin and Shtrikman 1963] Z. Hashin and S. Shtrikman, "A variational approach to the theory of the elastic behaviour of multiphase materials", *J. Mech. Phys. Solids* **11**:2 (1963), 127–140.

[Hill 1952] R. Hill, "The elastic behaviour of a crystalline aggregate", *Proc. Phys. Soc. A* **65**:5 (1952), 349–354.

[Hill 1963] R. Hill, "Elastic properties of reinforced solids: some theoretical principles", *J. Mech. Phys. Solids* **11**:5 (1963), 357–372.

[Hill 1964] R. Hill, "Theory of mechanical properties of fibre-strengthened materials, I: Elastic behaviour", *J. Mech. Phys. Solids* **12** (1964), 199–212.

[Kadic et al. 2012] M. Kadic, T. Bückmann, N. Stenger, M. Thiel, and M. Wegener, "On the practicability of pentamode mechanical metamaterials", *Appl. Phys. Lett.* **100**:19 (2012), art. id. 191901, 5 pp.

[Kochmann and Milton 2014] D. M. Kochmann and G. W. Milton, "Rigorous bounds on the effective moduli of composites and inhomogeneous bodies with negative-stiffness phases", *J. Mech. Phys. Solids* **71** (2014), 46–63.

[Kohn and Lipton 1988] R. V. Kohn and R. Lipton, "Optimal bounds for the effective energy of a mixture of isotropic, incompressible, elastic materials", *Arch. Rational Mech. Anal.* **102**:4 (1988), 331–350.

[Kohn and Wirth 2014] R. V. Kohn and B. Wirth, "Optimal fine-scale structures in compliance minimization for a uniaxial load", *Proc. Roy. Soc. A* **470**:2170 (2014), art. id. 20140432, 16 pp.

[Kohn and Wirth 2016] R. V. Kohn and B. Wirth, "Optimal fine-scale structures in compliance minimization for a shear load", *Comm. Pure Appl. Math.* **69**:8 (2016), 1572–1610.

[Lipton 1988] R. Lipton, "On the effective elasticity of a two-dimensional homogenised incompressible elastic composite", *Proc. Roy. Soc. Edinburgh Sect. A* **110**:1-2 (1988), 45–61.

[Lipton 1991] R. Lipton, "On the behavior of elastic composites with transverse isotropic symmetry", *J. Mech. Phys. Solids* **39**:5 (1991), 663–681.

[Lipton 1992] R. Lipton, "Bounds and perturbation series for incompressible elastic composites with transverse isotropic symmetry", *J. Elasticity* **27**:3 (1992), 193–225.

[Lipton 1994] R. Lipton, "Optimal bounds on effective elastic tensors for orthotropic composites", *Proc. Roy. Soc. London Ser. A* **444**:1921 (1994), 399–410.

[Liu et al. 2007] L. Liu, R. D. James, and P. H. Leo, "Periodic inclusion-matrix microstructures with constant field inclusions", *Metal. Mater. Trans. A* **38**:4 (2007), 781–787.

[Lurie et al. 1982] K. A. Lurie, A. V. Cherkaev, and A. V. Fedorov, "Regularization of optimal design problems for bars and plates, II", *J. Optim. Theory Appl.* **37**:4 (1982), 523–543.

[Maxwell 1873] J. C. Maxwell, *A treatise on electricity and magnetism, I*, Clarendon, Oxford, 1873.

[Milton 1980] G. W. Milton, "Bounds on the complex dielectric constant of a composite material", *Appl. Phys. Lett.* **37**:3 (1980), 300–302.

[Milton 1981a] G. W. Milton, "Bounds on the complex permittivity of a two-component composite material", *J. Appl. Phys.* **52**:8 (1981), 5286–5293.

[Milton 1981b] G. W. Milton, "Bounds on the electromagnetic, elastic, and other properties of two-component composites", *Phys. Rev. Lett.* **46**:8 (1981), 542–545.

[Milton 1981c] G. W. Milton, "Concerning bounds on the transport and mechanical properties of multicomponent composite materials", *Appl. Phys. A* **26**:2 (1981), 125–130.

[Milton 1985] G. W. Milton, *Some exotic models in statistical physics, I: The coherent potential approximation is a realizable effective medium scheme, II: Anomalous first-order transitions (composites, phases, fractals)*, Ph.D. thesis, Cornell University, 1985, available at http://tinyurl.com/miltonphd.

[Milton 1986] G. W. Milton, "Modelling the properties of composites by laminates", pp. 150–174 in *Homogenization and effective moduli of materials and media* (Minneapolis, MN, 1984/1985), edited by J. L. Ericksen et al., The IMA Volumes in Mathematics and its Applications **1**, Springer, New York, 1986.

[Milton 1990] G. W. Milton, "On characterizing the set of possible effective tensors of composites: the variational method and the translation method", *Comm. Pure Appl. Math.* **43**:1 (1990), 63–125.

[Milton 1994] G. W. Milton, "A link between sets of tensors stable under lamination and quasiconvexity", *Comm. Pure Appl. Math.* **47**:7 (1994), 959–1003.

[Milton 2002] G. W. Milton, *The theory of composites*, Cambridge Monographs on Applied and Computational Mathematics **6**, Cambridge Univ. Press, 2002.

[Milton and Berryman 1997] G. W. Milton and J. G. Berryman, "On the effective viscoelastic moduli of two-phase media, II: Rigorous bounds on the complex shear modulus in three dimensions", *Proc. Roy. Soc. London Ser. A* **453**:1964 (1997), 1849–1880.

[Milton and Camar-Eddine 2016] G. W. Milton and M. Camar-Eddine, "Complete characterization of the possible (average strain, average stress, volume fraction) triplets that can occur in 3-d printed materials", to be submitted, 2016.

[Milton and Cherkaev 1995] G. W. Milton and A. V. Cherkaev, "Which elasticity tensors are realizable?", *J. Eng. Mater. Tech.* **117**:4 (1995), 483–493.

[Milton and Kohn 1988] G. W. Milton and R. V. Kohn, "Variational bounds on the effective moduli of anisotropic composites", *J. Mech. Phys. Solids* **36**:6 (1988), 597–629.

[Milton and McPhedran 1982] G. W. Milton and R. C. McPhedran, "A comparison of two methods for deriving bounds on the effective conductivity of composites", pp. 183–193 in *Macroscopic properties of disordered media: proceedings of a conference held at the Courant Institute*, edited by R. Burridge et al., Lecture Notes in Physics **154**, Springer, Berlin, 1982.

[Milton and Phan-Thien 1982] G. W. Milton and N. Phan-Thien, "New bounds on effective elastic moduli of two-component materials", *Proc. Roy. Soc. London Ser. A* **380**:1779 (1982), 305–331.

[Milton et al. 2003] G. W. Milton, S. K. Serkov, and A. B. Movchan, "Realizable (average stress, average strain) pairs in a plate with holes", *SIAM J. Appl. Math.* **63**:3 (2003), 987–1028.

[Milton et al. 2017] G. W. Milton, D. Harutyunyan, and M. Briane, "Towards a complete characterization of the effective elasticity tensors of mixtures of an elastic phase and an almost rigid phase", *Math. Mech. Complex Syst.* **5**:1 (2017), 95–113.

[Nemat-Nasser and Hori 1993] S. Nemat-Nasser and M. Hori, *Micromechanics: overall properties of heterogeneous materials*, North-Holland Series in Applied Mathematics and Mechanics **37**, North-Holland, Amsterdam, 1993.

[Nemat-Nasser and Hori 1998] S. Nemat-Nasser and M. Hori, *Micromechanics: overall properties of heterogeneous materials*, 2nd ed., Elsevier, Amsterdam, 1998.

[Norris 1985] A. N. Norris, "A differential scheme for the effective moduli of composites", *Mech. Mater.* **4**:1 (1985), 1–16.

[Phani and Hussein 2017] A. S. Phani and M. I. Hussein (editors), *Dynamics of lattice materials*, Wiley, New York, 2017.

[Roscoe 1973] R. Roscoe, "Isotropic composites with elastic or viscoelastic phases: general bounds for the moduli and solutions for special geometries", *Rheol. Acta* **12**:3 (1973), 404–411.

[Seppecher et al. 2011] P. Seppecher, J.-J. Alibert, and dell'Isola, Francesco, "Linear elastic trusses leading to continua with exotic mechanical interactions", *J. Phys. Conf. Ser.* **319**:1 (2011), art. id. 012018, 13 pp.

[Sigmund 1994] O. Sigmund, "Materials with prescribed constitutive parameters: an inverse homogenization problem", *Internat. J. Solids Structures* **31**:17 (1994), 2313–2329.

[Sigmund 1995] O. Sigmund, "Tailoring materials with prescribed elastic properties", *Mech. Mater.* **20**:4 (1995), 351–368.

[Sigmund 2000] O. Sigmund, "A new class of extremal composites", *J. Mech. Phys. Solids* **48**:2 (2000), 397–428.

[Tartar 1979] L. Tartar, "Estimation de coefficients homogénéisés", pp. 364–373 in *Computing methods in applied sciences and engineering: proceedings of the Third International Symposium* (Versailles, 1977), edited by R. Glowinski and J.-L. Lions, Lecture Notes in Mathematics **704**, Springer, Berlin, 1979. Translated in *Topics Math. Model. Comp. Mater.* **31** (1997), 9–20.

[Tartar 1985] L. Tartar, "Estimations fines des coefficients homogénéisés", pp. 168–187 in *Ennio De Giorgi colloquium* (Paris, 1983), edited by P. Krée, Research Notes in Mathematics **125**, Pitman, Boston, 1985.

[Tartar 1989] L. Tartar, "H-measures and small amplitude homogenization", pp. 89–99 in *Random media and composites* (Leesburg, VA, 1988), edited by R. V. Kohn and G. W. Milton, SIAM, Philadelphia, PA, 1989.

[Tartar 1990] L. Tartar, "H-measures, a new approach for studying homogenisation, oscillations and concentration effects in partial differential equations", *Proc. Roy. Soc. Edinburgh Sect. A* **115**:3-4 (1990), 193–230.

[Tartar 2009] L. Tartar, *The general theory of homogenization*, Lecture Notes of the Unione Matematica Italiana **7**, Springer, Berlin, 2009.

[Torquato 2002] S. Torquato, *Random heterogeneous materials: microstructure and macroscopic properties*, Interdisciplinary Applied Mathematics **16**, Springer, New York, 2002.

[Vigdergauz 1986] S. B. Vigdergauz, "Effective elastic parameters of a plate with a regular system of equal-strength holes", *Inzh. Zh. Mekh. Tverd. Tela* **21**:2 (1986), 165–169. In Russian.

[Vigdergauz 1994] S. Vigdergauz, "Two-dimensional grained composites of extreme rigidity", *J. Appl. Mech.* **61**:2 (1994), 390–394.

[Vigdergauz 1996] S. Vigdergauz, "Rhombic lattice of equi-stress inclusions in an elastic plate", *Quart. J. Mech. Appl. Math.* **49**:4 (1996), 565–580.

[Vigdergauz 1999] S. Vigdergauz, "Energy-minimizing inclusions in a planar elastic structure with macroisotropy", *Struct. Optimiz.* **17**:2 (1999), 104–112.

[Walpole 1966] L. J. Walpole, "On bounds for the overall elastic moduli of inhomogeneous systems, I", *J. Mech. Phys. Solids* **14**:3 (1966), 151–162.

[Willis 1977] J. R. Willis, "Bounds and self-consistent estimates for the overall properties of anisotropic composites", *J. Mech. Phys. Solids* **25**:3 (1977), 185–202.

[Zhikov 1988] V. V. Zhikov, "Estimates for the trace of an averaged tensor", *Dokl. Akad. Nauk SSSR* **299**:4 (1988), 796–800. In Russian; translated in *Soviet Math. Dokl.* **37**:2 (1988), 456–459.

[Zhikov 1991a] V. V. Zhikov, "Estimates for an averaged matrix and an averaged tensor", *Uspekhi Mat. Nauk* **46**:3(279) (1991), 49–109. In Russian; translated in *Russ. Math. Surv.* **46**:3 (1991), 65–136; errata in **47**:1 (1992) 278.

[Zhikov 1991b] V. V. Zhikov, "Estimates for the trace of an averaged tensor", *Ukrain. Mat. Zh.* **43**:6 (1991), 745–755. In Russian; translated in *Ukrainian Math. J.* **43**:6 (1992), 694–704.

Received 10 Jun 2016. Revised 11 Oct 2016. Accepted 14 Nov 2016.

GRAEME W. MILTON: milton@math.utah.edu
Department of Mathematics, University of Utah, 155 South 1400 East Room 233, Salt Lake City, UT 84112-0090, United States

MARC BRIANE: mbriane@insa-rennes.fr
Institut de Recherche Mathématique de Rennes, INSA de Rennes, 20 Avenue des Buttes de Coësmes, CS 70839, 35708 Rennes Cedex 7, France

DAVIT HARUTYUNYAN: davith@math.utah.edu
Department of Mathematics, University of Utah, 155 South 1400 East Room 233, Salt Lake City, UT 84112-0090, United States

MATHEMATICS AND MECHANICS OF COMPLEX SYSTEMS
Vol. 5, No. 1, 2017
dx.doi.org/10.2140/memocs.2017.5.95

TOWARDS A COMPLETE CHARACTERIZATION OF THE EFFECTIVE ELASTICITY TENSORS OF MIXTURES OF AN ELASTIC PHASE AND AN ALMOST RIGID PHASE

GRAEME W. MILTON, DAVIT HARUTYUNYAN AND MARC BRIANE

The set GU_f of possible effective elastic tensors of composites built from two materials with positive definite elasticity tensors $\boldsymbol{C}_1$ and $\boldsymbol{C}_2 = \delta\boldsymbol{C}_0$ comprising the set $U = \{\boldsymbol{C}_1, \delta\boldsymbol{C}_0\}$ and mixed in proportions f and $1-f$ is partly characterized in the limit $\delta \to \infty$. The material with tensor $\boldsymbol{C}_2$ corresponds to a material which (for technical reasons) is almost rigid in the limit $\delta \to \infty$. This paper, and the underlying microgeometries, has many aspects in common with the companion paper "On the possible effective elasticity tensors of 2-dimensional and 3-dimensional printed materials". The chief difference is that one has a different algebraic problem to solve: determining the subspaces of stress fields for which the thin walled structures can be rigid, rather than determining, as in the companion paper, the subspaces of strain fields for which the thin walled structure is compliant. Recalling that GU_f is completely characterized through minimums of sums of energies, involving a set of applied strains, and complementary energies, involving a set of applied stresses, we provide descriptions of microgeometries that in appropriate limits achieve the minimums in many cases. In these cases the calculation of the minimum is reduced to a finite-dimensional minimization problem that can be done numerically. Each microgeometry consists of a union of walls in appropriate directions, where the material in the wall is an appropriate p-mode material that is almost rigid to $6-p \leq 5$ independent applied stresses, yet is compliant to any strain in the orthogonal space. Thus the walls, by themselves, can support stress with almost no deformation. The region outside the walls contains "Avellaneda material", which is a hierarchical laminate that minimizes an appropriate sum of elastic energies.

1. Introduction

This paper is a companion to "On the possible effective elasticity tensors of 2-dimensional and 3-dimensional printed materials" [Milton et al. 2017], which gives a partial characterization of the set GU_f of effective elasticity tensors that can be

Communicated by Robert P. Lipton.
MSC2010: 74Q20, 35Q74.
Keywords: elasticity G-closure, composites, metamaterials.

produced in the limit $\delta \to 0$ if we mix in prescribed proportions f and $1-f$ two materials with positive definite and bounded elasticity tensors $\boldsymbol{C}_1$ and $\boldsymbol{C}_2 = \delta\boldsymbol{C}_0$. Here we consider the opposite limit $\delta \to \infty$, which corresponds to mixing in prescribed proportions an elastic phase and an almost rigid phase. Our results are summarized in the theorem in the conclusion section. For a complete introduction and summary of previous results the reader is urged to read at least the first three sections of the companion paper. The essential ideas presented here are much the same as those contained in the companion paper. However, the algebraic problem relevant to this paper, of determining when the set of walls can support a set of stress fields, is quite different from the algebraic problem encountered in the companion paper of determining when the set of walls is compliant to a set of strain fields.

The microstructures we consider involve taking three limits. First, as they have structure on multiple length scales, the homogenization limit where the ratio between length scales goes to infinity needs to be taken. Second, the limit $\delta \to \infty$ needs to be taken. Third, as the structure involves walls of width ϵ, which are very stiff to certain applied stresses, the limit $\epsilon \to 0$ needs to be taken so the contribution to the elastic energy of these walls goes to zero, when the structure is compliant to an applied strain. The limits should be taken in this order, as, for example, standard homogenization theory is justified only if δ is positive and finite, so we need to take the homogenization limit before taking the limit $\delta \to \infty$.

As in the companion paper we emphasize that our analysis is valid only for linear elasticity, and ignores nonlinear effects such as buckling, which may be important even for small deformations. It is also important to emphasize that to apply our results when phase 2 is perfectly rigid (rather than almost rigid) requires special care. Indeed, if phase 2 is perfectly rigid, then many of the microgeometries considered here do not permit the kind of motions that are permitted for any finite value of δ, no matter how large. In particular, the structures considered in Figures 6, 8, and 9(d) of the companion paper would be completely rigid if phase 2 was perfectly rigid. To permit the required motions, one has to first replace the rigid phase 2 with a composite with a small amount of phase 1 as the matrix phase, so that its effective elasticity tensor is finite but approaches infinity as the proportion of phase 1 in it tends to zero. The microgeometry in this composite needs to be much smaller than the scales in the geometries discussed here, which would involve mixtures of it and phase 1.

2. Characterizing G closures through sums of energies and complementary energies

Cherkaev and Gibiansky [1992; 1993] found that bounding sums of energies and complementary energies could lead to very useful bounds on G-closures. It was subsequently proved in [Francfort and Milton 1994; Milton 1994] that minimums over $\boldsymbol{C}_* \in GU_f$ of such sums of energies and complementary energies completely characterize GU_f in much the same way that Legendre transforms characterize convex sets: the stability under lamination of GU_f is what allows one to recover GU_f from the values of these minimums (see also Chapter 30 in [Milton 2002]). Specifically, in the case of 3-dimensional elasticity, the set GU_f is completely characterized if we know the seven "energy functions",

$$
\begin{aligned}
W_f^0(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0, \boldsymbol{\sigma}_5^0, \boldsymbol{\sigma}_6^0) &= \min_{\boldsymbol{C}_* \in GU_f} \sum_{j=1}^{6} \boldsymbol{\sigma}_j^0 : \boldsymbol{C}_*^{-1} \boldsymbol{\sigma}_j^0, \\
W_f^1(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0, \boldsymbol{\sigma}_5^0, \boldsymbol{\epsilon}_1^0) &= \min_{\boldsymbol{C}_* \in GU_f} \left[\boldsymbol{\epsilon}_1^0 : \boldsymbol{C}_* \boldsymbol{\epsilon}_1^0 + \sum_{j=1}^{5} \boldsymbol{\sigma}_j^0 : \boldsymbol{C}_*^{-1} \boldsymbol{\sigma}_j^0 \right], \\
W_f^2(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0) &= \min_{\boldsymbol{C}_* \in GU_f} \left[\sum_{i=1}^{2} \boldsymbol{\epsilon}_i^0 : \boldsymbol{C}_* \boldsymbol{\epsilon}_i^0 + \sum_{j=1}^{4} \boldsymbol{\sigma}_j^0 : \boldsymbol{C}_*^{-1} \boldsymbol{\sigma}_j^0 \right], \\
W_f^3(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0) &= \min_{\boldsymbol{C}_* \in GU_f} \left[\sum_{i=1}^{3} \boldsymbol{\epsilon}_i^0 : \boldsymbol{C}_* \boldsymbol{\epsilon}_i^0 + \sum_{j=1}^{3} \boldsymbol{\sigma}_j^0 : \boldsymbol{C}_*^{-1} \boldsymbol{\sigma}_j^0 \right], \qquad (2\text{-}1) \\
W_f^4(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0) &= \min_{\boldsymbol{C}_* \in GU_f} \left[\sum_{i=1}^{4} \boldsymbol{\epsilon}_i^0 : \boldsymbol{C}_* \boldsymbol{\epsilon}_i^0 + \sum_{j=1}^{2} \boldsymbol{\sigma}_j^0 : \boldsymbol{C}_*^{-1} \boldsymbol{\sigma}_j^0 \right], \\
W_f^5(\boldsymbol{\sigma}_1^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0, \boldsymbol{\epsilon}_5^0) &= \min_{\boldsymbol{C}_* \in GU_f} \left[\left(\sum_{i=1}^{5} \boldsymbol{\epsilon}_i^0 : \boldsymbol{C}_* \boldsymbol{\epsilon}_i^0 \right) + \boldsymbol{\sigma}_1^0 : \boldsymbol{C}_*^{-1} \boldsymbol{\sigma}_1^0 \right], \\
W_f^6(\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0, \boldsymbol{\epsilon}_5^0, \boldsymbol{\epsilon}_6^0) &= \min_{\boldsymbol{C}_* \in GU_f} \sum_{i=1}^{6} \boldsymbol{\epsilon}_i^0 : \boldsymbol{C}_* \boldsymbol{\epsilon}_i^0.
\end{aligned}
$$

In fact, Milton and Cherkaev [1995] showed it suffices to know these functions for sets of applied strains $\boldsymbol{\epsilon}_i^0$ and applied stresses $\boldsymbol{\sigma}_j^0$ that are mutually orthogonal:

$$
(\boldsymbol{\epsilon}_i^0, \boldsymbol{\sigma}_j^0) = 0, \quad (\boldsymbol{\epsilon}_i^0, \boldsymbol{\epsilon}_k^0) = 0, \quad (\boldsymbol{\sigma}_j^0, \boldsymbol{\sigma}_\ell^0) = 0, \quad \text{for all } i, j, k, \ell \text{ with } i \neq j,\ i \neq k,\ j \neq \ell. \qquad (2\text{-}2)
$$

The terms appearing in the minimums have a physical significance. For example, in the expression for W_f^2,

$$
\sum_{i=1}^{2} \boldsymbol{\epsilon}_i^0 : \boldsymbol{C}_* \boldsymbol{\epsilon}_i^0 + \sum_{j=1}^{4} \boldsymbol{\sigma}_j^0 : \boldsymbol{C}_*^{-1} \boldsymbol{\sigma}_j^0 \qquad (2\text{-}3)
$$

has the physical interpretation of being the sum of energies per unit volume stored in the composite with effective elasticity tensor $\boldsymbol{C}_*$ when successively subjected to the two applied strains $\boldsymbol{\epsilon}_1^0$ and $\boldsymbol{\epsilon}_2^0$ and then to the four applied stresses $\boldsymbol{\sigma}_1^0$, $\boldsymbol{\sigma}_2^0$, $\boldsymbol{\sigma}_3^0$, and $\boldsymbol{\sigma}_4^0$. To distinguish the terms $\boldsymbol{\epsilon}_i^0 : \boldsymbol{C}_*\boldsymbol{\epsilon}_i^0$ and $\boldsymbol{\sigma}_j^0 : \boldsymbol{C}_*^{-1}\boldsymbol{\sigma}_j^0$, the first is called an energy (it is really an energy per unit volume associated with the applied strain $\boldsymbol{\epsilon}_i^0$) and the second is called a complementary energy, although it too physically represents an energy per unit volume associated with the applied stress $\boldsymbol{\sigma}_j^0$.

For well-ordered materials with $\boldsymbol{C}_2 \geq \boldsymbol{C}_1$ (or the reverse), Avellaneda [1987] showed that there exist sequentially layered laminates of finite rank having an effective elasticity tensor $\boldsymbol{C}_* = \boldsymbol{C}_f^A(\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0, \boldsymbol{\epsilon}_5^0, \boldsymbol{\epsilon}_6^0)$ (not to be confused with the elasticity tensor $\boldsymbol{C}_* = \widetilde{\boldsymbol{C}}_f^A(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0, \boldsymbol{\sigma}_5^0, \boldsymbol{\sigma}_6^0)$ used in the companion paper) that attains the minimum in the above expression for $W_f^6(\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0, \boldsymbol{\epsilon}_5^0, \boldsymbol{\epsilon}_6^0)$:

$$W_f^6(\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0, \boldsymbol{\epsilon}_5^0, \boldsymbol{\epsilon}_6^0) = \sum_{i=1}^{6} \boldsymbol{\epsilon}_i^0 : \boldsymbol{C}_f^A(\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0, \boldsymbol{\epsilon}_5^0, \boldsymbol{\epsilon}_6^0)\boldsymbol{\epsilon}_i^0. \tag{2-4}$$

The effective tensor $\boldsymbol{C}_* = \boldsymbol{C}_f^A(\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0, \boldsymbol{\epsilon}_5^0, \boldsymbol{\epsilon}_6^0)$ of the Avellaneda material is found by finding a combination of the parameters entering the formula for the effective tensor of sequentially layered laminates that minimizes the sum of six elastic energies. In general this has to be done numerically, but it suffices to consider laminates of rank at most 6 if $\boldsymbol{C}_1$ is isotropic [Francfort et al. 1995], or, using an argument of Avellaneda [1987], to consider laminates of rank at most 21 if $\boldsymbol{C}_1$ is anisotropic (see Section 2 in the companion paper).

In the case of 2-dimensional elasticity, the set GU_f is similarly completely characterized if we know the 4 "energy functions",

$$\begin{aligned}
W_f^0(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0) &= \min_{\boldsymbol{C}_* \in GU_f} \sum_{j=1}^{3} \boldsymbol{\sigma}_j^0 : \boldsymbol{C}_*^{-1}\boldsymbol{\sigma}_j^0, \\
W_f^1(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\epsilon}_1^0) &= \min_{\boldsymbol{C}_* \in GU_f} \left[\boldsymbol{\epsilon}_1^0 : \boldsymbol{C}_*\boldsymbol{\epsilon}_1^0 + \sum_{j=1}^{2} \boldsymbol{\sigma}_j^0 : \boldsymbol{C}_*^{-1}\boldsymbol{\sigma}_j^0\right], \\
W_f^2(\boldsymbol{\sigma}_1^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0) &= \min_{\boldsymbol{C}_* \in GU_f} \left[\left(\sum_{i=1}^{2} \boldsymbol{\epsilon}_i^0 : \boldsymbol{C}_*\boldsymbol{\epsilon}_i^0\right) + \boldsymbol{\sigma}_1^0 : \boldsymbol{C}_*^{-1}\boldsymbol{\sigma}_1^0\right], \\
W_f^3(\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0) &= \min_{\boldsymbol{C}_* \in GU_f} \sum_{i=1}^{3} \boldsymbol{\epsilon}_i^0 : \boldsymbol{C}_*\boldsymbol{\epsilon}_i^0.
\end{aligned} \tag{2-5}$$

Again $W_f^3(\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0)$ is attained for an "Avellaneda material" consisting of a sequentially layered laminate geometry having an effective tensor $\boldsymbol{C}_* = \boldsymbol{C}_f^A(\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0)$, i.e.,

$$W_f^3(\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0) = \sum_{i=1}^{3} \boldsymbol{\epsilon}_i^0 : \boldsymbol{C}_f^A(\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0)\boldsymbol{\epsilon}_i^0. \tag{2-6}$$

The effective tensor $\boldsymbol{C}_* = \boldsymbol{C}_f^A(\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0)$ of the Avellaneda material is found by finding a combination of the parameters entering the formula for the effective tensor of sequentially layered laminates that minimizes the sum of three elastic energies. In general this has to be done numerically, but it suffices to consider laminates of rank at most three if $\boldsymbol{C}_1$ is isotropic [Avellaneda and Milton 1989], or, using an argument of Avellaneda [1987], to consider laminates of rank at most 6 if $\boldsymbol{C}_1$ is anisotropic (see Section 2 in the companion paper).

3. Microgeometries which are associated with sharp bounds on many sums of energies and complementary energies

The analysis here of mixtures of an almost rigid phase mixed with an elastic phase is very similar to the analysis in the companion paper for mixtures of an extremely compliant phase and an elastic phase. The roles of stresses and strains are interchanged and now the challenge is to identify matrix pencils that are spanned by matrices with zero determinant, rather than symmetrized rank 1 matrices. We now have the inequalities

$$\begin{aligned}
0 &\le W_f^0(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0, \boldsymbol{\sigma}_5^0, \boldsymbol{\sigma}_6^0),\\
\boldsymbol{\epsilon}_1^0 : [\boldsymbol{C}_f^A(0,0,0,0,0,\boldsymbol{\epsilon}_1^0)]\boldsymbol{\epsilon}_1^0 &\le W_f^1(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0, \boldsymbol{\sigma}_5^0, \boldsymbol{\epsilon}_1^0),\\
\sum_{i=1}^{2}\boldsymbol{\epsilon}_i^0 : [\boldsymbol{C}_f^A(0,0,0,0,\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0)]\boldsymbol{\epsilon}_i^0 &\le W_f^2(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0),\\
\sum_{i=1}^{3}\boldsymbol{\epsilon}_i^0 : [\boldsymbol{C}_f^A(0,0,0,\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0,\boldsymbol{\epsilon}_3^0)]\boldsymbol{\epsilon}_i^0 &\le W_f^3(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0),\\
\sum_{i=1}^{4}\boldsymbol{\epsilon}_i^0 : [\boldsymbol{C}_f^A(0,0,\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0,\boldsymbol{\epsilon}_3^0,\boldsymbol{\epsilon}_4^0)]\boldsymbol{\epsilon}_i^0 &\le W_f^4(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0),\\
\sum_{i=1}^{5}\boldsymbol{\epsilon}_i^0 : [\boldsymbol{C}_f^A(0,\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0,\boldsymbol{\epsilon}_3^0,\boldsymbol{\epsilon}_4^0,\boldsymbol{\epsilon}_5^0)]\boldsymbol{\epsilon}_i^0 &\le W_f^5(\boldsymbol{\sigma}_1^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0, \boldsymbol{\epsilon}_5^0).
\end{aligned} \tag{3-1}$$

The first inequality is clearly sharp, being attained when the composite consists of islands of phase 1 surrounded by a phase 2 (so that $(\boldsymbol{C}_*)^{-1}$ approaches 0 as $\delta \to \infty$). Again the objective is to show that many of the other inequalities are also sharp in the limit $\delta \to \infty$, at least when the spaces spanned by the applied stresses $\boldsymbol{\sigma}_j^0$ for $j = 1, 2, \ldots, 6-p$ satisfy certain properties. This space of applied stresses associated with W_f^p has dimension $6-p$ and its orthogonal complement defines the p-dimensional space $\mathcal{V}_p$.

The recipe for doing this is to simply insert into a relevant Avellaneda material a microstructure occupying a thin walled region containing a p-mode material, such that the walled structure, by itself, is very stiff when the applied stress lies in the $(6-p)$-dimensional subspace spanned by the $\boldsymbol{\sigma}_j^0$, yet allows strains in the orthogonal p-dimensional subspace $\mathcal{V}_p$ spanned by the $\boldsymbol{\epsilon}_i^0$. We say a composite with effective tensor $\boldsymbol{C}_*$ built from the two materials $\boldsymbol{C}_1$ and $\boldsymbol{C}_2 = \delta\boldsymbol{C}_0$ is very stiff to a stress $\boldsymbol{\sigma}_j^0$ if the complementary energy $\boldsymbol{\sigma}_j^0 : \boldsymbol{C}_*^{-1}\boldsymbol{\sigma}_j^0$ goes to zero as $\delta \to \infty$, and allows a strain $\boldsymbol{\epsilon}_i^0$ if the elastic energy $\boldsymbol{\epsilon}_i^0 : \boldsymbol{C}_*\boldsymbol{\epsilon}_i^0$ has a finite limit as $\delta \to \infty$. These p-mode materials have exactly the same construction as that specified in Section 8 of the companion paper, only now the region that was occupied by the elastic phase is now occupied by the rigid phase, and the material that was occupied by the extremely compliant phase (which becomes void in the limit $\delta \to 0$) is occupied by the elastic phase. If we happened to choose $\boldsymbol{C}_0 = \boldsymbol{C}_1$, all the moduli (and effective moduli) are simply rescaled, i.e., for any δ, and in particular for large values of δ, if a mixture of two materials with effective tensors $\boldsymbol{C}_1$ and $\boldsymbol{C}_1/\delta$ has effective tensor $\boldsymbol{C}_*$, then when rescaling the elasticity tensors of the two phases to $\delta\boldsymbol{C}_1$ and $\boldsymbol{C}_1$, the resulting effective elasticity tensor will be $\delta\boldsymbol{C}_*$. Thus, the analysis of the response of the p-mode materials is essentially the same as in the companion paper. Exactly the same trial fields can be chosen to bound the response of the p-mode material. Hence we do not repeat this analysis but instead the reader is referred to Section 8 of the companion paper.

The subspace orthogonal to $\mathcal{V}_p$ is now required to be spanned by matrices $\boldsymbol{\sigma}^{(k)}$, for $k = 1, \dots, 6-p$, such that

$$\boldsymbol{\sigma}^{(k)}\boldsymbol{n}_k = 0 \tag{3-2}$$

for some unit vector $\boldsymbol{n}_k$. Thus the identifying feature of these matrices $\boldsymbol{\sigma}^{(k)}$ is that they have zero determinant, and then $\boldsymbol{n}_k$ can be chosen as a null vector of $\boldsymbol{\sigma}^{(k)}$. The existence of such matrices $\boldsymbol{\sigma}^{(k)}$ is proved in Section 4. The proof uses small perturbations of the applied stresses and strains. But, due the continuity of the energy functions W_f^k established in Section 5, the small perturbations do not modify the generic result. The vectors $\boldsymbol{n}_k$ determine the orientation of the walls in the structure since a set of walls orthogonal to $\boldsymbol{n}$ can support any stress $\boldsymbol{\sigma}$ such that $\boldsymbol{\sigma}\boldsymbol{n} = 0$.

To define the thin walled structure, introduce the periodic function $H_c(x)$ with period 1 which takes the value 1 if $x - [x] \le c$, where $[x]$ is the greatest integer less than x, and $c \in [0, 1]$ gives the relative thickness of each wall. Then for the unit vectors $\boldsymbol{n}_1, \boldsymbol{n}_2, \dots, \boldsymbol{n}_{6-p}$ appearing in (3-2), and for a small relative thickness $c = \epsilon$, define the characteristic functions

$$\eta_k(\boldsymbol{x}) = H_\epsilon(\boldsymbol{x} \cdot \boldsymbol{n}_k + k/p). \tag{3-3}$$

This characteristic function defines a series of parallel walls, as shown on the left in Figure 1, each perpendicular to the vector $\boldsymbol{n}_j$, where $\eta_j(\boldsymbol{x}) = 1$ in the wall material. The additional shift term k/p in (3-3) ensures the walls associated with

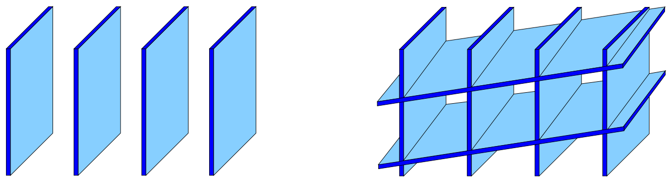

Figure 1. Example of walled structures. On the left we have a "rank 1" walled structure and on the right a "rank 2" walled structure. The generalization to walled structures of any rank is obvious, and precisely defined by the characteristic function (3-4) that is 0 in the walls, and 1 in the remaining material.

k_1 and k_2 do not intersect when it happens that $\boldsymbol{n}_{k_1} = \boldsymbol{n}_{k_2}$, at least when ϵ is small. We emphasize that ϵ is not a homogenization parameter, but rather represents a volume fraction of walls.

Now define the characteristic function

$$\chi_*(\boldsymbol{x}) = \prod_{k=1}^{p} (1 - \eta_k(\boldsymbol{x})). \tag{3-4}$$

If $p \leq 3$, this is usually a periodic function of $\boldsymbol{x}$, an exception being if $p = 3$ and there are no nonzero integers z_1, z_2, and z_3 such that $z_1\boldsymbol{n}_1 + z_2\boldsymbol{n}_2 + z_3\boldsymbol{n}_3 = 0$. More generally, $\chi_*(\boldsymbol{x})$ is a quasiperiodic function of $\boldsymbol{x}$. The walled structure is where $\chi_*(\boldsymbol{x})$ takes the value 0. In the case $p = 2$ the walled structure is illustrated on the right in Figure 1.

The walled structure is where $\chi_*(\boldsymbol{x})$ given by (3-4) takes the value 0. Inside it we put a p-mode material with effective tensor $\boldsymbol{C}_*^2 = \boldsymbol{C}_*(\mathcal{V}_p)$ that allows any applied strain $\boldsymbol{\epsilon}^0$ in the space $\mathcal{V}_p$ but which is very stiff to any stress $\boldsymbol{\sigma}^0$ orthogonal to the space $\mathcal{V}_p$. Using the six matrices

$$\boldsymbol{v}_1 = \boldsymbol{\sigma}_1^0/|\boldsymbol{\sigma}_1^0|, \ldots, \boldsymbol{v}_{6-p} = \boldsymbol{\sigma}_{6-p}^0/|\boldsymbol{\sigma}_{6-p}^0|, \boldsymbol{v}_{7-p} = \boldsymbol{\epsilon}_1^0/|\boldsymbol{\epsilon}_1^0|, \ldots, \boldsymbol{v}_6 = \boldsymbol{\epsilon}_p^0/|\boldsymbol{\epsilon}_p^0| \tag{3-5}$$

as our basis for the 6-dimensional space of 3×3 symmetric matrices, the compliance tensor $[\boldsymbol{C}_*(\mathcal{V}_p)]^{-1}$ in this basis takes the limiting form

$$\lim_{\delta \to \infty} [\boldsymbol{C}_*(\mathcal{V}_p)]^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & \boldsymbol{B} \end{pmatrix}, \tag{3-6}$$

where $\boldsymbol{B}$ represents a (strictly) positive definite $p \times p$ matrix and the 0 on the diagonal represents the $(6-p) \times (6-p)$ zero matrix. Inside the walled structure, where $\chi_*(\boldsymbol{x}) = 1$, we put the Avellaneda material with effective elasticity tensor

$$\boldsymbol{C}_*^1 = \boldsymbol{C}_f^A(0, \ldots, 0, \boldsymbol{\epsilon}_1^0, \ldots, \boldsymbol{\epsilon}_p^0).$$

In a variational principle similar to (4-4) in the companion paper (i.e., treating the Avellaneda material and the p-mode material both as homogeneous materials with effective tensors $\boldsymbol{C}_*^1 = \boldsymbol{C}_f^A$ and $\boldsymbol{C}_*^2 = \boldsymbol{C}_*(\mathcal{V}_p)$, respectively) we choose trial strain fields that are constant,

$$\underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x}) = \boldsymbol{\epsilon}_i^0, \quad \text{for } i = 1, 2, \dots, p, \tag{3-7}$$

thus trivially fulfilling the differential constraints, and trial stress fields of the form

$$\underline{\boldsymbol{\sigma}}_j(\boldsymbol{x}) = \sum_{k=1}^{6-p} \boldsymbol{\sigma}_{j,k}\eta_k(\boldsymbol{x})/\epsilon, \tag{3-8}$$

which are required to have the average values

$$\boldsymbol{\sigma}_j^0 = \langle \underline{\boldsymbol{\sigma}}_j \rangle = \sum_{k=1}^{6-p} \boldsymbol{\sigma}_{j,k}, \tag{3-9}$$

and the matrices $\boldsymbol{\sigma}_{i,j}$ are additionally required to lie in the space orthogonal to $\mathcal{V}_p$ (so they cost very little energy) and satisfy

$$\boldsymbol{\sigma}_{j,k} = c_{j,k}\boldsymbol{\sigma}^{(k)}, \tag{3-10}$$

for some choice of parameters $c_{j,k}$ to ensure that $\boldsymbol{\sigma}_{j,k}\boldsymbol{n}_k = 0$ and hence that $\underline{\boldsymbol{\sigma}}_j(\boldsymbol{x})$ satisfies the differential constraints of a stress field — this requires $\underline{\boldsymbol{\sigma}}_j(\boldsymbol{x})\boldsymbol{n}_k$ to be continuous across any interface with normal $\boldsymbol{n}_k$. Additionally, the $c_{j,k}$ in (3-10) should be chosen so the $\boldsymbol{\sigma}_j^0$ given by (3-9) are orthogonal.

To find upper bounds on the energy associated with this trial stress field, first consider those parts of the walled structure that are outside of any junction regions, i.e., where for some k we have $\eta_k(\boldsymbol{x}) = 1$, while $\eta_s(\boldsymbol{x}) = 0$ for all $s \neq k$. An upper bound for the volume fraction occupied by the region where $\eta_k(\boldsymbol{x}) = 1$ while $\eta_s(\boldsymbol{x}) = 0$ for all $s \neq k$ is of course ϵ, as this represents the volume of the region where $\eta_k(\boldsymbol{x}) = 1$. The associated energy per unit volume of the trial stress field in those parts of the walled structure that are outside of any junction regions is bounded above by

$$\sum_{k=1}^{6-p} \boldsymbol{\sigma}_{j,k} : [\boldsymbol{C}_*(\mathcal{V}_p)]^{-1}\boldsymbol{\sigma}_{j,k}/\epsilon. \tag{3-11}$$

With an appropriate choice of multimode material, one can construct bounded trial stress fields that are essentially concentrated in phase 2, and consequently, $\boldsymbol{\sigma}_{j,k} : [\boldsymbol{C}_*(\mathcal{V}_p)]^{-1}\boldsymbol{\sigma}_{j,k}$ is bounded above by a quantity proportional to $1/\delta$. Our assumption that we take the limit $\delta \to \infty$ before taking the limit $\epsilon \to 0$ ensures that $1/(\delta\epsilon) \to 0$, and thus ensures that the quantity (3-11) goes to zero in this limit.

Next, consider those junction regions where only two walls meet, i.e., where for some k_1 and $k_2 > k_1$, $\boldsymbol{x}$ is such that $\eta_{k_1}(\boldsymbol{x}) = \eta_{k_2}(\boldsymbol{x}) = 1$ while $\eta_s(\boldsymbol{x}) = 0$ for all s not equal to k_1 or k_2. Provided $\boldsymbol{n}_{k_1} \neq \boldsymbol{n}_{k_2}$, an upper bound for the volume fraction occupied by each such junction region is ϵ^2. Then the associated energy per unit volume of the trial stress field in these junction regions where only two walls meet is bounded above by

$$\sum_{k_1=1}^{6-p} \sum_{k_2=k_1+1}^{6-p} (\boldsymbol{\sigma}_{i,k_1} + \boldsymbol{\sigma}_{i,k_2}) : [\boldsymbol{C}_*(\mathcal{V}_p)]^{-1} (\boldsymbol{\sigma}_{j,k_1} + \boldsymbol{\sigma}_{j,k_2}). \tag{3-12}$$

Thus, the powers of ϵ cancel and this energy density goes to zero if the multimode material is easily compliant to the strains $\boldsymbol{\sigma}_{j,k_1} + \boldsymbol{\sigma}_{j,k_2}$ for all k_1 and k_2 with $k_2 > k_1$.

Finally, consider those junction regions where three or more walls meet, i.e., for some k_1, $k_2 > k_1$, and $k_3 > k_2$, $\boldsymbol{x}$ is such that $\eta_{k_i}(\boldsymbol{x}) = 1$ for $i = 1, 2, 3$. For a given choice of k_1, $k_2 > k_1$, and $k_3 > k_2$ such that the three vectors $\boldsymbol{n}_{k_1}$, $\boldsymbol{n}_{k_2}$, and $\boldsymbol{n}_{k_3}$ are not coplanar, an upper bound for the volume fraction occupied by this region is ϵ^3. In the case that the three vectors $\boldsymbol{n}_{k_1}$, $\boldsymbol{n}_{k_2}$, and $\boldsymbol{n}_{k_3}$ are coplanar, we can ensure that the volume fraction occupied by this region is ϵ^3 or less by appropriately translating one or two walled structures, i.e., by replacing $\eta_{k_m}(\boldsymbol{x})$ with $\eta_{k_m}(\boldsymbol{x} + \alpha_i \boldsymbol{n}_{k_m})$ for $m = 2, 3$, for an appropriate choice of α_2 and α_3 between 0 and 1. Since the energy density of the trial field in these regions scales as $\epsilon^3/\epsilon^2 = \epsilon$, we can ignore this contribution in the limit $\epsilon \to 0$ as it goes to zero too.

From this analysis of the energy densities associated with the trial fields it follows that one does not necessarily need the pentamode, quadramode, trimode, bimode, and unimode materials as appropriate for the material inside the walled structure. Instead, by modifying the construction, it suffices to use only pentamode and quadramode materials. In the walled structure we now put pentamode materials in those sections where for some k, we have $\eta_k(\boldsymbol{x}) = 1$ while $\eta_{k'}(\boldsymbol{x}) = 0$ for all $k' \neq k$. Each pentamode material is very stiff to the single stress $\boldsymbol{\sigma}^{(k)}$ appropriate to the wall under consideration. In each junction region of the walled structure where $\eta_{k_1}(\boldsymbol{x}) = \eta_{k_2}(\boldsymbol{x}) = 1$ for some $k_1 \neq k_2$ while $\eta_k(\boldsymbol{x}) = 0$ for all k not equal to k_1 or k_2, we put a quadramode material which is very stiff to any stress in the subspace spanned by $\boldsymbol{\sigma}^{(k_1)}$ and $\boldsymbol{\sigma}^{(k_2)}$ as appropriate to the junction region under consideration. In the remaining junction regions of the walled structure (where three or more walls intersect) we put phase 1. The contribution to the average energy of the fields in these regions vanishes as $\epsilon \to 0$ as discussed above.

By these constructions we effectively obtain materials with elasticity tensors $\boldsymbol{C}_*$ such that

$$\lim_{\delta \to \infty} (\boldsymbol{C}_*)^{-1} = \Pi_p (\boldsymbol{C}_f^A)^{-1} \Pi_p, \tag{3-13}$$

where $\boldsymbol{I}$ is the fourth-order identity matrix, Π_p is the fourth-order tensor that is the projection onto the space $\mathcal{V}_p$, and $\boldsymbol{C}_f^A$ is the relevant Avellaneda material. In the basis (3-5), Π_p is represented by the 6×6 matrix that has the block form

$$\Pi_p = \begin{pmatrix} 0 & 0 \\ 0 & \boldsymbol{I}_p \end{pmatrix}, \tag{3-14}$$

where $\boldsymbol{I}_p$ represents the $p \times p$ identity matrix and the 0 on the diagonal represents the $(6-p) \times (6-p)$ zero matrix.

In the case $d = 2$ the analysis simplifies in the obvious way. We have the inequalities

$$\begin{aligned} 0 &\le W_f^0(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0), \\ \boldsymbol{\epsilon}_1^0 : [\boldsymbol{C}_f^A(0, 0, \boldsymbol{\epsilon}_1^0)]\boldsymbol{\epsilon}_1^0 &\le W_f^1(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\epsilon}_1^0), \\ \sum_{i=1}^{2} \boldsymbol{\epsilon}_i^0 : \boldsymbol{C}_f^A(0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0)\boldsymbol{\epsilon}_i^0 &\le W_f^2(\boldsymbol{\sigma}_1^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0), \end{aligned} \tag{3-15}$$

the first one of which is sharp in the limit $\delta \to \infty$ being attained when the material consists of islands of phase 1 surrounded by phase 2. The recipe for showing that the bound (3-15) on $W_f^1(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\epsilon}_1^0)$ is sharp for certain values of $\boldsymbol{\sigma}_1^0$ and $\boldsymbol{\sigma}_2^0$ and that the bound (2-5) on $W_f^2(\boldsymbol{\sigma}_1^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0)$ is sharp for certain values of $\boldsymbol{\sigma}_1^0$ is almost exactly the same as in the 3-dimensional case: insert into the Avellaneda material a thin walled structure of unimode and bimode materials, respectively, so that it is very stiff to any stress in the space spanned by $\boldsymbol{\sigma}_1^0$ and $\boldsymbol{\sigma}_2^0$ in the case of W_f^1, or so that it is very stiff to the stress $\boldsymbol{\sigma}_1^0$ in the case of W_f^2.

4. The algebraic problem: characterizing those symmetric matrix pencils spanned by zero determinant matrices

Now we are interested in the following question: *Given k linearly independent symmetric $d \times d$ matrices $\boldsymbol{A}_1, \boldsymbol{A}_2, \dots, \boldsymbol{A}_k$, find necessary and sufficient conditions such that there exists linearly independent matrices $\{\boldsymbol{B}_i\}_{i=1}^k$ spanned by the basis elements $\boldsymbol{A}_i$ such that* $\det(\boldsymbol{B}_i) = 0$. *It is assumed that $d = 2$ or 3 and $1 \le k \le k_d$, where $k_2 = 2$ and $k_3 = 5$.* Here we are working in the generic situation, i.e., we prove the algebraic result for a dense set of matrices. The continuity result of Section 5 will allow us to conclude for the whole set of matrices. Actually, the proof below also shows that the algebraic result holds for the complement of a zero measure set of matrices.

Theorem 4.1. *The above problem is solvable if and only if the matrices $\boldsymbol{A}_i$ for $i = 1, \dots, k$ satisfy the following conditions*:

(i) $$\det(\boldsymbol{A}_1) = 0, \quad \textit{if } k = 1,\ d = 2, 3. \tag{4-1}$$

(ii) $$(\alpha_1\gamma_2+\alpha_2\gamma_1-2\beta_1\beta_2)^2 > 4\det(\boldsymbol{A}_1)\det(\boldsymbol{A}_2), \quad \textit{if } k=d=2, \tag{4-2}$$

where
$$\boldsymbol{A}_i = \begin{pmatrix}\alpha_i & \beta_i\\ \beta_i & \gamma_i\end{pmatrix}. \tag{4-3}$$

(iii) $$\begin{aligned}\triangle = 18\det(\boldsymbol{A}_1)\det(\boldsymbol{A}_2)S_1S_2-4S_1^3\det(\boldsymbol{A}_2)+S_1^2S_2^2-4S_2^3\det(\boldsymbol{A}_1)\\ -27\det(\boldsymbol{A}_1)^2\det(\boldsymbol{A}_2)^2>0, \quad \textit{if } k=2,\ d=3,\end{aligned} \tag{4-4}$$

where $S_i=\sum_{j=1}^3 s_{ij}$ *for* $i=1,2$ *and* s_{ij} *is the determinant of the matrix obtained by replacing the* j*-th row of* $\boldsymbol{A}_i$ *by the* j*-th row of* $\boldsymbol{A}_{i+1}$*, where by convention we have* $\boldsymbol{A}_3=\boldsymbol{A}_1$.

(iv) $$\textit{Always solvable if } k\ge 3,\ d=3. \tag{4-5}$$

Remark. In fact, the condition (4-1), that $\det(\boldsymbol{A}_1)=0$, could be excluded since we are considering the generic case. It is inserted because we can treat it explicitly.

Proof. We consider all the cases separately.

Case (i): $k=1$. In this case one must obviously have $\det(\boldsymbol{A}_1)=0$.

Case (ii): $k=2$, $d=2$. We can without loss of generality assume that (by small perturbations) $\det(\boldsymbol{A}_i)\neq 0$ for $i=1,2$. For $\eta,\mu\in\mathbb{R}^2$, denote $\boldsymbol{A}(\eta,\mu)=\eta\boldsymbol{A}_1+\mu\boldsymbol{A}_2$, and thus for the equality

$$\det(\boldsymbol{A}(\eta,\mu))=\det(\boldsymbol{A}_1)\eta^2+(\alpha_1\gamma_2+\alpha_2\gamma_1-2\beta_1\beta_2)\eta\mu+\det(\boldsymbol{A}_2)\mu^2 \tag{4-6}$$

to happen, one must first of all have $\mu\neq 0$; thus, dividing by μ^2 and setting $t=\eta/\mu$, we get that the quadratic equation

$$\frac{1}{\mu^2}\det(\boldsymbol{A}(\eta,\mu))=\det(\boldsymbol{A}_1)t^2+(\alpha_1\gamma_2+\alpha_2\gamma_1-2\beta_1\beta_2)t+\det(\boldsymbol{A}_2)=0 \tag{4-7}$$

must have two different solutions, i.e., the discriminant is strictly positive, which amounts to exactly (4-2).

Case (iii): $k=2$, $d=3$. Again, we can without loss of generality assume that $\det(\boldsymbol{A}_i)\neq 0$ for $i=1,2$. Set again $\boldsymbol{A}(\eta,\mu)=\eta\boldsymbol{A}_1+\mu\boldsymbol{A}_2$; thus we must have that the equation

$$\det(\boldsymbol{A}(\eta,\mu))=\det(\boldsymbol{A}_1)\eta^3+S_1\eta^2\mu+S_2\eta\mu^2+\det(\boldsymbol{A}_2)\mu^3=0 \tag{4-8}$$

has at least two different real roots, which by Cardan's condition gives

$$\begin{aligned}\triangle = 18\det(\boldsymbol{A}_1)\det(\boldsymbol{A}_2)S_1S_2-4S_1^3\det(\boldsymbol{A}_2)+S_1^2S_2^2\\ -4S_2^3\det(\boldsymbol{A}_1)-27\det(\boldsymbol{A}_1)^2\det(\boldsymbol{A}_2)^2>0,\end{aligned} \tag{4-9}$$

which is exactly (4-4).

Case (iv): $k \geq 3$, $d = 3$. Let us consider the case $k = 3$ first. Let us show that we can assume, without loss of generality, that $\det(\boldsymbol{A}_1) = \det(\boldsymbol{A}_2) = 0$, by proving that there exist numbers $\eta_i \neq 0$ for $i = 1, 2$ such that the matrices $\boldsymbol{B}_1 = \eta_1 \boldsymbol{A}_1 + \boldsymbol{A}_2$ and $\boldsymbol{B}_2 = \eta_2 \boldsymbol{A}_1 + \boldsymbol{A}_3$ have zero determinant. Indeed, we assume without loss of generality that $\det(\boldsymbol{A}_i) \neq 0$ for $i = 1, 2, 3$. We would then like to have

$$\det(\boldsymbol{B}_1)(\eta_1) = \eta_1^3 \det(\boldsymbol{A}_1) + \eta_1^2(\,\cdot\,) + \eta_1(\,\cdot\,) + \det(\boldsymbol{A}_2) = 0, \tag{4-10}$$

which has a nonzero root η_1, being a cubic equation with $\det(\boldsymbol{B}_1)(0) = \det(\boldsymbol{A}_2) \neq 0$. Similarly, the equation $\det(\boldsymbol{B}_2)(\eta_2) = 0$ has a nonzero solution η_2. The matrices $\boldsymbol{B}_1$, $\boldsymbol{B}_2$, and $\boldsymbol{A}_1$ are linearly independent, because the linear independence of $\boldsymbol{B}_1$, $\boldsymbol{B}_2$, and $\boldsymbol{A}_1$ is equivalent to the condition

$$\det \begin{pmatrix} \eta_1 & 1 & 0 \\ \eta_2 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = 1 \neq 0. \tag{4-11}$$

Assume now that $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, and $\boldsymbol{A}_3$ are linearly independent and

$$\det(\boldsymbol{A}_1) = \det(\boldsymbol{A}_2) = 0. \tag{4-12}$$

For any $\eta, \mu \in \mathbb{R}$, consider the matrix

$$\boldsymbol{B}_3 = \boldsymbol{B}(\eta, \mu) = \boldsymbol{A}_3 + \eta \boldsymbol{A}_1 + \mu \boldsymbol{A}_2.$$

It is clear that the triple $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, $\boldsymbol{B}_3$ is linearly independent, so we would like to show that there exist $\eta, \mu \in \mathbb{R}$, such that $\det(\boldsymbol{B}_3) = 0$. Assume, by contradiction, that

$$\det(\boldsymbol{B}_3) \neq 0, \quad \text{for all } \eta, \mu \in \mathbb{R}. \tag{4-13}$$

Let us then show that the condition (4-13) implies that $c_1 = c_2 = 0$, where, taking into account the condition (4-12), we have that

$$\det(\boldsymbol{B}_3) = c_1\eta^2\mu + c_2\eta\mu^2 + c_3\eta\mu + c_4\eta^2 + c_5\mu^2 + c_6\eta + c_7\mu + \det(\boldsymbol{A}_3). \tag{4-14}$$

Indeed, if $c_1 \neq 0$, then taking $\eta = \mu^2$ we get that the equation $\det(\boldsymbol{B}(\mu^2, \mu)) = 0$ would have a solution $\mu \in \mathbb{R}$, being a fifth-order equation; thus, we get $c_1 = c_2 = 0$. Next, by perturbing the elements of $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ if necessary, we can reach the situation where no entries and second-order minors of both $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ vanish, by first reaching the situation where $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ have no zero entries. If we now perturb any ij and ik elements of $\boldsymbol{A}_1$ by small numbers ϵ and δ, where $j \neq k$, then to keep the condition $\det(\boldsymbol{A}_1) = 0$, we must have ϵ and δ satisfying

$$\epsilon \cdot \operatorname{cof}_{ij}(\boldsymbol{A}_1) + \delta \cdot \operatorname{cof}_{ik}(\boldsymbol{A}_1) = 0. \tag{4-15}$$

On the other hand, the condition $c_2 = 0$ must not be violated by that perturbation, thus we must have as well

$$\epsilon \cdot \mathrm{cof}_{ij}(\boldsymbol{A}_2) + \delta \cdot \mathrm{cof}_{ik}(\boldsymbol{A}_2) = 0. \tag{4-16}$$

The last two conditions then imply that the cofactor matrix $\mathrm{cof}\,\boldsymbol{A}_1$ is a multiple of the cofactor matrix $\mathrm{cof}\,\boldsymbol{A}_2$, i.e.,

$$\mathrm{cof}(\boldsymbol{A}_2) = t \cdot \mathrm{cof}(\boldsymbol{A}_1), \quad t \neq 0. \tag{4-17}$$

Again, a small perturbation of the 11 and 12 elements of $\boldsymbol{A}_1$ by ϵ and δ satisfying (4-15) with $i = j = 1, k = 2$ does not violate the condition $\det(\boldsymbol{A}_1) = 0$, thus it must not violate the condition (4-16). Observe that the above perturbation does not change the cofactor $\mathrm{cof}_{11}(\boldsymbol{A}_1)$, but it changes the cofactor element $\mathrm{cof}_{33}(\boldsymbol{A}_1)$, which means that the desired condition $\det(\boldsymbol{B}_3) = 0$ can be reached by small perturbations. The case $k = d = 3$ is now done.

Assume now $k \geq 4$ and $d = 3$. By the previous step, in the space spanned by $\boldsymbol{A}_1$, $\boldsymbol{A}_2$, and $\boldsymbol{A}_3$ there are three matrices $\boldsymbol{A}'_1$, $\boldsymbol{A}'_2$, and $\boldsymbol{B}_3 = \boldsymbol{A}_3 + \eta_3 \boldsymbol{A}'_1 + \mu_3 \boldsymbol{A}'_2$ that are linearly independent with zero determinant. Then, again by the previous step, we can find linearly independent matrices $\boldsymbol{B}_1, \dots, \boldsymbol{B}_k$ that have the form $\boldsymbol{B}_1 = \boldsymbol{A}'_1$, $\boldsymbol{B}_2 = \boldsymbol{A}'_2$, and $\boldsymbol{B}_i = \boldsymbol{A}_i + \eta_i \boldsymbol{A}'_1 + \mu_i \boldsymbol{A}'_2$ for $3 \leq i \leq k$ and that are linearly independent and have zero determinant. $\square$

5. Continuity of the energy functions

It follows from the preceding analysis that we can determine the three energy functions

$$W_f^1(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0, \boldsymbol{\sigma}_5^0, \boldsymbol{\epsilon}_1^0),$$
$$W_f^2(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0),$$
$$W_f^3(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0)$$

in the limit $\delta \to \infty$ for almost all combinations of applied fields. Here we establish that these energy functions are continuous functions of the applied fields in the limit $\delta \to \infty$, and therefore we obtain expressions for the energy functions for all combinations of applied fields in this limit.

Recall that the set GU_f is characterized by its W-transform. For example, part of it is described by the function

$$W_f^2(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0) = \min_{\boldsymbol{C}_* \in GU_f} \left[\sum_{i=1}^{2} \boldsymbol{\epsilon}_i^0 : \boldsymbol{C}_* \boldsymbol{\epsilon}_i^0 + \sum_{j=1}^{4} \boldsymbol{\sigma}_j^0 : \boldsymbol{C}_*^{-1} \boldsymbol{\sigma}_j^0 \right]. \tag{5-1}$$

Here we want to show that such energy functions are continuous in their arguments. Let the compliance tensor $[\boldsymbol{C}_*(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0)]^{-1}$ be a minimizer of (5-1), and suppose we perturb the applied stress fields $\boldsymbol{\sigma}_j^0$ by $\delta\boldsymbol{\sigma}_j^0$ and the applied strain fields $\boldsymbol{\epsilon}_i^0$ by $\delta\boldsymbol{\epsilon}_i^0$. Now consider the walled material with a geometry described by the characteristic function

$$\chi_w(\boldsymbol{x}) = \prod_{k=1}^{3}(1 - H_{\epsilon'}(\boldsymbol{x}\cdot\boldsymbol{n}_k)), \tag{5-2}$$

where $\boldsymbol{n}_1$, $\boldsymbol{n}_2$, and $\boldsymbol{n}_3$ are the three orthogonal unit vectors

$$\boldsymbol{n}_1 = \begin{pmatrix}1\\0\\0\end{pmatrix}, \quad \boldsymbol{n}_2 = \begin{pmatrix}0\\1\\0\end{pmatrix}, \quad \boldsymbol{n}_3 = \begin{pmatrix}0\\0\\1\end{pmatrix}, \tag{5-3}$$

and ϵ' is a small parameter that gives the thickness of the walls. Inside the walls, where $\chi_w(\boldsymbol{x}) = 0$, we put an isotropic composite of phase 1 and phase 2, mixed in the proportions f and $1-f$ with isotropic effective elasticity tensor $\boldsymbol{C}(\kappa_0, \mu_0)$, where κ_0 is the effective bulk modulus and μ_0 is the effective shear modulus, which are assumed to have finite limits as $\delta \to \infty$. (The isotropic composite could consist of islands of void surrounded by phase 1.) Outside the walls, where $\chi_w(\boldsymbol{x}) = 1$, we put the material that has effective compliance tensor

$$[\boldsymbol{C}_*^1]^{-1} = [\boldsymbol{C}_*(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0)]^{-1}.$$

Let $\boldsymbol{C}_*'$ be the effective tensor of the composite. We have the variational principle

$$\begin{aligned}
&\sum_{i=1}^{2}(\boldsymbol{\epsilon}_i^0 + \delta\boldsymbol{\epsilon}_i^0) : \boldsymbol{C}_*'(\boldsymbol{\epsilon}_i^0 + \delta\boldsymbol{\epsilon}_i^0) + \sum_{j=1}^{4}(\boldsymbol{\sigma}_j^0 + \delta\boldsymbol{\sigma}_j^0) : (\boldsymbol{C}_*')^{-1}(\boldsymbol{\sigma}_j^0 + \delta\boldsymbol{\sigma}_j^0) \\
&\quad = \min_{\underline{\boldsymbol{\epsilon}}_1,\underline{\boldsymbol{\epsilon}}_2,\underline{\boldsymbol{\epsilon}}_3,\underline{\boldsymbol{\epsilon}}_4,\underline{\boldsymbol{\sigma}}_1,\underline{\boldsymbol{\sigma}}_2} \Big\langle \sum_{i=1}^{2} \underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x}) : [\chi_w(\boldsymbol{x})\boldsymbol{C}_*^1 + (1-\chi_w(\boldsymbol{x}))\boldsymbol{C}(\kappa_0, \mu_0)]\underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x}) \\
&\qquad\qquad + \sum_{j=1}^{4} \underline{\boldsymbol{\sigma}}_j(\boldsymbol{x}) : [\chi_w(\boldsymbol{x})\boldsymbol{C}_*^1 + (1-\chi_w(\boldsymbol{x}))\boldsymbol{C}(\kappa_0, \mu_0)]^{-1}\underline{\boldsymbol{\sigma}}_j(\boldsymbol{x}) \Big\rangle,
\end{aligned} \tag{5-4}$$

where the minimum is over fields subject to the appropriate average values and differential constraints. We choose constant trial stress fields

$$\underline{\boldsymbol{\sigma}}_j(\boldsymbol{x}) = \boldsymbol{\sigma}_j^0 + \delta\boldsymbol{\sigma}_j^0, \quad j = 1, 2, 3, 4, \tag{5-5}$$

and trial strain fields

$$\underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x}) = \boldsymbol{\epsilon}_i^0 + \delta\underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x}), \quad i = 1, 2, \tag{5-6}$$

where $\delta\underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x})$ has average value $\delta\boldsymbol{\epsilon}_i^0$ and is concentrated in the walls. Specifically, if $\{\delta\boldsymbol{\epsilon}_i^0\}_{k\ell}$ denote the matrix elements of $\delta\boldsymbol{\epsilon}_i^0$, and letting

$$\delta\boldsymbol{\epsilon}_i^1 = \begin{pmatrix} \{\delta\boldsymbol{\epsilon}_i^0\}_{11} & \{\delta\boldsymbol{\epsilon}_i^0\}_{12} & 0 \\ \{\delta\boldsymbol{\epsilon}_i^0\}_{21} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\delta\boldsymbol{\epsilon}_j^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \{\delta\boldsymbol{\epsilon}_i^0\}_{22} & \{\delta\boldsymbol{\epsilon}_i^0\}_{23} \\ 0 & \{\delta\boldsymbol{\epsilon}_i^0\}_{32} & 0 \end{pmatrix}, \tag{5-7}$$

$$\delta\boldsymbol{\epsilon}_j^3 = \begin{pmatrix} 0 & 0 & \{\delta\boldsymbol{\epsilon}_i^0\}_{13} \\ 0 & 0 & 0 \\ \{\delta\boldsymbol{\epsilon}_i^0\}_{31} & 0 & \{\delta\boldsymbol{\epsilon}_i^0\}_{33} \end{pmatrix},$$

then we choose

$$\delta\underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x}) = \sum_{k=1}^{3} \delta\boldsymbol{\epsilon}_i^k H_{\epsilon'}(\boldsymbol{x}\cdot\boldsymbol{n}_k)/\epsilon', \tag{5-8}$$

which has the required average value $\delta\boldsymbol{\sigma}_j^0$ and satisfies the differential constraints appropriate to a strain field because $\delta\boldsymbol{\epsilon}_i^k = \boldsymbol{a}_{i,k}\boldsymbol{n}_k^T + \boldsymbol{n}_k\boldsymbol{a}_{i,k}^T$ for some vector $\boldsymbol{a}_{i,k}$.

Hence, there exist constants α and β such that for sufficiently small ϵ' and for sufficiently small variations $\delta\boldsymbol{\sigma}_j^0$ and $\delta\boldsymbol{\epsilon}_i^0$ in the applied fields, we have

$$\begin{aligned}\Big\langle \sum_{i=1}^{2} &\underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x}) : [\chi_w(\boldsymbol{x})\boldsymbol{C}_*^1 + (1-\chi_w(\boldsymbol{x}))\boldsymbol{C}(\kappa_0,\mu_0)]\underline{\boldsymbol{\epsilon}}_i(\boldsymbol{x}) \\ &+ \sum_{j=1}^{4} \underline{\boldsymbol{\sigma}}_j(\boldsymbol{x}) : [\chi_w(\boldsymbol{x})\boldsymbol{C}_*^1 + (1-\chi_w(\boldsymbol{x}))\boldsymbol{C}(\kappa_0,\mu_0)]^{-1}\underline{\boldsymbol{\sigma}}_j(\boldsymbol{x})\Big\rangle \\ &\le W_f^2(\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0,\boldsymbol{\sigma}_3^0,\boldsymbol{\sigma}_4^0,\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0) + \alpha\epsilon' + \beta K/\epsilon',\end{aligned} \tag{5-9}$$

where K represents the norm

$$K = \sqrt{\sum_{i=1}^{2}\delta\boldsymbol{\epsilon}_i^0 : \delta\boldsymbol{\epsilon}_i^0 + \sum_{j=1}^{4}\delta\boldsymbol{\sigma}_j^0 : \delta\boldsymbol{\sigma}_j^0} \tag{5-10}$$

of the field variations. Choosing $\epsilon' = \sqrt{\beta K/\alpha}$ to minimize the right-hand side of (5-9), we obtain

$$\begin{aligned}W_f^2(\boldsymbol{\sigma}_1^0+\delta\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0+\delta\boldsymbol{\sigma}_2^0,\boldsymbol{\sigma}_3^0+\delta\boldsymbol{\sigma}_3^0,\boldsymbol{\sigma}_4^0+\delta\boldsymbol{\sigma}_4^0,\boldsymbol{\epsilon}_1^0+\delta\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0+\delta\boldsymbol{\epsilon}_2^0) \\ \le W_f^2(\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0,\boldsymbol{\sigma}_3^0,\boldsymbol{\sigma}_4^0,\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0) + 2\sqrt{\alpha\beta K}.\end{aligned} \tag{5-11}$$

Clearly the right-hand side approaches $W_f^2(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0)$ as $K \to 0$. On the other hand, by repeating the same argument with the roles of

$$W_f^2(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0)$$

and

$$W_f^2(\boldsymbol{\sigma}_1^0 + \delta\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0 + \delta\boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0 + \delta\boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0 + \delta\boldsymbol{\sigma}_4^0, \boldsymbol{\epsilon}_1^0 + \delta\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0 + \delta\boldsymbol{\epsilon}_2^0)$$

reversed, and with the compliance tensor

$$[\boldsymbol{C}_*(\boldsymbol{\sigma}_1^0 + \delta\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0 + \delta\boldsymbol{\sigma}_2^0, \boldsymbol{\epsilon}_1^0 + \delta\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0 + \delta\boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0 + \delta\boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0 + \delta\boldsymbol{\epsilon}_4^0)]^{-1}$$

replacing the compliance tensor

$$[\boldsymbol{C}_*(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0)]^{-1},$$

we deduce that

$$\begin{aligned} &W_f^2(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0) \\ &\quad \le W_f^2(\boldsymbol{\sigma}_1^0 + \delta\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0 + \delta\boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0 + \delta\boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0 + \delta\boldsymbol{\sigma}_4^0, \boldsymbol{\epsilon}_1^0 + \delta\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0 + \delta\boldsymbol{\epsilon}_2^0) \\ &\qquad\qquad + 2\sqrt{\alpha\beta K}. \end{aligned} \tag{5-12}$$

This, together with (5-11), establishes the continuity of $W_f^2(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\sigma}_3^0, \boldsymbol{\sigma}_4^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0)$. The continuity of the other energy functions follows by the same argument.

6. Conclusion

We have established the following theorems.

Theorem 6.1. *Consider composites in three dimensions of two materials with positive definite elasticity tensors $\boldsymbol{C}_1$ and $\boldsymbol{C}_2 = \delta\boldsymbol{C}_0$ mixed in proportions f and $1-f$. Let the seven energy functions W_f^k, for $k = 0, 1, \dots, 6$, that characterize the set GU_f (with $U = (\boldsymbol{C}_1, \delta\boldsymbol{C}_0)$) of possible elastic tensors be defined by (2-1). These energy functions involve a set of applied strains $\boldsymbol{\epsilon}_i^0$ and applied stresses $\boldsymbol{\sigma}_j^0$ meeting the orthogonality condition (2-2). The energy function W_f^6 is given by*

$$W_f^6(\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0, \boldsymbol{\epsilon}_5^0, \boldsymbol{\epsilon}_6^0) = \sum_{i=1}^{6} \boldsymbol{\epsilon}_i^0 : \boldsymbol{C}_f^A(\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0, \boldsymbol{\epsilon}_5^0, \boldsymbol{\epsilon}_6^0)\boldsymbol{\epsilon}_i^0 \tag{6-1}$$

(as established by Avellaneda [1987]), *where $\boldsymbol{C}_f^A(\boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0, \boldsymbol{\epsilon}_3^0, \boldsymbol{\epsilon}_4^0, \boldsymbol{\epsilon}_5^0, \boldsymbol{\epsilon}_6^0)$ is the effective elasticity tensor of an Avellaneda material that is a sequentially layered laminate with the minimum value of the sum of elastic energies*

$$\sum_{i=1}^{6} \boldsymbol{\epsilon}_j^0 : \boldsymbol{C}_*\boldsymbol{\epsilon}_j^0. \tag{6-2}$$

Again some of the applied stresses $\boldsymbol{\sigma}_j^0$ *or applied strains* $\boldsymbol{\epsilon}_i^0$ *could be zero. Additionally, we have*

$$\begin{aligned}
&\lim_{\delta\to\infty} W_f^0(\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0,\boldsymbol{\sigma}_3^0,\boldsymbol{\sigma}_4^0,\boldsymbol{\sigma}_5^0,\boldsymbol{\sigma}_6^0)=0,\\
&\lim_{\delta\to\infty} W_f^1(\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0,\boldsymbol{\sigma}_3^0,\boldsymbol{\sigma}_4^0,\boldsymbol{\sigma}_5^0,\boldsymbol{\epsilon}_1^0)=\boldsymbol{\epsilon}_1^0:[\boldsymbol{C}_f^A(0,0,0,0,0,\boldsymbol{\epsilon}_1^0)]\boldsymbol{\epsilon}_1^0,\\
&\lim_{\delta\to\infty} W_f^2(\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0,\boldsymbol{\sigma}_3^0,\boldsymbol{\sigma}_4^0,\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0)=\sum_{i=1}^{2}\boldsymbol{\epsilon}_i^0:[\boldsymbol{C}_f^A(0,0,0,0,\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0)]\boldsymbol{\epsilon}_i^0,\\
&\lim_{\delta\to\infty} W_f^3(\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0,\boldsymbol{\sigma}_3^0,\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0,\boldsymbol{\epsilon}_3^0)=\sum_{i=1}^{3}\boldsymbol{\epsilon}_i^0:[\boldsymbol{C}_f^A(0,0,0,\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0,\boldsymbol{\epsilon}_3^0)]\boldsymbol{\epsilon}_i^0,
\end{aligned}\tag{6-3}$$

for all combinations of applied stresses $\boldsymbol{\sigma}_j^0$ *and applied strains* $\boldsymbol{\epsilon}_i^0$. *In the case that* $\det(\boldsymbol{\sigma}_1^0)=0$, *we have*

$$\lim_{\delta\to\infty} W_f^5(\boldsymbol{\sigma}_1^0,\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0,\boldsymbol{\epsilon}_3^0,\boldsymbol{\epsilon}_4^0,\boldsymbol{\epsilon}_5^0)=\sum_{i=1}^{5}\boldsymbol{\epsilon}_i^0:[\boldsymbol{C}_f^A(0,\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0,\boldsymbol{\epsilon}_3^0,\boldsymbol{\epsilon}_4^0,\boldsymbol{\epsilon}_5^0)]\boldsymbol{\epsilon}_i^0,\tag{6-4}$$

while, when $f(t)=\det(\boldsymbol{\sigma}_1^0+t\boldsymbol{\sigma}_2^0)$ *has at least two roots (the condition for which is given by* (4-4)),

$$\lim_{\delta\to\infty} W_f^4(\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0,\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0,\boldsymbol{\epsilon}_3^0,\boldsymbol{\epsilon}_4^0)=\sum_{i=1}^{4}\boldsymbol{\epsilon}_i^0:[\boldsymbol{C}_f^A(0,0,\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0,\boldsymbol{\epsilon}_3^0,\boldsymbol{\epsilon}_4^0)]\boldsymbol{\epsilon}_i^0.\tag{6-5}$$

Theorem 6.2. *For* 2*-dimensional composites, the four energy functions* W_f^k, *for* $k=0,1,2,3$, *are defined by* (2-5), *and these characterize the set* GU_f, *with* $U=(\boldsymbol{C}_1,\delta\boldsymbol{C}_0)$, *of possible elastic tensors* $\boldsymbol{C}_*$ *of composites of two phases with positive definite elasticity tensors* $\boldsymbol{C}_1$ *and* $\boldsymbol{C}_2=\delta\boldsymbol{C}_0$. *These energy functions involve a set of applied strains* $\boldsymbol{\epsilon}_i^0$ *and applied stresses* $\boldsymbol{\sigma}_j^0$ *meeting the orthogonality condition* (2-2). *The energy function* W_f^3 *is given by*

$$W_f^3(\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0,\boldsymbol{\epsilon}_3^0)=\sum_{i=1}^{3}\boldsymbol{\epsilon}_i^0:\boldsymbol{C}_f^A(\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0,\boldsymbol{\epsilon}_3^0)\boldsymbol{\epsilon}_i^0\tag{6-6}$$

(as proved by Avellaneda [1987]), *where* $\boldsymbol{C}_f^A(\boldsymbol{\epsilon}_1^0,\boldsymbol{\epsilon}_2^0,\boldsymbol{\epsilon}_3^0)$ *is the effective elasticity tensor of an Avellaneda material that is a sequentially layered laminate with the minimum value of the sum of elastic energies*

$$\sum_{j=1}^{3}\boldsymbol{\epsilon}_j^0:\boldsymbol{C}_*\boldsymbol{\epsilon}_j^0.\tag{6-7}$$

We also have the trivial result that

$$\lim_{\delta\to\infty} W_f^0(\boldsymbol{\sigma}_1^0,\boldsymbol{\sigma}_2^0,\boldsymbol{\sigma}_3^0)=0.\tag{6-8}$$

When $\det \boldsymbol{\sigma}_1^0 = 0$, *we have*

$$\lim_{\delta\to\infty} W_f^2(\boldsymbol{\sigma}_1^0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0) = \sum_{i=1}^{2} \boldsymbol{\epsilon}_i^0 : [\boldsymbol{C}_f^A(0, \boldsymbol{\epsilon}_1^0, \boldsymbol{\epsilon}_2^0)]\boldsymbol{\epsilon}_i^0, \tag{6-9}$$

while when $f(t) = \det(\boldsymbol{\sigma}_1^0 + t\boldsymbol{\sigma}_2^0)$ *has exactly two roots* (*the condition for which is given by* (4-2)),

$$\lim_{\delta\to\infty} W_f^1(\boldsymbol{\sigma}_1^0, \boldsymbol{\sigma}_2^0, \boldsymbol{\epsilon}_1^0) = \boldsymbol{\epsilon}_1^0 : [\boldsymbol{C}_f^A(0, 0, \boldsymbol{\epsilon}_1^0)]\boldsymbol{\epsilon}_1^0. \tag{6-10}$$

These theorems, and the accompanying microstructures, help define what sort of elastic behaviors are theoretically possible in 2- and 3-dimensional materials consisting of a very stiff phase and an elastic phase (possibly anisotropic, but with fixed orientation). They should serve as benchmarks for the construction of more realistic microstructures that can be manufactured. We have found the minimum over all microstructures of various sums of energies and complementary energies.

It remains an open problem to find expressions for the energy functions in the cases not covered by these theorems. Notice that for 3-dimensional composites the function W_f^5 is only determined when the special condition $\det(\boldsymbol{\sigma}_1^0) = 0$ is satisfied exactly. Similarly, for 2-dimensional composites the function W_f^2 is only determined when the special condition $\det \boldsymbol{\sigma}_1^0 = 0$ is satisfied exactly. Thus these functions are only known on a set of zero measure.

Even for an isotropic composite with a bulk modulus κ_* and a shear modulus μ_*, the set of all possible pairs (κ_*, μ_*) is still not completely characterized either in the limit $\delta \to \infty$. In these limits the bounds of Berryman and Milton [1988] and Cherkaev and Gibiansky [1993] decouple and provide no extra information beyond that provided by the Hashin–Shtrikman–Hill bounds [Hashin and Shtrikman 1963; Hashin 1965; Hill 1963; 1964]. While the results of this paper show that in the limit $\delta \to \infty$ one can obtain 3-dimensional structures attaining the Hashin–Shtrikman–Hill lower bound on κ_*, while having $\mu_* = \infty$, it is not clear what the minimum value for μ_* is, given that $\kappa_* = \infty$, nor is it clear in two dimensions what the minimum value of κ_* is when $\mu_* = \infty$.

Acknowledgements

The authors thank the National Science Foundation for support through grant DMS-1211359. M. Briane wishes to thank the Department of Mathematics of the University of Utah for his stay during March 25–April 3, 2016.

References

[Avellaneda 1987] M. Avellaneda, "Optimal bounds and microgeometries for elastic two-phase composites", *SIAM J. Appl. Math.* **47**:6 (1987), 1216–1228.

[Avellaneda and Milton 1989] M. Avellaneda and G. W. Milton, "Bounds on the effective elasticity tensor of composites based on two-point correlations", pp. 89–93 in *Composite material technology, 1989: presented at the Twelfth Annual Energy-Sources Technology Conference and Exhibition* (Houston, 1989), American Society of Mechanical Engineers, Petroleum Division **24**, American Society of Mechanical Engineers, New York, 1989.

[Berryman and Milton 1988] J. G. Berryman and G. W. Milton, "Microgeometry of random composites and porous media", *J. Phys. D* **21**:1 (1988), 87–94.

[Cherkaev and Gibiansky 1992] A. V. Cherkaev and L. V. Gibiansky, "The exact coupled bounds for effective tensors of electrical and magnetic properties of two-component two-dimensional composites", *Proc. Roy. Soc. Edinburgh Sect. A* **122**:1-2 (1992), 93–125.

[Cherkaev and Gibiansky 1993] A. V. Cherkaev and L. V. Gibiansky, "Coupled estimates for the bulk and shear moduli of a two-dimensional isotropic elastic composite", *J. Mech. Phys. Solids* **41**:5 (1993), 937–980.

[Francfort and Milton 1994] G. A. Francfort and G. W. Milton, "Sets of conductivity and elasticity tensors stable under lamination", *Comm. Pure Appl. Math.* **47**:3 (1994), 257–279.

[Francfort et al. 1995] G. Francfort, F. Murat, and L. Tartar, "Fourth-order moments of nonnegative measures on S^2 and applications", *Arch. Rational Mech. Anal.* **131**:4 (1995), 305–333.

[Hashin 1965] H. Hashin, "On elastic behavior of fibre reinforced materials of arbitrary transverse phase geometry", *J. Mech. Phys. Solids* **13**:3 (1965), 119–134.

[Hashin and Shtrikman 1963] Z. Hashin and S. Shtrikman, "A variational approach to the theory of the elastic behaviour of multiphase materials", *J. Mech. Phys. Solids* **11**:2 (1963), 127–140.

[Hill 1963] R. Hill, "Elastic properties of reinforced solids: some theoretical principles", *J. Mech. Phys. Solids* **11**:5 (1963), 357–372.

[Hill 1964] R. Hill, "Theory of mechanical properties of fibre-strengthened materials, I: Elastic behaviour", *J. Mech. Phys. Solids* **12** (1964), 199–212.

[Milton 1994] G. W. Milton, "A link between sets of tensors stable under lamination and quasiconvexity", *Comm. Pure Appl. Math.* **47**:7 (1994), 959–1003.

[Milton 2002] G. W. Milton, *The theory of composites*, Cambridge Monographs on Applied and Computational Mathematics **6**, Cambridge Univ. Press, 2002.

[Milton and Cherkaev 1995] G. W. Milton and A. V. Cherkaev, "Which elasticity tensors are realizable?", *J. Eng. Mater. Tech.* **117**:4 (1995), 483–493.

[Milton et al. 2017] G. W. Milton, M. Briane, and D. Harutyunyan, "On the possible effective elasticity tensors of 2-dimensional and 3-dimensional printed materials", *Math. Mech. Complex Syst.* **5**:1 (2017), 41–94.

Received 12 Jun 2016. Revised 11 Oct 2016. Accepted 14 Nov 2016.

GRAEME W. MILTON: milton@math.utah.edu
Department of Mathematics, University of Utah, 155 South 1400 East Room 233, Salt Lake City, UT 84112-0090, United States

DAVIT HARUTYUNYAN: davith@math.utah.edu
Department of Mathematics, University of Utah, 155 South 1400 East Room 233, Salt Lake City, UT 84112-0090, United States

MARC BRIANE: mbriane@insa-rennes.fr
Institut de Recherche Mathématique de Rennes, INSA de Rennes, 20 Avenue des Buttes de Coësmes, CS 70839, 35708 Rennes Cedex 7, France

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the submission page.

Originality. Submission of a manuscript acknowledges that the manuscript is original and and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles in MEMOCS are usually in English, but articles written in other languages are welcome.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and a Mathematics Subject Classification or a Physics and Astronomy Classification Scheme code for the article, and, for each author, postal address, affiliation (if appropriate), and email address if available. A home-page URL is optional.

Format. Authors are encouraged to use LaTeX and the standard amsart class, but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should normally be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages — Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc. — allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with as many details as you can about how your graphics were generated.

Bundle your figure files into a single archive (using zip, tar, rar or other format of your choice) and upload on the link you been provided at acceptance time. Each figure should be captioned and numbered so that it can float. Small figures occupying no more than three lines of vertical space can be kept in the text ("the curve looks like this:"). It is acceptable to submit a manuscript with all figures at the end, if their placement is specified in the text by means of comments such as "Place Figure 1 here". The same considerations apply to tables.

White Space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

MATHEMATICS AND MECHANICS OF COMPLEX SYSTEMS
vol. 5 no. 1 2017

MEMOCS is a journal of the International Research Center for the Mathematics and Mechanics of Complex Systems at the Università dell'Aquila, Italy.

$\mathcal{M} \cap \mathcal{M}$