vol. 6   no. 3   2018

MATHEMATICS AND MECHANICS
of
Complex Systems

MATHEMATICS AND MECHANICS OF COMPLEX SYSTEMS

msp.org/memocs

THE VARIATIONAL STRUCTURE
OF CLASSICAL PLASTICITY

GIANPIETRO DEL PIERO

A unified approach to classical plasticity, including metal plasticity, geomaterials, and crystal plasticity, is presented. A distinctive feature of this approach is that the basic constitutive elements (yield criterion, flow rule, consistency condition, and hardening rule), instead of being assumed on a phenomenological basis or deduced from ad hoc principles, are obtained directly from the stationarity of the energy. The plastic continuum is regarded as a particular micromorphic continuum, and its energy has the form resulting from a homogenization procedure introduced in the theory of structured deformations. This form of the energy requires an additive decomposition of the deformation gradient, in place of the multiplicative decomposition usually adopted in finite plasticity. It is shown by examples that many of the models adopted in classical plasticity can be obtained from ad hoc specifications of the energy.

1. Introduction

Plasticity is a branch of continuum mechanics characterized by the presence of a state variable, the plastic strain, which describes rearrangements of the material structure at the microscopic level. With the progress of microstructural multiscale theories, it became important to specify the nature of the continuum in which a plasticity model is embedded. For *classical plasticity* the underlying continuum is the *classical continuum*, that is, a continuum whose external power is produced by body forces and surface tractions alone. This excludes nonlocal models such as *gradient plasticity*, in which the plastic strain is supposed to produce an extra power when multiplied by microscopic external forces, and the latter produce an extra stress measure plus a hyperstress represented by a third-order tensor.[1] Rate-dependent theories, and in particular viscoplasticity, are also excluded from the present treatment.

---

[1]For this and other models of classical and nonclassical plasticity, see the book by Gurtin, Fried, and Anand [Gurtin et al. 2010, Part XV] and the references therein.

The bases of classical plasticity were laid down in the 1940s. Of the basic constitutive elements of the theory,

- *the yield condition*,
- *the flow rule*,
- *the consistency condition*, and
- *the hardening rule*,

the first three had been fixed by the end of that decade, and the fourth followed a few years later.[2] Though these elements take origin from experimental observation, several efforts were made to relate them to general principles, in order to show that plastic response is not a caprice of nature, but obeys a precise mathematical structure. In this spirit, Prager [1949] proved that the flow rule, in the form of a *normality law*, is a consequence of the uniqueness of the solution of the incremental equilibrium problem. In a paper published just before, Hill [1948] had proved that uniqueness is, in turn, a consequence of a *principle of maximum plastic work.*

Soon afterwards, Drucker [1952] showed that normality is also a consequence of a *quasithermodynamic postulate* of material stability. Actually this was not progress, since Drucker's postulate came out to be more restrictive than Hill's principle. Progress was also not brought by the postulate of Il'yushin [1961] which, though less restrictive in general, in the case of classical plasticity is equivalent to Drucker's [Lucchesi and Podio-Guidugli 1990]. On the contrary, some progress came with the introduction of supplementary state variables.[3] This opened the way to the study of *nonassociated plasticity*, which is in contradiction with Drucker's postulate, since by its own definition nonassociated plasticity does not obey the normality law.[4]

Several variational principles were formulated at the earlier stages of the theory.[5] In the development of such principles, a turning point was marked by the introduction of the concept of plastic potential.[6] In the broader context of classical continuum mechanics, related concepts of dissipation function and dissipation potential were introduced by Ziegler [1963] and Moreau [1970; 1974]. Incremental minimum principles involving a strain energy made of the sum of an elastic energy and a dissipation potential were formulated by Fedelich and Ehrlacher [1989]

---

[2]See Prager's overview [1949]. Prager's kinematical hardening model [1955] was introduced in the mid 1950s.

[3]See [Lemaitre and Chaboche 1990, p. 193] or [Ziegler 1983, §14].

[4]The interest in nonassociated plasticity was stimulated by the study of geomaterials, such as soils, concrete, and stones; see, e.g., [Vermeer and de Borst 1984].

[5]For the variational principles formulated before the 1950s, see Hill's book [1950].

[6]See [Hill 1950] for plasticity, and [Rice 1971] for viscoplasticity. Later, the existence of a special type of potential, called *maximal responsive*, was proved to be equivalent to Hill's principle; see the article by Eve, Reddy, and Rockafellar [Eve et al. 1990, Theorem 4.1].

and by Petryk [2003]. Subsequent contributions by Dal Maso et al. [2006; 2008], Mielke [Carstensen et al. 2002; Mielke 2003], and their schools marked substantial progress in this direction.

The variational approach adopted in the present paper, formally similar to those in [Fedelich and Ehrlacher 1989; Petryk 2003], has the peculiarity of deducing all constitutive elements from the stationarity condition on the energy functional. Indeed, this condition determines the incremental response law without any supplementary assumption, such as the existence or convexity of the elastic range, or the form of the flow rule.[7]

After fixing the incremental response law, the incremental equilibrium problem can be formulated. This is the problem of determining the small deformations from a given equilibrium placement, due to a conveniently small load increment. Though the deformations to be determined are small, it is sometimes convenient to formulate the problem in large deformations.[8] This is the case, for example, when one has in mind to approximate a problem with large load increments by a sequence of problems with small load increments.

The present study is restricted to stationarity, that is, to the condition of nonnegativeness of the first variation of the energy. In this way only equilibrium conditions are obtained, without any information about stability. This is a serious limitation. Indeed, a stability analysis would show that in classical plasticity a softening response is unstable, because the plastic strain localizes on arbitrarily small regions of the body. Initially, this led to considering softening materials as inadmissible. But this viewpoint, consecrated by Drucker's postulate, conflicted with the evidence of the softening response exhibited by many real materials. Later, it was realized that softening can be described by adding to the energy a nonlocal stabilizing term, depending on the gradient of the plastic strain rate [Aifantis 1984; Bažant et al. 1984].[9]

In this paper some preliminary definitions, including a new ad hoc notation for homogeneous maps, form Section 2, and the transitions from the equilibrium problem in finite deformation to the evolution problem and from this one to the incremental equilibrium problem are briefly illustrated in Section 3. This section deals with *two-scale*, or *micromorphic*, continua, of which the plastic continua

---

[7]The idea of deducing the constitutive properties from two scalar potentials, an elastic energy and a dissipation function, had already been exploited by Collins, Houlsby, and coworkers. They initially applied it to geomaterials [Collins and Houlsby 1997], and then to general rate-independent dissipative materials [Houlsby and Purzrin 2000; Collins 2003].

[8]The linearized equations for large deformations differ from those of the linear theory; see Section 3.2 below.

[9]More recent one-dimensional analyses of the softening response in the proximity of fracture can be found in [Pham et al. 2011a; 2011b] for damage and in [Del Piero 2013; Del Piero et al. 2013] for plasticity.

are a subclass. For them, in Section 3.4 it is shown that the stationarity of the energy determines the response law relating the Piola stress to the elastic part of the deformation, and produces a local stationarity condition which will be shown to fully characterize the constitutive response.

In Section 4 the incremental problem is reformulated as a minimum problem for the energy. The selected form of the energy is based on the additive decomposition of the deformation gradient into an elastic and a plastic part, a rather unusual choice in the context of large deformations. As shown in Section 4.1, this choice leads to a particular form of the plastic strain rate, in which a symmetric *plastic stretching* is followed by a rotation to be determined by a constitutive assumption. In most models this rotation is taken equal to the identity. An exception is the *crystal plasticity* model discussed in Section 7.[10] An indifference argument developed in Section 4.2 shows that the plastic part of the energy is independent of this rotation.

In the crucial Section 5, the constitutive elements of the theory are deduced from the local stationarity condition, now called *plastic stationarity condition*. A major result in this paper is that this condition determines a *bounding map* in the stress space, related to the directional derivatives of a dissipation potential. The values taken by this map in different directions impose directional limit values for the stress. This leads to the definition of an *elastic range*, a region in the stress space which by its own construction turns out to be closed and convex. From the plastic stationarity condition it also follows that a plastic strain rate can only occur if the stress is a boundary point of the elastic range, and that its direction belongs to the normal cone at that boundary point. This is the *normality law* which determines the *associated flow rule*. Thus, *nonassociated flow rules are not provided by the variational procedure*.

In Section 6 some well known plasticity models are reobtained assuming particular forms of the plastic energy. Section 6.1 deals with the three basic models of perfect, kinematic, and dilatational plasticity, which include kinematic and isotropic hardening as special cases. The assumption of isochoricity of the plastic strain rate is studied in Section 6.2. It is well known that this assumption gives the opportunity of taking the hydrostatic pressure as a supplementary state variable.[11] As a consequence of stationarity, this extra variable generates a pressure-dependent family of elastic ranges, such that normality holds for each member of the family. This makes possible to include in the present scheme some plasticity models usually described by nonassociated flow rules.

---

[10]In the models based on the multiplicative decomposition, an equivalent constitutive assumption is the assumption of *plastic irrotationality*, by which the *plastic spin*, which is the skew-symmetric part of the plastic strain rate, is set to zero. Here, too, crystal plasticity is an exception.

[11]See, e.g., [Srinivasa 2010; Vermeer and de Borst 1984, Ziegler 1983, §17.6; Ziegler and Wehrli 1987, §VII.A].

The isotropic case, in which the plastic energy is independent of the direction of the plastic strain rate, is investigated in Section 6.3. This case includes the energies of Drucker and Prager and of Mises. These energies are used as paradigms to compare the plastic behaviors of metals and geomaterials. The fact that isochoric plasticity cannot describe some plastic volume changes observed in geomaterials, such as the dilatancy of soils, motivates the *Cam-clay model* summarized in Section 6.4, specifically conceived to describe such phenomena [Roscoe and Poorooshasb 1963; Roscoe et al. 1958].

Finally, Section 7 deals with the plasticity of crystals. This is a special case of isochoric plasticity, in which the plastic strain has the form of slips occurring on predetermined slip planes. As said above, this is the only case considered in this paper in which the form of the plastic strain is given a priori, so that there is no need of specifying constitutively any rotation. The single-slip and the multislip models are illustrated in Sections 7.1 and 7.2, respectively. In Section 7.3, the periodic energies used to study the two-level shear of single crystals are considered. Within the exception constituted by crystal plasticity, this model exhibits the further exception that the plastic potential which governs the evolution of the plastic strain need not be nonsmooth. Instead of being diffused along the whole process, the plastic dissipation concentrates on singular instability events of catastrophic nature. This opens perspectives of revision of the bases of classical plasticity, including the revisitation of the concept of elastic range and of the other constitutive elements. For reasons of brevity, only a mention of such perspectives can be made here.

It is the present author's opinion that the possibility of treating in a unified way many models reproducing the behavior of materials of different natures, just acting on the shape of the plastic energy and then looking at the consequences of the plastic stationarity condition, is a paramount advantage of the variational approach.

2. Notation and preliminaries

2.1. *Linear spaces and linear maps.* By *linear space* we mean a finite-dimensional vector space endowed with an inner product. A *linear map* from a linear space $\mathscr{A}$ to a linear space $\mathscr{B}$ is a map $\ell : \mathscr{A} \to \mathscr{B}$ such that

$$\ell(\alpha H + \beta K) = \alpha \ell(H) + \beta \ell(K) \quad \text{for all } H, K \in \mathscr{A} \text{ and } \alpha, \beta \in \mathbb{R}. \quad (2\text{-}1)$$

The set of all linear maps of $\mathscr{A}$ into $\mathscr{B}$ is a linear space which will be denoted by $\mathscr{L}(\mathscr{A}, \mathscr{B})$.

If $\mathscr{B}$ is the real line $\mathbb{R}$, the elements of $\mathscr{L}(\mathscr{A}, \mathbb{R})$ are the *linear functionals on $\mathscr{A}$*. By the representation theorem of linear functionals, $\mathscr{L}(\mathscr{A}, \mathbb{R})$ is isomorphic to $\mathscr{A}$. That is, with every $\ell \in \mathscr{L}(\mathscr{A}, \mathbb{R})$ one can associate a unique element H of $\mathscr{A}$ such

that

$$\ell(K) = H \cdot K \quad \text{for all } K \in \mathscr{A}, \tag{2-2}$$

with $(\,\cdot\,)$ the inner product of $\mathscr{A}$.[12]

For $\mathscr{A}$ fixed and $\mathscr{B} = \mathscr{A}$, we write $\mathscr{L}$ in place of $\mathscr{L}(\mathscr{A}, \mathscr{A})$ and we call its elements *second-order tensors*. The linear maps of $\mathscr{L}$ into itself are called *fourth-order tensors*.

A *subspace* of a linear space is a subset which is itself a linear space. Proper subspaces of $\mathscr{L}$ are the set $\mathscr{S}$ of all symmetric tensors and the set $\mathscr{W}$ of all skew-symmetric tensors.

2.2. *Homogeneous and bihomogeneous maps.*

In plasticity, an important role is played by homogeneous and bihomogeneous maps. Therefore, it does not seem inappropriate to recall some basic definitions and to introduce some ad hoc notation. A map $\mathfrak{h}$ from a linear space $\mathscr{A}$ to a linear space $\mathscr{B}$ is *homogeneous of order one*, in short, *homogeneous*, if

$$\mathfrak{h}(tH) = t\mathfrak{h}(H) \quad \text{for all } H \in \mathscr{A} \text{ and } t \geq 0. \tag{2-3}$$

The set $\mathscr{H}(\mathscr{A}, \mathscr{B})$ of all homogeneous maps from $\mathscr{A}$ to $\mathscr{B}$ is a vector space, with obvious definitions of the sum and multiplication by a scalar:

$$(\mathfrak{h} + \mathfrak{l})(H) = \mathfrak{h}(H) + \mathfrak{l}(H), \quad (\alpha\mathfrak{h})(H) = \alpha\mathfrak{h}(H) \quad \text{for all } \alpha \in \mathbb{R} \text{ and } H \in \mathscr{A}. \tag{2-4}$$

If a homogeneous map is additive,

$$\mathfrak{h}(H + K) = \mathfrak{h}(H) + \mathfrak{h}(K) \quad \text{for all } H, K \in \mathscr{A}, \tag{2-5}$$

it is linear. Therefore, $\mathscr{L}(\mathscr{A}, \mathscr{B})$ is a proper subspace of $\mathscr{H}(\mathscr{A}, \mathscr{B})$.

If $\mathscr{B}$ is the real line $\mathbb{R}$, the elements of $\mathscr{H}(\mathscr{A}, \mathbb{R})$ are the *homogeneous functionals on $\mathscr{A}$*. In this case we write

$$\mathfrak{h} \triangleright K \quad \text{in place of} \quad \mathfrak{h}(K). \tag{2-6}$$

The operator "$\triangleright$" maps the elements of $(\mathscr{H}(\mathscr{A}, \mathbb{R}) \times \mathscr{A})$ into the real numbers. In particular, by (2-2), a linear $\mathfrak{h}$ can be identified with an element H of $\mathscr{A}$. In this case, the pairing reduces to the inner product of $\mathscr{A}$:

$$\mathfrak{h} \in \mathscr{L}(\mathscr{A}, \mathbb{R}) \implies \mathfrak{h} \triangleright K = H \cdot K \quad \text{for all } K \in \mathscr{A}. \tag{2-7}$$

A *bihomogeneous map* on $\mathscr{A}$ is a homogeneous map $\mathfrak{K}$ from $\mathscr{A}$ to $\mathscr{H}(\mathscr{A}, \mathbb{R})$. For any such map, the homogeneous functional on $\mathscr{A}$ obtained applying $\mathfrak{K}$ to $K \in \mathscr{A}$

---

[12]See, e.g., [Halmos 1942, §67].

and the real number obtained by applying this homogeneous functional to $H \in \mathscr{A}$ are denoted by

$$\mathfrak{K}\{K\}, \qquad \mathfrak{K}\{K\} \triangleright H, \tag{2-8}$$

respectively. For every bihomogeneous map we have

$$\mathfrak{K}\{\lambda K\} \triangleright \mu H = \lambda \mu \mathfrak{K}\{K\} \triangleright H \quad \text{for all } \lambda, \mu \geq 0. \tag{2-9}$$

If $\mathfrak{K}$ is linear, we write $\mathfrak{K}[K]$ in place of $\mathfrak{K}\{K\}$. If $\mathfrak{K}$ is linear and $\mathfrak{K}[K]$ is linear for all K, then $\mathfrak{K}$ is a linear map from $\mathscr{A}$ to itself, that is, a fourth-order tensor $\mathbb{K}$. In this case, we have

$$\mathfrak{K}\{K\} \triangleright H = \mathbb{K}[K] \cdot H. \tag{2-10}$$

2.3. Directional derivatives. Let $\mathscr{A}$ be a linear space, and let ϕ be a map of $\mathscr{L} = \mathscr{L}(\mathscr{A}, \mathscr{A})$ into the real line $\mathbb{R}$. For $A \in \mathscr{L}$ and $H \in \mathscr{L} \setminus \{0\}$, the limit

$$\check{\nabla}\phi(A) \triangleright H = \lim_{\varepsilon \to 0^+} \frac{\phi(A + \varepsilon H) - \phi(A)}{\varepsilon} \tag{2-11}$$

is the *directional derivative* at A in the direction H. If this limit exists for all H, we say that ϕ is *Gâteaux differentiable* at A. From its very definition it is clear that $\check{\nabla}\phi(A)$ is a homogeneous functional, that is, an element of $\mathscr{H}(\mathscr{L}, \mathbb{R})$. If $\check{\nabla}\phi(A)$ is additive, it reduces to the ordinary derivative $\nabla\phi(A)$, and ϕ is said to be Fréchet differentiable, in short, *differentiable*, at A.

The map ϕ is *twice Gâteaux differentiable* at A if $\check{\nabla}\phi$ is Gâteaux differentiable in a neighborhood of A and the limit

$$\check{\nabla}^2\phi(A)\{H\} \triangleright K = \lim_{\varepsilon \to 0^+} \frac{\check{\nabla}\phi(A + \varepsilon H) \triangleright K - \check{\nabla}\phi(A) \triangleright K}{\varepsilon} \tag{2-12}$$

exists for all $H \in \mathscr{L} \setminus \{0\}$. In this case $\check{\nabla}^2\phi(A)$ is said to be the *second directional derivative at A in the direction H*. It is clear that $\check{\nabla}^2\phi(A)\{H\}$ is an element of $\mathscr{H}(\mathscr{L}, \mathbb{R})$, and that $\check{\nabla}^2\phi(A)$ is a homogeneous map from $\mathscr{L}$ to $\mathscr{H}(\mathscr{L}, \mathbb{R})$, that is, a bihomogeneous map on $\mathscr{L}$. If $\check{\nabla}^2\phi(A)$ is linear in both H and K, it reduces to the second Fréchet derivative $\nabla^2\phi(A)$, which is identified with a fourth-order tensor. Here we shall be interested in the case $H = K$, in which the expansions

$$\phi(A + \varepsilon H) = \phi(A) + \varepsilon \check{\nabla}\phi(A) \triangleright H + \tfrac{1}{2}\varepsilon^2 \check{\nabla}^2\phi(A)\{H\} \triangleright H + o(\varepsilon^2),$$
$$\check{\nabla}\phi(A + \varepsilon H) = \check{\nabla}\phi(A) + \varepsilon \check{\nabla}^2\phi(A)\{H\} + o(\varepsilon) \tag{2-13}$$

are direct consequences of the definitions.

2.4. Nonsmooth potentials and dissipation potentials. A function ϕ from $\mathscr{L}$ to the real line

(i) *with $\phi(A) = 0$,*

(ii) *twice differentiable at* $\mathscr{L} \setminus \{A\}$,

(iii) *twice Gâteaux differentiable at* A,

will be called a *nonsmooth potential from* A. A *smooth potential* is the special case in which ϕ is also twice differentiable at A. A nonsmooth potential

(iv) *strictly increasing along every direction* H,

$$\lambda > \mu \geq 0 \quad \Longrightarrow \quad \phi(A + \lambda H) > \phi(A + \mu H) \quad \text{for all } H \in \mathscr{L} \setminus \{0\}, \quad (2\text{-}14)$$

will be called a *dissipation potential*. We point out that (iv) implies

$$\check{\nabla}\phi(A) \triangleright H > 0 \quad \text{for all } H \in \mathscr{L} \setminus \{0\}. \quad (2\text{-}15)$$

A dissipation potential is *homogeneous* from $A \in \mathscr{L}$ if[13]

$$\phi(A + \varepsilon H) = \varepsilon\phi(A + H) \quad \text{for all } H \in \mathscr{L} \text{ and } \varepsilon \geq 0. \quad (2\text{-}16)$$

For a homogeneous potential from A, from $(2\text{-}13)_1$ we have

$$\phi(A + H) = \check{\nabla}\phi(A) \triangleright H, \qquad \check{\nabla}^2\phi(A)\{H\} \triangleright H = 0 \quad \text{for all } H \in \mathscr{L}. \quad (2\text{-}17)$$

3. The incremental problem

3.1. *From the equilibrium problem to the evolution problem.* Let Ω_R be the region occupied by a continuous body in the reference placement. Suppose that on Ω_R is prescribed a system of external loads, consisting of a body force field b_R at the interior of Ω_R and of a surface traction field s_R on a given portion $\partial_s\Omega_R$ of the boundary $\partial\Omega_R$.

A stress field T_R on Ω_R is said to be *equilibrated* with the given loads if it satisfies the *virtual work equation*[14]

$$\int_{\Omega_R} T_R \cdot \nabla v \, dV_R = \int_{\Omega_R} b_R \cdot v \, dV_R + \int_{\partial_s\Omega_R} s_R \cdot v \, dA_R \quad (3\text{-}1)$$

for all vector fields v on Ω_R which vanish on $\partial\Omega_R \setminus \partial_s\Omega_R$. The *Piola stress tensor* T_R is related to the deformation gradient ∇f by a *response law*. The *equilibrium problem* consists of finding the deformation f from Ω_R produced by the given loads. The *weak formulation* of the problem is obtained substituting the response law into the virtual work equation and imposing the boundary condition of place

$$f(x) = \hat{f}(x), \quad x \in \partial\Omega_R \setminus \partial_s\Omega_R. \quad (3\text{-}2)$$

---

[13]See, e.g., [Eve et al. 1990; Martin and Reddy 1993]. Here, nonhomogeneous potentials appear in the dilatational plasticity models in Section 6. A dissipation potential is also frequently assumed to be convex, but this assumption is not essential; see, e.g., [Petryk 2003, Remark 1]. In fact, nonconvexity is required to describe the softening response; see Section 6.1 below.

[14]See, e.g., [Ciarlet 1988, §2.6]. For simplicity, the variable of integration x is omitted.

For an elastic body, the response law is a functional relation between the punctual values of T_R and ∇f:

$$T_R(x) = \mathcal{G}(\nabla f(x)), \quad x \in \Omega_R. \tag{3-3}$$

For nonelastic bodies, the response law is less simple. In the class of *simple materials* it is assumed that, at each point x, the stress $T_R(x)$ is determined by the past history of the gradient of f at x [Truesdell and Noll 1965, §28; Noll 1972]. By consequence, the final deformation does not only depend on the final load (b_R, s_R), but also on the loading path along which this load has been reached.

To follow the evolution of the load, the equilibrium problem is replaced by an *evolution problem*, which can be formulated as follows. Let an initial placement f_0 of the body, an initial stress T_{R0}, a loading path $t \mapsto (b_{Rt}, s_{Rt})$, and a family $t \mapsto \hat{f}_t$ of constraints be given.[15] The initial placement is required to satisfy the constraint $\hat{f}_0$ on the constrained part of the boundary, and the initial stress is required to be equilibrated with the initial load (b_{R0}, s_{R0}). The problem consists of determining a deformation process $t \mapsto f_t$ such that, for each $t > 0$, f_t satisfies the constraint $\hat{f}_t$ and the stress T_{Rt} is equilibrated with the loads (b_{Rt}, s_{Rt}).

3.2. *From the evolution problem to the incremental problem.* In the evolution problem, t varies over a finite interval $[0, t^\dagger]$. In the *incremental equilibrium problem*, this interval is restricted by taking the limit of $t^\dagger$ to 0^+. Then, from the expansions

$$b_{Rt} = b_{R0} + t\delta b_R + o(t), \quad s_{Rt} = s_{R0} + t\delta s_R + o(t), \quad T_{Rt} = T_{R0} + t\delta T_R + o(t), \tag{3-4}$$

the incremental version of the virtual work equation (3-1)

$$\int_{\Omega_R} \delta T_R \cdot \nabla v \, dV_R = \int_{\Omega_R} \delta b_R \cdot v \, dV_R + \int_{\partial_s \Omega_R} \delta s_R \cdot v \, dA_R \tag{3-5}$$

follows. The weak formulation of the problem is completed by the prescription of an incremental constraint $\hat{u}$ on $\partial\Omega_R \setminus \partial_s \Omega_R$ and of an incremental response law of the form

$$\delta T_R = \mathfrak{h}(\nabla u), \tag{3-6}$$

where $u(x) = f(x) - x$ is the displacement vector and $\mathfrak{h}$ is a homogeneous map depending on a set of variables which define the current *state* of each specific continuum.[16]

---

[15]The parameter t is an *internal time* which describes the deformation path. It need not be related to the physical time.

[16]This map is an *evolution function* in the sense of W. Noll's *new theory of simple materials* [Noll 1972].

For the solution of the evolution problem, a natural strategy is to subdivide the interval $[0, t^\dagger]$ into subintervals, and to solve the incremental problem on each subinterval, using the solution of each problem as the input for the next one. It is expected that, in the presence of sufficient regularity, the solution to the evolution problem would be achieved in the limit, when the lengths of all subintervals tend uniformly to zero. There is, however, a difficulty in passing from a subinterval to the next. Indeed, in this passage the initial deformation f_0, the loads b_R, s_R, and the constraint $\hat{u}$ change. By consequence, the response function $\mathfrak{h}$ and the Piola stress T_R must be updated. This is perhaps the most difficult part of the solution procedure. It will be not considered in the present paper, which is restricted to the formulation of a single incremental problem.

The assumption of infinitesimal deformations, by which these changes are neglected, drastically simplifies the problem. However, since the increments obey different laws in small and in large deformations, the incremental problem is different in the two cases. Here, only the formulation in the context of large deformations will be considered.

3.3. *The energetic formulation.* An alternative formulation of the equilibrium problem consists of transforming it into a minimum problem for the total energy of the body. In this approach the equilibrium placements of the body are characterized by the stationarity of the energy, that is, by the nonnegativeness of the first variation. As already said in the Introduction, the stability of the equilibrium, which is decided by the sign of the second variation, will not be considered here.

The energy of a body is assumed to be the sum of an internal strain energy plus the energy communicated by the external loads:

$$\mathcal{E}_{\text{tot}} = \mathcal{E}_{\text{int}} + \mathcal{E}_{\text{ext}}. \tag{3-7}$$

A *classical continuum* is a continuum with an external energy of the form[17]

$$\mathcal{E}_{\text{ext}}(b_R, s_R, u) = -\int_{\Omega_R} b_R \cdot u \, dV_R - \int_{\partial\Omega_R} s_R \cdot u \, dA_R. \tag{3-8}$$

The internal energy has the form of a volume integral, with an energy density depending on the strain measures which characterize each specific continuum. In particular, the class of *two-scale*, or *micromorphic*, continua is characterized by two strain measures, the macroscopic deformation f and the local microscopic deformation F.[18] For a two-scale deformation (f, F) we assume an internal energy

---

[17]In a nonclassical, or *generalized* continuum, the external energy has supplementary terms involving external actions associated with the state variables.

[18]Here *local* means that the function F is not in general the gradient of a function defined over the whole body.

of the form

$$\mathcal{E}_{\mathrm{int}}(f, F) = \int_{\Omega_R} (\varphi(F) + \phi(F^d))\, dV_R, \qquad (3\text{-}9)$$

where

$$F^d = \nabla f - F \qquad (3\text{-}10)$$

is the deformation due to the *disarrangements* [Owen 1995] which occur at the microscopic scale, and the energy densities φ and ϕ are a smooth and a nonsmooth potential, respectively.

This energy is the relaxed limit of the energies associated with the discontinuities in a sequence of piecewise continuous macroscopic deformations approximating (f, F).[19] The existence of such approximating sequences for any two-scale deformation is ensured by the approximation theorem of the theory of *structured deformations* [Del Piero and Owen 1993].[20]

3.4. *Incremental minimization.* Let $t \mapsto (f_t, F_t)$ be a solution of the incremental problem from an initial placement which, for convenience, we take as the reference placement. Then f_0 and F_0 are the identity map ι and the identity tensor I, respectively, and f_t and F_t have the expansions

$$f_t = \iota + t\nabla u + o(t), \qquad F_t = I + tL + o(t). \qquad (3\text{-}11)$$

The region Ω_R is now the region Ω_0 occupied by the body at $t = 0$, and the loads and stresses (3-4) are now

$$b_t = b_0 + t\delta b_R + o(t), \quad s_t = s_0 + t\delta s_R + o(t), \quad T_t = T_0 + t\delta T_R + o(t). \quad (3\text{-}12)$$

For $F_t^d = \nabla f_t - F_t$ we have

$$F_0^d = 0, \qquad F_t^d = tL^d + o(t), \qquad L^d = \nabla u - L, \qquad (3\text{-}13)$$

----

[19]The dependence of the relaxed energy on the pair (F, F^d) was proved by Choksi and Fonseca [1997] for a model of defective crystals. They also found that the sum decomposition (3-9) holds in some special cases, and conjectured that it may hold in general. In [Del Piero 2001; Owen 2000] their conjecture was proved to be true in one dimension. Later this result was extended to three dimensions, under different regularity assumptions. In the papers [Baía et al. 2012; Owen and Paroni 2015; Šilhavý 2017], ϕ was found to be a homogeneous potential. For plastic materials, we shall see later that the homogeneous potentials can only describe a perfectly plastic response. To describe work-hardening, more general potentials have been obtained either assuming special forms of ϕ [Deseri and Owen 2002] or taking the relaxed limit of approximating sequences for *second-order structured deformations* [Barroso et al. 2017].

[20]That F^d is an independent energetic variable is due to the fact that, in the limit, the discontinuities in the piecewise continuous approximating deformations have a volume density, which is exactly F^d; see [Del Piero and Owen 1995].

and, for the energy densities φ and ϕ,

$$\varphi(F_t) = t\nabla\varphi(I) \cdot L + \tfrac{1}{2}t^2\nabla^2\varphi(I)[L] \cdot L + o(t^2),$$
$$\phi(F_t^d) = t\check{\nabla}\phi(0) \triangleright L^d + \tfrac{1}{2}t^2\check{\nabla}^2\phi(0)[L^d] \triangleright L^d + o(t^2). \tag{3-14}$$

For fixed $t > 0$, the solution (f_t, F_t) is the minimum of the total energy in the class of all deformations $(f_\varepsilon, F_\varepsilon)$ such that

$$f_\varepsilon = f_t + \varepsilon v, \qquad F_\varepsilon = F_t + \varepsilon\mathcal{L}, \qquad F_\varepsilon^d = F_t^d + \varepsilon\mathcal{L}^d, \tag{3-15}$$

with $(v, \mathcal{L})$ a field of virtual velocities on Ω_0 and $\varepsilon > 0$ a scalar parameter. The total energy of $(f_\varepsilon, F_\varepsilon)$ is

$$\mathcal{E}_{\text{tot}}(f_\varepsilon, F_\varepsilon) = \mathcal{E}_{\text{tot}}(f_t, F_t) + \varepsilon \int_{\Omega_0} (\nabla\varphi(F_t) \cdot \mathcal{L} + \check{\nabla}\phi(F_t^d) \triangleright \mathcal{L}^d) \, dV_0$$
$$- \varepsilon\left(\int_{\Omega_0} b_t \cdot v \, dV_0 + \int_{\partial\Omega_0} s_t \cdot v \, dA_0\right) + o(\varepsilon). \tag{3-16}$$

The energy is said to be *stationary* at (f, F) if

$$\lim_{\varepsilon\to 0^+} \frac{1}{\varepsilon}(\mathcal{E}_{\text{tot}}(f_\varepsilon, F_\varepsilon) - \mathcal{E}_{\text{tot}}(f_t, F_t)) \geq 0 \tag{3-17}$$

for all perturbations $(v, \mathcal{L})$. In particular, for $\mathcal{L} = \nabla v$ we have $\mathcal{L}^d = 0$, and for $v = 0$ on $\partial\Omega_0 \setminus \partial_s\Omega_0$ we get

$$\int_{\Omega_0} \nabla\varphi(F_t) \cdot \nabla v \, dV_0 = \int_{\Omega_0} b_t \cdot v \, dV_0 + \int_{\partial_s\Omega_0} s_t \cdot v \, dA_0, \tag{3-18}$$

with the equality sign because the inequality holds for both v and $-v$. A comparison with the virtual work equation (3-1) then leads to the identification

$$T_t = \nabla\varphi(F_t) \tag{3-19}$$

of $\nabla\varphi$ with a stress field T_t equilibrated with the loads (b_t, s_t). Therefore, *at every stationarity point of the energy there is a stress field T_t equilibrated with the given loads*. In particular, the same relation at $t = 0$

$$T_0 = \nabla\varphi(I) \tag{3-20}$$

provides the stress T_0 associated with the initial deformation $F_0 = I$.

As a consequence of (3-18), all terms in v cancel from (3-16) with $\mathcal{L}$ replaced by $\nabla v - \mathcal{L}^d$. Then (3-17) reduces to the inequality

$$\int_{\Omega_0} \check{\nabla}\phi(F_t^d) \triangleright \mathcal{L}^d \, dV_0 - \int_{\Omega_0} T_t \cdot \mathcal{L}^d \, dV_0 \geq 0. \tag{3-21}$$

The punctual inequality

$$T_t \cdot \mathcal{L}^d \leq \check{\nabla}\phi(F_t^d) \triangleright \mathcal{L}^d \quad \text{for all } \mathcal{L}^d \in \mathcal{L} \tag{3-22}$$

follows almost everywhere in Ω_R.[21] This is a necessary condition for the stationarity of the energy.

Since the left-hand side of (3-21) is the first-order approximation of the difference between the energy of $(f_\varepsilon, F_\varepsilon)$ and the energy of the minimizer (f_t, F_t), the minimum of this difference is zero and is attained for $\mathcal{L}^d = L^d$. Then for $\mathcal{L}^d = L^d$ inequality (3-22) is satisfied as an equality

$$T_t \cdot L^d = \check{\nabla}\phi(F_t^d) \triangleright L^d. \tag{3-23}$$

4. Incremental minimization in plasticity

Classical plasticity deals with continua which are both classical and micromorphic. That is, the deformations are two-scale deformations (f, F), and the total energy is the sum of the energies (3-8) and (3-9). The difference $\nabla f - F$ is the *plastic strain* F^d, and inequality (3-22) is the *plastic stationarity condition*. In what follows, this condition is used to characterize the incremental response law. This requires a preliminary definition of the set of the admissible plastic strain rates and the determination of the restrictions imposed on the energy densities φ and ϕ by material indifference.

4.1. *Admissible plastic strain rates.* The definition $F^d = \nabla f - F$ of the plastic strain can be read as the *additive decomposition*

$$\nabla f = F + F^d \tag{4-1}$$

of the macroscopic deformation gradient into the sum of an elastic and a plastic part.[22] The choice of the additive decomposition has notable consequences on the structure of the plastic strain rate and, consequently, on the restrictions on the plastic potential ϕ imposed by material indifference.

Consider a deformation process $t \mapsto (f_t, F_t)$ with the initial placement as reference placement. In the polar decomposition $F_t^d = R_t^d U_t^d$ of the plastic strain,

---

[21]Indeed, if for some $\mathcal{L}^d$ the punctual inequality were violated on a region of positive volume, the integral inequality would be violated by any perturbation with support in that region and with $\mathcal{L}^d(x)$ parallel to $\mathcal{L}^d$.

[22]This decomposition is standard in infinitesimal plasticity, while in finite plasticity the Kröner–Lee multiplicative decomposition $\nabla f = F^e F^p$ is largely preferred. The reason for the present choice of the additive decomposition is the form (3-9) assumed for the energy of a micromorphic continuum. For a detailed discussion on additive and multiplicative decompositions in plasticity, see the present author's paper [Del Piero 2018].

at $t = 0$ we have $F_0^d = 0$. Then U_0^d is zero, R_0^d is indeterminate, and R_t^d and U_t^d have the expansions[23]

$$R_t^d = R_0^d (I + t W^d) + o(t), \qquad U_t^d = t D^d + o(t). \tag{4-2}$$

Then, by consequence,

$$L^d = R_0^d D^d. \tag{4-3}$$

Thus, in the additive decomposition the *plastic strain rate* L^d and the *plastic stretching* D^d differ by the rotation R_0^d. This rotation is determined by a constitutive assumption.[24] Then either a set $\mathscr{S}_0^d$ of *admissible plastic stretchings* for the initial deformation is selected a priori and the set of the *admissible plastic strain rates* $\mathscr{L}_0^d = R_0^d \mathscr{S}_0^d$ is deduced, or vice versa.

The most common constitutive choice is $R_0^d = I$, which corresponds to assuming the symmetry of the plastic strain rate.[25] In the following we keep this choice. The important exception of crystal plasticity is treated separately in Section 7.

4.2. *Indifference requirements.* Like every scalar-valued function, the potentials φ, ϕ must be invariant under distance-preserving changes of placement.[26] This results in the indifference conditions

$$\varphi(F) = \varphi(QF), \qquad \phi(F^d) = \phi(QF^d) \tag{4-4}$$

to be satisfied by all F, F^d in $\mathscr{L}$ and by all proper orthogonal tensors Q.[27] For the function φ, from the polar decomposition $F = RU$ of F, taking $Q = R^T$ we get the condition

$$\varphi(F) = \varphi(U), \tag{4-5}$$

---

[23]In the multiplicative decomposition, F_t^p is defined in an *intermediate placement* at which both F_0^p and R_0^p are equal to I and therefore are nonsingular. In the additive decomposition there is no such placement, since both f and F are defined on the reference placement.

[24]In the multiplicative decomposition, $F_t^p = I + t L^p$ implies $R_0^p = I$. Nevertheless, the decomposition is determined only to within an indeterminate rotation; see [Dashner 1986; Green and Naghdi 1971; Lubarda and Lee 1981] and the paper [Del Piero 2018] by the present author. Therefore, a constitutive assumption on a rotation has to be made anyway. Note also that, if $R_t^p = I + t W^p$ and $U_t^p = I + t D^p$, the *plastic stretching* D^p and the *plastic spin* W^p are the symmetric and skew-symmetric parts of L^p, respectively. This property is not preserved in the additive decomposition.

[25]In the multiplicative decomposition, the corresponding constitutive choice is the *irrotationality assumption* $W^p = 0$ [Gurtin and Anand 2005; Reddy 2011; Reddy et al. 2008]. This assumption has been made, tacitly or explicitly, in classical textbooks [Hill 1950; Prager 1949; 1955], as well as in some more recent models of gradient plasticity [Fleck and Hutchinson 2001; Gurtin and Anand 2005; Gudmundson 2004].

[26]See, e.g., [Truesdell and Noll 1965, §17].

[27]The transformation law for F^d is a consequence of the laws $\nabla f \to Q \nabla f$ and $F \to QF$. In the multiplicative decomposition, while the transformation law $F^e \to QF^e$ was universally accepted, the question of the transformation law to be adopted for F^p has a long and controversial history [Dashner 1986; Green and Naghdi 1971; Lubarda and Lee 1981; Gurtin et al. 2010, §95.2].

according to which φ is insensitive to the rotation R which follows the pure stretch U in the microscopic deformation. This condition can be used to determine an indifference restriction on φ at $F = I$. Indeed, for R and U consider the expansions

$$R = I + \varepsilon W + \tfrac{1}{2}\varepsilon^2(W^2 + \widetilde{W}) + o(\varepsilon^2), \qquad U = I + \varepsilon D + \tfrac{1}{2}\varepsilon^2\widetilde{D} + o(\varepsilon^2), \quad (4\text{-}6)$$

with $D, \widetilde{D}$ symmetric and $W, \widetilde{W}$ skew-symmetric. Then,

$$F = RU = I + \varepsilon(W + D) + \tfrac{1}{2}\varepsilon^2(W^2 + \widetilde{W} + \widetilde{D} + 2WD) + o(\varepsilon^2), \qquad (4\text{-}7)$$

and therefore,

$$\varphi(F) = \nabla\varphi(I) \cdot (\varepsilon(W + D) + \tfrac{1}{2}\varepsilon^2(W^2 + \widetilde{W} + \widetilde{D} + 2WD))$$
$$+ \tfrac{1}{2}\varepsilon^2\mathbb{C}[W + D] \cdot (W + D) + o(\varepsilon^2), \quad (4\text{-}8)$$

where $\mathbb{C}$ is the fourth-order tensor

$$\mathbb{C} = \nabla^2\varphi(I), \qquad (4\text{-}9)$$

which, by the interchangeability of the order of differentiation, has the symmetry property

$$\mathbb{C}[H] \cdot K = \mathbb{C}[K] \cdot H. \qquad (4\text{-}10)$$

The expansion of $\varphi(U)$ is the same as (4-8), with $W = 0$. Subtracting the two expansions, from (4-5) we get

$$\nabla\varphi(I) \cdot (\varepsilon W + \varepsilon^2\widetilde{W}) + \varepsilon^2(\mathbb{C}[W] - W\nabla\varphi(I)) \cdot (\tfrac{1}{2}W + D) = 0, \qquad (4\text{-}11)$$

and due to the arbitrariness of ε, W, and D, the separate conditions

$$\nabla\varphi(I) \cdot W = 0,$$
$$(\mathbb{C}[W] - W\nabla\varphi(I)) \cdot W = 0, \quad \text{for all } W \in \mathscr{W} \text{ and } D \in \mathscr{S}, \qquad (4\text{-}12)$$
$$(\mathbb{C}[W] - W\nabla\varphi(I)) \cdot D = 0,$$

follow. The last two conditions merge in the single condition

$$\mathbb{C}[W] = W\nabla\varphi(I) \quad \text{for all } W \in \mathscr{W}, \qquad (4\text{-}13)$$

according to which the restriction of $\mathbb{C}$ to $\mathscr{W}$ is determined by $\nabla\varphi(I)$.

For the nonsmooth potential ϕ, for f_t and F_t as in (3-11) and from $(3\text{-}13)_2$ and (4-3) we have

$$\phi(F_t^d) = \phi(tL^d) + o(t) = \phi(tR_0^d D^d) + o(t), \qquad (4\text{-}14)$$

and from the indifference condition $(4\text{-}4)_2$ the condition

$$\phi(F_t^d) = \phi(tD^d) + o(t) \qquad (4\text{-}15)$$

follows. Then from the representations

$$\phi(F_t^d) = t\check{\nabla}\phi(0) \triangleright L^d + \tfrac{1}{2}t^2\check{\nabla}^2\phi(0)\{L^d\} \triangleright L^d,$$
$$\phi(F_t^d) = t\check{\nabla}\phi(0) \triangleright D^d + \tfrac{1}{2}t^2\check{\nabla}^2\phi(0)\{D^d\} \triangleright D^d, \qquad (4\text{-}16)$$

we get

$$\check{\nabla}\phi(0) \triangleright L^d = \check{\nabla}\phi(0) \triangleright D^d, \qquad \check{\nabla}^2\phi(0)\{L^d\} \triangleright L^d = \check{\nabla}^2\phi(0)\{D^d\} \triangleright D^d. \quad (4\text{-}17)$$

Since $R_0^d = I$ implies $L^d = D^d$, the indifference conditions on ϕ are trivially satisfied assuming $R_0^d = I$. Effective restrictions on ϕ follow from any other constitutive choice of R_0^d.

5. Construction of the incremental response law

The purpose of this section is to deduce the incremental response law from the plastic stationarity condition (3-22). From (3-19), (3-12)$_3$, and (3-14)$_1$ we have

$$T_0 + t\delta T = \nabla\varphi(I) + t\nabla^2\varphi(I)[L] + o(t), \qquad (5\text{-}1)$$

and from (4-9), (3-13)$_3$, and (4-3) with $R_0^d = I$,

$$\delta T = \mathbb{C}[L] = \mathbb{C}[\nabla u - L^d] = \mathbb{C}[\nabla u - D^d]. \qquad (5\text{-}2)$$

By comparison with the incremental stress-strain relation (3-6), we get

$$\mathbb{C}[\nabla u - D^d] = \mathfrak{h}(\nabla u). \qquad (5\text{-}3)$$

This is a homogeneous map $\mathfrak{g}$ from ∇u to D^d

$$D^d = \mathfrak{g}(\nabla u). \qquad (5\text{-}4)$$

Therefore, the determination of the incremental stress-strain relation (3-6) is reduced to the determination of the map $\mathfrak{g}$. With this goal in mind, we proceed to the characterization of the basic constitutive elements of the theory, in the order in which they are listed in the Introduction.

5.1. *The yield condition.* With the identification $\mathcal{L}^d = \mathcal{D}^d$ due to the assumption $R_0^d = I$, the plastic stationarity condition (3-22) at $t = 0$ becomes

$$T_0 \cdot \mathcal{D}^d \leq \check{\Phi}_0 \triangleright \mathcal{D}^d, \quad \check{\Phi}_0 = \check{\nabla}\phi(0), \qquad (5\text{-}5)$$

for all $\mathcal{D}^d$ in a given set $\mathscr{S}_0^d$ of admissible plastic stretchings.[28] This inequality imposes the upper bound $\check{\Phi}_0 \triangleright \mathcal{D}^d$ on the projection of T_0 in the direction of $\mathcal{D}^d$.

---

[28]Usually, $\mathscr{S}_0^d$ is assumed to be a cone in $\mathscr{S}$, that is, a subset of $\mathscr{S}$ such that $\mathcal{D}^d \in \mathscr{S}_0^d \implies \lambda\mathcal{D}^d \in \mathscr{S}_0^d$ for all $\lambda > 0$.

Then, $\check{\Phi}_0$ is a *bounding map* for T_0. Indeed, inequality (5-5) says that for every direction $N = \mathcal{D}^d/|\mathcal{D}^d|$ in $\mathscr{S}_0^d$ the stress T_0 must belong to the set [29]

$$\mathcal{H}_0^N = \{T \in \mathscr{S} \mid T \cdot N \le \check{\Phi}_0 \triangleright N\}, \tag{5-6}$$

which is the closed half-space of $\mathscr{S}$ bounded by the hyperplane with exterior unit normal N, placed at the distance $\check{\Phi}_0 \triangleright N$ from the origin. Since this holds for all directions in $\mathscr{S}_0^d$, T_0 must belong to the intersection

$$\mathcal{C}_0 = \bigcap_{N \in \mathscr{S}_0^d, \, |N|=1} \mathcal{H}_0^N \tag{5-7}$$

of all $\mathcal{H}_0^N$. Thus, a first consequence of (3-22) is the *yield condition*[30]

$$T_0 \in \mathcal{C}_0. \tag{5-8}$$

Since all $\mathcal{H}_0^N$ are closed convex sets, their intersection $\mathcal{C}_0$ is also a closed convex set. Moreover, if ϕ is a dissipation potential, then $\check{\Phi}_0 \triangleright N > 0$ by (2-15), and the inequality in (5-6) is strict at $T = 0$. That is, the null tensor is an interior point of $\mathcal{H}_0^N$. Since this holds for all N, the point $T = 0$ also belongs to the interior of $\mathcal{C}_0$.

The boundary of $\mathcal{C}_0$ consists of all $T_0 \in \mathscr{S}$ at which the inequality in (5-6) is satisfied as an equality:

$$T_0 \in \partial\mathcal{C}_0 \quad \Longleftrightarrow \quad \text{there is } N_0 \in \mathscr{S}_0^d \text{ such that } T_0 \cdot N_0 = \check{\Phi}_0 \triangleright N_0. \tag{5-9}$$

By (5-6), T_0 also belongs to the boundary of the half-space $\mathcal{H}_0^{N_0}$. Since $\mathcal{C}_0$ is included in $\mathcal{H}_0^{N_0}$, it follows that $T \cdot N_0 \le \check{\Phi}_0 \triangleright N_0$ for all $T \in \mathcal{C}_0$. Then subtracting from the previous equation we get

$$(T_0 - T) \cdot N_0 \ge 0 \quad \text{for all } T \in \mathcal{C}_0; \tag{5-10}$$

that is, N_0 belongs to the normal cone to $\mathcal{C}_0$ at T_0.

From (5-6) and (5-7) we see that $\mathcal{C}_0$ is determined by ϕ. Conversely, for a given $\mathcal{C}_0$ the relation (5-10) associates with every direction N_0 a (possibly nonunique)

---

[29]We recall that $T_0 = \nabla\varphi(I)$ is a symmetric tensor by the indifference condition (4-12)$_1$.

[30]In the terminology of convex analysis, inequality (5-2) says that the map $\check{\Phi}_0$ is subdifferentiable at $\mathcal{D}^d = 0$ and T_0 belongs to the subdifferential of $\check{\Phi}_0$ at 0. Moreover, if $\mathscr{S}_0^d$ is a proper subset of $\mathscr{S}$, then $\mathcal{C}_0$ is unbounded, and $\mathscr{S}_0^d$ and the recession cone of $\mathcal{C}_0$ are polar to each other [Rockafellar 1970, Theorem 14.6]. In particular, if $\mathscr{S}_0^d$ is a subspace, then the recession cone of $\mathcal{C}_0$ is its orthogonal complement. Equation (5-11) below also says that $\check{\Phi}_0$ is the support function, that is, the conjugate of the indicator function, of $\mathcal{C}_0$; see [Eve et al. 1990] or [Rockafellar 1970, Theorem 13.2]. Though the formalism of convex analysis fully captures the mathematical structure of classical plasticity, I prefer to keep the present exposition at a more elementary level.

boundary point T_0 of $\mathcal{C}_0$ such that

$$T_0 \cdot N_0 = \sup_{T \in \mathcal{C}_0} T \cdot N_0. \tag{5-11}$$

Then $\check{\Phi}_0 \triangleright N_0$ is specified by (5-9). Since this can be repeated for all directions N, the whole restriction of $\check{\Phi}_0$ to $\mathcal{S}_0^d$ is specified in this way. Thus, there is a one-to-one correspondence between the closed convex sets $\mathcal{C}_0$ of $\mathcal{S}$ and the homogeneous maps $\check{\Phi}_0$ on $\mathcal{S}_0^d$.[31] On the contrary, the correspondence between directions N in $\mathcal{S}_0^d$ and boundary points of $\mathcal{C}_0$ need not be one-to-one.[32]

5.2. *The flow rule.* At $t = 0$ and for $R_0^d = I$, (4-17)$_1$ holds and (3-23) takes the form

$$T_0 \cdot D^d = \check{\Phi}_0 \triangleright D^d. \tag{5-12}$$

By consequence, a nonnull plastic stretching may occur only if T_0 is a boundary point of $\mathcal{C}_0$. The *flow rule* is a law prescribing the direction of D^d at each boundary point. For example, the *associated flow rule* states that D^d belongs to the normal cone of $\mathcal{C}_0$ at T_0. This is indeed the rule provided by the variational approach, since (5-12) says precisely that D^d belongs to the normal cone at T_0. Calling N_0 the unit tensor $D^d/|D^d|$, the *normality rule*

$$D^d = \begin{cases} \lambda^d N_0, \lambda^d \geq 0 & \text{if } T_0 \in \partial\mathcal{C}_0, \\ 0 & \text{if } T_0 \in \mathcal{C}_0 \setminus \partial\mathcal{C}_0 \end{cases} \tag{5-13}$$

is obtained. The fact that D^d is zero at the interior point justifies the name of *elastic range* given to $\mathcal{C}_0$.

The product $T_0 \cdot D^d$ is the *plastic power*. According to (5-11), we have

$$(T_0 - T) \cdot N_0 \geq 0 \quad \text{for all } T \in \mathcal{C}_0. \tag{5-14}$$

This is Hill's *principle of maximum plastic work.*[33] The inequality is strict for $T = 0$, because the origin of $\mathcal{S}$ is an interior point of $\mathcal{C}_0$. Then the strict inequality

$$T_0 \cdot D^d > 0 \tag{5-15}$$

---

[31]If ϕ is a homogeneous map, then $\check{\Phi}_0 \triangleright N$ is equal to $\phi(N)$ by (2-17)$_1$. Therefore, there is a one-to-one correspondence between closed convex sets and homogeneous dissipation potentials. Eve et al. [1990] called *canonical yield function* the homogeneous dissipation potential associated with $\mathcal{C}_0$.

[32]For example, if $\mathcal{C}_0$ is polyhedral, the normal to a face is a normal at the infinitely many points which belong to that face, and at each point on an edge or vertex of $\mathcal{C}_0$ there are the infinitely many normals which form the normal cone at that point.

[33]According to Srinivasa [2010], the existence and convexity of the elastic range is a consequence of Ziegler's *principle of maximum rate of dissipation* [1963], of which Hill's principle is a special case. That the elastic range is a derived concept was also pointed out by Martin and Reddy [1993]. In the present context, elastic range and Hill's principle are both derived concepts, since both are consequences of the plastic stationarity condition.

holds. This is the *dissipation inequality* which establishes the dissipative character of the plastic power $T_0 \cdot D^d$.

5.3. *The consistency condition.* Equation (3-23) states that, in the limit $t \to 0^+$, the stress T_t and the bounding map $\check{\nabla}\phi(F_t^d)$ have the same projection in the direction of the plastic stretching D^d. This is the *consistency condition*, which says that for $t \to 0^+$ and $D^d \neq 0$ the stress T_t is a boundary point of the elastic range $\mathcal{C}_t$ at t [Prager 1949].[34]

With the identification $L^d = D^d$ due to the assumption $R_0^d = I$, for $\mathcal{L}^d = L^d$ from (3-23) and (5-12) we have

$$\delta T \cdot D^d = \check{\mathbb{K}}\{D^d\} \triangleright D^d, \quad \check{\mathbb{K}} = \check{\nabla}^2\phi(0), \tag{5-16}$$

and since $D^d = \lambda^d N_0$ we get

$$\lambda^d(\delta T \cdot N_0 - \lambda^d \check{\mathbb{K}}\{N_0\} \triangleright N_0) = 0. \tag{5-17}$$

This *complementarity condition*, which states that either λ^d or the term within parentheses is zero, is the mathematical form taken by the consistency condition.

5.4. *The incremental response law.* Due to the normality rule (5-13), the incremental relation (5-2) takes the form

$$\delta T = \mathbb{C}[\nabla u] - \lambda^d \mathbb{C}[N_0]. \tag{5-18}$$

On the right-hand side, only the *plastic multiplier* λ^d is unknown. For its determination we have at our disposal the inequality $\lambda^d \geq 0$ and the complementarity condition (5-17). Together with (5-18) multiplied by N_0,

$$\delta T \cdot N_0 = \mathbb{C}[\nabla u] \cdot N_0 - \lambda^d \mathbb{C}[N_0] \cdot N_0, \tag{5-19}$$

they form a system in the unknowns $\delta T \cdot N_0$ and λ^d. If $\mathbb{C}[\nabla u] \cdot N_0 < 0$, the last equation implies $\delta T \cdot N_0 < 0$, and then $\lambda^d = 0$ by the complementarity condition. If $\lambda^d > 0$, from (5-17) and (5-19) by elimination of $\delta T \cdot N_0$ we get

$$\lambda^d(\mathbb{C}[N_0] \cdot N_0 + \check{\mathbb{K}}\{N_0\} \triangleright N_0) = \mathbb{C}[\nabla u] \cdot N_0. \tag{5-20}$$

Then, assuming

$$\mathbb{C}[N_0] \cdot N_0 + \check{\mathbb{K}}\{N_0\} \triangleright N_0 > 0, \tag{5-21}$$

it follows that $\mathbb{C}[\nabla u] \cdot N_0 \geq 0$ for $\lambda^d \geq 0$. Then the problem has the unique solution

$$\lambda^d(N_0) = \frac{\langle \mathbb{C}[N_0] \cdot \nabla u \rangle}{\mathbb{C}[N_0] \cdot N_0 + \check{\mathbb{K}}\{N_0\} \triangleright N_0}, \tag{5-22}$$

---

[34]Since T_t is not symmetric in general, $\mathcal{C}_t$ is not, in general, a subset of $\mathscr{S}$. The symmetry of T_t and the inclusion of $\mathcal{C}_t$ in $\mathscr{S}$ can be recovered by a change of reference placement, taking at each t the current placement as reference placement.

where $\langle \alpha \rangle = \max\{\alpha, 0\}$ denotes the positive part of a real number α, and $\mathbb{C}[\nabla u] \cdot N_0 = \mathbb{C}[N_0] \cdot \nabla u$ by the symmetry property (4-10). This is a homogeneous relation between $\lambda^d(N_0)$ and ∇u, and its substitution into (5-18) provides the desired incremental law (3-6).

5.5. *Elastic unloading.* A strain rate ∇u is said to determine a regime of

$$\begin{cases} loading & \text{if } \mathbb{C}[N_0] \cdot \nabla u > 0, \\ unloading & \text{if } \mathbb{C}[N_0] \cdot \nabla u < 0, \\ neutral\ loading & \text{if } \mathbb{C}[N_0] \cdot \nabla u = 0. \end{cases} \qquad (5\text{-}23)$$

By (5-22), the response law can be split into two linear relations, one for loading and one for unloading. At loading, from (5-22) we have

$$\delta T = \mathbb{L}[\nabla u], \qquad (5\text{-}24)$$

with

$$\mathbb{L} = \mathbb{C} - \frac{\mathbb{C}[N_0] \otimes \mathbb{C}[N_0]}{\mathbb{C}[N_0] \cdot N_0 + \breve{\mathbb{K}}\{N_0\} \triangleright N_0}, \qquad (5\text{-}25)$$

and at unloading we have

$$\delta T = \mathbb{C}[\nabla u]. \qquad (5\text{-}26)$$

The latter is an *elastic* law, corresponding to $\lambda^d = 0$. It describes the phenomenon of *elastic unloading*, typical of plasticity.

5.6. *The hardening rule.* In the loading regime, for $\lambda^d > 0$ and for sufficiently small $t > 0$, the stress $T_t = T_0 + t\delta T + o(t)$ is placed outside, inside, or on the boundary of the elastic range $\mathcal{C}_0$, depending on the sign of the product $\delta T \cdot N_0$. Since by the consistency condition T_t is a boundary point of the elastic range $\mathcal{C}_t$, the three possible locations of T_t correspond to an enlargement, to a contraction, or to invariance of the elastic range at T_0. In the three cases we say that the response at loading is *hardening, softening*, and *perfectly plastic*, respectively. From (5-24) and (5-25) we get the *hardening rule*[35]

$$\delta T \cdot N_0 = \mathbb{L}[\nabla u] \cdot N_0 = h(N_0)\mathbb{C}[\nabla u] \cdot N_0, \qquad (5\text{-}27)$$

where

$$h(N_0) = \frac{\breve{\mathbb{K}}\{N_0\} \triangleright N_0}{\mathbb{C}[N_0] \cdot N_0 + \breve{\mathbb{K}}\{N_0\} \triangleright N_0} \qquad (5\text{-}28)$$

is the *hardening modulus*. Note that $h(N_0) < 1$ if $\mathbb{C}$ restricted to $\mathscr{S}$ is positive definite and that, the denominator being positive by assumption (5-21), $h(N_0)$ has the same sign of $\breve{\mathbb{K}}\{N_0\} \triangleright N_0$. Then in the loading regime, in which $\mathbb{C}[\nabla u] \cdot N_0$

---

[35]See, e.g., [Lemaitre and Chaboche 1990; Lubliner 1990].

is positive, the response in the direction N_0 is hardening, softening, or perfectly plastic if $h(N_0)$ is positive, negative, or zero, respectively.

6. Some models of classical plasticity

After establishing the incremental response law, we now show how the models of classical plasticity can be obtained taking particular forms of the plastic energy density ϕ. Throughout this section we assume $R_0^d = I$, that is, $L^d = D^d$.

6.1. *The three basic models.* We assume that the elastic energy density φ is a smooth potential and that the plastic energy density ϕ is a nonsmooth potential from $F^d = 0$. We consider three basic models of classical plasticity: perfect, kinematic, and dilatational plasticity, which correspond to the special forms of ϕ

$$\begin{cases} \textit{perfect plasticity} & \phi = \phi^h, \\ \textit{kinematic plasticity} & \phi = \psi + \phi^h, \\ \textit{dilatational plasticity} & \phi = \phi^{nh}, \end{cases} \tag{6-1}$$

with ϕ^h a homogeneous dissipation potential, ϕ^{nh} a nonhomogeneous dissipation potential, and ψ a smooth potential.

6.1.1. *Perfect plasticity.* If ϕ^h is a homogeneous potential from $A = 0$, from (2-17)$_1$ we have $\phi^h(\mathcal{D}^d) = \check{\nabla}\phi^h(0) \triangleright \mathcal{D}^d$, and inequality (5-5) reduces to

$$T_0 \cdot \mathcal{D}^d \le \phi^h(\mathcal{D}^d). \tag{6-2}$$

Then the boundary of each generating half-space $\mathcal{H}_0^N$ is placed at the distance $\phi^h(N)$ from the origin. Moreover, from (2-17)$_2$ we have

$$\check{\mathbb{K}}\{N\} \triangleright N = 0; \tag{6-3}$$

that is, the hardening modulus (5-28) is zero. The hardening rule (5-27) then reduces to

$$\delta T \cdot N_0 = 0, \tag{6-4}$$

in agreement with the definition of a perfectly plastic response given in the preceding section. Therefore,

$$T_t \cdot N_0 = (T_0 + t\delta T) \cdot N_0 = T_0 \cdot N_0. \tag{6-5}$$

Since T_0 is a boundary point of $\mathcal{C}_0$ and T_t is a boundary point of $\mathcal{C}_t$ by the consistency condition, this equation shows that for every direction N_0 in $\mathscr{S}_0^d$ the generating half-space $\mathcal{H}^{N_0}$ for $\mathcal{C}_0$ is also a generating half-space for the projection $\mathcal{C}_t^S$ of $\mathcal{C}_t$ on $\mathscr{S}$.[36] Then $\mathcal{C}_0$ and $\mathcal{C}_t^S$ have the same generating half-spaces; that is, $\mathcal{C}_t^S = \mathcal{C}_0$.

6.1.2. *Kinematic plasticity.* If ϕ is the sum of a smooth potential ψ and a homogeneous dissipation potential ϕ^h, from (4-15) we have

$$\phi(F_t^d) = \psi(tD^d) + \phi^h(tD^d) + o(t). \tag{6-6}$$

Let us use the differentiability of ψ to introduce the tensor

$$T_{Bt} = \nabla\psi(tD^d), \tag{6-7}$$

for which, from the power expansions

$$T_{Bt} = T_{B0} + t\delta T_B + o(t), \qquad \nabla\psi(tD^d) = \nabla\psi(0) + t\nabla^2\psi(0)[D^d] + o(t), \tag{6-8}$$

we have

$$T_{B0} = \nabla\psi(0), \qquad \delta T_B = \mathbb{D}[D^d], \quad \mathbb{D} = \nabla^2\psi(0). \tag{6-9}$$

Inequality (5-5) then takes the form

$$(T_0 - T_{B0}) \cdot N \le \phi^h(N) \quad \text{for all } N \in \mathscr{S}_0^d, |N| = 1. \tag{6-10}$$

This tells us that the boundary of each generating half-space $\mathcal{H}_0^N$ of $\mathcal{C}_0$ is placed at the distance $\phi^h(N)$ from the projection T_{B0}^S of T_{B0} on $\mathscr{S}$. This distance is the same as in perfect plasticity, but now is measured from the point T_{B0}^S and not from the origin. Moreover, since in the identities

$$\mathbb{K}\{N_0\} \triangleright N_0 = \check{\nabla}^2\phi(0)\{N_0\} \triangleright N_0 = \mathbb{D}[N_0] \cdot N_0 + \check{\nabla}^2\phi^h(0)\{N_0\} \triangleright N_0, \tag{6-11}$$

the last term is zero by (2-17)$_2$, the plastic multiplier (5-22) and the hardening modulus (5-28) take the form

$$\lambda^d(N_0) = \frac{\mathbb{C}[\nabla u] \cdot N_0}{(\mathbb{C} + \mathbb{D})[N_0] \cdot N_0}, \qquad h(N_0) = \frac{\mathbb{D}[N_0] \cdot N_0}{(\mathbb{C} + \mathbb{D})[N_0] \cdot N_0}. \tag{6-12}$$

Then, from (5-27) and (6-9)$_2$,

$$(\delta T - \delta T_B) \cdot N_0 = h(N_0) \, \mathbb{C}[\nabla u] \cdot N_0 - \lambda^d(N_0) \, \mathbb{D}[N_0] \cdot N_0 = 0. \tag{6-13}$$

This is the same relation (6-4) of perfect plasticity with T_t replaced by $(T_t - T_{Bt})$. It says that the generating half-spaces for $\mathcal{C}_t^S$ are the generating half-spaces $\mathcal{H}^{N_0}$ for $\mathcal{C}_0$ translated of δT_B^S, that is, that $\mathcal{C}_t^S = \mathcal{C}_0 + \{\delta T_B^S\}$. The fact that the translation of the

---

[36]We recall that, in general, $\mathcal{C}_t$ is not included in $\mathscr{S}$; see Footnote 34. Here and in the following, the superscript S denotes the projection on $\mathscr{S}$ symmetric part of a tensor. In particular, A^S denotes the symmetric part of a tensor A.

elastic range is controlled by the translation of T_B motivates the name *backstress tensor* attributed to T_B.

In kinematic plasticity, the response in a direction N_0 is hardening, softening, or perfectly plastic depending on the sign of $h(N_0)$. For $h(N_0)$ positive, we have Prager's *kinematic hardening* model [1955].[37]

6.1.3. *Dilatational plasticity.* As discussed in Section 5.6, to a positive hardening modulus corresponds an enlargement of the elastic range

$$h(N_0) > 0 \quad \text{for all } N_0 \in \mathscr{S}^d \quad \Longrightarrow \quad \mathcal{C}_0 \subset \mathcal{C}_t^S. \tag{6-14}$$

By (5-21), under assumption (5-28), $h(N_0)$ is positive if $\check{\mathbb{K}}\{N_0\} \rhd N_0$ is positive. If ϕ is a nonhomogeneous dissipation potential ϕ^{nh}, then $\check{\mathbb{K}} = \check{\nabla}^2 \phi^{nh}(0)$. Therefore, a hardening response occurs in all directions N_0 if ϕ^{nh} is strictly convex. *Isotropic hardening* is the special case of $h(N_0)$ independent of N_0.

If $h(N_0)$ is positive at $t = 0$, it remains positive for sufficiently small t. For all such t there is no way of producing a contraction, that is, the expansion of the elastic range is irreversible.

In conclusion, comparing $\mathcal{C}_t^S$ and $\mathcal{C}_0$ we see that a homogeneous dissipation potential leaves $\mathcal{C}_0$ unchanged, a smooth potential produces a rigid translation, and a nonhomogeneous dissipation potential produces a dilatation or a contraction, depending on the sign of the hardening modulus. More complex evolutions of the elastic range can be described by combinations of the potentials considered here.[38]

6.2. *Isochoric plasticity.* The experiments show that for many materials the plastic strain rate is practically isochoric. The presence of this internal constraint modifies the form of the energy, and this requires a nontrivial reformulation of the theory.

6.2.1. *The isochoricity constraint.* We still consider two-scale deformations (f, F) and take the difference $(\nabla f - F)$ as the plastic deformation F^d. *Isochoricity* is the assumption that the volume changes in the macroscopic and microscopic deformations are the same:

$$\det F = \det \nabla f. \tag{6-15}$$

In a deformation process $t \mapsto (f_t, F_t)$ from $(\imath, I)$, from the expansions (3-11),

$$\det \nabla f_t = 1 + t I \cdot \nabla u + o(t), \qquad \det F_t = 1 + t I \cdot L + o(t), \tag{6-16}$$

---

[37]The dependence of the backstress tensor on the differentiable part of the plastic strain energy was pointed out by Aifantis [1987]. See also [Anand and Gurtin 2003].

[38]In [Gurtin et al. 2010, p. 421], it has been pointed out that "for many metals, the actual strain-hardening behavior... may be approximated by a combination of nonlinear isotropic hardening and nonlinear kinematic hardening." In fact, the possibility of more general combinations of kinematic and dilatational plasticity emerges from the present analysis.

and since $\nabla u - L = L^d$, isochoricity results in the constraint

$$I \cdot L^d = 0. \tag{6-17}$$

Consider the decomposition of a second-order tensor A into the sum of a hydrostatic and a deviatoric part:

$$A = A^H + A^D, \quad A^H = \tfrac{1}{3}(I \cdot A)I, \quad I \cdot A^D = 0. \tag{6-18}$$

They are the perpendicular projections of A on the *hydrostatic axis* and on the *deviatoric hyperplane*,

$$\mathscr{L}^H = \{A \in \mathscr{L} \mid A = \alpha I, \ \alpha \in \mathbb{R}\}, \qquad \mathscr{L}^D = \{A \in \mathscr{L} \mid I \cdot A = 0\}, \tag{6-19}$$

respectively. Then the constraint (6-17) can be written in any of the equivalent forms

$$L^d \in \mathcal{L}^D, \qquad L^d = L^{dD}, \qquad L^{dH} = 0. \tag{6-20}$$

6.2.2. *The plastic strain energy.* The elastic energy density φ is a function of the variable F, which is not restricted by the constraint (6-17). Then φ is not constrained as well. On the contrary, the plastic energy density $\tilde{\phi}$ depends on the plastic strain rate L^d, which is now assumed to be deviatoric. It is then appropriate to assume a dependence on the part of the stress which does no work in any deformation process which satisfies the constraint, that is, on the hydrostatic stress T^H. Here we assume the separate dependence

$$\tilde{\phi}(L^d, p) = \phi(L^d)\psi(p), \quad L^d \in \mathcal{L}^D, \ p \in \mathbb{R}, \tag{6-21}$$

where ϕ is a dissipation potential, ψ is a smooth potential, and p is the *hydrostatic pressure*

$$p = -\tfrac{1}{3}I \cdot T, \tag{6-22}$$

related to the hydrostatic stress by

$$T^H = -pI. \tag{6-23}$$

6.2.3. *The plastic stationarity condition.* At each point x, consider a deformation process $t \mapsto (f_t, F_t)$ from the reference placement, and a pressure process $t \mapsto p_t$. To minimize the total energy at the time t, take a family $\varepsilon \mapsto (\varepsilon v, \varepsilon \mathcal{L})$ of perturbations and consider the perturbed process $t \mapsto (f_\varepsilon, F_\varepsilon)$, with f_ε and F_ε as in (3-15). The perturbed plastic energy density is $\phi(F_\varepsilon^d)\psi(p_t)$, with p_t the pressure at t and

$$\phi(F_\varepsilon^d) = \phi(F_t^d) + \varepsilon \check{\nabla}\phi(F_t^d) \triangleright \mathcal{L}^d. \tag{6-24}$$

With the expansions (3-11) of f_t and F_t, the perturbed total energy takes the form

$$\mathcal{E}_{\text{tot}}(f_\varepsilon, F_\varepsilon, p_t) = \mathcal{E}_{\text{tot}}(f_t, F_t, p_t) + \varepsilon \int_{\Omega_0} (\nabla\varphi(F_t) \cdot \mathcal{L} + \psi(p_t)\check{\nabla}\phi(F_t^d) \rhd \mathcal{L}^p) \, dV_0$$

$$- \varepsilon \left(\int_{\Omega_0} b_t \cdot v \, dV_0 + \int_{\partial\Omega_0} s_t \cdot v \, dA_0 \right) + o(\varepsilon). \quad (6\text{-}25)$$

From the stationarity condition (3-17), for $\mathcal{L}^d = 0$ we have $\mathcal{L} = \nabla v$, and the identification (3-19) of $\nabla\varphi(F_t)$ with the Piola stress T_t follows. Moreover, by elimination of all terms in v and subsequent localization, after recalling that $R_0^d = I$ implies $L^d = D^d$ we get the counterpart of the plastic stationarity condition (3-22),

$$T_t^D \cdot \mathcal{D}^d \leq \psi(p_t)\check{\nabla}\phi(tD^d) \rhd \mathcal{D}^d, \quad (6\text{-}26)$$

to be satisfied by all $\mathcal{D}^d$ belonging to a given subset $\mathscr{S}_0^d$ of $\mathscr{S}^D$. With the same motivation used to establish (3-23) in the unconstrained case, for $\mathcal{D}^d = D^d$ inequality (6-26) reduces to the equality

$$T_t^D \cdot D^d = \psi(p_t)\check{\nabla}\phi(tD^d) \rhd D^d. \quad (6\text{-}27)$$

6.2.4. *The evolution law for the hydrostatic pressure.* From (3-19) we still have the incremental stress-strain relation (5-2), with $\mathbb{C}$ the elastic tensor $\nabla^2\varphi(I)$. For this tensor, the identities

$$\mathbb{C}[A] = \mathbb{C}[A^H + A^D] = (\mathbb{C}[A])^H + (\mathbb{C}[A])^D \quad (6\text{-}28)$$

hold for all A in $\mathscr{L}$. For simplicity, we focus on the special case in which $\mathbb{C}$ maps the hydrostatic tensors into hydrostatic tensors[39]

$$\mathbb{C}[A^H] = (\mathbb{C}[A])^H. \quad (6\text{-}29)$$

Then (6-28) implies $\mathbb{C}[A^D] = (\mathbb{C}[A])^D$; that is, $\mathbb{C}$ also maps the deviators into deviators. Since the hydrostatic tensors are those parallel to the identity tensor I, from (6-29) and the linearity of $\mathbb{C}$ it follows that

$$\mathbb{C}[A^H] = 3kA^H, \quad (6\text{-}30)$$

with k a positive material constant, called the *bulk modulus.*

For $N_0 \in \mathscr{S}^D$, in the incremental stress-strain relation (5-18) we have

$$\delta T = \mathbb{C}[\nabla u^H] + \mathbb{C}[\nabla u^D - \lambda^d N_0] = 3k\nabla u^H + \mathbb{C}[\nabla u^D - \lambda^d N_0], \quad (6\text{-}31)$$

and by effect of assumption (6-29) this equation splits into two parts

$$\delta T^H = 3k\nabla u^H, \qquad \delta T^D = \mathbb{C}[\nabla u^D - \lambda^d N_0]. \quad (6\text{-}32)$$

---

[39]This relation is satisfied by all orthotropic materials with cubic symmetry, and in particular by all isotropic materials; see, e.g., [Gurtin 1972, §26].

For the hydrostatic part, recalling the definition (6-22) of the hydrostatic pressure,

$$\delta p = -\tfrac{1}{3} I \cdot \delta T = -\tfrac{1}{3} I \cdot \delta T^H = -kI \cdot \nabla u^H = -kI \cdot \nabla u. \qquad (6\text{-}33)$$

This is the evolution law for the hydrostatic pressure, according to which p is determined by the hydrostatic part of the macroscopic deformation. We now proceed to the determination of the evolution law for the deviatoric stress.

6.2.5. *Yield conditions, elastic ranges, and admissible stresses.* At $t = 0$, from (6-26) we have

$$T_0^D \cdot \mathcal{D}^d \le \psi(p_0)\check{\Phi}_0 \triangleright \mathcal{D}^d, \qquad (6\text{-}34)$$

with $\check{\Phi}_0$ as in (5-5). This inequality says that T_0^D, which is symmetric by the indifference condition (4-12)$_1$, belongs to a family of closed half-spaces $\mathcal{H}_{p_0}^N$ of $\mathscr{S}^D$, with normals N in $\mathscr{S}_0^d$ and with distance from the origin proportional to $\psi(p_0)$. The intersection $\mathcal{C}_{p_0} = \psi(p_0)\mathcal{C}_0$ of all $\mathcal{H}_{p_0}^N$ is the *elastic range associated with the pressure p_0*, and the condition

$$T_0^D \in \psi(p_0)\,\mathcal{C}_0 \qquad (6\text{-}35)$$

is the *yield condition for the pressure p_0*. Thus, with the initial deformation (ι, I) is associated a family of elastic ranges $p \mapsto \mathcal{C}_p = \psi(p)\mathcal{C}_0$, in which each $\mathcal{C}_p$ is a homothetic transformation of $\mathcal{C}_0$ with center at the origin and ratio $\psi(p)$.

In the space $\mathscr{S}$, each $\mathcal{C}_p$ belongs to the hyperplane lying at the (signed) distance p from $\mathscr{S}^D$. Then the $\mathcal{C}_p$ are pairwise disjoint subsets of $\mathscr{S}$. Their union $\mathcal{K}_0$ is the *set of all admissible stresses* for the initial deformation.[40]

6.2.6. *Flow rule and consistency condition.* At $t = 0$, (6-27) reduces to

$$T_0^D \cdot D^d = \psi(p_0)\check{\Phi}_0 \triangleright D^d. \qquad (6\text{-}36)$$

Since both D^d and $\mathcal{C}_{p_0}$ are in the space of the symmetric deviators, this equation tells us that T_0^D is a boundary point of $\mathcal{C}_{p_0}$ relative to this space, and that D^d belongs to the normal cone of $\mathcal{C}_{p_0}$ at T_0^D. Thus, the normality rule

$$D^p = \begin{cases} \lambda^d N_0, \lambda^d \ge 0 & \text{if } T_0^D \in \partial\mathcal{C}_{p_0}, \\ 0 & \text{if } T_0^D \in \mathcal{C}_{p_0} \setminus \partial\mathcal{C}_{p_0} \end{cases} \qquad (6\text{-}37)$$

is established.

---

[40]The idea of considering as the set of admissible stresses a one-parameter family of elastic ranges is not new. Families of yield regions enclosed by a *bounding surface* [Dafalias and Popov 1975] were used to describe the work-hardening of metals [Phillips and Sierakowski 1965] and their response to cyclic loading [Mróz 1969]. Later, the same idea was applied to geomaterials by Vermeer and de Borst [1984] and by Roscoe and coworkers [Roscoe and Burland 1968; Roscoe and Poorooshasb 1963]. See also [Ziegler and Wehrli 1987, p. 223].

According to this rule, the plastic strain rate is zero at all interior points of $\mathcal{C}_{p_0}$. Since the interior points of $\mathcal{C}_{p_0}$ are also interior points of the set $\mathcal{K}_0$ of the admissible stresses, one is tempted to consider $\mathcal{K}_0$ as the real elastic range. This choice, which has been actually made in several plasticity models, is incompatible with the normality law, because the direction N_0 is normal to $\mathcal{C}_{p_0}$ but not, in general, to $\mathcal{K}_0$. This inconvenience has been circumvented by rejecting the normality law, that is, by adopting a nonassociated flow rule. Here we have shown that this choice contrasts with the result provided by energy minimization.[41]

For $\delta p \neq 0$, D^p cannot be zero because (6-27) must be satisfied at $t > 0$. From this equation, with the expansions

$$T_t^D = T_0^D + t\delta T^D + o(t),$$
$$\check{\nabla}\phi(tD^p) = \check{\Phi}_0 + t\,\check{\mathbb{K}}\{D^p\} + o(t), \qquad (6\text{-}38)$$
$$\psi(p_t) = \psi(p_0) + \psi'(p_0)\,\delta p + o(t),$$

we get the separate conditions

$$T_0^D \cdot D^d = \psi(p_0)\check{\Phi}_0 \triangleright D^d,$$
$$\delta T^D \cdot D^d = \psi(p_0)\check{\mathbb{K}}\{D^d\} \triangleright D^d + \psi'(p_0)\,\delta p\check{\Phi}_0 \triangleright D^p. \qquad (6\text{-}39)$$

The first equation is (6-34). Substituting it into the second equation, recalling the normality rule $D^d = \lambda^d N_0$ and setting

$$\kappa(p_0) = \psi'(p_0)/\psi(p_0), \qquad (6\text{-}40)$$

we get the *consistency condition for isochoric plasticity*

$$\lambda^d\big(\delta T^D \cdot N_0 - \lambda^d\psi(p_0)\check{\mathbb{K}}\{N_0\} \triangleright N_0 - \kappa(p_0)(T_0^D \cdot N_0)\,\delta p\big) = 0. \qquad (6\text{-}41)$$

6.2.7. *The incremental response law and the hardening rule.* Comparing the consistency condition with (6-32)$_2$ multiplied by N_0 and recalling the expression (6-33) of δp, we find

$$\lambda^d(N_0, p_0) = \frac{\langle\mathbb{C}[N_0] \cdot \nabla u^D\rangle - \kappa(p_0)(T_0^D \cdot N_0)\,\delta p}{\mathbb{C}[N_0] \cdot N_0 + \psi(p_0)\check{\mathbb{K}}\{N_0\} \triangleright N_0}. \qquad (6\text{-}42)$$

---

[41] Another inconvenience of the nonassociated flow rules emerges from a comparison between isochoric plasticity and kinematic hardening. In both cases there is a family $t \mapsto \mathcal{C}_p$ of elastic ranges, and their union is the set $\mathcal{K}_0$ of the admissible stresses. However, in kinematic hardening the regions $\mathcal{C}_p$ are not pairwise disjoint. Then a boundary point of a region can be an interior point of another region, and therefore of $\mathcal{K}_0$. It is then inconceivable to restrict the plastic stretching to the boundary points of $\mathcal{K}_0$, since a nonnull plastic stretching is allowed at the boundary points of all $\mathcal{C}_p$. On the contrary, this is perfectly conceivable in isochoric plasticity, in which the boundary points of all $\mathcal{C}_p$ are also boundary points of $\mathcal{K}_0$. Thus, the choice of a nonassociated flow rule is possible in the second case but not in the first. Since there is no reason for choosing different flow rules in the two cases, the motivations for choosing nonassociated flow rules in the second case are not clear.

Then the incremental response law is (6-32)$_2$ with λ^d as above.

At loading, $\mathbb{C}[N_0] \cdot \nabla u^D > 0$, the response is *hardening*, *softening*, or *perfectly plastic* if the product $\delta T^D \cdot N_0$ is positive, negative, or zero, respectively. From (6-41) and (6-42) we have

$$\delta T^D \cdot N_0 = h(N_0, p_0)\mathbb{C}[N_0] \cdot \nabla u^D + (1 - h(N_0, p_0))\kappa(p_0)(T_0^D \cdot N_0)\delta p, \quad (6\text{-}43)$$

with the *hardening modulus*

$$h(N_0, p_0) = \frac{\psi(p_0)\,\check{\mathbb{K}}\{N_0\} \rhd N_0}{\mathbb{C}[N_0] \cdot N_0 + \psi(p_0)\,\check{\mathbb{K}}\{N_0\} \rhd N_0}. \quad (6\text{-}44)$$

Note that, assuming both the denominator and $\mathbb{C}[N_0] \cdot N_0$ are positive, we have $h(N_0, p_0) < 1$.

From (6-43) we see that $h(N_0, p_0)$ measures the evolution of the elastic range $\mathcal{C}_{p_0}$ at constant pressure, that is, under a purely deviatoric loading $\mathbb{C}[N_0] \cdot \nabla u^D > 0$. In this case the sign of $h(N_0, p_0)$, which assuming $\psi(p_0) > 0$ is the same of the sign of $\check{\mathbb{K}}\{N_0\} \rhd N_0$, is positive, negative, or zero when $\mathcal{C}_{p_0}$ expands, shrinks, or remains unchanged. In the three cases the response is hardening, softening, or perfectly plastic, respectively.

Under a purely hydrostatic loading, that is, for $\nabla u^D = 0$, from (6-43) we see that the pressure change δp determines a change δT^D of the deviatoric stress, such that

$$\delta T^D \cdot N_0 = (1 - h(N_0, p_0))\,\kappa_0(p_0)(T_0^D \cdot N_0)\delta p. \quad (6\text{-}45)$$

This is the change of T^D required to keep the total stress T on the boundary of the elastic range during its evolution from $\mathcal{C}_{p_0}$ to $\mathcal{C}_{p_0+\delta p}$. Since $h(N_0, p_0) < 1$ and $T_0^D \cdot N_0 > 0$ by the dissipation inequality (5-15), $\delta T^D \cdot N_0$ has the same sign of δp if $\kappa_0(p_0) > 0$ and the opposite sign if $\kappa_0(p_0) < 0$. In the first case, a positive δp produces a dilatation of the elastic range, and in the second case it produces a contraction.

6.3. *Isotropic isochoric plasticity.* A special case of isochoric plasticity is the case in which the plastic energy density (6-21) depends only on the modulus of L^d:

$$\tilde{\phi}(tL^d, p_t) = \phi(t|L^d|)\psi(p_t). \quad (6\text{-}46)$$

The linearized forms of this energy corresponding to the Drucker–Prager and to the Mises yield conditions are helpful for understanding some basic differences between the plastic behaviors of metals and of geomaterials.

6.3.1. *The yield condition.* For an energy of the form (6-46) and with ϕ such that $\check{\Phi}_0 \rhd \mathcal{D}^d = |\mathcal{D}^d|$, the plastic stationarity condition (6-34) reduces to

$$T_0^D \cdot \mathcal{D}^p \le \psi(p_0)|\mathcal{D}^p|, \quad (6\text{-}47)$$

and for any direction $N = \mathcal{D}^p / |\mathcal{D}^p|$ in $\mathscr{S}^{dD}$ we have

$$T_0^D \cdot N \le \psi(p_0). \tag{6-48}$$

This inequality tells us that the generating half-spaces $\mathcal{H}_{p_0}^N$ of the elastic range $\mathcal{C}_{p_0}$ are all placed at the same distance $\psi(p_0)$ from the origin. If $\mathscr{S}^{dD} = \mathscr{S}^D$, then $\mathcal{C}_{p_0}$ is the ball of $\mathscr{S}^D$ centered at the origin and with radius $\psi(p_0)$, and at every boundary point T_0^D of $\mathcal{C}_{p_0}$ the normal N_0 is parallel to T_0^D:

$$T_0^D = \psi(p_0)N_0. \tag{6-49}$$

Then, by the normality rule, T_0^D has the same direction as $D^p = \lambda^d N_0$.

6.3.2. *The Drucker–Prager and the Mises conditions.* For the energies of the form (6-46), consider the special case in which

$$\phi(|L^d|) = |L^d|, \qquad \psi(p) = \alpha p + \beta, \quad \alpha, \beta > 0. \tag{6-50}$$

For such energies the radius $\psi(p)$ of $\mathcal{C}_p$ vanishes for $p = -\beta/\alpha$. Then the set $\mathcal{K}_0$ of the admissible stresses

$$\mathcal{K}_0 = \{T \in \mathscr{S} \mid T = T^D - pI, \; p \ge -\beta/\alpha, \; |T^D| \le \alpha p + \beta\} \tag{6-51}$$

is a circular cone of $\mathscr{S}$ with vertex at $T = (\beta/\alpha)I$ and with axis on the hydro-static axis. This is the set of the admissible stresses of the *Drucker–Prager yield condition*.

In the limit case $\alpha = 0$, the radius becomes equal to β, and $\mathcal{K}_0$ becomes the cylinder

$$\mathcal{K}_0 = \{T \in \mathscr{S} \mid T = T^D - pI, \; p \in \mathbb{R}, \; |T^D| \le \beta\}, \tag{6-52}$$

which is the set of the admissible stresses of the *Mises yield condition*.

6.3.3. *The plastic behavior of metals and geomaterials.* When comparing the be-havior of metals and geomaterials, one sees that the latter exhibit a much lower strength in tension than in compression, while for metals the two strengths are of the same order of magnitude. Moreover, under increasing pressure the elastic range enlarges in geomaterials, and remains almost constant in metals.

These differences are captured by isotropic energies of the form (6-50). Indeed, in the Drucker–Prager cone the hydrostatic stress in tension cannot exceed the value $-\beta/\alpha$ attained at the vertex of the cone but is unbounded in compression, and the size of the elastic range increases with p_R. In Mises's cylinder, both tensile and compressive hydrostatic stresses are unbounded, and the size of the elastic range does not depend on the pressure. Therefore, the cone and the cylinder seem to be well suited to describe, at least in a first approximation, the behavior of geomaterials and of metals, respectively.

That the Mises criterion is appropriate to metals is almost universally recognized.[42] Large consensus was also initially met on adopting yield conditions of the Drucker–Prager type for geomaterials. In particular, favored by its analogy with Coulomb's theory of friction, the Mohr–Coulomb yield condition[43] was applied extensively in soil mechanics [Lubliner 1990, §6.1.3].[44] These models were coupled with normality in the deviatoric plane, and since the elastic range was identified with the set of the admissible stresses, this led to the adoption of nonassociated flow rules. In the present paper, in line with some earlier proposals [Ziegler and Wehrli 1987, p. 223; Srinivasa 2010], the associated character of these flow rules has been recovered regarding the hydrostatic pressure as an extra state variable. In this way, the assumption of normality in the deviatoric space is fully legitimated from the variational viewpoint.

Unfortunately, experimental evidence turned against the isochoricity assumption, since most geomaterials exhibit a form of inelastic volume change called *dilatancy*,[45] which fits neither normality with respect to the deviatoric plane nor normality with respect to the set of the admissible stresses.[46]

6.4. *The Cam-clay model.*

An experiment-based model expressly conceived to describe the dilatancy of soils is the *Cam-clay model* [Roscoe et al. 1958; Roscoe and Poorooshasb 1963].[47] Here we show that this is a special case of the dilatational model described in Section 6.1.3.

6.4.1. *The plastic energy.*

In the Cam-clay model we keep the irrotationality condition $R_0^d = I$ by which, due to condition (4-15), the plastic energy ϕ reduces to a function of the plastic stretching D^d. The assumed form of ϕ is

$$\phi(D^d) = \phi(\tilde{p}, \tilde{q}), \tag{6-53}$$

where

$$\tilde{p} = -\tfrac{1}{3} I \cdot D^d = -\tfrac{1}{3} I \cdot D^{dH}, \qquad \tilde{q} = |D^d + \tilde{p} I| = |D^{dD}|. \tag{6-54}$$

---

[42]"…Most modern discussions of plasticity (of metals) are based on generalizations and structural variations of the theory of Lévy, Mises, Prandtl, and Reuss…" [Gurtin et al. 2010, §76.1].

[43]For this condition, which can be regarded as a variant of the Drucker–Prager condition, see, e.g., [Lubliner 1990, §3.3.3].

[44]For other models of the Drucker–Prager type, see [Ziegler and Wehrli 1987, §VII.A, §VII.B].

[45]For plastic volume changes in geomaterials, see, e.g., [Lubliner 1990; Vermeer and de Borst 1984]. For metals, unexpected plastic volume changes were revealed by the experiments reported in [Wilson 2002].

[46]The volume increase due to normality with respect to $\mathcal{K}_0$ is measured by the *friction angle*, and the volume increase due to normality with respect to $\mathcal{C}_p$ is zero. The experiments show that the angle which measures the volume increase due to dilatancy has an intermediate value which is far both from the friction angle and from zero; see, e.g., [Lubliner 1990, §6.1.3], or [Vermeer and de Borst 1984, §2].

[47]The model was modified in [Roscoe and Burland 1968].

The dependence on the modulus of D^{dD} describes an isotropic response to the deviatoric strain rates.

6.4.2. *The yield condition.* For an energy of the form (6-53), the elastic range $\mathcal{C}_0$ is a solid of revolution with axis on the hydrostatic axis. It is assumed that $\mathcal{C}_0$ is represented in the plane $(\tilde{p}, \tilde{q})$ by the line[48]

$$\tilde{q} = \chi(\tilde{p}), \qquad \tilde{p} \in (0, p^\dagger), \tag{6-55}$$

with

$$\chi(\tilde{p}) \in (0, q^\dagger), \quad \chi(0) = \chi(p^\dagger) = 0, \quad p^\dagger, q^\dagger < +\infty. \tag{6-56}$$

Let us determine the homogeneous dissipation potential ϕ^h associated with this elastic range.[49] Let e^p, e^q be unit vectors in the directions of the axes $\tilde{p}$ and $\tilde{q}$. At a boundary point $(\tilde{p}, \tilde{q})$ of $\mathcal{C}_0$, the unit normal is

$$n(\tilde{p}) = \frac{e^q - \chi'(\tilde{p}) e^p}{\sqrt{1 + \chi'^2(\tilde{p})}}, \tag{6-57}$$

and

$$d(\tilde{p}) = (\tilde{p} e^p + \chi(\tilde{p}) e^q) \cdot n(\tilde{p}) = \frac{\chi(\tilde{p}) - \tilde{p}\chi'(\tilde{p})}{\sqrt{1 + \chi'^2(\tilde{p})}} \tag{6-58}$$

is the distance of $(\tilde{p}, \tilde{q})$ from the origin, measured in the direction of $n(\tilde{p})$. This is the value taken by the homogeneous potential ϕ^h at the boundary points of $\mathcal{C}_0$:[50]

$$\phi^h(\tilde{p}, \tilde{q}) = \phi^h(\tilde{p}, \chi(\tilde{p})) = d(\tilde{p}). \tag{6-59}$$

6.4.3. *Flow rule and hardening rule.* The flow rule assumed in the Cam-clay model is the normality rule. Accordingly, the plastic stretching at the boundary point $(\tilde{p}, \chi(\tilde{p}))$ of $\mathcal{C}_0$ has the direction $n(\tilde{p})$. Since $n(\tilde{p})$ has a nonnull component in the direction e^p, this implies that the plastic stretching is not isochoric. In particular, the volume change is positive (*dilatancy*) at points at which $\chi'(\tilde{p})$ is negative.

The homogeneous potential (6-59) corresponds to a perfectly plastic response. To get a hardening response, it is necessary to add to ϕ^h a nonhomogeneous part. For the quadratic potential

$$\phi(\lambda n(\tilde{p})) = (\lambda + \tfrac{1}{2}k\lambda^2)d(\tilde{p}), \quad \lambda \geq 0, \tag{6-60}$$

we have

$$\check{\nabla}\phi(0) \triangleright n(\tilde{p}) = d(\tilde{p}), \qquad \check{\nabla}^2\phi(0)\{n(\tilde{p})\} \triangleright n(\tilde{p}) = kd(\tilde{p}). \tag{6-61}$$

---

[48]This line is a parabolic arc in the original Cam-clay model [Roscoe et al. 1958; Roscoe and Poorooshasb 1963] and a half-ellipse in the modified model [Roscoe and Burland 1968].

[49]Here we take advantage of the one-to-one correspondence between elastic ranges and homogeneous dissipation potentials; see Footnote 31.

[50]See Section 5.1.

Then denoting by $\tilde{t}_0 = (\tilde{p}, \tilde{q})$ and by $\delta\tilde{t}$ the representatives of T_0 and δT on the plane $(\tilde{p}, \tilde{q})$, from the plastic stationarity condition (5-12) and from the consistency condition (5-17), we get

$$\tilde{t}_0 \cdot n(\tilde{p}) = d(\tilde{p}), \qquad \delta\tilde{t} \cdot n(\tilde{p}) = k\lambda^d d(\tilde{p}), \qquad (6\text{-}62)$$

respectively. That is, for every direction n the generating half-planes $\mathcal{H}_0^n$ of $\mathcal{C}_0$ undergo a translation proportional to the distance $d(\tilde{p})$ of the boundary point $\tilde{t}_0$ from the origin. This defines a new elastic range, obtained from $\mathcal{C}_0$ by a homothetic transformation with center at the origin and ratio $2k\lambda^d$. This is the evolution law for the elastic range currently assumed in the literature.[51] Clearly, a positive k corresponds to hardening and a negative k corresponds to softening.

7. Crystal plasticity

In materials with a crystalline structure, the plastic deformation consists of microslips occurring along some preferred directions on some preferred slip planes. A *plastic slip* is a plastic deformation of the form

$$L^d = \lambda^d s \otimes m, \quad \lambda^d \geq 0, \qquad (7\text{-}1)$$

where m is the unit normal to the slip plane, λ^d is the slip intensity, and s, the slip direction, is a unit vector in the slip plane. The pair (s, m) is a *slip system*, and the tensor $(s \otimes m)$ is the corresponding *Schmid tensor*. By the orthogonality of s and m, L^d is a deviator. Then the plastic deformation is isochoric; that is, crystal plasticity is a special case of isochoric plasticity.[52]

Below, I consider first the case of a single slip system, and then the case of a finite number of slip systems. Finally, I describe the *two-level shear* model based on the theory of structured deformations, in which the macroscopic features of plastic response can be reproduced without the use of nonsmooth potentials.

7.1. The single-slip model. The plastic slip (7-1) is a deformation of the form (4-3), $L^d = R_0^d D^d$, with[53]

$$R_0^d m = s, \qquad D^d = \lambda^d m \otimes m, \quad \lambda^d \geq 0. \qquad (7\text{-}2)$$

What distinguishes this model from the preceding ones is precisely that the constitutive assumption $R_0^d = I$ is replaced by an assumption on the direction of L^d

---

[51] See, e.g., [Schofield and Wroth 1968].

[52] In crystal plasticity the energy is usually assumed to be independent of the hydrostatic pressure, that is, the function $\psi(p)$ in (6-21) is taken equal to one. For basic reference to the plasticity of crystals, see, e.g., [Gurtin et al. 2010, §102.1].

[53] In principle, R_0^d may be any rotation which maps m into s. For definiteness, one may add the prescription $R_0^d s = -m$, which characterizes R_0^d as the rotation of amount $\pi/2$ about an axis perpendicular to both s and m.

dictated by the geometry of the crystal.[54] Due to this peculiarity, the procedure for the determination of the incremental response law slightly departs from the path followed in the previous sections.

7.1.1. *The plastic stationarity condition.* From (7-1) and (7-2) and from the indifference conditions (4-14) and (4-15), neglecting the terms of order $o(t)$, for the plastic strain energy we have

$$\phi(F_t^d) = \phi(tL^d) = \phi(t\lambda^d s \otimes m) = \phi(t\lambda^d R_0^d m \otimes m) = \phi(t\lambda^d m \otimes m), \quad (7\text{-}3)$$

and since m is fixed we may set

$$\phi(F_t^d) = \phi_m(t\lambda^d). \quad (7\text{-}4)$$

Thus, the only admissible plastic stretchings are those directed as $m \otimes m$, and the plastic energy is a function, depending on m, of their modulus $t\lambda^d$. Then from the definition of directional derivative we have

$$\nabla\phi(F_t^d) \triangleright s \otimes m = \lim_{\varepsilon \to 0^+} \frac{\phi((t\lambda^d + \varepsilon)\, s \otimes m) - \phi(t\lambda^d s \otimes m)}{\varepsilon}$$

$$= \lim_{\varepsilon \to 0^+} \frac{\phi_m(t\lambda^d + \varepsilon) - \phi_m(t\lambda^d)}{\varepsilon} = \phi_m'(t\lambda^d), \quad (7\text{-}5)$$

and the plastic stationarity condition (3-22) takes the form

$$\tau_t \le \phi_m'(t\lambda^d), \quad (7\text{-}6)$$

where the *resolved shear stress*

$$\tau_t = T_t \cdot s \otimes m = T_t m \cdot s \quad (7\text{-}7)$$

is the tangential component in the direction s of the stress vector $T_t m$ acting on the slip plane.

7.1.2. *Yield condition, flow rule, and consistency condition.* According to (7-1) and (7-2), the only direction allowed for the plastic stretching is $m \otimes m$. That is, the set of the admissible plastic stretchings is the half-line

$$\mathcal{S}_0^d = \{\lambda m \otimes m \mid \lambda \ge 0\}. \quad (7\text{-}8)$$

Then there is only one generating half-space $\mathcal{H}_0^{m \otimes m}$. It coincides with the elastic range $\mathcal{C}_0$, and the flow rule trivially says that the plastic stretching is directed as $m \otimes m$. In view of the expansions

$$\tau_t = \tau_0 + t\delta\tau + o(t), \qquad \phi_m'(t\lambda^d) = \phi_m'(0) + t\lambda^d \phi_m''(0) + o(t), \quad (7\text{-}9)$$

---

[54]Here only the case of L^d directed as $s \otimes m$ is discussed. A more extended analysis in the context of the multiplicative decomposition was made by Reina and coauthors [Reina and Conti 2014; Reina et al. 2016].

from the plastic stationarity condition at $t = 0$ we have the inequality

$$\tau_0 \leq \phi'_m(0), \tag{7-10}$$

which shows that $\phi'_m(0)$ is the distance of the boundary of $\mathcal{C}_0$ from the origin. It also tells us that no plastic slip occurs as long as the resolved shear stress is below the *activation threshold* $\phi'_m(0)$. When this threshold is attained, a plastic slip may occur in the direction s. In this case (7-10) is satisfied as an equality, and from (7-6) and (7-9) we have

$$\delta\tau \leq \lambda^d \phi''_m(0). \tag{7-11}$$

This inequality becomes an equality when $\lambda^d > 0$. Then we have the consistency condition

$$\lambda^d (\delta\tau - \lambda^d \phi''_m(0)) = 0. \tag{7-12}$$

7.1.3. *Incremental response law and hardening rule.* From (7-12) and from the incremental stress-strain relation (5-2)

$$\delta\tau = \delta T \cdot (s \otimes m) = \mathbb{C}[\nabla u - \lambda^d s \otimes m] \cdot (s \otimes m), \tag{7-13}$$

by elimination of $\delta\tau$ we get

$$\lambda^d = \frac{\langle \mathbb{C}[\nabla u] \cdot (s \otimes m)\rangle}{\mathbb{C}[s \otimes m] \cdot (s \otimes m) + \phi''_m(0)}, \tag{7-14}$$

and substitution into (7-13) provides the incremental response law. In particular, for $\lambda^d > 0$ we have

$$\delta\tau = h_m \mathbb{C}[\nabla u] \cdot (s \otimes m), \tag{7-15}$$

with the hardening modulus

$$h_m = \frac{\phi''_m(0)}{\mathbb{C}[s \otimes m] \cdot (s \otimes m) + \phi''_m(0)}. \tag{7-16}$$

The numerator $\phi''_m(0)$ is positive if ϕ_m is strictly convex. On the contrary, nothing can be said a priori about the positiveness of the denominator. Indeed, because of the indifference condition (4-13), one can control only the restriction of $\mathbb{C}$ to the symmetric tensors. For all previous models this was enough, because only this restriction appears in the expressions (5-22) and (6-42) of λ^d. Here, from the indifference condition (4-13) we have

$$\mathbb{C}[s \otimes m] \cdot (s \otimes m) = \mathbb{C}[(s \otimes m)^S] \cdot (s \otimes m)^S$$
$$+ 2(s \otimes m)^W T_0 \cdot (s \otimes m)^S + (s \otimes m)^W T_0 \cdot (s \otimes m)^W. \tag{7-17}$$

The two terms involving T_0 can be transformed into

$$-T_0 \cdot (s \otimes m)^W (2(s \otimes m)^S + (s \otimes m)^W) = \tfrac{1}{4} T_0 \cdot (3m \otimes m - s \otimes s). \tag{7-18}$$

Denoting by c_0 the smallest eigenvalue of $\mathbb{C}$ restricted to $\mathscr{S}$ and setting

$$\sigma_m = T_0 \cdot (m \otimes m), \qquad \sigma_s = T_0 \cdot (s \otimes s), \tag{7-19}$$

we get

$$\mathbb{C}[s \otimes m] \cdot (s \otimes m) \geq \tfrac{1}{2}c_0 + \tfrac{3}{4}\sigma_m - \tfrac{1}{4}\sigma_s. \tag{7-20}$$

Therefore, $\mathbb{C}[s \otimes m] \cdot (s \otimes m)$ is positive if $c_0 > 0$ and σ_m and σ_s are small with respect to c_0, and if $c_0 > 0$, the normal stress σ_m acting on the slip plane is tensile, and the in-plane normal stress σ_s is compressive.

7.2. *The multislip model.* A *multislip system* is defined as a finite set of slip systems (s^α, m^α).[55] For each of them the plastic deformation has the form

$$L^{d\alpha} = \lambda^{d\alpha} s^\alpha \otimes m^\alpha, \quad \lambda^{d\alpha} \geq 0, \tag{7-21}$$

and the corresponding energy is

$$\phi^\alpha(L^{d\alpha}) = \phi^\alpha(\lambda^{d\alpha} m^\alpha \otimes m^\alpha) = \phi_{m^\alpha}(\lambda^{d\alpha}). \tag{7-22}$$

The total plastic deformation and the total energy are

$$L^p = \sum_\alpha \lambda^{d\alpha} s^\alpha \otimes m^\alpha, \qquad \phi(L^d) = \sum_\alpha \phi_{m^\alpha}(\lambda^{d\alpha}), \tag{7-23}$$

respectively. The gradient of ϕ is the homogeneous map which with every $\mathcal{L}^d = \sum_\beta \mu^\beta s^\beta \otimes m^\beta$ associates the number

$$\nabla\phi(F_t^d) \triangleright \mathcal{L}^d = \nabla\phi\left(\sum_\alpha \lambda^{d\alpha} s^\alpha \otimes m^\alpha\right) \triangleright \sum_\beta \mu^\beta s^\beta \otimes m^\beta$$
$$= \sum_\alpha \mu^\alpha \phi'_{m^\alpha}(\lambda^{d\alpha}), \tag{7-24}$$

with the last identification preformed proceeding as in (7-5). The plastic stationarity condition (3-22) then takes the form

$$\sum_\alpha \mu^\alpha \tau_t^\alpha \leq \sum_\alpha \mu^\alpha \phi'_{m^\alpha}(t\lambda^{d\alpha}), \quad \mu^\alpha \geq 0, \tag{7-25}$$

where

$$\tau_t^\alpha = T_t \cdot (s^\alpha \otimes m^\alpha) \tag{7-26}$$

is the resolved shear stress for the slip system (s^α, m^α). At $t = 0$, taking all μ^α equal to zero except one, we get n inequalities

$$\tau_0^\alpha \leq \phi'_{m^\alpha}(0), \tag{7-27}$$

---

[55]This model is based on Koiter's model of *singular yield surfaces* [1953]. See also [Martin and Reddy 1993].

one for each slip system. Each of them defines a half-space $\mathcal{H}_0^{N^\alpha}$ of $\mathscr{S}^D$ with normal $N^\alpha = m^\alpha \otimes m^\alpha$ and with boundary at the distance $\phi'_{m^\alpha}(0)$ from the origin. The intersection of the $\mathcal{H}_0^{N^\alpha}$ is the elastic range $\mathcal{C}_0$. It is a polyhedral convex subset of $\mathscr{S}^D$, whose faces have the normals N^α and whose vertices have normal cones consisting of positive combinations of the N^α.

For each α, the expansions (7-9) hold. When for some τ_0^α the activation threshold $\phi'_{m^\alpha}(0)$ is reached, by the consistency condition in analogy with (7-12) we have

$$\lambda^{d\alpha}(\delta\tau^\alpha - \lambda^{d\alpha}\phi''_{m^\alpha}(0)) = 0. \tag{7-28}$$

If T_0 is on a face of $\mathcal{C}_0$, this occurs only for the α corresponding to that face. If T_0 is on a vertex of $\mathcal{C}_0$, this occurs for the α corresponding to the faces which concur at that vertex. Denoting by $\hat{\alpha}(T_0)$ the set of such α, the normal cone at T_0 is

$$D^d = \sum_{\alpha \in \hat{\alpha}(T_0)} \lambda^{d\alpha} N^\alpha, \tag{7-29}$$

and the incremental relation (5-2) takes the form

$$\delta T = \mathbb{C}[\nabla u] - \sum_{\alpha \in \hat{\alpha}(T_0)} \lambda^{d\alpha}\mathbb{C}[N^\alpha]. \tag{7-30}$$

Then for all β in $\hat{\alpha}(T_0)$,

$$\delta\tau^\beta = \delta T \cdot (s^\beta \otimes m^\beta) = \left(\mathbb{C}[\nabla u] - \sum_{\alpha \in \hat{\alpha}(T_0)} \lambda^{d\alpha}\mathbb{C}[s^\alpha \otimes m^\alpha]\right) \cdot (s^\beta \otimes m^\beta). \tag{7-31}$$

Together with the constraints $\lambda^{d\alpha} \geq 0$ and the complementarity conditions (7-28), these equations form a linear complementarity problem of the dimension of $\hat{\alpha}(T_0)$. In particular, if T_0 belongs to the relative interior of a face of $\mathcal{C}_{p_0}$, this dimension is one and we are back to the one-dimensional problem of the previous sections.

If the dimension is larger than one, the existence and uniqueness of the solution is guaranteed if the matrix $\{\mathbb{C}[s^\alpha \otimes m^\alpha] \cdot (s^\beta \otimes m^\beta)\}$ is positive definite. In this case the problem can be solved with the Gauss–Seidel iterative method or with any other nonlinear programming algorithm. But, like in the case of a single slip system, the positive definiteness is ensured only for suitable values of the initial stress T_0.

7.3. *Periodic energies and two-level shears.* In single crystals, a relative translation of an atomic unit along a slip plane maps the two halves of a crystal into a placement energetically indistinguishable from the initial one.[56] This suggests the

---

[56]See, e.g., [Gurtin et al. 2010, Figure 102.2].

use of periodic energies. In general, the energy ϕ is assumed to be of the form

$$\phi = \psi + \phi^{nh}, \tag{7-32}$$

with ψ a smooth periodic potential and ϕ^{nh} a nonhomogeneous nonsmooth potential. Here we consider the simple case of $\phi^{nh} = 0$.[57]

A *two-level shear* is a deformation process $t \mapsto (f_t, F_t)$ in which the macroscopic deformation f_t is a simple shear

$$f_t(x) = x + \bar{\gamma} t x \, s \otimes m, \quad \bar{\gamma} > 0, \tag{7-33}$$

and the deformation due to the disarrangements[58] is a single slip

$$F_t^d(x) = \gamma_t^d s \otimes m. \tag{7-34}$$

When ϕ reduces to a smooth potential ψ, the directional derivative $\check{\nabla}\phi$ reduces to the ordinary derivative $\nabla\psi$. Moreover, in the virtual strain rates $\mathcal{L}^d = \alpha s \otimes m$ the multiplier α is not anymore constrained to be positive. Then the stationarity condition (3-22) reduces to the equality

$$T_t \cdot (s \otimes m) = \nabla\psi(\gamma_t^d s \otimes m) \cdot (s \otimes m). \tag{7-35}$$

The left-hand side is the resolved shear stress (7-7). For the right-hand side, using the indifference condition (7-3) we define

$$\psi(F_t^d) = \psi(\gamma_t^d m \otimes m) = \psi_m(\gamma_t^d), \tag{7-36}$$

and proceeding as in (7-5) we find that the right-hand side of (7-35) is equal to $\psi_m'(\gamma_t^d)$. Then (7-35) reduces to

$$\tau_t = \psi_m'(\gamma_t^d). \tag{7-37}$$

On the other hand, for the energy without disarrangements φ, from (3-19), (7-33), and (7-34) we have

$$T_t = \nabla\varphi(F_t) = \nabla\varphi(\nabla f_t - F_t^d) = \nabla\varphi(I + (\bar{\gamma} t - \gamma_t^d) s \otimes m). \tag{7-38}$$

Multiplying by $(s \otimes m)$, on the left side we get again τ_t. Then, after defining

$$\Phi_m^s(\bar{\gamma} t - \gamma_t^d) = \nabla\varphi(I + (\bar{\gamma} t - \gamma_t^d) s \otimes m) \cdot (s \otimes m), \tag{7-39}$$

we get

$$\tau_t = \Phi_m^s(\bar{\gamma} t - \gamma_t^d). \tag{7-40}$$

---

[57] For two-scale continua, the use of smooth potentials in the study of plasticity and fracture was initiated in [Choksi et al. 1999, §4, §5], and continued by Deseri and Owen [2002] for the plasticity of single crystals. The specular case $\phi = \phi^{nh}$ of a purely nonsmooth potential was advanced in [Del Piero 1998; 2018], without subsequent developments.

[58] In the presence of a smooth potential, I prefer to avoid calling F_t^d a plastic deformation.

By elimination of γ_t^d between this equation and (7-37), a relation between the shear stress τ_t and the macroscopic shear $\bar{\gamma}t$ is obtained. For example, in the case of Φ_m^s linear and ψ_m trigonometric of period p,

$$\Phi_m^s(\xi) = k\xi, \qquad \psi_m(\xi) = \frac{k^d p}{2\pi}\left(1 - \cos\frac{2\pi\xi}{p}\right), \tag{7-41}$$

with k and k^d positive material constants, we get

$$\bar{\gamma}t = \frac{\tau_t}{k} + \frac{p}{2\pi}\sin^{-1}\left(\frac{\tau_t}{k^d}\right). \tag{7-42}$$

This determines a curve $\bar{\gamma}t = \mathcal{F}(\tau_t)$, in which $\mathcal{F}$ is a multivalued function with domain $(-k^d, k^d)$. The slope of the curve is

$$\frac{d\bar{\gamma}t}{d\tau_t} = \frac{1}{k} \pm \frac{p}{2\pi\sqrt{k^{d2} - \tau_t^2}}, \tag{7-43}$$

with the plus sign for the branch from the origin to its first intersection with the line $\tau_t = k^d$, the minus sign for the following branch up to the first intersection with the line $\tau_t = -k^d$, and so on.

The slope of the first branch is positive and increases from $1/k + p/2\pi k^d$ at $\tau_t = 0$ to $+\infty$ at $\tau_t = k^d$. The slope of the second branch increases from $-\infty$ at $\tau_t = k^d$, to zero at[59]

$$\tau_t = \sqrt{k^{d2} - \frac{k^2 p^2}{4\pi^2}}. \tag{7-44}$$

At this point the curve attains a local maximum. Then for a further increase of $\bar{\gamma}t$ there is no solution near the maximum point, and equilibrium for the increased $\bar{\gamma}t$ can be attained only jumping to another branch of the curve. This jump is a form of *catastrophic instability*, consisting of a sudden decrease of τ_t at constant $\bar{\gamma}t$.

Taking the macroscopic shear $\bar{\gamma}t$ as independent variable, the stress-strain response curve is $\tau_t = \mathcal{F}^{-1}(\bar{\gamma}t)$. The function $\mathcal{F}^{-1}$ is periodic with period p, since if $\bar{\gamma}t$ is a solution of (7-42) for some τ_t, then $\bar{\gamma}t + np$ is a solution for the same τ_t for all n. In the equilibrium branch starting from $(\tau_t, \bar{\gamma}t) = (0, 0)$, τ_t increases with $\bar{\gamma}t$ up to the upper limit k^d, and then decreases up to the lower limit $-k^d$. When in the descending branch the local maximum of $\mathcal{F}$ is attained, γ_t suddenly jumps downward at constant $\bar{\gamma}t$, to reach the next ascending branch of $\mathcal{F}^{-1}$. Due to the periodicity of $\mathcal{F}^{-1}$, the same jump occurs at each period. The resulting stress-strain

---

[59]This value is attained only if $2\pi k^d > kp$. But this condition is satisfied in practice, since p is a very small length, of the order of the interatomic distance [Choksi et al. 1999, §4].

diagram then shows an initial increase of γ_t from zero to k^d, followed by a periodic oscillation between k^d and the value at which the maximum of $\mathcal{F}$ is attained.[60]

Macroscopically, this looks like a typical elastic-perfectly plastic response consisting of an initial growth followed by a horizontal plateau. A peculiarity of this model is that the plateau is made of microscopic oscillations.[61] By consequence, while in all previous models yielding is activated when a threshold determined by a nonsmooth potential is attained, in the present model yielding is due to catastrophic instability. When ϕ is a smooth potential there is no dissipation associated with the equilibrium branches, the whole dissipation being concentrated on the jumps.[62]

Closure

I conclude with a summary of the main results.

  (i) By imposing the nonnegativeness of the first variation of the energy, all constitutive elements of the theory become dependent on a single punctual inequality, the plastic stationarity condition.

 (ii) From this condition the yield condition, the flow rule, the hardening rule, and the incremental response law can be deduced without any additional assumption.

(iii) A plastic energy made of a smooth potential plus a homogeneous dissipation potential determines a model of kinematic plasticity which includes Prager's kinematic hardening as a special case.

(iv) A plastic energy with the properties of a nonhomogeneous dissipation potential determines a model of dilatational hardening which includes isotropic hardening as a special case.

 (v) The assumption of isochoricity of the plastic strain rate leads to the definition of a one-parameter family of pairwise disjoint elastic ranges, the parameter being the hydrostatic pressure. In each elastic range the normality rule holds. This renders unnecessary the nonassociated flow rules introduced by the many authors for which the elastic range is the union of all individual elastic ranges.

(vi) The Cam-clay model for soils is a particular case of isotropic dilatational plasticity.

---

[60]See, e.g., [Choksi et al. 1999, Figure 6].

[61]This may reproduce the oscillations exhibited by testing machines operating with the "hard device", that is, by controlling the deformation. The oscillations are much less sensible in the "soft" testing machines, which control the applied load. This difference is well reproduced by the present model, see Figures 6 and 8 in [Choksi et al. 1999].

[62]See [Choksi et al. 1999, §4]. This alternative description of yielding led the present author to distinguish two types of yielding, which he called *reversible* and *irreversible* [Del Piero 1998; 2013].

(vii) Crystal plasticity is a special case of isochoric plasticity, with a particular form of the plastic stretching. The two-level shear model is a special case of crystal plasticity, with diffuse plastic dissipation replaced by concentrated catastrophic instability.

References

[Aifantis 1984] E. C. Aifantis, "On the microstructural origin of certain inelastic models", *J. Eng. Mater. Technol.* **106**:4 (1984), 326–330.

[Aifantis 1987] E. C. Aifantis, "The physics of plastic deformation", *Int. J. Plasticity* **3**:3 (1987), 211–247.

[Anand and Gurtin 2003] L. Anand and M. E. Gurtin, "A theory of amorphous solids undergoing large deformations, with application to polymeric glasses", *Int. J. Solids Struct.* **40**:6 (2003), 1465–1487.

[Baía et al. 2012] M. Baía, J. Matias, and P. M. Santos, "A relaxation result in the framework of structured deformations in a bounded variation setting", *Proc. R. Soc. Edinburgh A* **142**:2 (2012), 239–271.

[Barroso et al. 2017] A. C. Barroso, J. Matias, M. Morandotti, and D. R. Owen, "Second-order structured deformations: relaxation, integral representation and applications", *Arch. Ration. Mech. Anal.* **225**:3 (2017), 1025–1072.

[Bažant et al. 1984] Z. P. Bažant, T. B. Belytschko, and T.-P. Chang, "Continuum theory for strain softening", *J. Eng. Mech.* **110**:12 (1984), 1666–1692.

[Carstensen et al. 2002] C. Carstensen, K. Hackl, and A. Mielke, "Non-convex potentials and microstructures in finite-strain plasticity", *Proc. R. Soc. London A* **458**:2018 (2002), 299–317.

[Choksi and Fonseca 1997] R. Choksi and I. Fonseca, "Bulk and interfacial energy densities for structured deformations of continua", *Arch. Ration. Mech. Anal.* **138**:1 (1997), 37–103.

[Choksi et al. 1999] R. Choksi, G. Del Piero, I. Fonseca, and D. Owen, "Structured deformations as energy minimizers in models of fracture and hysteresis", *Math. Mech. Solids* **4**:3 (1999), 321–356.

[Ciarlet 1988] P. G. Ciarlet, *Mathematical elasticity, I*, Studies in Math. and Its Appl. **20**, North-Holland, Amsterdam, 1988.

[Collins 2003] I. F. Collins, "A systematic procedure for constructing critical state models in three dimensions", *Int. J. Solids Struct.* **40**:17 (2003), 4379–4397.

[Collins and Houlsby 1997] I. F. Collins and G. T. Houlsby, "Application of thermomechanical principles to the modelling of geotechnical materials", *Proc. R. Soc. London A* **453** (1997), 1975–2001.

[Dafalias and Popov 1975] Y. F. Dafalias and E. P. Popov, "A model of nonlinearly hardening materials for complex loading", *Acta Mech.* **21**:3 (1975), 173–192.

[Dal Maso et al. 2006] G. Dal Maso, A. DeSimone, and M. G. Mora, "Quasistatic evolution problems for linearly elastic-perfectly plastic materials", *Arch. Ration. Mech. Anal.* **180**:2 (2006), 237–291.

[Dal Maso et al. 2008] G. Dal Maso, A. DeSimone, M. G. Mora, and M. Morini, "Globally stable quasistatic evolution in plasticity with softening", *Netw. Heterog. Media* **3**:3 (2008), 567–614.

[Dashner 1986] P. A. Dashner, "Invariance considerations in large strain elasto-plasticity", *J. Appl. Mech.* **53**:1 (1986), 55–60.

[Del Piero 1998] G. Del Piero, "Towards a unified approach to fracture, yielding, and damage", pp. 679–692 in *Continuum models and discrete systems* (Istanbul, 1998), edited by E. Inan and K. Z. Markov, World Sci. Publ., River Edge, NJ, 1998.

[Del Piero 2001] G. Del Piero, "The energy of a one-dimensional structured deformation", *Math. Mech. Solids* **6**:4 (2001), 387–408.

[Del Piero 2013] G. Del Piero, "A variational approach to fracture and other inelastic phenomena", *J. Elasticity* **112**:1 (2013), 3–77.

[Del Piero 2018] G. Del Piero, "On the decomposition of the deformation gradient in plasticity", *J. Elasticity* **131**:1 (2018), 111–124.

[Del Piero and Owen 1993] G. Del Piero and D. R. Owen, "Structured deformations of continua", *Arch. Ration. Mech. Anal.* **124**:2 (1993), 99–155.

[Del Piero and Owen 1995] G. Del Piero and D. R. Owen, "Integral-gradient formulae for structured deformations", *Arch. Ration. Mech. Anal.* **131**:2 (1995), 121–138.

[Del Piero et al. 2013] G. Del Piero, G. Lancioni, and R. March, "A diffuse cohesive energy approach to fracture and plasticity: the one-dimensional case", *J. Mech. Mater. Struct.* **8**:2-4 (2013), 109–151.

[Deseri and Owen 2002] L. Deseri and D. R. Owen, "Energetics of two-level shears and hardening of single crystals", *Math. Mech. Solids* **7**:2 (2002), 113–147.

[Drucker 1952] D. C. Drucker, "A more fundamental approach to plastic stress-strain relations", pp. 487–491 in *Proc. of the First US National Congress of Applied Mechanics* (Chicago, 1951), edited by E. Sternberg, Amer. Soc. Mech. Engineers, New York, 1952.

[Eve et al. 1990] R. A. Eve, B. D. Reddy, and R. T. Rockafellar, "An internal variable theory of elastoplasticity based on the maximum plastic work inequality", *Quart. Appl. Math.* **48**:1 (1990), 59–83.

[Fedelich and Ehrlacher 1989] B. Fedelich and A. Ehrlacher, "Sur un principe de minimum concernant des matériaux à comportement indépendant du temps physique", *C. R. Acad. Sci. Paris* **308**:16 (1989), 1391–1394.

[Fleck and Hutchinson 2001] N. A. Fleck and J. W. Hutchinson, "A reformulation of strain gradient plasticity", *J. Mech. Phys. Solids* **49**:10 (2001), 2245–2271.

[Green and Naghdi 1971] A. E. Green and P. M. Naghdi, "Some remarks on elastic-plastic deformation at finite strain", *Int. J. Eng. Sci.* **9**:12 (1971), 1219–1229.

[Gudmundson 2004] P. Gudmundson, "A unified treatment of strain gradient plasticity", *J. Mech. Phys. Solids* **52**:6 (2004), 1379–1406.

[Gurtin 1972] M. E. Gurtin, *The linear theory of elasticity*, Band VIa/2, Handbuch der Physik, Springer, 1972.

[Gurtin and Anand 2005] M. E. Gurtin and L. Anand, "The decomposition $F = F^e F^p$, material symmetry, and plastic irrotationality for solids that are isotropic-viscoplastic or amorphous", *Int. J. Plasticity* **21**:9 (2005), 1686–1719.

[Gurtin et al. 2010] M. E. Gurtin, E. Fried, and L. Anand, *The mechanics and thermodynamics of continua*, Cambridge Univ. Press, 2010.

[Halmos 1942] P. R. Halmos, *Finite dimensional vector spaces*, Annals of Math. Studies **7**, Princeton Univ. Press, 1942.

[Hill 1948] R. Hill, "A variational principle of maximum plastic work in classical plasticity", *Quart. J. Mech. Appl. Math.* **1** (1948), 18–28.

[Hill 1950] R. Hill, *The mathematical theory of plasticity*, Oxford Univ. Press, 1950.

[Houlsby and Purzrin 2000] G. T. Houlsby and A. M. Purzrin, "A thermomechanical framework for constitutive models for rate-independent dissipative materials", *Int. J. Plasticity* **16**:9 (2000), 1017–1047.

[Il'yushin 1961] A. A. Il'yushin, "On the postulate of plasticity", *Prikl. Mat. Meh.* **25**:3 (1961), 503–507. In Russian; translated in *J. Appl. Math. Mech.* **25**:3 (1961), 746–752.

[Koiter 1953] W. T. Koiter, "Stress-strain relations, uniqueness and variational theorems for elastic-plastic materials with a singular yield surface", *Quart. Appl. Math.* **11** (1953), 350–354.

[Lemaitre and Chaboche 1990] J. Lemaitre and J.-L. Chaboche, *Mechanics of solid materials*, Cambridge Univ. Press, 1990.

[Lubarda and Lee 1981] V. A. Lubarda and E. H. Lee, "A correct definition of elastic and plastic deformation and its computational significance", *J. Appl. Mech.* **48**:1 (1981), 35–40. Erratum in **48**:3 (1981), 686.

[Lubliner 1990] J. Lubliner, *Plasticity theory*, MacMillan, New York, 1990.

[Lucchesi and Podio-Guidugli 1990] M. Lucchesi and P. Podio-Guidugli, "Equivalent dissipation postulates in classical plasticity", *Meccanica* **25**:1 (1990), 26–31.

[Martin and Reddy 1993] J. B. Martin and B. D. Reddy, "Piecewise smooth dissipation and yield functions in plasticity", *Meccanica* **28**:3 (1993), 169–175.

[Mielke 2003] A. Mielke, "Energetic formulation of multiplicative elasto-plasticity using dissipation distances", *Contin. Mech. Thermodyn.* **15**:4 (2003), 351–382.

[Moreau 1970] J.-J. Moreau, "Sur les lois de frottement, de plasticité et de viscosité", *C. R. Acad. Sci. Paris* **271** (1970), 608–611.

[Moreau 1974] J.-J. Moreau, "On unilateral constraints, friction and plasticity", pp. 171–322 in *New variational techniques in mathematical physics* (Bressanone, 1973), edited by G. Capriz and G. Stampacchia, CIME Summer Schools **63**, Cremonese, Rome, 1974.

[Mróz 1969] Z. Mróz, "An attempt to describe the behavior of metals under cyclic loads using a more general workhardening model", *Acta Mech.* **7**:2-3 (1969), 199–212.

[Noll 1972] W. Noll, "A new mathematical theory of simple materials", *Arch. Ration. Mech. Anal.* **48** (1972), 1–50.

[Owen 1995] D. R. Owen, "Disarrangements in continua and the geometry of microstructure", pp. 67–81 in *Recent advances in elasticity, viscoelasticity and inelasticity*, edited by K. R. Rajagopal, Series on Adv. Math. Appl. Sci. **26**, World Sci., Singapore, 1995.

[Owen 2000] D. R. Owen, "Structured deformations, II", lecture notes 58, Ravello Summer School of the Gruppo Nazionale di Fisica Matematica, 2000, Available at https://tinyurl.com/owensd.

[Owen and Paroni 2015] D. R. Owen and R. Paroni, "Optimal flux densities for linear mappings and the multiscale geometry of structured deformations", *Arch. Ration. Mech. Anal.* **218**:3 (2015), 1633–1652.

[Petryk 2003] H. Petryk, "Incremental energy minimization in dissipative solids", *C. R. Acad. Sci. Paris* **331**:7 (2003), 469–474.

[Pham et al. 2011a] K. Pham, H. Amor, J.-J. Marigo, and C. Maurini, "Gradient damage models and their use to approximate brittle fracture", *Int. J. Damage Mech.* **20**:4 (2011), 618–652.

[Pham et al. 2011b] K. Pham, J.-J. Marigo, and C. Maurini, "The issues of the uniqueness and the stability of the homogeneous response in uniaxial tests with gradient damage models", *J. Mech. Phys. Solids* **59**:6 (2011), 1163–1190.

[Phillips and Sierakowski 1965] A. Phillips and R. L. Sierakowski, "On the concept of the yield surface", *Acta Mech.* **1**:1 (1965), 29–35.

[Prager 1949] W. Prager, "Recent developments in the mathematical theory of plasticity", *J. Appl. Phys.* **20**:3 (1949), 235–241.

[Prager 1955] W. Prager, *Probleme der Plastizitätstheorie*, Birkhäuser, Basel, 1955.

[Reddy 2011] B. D. Reddy, "The role of dissipation and defect energy in variational formulations of problems in strain-gradient plasticity, I: Polycrystalline plasticity", *Contin. Mech. Thermodyn.* **23**:6 (2011), 527–549. Erratum in **24**:1 (2012), 79.

[Reddy et al. 2008] B. D. Reddy, F. Ebobisse, and A. McBride, "Well-posedness of a model of strain gradient plasticity for plastically irrotational materials", *Int. J. Plasticity* **24**:1 (2008), 55–73.

[Reina and Conti 2014] C. Reina and S. Conti, "Kinematic description of crystal plasticity in the finite kinematic framework: a micromechanical understanding of $F = F^e F^p$", *J. Mech. Phys. Solids* **67** (2014), 40–61.

[Reina et al. 2016] C. Reina, A. Schlömerkemper, and S. Conti, "Derivation of $F = F^e F^p$ as the continuum limit of crystalline slip", *J. Mech. Phys. Solids* **89** (2016), 231–254.

[Rice 1971] J. R. Rice, "Inelastic constitutive relations for solids: an internal-variable theory and its application to metal plasticity", *J. Mech. Phys. Solids* **19**:6 (1971), 433–455.

[Rockafellar 1970] R. T. Rockafellar, *Convex analysis*, Princeton Math. Series **28**, Princeton Univ. Press, 1970.

[Roscoe and Burland 1968] K. H. Roscoe and J. B. Burland, "On the generalized stress-strain behaviour of 'wet' clay", pp. 535–609 in *Engineering plasticity* (Cambridge, 1968), edited by J. Heyman and F. A. Leckie, Cambridge Univ. Press, 1968.

[Roscoe and Poorooshasb 1963] K. H. Roscoe and H. B. Poorooshasb, "A theoretical and experimental study of strains in triaxial compression tests on normally consolidated clays", *Géotechnique* **13**:1 (1963), 12–38.

[Roscoe et al. 1958] K. H. Roscoe, A. N. Schofield, and C. P. Wroth, "On the yielding of soils", *Géotechnique* **8**:1 (1958), 22–53.

[Schofield and Wroth 1968] A. N. Schofield and P. Wroth, "Cam-clay and the critical state concept", Chapter 6, pp. 93–115 in *Critical state soil mechanics*, McGraw-Hill, New York, 1968.

[Šilhavý 2017] M. Šilhavý, "The general form of the relaxation of a purely interfacial energy for structured deformations", *Math. Mech. Complex Syst.* **5**:2 (2017), 191–215.

[Srinivasa 2010] A. R. Srinivasa, "Application of the maximum rate of dissipation criterion to dilatant, pressure dependent plasticity models", *Int. J. Eng. Sci.* **48**:11 (2010), 1590–1603.

[Truesdell and Noll 1965] C. Truesdell and W. Noll, *The non-linear field theories of mechanics*, Band III/3, Handbuch der Physik, Springer, 1965.

[Vermeer and de Borst 1984] P. A. Vermeer and R. de Borst, "Non-associated plasticity for soils, concrete and rock", *Heron* **29**:3 (1984), 1–64.

[Wilson 2002] C. D. Wilson, "A critical reexamination of classical metal plasticity", *J. Appl. Mech.* **69**:1 (2002), 63–68.

[Ziegler 1963] H. Ziegler, "Some extremum principles in irreversible thermodynamics with application to continuum mechanics", pp. 91–193 in *Progress in solid mechanics, IV*, edited by I. N. Sneddon and R. Hill, North-Holland, Amsterdam, 1963.

[Ziegler 1983] H. Ziegler, *An introduction to thermomechanics*, 2nd ed., North-Holland Series in Appl. Math. and Mech. **21**, North-Holland, Amsterdam, 1983.

[Ziegler and Wehrli 1987] H. Ziegler and C. Wehrli, "The derivation of constitutive relations from the free energy and the dissipation function", pp. 183–237 Adv. in Appl. Mech. **25**, Academic Press, Orlando, FL, 1987.

GIANPIETRO DEL PIERO: dlpgpt@unife.it
Dipartimento di Ingegneria, Università di Ferrara, Ferrara, Italy

and

*Centro Internazionale di Ricerca per la Matematica & Meccanica dei Sistemi Complessi,
Università dell'Aquila, L'Aquila, Italy*
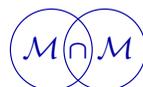
FAR-REACHING HELLENISTIC GEOGRAPHICAL KNOWLEDGE HIDDEN IN PTOLEMY'S DATA

Lucio Russo

The paper summarizes and discusses the main theses exposed in a previous book (*L'America dimenticata*, Mondadori Università, 2013; in Italian) in light of more recent results. Specifically, the work addresses the problem of explaining the origin of the systematic error on longitudes in Ptolemy's *Geographia* and its logical relation with the reduced estimate for the dimension of the Earth given there. The thesis is sustained that, contrary to a frequently advanced conjecture, the shrinking of the dimension of the Earth is a consequence of a scale error in longitudes, which, in turn, was originated by a misidentification of the Islands of the Blessed. The location of the Islands of the Blessed according to the source of Ptolemy is identified in the Caribbean. The analysis of a passage of Pliny provides an independent and quantitative confirmation of the proposed identification, which sheds new light on possible contact among civilizations.

1. The shrinking of the Earth and the dilation of longitudes in Ptolemy's *Geographia*

It is well known that Eratosthenes, in the 3rd century BC, measured the circumference of the Earth, obtaining the value of 252,000 stadia (corresponding to 700 stadia per degree). Four centuries later Ptolemy, accepting a value suggested by Posidonius in the 1st century BC, estimated the same circumference in 180,000 stadia (500 stadia per degree). The method employed by Eratosthenes is described, in outline at least,[1] by Cleomedes (*Caelestia*, I, 7, 48–120), while we have no information about the origin of the smaller value. Almost all scholars have always believed that Eratosthenes and Ptolemy were using the same stadium, even if this

---

[1]For a discussion of Eratosthenes' method see [Russo 2004, 273–276]; concerning the possible simplifications introduced by Cleomedes in his popular account, see [Russo 2013a, 71–76], which also contains an attempt to reconstruct the original computation by Eratosthenes.

has been called into question by certain historians.[2]

Much more controversial has been the actual value of the stadium considered by the two. Indeed, many scholars[3] have accepted the value of 157.5 m, deduced from a passage in Pliny,[4] while several others have maintained that the Olympic stadium (ca. 185 m) was employed.[5] Other values have also been proposed.[6] Assuming the value of 157.5 m, Eratosthenes would have made an error of less than 1%, whereas for Ptolemy and Posidonius the error is large, roughly 40%. While these error estimates rely on the determination of the value of the stadium, another error in Ptolemy's *Geographia*, which can be studied independently of the question of length units, is the systematic error in longitudes. It is well known that longitude differences were systematically dilated by Ptolemy; in particular he grossly over-estimated the longitudinal amplitude of the oikoumene (the inhabited part of the world).

In [Russo 2013a] the longitudes reported by Ptolemy were used to determine the "effective" value of Ptolemy's stadium. The main result of that work was an independent validation of the substantial exactness of the value 157.5 deduced from Pliny, which implies that the estimate for the Earth's circumference by Ptolemy was

---

[2]A few scholars have suggested that Ptolemy (and Posidonius before him) replaced Erastosthenes' stadium with a new stadium corresponding to about 222 meters, so that the two measures of the circumference of the Earth should coincide [Gossellin 1790; Jomard1817; Valerio 2013]. This conjecture is contradicted by the explicit statement by Strabo that Posidonius' measure of 180,000 stadia had reduced the dimension of the Earth (*Geographica*, II, ii, 2). Furthermore, if Ptolemy was using a stadium 40% longer than that of Eratosthenes, one would be hard pressed to explain why he dilated in (approximately) the same measure the numerical values of differences of longitude.

[3]See, among others, [Letronne 1851, 104–119, 212–246; Hultsch 1882, 60–63; Tannery 1893, 109–110; Dreyer 1953, 175; Miller 1919, 6–7; Oxé 1963, 269–270; Aujac 1966, 176–179; Fraser 1972, II, 599, n. 312; Stückelberger 1988, 188; Dutka 1993/94, 63–64; Meuret 1998, 163–164, Tupikova and Geus 2013, 21.]

[4]Pliny writes: "schoenus patet Eratosthenis ratione stadia XL, hoc est p. $\overline{v}$" (*Naturalis Historia*, XII, 53). This sentence, using the known value of the schoenus, gives for the stadium the value of 157.5 meters. It is true that in another passage (*N.H.*, II, 247) Pliny translates Eratosthenes' result into 31500 milia passuum (a calculation involving the use of the Olympic stadium of 185 meters), but this circumstance enhances, in my opinion, the role of the words "Eratosthenis ratione" used in the first case. Indeed, the Olympic stadium was certainly widely used and its automatic use by Pliny is not surprising, while his need to clarify, in the other case, that he is considering the particular unit introduced by Eratosthenes, gives us a precious testimony. Given the enormous influence of Eratosthenes' *Geographica*, it is not surprising that the new "stadium" introduced by him was adopted as the standard unit in geographical treatises, while the Olympic stadium remained in use for other purposes.

[5]See, among others, [Columba 1895, 63–68; Czwalina 1925, 295; Dicks 1960, 42–46; Rawlins 1981, 218; Pothecary 1995, 49–67; Berggren and Jones 2000, 14, footnote 10].

[6]Most of them are analyzed in [Tupikova and Geus 2013, 20–22] (where, however, the values outside the interval 148–180m are considered quite implausible and the extreme possibility mentioned in note 2 above is not considered worthy of mention).

affected by a very large error. That paper was based on the analysis of longitudes of a sample of 80 cities, chosen as the most renowned among those in the part of the world best- nown in Hellenistic times. Plotting Ptolemy's longitudes against the values currently accepted for them, a graph was obtained which is approximated remarkably well by means of a linear regression. The equation of the regression line is

$$y = 1.428x + 17.06 \qquad (1)$$

and the coefficient of determination is $R^2 = 0.9935$. More recently, in [Shcheglov 2014] a larger sample of 245 locations (including some river's mouths and capes) was considered, yielding very similar results. Indeed, in that study the equation of the regression line is

$$y = 1.4279x + 16.425, \qquad (2)$$

and the relative coefficient of determination is $R^2 = 0.9874$. This result confirms that of [Russo 2013a], and the slightly lower value of R^2 (still very close to 1) is likely linked to the inclusion in the sample of locations from lesser known regions such as Northern Europe and India. In the two papers the slope of the regression line is virtually identical,[7] and it is close to the ratio of 1.4 between Ptolemy's and Eratosthenes' values for the Earth's circumference. We know that Ptolemy generally deduced differences of longitudes from distances expressed in stadia along circles of latitude (mostly taken from Eratosthenes, who in his geographical work, instead of degrees of longitude, had reported distances between meridians along a particular parallel of latitude). It is therefore not surprising that Ptolemy's longitude differences were dilated in such a way to compensate, for such distances, for the error in the dimension of the Earth. (We will return later to the slight difference between 1.4 and the value of the regression coefficient.) Hence it is very likely that a logical link does exist between the error on the dimension of the Earth and the error on the differences of longitudes.

Some scholars have interpreted this link by proposing the following implication:

A* : *Ptolemy assumed a wrong measure for the Earth's circumference. As a consequence he systematically deduced dilated longitudes from his data involving distances along circles of latitude.*[8]

---

[7]The method used in [Russo 2013a] for estimating the value of the stadium was actually based on the regression coefficient, so a validation of the value of this coefficient by means of the much larger set of locations considered in [Shcheglov 2014] provides a sounder statistical basis for that estimate.

[8]This implication is maintained, in particular, in [Rawlins 1985; Rawlins 2008; Tupikova and Geus 2013; Tupikova 2013]. Strangely enough, in Shcheglov 2014 the same opinion is ascribed also to [Russo 2013a] (where, in this regard, it is only written that "the distortion operated by Ptolemy on the longitudes is not independent of the new value he had assumed for the length of the Earth's circumference") and even to [Russo 2013b], where the implication is explicitly criticized.

This implication is not, however, a necessary consequence of the link we mentioned before and in the next section we will show that actually several arguments allow us to discard it.

2. Rereading the relationship between the two errors by Ptolemy

A first argument against A* consists in observing that it leaves unsolved the problem of the origin of the error in the Earth's dimension. The value obtained by Eratosthenes was indeed well known, and it is transmitted by ten different extant sources, dating from the I century BC to the V century AD.[9] Posidonius gave two values for the Earth's circumference. For one of them (240,000 stadia), not too far from the one by Eratosthenes, we know the actual procedure by which it was obtained (Cleomedes, *Caelestia*, I, 7, 1–47). Had the other one (180,000 stadia) been a result of some measurement procedure, one could wonder about three things:

1. Why is there no source describing it?

2. How could it have produced such a large error?

3. Why should such a grossly wrong estimate, which no extant source considers worth of being described, have prevailed over the value produced by the celebrated measurement by Eratosthenes?

A further strong argument against A* was provided in [Shcheglov 2014]. Namely, if all differences of longitudes were dilated because of an error in the assumed dimension of the Earth, we should have more or less the same dilation in different regions (at least for the better known ones). In contrast, splitting his sample in nine subsets corresponding to different geographic areas, Shcheglov verified that linear regressions performed on the different subsets give substantially different values for the slope, in every case with the coefficient R^2 very close to 1.

Those arguments are largely sufficient, in my opinion, to reject implication A*.

3. How did Ptolemy actually work?

Ptolemy, in his *Geographia*, states that the oikoumene is spread over 180° of longitude, from the westernmost locations (four of the six Islands of the Blessed, the other two being one degree further east) to the easternmost ones (some towns in

---

[9]Strabo, *Geographica*, II, v, 7, 34; Geminus, *Introduction to the Phenomena*, XVI, 6; Macrobius, *Commentarii in Somnium Scipionis*, I, xx, 20; Vitruvius, *De Architectura*, I, vi, 9; Pliny the Elder, *Naturalis Historia*, II, 247; Censorinus, *De Die Natali*, xiii, 5; Theon of Smyrna, *De Utilitate Mathematicae*, 124, 10–12 (ed. Hiller); Heron of Alexandria, *Dioptra*, xxxv, 302, 10–17 (ed. Schöne); Martianus Capella, *De Nuptiis Philologiae et Mercurii*, VI, 596; and the already mentioned Cleomedes (*Caelestia*, I, 7, 48–120). This last author is the only one giving a round value of 250,000.

China). He claims (*Geographia*, I, 12) to have determined the amplitude of the oikoumene in pieces, by considering a path from the meridian of the Islands of the Blessed to the Metropolis of the Seres (the capital of China). The path is formed from twelve portions, whose relative lengths are derived by heavily modifying, with very rough arguments,[10] data given by Marinus of Tyre. The values obtained after this procedure are finally translated into differences of longitude assuming the aforementioned estimate for the Earth's circumference: 500 stadia per degree. The result is 177° 15′. Since the easternmost locations known to Ptolemy are just 2° 45′ to the east of the Metropolis of the Seres, he eventually gets a total of 180°. Clearly Ptolemy is fiddling with the numbers: the likelihood is negligible that he'd have reached this round value by accident after arbitrarily modifying twelve terms in a sum.

At the same time, Ptolemy himself, at the beginning of his *Geographia*, explains his actual method, when he points out the difference between the subject of his work and chorography:

> The goal of chorography is an impression of a part, as when one makes an image of just an ear or an eye; but [the goal] of geography is a general view, analogous to making a portrait of the whole head. That is, whenever a portrait is to be made, one has to fit in the main parts [of the head] in a determined pattern and an order of priority. Furthermore the [surfaces] that are going to hold the drawings ought to be of a suitable size for the spacing of the visual rays at an appropriate distance, whether the drawing be of whole or part, so that everything will be grasped by the sense [of sight].[11]

From the metaphor it is clear that Ptolemy, having to represent the whole oikoumene, wants first to fix the positions of some key locations, and in particular its global dimension, and only afterwards to add the coordinates of all the remaining localities.

One may wonder in what sense the positions of these key locations were fixed — whether by means of their angular coordinates or their relative distances. In this regard, it can be recalled that Ptolemy himself explains that the astronomical method,

---

[10]For instance, dividing a value by two to account for tortuous roads.

[11]Ἔχεται δὲ τὸ μὲν χωρογραφικὸν τέλος τῆς ἐπὶ μέρους προσβολῆς, ὡς ἂν εἴ τις οὓς μόνον ἢ ὀφθαλμὸν μιμοῖτο, τὸ δὲ γεωγραφικὸν τῆς καθόλου θεωρίας κατὰ τὸ ἀνάλογον τοῖς ὅλην τὴν κεφαλὴν ὑπογραφομένοις. Πάσαις γὰρ ταῖς ὑποτεθειμέναις εἰκόσι τῶν πρώτων μερῶν ἀναγκαίως καὶ προηγουμένως ἐφαρμοζομένων, καὶ ἔτι τῶν δεξομένων τὰς γραφὰς συμμέτρων ὀφειλόντων εἶναι ταῖς ἐξ ἀποχῆς αὐτάρχως τῶν ὄψεων διαστάσεσιν, ἐάν τε τέλειον ᾖ τὸ γραφόμενον ἐάν τ'ἐπὶ` μέρους, ἵν'ἄπαν αισθητῶς παραλαμβάνηται [. . .] (Ptolemy, *Geographia*, I, 1, 2–3; translation adapted from [Berggren and Jones 2000]).

from which only angular measures can be derived, is to be preferred to the survey-
ing which could provide distances in stadia:

> The surveying component is that which indicates the relative positions
> of localities solely through measurement of distances; the astronomical
> component [is that which does the same] by means of the phenomenon
> [obtained] from astronomical sighting and shadow-casting instruments.
> Astronomical observation is a self-sufficient thing and less subject to
> error, while surveying is cruder and incomplete without [astronomical
> observation].[12]

The order in which the operations were performed is indeed explicitly indicated
in Ptolemy's *Geographia*, in the title of chapter I, 4:

> That it is necessary to give a priority to the [astronomical] phenomena
> over [data] from records of travel.[13]

Shortly afterward he points out:

> It would therefore also be reasonable for one who intended to practice
> geography following these [principles] to give priority in his map to the
> [features] that have been obtained through the more accurate observa-
> tions, as foundations, so to speak, but to fit [the features] that come from
> the other [kinds of data] to these, until their positions with respect to each
> other and to the first [features] stand as much as possible in agreement
> with those reports that are less subject to error.[14]

From these passages, we can deduce that Ptolemy first fixed a few longitudes
known by astronomical methods (in particular those of some extreme points deter-
mining the amplitude in degrees of the oikoumene), and then he interpolated the
longitudes of the intermediate locations by using known distances along circles of
latitude. Therefore, the first error among the two mentioned in the previous section
has necessarily to be the one on longitudes. Hence, as already said, A* has to be
discarded. Since, on the other hand, Ptolemy's scale error on longitudes, combined

---

[12]γεωμετρικὸν μὲν τὸ διὰ ψιλῆς τῆς ἀναμετρήσεως τῶν διαστάσεων τὰς πρὸς ἀλλήλους
θέσεις τῶν τόπων ἐμφανίζον, μετεωροσκοπικὸν δὲ τὸ διὰ τῶν φαινομένων ἀπὸ τῶν ἀστρολάβων
καὶ σκιοθήρων ὀργάνων· τοῦτο μὲν, ὡς αὐτοτελές τι καὶ ἀδιστακτότερον, ἐκεῖνο δὲ, ὡς
ὁλοσχερέστερον καὶ τούτου προσδεόμενον. (Ptolemy, *Geographia*, I, 1, 2; translation by Berggren
and Jones.)

[13]Ὅτι δεῖ τὰ ἐκ τῶν φαινομένων προϋποτίθεσθαι τῶν ἐκ τῆς περιοδικῆς ἱστορίας (Ptolemy,
Geographia, I, 4, T; translation by Berggren and Jones).

[14]εὔλογον ἂν εἴη καὶ τὸν τούτοις ἀκολούθως γεωγραφήσοντα τὰ μὲν διὰ τῶν ἀκριβεστέρων
τηρήσεων εἰλημμένα προϋποτίθεσθαι τῇ καταγραφῇ καθάπερ θεμελίους, τὰ δ' ἀπὸ τῶν ἄλλων
ἐφαρμόζειν τούτοις, ἕως ἂν αἱ πρὸς ἄλληλα θέσεις αὐτῶν μετὰ τῶν πρὸς τὰ πρῶτα τηρῶσιν
ὡς ἔνι μάλιστα συμφώνως τὰς ἀδιστακτοτέρας τῶν παραδόσεων. (Ptolemy, *Geographia*, I, 4, 1;
adapted from the translation by Berggren and Jones.)

with a set of right distances along circles of latitude, necessarily implies his error on Earth's dimension, we are led to propose the following implication:

A** : *Ptolemy's error on the amplitude in longitude of the oikoumene was the cause of his systematic dilation of the differences of longitude and of his deduction, from the known data for the distances along circles of latitude, of an underestimation of the dimension of the Earth.*

As for the source of the scale error on longitudes, we start by observing that, since Ptolemy emphasizes the importance of astronomical data provided by Hipparchus (*Geographia*, I, 4, 1), we can conjecture that he took from him also the value of 180° for the amplitude of the oikoumene. An argument supporting this is provided once again by Ptolemy himself, in his Almagest, when he states that the oikoumene occupies approximately one fourth of the earth surface bounded by half of the equator and the halves of two opposite meridians. The amplitude of 180° is justified in the following passage:

> In the case of longitude (that is in the east-west direction) the main proof is that observations of the same eclipse (especially a lunar eclipse) by those at the extreme western and extreme eastern regions of our part of the inhabited world (which occur at the same [absolute] time) never differ by more than twelve equinoctial hours [in local time]; and the quarter [of the earth] contains a twelve-hour interval in longitude, since it is bounded by one of the two halves of the equator. [15]

The reference to the method proposed by Hipparchus for measuring differences of longitude from local times of lunar eclipses[16] gives a clue in support of our hypothesis, which is also consistent with the fact that the amplitude of 180° for the oikoumene was already transmitted, before Ptolemy and Marinus, by Strabo. Indeed, shortly after having recalled the method used by Hipparchus for determining differences of longitude (*Geographica*, I, 1, 12), Strabo observes that people living in the extreme eastern regions are in a sense the antipodes of those living at the extreme west of Iberia (*Geographica*, I, 1, 13).

If the amplitude of the oikoumene was taken from Hipparchus, and was at the origin of the wrong estimate of the Earth's circumference, it is understandable that

---

[15]ἐπὶ δὲ τοῦς μήκους, τουτέστιν τῆς ἀπὸ ἀνατολῶν πρὸς δυσμὰς παρόδου, διὰ τοῦ τὰς αυτὰς ἐκλείψεις, μάλιστα δὲ τὰς σεληνιακὰς, παρά τε τοῖς ἐπ' ἄκρων τῶν ἀνατολικῶν μερῶν τῆς καθ'ἡμᾶς οἰκουμένης οἰκοῦσι καὶ παρὰ τοῖς ἐπ' ἄκρων τῶν δυτικῶν κατὰ τὸν αυτὸν χρόνον θεορουμένας μὴ πλέον δώδεκα προτερεῖν ἢ ὑστερεῖν ὡρ˜ν ἰσημερινῶν αὐτοῦ κατὰ μῆκος τοῦ τεταρτημορίου δωδεκάωρον διάστημα περιέχοντος, ἐπειδήπερ ὑφ'ἑνὸς τῶν τοῦ ἰσημερινοῦ ἡμικυκλίων ἀφορίζεται. (Ptolemy, *Almagest*, II, 1, p.88 Heiberg, 10–19; translation by G. J. Toomer.)

[16]Obviously it is not possible to observe the same lunar eclipse from two opposite semi-meridians (or not without assuming unrealistic observations made from points close to a pole), but it *is* possible to obtain the total longitude between them by summing two (or more) smaller longitude differences.

this last error was already made by Posidonius, whose main source in astronomical and geographical matters was probably Hipparchus himself. [17] Furthermore, the rejection of the celebrated result by Eratosthenes becomes understandable if it was girded by a belief that one was following Hipparchus, the famous scientist whose successful criticism of Eratosthenes was well known.

Finally, we can notice that the discrepancy among the ratio 1.4 between the two estimates for the Earth's circumference and the value 1.428 of the regression coefficient concerning the longitude dilation, though small, is not negligible and hardly compatible with A*, in view of the high value of R^2. This discrepancy becomes easily understandable by adopting the implication A**. A contraction of Eratosthenes' value of 700 stadia per degree by a factor 1.428 leads indeed to a value of 490.2 stadia per degree, but it is reasonable that Ptolemy (and Posidonius before him), having to replace an estimate given by a round figure,[18] wanted to select an equally round figure, therefore choosing 500.

4. Global and local errors

Let us go back to chapter I, 4 of Ptolemy's *Geographia*:

> [...] most distances, especially the east-west ones, have come down to us in a less precise form, not through the negligence of those who devoted themselves to research, but perhaps because they had not yet understood the usefulness of more scientific methods and because they had not observed many lunar eclipses at the same time in different places (such as the one that was seen in Arbela at the fifth hour and at Carthage at the second hour), from which it would have been clear how many equinoctial time units separated the localities to the east or west.[19]

The mention of people who had "not yet" (μηδέπω) understood the usefulness of the method based on lunar eclipses for determining the longitudes implies that

---

[17]We recall that Hipparchus worked in the same Rhodes where Posidonius, a generation later, established his school.

[18]The aforementioned passage from Pliny about the stadium according to the "Eratosthenis ratio" suggests that in the case of Eratosthenes the round figure could have been the result of the definition of the new stadium as a convenient submultiple of the circumference of the Earth (252,000 is a particularly convenient number, because it is divisible by all numbers from 1 to 10, whose least common multiple is in fact 2,520).

[19]τὰ δὲ πλεῖστα τῶν διαστημάτων καὶ μάλιστα τῶν πρὸς ἀνατολὰς ἢ δυσμὰς ὁλοσχερεστέρας ἔτυχε παραδόσεως, οὐ ῥαθυμίᾳ τῶν ἐπιβαλόντων ταῖς ἱστορίαις, ἀλλ᾽ ἴσως τῷ μηδέπω τὸ πρόχειρον κατειλῆφθαι τῆς μαθηματικωτέρας ἐπισκέψεως, καὶ διὰ τὸ μὴ πλείους τῶν ὑπὸ τὸν αὐτὸν χρόνον ἐν διαφόροις τόποις τετηρημένων σεληνιακῶν ἐκλείψεων, ὡς τὴν μὲν Ἀρβήλοις πέμπτης ὥρας φανεῖσαν, ἐν δὲ Καρχηδόνι δευτέρας, ἀναγραφῆς ἠξιῶσθαι, ἐξ ὧν ἐφαίνετ᾽ ἂν πόσους ἀπέχουσιν ἀλλήλων οἱ τόποι χρόνους ἰσημερινοὺς πρὸς ἀνατολὰς ἢ δυσμάς· (Ptolemy, Geographia, I, 4, 1).

such usefulness was clear to later scholars. Ptolemy indeed had used such astro-
nomical method in the case of Arbela and Carthage[20] and most probably, as already
observed, in some other cases.

Suppose that Ptolemy had obtained all his longitudes by dividing the (wrong)
amplitude of the oikoumene assumed by him in parts proportional to the known
distances along circles of latitude. In this case, and in the absence of large sys-
tematical errors in distances, we should expect that Ptolemy's longitudes were
well approximated everywhere by the same linear function of the actual ones, i.e.,
all differences of longitude are dilated almost exactly in the same proportion the
oikoumene as a whole is.

But if instead, as suggested by Ptolemy himself, the interpolation procedure
started after the insertion of some milestone astronomical data — which would be
very unlikely to fit well with the linear relation above[21] — then the set of all the
locations considered would have been broken into regional subsets such that:

(a) In each subset the longitudes, being obtained with an interpolation procedure,
    are very close to a linear function of the actual longitudes, so that we should
    get in any case very high values of the coefficients of determination R^2.

(b) The regression coefficients corresponding to different subsets should differ
    considerably from each other.

(c) The longitudes inserted on an astronomical basis should mark the fracture
    points between different subsets.

(d) The deviations of the different regression coefficients from the global dilation
    ratio should compensate each other in such a way that the set of all the longi-
    tudes can be well approximated by a linear regression with a coefficient equal
    to such dilation ratio.

To these points, one can add a further observation. The data given by Ptolemy
are interconnected by a very complex chain of relations. In particular, he states
(*Geographia*, I, 4, 1) that in some cases he has the information that two locations
are on the same meridian. Considering this kind of interconnection in relation with
the previous reasoning, it is clear that:

(e) There exists the possibility that two sets of data covering more or less the same
    longitude area (but coming from different latitudes), are broken in subsets in
    such a way that the fracture points are approximately at the same longitude
    for both.

This scenario, suggested by Ptolemy's exposition, describes well the features of
the longitudes reported in his *Geographia*, as is apparent from the analysis made

---

[20]See next page.

[21]Of course we cannot exclude the possibility that Ptolemy discarded data too far from it.

in [Shcheglov 2014]. This analysis, which allowed Shcheglov to correctly discard thesis A*, is not only consistent with A**, but gives a strong argument in its support once one takes into account that, in addition to the systematic dilation, significant and well characterized errors due to the insertion of some longitudes deduced from astronomical data are to be expected. In this regard, the fact that the sample chosen by Shcheglov as a whole leads to a value of R^2 that is close to 1 (0.9874) should not be, in my opinion, dismissed as "deceptive" (as Shcheglov does), but provides a key element for the reconstruction of the overall framework.

We cannot hope to identify all the cases in which Ptolemy used astronomical data concerning longitudes, but in one significant case it is possible to reasonably conjecture the origin of a local error. Let us start by observing that the presence of local regression coefficients which largely differ from the global one is especially significant in areas that were well-known in Hellenistic times. In this regard, the data coming form the south and north coasts of the Mediterranean sea are particularly relevant. Analyzing the data shown in Figure 6 of [Shcheglov 2014], one can see that both the southern and the northern coasts of the Mediterranean Sea can be divided in two parts, eastern and western, which feature substantially different values for the regression coefficient. Specifically, the regression coefficients are 1.81 and 1.19 for the western and eastern part of the south coast; 1.77 and 1.15 for the western and eastern part of the north coast. It is also striking that the fracture points between the two different slopes for the north and south coasts[22] are situated at approximately the same longitude — in modern terms, slightly more than 10°E. A good explanation for that fact would be that, for some location at a longitude slightly more than 10°E, Ptolemy had an astronomical datum that was inconsistent with the systematic dilation of longitudes. It is then perhaps not by chance that that longitude corresponds very well to one of the endpoints (Carthage, 10°19′) of the only longitude interval for which Ptolemy explicitly states that he possessed an astronomical datum. The value given by Ptolemy for the difference in longitude between Arbela and Carthage (45°10′) in fact agrees very closely with the one corresponding to the difference of three hours mentioned in the previous passage. On the other hand, this value is dilated by a factor of approximately 1.30 with respect to the actual difference, a value which significantly differs from the global coefficient of regression.[23] If one considers that the longitude given by Ptolemy for Arbela places it almost exactly on the global regression line, this strongly suggests that Ptolemy, taking into account the astronomical datum for the

---

[22]See point (e) on the previous page.

[23]We notice that, as a consequence, the longitudinal distance of Carthage from the Pillars of Heracles is dilated by Ptolemy by the factor 1.74, in good agreement with the regression coefficients (1.81 and 1.77) found by Shcheglov for the western part of the Mediterranean Sea.

determination of the longitude of Carthage,[24] may have had to adjust the other data which were linked to it by relations between distances or other more complex kinds of interconnections. This hypothesis can explain:

 (i) the different slopes observed for eastern and western areas;

 (ii) the high values of R^2 for both subsets;

(iii) the fact that the regression coefficient for the union of the considered subsets agrees with a good approximation with the global dilation ratio;[25]

(iv) the fact that the northern and southern Mediterranean coasts each break into different slope subsets more or less at the same longitude.[26]

A further confirmation of this reconstruction comes from a comparison between the global regression coefficient (1.428) and the ratio between the actual longitudinal amplitude for the Ptolemy's oikoumene and the value of 180° accepted by him.

The westernmost locations considered by Ptolemy are the Islands of the Blessed, which he identified with the Canary Islands.[27]

In order to estimate the global dilation of the oikoumene, on the other extreme we consider the Sera Metropolis (today's Xi'an[28]), since more eastern locations

---

[24]The poor accuracy of this particular value is easily understandable if we take into account that it relies on non-scientific descriptions of the eclipse, dating to the fourth century BC, two centuries before that Hipparchus suggested that this kind of data could be used to determine differences of longitude.

[25]A linear regression performed over the whole set of locations indicated in [Shcheglov 2014] as the "north coast" and "south coast" of the Mediterranean gives the line $y = 1.437x + 17.00$, with $R^2 = 0.980$. The very low values of the regression coefficients found by Shcheglov for the eastern part of the Mediterranean Sea (1.19 and 1.15) are thus the right ones to compensate the overestimation of the dilation on the other side.

[26]See again point (e) above.

[27]This identification was usual in imperial times. It appears for the first time implicitly in the work of Pomponius Mela, who places the Islands of the Blessed in front of the Atlas Mountain (*De chorographia*, III, 101–102). In Pliny's *Naturalis Historia* (VI, 202–203) the identification is even clearer, since he places them in front of Mauretania and gives with a good approximation their distances from Gades (Cadiz). As for Ptolemy (who hardly dares to question knowledge widely accepted in his days), he gives names and coordinates of six "Islands of the Blessed" (Μακάρων νῆσοι) (*Geographia*, IV, 6, 34). Their identification with the Canaries is implied by their longitude, and more importantly by the names of the islands: three of Ptolemy's names are obvious Greek correspondents of latin names given by Pliny: Canaria, Junonia and Pluvialia. The slight discrepancy between the number given by Ptolemy (six) and the actual number (seven) of the major islands of the archipelago can be explained in many ways. Almost all scholars agree with this identification; see, for instance, [Stückelberger and Graßhoff 2006, 455, footnote 200]. Nevertheless, a few scholars have questioned the identification with the Canaries, on the basis of the latitude given by Ptolemy, which is very far from that of the Canaries. We shall return to this point.

[28]For the identification of Sera Metropolis with Xi'an see [Stückelberger and Graßhoff 2006, 669, note 229].

are not unanimously identified today. Since the actual longitudes of Xi'an and the Canary Islands [29] are respectively 108°54′E and 15°W, and the corresponding values given by Ptolemy are 177°15′ and 0°, the ratio between their difference of longitude given by Ptolemy and the corresponding actual value is approximately 1.43, and thus very close to the regression coefficient given in [Russo 2013a] and [Shcheglov 2014].

Of course a detailed reconstruction of the procedure followed by Ptolemy is very difficult, and probably impossible, since it should also take into account the presence of local errors due to the different level of geographical knowledge for the various regions, and the complex interconnections between his data, already alluded to. Nevertheless, as we saw, some firm points could be established, the most important of which is a general dilation of all the differences of longitude by a mean factor equal to the dilation of the whole oikoumene. The implication A** is thus confirmed.

However, we still did not tackle the point on which our first criticism to A* was based, i.e, its failure to explain the origin of the double mistake, which is left by that hypothesis in the darkest obscurity. We will devote the next section to test the proposed thesis A** against this last problem.

5. The origin of the error on longitudes by Ptolemy

We conjectured that the value of 180° for the amplitude of the oikoumene, i.e., the difference in longitude between the Islands of the Blessed and the easternmost regions, was taken from a Hellenistic source (most probably Hipparchus, who may well have intended it as a rounded value). On the other hand we know that Hellenistic scientists, and in particular Hipparchus, had accepted the value of Eratosthenes for the Earth's circumference (Strabo, *Geographica*, I, iv, 1; II, v, 7; II, v, 34), while the smaller estimate, grossly wrong, was first introduced (as far as we know) by Posidonius (Strabo, *Geographica*, II, ii, 2). Why does this value of 180°, which according to Hipparchus was consistent with the measurement by Eratosthenes, coexist with a much "smaller" Earth in (among others) Ptolemy? Since there is substantial agreement in the distances along the circles of latitude between Ptolemy's and Eratosthenes' data, it is clear that the only possibility is that there was a misinterpretation on the identification of one of the two extrema of the oikoumene. Two arguments allow us to exclude that the issue involved locations in the far East:

---

[29]As the modern value for the Canary Islands we take the value of 15°W, which is the best rounded value for the average longitude.

1. It is a priori easier to misinterpret the identification of oceanic islands, for which no close locations are known for very large distances, than that of a city which is reachable through a series of intermediate locations.

2. To produce an error of the order of magnitude of that made by Ptolemy, the original location of some town in China assumed by his source as the eastern extremum of the oikoumene had to be very far into the Pacific Ocean.

The only remaining possibility, therefore, is that the two errors were originated by an erroneous identification of the Islands of the Blessed.

The previous reasoning, which up to now was intended at analyzing the origin of the errors in Ptolemy's *Geographia*, can also disclose new possibilities in a more general (and possibly more relevant) issue. We are indeed led to the conclusion that the Islands of the Blessed, to which the Hellenistic source of Ptolemy (most probably Hipparchus) referred, can be found approximately on the opposite semi-meridian with respect to the more eastern locations cited in the *Geographia*. This entails the need to consider longitudes that are much farther on the west than those believed as known in Hellenistic times. We arrived at this conclusion by elimination, after having considered and discarded every possible alternative. Therefore, in my opinion, we should seriously consider the extreme consequences of the previous reasoning. As a renowned investigator used to say, when you have eliminated the impossible, whatever remains, however improbable, must be the truth.

6. The Islands of the Blessed and the Caribbean

The elimination process illustrated before conduced us to the conclusion that the first cause for the errors made by Ptolemy was the wrong identification of the Islands of the Blessed with the Canary Islands. One may wonder how such a misinterpretation of the sources was possible. It is therefore useful, in my opinion, to briefly recall some historical processes which played a key role in this connection.

Mentions of travels in the Atlantic Ocean are not negligible in ancient sources.[30] However, between the Hellenistic age and Ptolemy, a significant loss of geographic knowledge concerning this Ocean occurred, and in a short time even well known descriptions of voyages started to be considered unreliable. The report of the famous expedition made by Pytheas, for instance, in which he described the iced sea and the midnight sun, was considered trustworthy by Hellenistic scientists such as Eratosthenes and Hipparchus,[31] but was later rejected by Strabo (*Geographica*, II, iv, 1; II, iv, 2; II, iii, 5), was not copied and transmitted anymore, and finally lost.

---

[30]For a review of classical sources on travels in the Atlantic Ocean see [Roller 2006].

[31]See, among other passages, Hipparchus, *In Arati et Eudoxi phaenomena commentariorum libri iii*, I, 4, 1; *Strabo, Geographica*, I, 1, 9; II, 1, 12; II, 4, 2.

The error made by Ptolemy according to our reconstruction was made possible by the combination of the loss of knowledge about the Atlantic Ocean with a general weakening of science in the Roman world with respect to the Hellenistic age. [32] Geography, in particular, was transformed radically, from the mathematically founded subject it was in Hellenistic times to the purely descriptive one it became in works like those by Strabo and Polybius. An error like the misidentification of an archipelago became clearly much more probable once the use of spherical coordinates (latitude and longitude) for the identification of the locations had been abandoned, as it was in the first century BC. The oldest evidence concerning the identification of the Islands of the Blessed with the Canaries dates indeed from the first century BC, [33] and then was inherited by Ptolemy when he attempted at reconstructing quantitatively the mathematical geography for the first time after the methodological crisis had occurred. Posidonius seems the most probable candidate as the source of the misidentification. He is, to our knowledge, the oldest source giving the length of 500 stadia per degree for the Earth's circumference, which tends to suggest that the transition from the "old" to the "new" value occurred in his work.

A quantitative analysis of the data given by Ptolemy strongly supports the idea of a mistaken identification of the archipelago. Indeed:

1. The latitude he gives for the Islands is wrong by about 15°, a huge error which can be regarded as incompatible with any data coming from real measurements and can be only explained as the result of a confusion between two different archipelagos.

2. The Canaries are spread over a total longitude of about four and a half degrees, while the archipelago considered by Ptolemy covers just one degree in west-east direction; moreover, the Canary Islands cover less than two degrees in north-south direction, while Ptolemy's Islands five and a half. In conclusion, both archipelagos have a strip-like shape, but they are approximately oriented in orthogonal directions.

A decisive test in order to check the proposed thesis is at this point possible. We can indeed compute the original longitude of the Islands of the Blessed simply taking a difference of longitude of 180° with respect of Ptolemy's eastern extreme locations, and verify whether we come close to some archipelago. The longitude of the eastern bound of Ptolemy's oikoumene can be computed by the regression line (1) given on page 183, solving the equation

$$180 = 1.428x + 17.06.$$

---

[32]On this point see [Russo 2004, passim].

[33]See note 27 above.
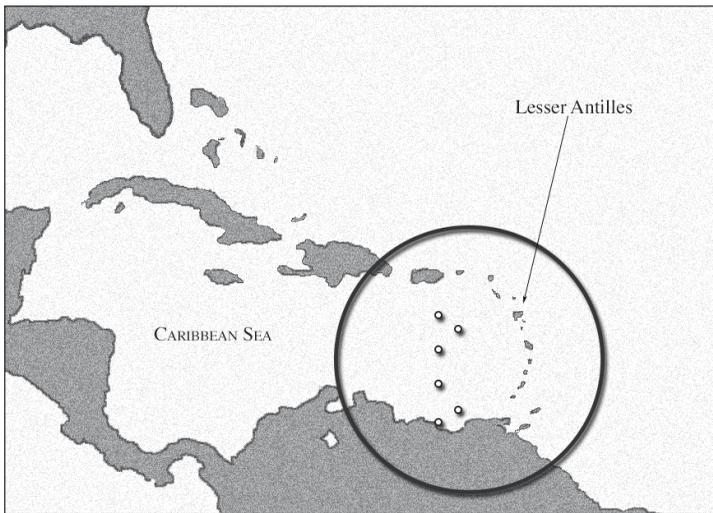
Figure 1. White dots indicate the original locations of the Islands of the Blessed as computed in the text.

In this way, we eventually get the following coordinates for the Islands:

65°54′W, 16°N;   64°54′W, 15°15′N;   65°54′W, 14°15′N;
65°54′W, 12°30′N;   64°54′W, 11°N;   65°54′W, 10°30′N.

These locations are marked by circles in Figure 1.

As the reader can see, not only do we come very close to an actual archipelago (formed by the Leeward Islands and the Windward Islands, in the Lesser Antilles), but also its dimension, shape and orientation correspond well to the dots. If we compute the eastern extreme using the regression line (2), given in [Shcheglov 2014], we obtain an even slightly better match with the Leeward and Windward Islands, since all the points move east by about 0.44 degrees.

The available descriptions of the Islands of the Blessed in the sources can provide further arguments in support of this proposed (and admittedly surprising) identification. In this regard, one can observe that there is a substantial consistency between the numerous descriptions we have dating from the archaic and classical periods, and that these descriptions are hardly reconcilable with the identification with the Canaries. The issue was examined in detail in [Manfredi 1993]; in his opinion, some striking characteristics mentioned frequently in the sources (and especially the presence of lush and evergreen foliage) can exclude the possibility that they referred to the Canaries, which are generally bleak, and would rather suggest islands in the Caribbean sea [Manfredi 1993, p. 204].

In the following section a further independent and quantitative confirmation of the proposed thesis will be provided.

7. A passage from Pliny

The following passage by Pliny has eluded, until now, any reasonable interpretation:

> Regarding the length and the breadth this is what I deem worthy of mention. For the whole circumference, Eratosthenes, a man highly regarded by all and surpassing others in every subtlety of learning, and especially in the present matter, gave the value of 252,000 stadia. [. . .] Hipparchus, a man to be admired for taking issue with him and for much more besides, then added to that number a little less than 26,000 stadia.[34]

The last statement by Pliny is contradicted by passages in Strabo to the effect that Hipparchus accepted the value given by Eratosthenes for the Earth's circumference (*Geographica*, I, iv, 1; II, v, 7; II, v, 34). The key to understanding the passage above, in my opinion, lies in its first words. Just before that passage Pliny's topic was the dimensions of the oikoumene. Since it makes no sense to talk of the length and the breadth of a spherical Earth, it is therefore very plausible that Pliny's source was still discussing the oikoumene when talking about the 26,000 stadia, and that Pliny misunderstood the reference as being about the earth's circumference because that was Eratosthenes' most celebrated measurement. The length and breadth of the oikoumene were typically discussed in geographical works, and we know from Strabo (*Geographica*, I, iv, 5) that Eratosthenes did calculate the length of the oikoumene along the parallel of Athens, getting the result of 77,800 stadia, of which 5,000 to the west of the Pillars of Heracles. Pliny's passage can therefore be explained if we conjecture that Hipparchus extended the oikoumene in longitude by adding just under 26,000 stadia to the value given by Eratosthenes. Since this extension could hardly have concerned the Pacific Ocean, we must conclude that the western boundary of the oikoumene according to Hipparchus was at 31,000 (26,000 + 5000) stadia from the Pillars of Heracles along the parallel of Athens. Since the cosine of the latitude of Athens is about 0.788, a degree of longitude along this parallel, according to Eratosthenes, corresponds to about $700 \times 0.788 = 552$ stadia, which yields a location approximately $56°10'$ the west of Gibraltar, corresponding to a longitude of $61°31'$ W. The corresponding semi-meridian is shown in

---

[34]De longitude ac latitude haec sunt, quae digna memoratu putem. Universum autem circuitum Eratosthenes, in omnium quidem literarum subtilitate, et in hac utique praeter ceteros sollers, quem cunctis probari video, CCLII milium stadiorium prodidit, [. . .] Hipparchus, et in coarguendo eo, et in reliqua omni diligentia mirus, adicit stadiorum paulo minus XXVI milia (Pliny, *Naturalis Historia*, II, 247).
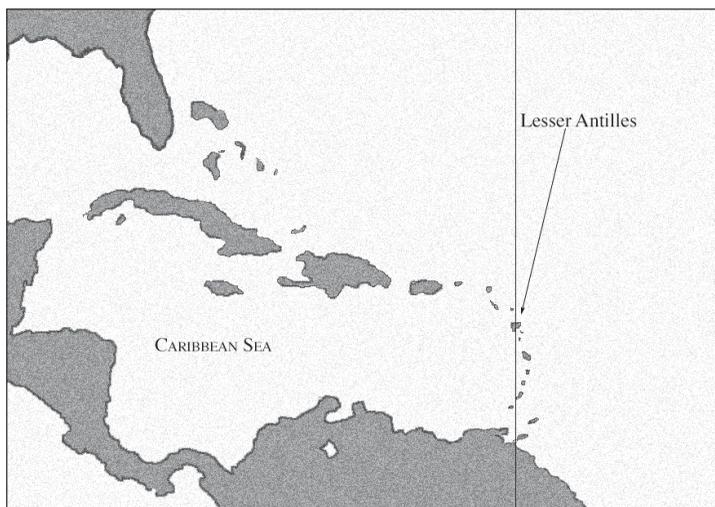
Figure 2. The westernmost longitude of the oikoumene according to Hipparchus, as recovered from Pliny's passage.

Figure 2, and represents a striking confirmation of the thesis that Hipparchus knew the coordinates of the Leeward and Windward Islands, and also of the correctness of our reconstruction of the meaning of the passage by Pliny. Moreover, it indirectly supports the idea that the source of Ptolemy on the Islands of the Blessed was Hipparchus. We notice that it is not surprising that the match between theoretical previsions and actual geographical data is much better in Figure 2 than in Figure 1, since the reconstruction underlying Figure 1 had a statistical basis and relied on the value of 180° for the amplitude of the oikoumene, which was possibly rounded off, while in the case of Figure 2 the method only uses one quantitative datum taken from the sources.

Of course, the addition of 26,000 stadia by Hipparchus should be considered not as a correction to the value given by Eratosthenes, but rather an update due to new geographical discoveries.

8. Conclusions

We want here to summarize the conclusions of the present paper and the arguments that can be considered in their support.

The main thesis can be expressed as follows:

Some source of Ptolemy's Geographia knew with remarkable precision the position of some locations in the Caribbean Sea, i.e. the Leeward and Windward Islands. These islands were the locations originally intended

as the "Islands of the Blessed". Their misidentification with the Canaries first occurred in the first century BC and was then accepted by Ptolemy.

This thesis, which is plausible in view of the loss of knowledge concerning the Atlantic Ocean and the lapse into disuse of spherical coordinates between Hipparchus and Ptolemy, is justified by abductive inference based on its ability to explain a number of facts for which so far no explanation had been offered:

1. The thesis is implied by a procedure performed by Ptolemy which is suggested by his own words and in turn explains both (i) the very high values for the determination coefficient (and the virtually identical regression coefficients) found in [Russo 2013a] and [Shcheglov 2014], and (ii) the significant differences between regression coefficients relative to different regions shown in [Shcheglov 2014].

2. It explains why the archaic and classical descriptions of the Islands of the Blessed match better with Caribbean islands rather than the Canaries (as discussed in [Manfredi 1993]).

3. It explains why Ptolemy's coordinates for the Canary Islands describe an elongated archipelago whose orientation is basically orthogonal to the actual one — in other words, as wrong as it could be).

4. It explains the huge error of 15° made by Ptolemy in the latitude of the Canary Islands.

5. It provides a simple, and quantitatively accurate, explanation of the systematic dilation of differences of longitudes operated by Ptolemy.

6. It explains as a simple consequence of the previous point the new estimate in the measure of the Earth accepted by Posidonius and Ptolemy — even more so because the ratio between the old and the new estimates is close to, but not identical with, the dilation coefficient.

7. It explains the striking match between the shape and the position of the Leeward and Windward Islands and the locations indicated by Ptolemy as the "Islands of the Blessed", once their original coordinates are reconstructed by means of the statistical approach herein considered.

8. It provides a simple (and quantitatively accurate) explanation for the passage of Pliny discussed above.

References

Aujac, Germaine (1966), *Strabon et la science de son temps*. Les sciences du monde, Paris.

Berggren, J. Lennart and Alexander Jones (2000), *Ptolemy's Geography: an annotated translation of the theoretical chapters*, Princeton and Oxford.

Columba, Gaetano Mario (1895), *Eratostene e la misurazione del meridiano terrestre*, Palermo.

Czwalina, Arthur (1925), *Die Gradmessung des Eratosthenes*, in *Hirts Literatur-Bericht* **20**, 29–295.

Dicks, D. R. (1960): *The geographical fragments of Hipparchus*, edited with an introduction and commentary, London.

Dreyer, John Louis Emil (1953), *A history of astronomy from Thales to Kepler*, 2nd ed., New York.

Dutka, Jacques (1993/94), "Eratosthenes' measurement of the Earth reconsidered", *Archive for the History of Exact Sciences* **46**, 55–66.

Fraser, Peter Marshall (1972), *Ptolemaic Alexandria*, 3 vols., Oxford.

Gossellin, Pascal François Joseph (1790), *Géographie des Grecs analysée*, Paris.

Hultsch, Friedrich Otto (1882), *Griechische und römische Metrologie*, 2nd ed., Berlin.

Jomard, Edme François (1817), *Mémoire sur le système métrique des anciens égyptiens, contenant des recherches sur leurs connoissances géométriques et sur les mesures des autres peuples de l'antiquité*, Paris.

Letronne, Antoine Jean (1851), *Recherches critiques, historiques et géographiques sur les fragments d'Héron d'Alexandrie ou du systéme métrique égyptien*, Paris.

Manfredi, Valerio Massimo (1993), *Le Isole Fortunate: topografia di un mito*, Roma 1993.

Meuret, Christophe (1998), "Outils mathématiques et données itineraires: reflexions sur evaluation de la circonférence terrestre chez Ptolémée", in: P. Arnaud / P. Counillon (eds), *Geographica Historica*, Bordeaux, Nice, 151–166.

Miller, Konrad (1919), *Die Erdmessung im Altertum und ihr Schicksal*, Stuttgart.

Oxé, August (1963), *Die Maßtafel des Julianus von Askalon*, in: *Rheinisches Museum für Philologie* **106**, 264–286.

Pothecary, Sarah (1995), "Strabo, Polybios, and the Stade", *Phoenix* **49**, 49–67.

Rawlins, Dennis (1981), "The Eratosthenes–Strabo Nile map: is it the earliest surviving instance of spherical cartography? did it supply the 5000 stades arc for eratosthenes experiment?", *Archive for the History of Exact Sciences* **26**, 211–219.

Rawlins, Dennis (1985), "Ancient geodesy: achievement and corruption", *Vistas in Astronomy* **28**, 255–268.

Rawlins, Dennis (2008), "The Ptolemy GEOGRAPHY's Secrets", *DIO* **14**, 33–58, http://www.dioi.org/vols/we0.pdf

Roller, Duane W. (2006), *Through the Pillars of Herakles: Greco-Roman explorations of the Atlantic*, New York and London.

Russo, Lucio (2004), *The forgotten revolution*, Springer, Heidelberg/New York.

Russo, Lucio (2013a), "Ptolemy's longitudes and Eratosthenes' measurement of the Earth's circumference", *Math. Mech. Complex Systems* **1**, pp. 67–79, DOI 10.2140/memocs.2013.1.67

Russo, Lucio (2013b), *L'America dimenticata: i rapporti tra le civilta' e un errore di Tolomeo*, 2nd ed., Mondadori Università, Milano.

Shcheglov, Dmitry (2014), "The error in longitude in Ptolemy's *Geography* revisited", *Cartographic Journal* **09/2014**. DOI: DOI 10.1179/1743277414Y.0000000098

Stückelberger, A. (1988), *Einführung in die antiken Naturwissenschaften*, Darmstadt.

Stückelberger, Alfred and Gerd Graßhoff (Eds.) (2006). *Klaudios Ptolemaios: Handbuch der Geographie. Griechisch – Deutsch. Einleitung, Text und Übersetzung*, 2 vols. + CD-ROM, Schwabe Verlag, Basel.

Tannery, Paul (1893), *Recherches sur l'histoire de l'astronomie ancienne*, Paris (repr. Hildesheim, New York, 1976).

Tupikova, Irina (2013), "Ptolemy's World Map and Eratosthenes's circumference of the Earth", in *Proceedings of the 26th International Cartographic Conference* (Dresden, August 2013), https://tinyurl.com/Tupikova-Ptolemy

Tupikova, Irina and Klaus Geus (2013), "The circumference of the Earth and Ptolemy's world map", Preprint 439, Max-Planck-Institut für Wissenschaftsgeschichte, Berlin, http://www.mpiwg-berlin.mpg.de/Preprints/P439.PDF

Valerio, Vladimiro (2013), "Spunti e osservazioni dal libro di Lucio Russo *L'America dimenticata: i rapporti tra le civiltà e un errore di Tolomeo*", *Geostorie* **21**:1–2, gennaio–agosto 2013, 77–97.

LUCIO RUSSO: russo@mat.uniroma2.it
Dipartimento di Matematica, Università degli Studi di Roma Tor Vergata, 00173 Roma, Italy
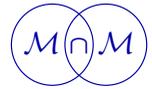
GENERATION OF SH-TYPE WAVES DUE TO SHEARING STRESS DISCONTINUITY IN AN ANISOTROPIC LAYER OVERLYING AN INITIALLY STRESSED ELASTIC HALF-SPACE

SANTOSH KUMAR AND DINBANDHU MANDAL

The paper investigates the generation of SH-type waves due to a sudden application of a stress discontinuity which moves after creation at the anisotropic layer of finite thickness overlying an initially stressed isotropic half-space. The displacements are obtained in exact form by the method due to Cagniard modified by de Hoop. Two cases of shearing stress discontinuities are considered. The numerical results are obtained for a particular model and discussed by plotting graphs for displacement component with the elapsed time of the disturbance for different values of initial stress and also for different values of initial time at which pulses arrive.

1. Introduction

The notion of initial stress is essential to the study of seismic wave propagation. Biot [1940] is largely responsible for the notion's introduction and initial applications to elastic wave propagation; in [Biot 1965] he further developed the notion of initial stress. Many authors have used that book as fundamental to the study of wave propagation in an initially stressed medium. Abd-Alla and Ahmed [1999] analyzed Love waves propagation in a non-homogeneous orthotropic elastic layer under initial stress overlying semi-infinite medium. Khurana [2001] considered the effect of initial stress on the propagation of Love wave. Further significant steps were taken in [Das and Dey 1968; 1970, Dey 1971; Dey and Addy 1978; Chattopadhyay and De 1981; Chattopadhyay and Kar 1978; Majhi et al. 2016; 2017] to cite but a few works.

In addition to initial stress, shearing stress discontinuity also plays a vital role in the study of seismic wave propagation. Pal [1983] considered the problem of generation and propagation of SH-type waves due to non-uniformly moving stress

discontinuity in layered anisotropic elastic half-space using Garvin's [1956] techniques, which are a modification of Cagniard's [1939] technique. Pal and Kumar [2000] considered the generation of SH waves by a moving stress discontinuity in an anisotropic soil layer over an elastic half-space using the Cagniard–de Hoop special reduction technique [de Hoop 1960]. Next de Hoop [2002] considered the reflection and transmission properties of an elastic interfacial bonding of two semi-infinite solids, investigated for the simplest possible case of a line-source excited two-dimensional SH-wave. Pal and Mandal [2014] studied the generation of SH-type waves due to sudden application of a stress discontinuity which moves after creation at the sandy layer of finite thickness overlying an isotropic and inhomogeneous elastic half-space. Mandal et al. [2014] considered the disturbance and propagation of SH-type waves in an anisotropic soil layer overlying an inhomogeneous elastic half-space by a moving stress discontinuity. All authors have considered the effect of initial stress and shearing stress discontinuity separately but haven't considered the initial stress and shearing stress discontinuity together.

In the present problem our intention is to investigate the two dimensional problem of generation of SH-type waves at the free surface of an anisotropic layer due to an impulsive stress discontinuity moving with uniform velocity along the interface of initially stressed isotropic medium. The displacement is calculated numerically for two particular distances on the surface for two different types of the discontinuity in the shearing stress for different value of initial stress. It involves Laplace and Fourier transform and the inversion is based on Garvin's [1956] method. The problem discussed may be of importance in connection with the propagation of cracks in the layer. Two cases of stress discontinuity are considered and the numerical results are shown graphically.

2. Formulation of problem

We consider an anisotropic elastic layer of thickness h with elastic constants L_1, N_1 and density ρ_1 over an initially stressed isotropic half-space with elastic constant μ_2 and density ρ_2. The interface of these two media is considered at $z = 0$ whereas free surface is at $z = -h$. Here, z axis is directed vertically downward and x axis is assumed in the direction of the propagation of wave with velocity c. For SH-type of waves the displacement does not depend on y and if (u, v, w) be the displacement at any point $P(x, y, z)$ into the medium then $u = w = 0$ and v are function of x, z and t. The two equations of motion are identically satisfied. The geometrical configuration is depicted in Figure 1.

The equation of motion for the anisotropic layer (Medium I) without body forces is given by

$$N_1 \frac{\partial^2 v_1}{\partial x^2} + L_1 \frac{\partial^2 v_1}{\partial z^2} = \rho_1 \frac{\partial^2 v_1}{\partial t^2}. \tag{1}$$
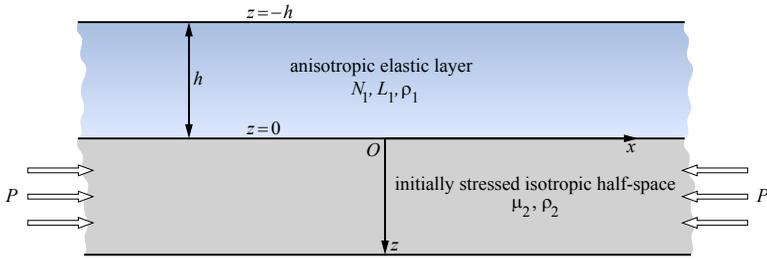
Figure 1. Geometry of problem.

The stress strain relation is given by $(\tau_{xy})_{\mathrm{I}} = N_1 \frac{\partial v_1}{\partial x}$ and $(\tau_{yz})_{\mathrm{I}} = L_1 \frac{\partial v_1}{\partial x}$.

The equation of motion for initially stressed isotropic half-space (Medium II) without body forces is given by

$$\frac{\partial (\tau_{xy})_{\mathrm{II}}}{\partial x} + \frac{\partial (\tau_{yz})_{\mathrm{II}}}{\partial z} - \frac{P}{2} \frac{\partial v_2}{\partial x^2} = \frac{\partial^2}{\partial t^2}(\rho_2 v_2). \qquad (2)$$

The stress strain relation is given by $(\tau_{xy})_{\mathrm{II}} = \mu_2 \frac{\partial v_2}{\partial x}$ and $(\tau_{yz})_{\mathrm{II}} = \mu_2 \frac{\partial v_2}{\partial z}$.

The boundary conditions are

$$(\tau_{yz})_{\mathrm{I}} = 0 \qquad \qquad \text{at } z = -h, \qquad (3)$$

$$v_1 = v_2 \qquad \qquad \text{at } z = 0, \qquad (4)$$

$$(\tau_{yz})_{\mathrm{I}} - (\tau_{yz})_{\mathrm{II}} = S(x, t) H(t) \quad \text{at } z = 0, \qquad (5)$$

where $S(x, t)$ is a function of x and t; $H(t)$ is the Heaviside unit function of time t.

3. Method of solution

The above problem can readily be solved by using the Laplace and Fourier transforms combined with the modified Cagniard–de Hoop [1960] method. The Laplace transform with respect to t and the Fourier transform with respect to x are defined by

$$\widetilde{\bar{v}} = \int_{-\infty}^{\infty} e^{-i\xi x} dx \int_{0}^{\infty} e^{-pt} v(x, z; t) \, dt. \qquad (6)$$

We can easily get the upper and lower layer with $v_2 \to 0$ as $z \to \infty$ in the form

$$\bar{v}_1(x, z; p) = \int_{-\infty}^{\infty} (A \cosh s_1 z + B \sinh s_1 z) \, e^{i\xi x} d\xi, \qquad (7)$$

$$\bar{v}_2 = \int_{-\infty}^{\infty} C e^{(i\xi x - s_2 z)} d\xi, \qquad (8)$$

where the constants A, B and C are to be determined from the boundary conditions

(3)–(5):

$$s_1 = \left(\phi_1^2\xi^2 + \frac{p^2}{\beta_1^2}\right)^{\frac{1}{2}}, \quad s_2 = \left(\phi_2^2\xi^2 + \frac{p^2}{\beta_2^2}\right)^{\frac{1}{2}}, \tag{9}$$

where $\beta_1^2 = \frac{L_1}{\rho_1}$, $\beta_2^2 = \frac{\mu_2}{\rho_2}$, $\phi_1^2 = \frac{N_1}{L_1}$ and $\phi_2^2 = 1 - \frac{P}{2\mu_2}$.

It follows from the boundary conditions (3) and (4) that

$$A = C, \quad A \cosh s_1 h = B \sinh s_1 h. \tag{10}$$

<u>Case I</u>. Let

$$S(x, t) = \begin{cases} Q, & a \le x \le b + Vt, \\ 0, & \text{elsewhere,} \end{cases} \tag{11}$$

where Q is constant.

This definition of stress discontinuity shows that it is created in the region $x = a$ to $x = b$ and then expands with the uniform velocity V in the x direction. In particular, when $a = b = 0$, the discontinuity is created at the origin and expands with uniform velocity V in the x direction.

From the boundary condition (5) one gets, with the help of (11),

$$BL_1s_1 + C\mu_2s_2 = \frac{Q}{2\pi p}\left[\frac{e^{-i\xi a} - e^{-i\xi b}}{i\xi} + \frac{e^{-i\xi b}}{i\xi + \frac{p}{V}}\right]. \tag{12}$$

Solving for A, B and C from (10) and (12), we get the displacement function at the free surface at $z = -h$ in the form

$$\bar{v}_1(x, -h; p)$$
$$= \frac{Q}{2\pi p}\int_{-\infty}^{\infty}\frac{e^{i\xi x}}{(L_1s_1 \sinh s_1h + \mu_2s_2 \cosh s_1h)}\left[\frac{e^{-i\xi a} - e^{-i\xi b}}{i\xi} + \frac{e^{-i\xi b}}{i\xi + \frac{p}{V}}\right]d\xi$$
$$= \frac{Q}{\pi p}\int_{-\infty}^{\infty}\frac{e^{(i\xi x - hs_1)}}{(L_1s_1 + \mu_2s_2)}\left[\frac{e^{-i\xi a} - e^{-i\xi b}}{i\xi} + \frac{e^{-i\xi b}}{i\xi + \frac{p}{V}}\right](1 - Ke^{-2hs_1})^{-1}d\xi, \tag{13}$$

where

$$K = \frac{L_1s_1 - \mu_2s_2}{L_1s_1 + \mu_2s_2} < 1, \tag{14}$$

represents the reflection coefficient of SH-waves incident from the sandy medium at the interface between two half-spaces. The coefficients of different power of K in series of (13) are associated with the pulses undergoing repeated reflection in the upper layer. Using the inverse Laplace transform, we can rewrite (13) in a convenient form:

$$v_1(x, -h; p) = L^{-1}(I_1 + I_2 + I_3) = L^{-1}(I_1) + L^{-1}(I_2) + L^{-1}(I_3), \tag{15}$$

where I_1, I_2 and I_3 are defined in the Appendix. The inverse Laplace transforms of I_1, I_2 and I_3 are obtained by following [Garvin 1956]; details are in the Appendix.

Equation (15) gives the exact value of the surface displacement field $v_1(x, -h, t)$ surface.

Case II. Let

$$S(x, t) = Qh\delta(x - Vt), \tag{16}$$

where Q is a constant and $\delta(x - Vt)$ is Dirac's delta function of argument $(x - Vt)$. A term h is included on the right-hand side of (16) so as to give S as the dimension of a stress.

The boundary condition (5) gives

$$BL_1 s_1 + C\mu_2 s_2 = \frac{Qh}{2\pi V \left(i\xi + \frac{p}{V} \right)}. \tag{17}$$

Solving for A, B and C from (10) and (17) one gets

$$\bar{v}_1(x, -h; p) = \frac{Qh}{\pi V} \int_{-\infty}^{\infty} \frac{e^{(i\xi x - hs_1)}(1 + Ke^{-2hs_1} + K^2 e^{-4hs_1} + \cdots)}{\left(i\xi + \frac{p}{V} \right)(L_1 s_1 + \mu_2 s_2)} \, d\xi \tag{18}$$

Proceeding similarly as in Case I we obtain

$$v_1(x, -h, t) = \frac{2Q\beta_1 h}{\pi L_1} \sum_{n=1,3,5,\ldots} \int_0^t G_n[\zeta_n(\lambda)] \, d\lambda, \tag{19}$$

where

$$G_n[\zeta_n(t)] = \text{Re}\left[\left\{ 1 + \phi_1^2 \zeta_n^2(t) \right\}^{\frac{1}{2}} + \mu \left(\varepsilon^2 + \phi_2^2 \zeta_n^2(t) \right)^{\frac{1}{2}} \right]^{-1} \times \left[\frac{\beta_1}{V} + i\zeta_n(t) \right]^{-1}$$

$$\times K^{\frac{n-1}{2}}[\zeta_n(t)] \frac{d\zeta_n(t)}{dt} H\left[t - (x^2 + n^2 h^2)^{\frac{1}{2}} \beta_1^{-1} \right] \tag{20}$$

and

$$\zeta_n(t) = \frac{\beta_1}{x^2 + n^2 h^2 \phi_1^2} \left[itx + nh\{ t^2 - (x^2 + n^2 h^2 \phi_1^2) \beta_1^{-2} \}^{\frac{1}{2}} \right], \quad n = 1, 3, 5, \ldots. \tag{21}$$

If the stress discontinuity is taken as $H(x) - H(x - Vt)$ in place of $\delta(x - Vt)$ the corresponding expression on the right-hand side of (18) will differ only by a constant factor from I_3 (with $a = b = 0$).

4. Numerical results and discussion

For numerical results, we have taken data for anisotropic layer and initially stressed half-space from [Babuska and Cara 1991]:

$$N_1 = 175 \, \text{GPa}, \quad L_1 = 202 \, \text{GPa}, \quad \rho_1 = 4408 \, \text{kg/m}^3$$

$$\mu_2 = 91.6 \, \text{GPa}, \quad \rho_2 = 3582 \, \text{kg/m}^3.$$

The values of $K_1 v_1(x, -h, t)$ for $x = 7h$ and $x = 14h$ have been plotted against $\tau_1 = \tau - \tau_0$, where τ_0 denotes the time at which the disturbance arrives at the point of observation with $K_1 = \frac{\pi L_1}{2Q\beta_1 h}$, $\tau = \frac{t\beta_1}{h}$ is the time of the disturbance to arrive from source to initial point. The value of τ at $x = 7h$ is $(7^2 + n^2\phi_1^2)^{1/2}$, $n = 1, 3, 5, \ldots$ and $\tau_0 = 7.06$ at $n = 1$ and the value of τ at $x = 14h$ is $(14^2 + n^2\phi_1^2)^{1/2}$, $n = 1, 3, 5, \ldots$ and $\tau_0 = 14.03$ at $n = 1$.

When $x = 7h$, for six initial values, we have

$$K_1 v_1(x, -h, t) = \sum_{n=1,3,5,7,9,11} A^0(\theta_n) \cosh^{-1}\left(\tau / \sqrt{7^2 + n^2\phi_1^2}\right) H\left(\tau - \sqrt{7^2 + n^2\phi_1^2}\right),$$

where

$$A^0(\theta_n) = \text{Re} \frac{\left[(1 - \phi_1^2 \cos^2\theta_n)^{\frac{1}{2}} - \mu(\varepsilon^2 - \phi_2^2\cos^2\theta_n)^{\frac{1}{2}}\right]^{\frac{n-1}{2}} \sin\theta_n}{\left[(1 - \phi_1^2\cos^2\theta_n)^{\frac{1}{2}} + \mu(\varepsilon^2 - \phi_2^2\cos^2\theta_n)^{\frac{1}{2}}\right]^{\frac{n-3}{2}} \left(\frac{\beta_1}{V} - \cos\theta_n\right)};$$

represents the reflection coefficient of SH-type waves incident from the anisotropic medium to initially stressed isotropic half-space. When $x = 14h$, for six initial values, we have the same expression for $K_1 v_1(x, -h, t)$, with 7^2 replaced by 14^2.

Figures 2 and 3 show graphs of the variation of displacement with elapsed time τ_1 for different values of initial stress and for different values of initial time at which pulses arrive.

In Figure 2, top, the graph is plotted for the disturbance effect for $x = 7h$ and $x = 14h$ and fixed value of $P = 1 \, \text{GPa}$. From the figure it can be observed that all the curves start from the origin with sharp changes in their slope and after sometime the curves get smooth. Also, it reflects that the disturbance is more prominent for early arrival of pulses i.e. for $x = 7h$ has more jumping effect than for $x = 14h$. If the place of observation has more distance from the source the impact of pluses is less. In Figure 2, bottom, the graph is plotted for the disturbance effect for $x = 7h$ and $x = 14h$ and fixed value of $P = 10 \, \text{GPa}$. The nature of the curves remains the same, but the magnitude of disturbance increases to a large extent as initial stress increases.

In Figure 3, top, the graph is plotted for the disturbance for $x = 7h$ and different values of P (1 GPa, 10 GPa and 100 GPa). The nature of the curves is oscillating
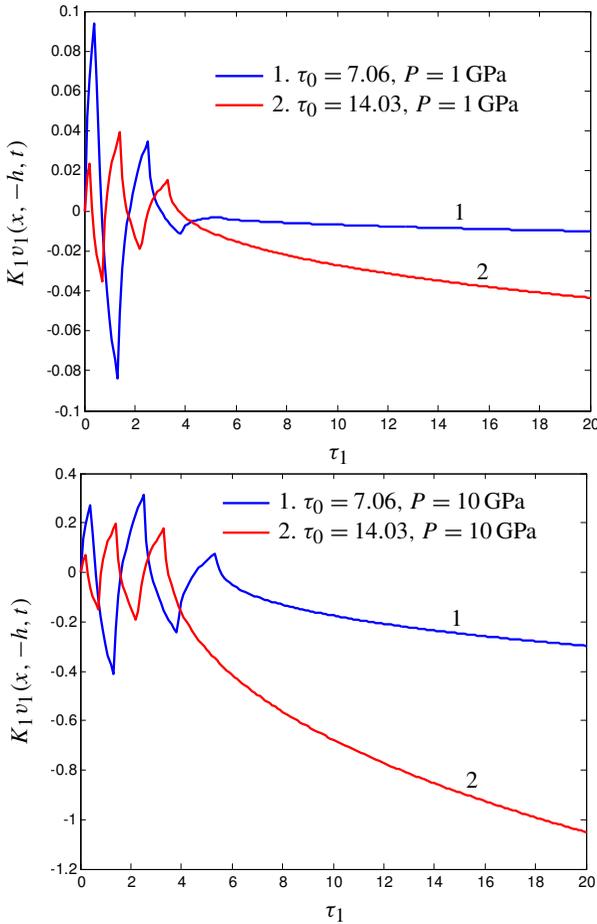
Figure 2. Variation of $K_1 v_1(x, -h, t)$ with τ_1 for $x = 7h$ and $x = 14h$, for $P = 1\,\mathrm{GPa}$ (top) and $P = 10\,\mathrm{GPa}$.

and after some time the curves become smooth and steady. As we increase the values of initial stress, the jumping effect increases. In Figure 3, bottom, the graph is plotted for the disturbance for $x = 14h$ and the same values of P. The nature of the curves remains the same, but the effect of disturbance is reduced to a large extent due to the late arrival of the pulses.

5. Conclusions

The generation of SH-type waves at the free surface of an anisotropic layer due to an impulsive stress discontinuity moving with uniform velocity along the interface of initially stressed isotropic medium has been considered. The displacement is calculated numerically for two particular distances on the surface for two different
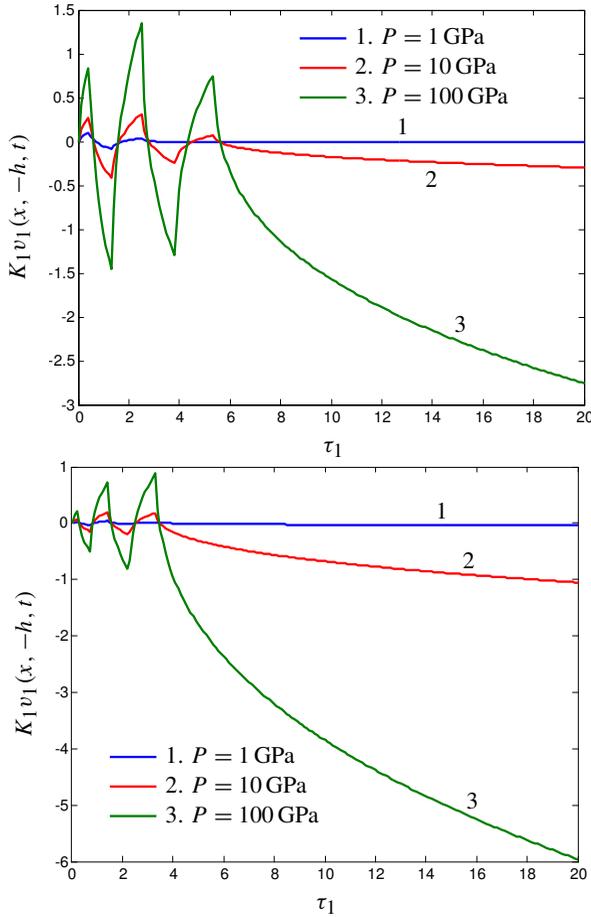
Figure 3. Variation of $K_1 v_1(x, -h, t)$ with τ_1 for $x = 7h$ (top) and $x = 14h$ (bottom) for different values of P.

types of the discontinuity in the shearing stress for different value of initial stress. It involves Laplace and Fourier transform and the inversion is based on Garvin's [1956] method. The numerical results are obtained for a particular model. From the figures it can be observed that initial stress and initial time at which pulses arrive has a significant effect. From the graph it is visible that that the displacement factor starts oscillating and after sometime it gets stable. The results are more comprising with the real scenario as we see that the disturbance arrives at the surface, it shakes the surface for a while and slowly gets stable. Also, from the figures it is visible that if the observer is nearer to the source then pulses arrive early and produce more disturbance and if the medium is highly pre-stressed then it produces more disturbance.

Appendix

We have

$$I_1 = \frac{Q}{\pi p} \int_{-\infty}^{\infty} \frac{(1 - Ke^{-2hs_1})^{-1}}{i\xi(L_1s_1 + \mu_2s_2)} e^{(i\xi x_1 - hs_1)} \, d\xi,$$

$$I_2 = \frac{Q}{\pi p} \int_{-\infty}^{\infty} \frac{(1 - Ke^{-2hs_1})^{-1}}{i\xi(L_1s_1 + \mu_2s_2)} e^{(i\xi x_2 - hs_1)} d\xi,$$

$$I_3 = \frac{Q}{\pi p} \int_{-\infty}^{\infty} \frac{(1 - Ke^{-2hs_1})^{-1}}{(i\xi + \frac{p}{V})(L_1s_1 + \mu_2s_2)} e^{(i\xi x_2 - hs_1)} \, d\xi,$$

with $x_1 = x - a$ and $x_2 = x - b$. In order to evaluate the Laplace inversion integral, we have used Garvin's method; see [Garvin 1956] for discussion of the contour integration and mapping.

Next for non-dimensionalisation, we substitute $\xi = \frac{\zeta p}{\beta_1}$, $\mu = \frac{\mu_2}{L_1}$, $\frac{\beta_1}{\beta_2} = \varepsilon$ in the integral above so that $s_1 = \frac{p}{\beta_1}(1 + \phi_1^2\zeta^2)^{\frac{1}{2}}$, $s_2 = \frac{p}{\beta_2}(\varepsilon^2 + \phi_2^2\zeta^2)^{\frac{1}{2}}$ and

$$K = \frac{(1 + \phi_1^2\zeta^2)^{\frac{1}{2}} - \mu(\varepsilon^2 + \phi_2^2\zeta^2)^{\frac{1}{2}}}{(1 + \phi_1^2\zeta^2)^{\frac{1}{2}} + \mu(\varepsilon^2 + \phi_2^2\zeta^2)^{\frac{1}{2}}}.$$

Thus we obtain

$$I_1 = \frac{2Q}{\pi p} \text{Im} \int_0^{\infty} \frac{e^{(i\xi x_1 - hs_1)}(1 + Ke^{-2hs_1} + K^2e^{-4hs_1} + \cdots)}{\xi(L_1s_1 + \mu_2s_2)} \, d\xi. \qquad (A.1)$$

The first term in I_1 is

$$I_{1,1} = \frac{2Q\beta_1}{\pi p L_1} \text{Im} \int_0^{\infty} \frac{\exp\left[-p\left\{-i\zeta x_1 + h\left(1 + \phi_1^2\zeta^2\right)^{\frac{1}{2}}\right\}/\beta_1\right]}{p\zeta\left[\left(1 + \phi_1^2\zeta^2\right)^{\frac{1}{2}} + \mu\left(\varepsilon^2 + \phi_2^2\zeta^2\right)^{\frac{1}{2}}\right]} \, d\zeta. \qquad (A.2)$$

The integrand (A.2) has singularities at $\zeta = 0, \pm\frac{i}{\phi_1}, \pm\frac{i\varepsilon}{\phi_2}$. Let $t = \left\{-i\zeta x_1 + h(1 + \phi_1^2\zeta^2)^{\frac{1}{2}}\right\}/\beta_1$. Then by inversion $\zeta(t) = \frac{\beta_1}{x_1^2 + \phi_1^2 h^2}\left[itx_1 + h\left\{t^2 - (x_1^2 + \phi_1^2 h^2)\beta_1^{-2}\right\}^{\frac{1}{2}}\right]$. The mapping of the ζ-plane into the t-plane is shown in Figure 4.

Making the reference to the Figure 4 and the paper of [Pal 1983], we find

$$L^{-1}I_{1,1} = \frac{2Q\beta_1}{\pi L_1} \int_0^t (t - \lambda_1)G_{1,1}[\zeta_{1,1}(\lambda_1)] \, d\lambda_1, \qquad (A.3)$$

where $L[tH(t)] = \frac{1}{p^2}$ and

$$G_{1,1}[\zeta_{1,1}(t)] = \text{Im}\left[\left(1 + \phi_1^2\zeta_{1,1}^2\right)^{\frac{1}{2}} + \mu\left(\varepsilon^2 + \phi_2^2\zeta_{1,1}^2\right)^{\frac{1}{2}}\right]^{-1} \zeta_{1,1}^{-1}(t)$$

$$\times \frac{d\zeta_{1,1}(t)}{dt} H\left[t - \left\{\eta\left(x_1^2 + h^2\phi_1^2\right)\right\}^{\frac{1}{2}} \beta_1^{-1}\right].$$
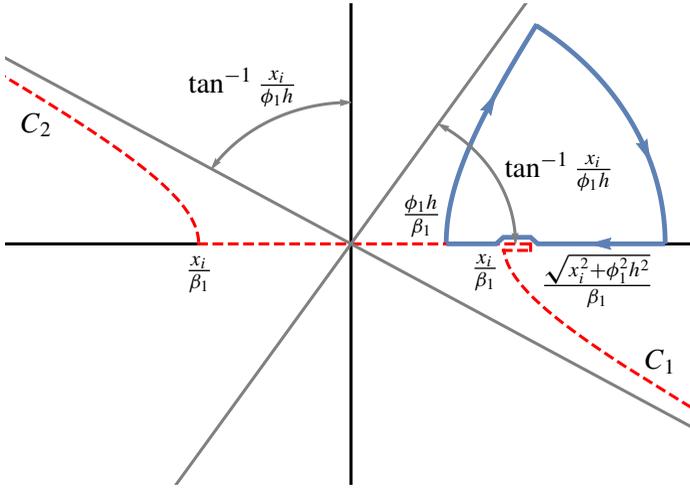
Figure 4. The t-plane showing the mapping and the contour of integration.

Since for $\frac{\phi_1 h}{\beta_1} < t < \frac{(x_1^2 + h^2 \phi_1^2)^{\frac{1}{2}}}{\beta_1}$, $\left[\left(1 + \phi_1^2 \zeta_{1,1}^2\right)^{\frac{1}{2}} + \mu\left(\varepsilon^2 + \phi_2^2 \zeta_{1,1}^2\right)^{\frac{1}{2}}\right]\zeta_{1,1}^{-1}(t) \times \frac{d\zeta_{1,1}(t)}{dt}$
is real. In general,

$$L^{-1}I_{1,n} = \frac{2Q\beta_1}{\pi L_1} \int_0^t (t - \lambda_1) G_{1,n}[\zeta_{1,n}(\lambda_1)] \, d\lambda_1, \tag{A.4}$$

where

$$G_{1,n}[\zeta_{1,n}(t)] = \text{Im}\left[\left\{1 + \phi_1^2 \zeta_{1,n}^2(t)\right\}^{\frac{1}{2}} + \mu\left(\varepsilon^2 + \phi_2^2 \zeta_{1,n}^2(t)\right)^{\frac{1}{2}}\right]^{-1} \zeta_{1,n}^{-1}(t)$$
$$\times K^{\frac{n-1}{2}}[\zeta_{1,n}(t)]\frac{d\zeta_{1,n}(t)}{dt} H\left[t - \left\{(x_1^2 + h^2\phi_1^2)\right\}^{\frac{1}{2}}\beta_1^{-1}\right],$$

and

$$\zeta_{1,n}(t) = \frac{\beta_1}{x_1^2 + n^2 h^2 \phi_1^2}\left[itx_1 + nh\left\{t^2 - (x_1^2 + n^2 h^2 \phi_1^2)\beta_1^{-2}\right\}^{\frac{1}{2}}\right], \quad n = 1, 3, 5, \ldots.$$

So that

$$L^{-1}I_1 = \sum_{n=1,3,5,\ldots} L^{-1}I_{1,n}. \tag{A.5}$$

Similarly

$$L^{-1}I_2 = \sum_{n=1,3,5,\ldots} L^{-1}I_{2,n}, \tag{A.6}$$

where x_1 is replaced by x_2.

Proceeding in the same way, we get

$$L^{-1}I_3 = \sum_{n=1,3,5,\dots} L^{-1}I_{3,n}, \tag{A.7}$$

where

$$L^{-1}I_{3,n} = \frac{2Q\beta_1}{\pi L_1} \int_0^t (t-\lambda_1) G_{3,n}[\zeta_{2,n}(\lambda_1)]\,d\lambda_1 \tag{A.8}$$

$$G_{3,n}[\zeta_{2,n}(t)] = \mathrm{Re}\left[\left\{ 1+\phi_1^2\zeta_{2,n}^2(t) \right\}^{\frac{1}{2}} + \mu\left(\varepsilon^2+\phi_2^2\zeta_{2,n}^2(t)\right)^{\frac{1}{2}} \right]^{-1} \times \left[\frac{\beta_1}{V} + i\zeta_{2,n}(t) \right]^{-1}$$
$$\times K^{\frac{n-1}{2}}[\zeta_{2,n}(t)] \frac{d\zeta_{2,n}(t)}{dt} H\left[t - \left(x_2^2+n^2h^2\phi_1^2\right)^{\frac{1}{2}} \beta_1^{-1} \right],$$

and $\zeta_{2,n}(t)$ is given by Appendix with x_2 in place of x_1.

Acknowledgement

References

[Abd-Alla and Ahmed 1999] A. M. Abd-Alla and S. M. Ahmed, "Propagation of Love waves in a non-homogeneous orthotropic elastic layer under initial stress overlying semi-infinite medium", *Appl. Math. Comput.* **106**:2 (1999), 265–275.

[Babuska and Cara 1991] V. Babuska and M. Cara, *Seismic anisotropy in the earth*, Kluwer, Dordrecht, 1991.

[Biot 1940] M. A. Biot, "The influence of initial stress on elastic wave", *J. Appl. Phys.* **8** (1940), 522–530.

[Biot 1965] M. A. Biot, *Mechanics of incremental deformation*, Wiley, New York, 1965.

[Cagniard 1939] L. Cagniard, *Reflection et refraction des ondes seismiques progressive*, Gauthier-Villars, Paris, 1939.

[Chattopadhyay and De 1981] A. Chattopadhyay and R. K. De, "Propagation of Love type waves in a visco elastic initially stressed layer overlying a visco-elastic half-space with irregular interface", *Rev. Roum. Sci. Tech. Mech. Appl.* **26**:3 (1981), 449–460.

[Chattopadhyay and Kar 1978] A. Chattopadhyay and B. K. Kar, "On the dispersion curves of Love types waves in an initially stressed crustal layer having an irregular interface", *Geophys. Res. Bull.* **16**:1 (1978), 13–23.

[Das and Dey 1968] S. C. Das and S. Dey, "Note on gravity waves", *Indian J. Engin. Math.* **1**:2 (1968), 155–160.

[Das and Dey 1970] S. C. Das and S. Dey, "Edge wave under initial stress", *Appl. Sci. Res.* **22**:1 (1970), 382–389.

[de Hoop 1960] A. T. de Hoop, "A modification of Cagniard's method for solving seismic pulse problem", *Appl. Sci. Res. B* **8**:1 (1960), 349–356.

[de Hoop 2002] A. T. de Hoop, "Reflection and transmission of a transient, elastic, line-source excited SH wave by a planer, elastic bonding surface in solid", *Int. J. Solids Struct.* **39**:21-22 (2002), 5379–5391.

[Dey 1971] S. Dey, "Wave propagation in two layered medium under initial stresses", *J. Pure Appl. Geophys. (Switzerland)* **90**:1 (1971), 38–52.

[Dey and Addy 1978] S. Dey and S. K. Addy, "Love waves under initial stresses in a visco-elastic medium overlying an elastic half-space", *Gerlands Beiträge Geophys.* **87**:4 (1978), 306–311.

[Garvin 1956] W. W. Garvin, "Exact transient solution of the Buried line source problem", *Proc. R. Soc. Lond. A* **234**:1199 (1956), 528–541.

[Khurana 2001] P. Khurana, "Love wave propagation in a prestressed medium", *Indian J. Pure Appl. Math.* **32**:8 (2001), 1201–1207.

[Majhi et al. 2016] S. Majhi, P. C. Pal, and S. Kumar, "Love waves in a layered functionally graded piezoelectric structure under initial stress", *Waves Random Complex Media* **26** (2016), 535–552.

[Majhi et al. 2017] S. Majhi, P. C. Pal, and S. Kumar, "Reflection and transmission of plane SH-waves in an initially stressed inhomogeneous anisotropic magnetoelastic medium", *J. Seismolog.* **21** (2017), 155–163.

[Mandal et al. 2014] P. C. Pal, D. Mandal, and S. Kumar, "Disturbance of SH-type waves due to moving stress discontinuity in an anisotropic soil layer overlying an inhomogeneous elastic half-space", *Sadhana* **39**:2 (2014), 451–465.

[Pal 1983] P. C. Pal, "Generation of SH-type waves due to non-uniformly moving stress-discontinuity in layered anisotropic elastic half-space", *Acta Mech.* **49**:3-4 (1983), 209–220.

[Pal and Kumar 2000] P. C. Pal and L. Kumar, "Generation of SH waves by a moving stress discontinuity in an anisotropic soil layer over an elastic half-space", *Acta Geophys.* **48**:4 (2000), 465–478.

[Pal and Mandal 2014] P. C. Pal and D. Mandal, "Generation of SH-type waves due to shearing stress discontinuity in a sandy layer overlying an isotropic and inhomogeneous elastic half-space", *Acta Geophys.* **62**:1 (2014), 44–58.

SANTOSH KUMAR: santosh453@gmail.com
Department of Mathematics, VVIT, Vidya Vihar Institute of Technology, Purnea 854303, India

DINBANDHU MANDAL: dinbandhumandal@gmail.com
Department of Mathematics, Chaibasa Engineering College, West Singhbhum 833215, India
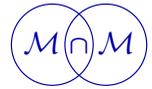
STRAIN GRADIENT AND GENERALIZED CONTINUA OBTAINED BY HOMOGENIZING FRAME LATTICES

HOUSSAM ABDOUL-ANZIZ AND PIERRE SEPPECHER

We determine the effective behavior of periodic structures made of welded elastic bars. Taking into account the fact that flexural and torsional stiffnesses are much smaller than the extensional one, we bypass classical homogenization formulas and obtain totally different types of effective energies. We work in the framework of linear elasticity. We give different examples of 2D or 3D microstructures which lead to generalized 1D, 2D, or 3D continua like the Timoshenko beam, Mindlin–Reissner plate, strain gradient, or Cosserat or micromorphic continua.

1. Introduction

Composite materials have proved to be so useful in structural design that homogenization techniques have received considerable attention over the past few decades both in mechanics and mathematics. More recently researchers realized that homogenization of composites made of very highly contrasted materials could lead to exotic effective behaviors. On the other hand the new manufacturing processes, which allow for extremely fine designs, gave birth to the new research field of metamaterials (or architected materials).

From the mathematical point of view, asymptotic homogenization of periodic media is now well founded. It consists of taking into account the fact that the size of the periodic cell is much smaller than the characteristic size of the considered

sample and of passing to the limit when the ratio ε between these two lengths tends to zero. This problem has been widely studied in static or dynamic cases, for conduction or elasticity problems, when the cell is made of a material with varying properties [Bensoussan et al. 1978; Sánchez-Palencia 1980; Allaire 1992; Oleĭnik et al. 1992]. Whenever the elastic energy of the material is a convex functional of the displacement field — and that is the case here where only linear elasticity is considered — a formula (see for instance [Allaire 1992, (3.6)] or [Milton 2002, p. 10]) gives the effective (i.e., limit) behavior of the medium in terms of a local minimization problem set in a rescaled cell. In this approach one lets ε tend to zero alone, while all other parameters of the system remain fixed. However, in many cases some other small parameters are present and the relative convergence speeds are crucial for the effective behavior of the material. The case when the cell is made of materials with very different properties is called "high contrast homogenization" [Cherednichenko et al. 2006]. Closely related is the case (which could be called the "infinitely high contrast" case) when holes are present (see for instance [Tartar 2009, Chapter 16]). But a small parameter can also derive from strong anisotropy or from geometric considerations. It is known that the effective behavior can then strongly differ [Pideri and Seppecher 1997; Bouchitté and Bellieud 2002; Bellieud and Gruais 2005; Bellieud et al. 2016; Bellieud 2017] from the initial behavior of the materials of which the structure is made. The first results in this direction were dealing with conduction problems, and a nonlocal limit energy was found [Tartar 1989; Khruslov 1991]. We are more interested by limit energies involving higher derivatives than the initial ones. Indeed materials with such energies are seldom found in nature [Barbagallo et al. 2017a; 2017b] and are expected to have a very special behavior [dell'Isola et al. 2015a; Eringen 2001; Mindlin 1965]. Their most distinctive feature is that they do not enter the framework of Cauchy stress theory (the internal mechanical interactions are not described by a Cauchy stress tensor) [dell'Isola et al. 2015b; 2016b; 2017]. However, such models are frequently used for regularizing the singularities which may arise in fracture, plasticity, interfaces, etc. (see for instance [Triantafyllidis and Aifantis 1986; Polizzotto and Borino 1998; Aifantis 1992; 1999; Mazière and Forest 2015; Sciarra et al. 2007; Yang and Misra 2012]).

Here we deal with static linear elasticity. In this framework a general closure result [Camar-Eddine and Seppecher 2003] states that all regular enough objective quadratic energies can be obtained through homogenization of highly contrasted media. In particular energies depending on the second gradient of the displacement (or equivalently on the strain gradient) or nonlocal energies [Bellieud and Gruais 2005] like energies associated to generalized continua [Forest 1999] can be obtained. But the result stated in [Camar-Eddine and Seppecher 2003] does not provide any reasonably applicable procedure for designing a microstructure with these exotic effective properties.

Note that we are interested here in the actual effective (i.e., limit) energy and not in corrections at order ε of a classical effective energy. The controversy about the sign of such corrective terms [Allaire et al. 2016; Le 2015] shows that they are difficult to interpret and to apply. Moreover, the fact that these corrective terms are present in conduction problems as well as in elasticity problems while it has been proved that no second gradient effect can appear in the limit energy for conductivity [Camar-Eddine and Seppecher 2002] shows that they are a very different notion.

A few structures have been described with a second gradient effective energy. Many of them [Pideri and Seppecher 1997; Bouchitté and Bellieud 2002; Bellieud et al. 2016; Bellieud 2017; Briane and Camar-Eddine 2007] lead to a couple-stress model, that is, to an energy depending only on the gradient of the skew-symmetric part of the gradient of the displacement [Toupin 1962; 1964; Mindlin 1963; Mindlin and Tiersten 1962; Bouyge et al. 2002] . Some discrete structures [Alibert et al. 2003; Alibert and Della Corte 2015; Boutin 1996; Seppecher et al. 2011; Boutin et al. 2017] lead to a more general second gradient energy.

In a recent paper [Abdoul-Anziz and Seppecher 2018] we have provided the first rigorous homogenization result in continuous elasticity which led to a general second gradient energy. We have considered periodic structures made by a single very stiff linear elastic material and void. The geometry of the structure consists of connected slender bars. They are so slender that the ratio of the section of these bars with respect to the size of the cell is comparable to ε. We have been able to prove that the 2D elasticity problem was, as expected, asymptotically similar to a frame lattice whose bars have a much smaller flexural stiffness than extensional stiffness. Then we have established a general formula for computing the effective energy of the medium. This result differs from the ones given in [Martinsson and Babuška 2007b; 2007a; Gonella and Ruzzene 2008] where very similar discrete systems are studied. The point is that, in these works, the orders of magnitude of the different types of interactions are assumed, as is generally done (see [Martinsson and Babuška 2007b, Remark 7.5], [Meunier et al. 2012, (2.7)], or [Braides and Gelli 2002]) not to interfere with the homogenization asymptotic process. In [Pastukhova 2005] or [Zhikov and Pastukhova 2003] the authors have assumed, like we do, that the ratio of the section of the bars with respect to the size of the cell is of order ε, but they considered like in [Zhikov 2002] a too soft material for obtaining generalized continuum limits.

In this paper, we start with a discrete lattice and extend the homogenization result of [Abdoul-Anziz and Seppecher 2018] to dimension 3. We also study concomitant homogenization and reduction of dimension in order to describe beam or plate models. We precisely describe the algebraic computation needed for making explicit the effective behavior of the considered lattices. Then we explore the wide variety of models which can be obtained. We feel that these examples provide academic

microscopic mechanisms enlightening the behavior of generalized materials. Our homogenization formula is a tool, which was up to now missing, for explaining how the strain or microdeformation can propagate in a strain gradient material or in a micromorphic material, respectively.

The paper is organized as follows. In Section 2 we fix the notation and the way of describing lattices. In Section 3 we recall the homogenization result and show that it can be recovered by using a formal expansion procedure. In Section 4 we present the algebraic computation needed for making explicit the effective energy in a sufficiently detailed way for enabling the reader to follow (and eventually check) the Octave/MATLAB package that we provide in [Abdoul-Anziz and Seppecher 2017]. Section 5 is devoted to the description of many examples, leading successively to beams, membranes, plates, and 3D materials. We recover classical models like Euler or Timoshenko beams, the Cosserat model for membranes, or the Kirchoff–Love or Mindlin–Reissner plate, but we also get strain gradient models with the possibility of mixing different effects.

2. Initial problem, description of the geometry, and notation

2.1. *The frame lattice.* In the physical space $\mathbb{R}^3$, we consider a periodic discrete lattice (see an example in Figure 1) defined by

- a bounded open domain $\underline{\Omega} \subset \mathbb{R}^3$,
- a small dimensionless parameter ε which we will let tend to zero (this parameter compares the size of the periodic cell εY of the lattice with the size of the macroscopic domain $\underline{\Omega}$),
- a prototype cell containing a finite number K of nodes, the positions of which are denoted $y_s \in \mathbb{R}^3$, $s \in \{1, \dots, K\}$,[1] and
- a family of N independent periodicity vectors t_α, $\alpha \in \{1, \dots, N\}$, with $1 \leq N \leq 3$.

The case $N = 3$ corresponds to standard 3D homogenization while the cases $N = 2$ and $N = 1$ correspond to 3D-2D and 3D-1D concomitant homogenization and reduction of dimension, respectively. They lead to plate or beam models.[2]

We assume with no loss of generality that the vector space $\mathbb{R}^N$ spanned by the vectors t_α coincides with the space spanned by the N first vectors of the canonical basis (e_1, e_2, e_3) of the physical space. The intersection of $\underline{\Omega}$ with $\mathbb{R}^N$ is denoted Ω, and we assume (a simple choice of the unit length) that its N-dimensional volume satisfies $|\Omega| = 1$.

---

[1]Note that lower-dimension cases $y_s \in \mathbb{R}$ or $y_s \in \mathbb{R}^2$ can be treated by simply embedding $\mathbb{R}$ or $\mathbb{R}^2$ in $\mathbb{R}^3$.

[2]Note that 2D-2D and 2D-1D are also treated by using the previous remark.
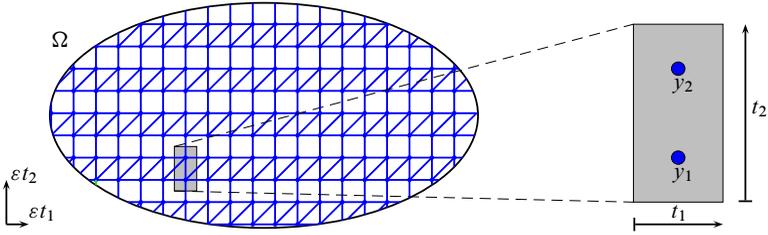
Figure 1. A prototype cell in a typical planar geometry when $K = 2$ and $N = 2$.

For $I = (i_1, \ldots, i_N) \in \mathbb{Z}^N$, consider the points $y_{I,s}^{\varepsilon} := \varepsilon(y_s + i_1 t_1 + \cdots + i_N t_N)$. We use $y_I^{\varepsilon} := (1/K) \sum_{s=1}^{K} y_{I,s}^{\varepsilon}$ as a reference point in the cell I. The nodes of the considered lattice are those nodes which lie sufficiently inside the domain $\underline{\Omega}$: more precisely the nodes $y_{I,s}^{\varepsilon}$ with $s \in \{1, \ldots, K\}$ and $I \in \mathcal{I}^{\varepsilon}$ where

$$\mathcal{I}^{\varepsilon} := \{I : y_I^{\varepsilon} \in \underline{\Omega}, \ d(y_I^{\varepsilon}, \partial\underline{\Omega}) > \sqrt{\varepsilon}\}. \tag{1}$$

The cardinal of this set, denoted N^{ε}, is of order ε^{-N}. In the sequel, for any field $\Phi_{I,s}$ defined at the nodes of the structure, we will denote by $\sum_I \Phi_{I,s}$ the mean values

$$\sum_I \Phi_{I,s} := \frac{1}{N^{\varepsilon}} \sum_{I \in \mathcal{I}^{\varepsilon}} \Phi_{I,s} \sim \varepsilon^N \sum_{I \in \mathcal{I}^{\varepsilon}} \Phi_{I,s}. \tag{2}$$

For any fixed cell I, the number of closest neighboring cells is $3^N - 1$. Counting the cell I itself, these cells are the cells $I \pm p$ with $p \in \mathcal{P}$ (the cardinal of $\mathcal{P}$ is $n = (3^N + 1)/2$). When $N = 1, 2$, or 3 we can choose

$$\mathcal{P} := \{0, 1\}, \tag{3}$$

$$\mathcal{P} := \{(0,0), (1,0), (0,1), (1,1), (1,-1)\}, \tag{4}$$

$$\mathcal{P} := \{(0,0,0), (1,0,0), (0,1,0), (1,1,0), (1,-1,0), (0,0,1), (0,1,1), (0,1,-1),$$
$$(1,0,1), (1,0,-1), (1,1,1), (1,1,-1), (1,-1,1), (1,-1,-1)\}, \tag{5}$$

respectively. In all cases, respecting the order given above, we identify $\mathcal{P}$ with $\{1, \ldots, n\}$. For any $p = (p_1, \ldots, p_N) \in \mathcal{P}$, we denote by $\boldsymbol{p} := p_1 t_1 + \cdots + p_N t_N$ the corresponding vector so that $y_{I+p,s}^{\varepsilon} = y_{I,s}^{\varepsilon} + \varepsilon \boldsymbol{p}$.

For any pair of distinct nodes $(y_{I,s}^{\varepsilon}, y_{I+p,s'}^{\varepsilon})$, we denote

$$\ell_{p,s,s'} := \varepsilon^{-1} \|y_{I+p,s'}^{\varepsilon} - y_{I,s}^{\varepsilon}\|, \quad \tau_{p,s,s'} := (y_{I+p,s'}^{\varepsilon} - y_{I,s}^{\varepsilon})/\|y_{I+p,s'}^{\varepsilon} - y_{I,s}^{\varepsilon}\|.$$

2.2. *Mechanical interactions.* To make precise the mechanical structure we are considering, we have to make precise the mechanical interactions between the nodes. The structures we want to model are periodic grids or frames made of

welded elastic bars. Essentially, the nodes behave like small rigid bodies and the interactions between these bodies can be divided in two parts. The extensional stiffness of one bar controls the relative displacements of its extremity nodes in the direction of the bar while the flexural and torsional stiffnesses control the relative rotations of its extremity nodes and the difference between these rotations and the global rotation of the bar.

Without loss of generality we assume that a cell is interacting only with its closest neighbors: indeed we can always choose a prototype cell large enough for this assumption to become true. Taking into account the symmetry, it is enough to fix the interactions between the nodes of cell I and between the nodes of cell I and half of its closest neighbors $I + p$ with $p \in \mathcal{P}$ (see Figure 2).
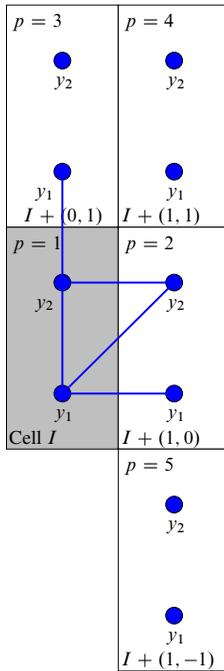
Figure 2. Fixing the interactions between a cell and its neighbors: here $K = 2$ and $N = 2$ and nonvanishing interactions are represented by blue lines, namely an internal ($p = 1$) interaction between nodes y_1 and y_2 ($a_{1,1,2} \neq 0$), three interactions with the cell on the right ($p = 2$) between node y_1 and nodes y_1 and y_2 and between nodes y_2 ($a_{2,1,1} \neq 0$, $a_{2,1,2} \neq 0$, and $a_{2,2,2} \neq 0$), and a last interaction with the cell above ($a_{3,2,1} \neq 0$). In this example there is no interaction with neighbor cells $p = 4$ nor $p = 5$, but such interactions could have been considered. Interactions with the cells below or on the left do exist. Owing to periodicity they do not need to be fixed.

• *Extensional interactions.* These interactions between the nodes of the lattice are determined by nK^2 nonnegative coefficients $a_{p,s,s'}$. We introduce the set of multi-indices corresponding to all pairs of nodes in interaction:

$$\mathcal{A} := \{(p, s, s') : p \in \mathcal{P}, \, 1 \le s \le K, \, 1 \le s' \le K, \, a_{p,s,s'} \ne 0\}.$$

For any displacement field U of the lattice, that is, a vector field $U_{I,s}$ defined on $\mathcal{I}^\varepsilon \times \{1, \ldots, K\}$, we call the "extension" between nodes (I, s) and $(I + p, s')$ the quantity

$$(\rho_U)_{I,p,s,s'} := \frac{U_{I+p,s'} - U_{I,s}}{\varepsilon} \cdot \tau_{p,s,s'}. \tag{6}$$

The extensional energy of the lattice has the form

$$E^\varepsilon(U) := \varepsilon^{-2} \sum_I \sum_{(p,s,s')} \frac{a_{p,s,s'}}{2} (\rho_U)^2_{I,p,s,s'}$$

$$= \varepsilon^{-2} \sum_I \sum_{(p,s,s') \in \mathcal{A}} \frac{a_{p,s,s'}}{2} (\rho_U)^2_{I,p,s,s'}. \tag{7}$$

• *Flexural/torsional interactions.* We attach to each node (I, s) of the structure a rigid motion: in addition to the displacement $U_{I,s}$, each node is endowed with a rotation vector[3] $\theta_{I,s}$. Let us introduce the vector

$$(\alpha_U)_{I,p,s,s'} := \tau_{p,s,s'} \times \frac{U_{I+p,s'} - U_{I,s}}{\varepsilon \ell_{p,s,s'}}.$$

As mechanical interactions need to be objective (i.e., invariant when adding both a constant value Φ to the field $\theta_{I,s}$ and the field $\Phi \times y^\varepsilon_{I,s}$ to the displacement field $U_{I,s}$), flexural/torsional interaction between nodes (I, s) and $(I + p, s')$ has to be a positive quadratic form of two vectors $\theta_{I,s} - (\alpha_U)_{I,p,s,s'}$ and $\theta_{I+p,s'} - (\alpha_U)_{I,p,s,s'}$. It can be represented by a nonnegative 6×6 matrix.[4]

Thus, flexural/torsional interactions are determined by $3nK^2$ nonnegative matrices $B_{p,s,s'}$, $C_{p,s,s'}$, $D_{p,s,s'}$ whose elements are 3×3 matrices, so that the flexural energy reads

$$F^\varepsilon(U, \theta) :=$$

$$\sum_I \sum_{(p,s,s')} \Big[(\theta_{I,s} - (\alpha_U)_{I,p,s,s'}) \cdot \frac{B_{p,s,s'}}{2} \cdot (\theta_{I,s} - (\alpha_U)_{I,p,s,s'})$$

$$+ (\theta_{I,s} - (\alpha_U)_{I,p,s,s'}) \cdot C_{p,s,s'} \cdot (\theta_{I+p,s'} - (\alpha_U)_{I,p,s,s'})$$

$$+ (\theta_{I+p,s'} - (\alpha_U)_{I,p,s,s'}) \cdot \frac{D_{p,s,s'}}{2} \cdot (\theta_{I+p,s'} - (\alpha_U)_{I,p,s,s'}) \Big]. \tag{8}$$

---

[3] Remember that we are in the framework of linear elasticity and that rotations are represented by skew-symmetric matrices which can be identified by vectors.

[4] Note that objectivity implies also that the rank of this matrix cannot exceed 5.

We assume that flexural/torsional interaction is present only in conjunction with extensional interaction:

$$a_{p,s,s'} > 0 \quad \Longleftrightarrow \quad \begin{pmatrix} B_{p,s,s'} & C_{p,s,s'} \\ C^t_{p,s,s'} & D_{p,s,s'} \end{pmatrix} > 0.$$

Our choice of the order of magnitude of these interactions needs some comment. We first emphasize that speaking of the order of magnitude of the stiffness of a structure makes sense only if we compare it to some force. In other words, making an assumption about the elastic rigidity is equivalent to making an assumption about the order of magnitude of the applied external forces.

Our aim is to consider structures for which classical homogenization would lead to a degenerated material. As usual a rescaling process is needed if one wants to capture a finite limit energy. Different assumptions can be made which correspond to different experiments. This is not surprising: the reader accustomed for instance to the 3D-2D or 3D-1D reduction of models for plates or beams knows that changing the assumptions about the order of magnitude of the elasticity stiffness of the material drastically changes the limit model. If the structure cannot resist some applied forces (like a membrane cannot resist transverse forces), it may resist them after a suitable scaling of the material properties (like the membrane model is replaced by the Kirchhoff–Love plate model). Simultaneously some mobility may disappear (like the Kirchhoff–Love plate becomes inextensible). Our choice of the order of magnitude of the extensional interactions means that the applied external forces are not sufficient to significantly extend the bonds between nodes. On the other hand, we have assumed that the flexural rigidities were much smaller than the extensional ones. This is unavoidable when considering structures in which mechanical interactions are due to slender parts. The chosen order of magnitude (ε^0) is critical. Other cases can be deduced from our results by letting in a further step (b_p, c_p, d_p) tend to zero or to infinity. The assumption that the ratio between bending and extension stiffnesses is comparable to the homogenization small parameter ε is essential: we emphasize that one cannot capture all interesting asymptotic effects by homogenizing the structure in a first step and letting the ratio bending stiffness/extension stiffness tend to zero in a second step; see [Martinsson and Babuška 2007b].

Example. Assume the lattice consists of slender cylinders joining the interacting nodes. Assume that all the cylinders have a circular basis of radius $r^\varepsilon = \beta \varepsilon^2$ and are made of a homogeneous isotropic elastic material with Young modulus $\mathcal{Y}$ and Poisson coefficient ν. Extension, bending, and torsion rigidities of an elastic cylindrical bar of radius r^ε are classical results of mechanics [Germain 1973]. Integrating along the bar, one can deduce the values of the interactions due to the elasticity of a bar of length ℓ^ε. We get

$$B_{p,s,s'} = D_{p,s,s'} = a_{p,s,s'}(f\,\mathrm{Id} + (t-f)\tau_{p,s,s'} \otimes \tau_{p,s,s'}),$$

$$2C_{p,s,s'} = a_{p,s,s'}(f\,\mathrm{Id} - (2t+f)\tau_{p,s,s'} \otimes \tau_{p,s,s'})$$

with

$$a_{p,s,s'} = \frac{\mathcal{Y}\pi\beta^2\varepsilon^7 N^\varepsilon}{\ell_{p,s,s'}}, \qquad f = \beta^2, \qquad t = \frac{\beta^2}{4(1+\nu)}.$$

This case satisfies our assumptions as soon as one assumes that the Young modulus of the material is of order ε^{N-7}. Note that f and t are of order 1 and thus $a_{p,s,s'}$, $B_{p,s,s'}$, and $C_{p,s,s'}$ have the same order. Indeed the difference of order between extensional and flexural/torsional stiffnesses, due to the slenderness $\beta\varepsilon$ of the bars, has already been taken into account by the factor ε^{-2} introduced in the definition of E^ε.

Example. The case of a 2D lattice can also be treated in our framework. It is enough to fix $y_{s,3} = 0$ for all $s \in \{1, \dots, K\}$ and $t_{\alpha,3} = 0$ for all $\alpha \leq N \leq 2$ and to focus only on planar displacements $U_{I,s,3} = 0$ and $\theta_{I,s,1} = \theta_{I,s,2} = 0$ at all nodes. Let us assume that the nodes are linked by slender rectangles of thickness $\beta\varepsilon^2$. Textbooks in mechanics give the extension and bending rigidities of a slender rectangle. We still can use the matrices $B_{p,s,s'}$, $C_{p,s,s'}$, and $D_{p,s,s'}$ defined in the previous example but modifying $a_{p,s,s'}$ and f in

$$a_{p,s,s'} = \frac{\mathcal{Y}\beta\varepsilon^3 N^\varepsilon}{\ell_{p,s,s'}}, \qquad f = \frac{4\beta^2}{3}.$$

Note that t plays no role in this example. This case satisfies again our assumptions as soon as one assumes that the Young modulus of the material is of order ε^{N-5}.

• *Boundary conditions.* We do not intend to study the way the different boundary conditions which could be imposed to our lattices pass to the limit. The richness [dell'Isola and Seppecher 1997; Seppecher et al. 2011] of the boundary conditions associated to generalized continua is such that trying to describe them in a general way is a real challenge. On the other hand, we cannot adopt the frequently used Dirichlet boundary conditions: indeed the lattices we consider generally present in the limit some inextensibility constraint and Dirichlet boundary conditions could lead to a trivial set of admissible deformations. So we consider here only free boundary conditions. So, in order to ensure uniqueness of the equilibrium solution, we impose a zero mean rigid motion:

$$\sum_I \frac{1}{K}\sum_{s=1}^K U_{I,s} = 0, \quad \sum_I \frac{1}{K}\sum_{s=1}^K \theta_{I,s} = 0. \tag{9}$$

• *Connectedness.* We are not interested in structures made of different unconnected

parts: we assume that mechanical interactions make a connected network. This has to be checked before using our results. This checking is generally obvious but is actually difficult to automate [Babuška and Sauter 2004].

3. Homogenization result

In a recent paper [Abdoul-Anziz and Seppecher 2018] we have rigorously derived the model (7)–(8) from a 2D linear elastic problem by analyzing the behavior of the slender substructures: we have then identified the effective energy through a Γ-convergence theorem (for a simple definition of this notion the reader can refer to [Braides 2002] or [Dal Maso 1993]) using tools of double scale convergence (see [Nguetseng 1989] or [Allaire 1992]). This has been done in the 2D-2D case. The extension of the proof to other dimensions does not need new arguments. We will provide here neither the proof which can be found in [Abdoul-Anziz and Seppecher 2018] nor the technical but straightforward extension to other dimensions. The goal of this paper is to explore the diversity of possible limit models. However, for the readers who do not desire to enter into the mathematical developments of [Abdoul-Anziz and Seppecher 2018], we show below that formal expansions of the kinematic variables actually give the right effective energy.

Assume that there exist smooth enough functions (u, v_s, w_s, θ_s) (for any $s \in \{1, \ldots, K\}$) such that

$$
\begin{aligned}
U_{I,s}^\varepsilon &:= u(y_I^\varepsilon) + \varepsilon v_s(y_I^\varepsilon) + \varepsilon^2 w_s(y_I^\varepsilon) + o(\varepsilon^2), \\
\theta_{I,s}^\varepsilon &:= \theta_s(y_I^\varepsilon) + o(1).
\end{aligned}
\tag{10}
$$

Then

$$
U_{I+p,s'}^\varepsilon - U_{I,s}^\varepsilon = \varepsilon \nabla u(y_I^\varepsilon) \cdot \boldsymbol{p} + \frac{\varepsilon^2}{2} \nabla\nabla u(y_I^\varepsilon) \cdot \boldsymbol{p} \cdot \boldsymbol{p} + \varepsilon(v_{s'}(y_I^\varepsilon) + \varepsilon \nabla v_{s'}(y_I^\varepsilon) \cdot \boldsymbol{p} - v_s(y_I^\varepsilon))
$$
$$
+ \varepsilon^2 (w_{s'}(y_I^\varepsilon) - w_s(y_I^\varepsilon)) + o(\varepsilon^2)
$$

and thus

$$
\begin{aligned}
\varepsilon^2 E^\varepsilon(U^\varepsilon) &= \sum_I \sum_{(p,s,s')\in\mathcal{A}} \frac{a_{p,s,s'}}{2} \left(\frac{U_{I+p,s'}^\varepsilon - U_{I,s}^\varepsilon}{\varepsilon} \cdot \tau_{p,s,s'} \right)^2 \\
&= \sum_I \sum_{(p,s,s')\in\mathcal{A}} \frac{a_{p,s,s'}}{2} ((\nabla u(y_I^\varepsilon) \cdot \boldsymbol{p} + v_{s'}(y_I^\varepsilon) - v_s(y_I^\varepsilon)) \cdot \tau_{p,s,s'})^2 + o(1) \\
&= \int_\Omega \sum_{(p,s,s')\in\mathcal{A}} \frac{a_{p,s,s'}}{2} ((\nabla u(x) \cdot \boldsymbol{p} + v_{s'}(x) - v_s(x)) \cdot \tau_{p,s,s'})^2 \, dx + o(1) \\
&= \bar{E}(v, \eta_u) + o(1),
\end{aligned}
$$

where we define for any functions v_s and $\eta_{p,s}$

$$\bar{E}(v,\eta) := \int_\Omega \sum_{p,s,s'} \frac{a_{p,s,s'}}{2}((\eta_{p,s'}(x) + v_{s'}(x) - v_s(x)) \cdot \tau_{p,s,s'})^2 \qquad (11)$$

and

$$\eta_u := \nabla u \cdot \boldsymbol{p}. \qquad (12)$$

As we are only interested by situations in which the energy $E^\varepsilon(U^\varepsilon) + F^\varepsilon(U^\varepsilon, \theta^\varepsilon)$ remains bounded, $\varepsilon^2 E^\varepsilon(U^\varepsilon)$ tends to zero and we get the constraint

$$\bar{E}(v, \eta_u) = 0. \qquad (13)$$

This implies that, for any $(p, s, s') \in \mathcal{A}$,

$$(v_{s'} - v_s + \nabla u \cdot \boldsymbol{p}) \cdot \tau_{p,s,s'} = 0,$$

from which we deduce that

$$(\nabla v_{s'} \cdot \boldsymbol{p} - \nabla v_s \cdot \boldsymbol{p} + \nabla\nabla u \cdot \boldsymbol{p} \cdot \boldsymbol{p}) \cdot \tau_{p,s,s'} = 0.$$

Using the two last equations, we get for any $(p, s, s') \in \mathcal{A}$

$$\varepsilon^{-2}(U^\varepsilon_{I+p,s'} - U^\varepsilon_{I,s}) \cdot \tau_{p,s,s'} = \left(\tfrac{1}{2}\nabla\nabla u(y^\varepsilon_I) \cdot \boldsymbol{p} \cdot \boldsymbol{p} + \nabla v_{s'}(y^\varepsilon_I) \cdot \boldsymbol{p} + (w_{s'}(y^\varepsilon_I) - w_s(y^\varepsilon_I))\right)$$
$$\cdot \tau_{p,s,s'} + o(1)$$

and thus

$$E^\varepsilon(U^\varepsilon) = \sum_I \sum_{(p,s,s')\in\mathcal{A}} \frac{a_{p,s,s'}}{2}\left(\frac{U^\varepsilon_{I+p,s'} - U^\varepsilon_{I,s}}{\varepsilon^2} \cdot \tau_{p,s,s'}\right)^2$$

$$= \int_\Omega \sum_{(p,s,s')\in\mathcal{A}} \frac{a_{p,s,s'}}{2}\left(\left(\tfrac{1}{2}\nabla\nabla u(x) \cdot \boldsymbol{p} \cdot \boldsymbol{p} + \nabla v_{s'}(x) \cdot \boldsymbol{p}\right.\right.$$
$$\left.\left. + (w_{s'}(x) - w_s(x))\right) \cdot \tau_{p,s,s'}\right)^2 dx + o(1)$$

$$= \bar{E}(w, \xi_{u,v}) + o(1)$$

where $\bar{E}$ is the functional defined in (11) and $\xi_{u,v}$ is the quantity

$$\xi_{u,v} := \tfrac{1}{2}\nabla\nabla u \cdot \boldsymbol{p} \cdot \boldsymbol{p} + \nabla v_{s'} \cdot \boldsymbol{p}. \qquad (14)$$

On the other hand

$$(\alpha_U)^\varepsilon_{I,p,s,s'} := \varepsilon^{-1}\tau_{p,s,s'} \times (U^\varepsilon_{I+p,s'} - U^\varepsilon_{I,s})$$
$$= \varepsilon^{-1}\tau_{p,s,s'} \times ((u(y^\varepsilon_{I+p}) - u(y^\varepsilon_I)) + \varepsilon(v_{s'}(y^\varepsilon_{I+p}) - v_s(y^\varepsilon_I)) + o(\varepsilon))$$
$$= \tau_{p,s,s'} \times (\nabla u(y^\varepsilon_I) \cdot \boldsymbol{p} + v_{s'}(y^\varepsilon_I) - v_s(y^\varepsilon_I)) + o(1).$$

Using also the fact that $\theta^\varepsilon_{I+p,s'} = \theta(y^\varepsilon_{I+p}) = \theta(y^\varepsilon_I) + o(1)$ we get

$$F^\varepsilon(U^\varepsilon, \theta^\varepsilon) :=$$

$$\sum_I \sum_{(p,s,s')} \left[(\theta^\varepsilon_{I,s} - (\alpha_U)^\varepsilon_{I,p,s,s'}) \cdot \frac{B_{p,s,s'}}{2} \cdot (\theta^\varepsilon_{I,s} - (\alpha_U)^\varepsilon_{I,p,s,s'}) \right.$$

$$+ (\theta^\varepsilon_{I,s} - (\alpha_U)^\varepsilon_{I,p,s,s'}) \cdot C_{p,s,s'} \cdot (\theta^\varepsilon_{I+p,s'} - (\alpha_U)^\varepsilon_{I,p,s,s'})$$

$$\left. + (\theta^\varepsilon_{I+p,s'} - (\alpha_U)^\varepsilon_{I,p,s,s'}) \cdot \frac{D_{p,s,s'}}{2} \cdot (\theta^\varepsilon_{I+p,s'} - (\alpha_U)^\varepsilon_{I,p,s,s'}) \right]$$

$$= \bar{F}(v, \eta_u, \theta) + o(1),$$

where we define

$$\bar{F}(v, \eta, \theta) := \int_\Omega \sum_{p,s,s'} \left[\left(\theta_s(x) - \frac{\tau_{p,s,s'}}{\ell_{p,s,s'}} \times (v_{s'}(x) - v_s(x) + \eta_{p,s'}(x)) \right) \right.$$

$$\cdot \frac{B_{p,s,s'}}{2} \cdot \left(\theta_s(x) - \frac{\tau_{p,s,s'}}{\ell_{p,s,s'}} \times (v_{s'}(x) - v_s(x) + \eta_{p,s'}(x)) \right)$$

$$+ \left(\theta_s(x) - \frac{\tau_{p,s,s'}}{\ell_{p,s,s'}} \times (v_{s'}(x) - v_s(x) + \eta_{p,s'}(x)) \right)$$

$$\cdot C_{p,s,s'} \cdot \left(\theta_{s'}(x) - \frac{\tau_{p,s,s'}}{\ell_{p,s,s'}} \times (v_{s'}(x) - v_s(x) + \eta_{p,s'}(x)) \right)$$

$$+ \left(\theta_{s'}(x) - \frac{\tau_{p,s,s'}}{\ell_{p,s,s'}} \times (v_{s'}(x) - v_s(x) + \eta_{p,s'}(x)) \right)$$

$$\left. \cdot \frac{D_{p,s,s'}}{2} \cdot \left(\theta_{s'}(x) - \frac{\tau_{p,s,s'}}{\ell_{p,s,s'}} \times (v_{s'}(x) - v_s(x) + \eta_{p,s'}(x)) \right) \right]. \quad (15)$$

To summarize, the effective energy is $\bar{E}(w, \xi_{u,v}) + \bar{F}(v, \eta_u, \theta)$ under the constraint $\bar{E}(v, \eta_u) = 0$. As w is an internal variable and as, in general, we also have no external action on θ, it is better to write the effective energy in terms of the macroscopic displacement u only:

Theorem 1. *The limit (effective) energy associated with the microscopic energy $E^\varepsilon + F^\varepsilon$ is*

$$\mathscr{E}(u) := \inf_{w,v,\theta} \{ \bar{E}(w, \xi_{u,v}) + \bar{F}(v, \eta_u, \theta) : \bar{E}(v, \eta_u) = 0 \}. \quad (16)$$

We remark that the constraint $\bar{E}(v, \eta_u) = 0$ may induce a constraint on the strain tensor $e(u)$ (i.e., the symmetric part of ∇u). Indeed we will see that the effective behavior of the considered structure is often subject to some constraints (like inextensibility in some direction, incompressibility, or even total rigidity).

We also expect that the effective energy corresponds to a strain gradient model. Indeed the second gradient of u enters the expression of $\bar{E}(w, \xi_{u,v})$ directly through the definition (14) of $\xi_{u,v}$. Moreover, the constraint $\bar{E}(v, \eta_u) = 0$ establishes a linear relation between v and ∇u; thus, the dependence of $\xi_{u,v}$ on the gradient of v can be a second source for strain gradient terms. However, it is not so simple to find structures for which such strain gradient effects arise and are not concealed by the constraint. In the next section we explain how to compute explicitly the limit energy, and we apply this procedure to many examples in Section 5.

Note that we prefer to describe the homogenized behavior of the considered structures in terms of the limit elastic energy only. Beyond the fact that it is very concise, it has the advantage of being written without considering any applied external forces. Indeed external forces have little to do with the constitutive law of the new material. Equilibrium equations under the action of a (reasonable) external force field f can then be obtained by simply writing the Euler equations of the minimization of the total energy $\mathscr{E}(u) - \int_\Omega f(x) \cdot u(x)\, dx$. Properties of Γ-convergence [Braides 2002; Dal Maso 1993] ensure that the equilibrium states of the considered structure converge towards this minimum.

4. Explicit computation of the homogenized stiffness matrices

Let us describe the algorithm which makes explicit the limit energy. We give here all the details needed for understanding the Octave/MATLAB software we provide in [Abdoul-Anziz and Seppecher 2017].

Note first that in the computations leading to Theorem 1, we have assumed $|Y| = 1$. As it is sometimes clearer to describe the structure using a prototype cell which does not satisfy $|Y| = 1$, all geometric quantities have to be rescaled (i.e., divided by $|Y|^{1/N}$) before using the following algorithm.

Both limit energies $\bar{E}(v, \eta_u)$ and $\bar{F}(v, \eta_u, \theta)$ are quadratic forms of their variables. A priori the variables v and θ are $K \times 3$ matrices $v_{s,i}$ and $\theta_{s,i}$ with $s \in \{1, \ldots, K\}$ and $i \in \{1, 2, 3\}$, while the variable η is an $n \times K \times 3$ tensor $\eta_{p,s,i}$ with $p \in \mathcal{P}$. From now on we identify them with $3K$ or $3nK$ vectors $v_{(s,i)}$, $\theta_{(s,i)}$, and $\eta_{(p,s,i)}$ without modifying the notation. In the same way $(\nabla u)_{i,\gamma}$ and $(\nabla\nabla u)_{i,\gamma,\gamma'}$ are identified with the $3N$ and $3N^2$ vectors $(\nabla u)_{(i,\gamma)}$ and $(\nabla\nabla u)_{(i,\gamma,\gamma')}$ without modifying the notation.

Step 1 (rewriting the energies (11) and (15) in canonical form).

$$\bar{E}(v, \eta) = \frac{1}{2} \int_\Omega v^t \cdot \boldsymbol{A} \cdot v + \eta^t \cdot \boldsymbol{B} \cdot \eta + 2v^t \cdot \boldsymbol{C} \cdot \eta,$$

$$\bar{F}(v, \eta, \theta) = \frac{1}{2} \int_\Omega v^t \cdot \boldsymbol{D} \cdot v + \eta^t \cdot \boldsymbol{E} \cdot \eta + \theta^t \cdot \boldsymbol{F} \cdot \theta + 2v^t \cdot \boldsymbol{G} \cdot \eta + 2\theta^t \cdot \boldsymbol{H} \cdot v + 2\theta^t \cdot \boldsymbol{J} \cdot \eta.$$

This is a simple assembly process. Indeed, denoting $\tilde{a}_{p,s,r,i,j} := a_{p,s,r}\tau_{p,s,r,i}\tau_{p,s,r,j}$,

$$
\begin{cases}
\boldsymbol{A}_{(s,i),(r,j)} := -\sum_{p\in\mathcal{P}}(\tilde{a}_{p,s,r,i,j} + \tilde{a}_{p,r,s,i,j}) & \text{if } r \neq s, \\
\boldsymbol{A}_{(s,i),(s,j)} := -\sum_{r\neq s}\boldsymbol{A}_{(s,i),(r,j)},
\end{cases}
$$

$$
\begin{cases}
\boldsymbol{B}_{(p,s,i),(q,r,j)} := 0 & \text{if } (q,r) \neq (p,s), \\
\boldsymbol{B}_{(p,s,i),(p,s,j)} := \sum_r \tilde{a}_{p,r,s,i,j}.
\end{cases}
$$

$$
\begin{cases}
\boldsymbol{C}_{(s,i),(p,r,j)} := -\tilde{a}_{p,s,r,i,j} & \text{if } r \neq s, \\
\boldsymbol{C}_{(s,i),(p,s,j)} := -\sum_{r\neq s}\boldsymbol{C}_{(s,i),(p,r,j)},
\end{cases}
$$

and using the Levi-Civita symbol ϵ and denoting

$$
\tilde{b}_{p,s,r,i,j} := \sum_{k,k',l,l'}\frac{1}{\ell_{p,s,r}^2}\epsilon_{i,k,l}\epsilon_{j,k',l'}(b+2c+d)_{p,s,r,k,k'}\tau_{p,s,r,l}\tau_{p,s,r,l'},
$$

we have

$$
\begin{cases}
\boldsymbol{D}_{(s,i),(r,j)} := -\sum_{p\in\mathcal{P}}(\tilde{b}_{p,s,r,i,j} + \tilde{b}_{p,r,s,i,j}) & \text{if } r \neq s, \\
\boldsymbol{D}_{(s,i),(s,j)} := -\sum_{r\neq s}\boldsymbol{D}_{(s,i),(r,j)},
\end{cases}
$$

$$
\begin{cases}
\boldsymbol{E}_{(p,s,i),(q,r,j)} := 0 & \text{if } (q,r) \neq (p,s), \\
\boldsymbol{E}_{(p,s,i),(p,s,j)} := \sum_r \tilde{b}_{p,r,s,i,j}.
\end{cases}
$$

$$
\begin{cases}
\boldsymbol{G}_{(s,i),(p,r,j)} := -\tilde{b}_{p,s,r,i,j} & \text{if } r \neq s, \\
\boldsymbol{G}_{(s,i),(p,s,j)} := -\sum_{r\neq s}\boldsymbol{G}_{(s,i),(p,r,j)},
\end{cases}
$$

$$
\begin{cases}
\boldsymbol{F}_{(s,i),(r,j)} := 2\sum_{p\in\mathcal{P}}c_{p,s,r,i,j} & \text{if } r \neq s, \\
\boldsymbol{F}_{(s,i),(s,j)} := \sum_{p\in\mathcal{P}}\big(2c_{p,s,s,i,j} + \sum_r(b_{p,s,r,i,j} + d_{p,r,s,i,j})\big),
\end{cases}
$$

$$
\begin{cases}
\boldsymbol{H}_{(s,i),(r,j)} := -\sum_{p\in\mathcal{P}}\sum_{k,l}(1/\ell_{p,s,r})\epsilon_{j,l,k}((b+c)_{p,s,r,l,i} + (d+c)_{p,r,s,l,i})\tau_{p,s,r,k} \\
\hspace{10cm} \text{if } r \neq s, \\
\boldsymbol{H}_{(s,i),(s,j)} := -\sum_{r\neq s}\boldsymbol{H}_{(s,i),(r,j)},
\end{cases}
$$

$$
\begin{cases}
\boldsymbol{J}_{(s,i),(p,r,j)} := \sum_{p\in\mathcal{P}}\sum_{k,l}(1/\ell_{p,s,r})\epsilon_{j,l,k}(b+c)_{p,s,r,l,i}\tau_{p,s,r,k} & \text{if } r \neq s, \\
\boldsymbol{J}_{(s,i),(p,s,j)} := \sum_{p\in\mathcal{P}}\sum_{k,l}(1/\ell_{p,s,s})\epsilon_{j,l,k}\big((b+c)_{p,s,s,l,i} \\
\hspace{6cm} + \sum_{r\neq s}(c+d)_{p,s,r,l,i}\big)\tau_{p,s,r,k}.
\end{cases}
$$

Step 2 (computing the constraint). Using the canonical form it is easy to compute the minimum of $\bar{E}(v,\eta_u)$ with respect to v. When the minimum is attained, v satisfies $\boldsymbol{A}\cdot v + \boldsymbol{C}\cdot\eta_u = 0$. The vector $\bar{v} := -\boldsymbol{A}^+\cdot\boldsymbol{C}\cdot\eta_u$ where $\boldsymbol{A}^+$ stands for the pseudoinverse of $\boldsymbol{A}$ is a possible solution[5] and the minimal value is $\int_\Omega \frac{1}{2}\eta_u^t\cdot\boldsymbol{X}\cdot\eta_u$ where

$$
\boldsymbol{X} := \boldsymbol{B} - \boldsymbol{C}^t\cdot\boldsymbol{A}^+\cdot\boldsymbol{C}. \tag{17}
$$

Note that the minimum with respect to w of $\bar{E}(w,\xi_{u,v})$ is computed in a similar way and becomes $\int_\Omega \frac{1}{2}(\xi_{u,v})^t\cdot\boldsymbol{X}\cdot\xi_{u,v}$.

---

[5]The properties of the Moore–Penrose pseudoinverse imply that this vector $\bar{v}$ belongs to the orthogonal to the kernel of A and so satisfies $\sum_k v_k = 0$.

Let us now introduce a linear operator L by setting, for any $p \in \mathcal{P}$, $s \in \{1, \ldots, K\}$, $1 \le i, j \le 3$, and $1 \le \gamma \le n$,

$$L_{(p,s,i),(j,\gamma)} := \delta_{i,j}(p \cdot t_\gamma), \tag{18}$$

where δ stands for the Kronecker delta so that

$$\eta_u = L \cdot \nabla u. \tag{19}$$

Setting $Q := L^t \cdot X \cdot L$ and $K := -A^+ \cdot C \cdot L$, the constraint $\bar{E}(v, \eta_u) = 0$ reads

$$Q \cdot \nabla u = 0, \tag{20}$$

$$v = K \cdot \nabla u + \tilde{v} \quad \text{with } \tilde{v} \in \text{Ker}(A). \tag{21}$$

Note that the matrix Q would have been the homogenized stiffness matrix of our structure if we had assumed a less stiff behavior of the interactions: we recover here results which have been obtained recently by [Martinsson and Babuška 2007b]. As we are, on the contrary, interested here in structures made by a very stiff material, we have to focus only on the kernel of Q. Objectivity implies that it contains at least the skew-symmetric matrices,[6] but in the most interesting cases it is much larger.

We introduce an orthonormal basis $(W^\xi)_{\xi=1}^d$ of $\text{Ker}(Q)$ ($N(5-N)/2 \le d \le 3N$). The matrix $P_{(i,\gamma),(j,\gamma')} := \sum_\xi W^\xi_{(i,\gamma)} W^\xi_{(j,\gamma')}$ represents the projection onto $\text{Ker}(Q)$, and constraint (20) reads

$$\nabla u = P \cdot \nabla u. \tag{22}$$

On the other hand, in order to represent $\tilde{v}$, we introduce a basis $(V^\xi)_{\xi=1}^{\tilde{d}}$ of $\text{Ker}(A)$ ($3 \le \tilde{d} \le 3K$): we set $\tilde{v} = \sum_{\xi=1}^{\tilde{d}} b_\xi(x) V^\xi$, that is, $\tilde{v} = V \cdot b$ with $V_{(s,i),\xi} := V^\xi_{(s,i)}$.

Remark. Surprisingly enough, the variable b which is introduced at this step may play an important role. It represents $\tilde{v}$ that is internal degrees of freedom (free only with respect to the highest-order energy). In classical homogenization $\tilde{v}$ is constant in the unit cell Y, and thus, its value has no effect on the homogenized energy. That is why one usually fixes, without loss of generality, the mean value of v to be zero. Here the situation is completely different: due to high contrast and for special geometries, the set in which $\tilde{v}$ lives may be much richer and the effect of this extra kinematic variable on the homogenized energy may be fundamental. This fact is illustrated by some examples in Section 5.

---

[6]By saying that a $3 \times N$ matrix M is skew-symmetric we mean that, for all $1 \le \gamma, \gamma' \le N$, $M_{\gamma,\gamma'} + M_{\gamma',\gamma} = 0$.

Step 3 (computing the extensional part of the energy). Owing to (20), we have $(\nabla v)_{s,i,\gamma} = \sum_{j,\gamma'} (K_{(s,i),(j,\gamma')} (\nabla\nabla u)_{j,\gamma',\gamma}) + (\nabla \tilde{v})_{s,i,\gamma}$. Using (22) and expressing $\tilde{v}$ in the base (V^ξ), we can rewrite (14) as

$$\xi_{u,v} = \boldsymbol{M} \cdot \nabla\nabla u + \boldsymbol{N} \cdot \nabla b$$

with

$$\boldsymbol{M}_{(p,s,i),(j,\gamma,\gamma')} = \sum_\zeta \sum_k (K_{(s,i),(k,\zeta)} + \tfrac{1}{2} \boldsymbol{p}_\zeta \delta_{i,k}) \boldsymbol{P}_{(k,\zeta),(j,\gamma)} \boldsymbol{p}_{\gamma'},$$

$$\boldsymbol{N}_{(p,s,i),(\xi,\gamma)} = V^\xi_{(s,i)} \boldsymbol{p}_\gamma.$$

The extensional energy $\inf_w \bar{E}(w, \xi_{u,v}) = \tfrac{1}{2} \int_\Omega (\xi_{u,v})^t \cdot \boldsymbol{X} \cdot \xi_{u,v}$ becomes

$$\frac{1}{2} \int_\Omega (\nabla\nabla u)^t \cdot \boldsymbol{M}^t \cdot \boldsymbol{X} \cdot \boldsymbol{M} \cdot \nabla\nabla u + (\nabla b)^t \cdot \boldsymbol{N}^t \cdot \boldsymbol{X} \cdot \boldsymbol{N} \cdot \nabla b + 2(\nabla\nabla u)^t \cdot \boldsymbol{M}^t \cdot \boldsymbol{X} \cdot \boldsymbol{N} \cdot \nabla b.$$

We prefer to rewrite it as the sum of two nonnegative terms:

$$\inf_w \bar{E}(w, \xi_{u,v}) = \frac{1}{2} \int_\Omega (\nabla\nabla u)^t \cdot \mathcal{R} \cdot \nabla\nabla u + (\nabla b + \mathcal{T} \cdot \nabla\nabla u)^t \cdot \mathcal{S} \cdot (\nabla b + \mathcal{T} \cdot \nabla\nabla u) \quad (23)$$

where

$$\mathcal{S} := \boldsymbol{N}^t \cdot \boldsymbol{X} \cdot \boldsymbol{N}, \quad \mathcal{T} := \mathcal{S}^+ \cdot \boldsymbol{N}^t \cdot \boldsymbol{X} \cdot \boldsymbol{M}, \quad \text{and} \quad \mathcal{R} := \boldsymbol{M}^t \cdot \boldsymbol{X} \cdot \boldsymbol{M} - \boldsymbol{M}^t \cdot \boldsymbol{X} \cdot \boldsymbol{N} \cdot \mathcal{T}.$$

Step 4 (computing the flexural part of the energy). We can also easily compute the minimum with respect to θ of $\bar{F}(v, \eta_u, \theta)$. When the minimum is attained, θ satisfies $\boldsymbol{F} \cdot \theta + \boldsymbol{H} \cdot v + \boldsymbol{J} \cdot \eta_u = 0$. The vector $\bar{\theta} := -\boldsymbol{F}^+ \cdot (\boldsymbol{H} \cdot v + \boldsymbol{J} \cdot \eta_u)$ is a possible solution, and the minimal value is

$$\inf_\theta \bar{F}(v, \eta_u, \theta) = \frac{1}{2} \int_\Omega v^t \cdot (\boldsymbol{D} - \boldsymbol{H}^t \cdot \boldsymbol{F}^+ \cdot \boldsymbol{H}) \cdot v + \eta_u^t \cdot (\boldsymbol{E} - \boldsymbol{J}^t \cdot \boldsymbol{F}^+ \cdot \boldsymbol{J}) \cdot \eta_u$$
$$+ 2v^t \cdot (\boldsymbol{G} - \boldsymbol{H}^t \cdot \boldsymbol{F}^+ \cdot \boldsymbol{J}) \cdot \eta_u.$$

Using (22), let us replace η_u by $\boldsymbol{L} \cdot \nabla u = \boldsymbol{L} \cdot \boldsymbol{P} \cdot \nabla u$ and v by $\boldsymbol{K} \cdot \boldsymbol{P} \cdot \nabla u + \boldsymbol{V} \cdot b$. We get

$$\inf_\theta \bar{F}(v, \eta_u, \theta) = \frac{1}{2} \int_\Omega b^t \cdot S \cdot b + (\nabla u)^t \cdot Z \cdot \nabla u + 2b^t \cdot Y \cdot \nabla u$$

with

$$S := \boldsymbol{V}^t \cdot (\boldsymbol{D} - \boldsymbol{H}^t \cdot \boldsymbol{F}^+ \cdot \boldsymbol{H}) \cdot \boldsymbol{V},$$
$$Z := \boldsymbol{P}^t \cdot \big(\boldsymbol{K}^t \cdot (\boldsymbol{D} - \boldsymbol{H}^t \cdot \boldsymbol{F}^+ \cdot \boldsymbol{H}) \cdot \boldsymbol{K} + \boldsymbol{L}^t \cdot (\boldsymbol{E} - \boldsymbol{J}^t \cdot \boldsymbol{F}^+ \cdot \boldsymbol{J}) \cdot \boldsymbol{L}$$
$$+ \boldsymbol{K}^t \cdot (\boldsymbol{G} - \boldsymbol{H}^t \cdot \boldsymbol{F}^+ \cdot \boldsymbol{J}) \cdot \boldsymbol{L} + \boldsymbol{L}^t \cdot (\boldsymbol{G}^t - \boldsymbol{J}^t \cdot \boldsymbol{F}^+ \cdot \boldsymbol{K}) \cdot \boldsymbol{L} \big) \cdot \boldsymbol{P},$$
$$Y := \boldsymbol{V}^t \cdot ((\boldsymbol{D} - \boldsymbol{H}^t \cdot \boldsymbol{F}^+ \cdot \boldsymbol{H}) \cdot \boldsymbol{K} + (\boldsymbol{G} - \boldsymbol{H}^t \cdot \boldsymbol{F}^+ \cdot \boldsymbol{J}) \cdot \boldsymbol{L}) \cdot \boldsymbol{P}.$$

Again we prefer to write this energy as the sum of two nonnegative terms:

$$\inf_{\theta} \overline{F}(v, \eta_u, \theta) = \frac{1}{2} \int_{\Omega} (\nabla u)^t \cdot \boldsymbol{R} \cdot \nabla u + (b + \boldsymbol{T} \cdot \nabla u)^t \cdot \boldsymbol{S} \cdot (b + \boldsymbol{T} \cdot \nabla u) \quad (24)$$

with

$$\boldsymbol{T} := \boldsymbol{S}^+ \cdot \boldsymbol{Y} \quad \text{and} \quad \boldsymbol{R} := \boldsymbol{Z} - \boldsymbol{Y}^t \cdot \boldsymbol{T}.$$

Collecting the results. The limit energy $\mathcal{E}$ obtained by collecting (23) and (24), namely

$$\mathcal{E} = \inf_{b} \frac{1}{2} \int_{\Omega} (\nabla\nabla u)^t \cdot \mathcal{R} \cdot \nabla\nabla u + (\nabla b + \mathcal{T} \cdot \nabla\nabla u)^t \cdot \mathcal{S} \cdot (\nabla b + \mathcal{T} \cdot \nabla\nabla u)$$
$$+ (\nabla u)^t \cdot \boldsymbol{R} \cdot \nabla u + (b + \boldsymbol{T} \cdot \nabla u)^t \cdot \boldsymbol{S} \cdot (b + \boldsymbol{T} \cdot \nabla u), \quad (25)$$

appears to be the integral of a quadratic form depending on the first and second gradients of the macroscopic displacement and on an extra kinematic variable b and its first gradient. The limit model is both a second gradient model (or strain gradient model) and a generalized continuum. From now on, we will call "microadjustment" the variable b.

In general we cannot go further because the microadjustment cannot be computed locally. The equilibrium equations are a coupled linear system of partial differential equations for u and b. This system is fixed as soon as the matrices $\boldsymbol{Q}$, $\mathcal{R}$, $\mathcal{S}$, $\mathcal{T}$, $\boldsymbol{R}$, $\boldsymbol{S}$, and $\boldsymbol{T}$ are fixed.

Step 5 (when possible, eliminating the extra kinematic variable). However, it is sometimes still possible to eliminate the microadjustment. That is the case when, for any field u, there exists a field $\bar{b}$ such that

$$\mathcal{S} \cdot (\nabla\bar{b} + \mathcal{T} \cdot \nabla\nabla u) = 0 \quad \text{and} \quad \boldsymbol{S} \cdot (\bar{b} + \boldsymbol{T} \cdot \nabla u) = 0. \quad (26)$$

Note that this operation would lead to serious difficulties if nonfree boundary conditions were considered. This field $\bar{b}$ clearly minimizes the energy and the homogenized energy reduces to

$$\mathscr{E}(u) = \frac{1}{2} \int_{\Omega} (\nabla\nabla u)^t \cdot \mathcal{R} \cdot \nabla\nabla u + (\nabla u)^t \cdot \boldsymbol{R} \cdot \nabla u \quad \text{under the constraint} \quad \boldsymbol{Q} \cdot \nabla u = 0. \quad (27)$$

Let us emphasize that the cases where the microadjustment b cannot be eliminated are also very relevant: we will see in the examples given in Section 5 that we recover many models of generalized continua which are widely used by mechanicians in practical situations.

Implementation. The algorithm we just described for determining the homogenized energy (25) or (27) is pure linear algebra dealing with very low-dimension matrices. It is very easy to implement in languages like Octave or MATLAB (an Octave/MATLAB package can be found in [Abdoul-Anziz and Seppecher 2017])

for getting numerical results or like Maxima for obtaining analytical results. As no optimization is needed, it can even be implemented in JavaScript (an online JavaScript tool is in development).

However, two points are not automated. Before using the algorithm, one has to manually check that the connectedness condition is satisfied. Then for using (27) one also has to check that equations (26) admit a solution; otherwise, one has to deal with the generalized continuum model given by (25).

In view of our results, the effective energy may correspond to a second gradient model (i.e., a strain gradient model) possibly coupled with an extra kinematic variable and subject to some first gradient constraints. However, few periodic structures exhibit such a behavior. Indeed most of them present a nondegenerate energy $\bar{E}$; in that case the strong constraint $\boldsymbol{Q} \cdot \nabla u = 0$ hides any second gradient effect. But even when $\boldsymbol{Q}$ is degenerate, it happens frequently that $\mathcal{R} = 0$: the model remains degenerate after rescaling. Different cases will be illustrated in the next section. We have no way, other than the algorithm we just described, for predicting whether a limit energy model will correspond to a classical model or a second gradient one or even a generalized continuum.

5. Examples

We apply the procedure described in the previous section to different 2D or 3D examples following the cases described in the introduction. To fix the ideas we always choose $a_{p,s,r} = 1$ whenever two nodes are interacting (i.e., when $a_{p,s,r} \neq 0$). Note that this assumption means that the sections of the bars differ when their lengths differ. We also always choose $f = 1$ (and, in the 3D case, $t = 0.25$). We classify our examples by the dimension $N = 1, 2, 3$ of the space in which the homogenized energy lives, leading thus to beams, membranes or plates, or 3D materials. In the cases $N = 1$ or $N = 2$ we successively consider 2D and 3D examples. We write the effective energy in terms of the components $e_{i,j}(u)$ of the strain tensor $e(u)$ and of the components $\partial^2 u_i / \partial x_j \, \partial x_k$ of the second gradient of the displacement. Translating the results in terms of the strain gradient is straightforward. Mind that the presence of constraints allows us to write different forms for the expression of the limit energy.

5.1. *Beams.* For sake of simplicity let us start by considering structures in $\mathbb{R}^2$ with one vector of periodicity ($N = 1$).

5.1.1. *2D Warren beam.* We consider the geometry (see Figure 3) $\Omega = (0, 1)$, $K = 2$, $y_1 = (0, 0)$, $y_2 = (0, 1)$, $t_1 = (1, 0)$, and $a_{1,1,2} = a_{2,1,1} = a_{2,2,2} = a_{2,1,2} = 1$; all other components $a_{p,s,s'}$ vanish. For this well known structure, the constraint (22) reads $e_{1,1}(u) = 0$: the beam is inextensible. A simple solution for condition (26)
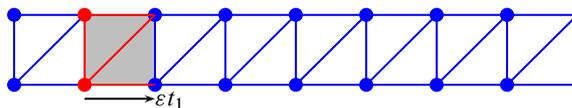
Figure 3. Warren beam.
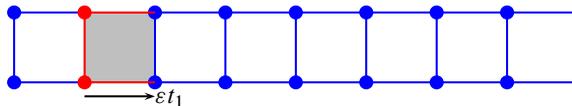
Figure 4. Square periodic beam.

is $b = 0$ and the limit energy reads

$$\mathscr{E}(u) = \frac{1}{2}\int_0^1 \lambda \left(\frac{\partial^2 u_2}{\partial x_1^2}\right)^2 dx_1$$

(with $\lambda = \frac{1}{2}$) and thus corresponds to an inextensible Euler–Bernoulli beam model.

Remark. The inextensibility constraint deserves a comment. This constraint is due to the order of magnitude we have assumed for the stiffness of the material the structure is made of. We already noticed that this assumption is actually an assumption for the order of magnitude of the forces which act on the structure. In other words we have implicitly assumed that the applied forces are not strong enough for extending the beam but that they are strong enough for bending it. Of course, different assumptions corresponding to different practical situations could be considered. The constraints which arise in all the following examples must be interpreted in this way.

5.1.2. *Square periodic beam.* The geometry is similar to the previous example (see Figure 4). We simply delete the diagonal bars by setting $a_{2,1,2} = 0$. The constraint (22) still reads $e_{1,1}(u) = 0$, but the homogenized energy now reads

$$\mathscr{E}(u) = \inf_b \frac{1}{2}\int_0^1 \left(\left(\frac{\partial b_2}{\partial x_1}\right)^2 + \left(\frac{\partial b_1}{\partial x_1}\right)^2 + 2\left(-b_1 + b_2 + \frac{\partial u_2}{\partial x_1}\right)^2\right) dx_1.$$

Denoting $\varphi := b_1 - b_2$ and minimizing with respect to $b_1 + b_2$ (by choosing $b_1 + b_2 = 0$), the limit energy becomes

$$\mathscr{E}(u) = \inf_\varphi \frac{1}{2}\int_0^1 \left(\lambda\left(\frac{\partial \varphi}{\partial x_1}\right)^2 + \zeta\left(\frac{\partial u_2}{\partial x_1} - \varphi\right)^2\right) dx_1$$

(with $\lambda = \frac{1}{2}$ and $\zeta = 2$) and thus corresponds to an inextensible Timoshenko beam model. It is well known that this model is nonlocal (in terms of u only) and that
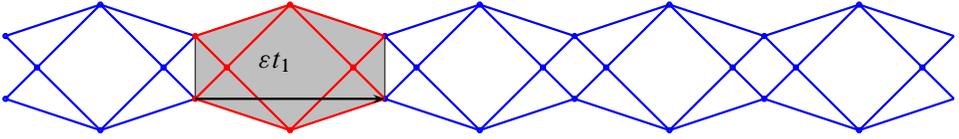
Figure 5. Pantographic beam P_{e_1,e_2}.

the extra kinematic variable φ cannot be locally eliminated. The remaining part of our "microadjustment" coincides with the Timoshenko extra variable usually interpreted as the "rotation of the section".

5.1.3. *Pantographic beam.* Structures based on a pantograph have been the first (and almost the only ones) to give a microscopic interpretation to the propagation of dilatation, a characteristic feature of complete second gradient models. The structures have been studied using formal homogenization techniques [Rahali et al. 2015; Seppecher et al. 2011; Madeo et al. 2017], Γ-convergence tools [Alibert et al. 2003; Alibert and Della Corte 2015], numerical simulations [Giorgio 2016; Turco et al. 2016a], and even experiments [dell'Isola et al. 2016a; Turco et al. 2016b]. Our procedure makes their study easy.

We consider a planar beam with a cell made of six nodes $y_1 = \frac{1}{6}e_2$, $y_2 = -\frac{1}{6}e_2$, $y_3 = \frac{1}{6}e_1$, $y_4 = \frac{1}{6}(3e_1 + 2e_2)$, $y_5 = \frac{1}{6}(3e_1 - 2e_2)$, and $y_6 = \frac{5}{6}e_1$; a periodicity vector $t_1 = e_1$; and $a_{1,1,3} = a_{1,1,4} = a_{1,2,3} = a_{1,2,5} = a_{1,3,4} = a_{1,3,5} = a_{1,4,6} = a_{1,5,6} = a_{2,4,1} = a_{2,5,2} = a_{2,6,1} = a_{2,6,2} = 1$; all other components of the matrices a_1 and a_2 vanish.

This beam (see Figure 5) which lies along the line $(0, e_1)$ and belongs to the plane (e_1, e_2) is denoted P_{e_1,e_2} for further purpose.

In that case (22) gives no constraint. A possible solution for condition (26) is $b = (4, 0, -1, -1, 1, 1)/4 \times \frac{\partial u_3}{\partial x_1}$. The limit energy reads

$$\mathscr{E}(u) = \frac{1}{2}\int_0^1 \left(\lambda \left(\frac{\partial^2 u_1}{\partial x_1^2}\right)^2 + \mu \left(\frac{\partial^2 u_2}{\partial x_1^2}\right)^2 + \zeta \left(\frac{\partial u_1}{\partial x_1}\right)^2 \right) dx_1$$

(with $\lambda = \frac{2}{23}$, $\mu = \frac{2}{63}$, and $\zeta = 324$). We recover the results obtained in [Seppecher et al. 2011; Alibert et al. 2003; Alibert and Della Corte 2015] where the exotic properties of this pantographic structure have been detailed. Its main specificity lies in the fact that a dilatation imposed in a part of the beam tends to spread on the whole beam. This phenomenon is due to the term $\left(\frac{\partial^2 u_1}{\partial x_1^2}\right)^2$ and damped by the term $\left(\frac{\partial u_1}{\partial x_1}\right)^2$. This competition endows the model with intrinsic length $\sqrt{\lambda/\zeta}$.

5.1.4. *3D Warren beam.* The previous examples deal with planar beams. In that case the energy is, of course, degenerate with respect to out-of-plane displacements. Let us give a single example of a Warren-type 3D beam leading to an
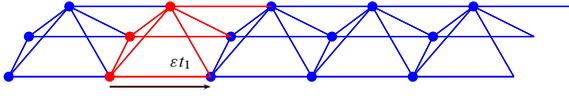
Figure 6. 3D Warren beam.

Euler–Bernoulli beam. The geometry (see Figure 6) is now $\Omega = (0, 1)$, $K = 3$, $y_1 = (0, 0, -\frac{1}{2})$, $y_2 = (0, 0, \frac{1}{2})$, $y_3 = (\frac{1}{2}, 0, \sqrt{2}/2)$, $t_1 = (1, 0)$, and $a_{1,1,2} = a_{1,1,3} = a_{1,2,3} = a_{2,1,1} = a_{2,2,2} = a_{2,3,3} = a_{2,3,1} = a_{2,3,2} = 1$; all other components of the matrices a_1 and a_2 vanish. Again the beam is not extensible ($e_{1,1} = 0$) and $b = 0$ is a possible solution for the microadjustment. The limit energy reads

$$\mathscr{E}(u) = \frac{1}{2} \int_0^1 \left(\lambda \left(\frac{\partial^2 u_2}{\partial x_1^2} \right)^2 + \zeta \left(\frac{\partial^2 u_3}{\partial x_1^2} \right)^2 \right) dx_1$$

(with $\lambda = \frac{1}{3}$ and $\zeta = \frac{1}{2}$) and corresponds to a nondegenerate Euler–Bernoulli beam. The bending stiffnesses in the two transverse directions are uncoupled. This is due to the symmetry of our structure.

5.2. *Membranes.*

5.2.1. *Regular triangle lattice.* The regular triangular truss (see Figure 7) is defined by a cell Y made of only one node ($K = 1$); two vectors $t_1 = (1, 0)$ and $t_2 = (-\frac{1}{2}, \sqrt{3}/2)$ for translating the cell; and five 1×1 matrices a_p, defining the interactions between the node of cell Y_I and the one of its neighbors Y_{I+p}, given by $a_1 = [0]$, $a_2 = [1]$, $a_3 = [1]$, $a_4 = [1]$, and $a_5 = [0]$.

We know that the constraint $\boldsymbol{Q} \cdot \nabla u = 0$ involves only the symmetric part of ∇u. In terms of $e(u)$ it reads

$$\frac{\sqrt{3}}{4} \begin{pmatrix} 3 & 1 & 0 \\ 1 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix} \cdot \begin{pmatrix} e_{1,1}(u) \\ e_{2,2}(u) \\ e_{1,2}(u) \end{pmatrix} = 0.$$

We already noticed that the matrix $\boldsymbol{Q}$ corresponds to the homogenized behavior which would have been obtained if assuming a smaller order of magnitude for the mechanical interactions. The result above is consistent with this remark and with
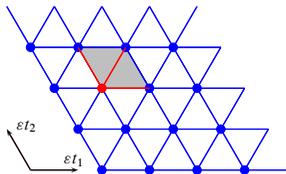
Figure 7. Regular triangle truss.
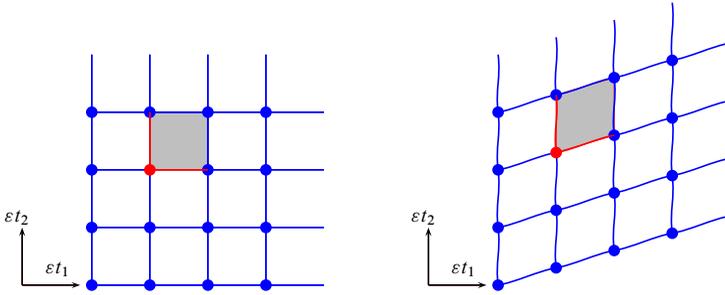
Figure 8. The regular square lattice and its admissible shear deformation.

the result given by [Martinsson and Babuška 2007b]. It corresponds, as expected, to a 2D isotropic material. Its Lamé coefficients are $\mu = \lambda = \sqrt{3}/4$, and its Poisson ratio is $\nu = \frac{1}{3}$. As the matrix above is nondegenerate, the constraint imposes the homogenized material to behave like a rigid body. As we have $\mathcal{E} = 0$ for rigid motions, there is no need for supplementary computations for the energy. We get the same uninteresting result for many structures (like, for instance, the Kagome (trihexagonal) lattice studied in [Leung and Guest 2007]). From now on, we will focus only on structures which have more degrees of mobility.

5.2.2. *Square grid.* The geometry of the regular square lattice (see Figure 8) is determined by a single node ($K = 1$), two vectors $t_1 = (1, 0)$ and $t_2 = (0, 1)$ for translating the cell, and five 1×1 matrices a_p defining the interactions between the node $y^\varepsilon_{I,1}$ and its neighbors $y^\varepsilon_{I+p,1}$ given by $a_1 = [0]$, $a_2 = [1]$, $a_3 = [1]$, $a_4 = [0]$, and $a_5 = [0]$.

Constraint (22) reads $e_{1,1}(u) = e_{2,2}(u) = 0$: the structure is inextensible in directions e_1 and e_2 and only shear is allowed. Microadjustment $b = 0$ is optimal, and the limit energy is

$$\mathcal{E}(u) = \frac{1}{2} \int_\Omega \lambda(e_{1,2}(u))^2 \, dx_1 \, dx_2$$

(with $\lambda = 6$). Contrarily to its 1D analog, this structure is a classical elastic material. It presents neither any second gradient effect nor generalized continuum effect.

5.2.3. *Square grid without constraints.* The reader may be frustrated by the fact that almost all our examples present a homogenized behavior subject to strong constraints. We show in this example that constraints can be avoided. Let us replace in the previous example the direct interactions by zigzags (see Figure 9): we consider a cell made of three nodes $y_1 = (0, 0)$, $y_2 = (0.5, 0.3)$, and $y_3 = (0.3, 0.5)$. All components of the five interaction 3×3 matrices a_p vanish but $a_{1,1,2} = a_{1,1,3} = a_{2,2,1} = a_{3,3,1} = 1$.
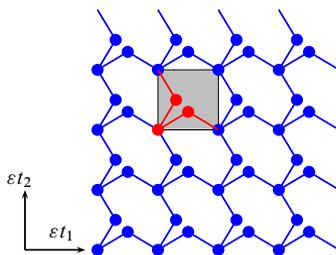
Figure 9. A square structure with unconstrained limit energy.
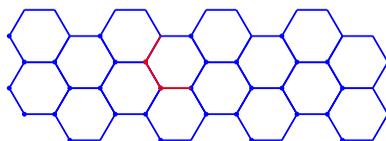
Figure 10. The honeycomb structure.

Constraint (22) disappears, and the limit energy reads

$$\mathcal{E}(u) = \frac{1}{2} \int_{\Omega} (\lambda(e_{1,1}(u))^2 + \lambda(e_{2,2}(u))^2 + \zeta(e_{1,2}(u))^2) \, dx_1 \, dx_2$$

(with $\lambda = \frac{50}{3}$ and $\zeta = 3$).

In the sequel, for sake of simplicity, we will not try to avoid all constraints: we let the reader check whether a suitable modification of the proposed structures could provide an unconstrained limit energy.

5.2.4. *Honeycomb structure.* The honeycomb structure (see Figure 10) is frequently put forward for its mechanical properties. It is defined by a cell Y made of two nodes ($K = 2$), two vectors $t_1 = (\frac{3}{2}, -\sqrt{3}/2)$ and $t_2 = (0, \sqrt{3})$ for translating the cell, and five 2×2 matrices a_p defining the interactions between the nodes of cell Y_I and the ones of its neighbors Y_{I+p}. All their components vanish but $a_{1,1,2} = a_{2,1,2} = a_{3,2,1} = 1$. Constraint (22) reads $e_{1,1}(u) + e_{2,2}(u) = 0$: the structure is incompressible. The microadjustment can be eliminated, and the limit energy is

$$\mathcal{E}(u) = \frac{1}{2} \int_{\Omega} \lambda \|e(u)\|^2 \, dx_1 \, dx_2$$

(with $\lambda = 9$). Contrarily to what was expected, this structure is a classical 2D elastic material which presents neither any second gradient effect nor generalized continuum effect. Incompressibility is its only specificity. This geometry has been studied in [Gibson and Ashy 1997; Davini and Ongaro 2011; Davini 2013; Dos Reis and Ganghoffer 2010]. Our result is in concordance with these results but differs due to different assumptions: in these works bending and extensional
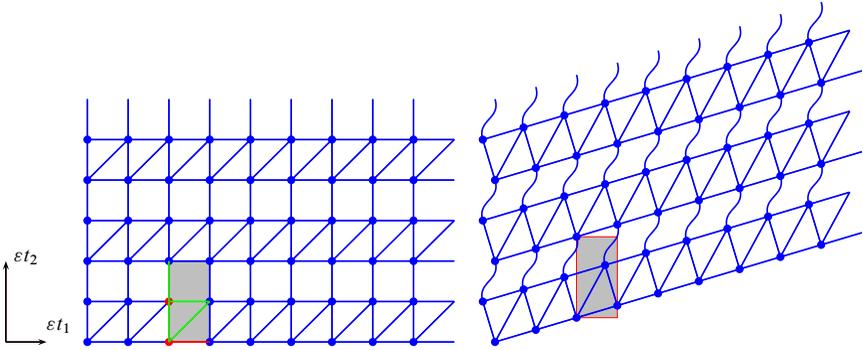
Figure 11. A layered structure.

stiffnesses have the same order of magnitude. It differs also from [Raoult et al. 2008; Le Dret and Raoult 2013] where nonlinearity has been taken into account but where bending stiffness has been chosen either weaker or stronger than we have.

5.2.5. *A couple-stress membrane.* We add a diagonal bar in one square cell over two in the square lattice described in Section 5.2.2 (see Figure 11). The lattice is now defined by a cell Y made of two nodes ($K = 2$) at points $y_1 = (0, 0)$ and $y_2 = (0, 1)$, the periodicity vectors $t_1 = (1, 0)$ and $t_2 = (0, 2)$, and five 2×2 matrices a_p. All components of these matrices vanish but $a_{1,1,2} = a_{2,1,1} = a_{2,2,2} = a_{2,1,2} = a_{3,2,1} = 1$.

This structure, when homogenized, is again subject to the constraint $e_{1,1}(u) = e_{2,2}(u) = 0$. An optimal microadjustment can be found, and the limit energy reads

$$\mathscr{E}(u) = \frac{1}{2} \int_\Omega \left(\lambda \left(\frac{\partial^2 u_2}{\partial x_1^2} \right)^2 + \zeta \left(e_{1,2}(u) \right)^2 \right) dx_1 \, dx_2$$

(with $\lambda = \frac{1}{8}$ and $\zeta = \frac{192}{5}$). From the mechanical point of view, the horizontal substructures behave like bending beams and their resistance to bending is responsible for the second gradient part of the limit energy. The model enters the framework of couple-stress models. Indeed, owing to the constraint, the energy can be rewritten

$$\mathscr{E}(u) = \frac{1}{2} \int_\Omega \left(\lambda \left(\frac{\partial}{\partial x_1} \left(\frac{\partial u_2}{\partial x_1} - \frac{\partial u_1}{\partial x_2} \right) \right)^2 + \zeta (e_{1,2}(u))^2 \right) dx_1 \, dx_2$$

and depends only on the gradient of the skew-symmetric part of the gradient of u.

As the energy can alternatively be written $\mathscr{E}(u) = \frac{1}{2} \int_\Omega (4\lambda (\partial e_{1,2}(u)/\partial x_1)^2 + \zeta (e_{1,2}(u))^2) \, dx_1 \, dx_2$, the model is clearly endowed with the internal length $\sqrt{4\lambda/\zeta}$.

5.2.6. *Pantographic membrane.* This structure is made by a connected array of pantographic structures quite similar to those studied in Section 5.1.3 (see Figures 12 and 13). It is defined by a cell Y made of six nodes ($K = 6$) at points $y_1 = (0, 1)$,
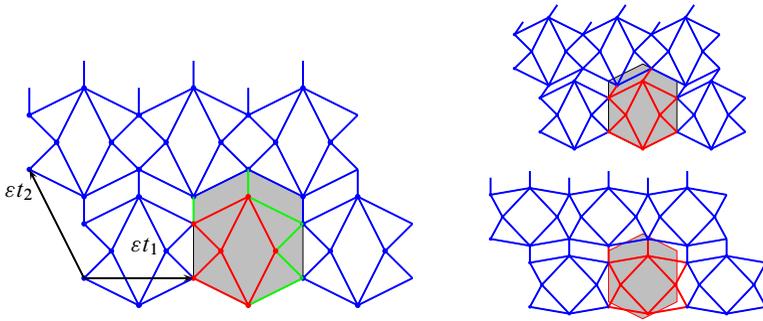
Figure 12. Pantographic membrane and its two admissible deformations (bending of the bars are not represented in the deformed configurations).
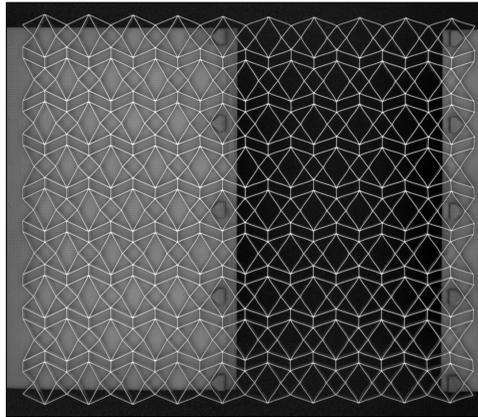
Figure 13. Experiment: traction of a pantographic membrane.

$y_2 = (0, -1)$, $y_3 = (1, 0)$, $y_4 = (2, 2)$, $y_5 = (2, -2)$ and $y_6 = (3, 0)$; two vectors $t_1 = (4, 0)$ and $t_2 = (-2, 4)$ for translating the cell; and five 6×6 matrices a_p defining the mechanical interactions. All components of these matrices vanish but

$$a_{1,1,3} = a_{1,1,4} = a_{1,2,3} = a_{1,2,5} = a_{1,3,4} = a_{1,3,5} = a_{1,4,6} = a_{1,5,6} = 1,$$
$$a_{2,4,1} = a_{2,5,2} = a_{2,6,1} = a_{2,6,2} = a_{3,1,5} = a_{4,4,2} = 1.$$

Constraint (22) reads $e_{2,2}(u) = 0$. Both horizontal dilatation and shear are admissible. Indeed, these macroscopic displacements, as shown in Figure 12, can be performed without extending any bar. A microadjustment satisfying (26) can be
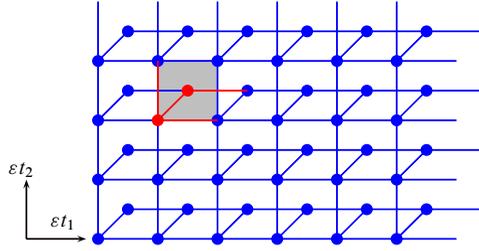
Figure 14. Planar structure leading to Cosserat model.

found, and the homogenized energy is

$$\mathcal{E}(u) = \frac{1}{2} \int_\Omega \left(\zeta (e_{1,1}(u))^2 + \sigma (e_{1,2}(u))^2 + \mu \left(\left(\frac{\partial^2 u_1}{\partial x_1^2} \right)^2 + \left(\frac{\partial^2 u_2}{\partial x_1^2} \right)^2 \right. \right.$$
$$\left. \left. + \lambda \left(\frac{\partial^2 u_1}{\partial x_1 \partial x_2} + \kappa \frac{\partial^2 u_2}{\partial x_1^2} \right)^2 \right) \right) dx_1 \, dx_2$$

(with $\lambda = \frac{484}{131}$, $\kappa = \frac{13}{44}$, $\mu = \frac{3}{44}$, $\sigma = 288$, and $\zeta = 144$). This model which has been studied in [Abdoul-Anziz and Seppecher 2018] is the prototype of complete second gradient models (indeed it does not enter the framework of couple-stress models because of the term $\left(\partial^2 u_1 / \partial x_1^2 \right)^2$). The very special behavior of this model has been described in [Seppecher et al. 2011]. Due to the strong anisotropy of the structure, it is difficult to distinguish the several intrinsic lengths contained in the model. Structures based on pantographic mechanisms have been intensively studied from theoretical [Placidi et al. 2016] but also numerical [Turco et al. 2017; Harrison 2016] and experimental [Placidi et al. 2017] points of view.

5.2.7. *A Cosserat model.* We consider the lattice described in Figure 14. It is a planar structure in which we have authorized crossing interactions. The periodic cell is made of two nodes at points $y_1 = (0, 0)$ and $y_2 = (0.5, 0.5)$; the periodicity vectors are $t_1 = (1, 0)$ and $t_2 = (0, 1)$. All components of the five 2×2 matrices a_p vanish but $a_{1,1,2} = a_{2,1,1} = a_{2,2,2} = a_{3,1,1} = 1$.

The constraint (22) is again $e_{1,1}(u) = e_{2,2}(u) = 0$, and only shear is admissible. The limit energy takes the form

$$\mathcal{E}(u) = \inf_\varphi \frac{1}{2} \int_\Omega \left[\zeta \left(\frac{\partial \varphi}{\partial x_1} \right)^2 + \gamma \left(\varphi - \frac{1}{2} \left(\frac{\partial u_2}{\partial x_1} - \frac{\partial u_1}{\partial x_2} \right) \right)^2 + \kappa (e_{1,2}(u))^2 \right] dx_1 \, dx_2$$

(with $\zeta = \frac{800}{729}$, $\gamma = \frac{1600}{333}$, and $\kappa = \frac{56}{9}$). The extra variable φ plays the role of a Cosserat variable. The reader can understand by considering Figure 14 that the rotation of the bars $[y_{I,1}^\varepsilon, y_{I,2}^\varepsilon]$ tends to be uniform and that it is coupled to the global displacement owing to the welding of the bars at each node.
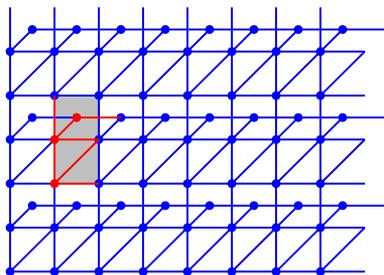
Figure 15. Planar structure leading to both second gradient and Cosserat effects.

5.2.8. *Second gradient and Cosserat effects together.* Let us combine in Figure 15 the structures of Figures 11 and 14: the periodic cell is now made of three nodes at points $y_1 = (0,0)$, $y_2 = (0,1)$, and $y_3 = (0.5, 1.5)$; the periodicity vectors $t_1 = (1,0)$ and $t_2 = (0,2)$; and five 3×3 matrices a_p. All components of these matrices vanish but $a_{1,1,2} = a_{1,2,3} = a_{2,1,1} = a_{2,1,2} = a_{2,2,2} = a_{2,3,3} = a_{3,2,1} = 1$.

The constraint (22) is still $e_{1,1}(u) = e_{2,2}(u) = 0$. Shear is admissible. The limit energy takes now the form

$$\mathcal{E}(u) = \inf_{\varphi} \frac{1}{2} \int_{\Omega} \left[\lambda \left(\frac{\partial^2 u_2}{\partial x_1^2} \right)^2 + \zeta \left(\frac{\partial \varphi}{\partial x_1} \right)^2 \right. $$
$$\left. + \gamma \left(\varphi - \frac{1}{2} \left(\frac{\partial u_2}{\partial x_1} - \frac{\partial u_1}{\partial x_2} \right) \right)^2 + \kappa (e_{1,2}(u))^2 \right] dx_1 \, dx_2,$$

thus mixing second gradient and Cosserat effects.

5.3. *Plates.* Up to now we have only considered planar structures which, of course, are completely degenerate with respect to transverse displacement. Let us now consider structures with a nonzero thickness.

5.3.1. *Kirchhoff–Love plate.* The considered lattice is made by two superposed regular triangular lattices (see Figure 16). It is defined by a cell Y made of two nodes ($K = 2$) at points $y_1 = (0,0,0)$ and $y_2 = (0,0,1)$, the periodicity vectors $t_1 = (1,0,0)$ and $t_2 = (-\frac{1}{2}, \sqrt{3}/2, 0)$, and five 2×2 matrices a_p. All components of these matrices vanish but $a_{1,1,2} = a_{2,1,1} = a_{2,2,2} = a_{2,1,2} = a_{2,2,1} = a_{3,1,1} = a_{3,2,2} = a_{3,1,2} = a_{3,2,1} = a_{4,1,1} = a_{4,2,2} = a_{4,1,2} = a_{4,2,1} = 1$.

The homogenized model is subject to the constraints $e_{11}(u) = e_{22}(u) = e_{12}(u) = 0$ (as a membrane, it is undeformable). The microadjustment $b = 0$ is optimal, and the limit energy reads

$$\mathcal{E}(u) = \frac{1}{2} \int_{\Omega} (\lambda \|\nabla \nabla u_3\|^2 + \zeta (\Delta u_3)^2) \, dx_1 \, dx_2$$
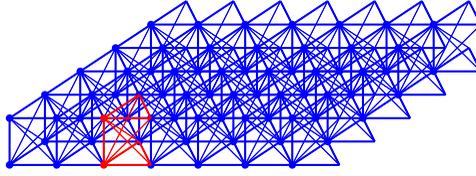
Figure 16. A Kirchoff–Love plate.
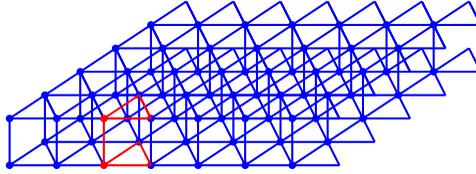
Figure 17. A Mindlin–Reissner plate.

(with $\lambda = \frac{1}{2}$ and $\zeta = \frac{1}{4}$) which corresponds to an isotropic inextensible Kirchhoff–Love bending plate.

5.3.2. *Mindlin–Reissner plate.* Let us consider the same geometry as in the previous example but where all diagonals joining the lower nodes to the upper ones are deleted (see Figure 17): it is enough to set $a_{2,1,2} = a_{2,2,1} = a_{3,1,2} = a_{3,2,1} = a_{4,1,2} = a_{4,2,1} = 0$.

The inextensibility constraint $e_{11}(u) = e_{22}(u) = e_{12}(u) = 0$ remains, but now the microadjustment b cannot be completely eliminated. The homogenized energy still involves two extra kinematic variables which can be written $\varphi = (\varphi_1, \varphi_2)$ and reads

$$\mathcal{E}(u) = \inf_{\varphi} \frac{1}{2} \int_{\Omega} (\lambda \|\nabla u_3 - \varphi\|^2 + \zeta \|e(\varphi)\|^2) \, dx_1 \, dx_2$$

(with $\lambda = \frac{9}{4}$ and $\zeta = \frac{1}{2}$). This corresponds to an isotropic Mindlin–Reissner plate [Reissner 1985; Sab and Lebée 2015]. Generally, in this theory, φ is interpreted as the rotation of the "fiber" which differs from the rotation of the "mid-surface".

5.3.3. *Generalized Mindlin–Reissner plate.* We are not limited to the extra kinematic variable φ introduced in the previous section. We get two such variables when considering three superposed triangular lattices instead of two (see Figure 18), for instance assuming that the lattice is defined by a cell Y made of three nodes at points $y_1 = (0, 0, 0)$, $y_2 = (0, 0, 1)$, and $y_3 = (0, 0, -2)$ and the periodicity vectors $t_1 = (1, 0, 0)$ and $t_2 = (-\frac{1}{2}, \sqrt{3}/2, 0)$. All components of the matrices a_p vanish but $a_{1,1,2} = a_{2,1,1} = a_{2,2,2} = a_{3,1,1} = a_{3,2,2} = a_{4,1,1} = a_{4,2,2} = a_{1,1,3} = a_{2,3,3} = a_{3,3,3} = a_{4,3,3} = 1$.
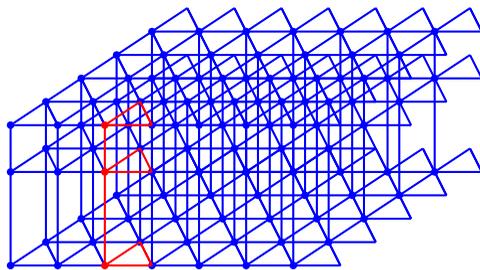
Figure 18. A generalized Mindlin–Reissner plate.
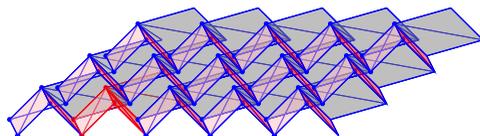
Figure 19. An "origami" plate.

In that case we obtain a model with two "rotation" vectors φ and ψ and an energy of type

$$\mathscr{E}(u) = \inf_{\varphi} \frac{1}{2} \int_{\Omega} (\lambda \|\nabla u_3 - \varphi\|^2 + \zeta \|\nabla u_3 - \psi\|^2 + Q(\varphi, \psi, \nabla\varphi, \nabla\psi)) \, dx_1 \, dx_2$$

where Q is a nonnegative quadratic form. It is not worth giving here the precise values of λ and ζ nor detailing Q.

Multiple layers could also be considered leading to more extra kinematic variables. These models correspond to the generalized Mindlin–Reissner plates recently described in [Lebée and Sab 2017].

Another way for generalizing Reissner models is to mix the structures described in Figures 18 and 14 in order to mix the in-plane Cosserat effect obtained in Section 5.2.7 and Reissner effect. Then we would get a model similar to Reissner but with φ living in $\mathbb{R}^3$ like described in [Altenbach and Eremeyev 2009].

5.3.4. *Origami-type plate.* We consider now a lattice made of four nodes at points $y_1 = (0, 0, 0)$, $y_2 = (1, 0, 1)$, $y_3 = (-1, 1, 0)$, and $y_4 = (0, 1, 1)$ with periodicity vectors $t_1 = (2, 0, 0)$ and $t_2 = (0, 2, 0)$; the interaction matrices are defined by $a_{1,1,2} = a_{1,1,3} = a_{1,2,3} = a_{1,2,4} = a_{1,3,4} = a_{2,2,1} = a_{2,2,3} = a_{2,4,1} = a_{2,4,3} = a_{3,3,1} = a_{3,3,2} = a_{3,3,2} = a_{3,4,2} = a_{4,4,1} = a_{5,2,3} = 1$ (all other components vanish). This simulates a Miura fold which is suspected to have exotic mechanical properties [Lebée and Sab 2012]: nodes correspond to wedges of the fold while interactions correspond to edges and diagonals of the faces (see Figure 19).

Constraint (22) reads $e_{1,2}(u) = 0$ and $e_{1,1}(u) = e_{2,2}(u)$. Microadjustment $b = 0$ is optimal, and the limit energy reads

$$\mathcal{E}(u) = \frac{1}{2} \int_\Omega (\lambda(\Delta u_3)^2 + \zeta(e_{1,1}(u) + e_{2,2}(u))^2) \, dx_1 \, dx_2$$

(with $\lambda = \frac{1}{64}$ and $\zeta = \frac{61}{9}$). As a membrane, only isotropic dilatation is admissible and no in-plane second gradient effects are present. As far as transverse displacements are concerned, the structure is degenerated: a curvature is possible with zero elastic energy provided the total curvature vanishes (this behavior is clearly visible when one manipulates this type of fold). From the mathematical point of view, compactness is not ensured and the homogenization result can only be applied when some extra confinement potential is present.

5.3.5. *Reinforced origami plate.* In the previous example the faces of the structure, made by a planar parallelogram with one diagonal, are very easy to bend. Let us reinforce each of them by adding an out-of-plane node and linking it to the four corners of the face. We add $y_5 = (0, 0, 1)$, $y_6 = (1, 0, 0)$, $y_7 = (-1, 1, 1)$, and $y_8 = (0, 1, 0)$, and we add the interactions $a_{1,1,5} = a_{1,2,5} = a_{1,3,5} = a_{1,4,5} = a_{1,2,6} = a_{1,4,6} = a_{2,6,1} = a_{2,6,3} = a_{1,4,8} = a_{1,3,8} = a_{3,8,1} = a_{3,8,2} = a_{1,3,7} = a_{5,2,7} = a_{3,7,1} = a_{2,4,7} = 1$. The constraint is unchanged, but the effective energy $\mathcal{E}(u)$ becomes

$$\frac{1}{2} \int_\Omega \left(\lambda \left((\Delta u_3)^2 + \left(\frac{\partial^2 u_3}{\partial x_1 \partial x_2} \right)^2 + \left(\frac{\partial^2 u_2}{\partial x_1 \partial x_2} \right)^2 + \left(\frac{\partial^2 u_2}{\partial x_1^2} \right)^2 \right) \right.$$
$$\left. + \mu \left(\frac{\partial^2 u_3}{\partial x_1^2} - \frac{\partial^2 u_3}{\partial x_2^2} \right)^2 + \zeta(e_{1,1}(u) + e_{2,2}(u))^2 \right) dx_1 \, dx_2$$

(with $\lambda = \frac{1}{64}$, $\mu = \frac{1}{192}$, and $\zeta \approx 14.06$). Now the plate is nondegenerate: transverse displacement is controlled. As far as in-plane displacement is concerned, the strain tensor takes the form $e = k \, \mathrm{Id}$ (note that compatibility conditions induce strong constraints for the second derivatives of k) and the corresponding part of the energy reads

$$\int_\Omega \left(\lambda \left(\frac{\partial k}{\partial x_1} \right)^2 + \zeta k^2 \right) dx_1 \, dx_2.$$

The effective membrane is endowed with the intrinsic length $\sqrt{\lambda/\zeta}$.

5.4. *Materials.* It is difficult to describe clearly and even more to draw periodic lattices with a 3D periodicity. Indeed the number of nodes and edges increases considerably. So we limit ourselves to studying the simple regular cubic lattice and the lattice obtained by replacing each "fiber" of this cubic lattice by a pantographic structure as described in Section 5.1.3.
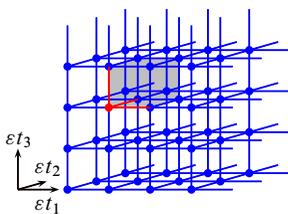
Figure 20. The regular cubic lattice.

5.4.1. *Cubic lattice.* Let us extend Section 5.2.2 to dimension 3 by considering a periodic lattice (see Figure 20) made by a single node ($K = 1$); three vectors $t_1 = (1, 0, 0)$, $t_2 = (0, 1, 0)$, and $t_3 = (0, 0, 1)$, for translating the cell; and fourteen 1×1 matrices a_p defining the interactions between the node $y_{I,1}^\varepsilon$ and its neighbors $y_{I+p,1}^\varepsilon$ given by $a_1 = [0]$, $a_2 = [1]$, $a_3 = [1]$, $a_4 = [1]$, and $a_p = [0]$ for $p > 4$.

Constraint (22) reads $e_{1,1}(u) = e_{2,2}(u) = e_{3,3}(u) = 0$: the structure is inextensible in directions e_1, e_2, and e_3. Only shear is allowed. Again $b = 0$ is an optimal microadjustment, and the limit energy is

$$\mathscr{E}(u) = \frac{1}{2} \int_\Omega \lambda \|e(u)\|^2 \, dx_1 \, dx_2 \, dx_3$$

(with $\lambda = 3$). This structure is a classical elastic material which presents no second gradient effect nor generalized continuum effect.

5.4.2. *Weaved pantographs.* We can see the structure of Section 5.4.1 as made by three families of parallel fibers. Now let us replace the fibers with direction e_1 by pantographic beams P_{e_1,e_3} and those with direction e_2 or e_3 by pantographic beams P_{e_2,e_1} or P_{e_3,e_2}, respectively. These beams share the common node $y_3 = 0$, so our new structure is made of a cell containing 16 nodes with 24 internal edges and 12 edges linking it to its neighbors (see Figure 21).

The effective material resulting from the homogenization of this structure is not subject to any constraint. Microadjustment $b = 0$ is still optimal, and the limit
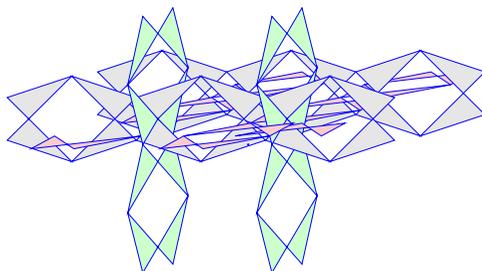
Figure 21. Weaved pantographs.

energy reads

$$\mathscr{E}(u) = \frac{1}{2} \int_\Omega \left[\lambda \left(\left(\frac{\partial^2 u_1}{\partial x_1^2} \right)^2 + \left(\frac{\partial^2 u_2}{\partial x_2^2} \right)^2 + \left(\frac{\partial^2 u_3}{\partial x_3^2} \right)^2 \right) \right.$$
$$+ \mu \left(\left(\frac{\partial^2 u_1}{\partial x_2^2} \right)^2 + \left(\frac{\partial^2 u_2}{\partial x_3^2} \right)^2 + \left(\frac{\partial^2 u_3}{\partial x_1^2} \right)^2 \right)$$
$$+ \xi ((e_{3,1}(u))^2 + (e_{2,3}(u))^2 + (e_{1,2}(u))^2)$$
$$\left. + \zeta ((e_{1,1}(u))^2 + (e_{2,2}(u))^2 + (e_{3,3}(u))^2) \right] dx_1 \, dx_2 \, dx_3,$$

with $\lambda = \frac{2}{23}$, $\mu = \frac{2}{63}$, $\zeta = 324$, and $\xi \approx 3.91$.

We obtain here a complete strain gradient 3D material. This example illustrates the huge variety of models which can be obtained by homogenizing lattice structures.

6. Conclusion

Let us conclude with some remarks.

Our starting point is a lattice made of welded bars with extensional, flexural, and torsional rigidities. The reader could think that, as bending stiffness is by itself a second gradient effect, it is the source of the effective second gradient effects. Surprisingly enough, it is not the case: second gradient effects are due to the extensional stiffness of the bars and to particular designs of the periodic cell while the bending stiffness of the bars is, on the contrary, the source of the first gradient effects in the homogenized energy. Ostoja-Starzewski [2002] has foreseen that lattices can be very useful for giving a micromechanical insight of nonclassical continua, but the role played by the nonextensional part of the mechanical interactions there is overestimated.

Strain gradient and micromorphic models are often presented as competing models. For some researchers, strain gradient models correspond simply to the limit case of micromorphic models in which the coupling between strain and microdeformation is infinitely strong. For other ones, generalized continuum models are regularizations of strain gradient models. Our results show that both effects appear generally together and at the same level.

It is also remarkable that, in our results, strain and strain gradient are never coupled. There is no fundamental reason which prevents such a coupling in a strain gradient model. Some symmetries could explain this absence of coupling [Auffray et al. 2009; Poncelet et al. 2017], but our general homogenization result does not ask for any symmetry in the design of the structure. The point is that strain gradient terms and classical strain terms come from two different sources

(extensional and flexural/torsional energies) which are assumed at the very beginning to be uncoupled. We think that considering nonhomogeneous or nonisotropic bars would likely lead to coupled models.

The general closure result [Camar-Eddine and Seppecher 2003] allows for effective models more exotic than the ones we have presented in this paper, for instance with an elastic energy depending on the third gradient of the displacement. Indeed we already mentioned that we can design our structures in order to get a degenerate effective energy. In that case, it is natural to rescale again the original energy by multiplying it by ε^{-2} (or equivalently to act with much lighter forces on the sample), and hope that the limit energy will become nondegenerate. Moreover, one would have to assume that the bars are still slenderer in order to get a compatible flexural energy. In that case, increasing the formal expansion (10) up to order 3, one should likely get third-order models. And, of course, the process can be pushed further. It is not clear whether one can get any reachable effective model by homogenizing frame lattices. Such inverse problem has been addressed in the dynamic case in [Carcaterra et al. 2015].

We have tried to get experimental evidence of second gradient effects (see Figure 13) for the structures described in Section 5.2.6. Up to now, our efforts have been unsuccessful. We think that the major reason for that is twofold: (i) geometrical nonlinearities arise very quickly in these microstructures and (ii) the limit model is extremely sensitive to design; indeed we have checked that a small modification of the position of one node of the periodic cell is enough to change the effective model from strain gradient model to a totally rigid body. Hence, the basic assumption of linear elasticity that current and initial configurations coincide is too strong and the extension of our study to nonlinear elasticity should be undertaken.

Acknowledgments

References

[Abdoul-Anziz and Seppecher 2017] H. Abdoul-Anziz and P. Seppecher, "Octave/MATLAB package for homogenizing periodic graph-based elastic structures", software, 2017, available at http://seppecher.univ-tln.fr/homogenizer/Octave/.

[Abdoul-Anziz and Seppecher 2018] H. Abdoul-Anziz and P. Seppecher, "Homogenization of periodic graph-based elastic structures", *J. Éc. Polytech. Math.* **5** (2018), 259–288.

[Aifantis 1992] E. C. Aifantis, "On the role of gradients in the localization of deformation and fracture", *Int. J. Eng. Sci.* **30**:10 (1992), 1279–1299.

[Aifantis 1999] E. C. Aifantis, "Strain gradient interpretation of size effects", *Int. J. Fracture* **95** (1999), 299–314.

[Alibert and Della Corte 2015] J.-J. Alibert and A. Della Corte, "Second-gradient continua as homogenized limit of pantographic microstructured plates: a rigorous proof", *Z. Angew. Math. Phys.* **66**:5 (2015), 2855–2870.

[Alibert et al. 2003] J.-J. Alibert, P. Seppecher, and F. dell'Isola, "Truss modular beams with deformation energy depending on higher displacement gradients", *Math. Mech. Solids* **8**:1 (2003), 51–73.

[Allaire 1992] G. Allaire, "Homogenization and two-scale convergence", *SIAM J. Math. Anal.* **23**:6 (1992), 1482–1518.

[Allaire et al. 2016] G. Allaire, M. Briane, and M. Vanninathan, "A comparison between two-scale asymptotic expansions and Bloch wave expansions for the homogenization of periodic structures", *SeMA J.* **73**:3 (2016), 237–259.

[Altenbach and Eremeyev 2009] H. Altenbach and V. A. Eremeyev, "On the linear theory of micropolar plates", *Z. Angew. Math. Mech.* **89**:4 (2009), 242–256.

[Auffray et al. 2009] N. Auffray, R. Bouchet, and Y. Bréchet, "Derivation of anisotropic matrix for bi-dimensional strain-gradient elasticity behavior", *Int. J. Solids Struct.* **46**:2 (2009), 440–454.

[Babuška and Sauter 2004] I. Babuška and S. A. Sauter, "Algebraic algorithms for the analysis of mechanical trusses", *Math. Comp.* **73**:248 (2004), 1601–1622.

[Barbagallo et al. 2017a] G. Barbagallo, A. Madeo, I. Azehaf, I. Giorgio, F. Morestin, and P. Boisse, "Bias extension test on an unbalanced woven composite reinforcement: experiments and modeling via a second-gradient continuum approach", *J. Compos. Mater.* **51**:2 (2017), 153–170.

[Barbagallo et al. 2017b] G. Barbagallo, A. Madeo, F. Morestin, and P. Boisse, "Modelling the deep drawing of a 3D woven fabric with a second gradient model", *Math. Mech. Solids* **22**:11 (2017), 2165–2179.

[Bellieud 2017] M. Bellieud, "Homogenization of stratified elastic composites with high contrast", *SIAM J. Math. Anal.* **49**:4 (2017), 2615–2665.

[Bellieud and Gruais 2005] M. Bellieud and I. Gruais, "Homogenization of an elastic material reinforced by very stiff or heavy fibers: non-local effects, memory effects", *J. Math. Pures Appl.* (9) **84**:1 (2005), 55–96.

[Bellieud et al. 2016] M. Bellieud, G. Geymonat, and F. Krasucki, "Asymptotic analysis of a linear isotropic elastic composite reinforced by a thin layer of periodically distributed isotropic parallel stiff fibres", *J. Elasticity* **122**:1 (2016), 43–74.

[Bensoussan et al. 1978] A. Bensoussan, J.-L. Lions, and G. Papanicolaou, *Asymptotic analysis for periodic structures*, Studies in Mathematics and its Applications **5**, North-Holland, 1978.

[Bouchitté and Bellieud 2002] G. Bouchitté and M. Bellieud, "Homogenization of a soft elastic material reinforced by fibers", *Asymptot. Anal.* **32**:2 (2002), 153–183.

[Boutin 1996] C. Boutin, "Microstructural effects in elastic composites", *Int. J. Solids Struct.* **33**:7 (1996), 1023–1051.

[Boutin et al. 2017] C. Boutin, F. dell'Isola, I. Giorgio, and L. Placidi, "Linear pantographic sheets: asymptotic micro-macro models identification", *Math. Mech. Complex Syst.* **5**:2 (2017), 127–162.

[Bouyge et al. 2002] F. Bouyge, I. Jasiuk, S. Boccara, and M. Ostoja-Starzewski, "A micromechanically based couple-stress model of an elastic orthotropic two-phase composite", *Eur. J. Mech. A Solid.* **21**:3 (2002), 465–481.

[Braides 2002] A. Braides, Γ-*convergence for beginners*, Oxford Lecture Series in Mathematics and its Applications **22**, Oxford University, 2002.

[Braides and Gelli 2002] A. Braides and M. S. Gelli, "Limits of discrete systems with long-range interactions", *J. Convex Anal.* **9**:2 (2002), 363–399.

[Briane and Camar-Eddine 2007] M. Briane and M. Camar-Eddine, "Homogenization of two dimensional elasticity problems with very stiff coefficients", *J. Math. Pures Appl.* (9) **88**:6 (2007), 483–505.

[Camar-Eddine and Seppecher 2002] M. Camar-Eddine and P. Seppecher, "Closure of the set of diffusion functionals with respect to the Mosco-convergence", *Math. Models Methods Appl. Sci.* **12**:8 (2002), 1153–1176.

[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.

[Carcaterra et al. 2015] A. Carcaterra, F. dell'Isola, R. Esposito, and M. Pulvirenti, "Macroscopic description of microscopically strongly inhomogenous systems: a mathematical basis for the synthesis of higher gradients metamaterials", *Arch. Ration. Mech. Anal.* **218**:3 (2015), 1239–1262.

[Cherednichenko et al. 2006] K. D. Cherednichenko, V. P. Smyshlyaev, and V. V. Zhikov, "Non-local homogenized limits for composite media with highly anisotropic periodic fibres", *Proc. Roy. Soc. Edinburgh Sect. A* **136**:1 (2006), 87–114.

[Dal Maso 1993] G. Dal Maso, *An introduction to Γ-convergence*, Progress in Nonlinear Differential Equations and their Applications **8**, Birkhäuser, 1993.

[Davini 2013] C. Davini, "Homogenization of linearly elastic honeycombs", *Math. Mech. Solids* **18**:1 (2013), 3–23.

[Davini and Ongaro 2011] C. Davini and F. Ongaro, "A homogenized model for honeycomb cellular materials", *J. Elasticity* **104**:1–2 (2011), 205–226.

[dell'Isola and Seppecher 1997] F. dell'Isola and P. Seppecher, "Edge contact forces and quasi-balanced power", *Meccanica* **32**:1 (1997), 33–52.

[dell'Isola et al. 2015a] F. dell'Isola, U. Andreaus, and L. Placidi, "At the origins and in the vanguard of peridynamics, non-local and higher-gradient continuum mechanics: an underestimated and still topical contribution of Gabrio Piola", *Math. Mech. Solids* **20**:8 (2015), 887–928.

[dell'Isola et al. 2015b] F. dell'Isola, P. Seppecher, and A. Della Corte, "The postulations *á la d'Alembert* and *á la Cauchy* for higher gradient continuum theories are equivalent: a review of existing results", *P. Roy. Soc. A Math. Phys.* **471**:2183 (2015), 20150415.

[dell'Isola et al. 2016a] F. dell'Isola, I. Giorgio, M. Pawlikowski, and N. L. Rizzi, "Large deformations of planar extensible beams and pantographic lattices: heuristic homogenization, experimental and numerical examples of equilibrium", *P. Roy. Soc. A Math. Phys.* **472**:2185 (2016), 20150790.

[dell'Isola et al. 2016b] F. dell'Isola, A. Madeo, and P. Seppecher, "Cauchy tetrahedron argument applied to higher contact interactions", *Arch. Ration. Mech. Anal.* **219**:3 (2016), 1305–1341.

[dell'Isola et al. 2017] F. dell'Isola, A. Della Corte, and I. Giorgio, "Higher-gradient continua: the legacy of Piola, Mindlin, Sedov and Toupin and some future research perspectives", *Math. Mech. Solids* **22**:4 (2017), 852–872.

[Dos Reis and Ganghoffer 2010] F. Dos Reis and J. F. Ganghoffer, "Discrete homogenization of architectured materials: implementation of the method in a simulation tool for the systematic prediction of their effective elastic properties", *Tech. Mech.* **30**:1–3 (2010), 85–109.

[Eringen 2001] A. C. Eringen, *Microcontinuum field theories, II: Fluent media*, Springer, 2001.

[Forest 1999] S. Forest, "Homogenization methods and the mechanics of generalized continua", pp. 35–48 in *Geometry, Continua and Microstructure*, edited by G. Maugin, Travaux en Cours **60**, Hermann, 1999.

[Germain 1973] P. Germain, *Cours de mécanique des milieux continus*, tome I: Théorie générale, Masson et Cie, 1973.

[Gibson and Ashy 1997] L. J. Gibson and M. F. Ashy, *Cellular solids: structure and properties*, 2nd ed., Cambridge University, 1997.

[Giorgio 2016] I. Giorgio, "Numerical identification procedure between a micro-Cauchy model and a macro-second gradient model for planar pantographic structures", *Z. Angew. Math. Phys.* **67**:4 (2016), 95.

[Gonella and Ruzzene 2008] S. Gonella and M. Ruzzene, "Homogenization and equivalent in-plane properties of two-dimensional periodic lattices", *Int. J. Solids Struct.* **45**:10 (2008), 2897–2915.

[Harrison 2016] P. Harrison, "Modelling the forming mechanics of engineering fabrics using a mutu-ally constrained pantographic beam and membrane mesh", *Compos. A Appl. S.* **81** (2016), 145–157.

[Khruslov 1991] E. Ya. Khruslov, "Homogenized models of composite media", pp. 159–182 in *Composite media and homogenization theory* (Trieste, Italy, 1990), edited by G. Dal Maso and G. F. Dell'Antonio, Progress in Nonlinear Differential Equations and their Applications **5**, Birkhäuser, 1991.

[Le 2015] D. T. Le, *Modèle d'endommagement à gradient: approche par homogénéisation*, Ph.D. thesis, Université Pierre et Marie Curie – Paris VI, 2015, available at https://tel.archives-ouvertes.fr/tel-01367480.

[Le Dret and Raoult 2013] H. Le Dret and A. Raoult, "Homogenization of hexagonal lattices", *Netw. Heterog. Media* **8**:2 (2013), 541–572.

[Lebée and Sab 2012] A. Lebée and K. Sab, "Homogenization of thick periodic plates: application of the bending-gradient plate theory to a folded core sandwich panel", *Int. J. Solids Struct.* **49**:19–20 (2012), 2778–2792.

[Lebée and Sab 2017] A. Lebée and K. Sab, "On the generalization of Reissner plate theory to laminated plates, I: Theory", *J. Elasticity* **126**:1 (2017), 39–66.

[Leung and Guest 2007] A. C. H. Leung and S. D. Guest, "Single member actuation of kagome lattice structures", *J. Mech. Mater. Struct.* **2**:2 (2007), 303–317.

[Madeo et al. 2017] A. Madeo, A. Della Corte, I. Giorgio, and D. Scerrato, "Modeling and design-ing micro- and nano-structured metamaterials: towards the application of exotic behaviors", *Math. Mech. Solids* **22**:4 (2017), 873–884.

[Martinsson and Babuška 2007a] P. G. Martinsson and I. Babuška, "Homogenization of materials with periodic truss or frame micro-structures", *Math. Models Methods Appl. Sci.* **17**:5 (2007), 805–832.

[Martinsson and Babuška 2007b] P.-G. Martinsson and I. Babuška, "Mechanics of materials with periodic truss or frame micro-structures", *Arch. Ration. Mech. Anal.* **185**:2 (2007), 201–234.

[Mazière and Forest 2015] M. Mazière and S. Forest, "Strain gradient plasticity modeling and finite element simulation of Lüders band formation and propagation", *Contin. Mech. Thermodyn.* **27**:1–2 (2015), 83–104.

[Meunier et al. 2012] N. Meunier, O. Pantz, and A. Raoult, "Elastic limit of square lattices with three-point interactions", *Math. Models Methods Appl. Sci.* **22**:11 (2012), 1250032.

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

[Mindlin 1963] R. D. Mindlin, "Influence of couple-stresses on stress concentrations", *Exp. Mech.* **3**:1 (1963), 1–7.

[Mindlin 1965] R. D. Mindlin, "Second gradient of strain and surface-tension in linear elasticity", *Int. J. Solids Struct.* **1**:4 (1965), 417–438.

[Mindlin and Tiersten 1962] R. D. Mindlin and H. F. Tiersten, "Effects of couple-stresses in linear elasticity", *Arch. Rational Mech. Anal.* **11** (1962), 415–448.

[Nguetseng 1989] G. Nguetseng, "A general convergence result for a functional related to the theory of homogenization", *SIAM J. Math. Anal.* **20**:3 (1989), 608–623.

[Oleĭnik et al. 1992] O. A. Oleĭnik, A. S. Shamaev, and G. A. Yosifian, *Mathematical problems in elasticity and homogenization*, Studies in Mathematics and its Applications **26**, North-Holland, 1992.

[Ostoja-Starzewski 2002] M. Ostoja-Starzewski, "Lattice models in micromechanics", *Appl. Mech. Rev.* **55**:1 (2002), 35–60.

[Pastukhova 2005] S. E. Pastukhova, "Homogenization of problems of elasticity theory on periodic box and rod frames of critical thickness", *J. Math. Sci.* **130**:5 (2005), 4954–5004.

[Pideri and Seppecher 1997] C. Pideri and P. Seppecher, "A second gradient material resulting from the homogenization of an heterogeneous linear elastic medium", *Contin. Mech. Thermodyn.* **9**:5 (1997), 241–257.

[Placidi et al. 2016] L. Placidi, E. Barchiesi, E. Turco, and N. L. Rizzi, "A review on 2D models for the description of pantographic fabrics", *Z. Angew. Math. Phys.* **67**:5 (2016), 121.

[Placidi et al. 2017] L. Placidi, U. Andreaus, and I. Giorgio, "Identification of two-dimensional pantographic structure via a linear D4 orthotropic second gradient elastic model", *J. Engrg. Math.* **103** (2017), 1–21.

[Polizzotto and Borino 1998] C. Polizzotto and G. Borino, "A thermodynamics-based formulation of gradient-dependent plasticity", *Eur. J. Mech. A Solids* **17**:5 (1998), 741–761.

[Poncelet et al. 2017] M. Poncelet, N. Auffray, C. Jailin, A. Somera, and C. Morel, "Experimental strain gradient evidence in non-central symmetric lattice", in *Generalized and microstructured continua* (Arpino, Italy, 2017), EUROMECH Colloquium **579**, European Mechanics Society, 2017.

[Rahali et al. 2015] Y. Rahali, I. Giorgio, J. F. Ganghoffer, and F. dell'Isola, "Homogenization à la Piola produces second gradient continuum models for linear pantographic lattices", *Internat. J. Engrg. Sci.* **97** (2015), 148–172.

[Raoult et al. 2008] A. Raoult, D. Caillerie, and A. Mourad, "Elastic lattices: equilibrium, invariant laws and homogenization", *Ann. Univ. Ferrara Sez. VII Sci. Mat.* **54**:2 (2008), 297–318.

[Reissner 1985] E. Reissner, "Reflections on the theory of elastic plates", *Appl. Mech. Rev.* **38**:11 (1985), 1453–1464.

[Sab and Lebée 2015] K. Sab and A. Lebée, *Homogenization of heterogeneous thin and thick plates*, Wiley, 2015.

[Sánchez-Palencia 1980] E. Sánchez-Palencia, *Nonhomogeneous media and vibration theory*, Lecture Notes in Physics **127**, Springer, 1980.

[Sciarra et al. 2007] G. Sciarra, F. dell'Isola, and O. Coussy, "Second gradient poromechanics", *Int. J. Solids Struct.* **44**:20 (2007), 6607–6629.

[Seppecher et al. 2011] P. Seppecher, J.-J. Alibert, and F. dell'Isola, "Linear elastic trusses leading to continua with exotic mechanical interactions", *J. Phys. Conf. Ser.* **319**:1 (2011), 012018.

[Tartar 1989] L. Tartar, "Nonlocal effects induced by homogenization", pp. 925–938 in *Partial differential equations and the calculus of variations*, vol. II, edited by F. Colombini et al., Progress in Nonlinear Differential Equations and their Applications **2**, Birkhäuser, 1989.

[Tartar 2009] L. Tartar, *The general theory of homogenization: a personalized introduction*, Lecture Notes of the Unione Matematica Italiana **7**, Springer, 2009.

[Toupin 1962] R. A. Toupin, "Elastic materials with couple-stresses", *Arch. Rational Mech. Anal.* **11** (1962), 385–414.

[Toupin 1964] R. A. Toupin, "Theories of elasticity with couple-stress", *Arch. Rational Mech. Anal.* **17** (1964), 85–112.

[Triantafyllidis and Aifantis 1986] N. Triantafyllidis and E. C. Aifantis, "A gradient approach to localization of deformation, I: Hyperelastic materials", *J. Elasticity* **16**:3 (1986), 225–237.

[Turco et al. 2016a] E. Turco, F. dell'Isola, A. Cazzani, and N. L. Rizzi, "Hencky-type discrete model for pantographic structures: numerical comparison with second gradient continuum models", *Z. Angew. Math. Phys.* **67**:4 (2016), 85.

[Turco et al. 2016b] E. Turco, M. Golaszewski, A. Cazzani, and N. L. Rizzi, "Large deformations induced in planar pantographic sheets by loads applied on fibers: experimental validation of a discrete Lagrangian model", *Mech. Res. Commun.* **76** (2016), 51–56.

[Turco et al. 2017] E. Turco, M. Golaszewski, I. Giorgio, and F. D'Annibale, "Pantographic lattices with non-orthogonal fibres: experiments and their numerical simulations", *Compos. B Eng.* **118** (2017), 1–14.

[Yang and Misra 2012] Y. Yang and A. Misra, "Micromechanics based second gradient continuum theory for shear band modeling in cohesive granular materials following damage elasticity", *Int. J. Solids Struct.* **49**:18 (2012), 2500–2514.

[Zhikov 2002] V. V. Zhikov, "Homogenization of elasticity problems on singular structures", *Izv. Math.* **66**:2 (2002), 299–365.

[Zhikov and Pastukhova 2003] V. V. Zhikov and S. E. Pastukhova, "Homogenization for elasticity problems on periodic networks of critical thickness", *Sb. Math.* **194**:5 (2003), 697–732.

HOUSSAM ABDOUL-ANZIZ: houssam@imath.fr
Institut de Mathématiques de Toulon, Université de Toulon, Toulon, France

PIERRE SEPPECHER: seppecher@imath.fr
Institut de Mathématiques de Toulon, Université de Toulon, Toulon, France
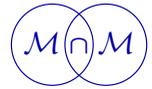
ON THE EFFECT OF PHASE TRANSITION
ON THE MANIFOLD DIMENSIONALITY:
APPLICATION TO THE ISING MODEL

ELENA LOPEZ, ADRIEN SCHEUER,
EMMANUELLE ABISSET-CHAVANNE AND FRANCISCO CHINESTA

Fields can be represented in a discrete manner from their values at some locations, the nodes when considering finite element descriptions. Thus, each discrete scalar solution can be considered as a point in $\mathbb{R}^N$ (N being the number of nodes used for approximating the scalar field). Most manifold learning techniques (linear and nonlinear) are based on the fact that those solutions define a slow manifold of dimension $n \ll N$ embedded in the space $\mathbb{R}^N$. This paper explores such a behavior in systems exhibiting phase transitions in order to analyze the evolution of the local dimensionality n when the system moves from one side of the critical behavior to the other. For that purpose we consider the Ising model.

1. Introduction

Physical models usually involve unknown continuous scalar or vector fields. If we consider without any loss of generality a model involving a scalar field approximated on a mesh or grid involving N nodes, the discrete solution can be represented as a point belonging to $\mathbb{R}^N$. However, existing correlations lead to solutions that, instead of filling the whole space $\mathbb{R}^N$, define a slow n-dimensional manifold, with $n \ll N$, embedded into $\mathbb{R}^N$.

This fact is at the origin of manifold learning approaches that consist of extracting the uncorrelated (latent) dimensions describing the slow manifold representing the original states into a reduced form. Principal component analysis (PCA) was specially designed to find a linear subspace of lower dimensionality than the original space; however, it fails when the manifold becomes extremely nonlinear. Nonlinear dimensionality reduction techniques were proposed for circumventing this limitation. Among the many existing techniques (an abundant literature is available on the topic, and the interested reader can refer to [Lee and Verleysen 2007] and the references therein) kernel-based PCA (kPCA) [Wang 2014; Schölkopf et al.

1998; 1999] and locally linear embedding (LLE) [Roweis and Saul 2000] will be considered in the present work. The latter can be considered as a particular case of the former, for a particular kernel choice [Zimmer et al. 2015].

These techniques that automatically extract the latent dimensions have been extensively and successfully applied in many domains; however, few works addressed the issue of phase transition and in particular the way in which that transition affects the manifold dimensionality and the possibility of characterizing the transition from a geometrical point of view. To address that issue, we consider in the present work a well established and widely studied equation exhibiting phase transition, the Ising model.

In the context of the microscopic theory of critical behavior [Koonin and Meredith 1990; Fisher 1965], the Ising model [Newell and Montroll 1953] addresses a ferromagnetic lattice in a quite simple statistical description, including phase transition. Ferromagnetic materials exhibit long-range spin ordering at the atomic level. When a magnetic field is applied to a ferromagnetic material, atomic spins align along the direction of the applied field. However, when the temperature becomes higher than the so-called critical temperature, a phase change occurs. The Ising model allows describing such materials [Myers 1997]. This model initially proposed by Lenz was solved in 1D by Ising [1925]. The 2D model without magnetic field was then developed by Onsager [1944], and it is this one that we are considering in the present work. Its solution that will be addressed by using the Monte Carlo method [Metropolis et al. 1953] concerns a binary property called spin, given each point of a 2D grid, and more specifically its time evolution with respect to the applied temperature. When reaching the so-called critical temperature, phase transition occurs. Even if complex variants of the Ising model were proposed [Bellettini et al. 2007; De Masi et al. 2009; 2008], in the present work we consider its simplest version.

Our feeling is that applying usual manifold learning techniques on the spin field itself, consisting of a binary variable defined at each grid (mesh) node, is neither the best nor the optimal representation because those techniques are based on the use of euclidean metrics and two similar microstructures can significantly differ when calculating the euclidean distance between both of them. For that reason, a first contribution of the present work is to propose a better microstructure representation much less sensitive to local distributions. In that sense, considering Fourier-based representations, where space dependencies are described from their frequency content, seems a much better alternative, in particular, the 2D fast Fourier transform (FFT) 2D-FFT that has been widely considered in image analysis to identify biological damage [Fung et al. 2010], extract ordered structures from microtomography [Jeulin and Moreaud 2008], analyze complex microstructures [Lebensohn et al. 2011; Zhu et al. 2018], etc.

When applying the FFT to solutions coming from the Ising model, one expects low frequencies at low temperatures, where long-range correlation exists, whereas at high temperatures, where entropic effects dominate, higher frequency contents are expected. However, at the transition in between, richer structures with its a priori higher dimensionality are expected. Thus, we would like to locally define the manifold dimensionality, and instead of using global manifold learning techniques, local extractors seem more appropriate. In particular local principal component analysis (lPCA) [Kambhatla and Leen 1997] seems specially appealing for that purpose. It is important to note that lPCA encountered limited interest as a manifold learning technique because of the difficulty of defining a continuous mapping throughout the manifold. However, it has been widely employed in other domains where the main goal, as in the present case, was to estimate the local dimensionality [Fukunaga and Olsen 1971].

The paper is structured as follows. Section 2 revisits the main concepts related to the Ising model, as well as its solution using the Metropolis-based Monte Carlo algorithm. Section 3 addresses the application of LLE for constructing the manifold on which data will be classified using the procedure proposed in Section 4. Finally, Section 5 presents and discusses different numerical solutions, proving the potential of the proposed approach.

2. The Ising model and its Monte Carlo solution

The Ising model describes a ferromagnetic behavior and is able to capture phase transitions [Niss 2005]. In this model, the discrete variables called "spins" are defined on a lattice. Each lattice site has only one spin with value either -1 (spin down) or $+1$ (spin up), interacting only with its nearest neighbors. By considering the 2D Ising model introduced by Onsager [1944], the Hamiltonian H reads

$$H = -J \sum_i \sum_{\substack{j \\ |i-j|=1}} S_i S_j, \qquad (1)$$

J being the dimensionless interaction strength. S_i and S_j are the spin states at lattice sites i and j, respectively. As can be noticed, the sum of products reduces to the nearest-neighbor pair of spins. Every spin will interact with four other spins (up, down, left, and right). If spins are aligned and $J > 0$, the energy of the system will be minimal.

The partition function Z can be computed from the Hamiltonian

$$Z = \sum_i e^{-H_i}, \qquad (2)$$

where the sum applies over all the particles in the lattice. Then, we compute the probability of finding the system in a particular state S from

$$p(S) = \frac{e^{-H(S)}}{Z}. \tag{3}$$

Some important outputs are the magnetization Ma

$$\mathrm{Ma} = \sum_S p(S) \sum_i S_i, \tag{4}$$

and the system energy E

$$E = \sum_S p(S) \beta H(S), \tag{5}$$

where $\beta = k_B T$, with the Boltzmann constant $k_B = 1.380658 \times 10^{-23}$ J/K and T the temperature.

Usually the solution of the Ising model is performed by using a Monte Carlo method. In this paper, we consider a variant of the Monte Carlo method, the so-called Metropolis algorithm [Fricke 2006]. For a given temperature, at each time step, a trial spin configuration is generated. The algorithm computes the system energy associated with the trial state. If the change of energy is negative, it means that the system evolution has brought the system to a state of lower energy, so we allow the change and put the spins in their new state. On the other hand, if the change of energy is positive, we allow the change with a probability given by $p(S)$ computed from (3).

3. Manifold construction

It is well known that microstructures do not allow simple reduced descriptions [Lopez et al. 2018]. In fact the main concern is how to quantify similarities or resemblances, and how to take profit of them.

In this paper we consider and analyze a route based on the use of the locally linear embedding (LLE) technique [Roweis and Saul 2000], a member of the large family of the so-called manifold learning techniques.

The remainder of this section describes the procedure directly on the problem we are interested in. The procedure consists of two steps: the analysis of the Ising samples in order to obtain a discriminative description of the microstructures (in this case as previously said, we do not work directly on the Ising microstructures but on their fast Fourier transform), followed by a dimensionality reduction able to discriminate the three temperature zones that characterize the Ising model behavior.

3.1. *Applying the fast Fourier transform on the Ising microstructures.* First we assume the existence of M microstructures $\mathcal{M}_m$, $m = 1, \ldots, M$, coming from the

Ising model defined on the domain ω. In what follows and without loss of generality, we consider 2D microstructures and the 2D fast Fourier transform (2D-FFT). Moreover, we assume the existence of two phases, the spins with value 1 and the others with value -1, occupying the domains ω_1^m and ω_{-1}^m, respectively, with $\omega_1^m \cup \omega_{-1}^m = \omega$, $m = 1, \ldots, M$. A regular mesh is associated with each microstructure consisting of N nodes ($N = N_n^2$, with N_n the number of nodes along the x and y directions). The coordinates of each node are $\boldsymbol{x}_i$, $i = 1, \ldots, N$ ($\boldsymbol{x}_i^T = (x_i, y_i)$).

For each microstructure $\mathcal{M}_m$ we define the phase field $\chi(\boldsymbol{x}; \mathcal{M}_m)$ as

$$\chi(\boldsymbol{x}; \mathcal{M}_m) = \begin{cases} 1 & \text{if } \boldsymbol{x} \in \omega_1^m, \\ -1 & \text{if } \boldsymbol{x} \in \omega_{-1}^m. \end{cases} \tag{6}$$

As is well known, these microstructure descriptions do not allow simple reduced descriptions. Thus, the objective is to geometrically characterize them in order to obtain reduced descriptions. Concerning geometrical characterization and microstructural descriptions, there are several tools proposed for signal processing purposes. Concerning the type of microstructures we are dealing with, it seems more appropriate to work in the frequency domain instead of using their physical space description. In particular we propose to apply the 2D-FFT to characterize the microstructures in the frequency domain.

We now apply the 2D-FFT on each microstructure $\mathcal{M}_m$, $m = 1, \ldots, M$, defined by its phase field $\chi(\boldsymbol{x}; \mathcal{M}_m)$. Thus, the 2D-FFT of the Ising microstructures can be represented in a discrete way from vectors $\hat{\boldsymbol{\chi}}^m$, $\mathcal{M}_m$, $m = 1, \ldots, M$. Vectors $\hat{\boldsymbol{\chi}}^m$ are defined in $\mathbb{R}^N$; i.e., the dimension coincides with the number of nodes considered in the discrete microscopic description.

3.2. *Nonlinear dimensionality reduction.* Each vector $\hat{\boldsymbol{\chi}}^m$ defines a point in a space of dimension N, and then the data set of M 2D-FFTs related to the M Ising microstructures represents a set of M points in $\mathbb{R}^N$. Hence, the question: do all these points belong to a particular low-dimensional manifold embedded in the high-dimensional space $\mathbb{R}^N$?

Imagine that, despite the impressive space dimension N, the M points belong to a curve, a surface, or a hypersurface of dimension $n \ll N$. When $N = 3$ a simple observation suffices to check if these points are located on a curve (1D manifold) or on a surface (2D manifold). However, when dealing with spaces of thousands of dimensions, simple visual observation is unsuitable.

Instead, appropriate techniques are needed to extract the underlying manifold (when it exists) when dealing with extremely multidimensional spaces. There is a variety of techniques to accomplish this task. The interested reader can refer to [Tenenbaum et al. 2000; Roweis and Saul 2000; Polito and Perona 2001; Wang 2014; Amsallem and Farhat 2008]. In this work we focus on the LLE (locally linear embedding) technology [Roweis and Saul 2000]. It proceeds as follows.

- Each point $\hat{\chi}^m$, $m = 1, \ldots, M$, is linearly reconstructed from its K-nearest neighbors. In principle K should be greater than the expected dimension n of the underlying manifold and the points should be close enough to ensure the validity of the linear approximation. In general, a large-enough number of neighbors K and a dense-enough sampling M ensure a satisfactory reconstruction as shown later. For each point $\hat{\chi}^m$ we can write the locally linear data reconstruction as

$$\hat{\chi}^m = \sum_{i \in \mathscr{S}_m} W_{mi} \hat{\chi}^i, \tag{7}$$

where W_{mi} are the unknown weights and $\mathscr{S}_m$ is the set of the K-nearest neighbors of $\hat{\chi}^m$.

Weights, grouped in vector W, result from the minimization of functional

$$\mathscr{F}(W) = \sum_{m=1}^{M} \left\| \hat{\chi}^m - \sum_{i=1}^{M} W_{mi} \hat{\chi}^i \right\|^2, \tag{8}$$

where here W_{mi} is zero if $\hat{\chi}^i$ does not belong to the set of K-nearest neighbors of $\hat{\chi}^m$.

The minimization of $\mathscr{F}(W)$ allows us to determine all the weights involved in all the locally linear data reconstruction.

- We suppose now that each linear patch around $\hat{\chi}^m$, for all m, is mapped into a lower-dimensional embedding space of dimension n, $n \ll N$. Because of the linear mapping of each patch, weights remain unchanged. The problem now becomes the determination of the coordinates of each point $\hat{\chi}^m$ when it is mapped into the low-dimensional space, $\xi^m \in \mathbb{R}^n$.

For this purpose a new functional $\mathscr{G}$ is introduced that depends on the searched coordinates $\xi^1, \ldots, \xi^M$:

$$\mathscr{G}(\xi^1, \ldots, \xi^M) = \sum_{m=1}^{M} \left\| \xi^m - \sum_{i=1}^{M} W_{mi} \xi^i \right\|^2, \tag{9}$$

where now the weights are known and the reduced coordinates ξ^m are unknown.

The minimization of functional $\mathscr{G}$ results in an $M \times M$ eigenvalue problem whose n-bottom nonzero eigenvalues define the set of orthogonal coordinates in which the manifold is mapped.

4. Discriminating criteria

Once the manifold composed by the reduced coordinates ξ^m is known, the goal is to cluster in some way the three temperature zones to be able to classify them

in low, transition, and high temperature. For this purpose, we propose a criterion based on a variant of the principal component analysis (PCA), the so-called local principal component analysis (lPCA) that allows us to compute the manifold local dimension, and then infer the dimension at low, high, and critical temperatures.

Principal components analysis (PCA) attempts to find a linear subspace of lower dimensionality than the original space. If data exhibit more complex structures which cannot be well represented in a linear subspace, standard PCA fails to accomplish the reduction. However, such reduction can be successfully performed by using nonlinear dimensionality reduction techniques, like the lPCA here considered. In what follows we revisit first the standard PCA before focusing on its local counterpart.

Let us consider n observed variables defining the vector (snapshot) $\boldsymbol{\xi} \in \mathbb{R}^n$. We assume that these variables are therefore not totally uncorrelated and, notably, that there exists a linear transformation $\boldsymbol{L}$ defining the vector $\boldsymbol{\xi}^{\mathrm{red}} \in \mathbb{R}^{n^{\mathrm{red}}}$, $n^{\mathrm{red}} \leq n$, that represents the so-called latent reduced variables, according to

$$\boldsymbol{\xi} = \boldsymbol{L}\boldsymbol{\xi}^{\mathrm{red}}. \tag{10}$$

We assume the existence of M different snapshots $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_M$ that can be stored in the columns of the $n \times M$ matrix $\mathbb{X}$. The associated $n^{\mathrm{red}} \times M$ reduced matrix $\mathbb{Y}$ contains the associated reduced vectors $\boldsymbol{\xi}_i^{\mathrm{red}}$, $i = 1, \ldots, M$.

PCA proceeds by enforcing the fact that latent variables must be as much as possible uncorrelated, and allows us to extract both the dimension n^{red} and the mapping $\boldsymbol{L}$. For that purpose the covariance matrix $\boldsymbol{C}_{xx}$,

$$\boldsymbol{C}_{xx} = E\big\{(\mathbb{X} - E\{\mathbb{X}\})(\mathbb{X} - E\{\mathbb{X}\})^T\big\}, \tag{11}$$

is factorized as

$$\boldsymbol{C}_{xx} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^T, \tag{12}$$

equivalent to applying the singular value decomposition (SVD) to $\mathbb{X}$. In (12), $\boldsymbol{V}$ contains the orthonormal eigenvectors and $\boldsymbol{\Lambda}$ is the diagonal matrix containing the eigenvalues (nonnegative real numbers), assumed to be in descending order. Thus, the n^{red} columns of $\boldsymbol{L}$ are the n^{red} first columns of $\boldsymbol{V}$ [Lee and Verleysen 2007].

From this summary of the PCA rationale, we can now briefly explain its local counterpart. For that purpose, we consider each snapshot $\boldsymbol{\xi}_i$, for $i = 1, \ldots, M$, and for each of them, its K-nearest neighbors. From those and by proceeding as just indicated, we can compute the local transformation matrix $\boldsymbol{L}_i$, $i = 1, \ldots, M$, as well as the local reduced dimensionality n_i^{red}.
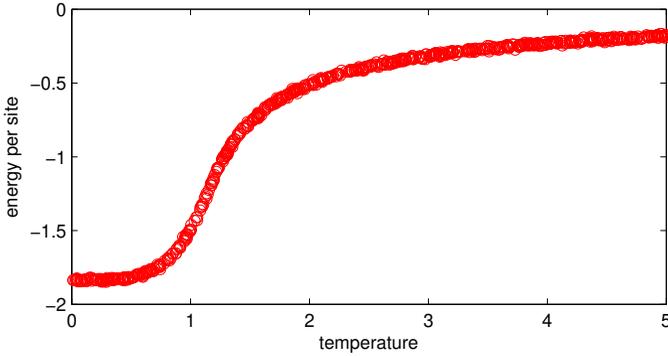
Figure 1. Energy versus temperature.

5. Numerical results

Ising microstructures have been generated using the Metropolis-based Monte Carlo method proposed in [Fricke 2006]. We have run $M = 741$ simulations to obtain our microstructure data set (snapshots), $\mathcal{M}_m$, $m = 1, \ldots, M$. The parameters considered in those simulations are a 200×200 grid ($N = 40000$ nodes) with the dimensionless temperature T randomly chosen between 0 and 5. Figure 1 shows the energy per site E of the final configuration of each simulation. The magnetization per site in turn is shown in Figure 2.

It is clear that a phase transition occurs in between $T = 1$ and $T = 1.5$. At lower temperatures, $T < 1$, the system tends to any of the two ground states $Ma = -1$ or $Ma = +1$. At higher temperatures, $T > 1.5$, the spins tend to align randomly, leading to an almost vanishing magnetization ($Ma \approx 0$).

To illustrate the Ising microstructures configuration, Figure 3 depicts microstructures associated with low, transition, and high temperature. In the one associated with low temperature, approximately half of the spins are up and the other half
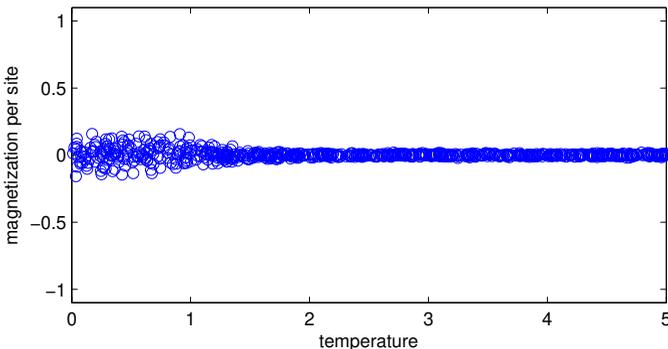
Figure 2. Magnetization versus temperature.
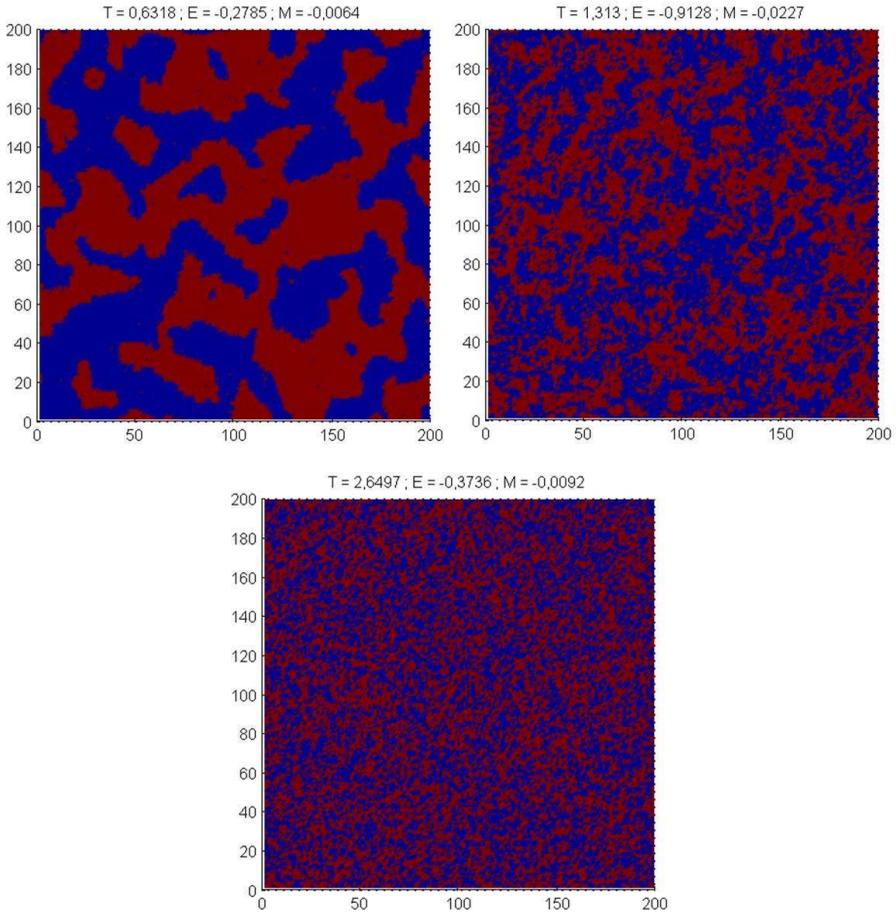
Figure 3. Low, transition, and high temperature microstructures (clockwise from top left).

down, with aligned spins forming a sort of clusters called "metastable states", with Ma $\approx$ 0. Around the transition temperature, the number of down spins has increased (up spins consequently decreasing) compared to the previous scenario. This means that the system has enough energy and the Metropolis algorithm easily accepts antiparallel spins. Finally, in the high-temperature state, almost the same number of down and up spins are found with a small characteristic length describing the phases distribution in the microstructures due to the high energy communicated to the system.

The M samples were described by their phase field vectors χ^m, $m = 1, \ldots, M$, each defined in $\mathbb{R}^N$. But as explained in Section 3.1, microstructures do not allow

reduced representations and that reason motivated the consideration of 2D-FFT on the *M* Ising microstructures in order to obtain the data reduced representation.

Figure 4 illustrates the 2D-FFT representation of three microstructures associated with low, transition, and high temperature, respectively, considered in Figure 3. In the low-temperature microstructure, we can appreciate the presence of very few significant frequencies, almost located at the center of the domain (low frequencies), and the rest of the domain is practically flat.
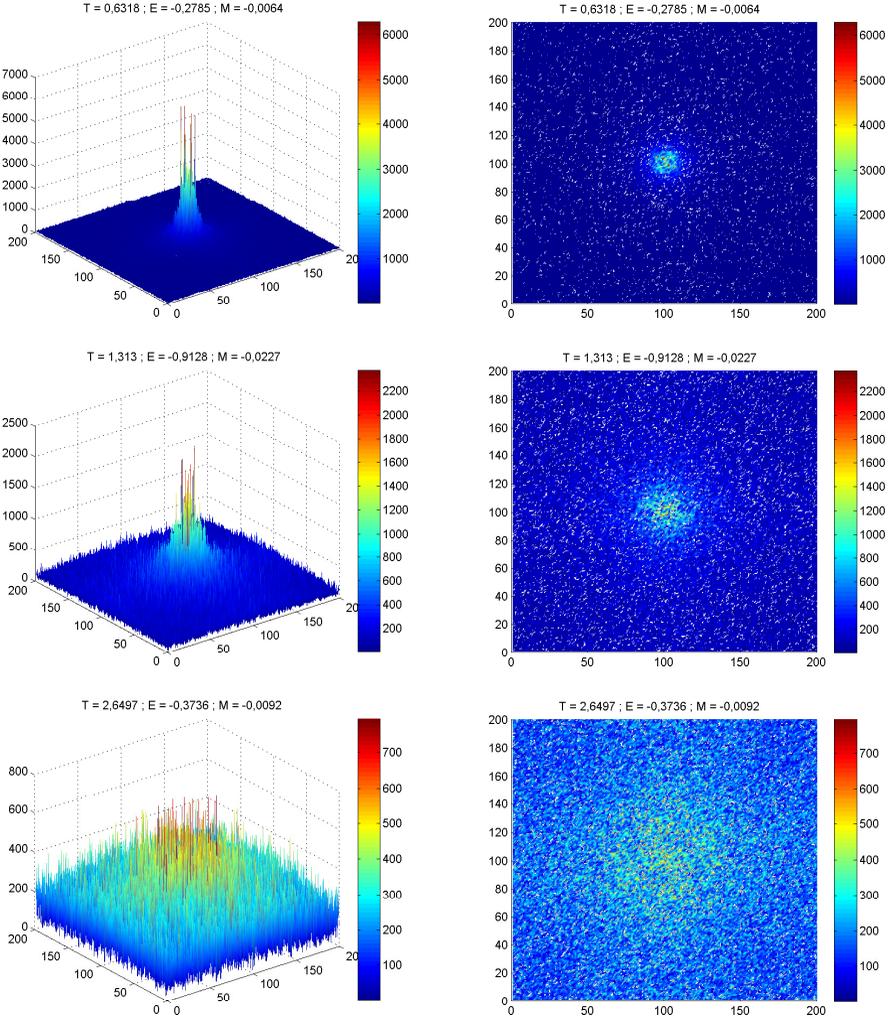
Figure 4. 3D (left) and 2D (right) representation of the 2D-FFT for low (top), transition (middle), and high (bottom) temperature microstructures.
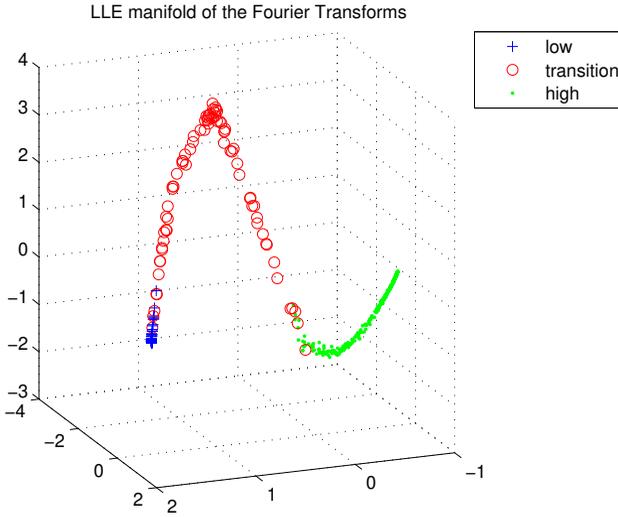
Figure 5. LLE manifold of the 2D-FFT on the Ising microstructures. Axes represent a 3D space in which the almost 1D manifold is embedded.

This result is intuitive since the spins of the microstructures are more clustered in the low-temperature samples. On the contrary, in the case of 2D-FFT applied to the high-temperature microstructures, we observe the opposite result. Here, spins are totally randomly distributed and consequently all frequencies are present. In the case of microstructures around the transition temperature, its 2D-FFT becomes a mix of the aforementioned cases, with dominant frequencies located at the center of the domain but now the rest of the domain is no longer flat, since it is populated by the contribution of nonnegligible higher frequencies.

We denote by $\hat{\chi}^m$, $m = 1, \ldots, M$, the 2D-FFT of the Ising microstructures. By applying the locally linear embedding technique on $\hat{\chi}^m$, as explained in Section 3.2, the weights involved in the linear data reconstruction are calculated as well as the reduced data. The performed analysis allows us to consider a reduced dimension $n = 3$ that moreover facilitates the solution's graphical representation. Figure 5 depicts the resulting points $\boldsymbol{\xi}^m \in \mathbb{R}^3$, $m = 1, \ldots, M$, from which one can realize that the manifold is almost 1D.

Now in order to properly check the local dimensionality, local principal component analysis is applied in locations belonging to low, high, and transition temperatures, whose associated regions can be clearly identified in Figure 5. After applying the lPCA in these three regions (the solution represented in the manifold of Figure 5), the highest eigenvalues are extracted determining the local manifold dimensionality. As shown in Figure 6, the local dimensionality of low and high
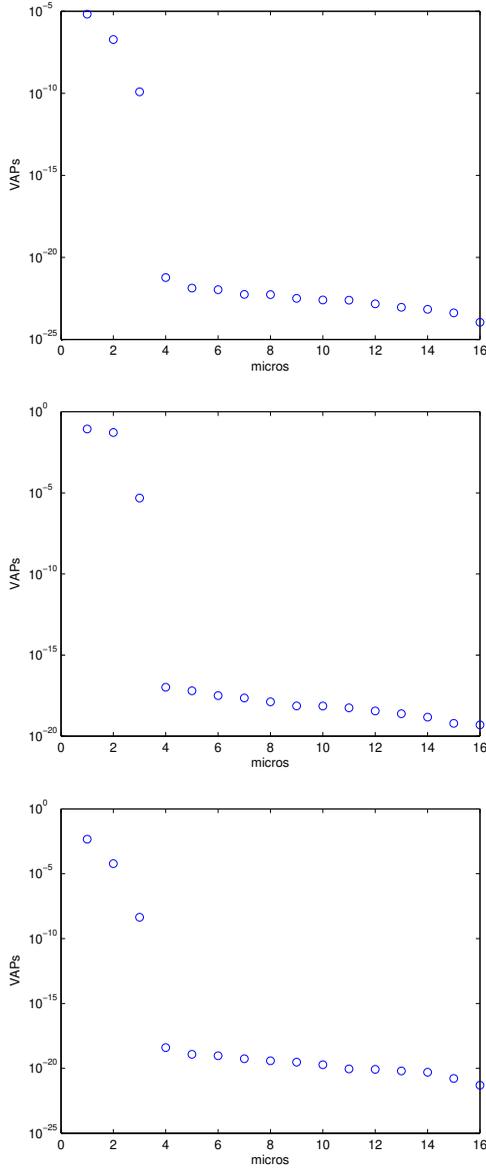
Figure 6. Local dimensionality of the low, transition, and high temperature zones (top to bottom).

temperatures is one because there is a difference of almost two orders of magnitude between the values of the first and the second eigenvalues, which means that the first is much more important than the second.

This result was expected, since looking at the manifold of Figure 5, it could be clearly seen that the points corresponding to these two zones defined a rather

1D manifold. In the transition region the first two eigenvalues have practically the same value and differ from the third by four orders of magnitude, which clearly means that the local dimensionality of this region is almost two. This result was also expected, since the points in this zone mix the 1D behaviors of low and high temperatures and being not collinear increases the dimensionality. Local dimension is not able to distinguish between the two phases of the system. Descriptions able to differentiate them constitute a work in progress.

6. Conclusions

This paper proposes a methodology to interpret phase transition from a geometrical point of view, from the local dimensionality of the manifold defined from different microscopic fields. For that purpose the Ising model is simulated by varying the temperature from one side to other of the critical temperature associated with the phase transition.

Using the Metropolis algorithm, a group of microstructures related to the solution of the Ising model has been created, encompassing samples corresponding to low, high, and transition temperatures. Unable to deal directly with the microstructures' characteristic function, the 2D-FFT has been applied to those because the frequency content is expected to exacerbate the difference between microstructures at low, high, and transition temperatures. The manifold was then created on the 2D-FFT representations of phases distribution, and then a local PCA was applied locally in the different regions to extract the local dimensionality.

The performed analysis reveals that the dimensionality slightly increases in the transition region where complex microstructures mixing low and high temperature patterns coexist. Below and above that transition temperature the solution seems to be explained by only one latent variable that in the present case can be associated with the temperature. In the transition region solutions seem a bit richer and cannot be explained by a single latent variable.

References

[Amsallem and Farhat 2008] D. Amsallem and C. Farhat, "Interpolation method for adapting reduced-order models and application to aeroelasticity", *AIAA J.* **46**:7 (2008), 1803–1813.

[Bellettini et al. 2007] G. Bellettini, A. De Masi, N. Dirr, and E. Presutti, "Tunneling in two dimensions", *Comm. Math. Phys.* **269**:3 (2007), 715–763.

[De Masi et al. 2008] A. De Masi, I. Merola, E. Presutti, and Y. Vignaud, "Potts models in the continuum: uniqueness and exponential decay in the restricted ensembles", *J. Stat. Phys.* **133**:2 (2008), 281–345.

[De Masi et al. 2009] A. De Masi, I. Merola, E. Presutti, and Y. Vignaud, "Coexistence of ordered and disordered phases in Potts models in the continuum", *J. Stat. Phys.* **134**:2 (2009), 243–306.

[Fisher 1965] M. E. Fisher, "The nature of critical points", pp. 1–159 in *Lectures in theoretical physics*, vol. VIIC, edited by W. E. Brittin, University of Colorado, 1965.

[Fricke 2006] T. Fricke, "Monte Carlo investigation of the Ising model", preprint, 2006, Available at https://www.physics.ohio-state.edu/~braaten/statphys/Ising_MatLab.pdf.

[Fukunaga and Olsen 1971] K. Fukunaga and D. R. Olsen, "An algorithm for finding intrinsic dimensionality of data", *IEEE T. Comput.* **C-20**:2 (1971), 176–183.

[Fung et al. 2010] D. T. Fung, J. B. Sereysky, J. Basta-Pljakic, D. M. Laudier, R. Huq, K. J. Jepsen, M. B. Schaffler, and E. L. Flatow, "Second harmonic generation imaging and Fourier transform spectral analysis reveal damage in fatigue-loaded tendons", *Ann. Biomed. Eng.* **38**:5 (2010), 1741–1751.

[Ising 1925] E. Ising, "Beitrag zur Theorie des Ferromagnetismus", *Z. Phys.* **31** (1925), 253–258.

[Jeulin and Moreaud 2008] D. Jeulin and M. Moreaud, "Segmentation of 2D and 3D textures from estimates of the local orientation", *Image Anal. Stereol.* **27**:3 (2008), 183–192.

[Kambhatla and Leen 1997] N. Kambhatla and T. K. Leen, "Dimension reduction by local principal component analysis", *Neural Comput.* **9**:7 (1997), 1493–1516.

[Koonin and Meredith 1990] S. E. Koonin and D. Meredith, *Computational physics: FORTRAN version*, Addison-Wesley, 1990.

[Lebensohn et al. 2011] R. A. Lebensohn, A. D. Rollett, and P. Suquet, "Fast Fourier transform-based modeling for the determination of micromechanical fields in polycrystals", *J. Mater.* **63**:3 (2011), 13–18.

[Lee and Verleysen 2007] J. A. Lee and M. Verleysen, *Nonlinear dimensionality reduction*, Springer, 2007.

[Lopez et al. 2018] E. Lopez, D. Gonzalez, J. V. Aguado, E. Abisset-Chavanne, E. Cueto, C. Binetruy, and F. Chinesta, "A manifold learning approach for integrated computational materials engineering", *Arch. Comput. Methods Eng.* **25**:1 (2018), 59–68.

[Metropolis et al. 1953] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, and A. H. Teller, "Equation of state calculations by fast computing machines", *J. Chem. Phys.* **21**:6 (1953), 1087–1092.

[Myers 1997] H. P. Myers, *Introductory solid state physics*, 2nd ed., Taylor & Francis, 1997.

[Newell and Montroll 1953] G. F. Newell and E. W. Montroll, "On the theory of the Ising model of ferromagnetism", *Rev. Modern Physics* **25** (1953), 353–389.

[Niss 2005] M. Niss, "History of the Lenz–Ising model 1920–1950: from ferromagnetic to cooperative phenomena", *Arch. Hist. Exact Sci.* **59**:3 (2005), 267–318.

[Onsager 1944] L. Onsager, "Crystal statistics, I: A two-dimensional model with an order-disorder transition", *Phys. Rev.* (2) **65** (1944), 117–149.

[Polito and Perona 2001] M. Polito and P. Perona, "Grouping and dimensionality reduction by locally linear embedding", pp. 1255–1262 in *Advances in Neural Information Processing Systems 14* (Vancouver, 2001), edited by T. G. Dietterich et al., MIT, 2001.

[Roweis and Saul 2000] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding", *Science* **290**:5500 (2000), 2323–2326.

[Schölkopf et al. 1998] B. Schölkopf, A. Smola, and K.-R. Müller, "Nonlinear component analysis as a kernel eigenvalue problem", *Neural Comput.* **10**:5 (1998), 1299–1319.

[Schölkopf et al. 1999] B. Schölkopf, A. Smola, and K.-R. Müller, "Kernel principal component analysis", Chapter 20, pp. 327–352 in *Advances in kernel methods*, edited by C. J. C. Burges et al., MIT, 1999.

[Tenenbaum et al. 2000] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction", *Science* **290**:5500 (2000), 2319–2323.

[Wang 2014] Q. Wang, "Kernel principal component analysis and its applications in face recognition and active shape models", preprint, 2014. arXiv

[Zhu et al. 2018] J. Zhu, R. Balieu, X. Lu, and N. Kringos, "Microstructure evaluation of polymer-modified bitumen by image analysis using two-dimensional fast Fourier transform", *Mater. Design* **137** (2018), 164–175.

[Zimmer et al. 2015] V. A. Zimmer, K. Lekadir, C. Hoogendoorn, A. F. Frangi, and G. Piella, "A framework for optimal kernel-based manifold embedding of medical image data", *Comput. Med. Imag. Grap.* **41** (2015), 93–107.

ELENA LOPEZ: elena.lopez-tomas@ec-nantes.fr
Institut de Calcul Intensif, École Centrale de Nantes, Nantes, France

ADRIEN SCHEUER: adrien.scheuer@ec-nantes.fr
Institut de Calcul Intensif, École Centrale de Nantes, Nantes, France

and

Institute of Information and Communication Technologies, Electronics and Applied Mathematics, Université catholique de Louvain, Louvain-la-Neuve, Belgium

EMMANUELLE ABISSET-CHAVANNE: emmanuelle.abisset-chavanne@ec-nantes.fr
Institut de Calcul Intensif, École Centrale de Nantes, Nantes, France

FRANCISCO CHINESTA: francisco.chinesta@ensam.eu
Procédés et Ingénierie en Mécanique et Matériaux, Arts et Métiers ParisTech, Paris, 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.