

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

vol. 4

no. 1

2016

MATHEMATICS AND MECHANICS  
*of*  
**Complex Systems**

# MATHEMATICS AND MECHANICS OF COMPLEX SYSTEMS

[msp.org/memocs](http://msp.org/memocs)

## EDITORIAL BOARD

ANTONIO CARCATERRA  
ERIC A. CARLEN  
FRANCESCO DELL'ISOLA  
RAFFAELE ESPOSITO  
ALBERT FANNJIANG  
GILLES A. FRANCFORT  
PIERANGELO MARCATI  
JEAN-JACQUES MARIGO  
PETER A. MARKOWICH  
MARTIN OSTOJA-STARZEWSKI  
PIERRE SEPPECHER  
DAVID J. STEIGMANN  
PAUL STEINMANN  
PIERRE M. SUQUET

Università di Roma "La Sapienza", Italia  
Rutgers University, USA  
(CO-CHAIR) Università di Roma "La Sapienza", Italia  
(TREASURER) Università dell'Aquila, Italia  
University of California at Davis, USA  
(CO-CHAIR) Université Paris-Nord, France  
Università dell'Aquila, Italy  
École Polytechnique, France  
DAMTP Cambridge, UK, and University of Vienna, Austria  
(CHAIR MANAGING EDITOR) Univ. of Illinois at Urbana-Champaign, USA  
Université du Sud Toulon-Var, France  
University of California at Berkeley, USA  
Universität Erlangen-Nürnberg, Germany  
LMA CNRS Marseille, France

## MANAGING EDITORS

MICOL AMAR  
CORRADO LATTANZIO  
ANGELA MADEO  
MARTIN OSTOJA-STARZEWSKI

Università di Roma "La Sapienza", Italia  
Università dell'Aquila, Italy  
Université de Lyon-INSA (Institut National des Sciences Appliquées), France  
(CHAIR MANAGING EDITOR) Univ. of Illinois at Urbana-Champaign, USA

## ADVISORY BOARD

ADNAN AKAY  
HOLM ALTENBACH  
MICOL AMAR  
HARM ASKES  
TEODOR ATANACKOVIĆ  
VICTOR BERDICHEVSKY  
GUY BOUCHITTÉ  
ANDREA BRAIDES  
ROBERTO CAMASSA  
MAURO CARFORE  
ERIC DARVE  
FELIX DARVE  
ANNA DE MASI  
GIANPIETRO DEL PIERO  
EMMANUELE DI BENEDETTO  
BERNOLD FIEDLER  
IRENE M. GAMBA  
DAVID Y. GAO  
SERGEY GAVRILYUK  
TIMOTHY J. HEALEY  
DOMINIQUE JEULIN  
ROGER E. KHAYAT  
CORRADO LATTANZIO  
ROBERT P. LIPTON  
ANGELO LUONGO  
ANGELA MADEO  
JUAN J. MANFREDI  
CARLO MARCHIORO  
GÉRARD A. MAUGIN  
ROBERTO NATALINI  
PATRIZIO NEFF  
ANDREY PIATNITSKI  
ERRICO PRESUTTI  
MARIO PULVIRENTI  
LUCIO RUSSO  
MIGUEL A. F. SANJUAN  
PATRICK SELVADURAI  
ALEXANDER P. SEYRANIAN  
MIROSLAV ŠILHAVÝ  
GUIDO SWEERS  
ANTOINETTE TORDÉSILLAS  
LEV TRUSKINOVSKY  
JUAN J. L. VELÁZQUEZ  
VINCENZO VESPRI  
ANGELO VULPIANI

Carnegie Mellon University, USA, and Bilkent University, Turkey  
Otto-von-Guericke-Universität Magdeburg, Germany  
Università di Roma "La Sapienza", Italia  
University of Sheffield, UK  
University of Novi Sad, Serbia  
Wayne State University, USA  
Université du Sud Toulon-Var, France  
Università di Roma Tor Vergata, Italia  
University of North Carolina at Chapel Hill, USA  
Università di Pavia, Italia  
Stanford University, USA  
Institut Polytechnique de Grenoble, France  
Università dell'Aquila, Italia  
Università di Ferrara and International Research Center MEMOCS, Italia  
Vanderbilt University, USA  
Freie Universität Berlin, Germany  
University of Texas at Austin, USA  
Federation University and Australian National University, Australia  
Université Aix-Marseille, France  
Cornell University, USA  
École des Mines, France  
University of Western Ontario, Canada  
Università dell'Aquila, Italy  
Louisiana State University, USA  
Università dell'Aquila, Italia  
Université de Lyon-INSA (Institut National des Sciences Appliquées), France  
University of Pittsburgh, USA  
Università di Roma "La Sapienza", Italia  
Université Paris VI, France  
Istituto per le Applicazioni del Calcolo "M. Picone", Italy  
Universität Duisburg-Essen, Germany  
Narvik University College, Norway, Russia  
Università di Roma Tor Vergata, Italy  
Università di Roma "La Sapienza", Italia  
Università di Roma "Tor Vergata", Italia  
Universidad Rey Juan Carlos, Madrid, Spain  
McGill University, Canada  
Moscow State Lomonosov University, Russia  
Academy of Sciences of the Czech Republic  
Universität zu Köln, Germany  
University of Melbourne, Australia  
École Polytechnique, France  
Bonn University, Germany  
Università di Firenze, Italia  
Università di Roma La Sapienza, Italia

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

Cover image: "Tangle" by © John Horigan; produced using the *Context Free* program ([contextfreetart.org](http://contextfreetart.org)).

PUBLISHED BY

 **mathematical sciences publishers**  
nonprofit scientific publishing  
<http://msp.org/>

© 2016 Mathematical Sciences Publishers



# GRADIENT MATERIALS WITH INTERNAL CONSTRAINTS

ALBRECHT BERTRAM AND RAINER GLÜGE

The concept of internal constraints is extended to gradient materials. Here, interesting constraints can be introduced, such as pseudorigid ones. The stresses and the hyperstresses will be given by constitutive equations only up to reactive parts, which do no work during any compatible motion of the body. For the inclusion of thermodynamical effects, the theory is generalized to the case of thermomechanical constraints. Here one obtains reactive parts of the stresses, heat flux, entropy, and energy, which do not contribute to the dissipation. Some critical remarks on the classical concept of internal constraints are finally given. A method to introduce internal constraints in a natural way is described to overcome some conceptual deficiencies of the classical concept.

## 1. Introduction

The theory of internal constraints as it is described in, e.g., [Truesdell and Noll 1965, Section 30] is a useful tool for the description of incompressible materials, inextensible composites, and many more material classes. It gives a conceptually sound basis upon which both theoretical and practical investigations can be developed. Particularly, it provides a change of the structure of the basic balance equations, which can be helpful for the construction of solutions of the field problem. This way, the only nonhomogeneous universal solutions for simple materials are those for constrained materials [Ericksen 1955].

On the other hand, there has been an increasing interest in nonclassical extensions of the concept of simple materials. Micromorphic, micropolar, and gradient materials are examples for a blossoming variety of new theories, which go far beyond the classical simple materials. In particular, the inclusion of the second deformation gradient opens the door for many challenging new perspectives for material modeling with the inclusion of internal length scales.

For gradient materials, one wants to introduce internal constraints other than the classical ones to again benefit from such extensions. The question arises whether such an extension is possible, or demands substantial alterations of the entire

---

**Communicated by Francesco dell’Isola.**

*MSC2010:* 74A30.

*Keywords:* internal constraints, gradient materials, pseudorigidity.

format. It turns out, and will be shown in the sequel, that such an extension is in fact straightforward once a theory of gradient materials has been constructed—at least within the mechanical context.

The extension to a thermomechanical format is more complicated. This is, however, necessary not only in order to investigate the compatibility of such constraints with the second law of thermodynamics, but also to study the temperature dependence of mechanical constraints.

There has been some discussion about a sound format for the inclusion of the thermodynamical variables into such a theory of internal constraints; see [Green et al. 1970; Trapp 1971; Andreussi and Podio Guidugli 1973; Gurtin and Podio Guidugli 1973; Bertram and Haupt 1976; Casey and Krishnaswamy 1998; Casey 2011; Bertram 2005].

The starting point for the present approach is a suggestion by [Trapp 1971; Bertram 2005], where a rate form of a thermomechanical constraint is assumed and the possibility for reactive parts of the stresses, heat fluxes, and energies is given that are not dissipative during any process that is compatible with the constraint. Again, the extension of this theory to gradient materials is straightforward.

At the end of this contribution, some critical remarks on the standard theory of mechanical constraints are given and a procedure to avoid these shortcomings is suggested.

*Notations.* Throughout the paper, a dot will denote a scalar product between tensors of arbitrary order. Vectors are denoted by small bold letters like  $\mathbf{j}$ ,  $\mathbf{m}$ ,  $\mathbf{n}$ . Second-order tensors are denoted by capital letters like  $\mathbf{J}$ ,  $\mathbf{S}$ ,  $\mathbf{T}$ , and third-order tensors by  $\mathbf{J}$ ,  $\mathbf{G}$ ,  $\mathbf{M}$ . Since there is not yet a standard notation for odd-order tensors, we have to introduce some operations. The first product we need is the pullback operation of a third-order tensor  $\mathbf{G}$  by a second-order tensor  $\mathbf{F}$ , defined as

$$\mathbf{F}^{-1} \circ \mathbf{G} = \mathbf{F}^{-1} \circ (G^{ijk} \mathbf{e}_i \otimes \mathbf{e}_j \otimes \mathbf{e}_k) := G^{ijk} (\mathbf{F}^T \mathbf{e}_j) \otimes (\mathbf{F}^{-1} \mathbf{e}_j) \otimes (\mathbf{F}^{-1} \mathbf{e}_k). \quad (1)$$

Another useful tool is the Rayleigh product between a second-order tensor  $\mathbf{F}$  and a tensor of arbitrary order  $\mathbf{J}$ , defined as

$$\mathbf{F} * \mathbf{J} := J^{i \dots j} (\mathbf{F} \mathbf{e}_j) \otimes \dots \otimes (\mathbf{F} \mathbf{e}_j). \quad (2)$$

For orthogonal tensors  $\mathbf{F}$  both products coincide.

## 2. Simple and gradient materials

A simple material is defined as one for which the Cauchy stresses are given by a history functional or a process functional of the form

$$\mathbf{T}(\mathbf{x}_0, t) = \mathcal{F}_1 \left\{ \chi(\mathbf{x}_0, \tau), \text{Grad } \chi(\mathbf{x}_0, \tau) \Big|_{\tau=0}^t \right\} \quad (3)$$

with  $\chi$  being the motion of the body in the time interval between some initial time  $\tau \equiv 0$  and some final time  $t$ , and  $\mathbf{x}_0$  a material point in the reference placement. If we submit this functional to the *principle of Euclidean invariance*, then we can find reduced forms which identically fulfill this principle. One example for a reduced form is the functional

$$\mathbf{S}(\mathbf{x}_0, t) = \mathfrak{H}_1 \{ \mathbf{C}(\mathbf{x}_0, \tau) |_{\tau=0}^t \}, \quad (4)$$

which assigns to each process in the right Cauchy–Green tensor  $\mathbf{C} = \mathbf{F}^T \mathbf{F}$  the second Piola–Kirchhoff stress tensor  $\mathbf{S} = J \mathbf{F}^{-1} \mathbf{T} \mathbf{F}^{-T}$ . Here,  $\mathbf{F}$  is the deformation gradient and  $J$  its determinant.

If we extend this format to second gradient materials<sup>1</sup> (see [Bertram 2015a]), the general form of the constitutive functional is

$$\mathbf{T}(\mathbf{x}_0, t) = \mathfrak{F} \{ \chi(\mathbf{x}_0, \tau), \text{Grad } \chi(\mathbf{x}_0, \tau), \text{Grad Grad } \chi(\mathbf{x}_0, \tau) |_{\tau=0}^t \}. \quad (5)$$

However, for second gradient materials we must expect the existence of a third-order hyperstress tensor  $\mathbf{G}$  for which a second constitutive functional of the form

$$\mathbf{G}(\mathbf{x}_0, t) = \mathfrak{G} \{ \chi(\mathbf{x}_0, \tau), \text{Grad } \chi(\mathbf{x}_0, \tau), \text{Grad Grad } \chi(\mathbf{x}_0, \tau) |_{\tau=0}^t \} \quad (6)$$

is needed. Reduced forms of these two functionals are, among others,

$$\mathbf{S}(\mathbf{x}_0, \tau) = \mathfrak{k} \{ \mathbf{C}(\mathbf{x}_0, \tau), \mathbf{K}(\mathbf{x}_0, \tau) |_{\tau=0}^t \}, \quad (7)$$

$$\mathbf{H}(\mathbf{x}_0, \tau) = \mathfrak{H} \{ \mathbf{C}(\mathbf{x}_0, \tau), \mathbf{K}(\mathbf{x}_0, \tau) |_{\tau=0}^t \}, \quad (8)$$

with the two third-order tensors called the *continuity tensor*

$$\mathbf{K} := \mathbf{F}^{-1} \text{Grad } \mathbf{F} \quad (9)$$

and the *material hyperstress tensor*

$$\mathbf{H} := \mathbf{F}^{-1} \circ J \mathbf{G}. \quad (10)$$

Both tensors are material tensors and hence invariant under rigid body motions. In [Bertram 2013; 2015a] examples for elastic and plastic materials within such a format are given.

A hyperelastic gradient material would be constituted by an energy function  $w(\mathbf{C}, \mathbf{K})$  such that the stresses are given by the potential relations

$$\mathbf{S} = 2\rho_0 \partial_{\mathbf{C}} w(\mathbf{C}, \mathbf{K}), \quad (11)$$

$$\mathbf{H} = \rho_0 \partial_{\mathbf{K}} w(\mathbf{C}, \mathbf{K}), \quad (12)$$

with the density in the reference placement  $\rho_0$ .

---

<sup>1</sup>Only these are considered here. Third- and higher-order gradients are beyond the scope of this paper.

### 3. Classical internal constraints

The classical theory of internal constraints after [Truesdell and Noll 1965] and several authors earlier and also later is based on two assumptions.

**Assumption 1a** (constraint equation for simple materials). *There are restrictions upon the possible deformations of the material element such that a scalar-valued function of the deformation gradient  $\mathbf{F}$  equals zero for all possible deformations,*

$$\gamma(\mathbf{F}) = 0. \quad (13)$$

**Assumption 2a** (principle of determinism for simple materials subject to internal constraints). *The stress is determined by the history of the deformation only to within an additive part that does no work in any possible motion satisfying the constraint.*

Truesdell and Noll [1965] put these assumptions in an axiomatic way, without giving any substantiation for them other than the plausibility of their consequences in particular applications.

If one applies the *principle of Euclidean invariance* (material objectivity)<sup>2</sup> to the material function  $\gamma$ , one can show that a function  $\gamma_{\text{red}}(\mathbf{C})$  of the right Cauchy–Green tensor  $\mathbf{C} = \mathbf{F}^T \mathbf{F}$  is objective and, hence, a reduced form.

Such a constraint would be considered as isotropic if it were invariant under arbitrary rotations, that is,

$$\gamma_{\text{red}}(\mathbf{C}) = \gamma_{\text{red}}(\mathbf{Q}\mathbf{C}\mathbf{Q}^T) \quad (14)$$

for all orthogonal tensors  $\mathbf{Q}$ . In this sense, incompressibility would be an isotropic constraint, while inextensibility in one direction is not isotropic.

By exploiting the second assumption, we start with an additive split of the Cauchy stresses into an extra part, which is determined by a constitutive functional, and a reactive part (for which no constitutive functional exists):

$$\mathbf{T} = \mathbf{T}_E + \mathbf{T}_R \quad (15)$$

so that the specific stress power of the latter,

$$\frac{1}{\rho} \mathbf{T}_R \cdot \mathbf{D} = 0 \quad (16)$$

vanishes for every process that is compatible with the constraint, where  $\mathbf{D}$  denotes the rate of stretching. We can alternatively express this assumption in terms of a

---

<sup>2</sup>For a precise introduction of this controversial issue, see [Bertram and Svendsen 2001; Bertram 2005], therein called the *principle of invariance under superimposed rigid body motions*.

reactive second Piola–Kirchhoff stress with  $\mathbf{S} = \mathbf{S}_E + \mathbf{S}_R$ :

$$\frac{1}{\rho_0} \mathbf{S}_R \cdot \mathbf{C}^\bullet = 0. \quad (17)$$

If the constraint equation is differentiable, we can bring it into the rate form

$$\gamma_{\text{red}}(\mathbf{C})^\bullet = \partial_{\mathbf{C}} \gamma_{\text{red}}(\mathbf{C}) \cdot \mathbf{C}^\bullet = 0. \quad (18)$$

If we multiply this equation by a Lagrangian multiplier  $\alpha$ , and subtract it from the constraint equation (17), we find that the second Piola–Kirchhoff reaction stress  $\mathbf{S}_R$  must have the representation

$$\mathbf{S}_R = \alpha \partial_{\mathbf{C}} \gamma_{\text{red}}(\mathbf{C}) \quad (19)$$

and the Cauchy reaction stress

$$\mathbf{T}_R = \alpha \mathbf{F} \partial_{\mathbf{C}} \gamma_{\text{red}}(\mathbf{C}) \mathbf{F}^T \quad (20)$$

with some scalar field  $\alpha$ .

As a normalization of the decomposition, we can pose the orthogonality condition

$$\mathbf{T}_R \cdot \mathbf{T}_E = 0. \quad (21)$$

This makes the decomposition unique. However, this is not necessary and often not even practical.

If there is more than one internal constraint (say  $N \leq 6$ ), one also has to allow for  $N$  reaction stresses, which can be superimposed onto the total stress as

$$\mathbf{T} = \mathbf{T}_E + \sum_{i=1}^N \alpha_i \mathbf{F} \partial_{\mathbf{C}} \gamma_{\text{red}i}(\mathbf{C}) \mathbf{F}^T \quad (22)$$

with  $N$  scalar fields  $\alpha_i$ . In the limit for six independent constraints,  $\mathbf{C}$  is completely constrained, the material is rigid, and the stresses are completely reactive. This case will be discussed in the last section of this paper.

#### 4. Nonclassical internal constraints

The question arises if one could extend this method to gradient materials. It turns out that such a generalization is straightforward.

It has been shown in [Forest and Sievert 2003; Bertram 2013; 2015a] that the specific stress power for a second gradient material can be brought into the Eulerian and Lagrangian forms

$$\frac{1}{\rho} (\mathbf{T} \cdot \text{grad } \mathbf{v} + \mathbf{G} \cdot \text{grad grad } \mathbf{v}) = \frac{1}{\rho_0} \left( \frac{1}{2} \mathbf{S} \cdot \mathbf{C}^\bullet + \mathbf{H} \cdot \mathbf{K}^\bullet \right), \quad (23)$$

respectively. We now generalize our two assumptions to gradient materials.

**Assumption 1b** (constraint equation for gradient materials). *There are restrictions upon the possible deformations of the material element such that a scalar-valued function of the motion, the deformation gradient, and the second gradient equals zero for all possible deformations*

$$\Gamma(\chi, \text{Grad } \chi, \text{Grad Grad } \chi) = 0. \quad (24)$$

If this is understood as a constitutive equation, it must fulfill the Euclidean invariance requirement. This leads to the reduced form of the nonclassical internal constraint

$$\Gamma_{\text{red}}(\mathbf{C}, \mathbf{K}) = 0. \quad (25)$$

If the constraint equation function is differentiable, we get the rate form of it as

$$\partial_{\mathbf{C}}\Gamma_{\text{red}} \cdot \mathbf{C}^\bullet + \partial_{\mathbf{K}}\Gamma_{\text{red}} \cdot \mathbf{K}^\bullet = 0. \quad (26)$$

**Assumption 2b** (principle of determinism for gradient materials subject to internal constraints). *The stresses and the hyperstresses are determined by the deformation process only to within additive parts that do no work in any possible motion satisfying the constraint.*

Accordingly, we have the decompositions  $\mathbf{T} = \mathbf{T}_E + \mathbf{T}_R$  and  $\mathbf{G} = \mathbf{G}_E + \mathbf{G}_R$  for the spatial stresses, and  $\mathbf{S} = \mathbf{S}_E + \mathbf{S}_R$  and  $\mathbf{H} = \mathbf{H}_E + \mathbf{H}_R$  for the material ones.

After [Assumption 2b](#), we have

$$\frac{1}{\rho}(\mathbf{T}_R \cdot \text{grad } \mathbf{v} + \mathbf{G}_R \cdot \text{grad grad } \mathbf{v}) = \frac{1}{\rho_0} \left( \frac{1}{2} \mathbf{S}_R \cdot \mathbf{C}^\bullet + \mathbf{H}_R \cdot \mathbf{K}^\bullet \right) = 0. \quad (27)$$

By subtracting an  $\alpha$ -multiple of the constraint equation in the rate form, we obtain

$$0 = \left[ \frac{1}{2\rho_0} \mathbf{S}_R - \alpha \partial_{\mathbf{C}}\Gamma_{\text{red}} \right] \cdot \mathbf{C}^\bullet + \left[ \frac{1}{\rho_0} \mathbf{H}_R - \alpha \partial_{\mathbf{K}}\Gamma_{\text{red}} \right] \cdot \mathbf{K}^\bullet, \quad (28)$$

so that the following equations must hold:

$$\mathbf{S}_R = \alpha 2\rho_0 \partial_{\mathbf{C}}\Gamma_{\text{red}}(\mathbf{C}, \mathbf{K}), \quad (29)$$

$$\mathbf{H}_R = \alpha \rho_0 \partial_{\mathbf{K}}\Gamma_{\text{red}}(\mathbf{C}, \mathbf{K}) \quad (30)$$

with a joint Lagrangian parameter  $\alpha$  which couples the two reactive stresses.

As a normalization of the decomposition, one can pose the orthogonality condition

$$\mathbf{H}_R \cdot \mathbf{H}_E + \frac{1}{4} \mathbf{S}_R \cdot \mathbf{S}_E = 0. \quad (31)$$

This is, however, not compulsory and perhaps not even practical.

A particular choice of the constraint equation would be to demand that certain components of  $\mathbf{K}$  must vanish. In such cases the corresponding components of  $\mathbf{F}$  must be constant in space.



For  $\mathbf{F}$  having nine independent components and the space having three linear independent directions,  $9 \times 3 = 27$  such constraints on  $\text{Grad } \mathbf{F}$  are possible. This, however, reduces to 18 independent constraints because of Schwarz's commutation law since the connection tensor has the right subsymmetry

$$\mathbf{K}_{ijk} = F_{ij,k} = \chi_{i,jk} = \chi_{i,kj} = \mathbf{K}_{ikj}. \quad (32)$$

In [Seppecher et al. 2011] one finds examples of materials with microstructures with such properties.

By imposing 18 independent constraints of this kind, the deformation gradient can only be constant in space. Bodies with this property have been investigated in the past under the label *homogeneous strains* — see [Sławianowski 1974; 1975] — or *pseudorrigidity* — see [Cohen 1981; Cohen and Muncaster 1984; Cohen and MacSithigh 1989; Antman and Marlow 1991; Casey 2004; 2006; 2007] — and critically commented upon by [Steigmann 2006]. However, these approaches are completely different from the present one, since there the homogeneity of strains is imposed on the body as a global constraint, while in the present approach we still assume local constraints as an extension of classical constraints to gradient materials.

## 5. Thermomechanical constraints

Not only in mechanics, but also more general in thermomechanics, the introduction of internal constraints is reasonable. In the literature, several suggestions have been made to generalize the mechanical concepts of constraints to thermodynamics; see [Green et al. 1970; Trapp 1971; Andreussi and Podio Guidugli 1973; Gurtin and Podio Guidugli 1973; Casey and Krishnaswamy 1998; Casey 2011; Bertram 2005]. For the extension of the theory of gradient materials to thermomechanics, see [Bertram 2005; 2015b].

In what follows, we extend the concept introduced in [Trapp 1971; Bertram 2005], where one also finds examples for thermomechanical constraints like temperature-dependent incompressibility or inextensibility.

**Assumption 1c** (constraint equation for thermomechanical gradient materials). *A thermomechanical internal constraint consists of four material functions,*

$$\mathbf{J}(\mathbf{C}, \mathbf{K}, \theta), \quad \mathbf{J}(\mathbf{C}, \mathbf{K}, \theta), \quad \mathbf{j}(\mathbf{C}, \mathbf{K}, \theta), \quad j(\mathbf{C}, \mathbf{K}, \theta), \quad (33)$$

*of the configuration and the temperature with values in third-order tensors, second-order tensors, vectors, and scalars, respectively, such that for all admissible thermo-kinematical processes the constraint equation*

$$\mathbf{J} \cdot \mathbf{K}^\bullet + \mathbf{J} \cdot \mathbf{C}^\bullet + \mathbf{j} \cdot \mathbf{g}_0 + j \cdot \theta^\bullet = 0 \quad (34)$$

*holds at each instant.*

Here,  $\theta$  denotes the temperature and  $\mathbf{g}_0$  is the material temperature gradient.

Since we made use of material variables, this constraint is already in a reduced form. Note that the first two terms, to which the equation is reduced in the isothermal case, correspond to the mechanical constraint (26) in its rate form, which is therefore included in this format as a special case.

Once again, we have to modify the *principle of determinism*. There have been several suggestions in the literature whether the reactive parts of the dependent variables shall not produce energy or entropy [Green et al. 1970; Trapp 1971; Gurtin and Podio Guidugli 1973], or neither of the two [Andreussi and Podio Guidugli 1973]. The following assumption is close to what Trapp [1971] suggested and follows [Bertram 2005].

**Assumption 2c** (principle of determinism for gradient materials with thermomechanical internal constraints). *The current values of the hyperstress, stress, heat flux, internal energy, and entropy are determined by the thermokinematical process only up to additive parts that are not dissipative during all admissible processes.*

Thus, we have the decompositions of the dependent variables into reactive parts and extra parts:

$$\begin{array}{ll}
 \text{hyperstress} & \mathbf{H} = \mathbf{H}_E + \mathbf{H}_R, \\
 \text{2nd Piola–Kirchhoff stress} & \mathbf{S} = \mathbf{S}_E + \mathbf{S}_R, \\
 \text{material heat flux} & \mathbf{q}_0 = \mathbf{q}_{0E} + \mathbf{q}_{0R}, \\
 \text{internal energy} & \varepsilon = \varepsilon_E + \varepsilon_R, \\
 \text{entropy} & \eta = \eta_E + \eta_R,
 \end{array} \tag{35}$$

and, consequently, also for the

$$\text{free energy} \quad \psi = \varepsilon_E + \varepsilon_R - \theta \eta_E - \theta \eta_R =: \psi_E + \psi_R, \tag{36}$$

where only the extra terms depend on the thermokinematical process.

The reactive parts shall not be dissipative in the sense of the Clausius–Duhem inequality:

$$\frac{1}{\rho_0} \left( \frac{1}{2} \mathbf{S}_R \cdot \mathbf{C}^\bullet + \mathbf{H}_R \cdot \mathbf{K}^\bullet \right) - \frac{1}{\theta \rho_0} \mathbf{q}_{0R} \cdot \mathbf{g}_0 - \psi_{R^\bullet} - \eta_R \theta^\bullet = 0 \tag{37}$$

for all admissible thermokinematical processes.

If we subtract from this equation an  $\alpha$ -multiple of the constraint equation (34), we get

$$\begin{aligned}
 \left( \frac{1}{\rho_0} \frac{1}{2} \mathbf{S}_R - \alpha \mathbf{J} \right) \cdot \mathbf{C}^\bullet + \left( \frac{1}{\rho_0} \mathbf{H}_R - \alpha \mathbf{J} \right) \cdot \mathbf{K}^\bullet - \left( \frac{1}{\theta \rho_0} \mathbf{q}_{0R} + \alpha \mathbf{j} \right) \cdot \mathbf{g}_0 \\
 - \psi_{R^\bullet} - (\eta_R + \alpha j) \theta^\bullet = 0
 \end{aligned} \tag{38}$$

for any real  $\alpha$ . Because of the independence of the terms in brackets of  $\mathbf{C}^\bullet$ ,  $\mathbf{K}^\bullet$ ,  $\mathbf{g}_0$  and  $\theta^\bullet$  in a particular material point, this is solved for all constrained materials

only by

$$\begin{aligned} \mathbf{H}_R &= \alpha \rho_0 \mathbf{J}(\mathbf{C}, \mathbf{K}, \theta), & \psi_R^\bullet &= 0, \\ \mathbf{S}_R &= 2\alpha \rho_0 \mathbf{J}(\mathbf{C}, \mathbf{K}, \theta), & \eta_R &= -\alpha j(\mathbf{C}, \mathbf{K}, \theta), \\ \mathbf{q}_{0R} &= -\alpha \rho_0 \theta \mathbf{J}(\mathbf{C}, \mathbf{K}, \theta), \end{aligned} \quad (39)$$

or, for the spatial versions of the reactive parts,

$$\begin{aligned} \mathbf{G}_R &= \alpha \rho \mathbf{F} \circ \mathbf{J}(\mathbf{C}, \mathbf{K}, \theta) & \text{with} & \quad \mathbf{G}_R = \mathbf{F} \circ J^{-1} \mathbf{H}_R, \\ \mathbf{T}_R &= 2\alpha \rho \mathbf{F} * \mathbf{J}(\mathbf{C}, \mathbf{K}, \theta) & \text{with} & \quad \mathbf{T}_R = \mathbf{F} * J^{-1} \mathbf{S}_R, \\ \mathbf{q}_R &= -\alpha \rho \theta \mathbf{F} * \mathbf{j}(\mathbf{C}, \mathbf{K}, \theta) & \text{with} & \quad \mathbf{q}_R = \mathbf{F} * J^{-1} \mathbf{q}_{0R}. \end{aligned} \quad (40)$$

With this form, for no real  $\alpha$  can a contradiction to the Clausius–Duhem inequality occur if the extra terms already fulfill it alone.

As a normalization of the decomposition, one can pose the orthogonality condition

$$\mathbf{H}_R \cdot \mathbf{H}_E + \frac{1}{4} \mathbf{S}_R \cdot \mathbf{S}_E + \mathbf{q}_{0R} \cdot \frac{\mathbf{q}_{0E}}{\theta^2} + \rho_0^2 \eta_R \cdot \eta_E = 0. \quad (41)$$

This is, however, not compulsory and perhaps not even practical. As the free energy is only determined up to a constant, we can principally assume  $\psi_R = 0$ .

If more than one constraint is active, then the reactive parts are simply additive superpositions of those resulting from each constraint alone.

## 6. Critical remarks

There have been some papers considering the approach of [Truesdell and Noll 1965], like [Bertram 1980; 1982; Podio Guidugli 1990; Antman and Marlow 1991; Carlson and Tortorelli 1996; Carlson et al. 2003], and in their majority they confirm it in the sequel. However, there is one point which needs more attention: the limit of rigidity.

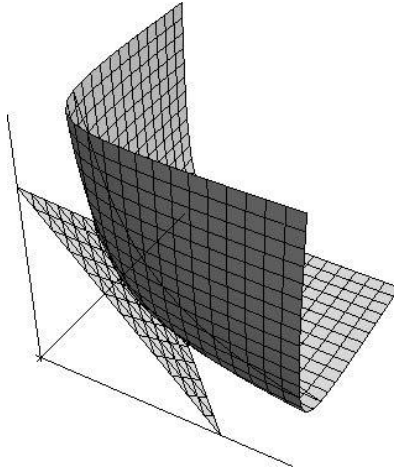
Truesdell and Noll [1965] claim that rigidity is described by  $\mathbf{C} \equiv \mathbf{I}$ . Since the space of all symmetric tensors is 6-dimensional, rigidity would correspond to 6 independent constraints and the total stress would be completely reactive in this case, since

$$\mathbf{S}_R = \sum_{i=1}^6 \alpha_i \partial_{\mathbf{C}} \gamma_{\text{red } i}(\mathbf{C}) \quad (42)$$

spans the whole space of symmetric second-order tensors. We now consider the case of two simultaneous internal constraints:

(1) (incompressibility)

$$\gamma_1(\mathbf{C}) = \det \mathbf{C} - 1 \implies \mathbf{S}_{R1} = \alpha_1 \mathbf{C}^{-1} \quad (\text{a hydrostatic pressure}) \quad (43)$$



**Figure 1.** Graphical representation of the two constraint manifolds (43) and (44) in the space of the eigenvalues of  $\mathbf{C}$ .

(2) (Bell-type [1973; 1985; 1996] constraint)<sup>3</sup>

$$\gamma_2(\mathbf{C}) = \text{tr } \mathbf{C} - 3 \implies \mathbf{S}_{R2} = \alpha_2 \mathbf{I}. \quad (44)$$

It is not our intention here to comment on the physical significance of this constraint.<sup>4</sup> Beatty and Hayes [1992] have investigated the geometrical properties of these constraints.<sup>5</sup> The constraint manifold of  $\gamma_2$  forms a plane triangle in the space of the eigenvalues of  $\mathbf{C}$ , while  $\gamma_1$  forms a curved hypersurface which touches the triangle in only one point; see Figure 1. So the constraint  $\gamma_2$  has only one isochoric (or unimodular) point, namely  $\mathbf{C} \equiv \mathbf{I}$ , which describes rigidity. In this point (and only there) they have a joint tangent plane (that coincides with the plane of  $\gamma_2$ ). In this point with  $\mathbf{C} \equiv \mathbf{I}$ , both  $\mathbf{S}_{R1}$  and  $\mathbf{S}_{R2}$  become pressures and  $\mathbf{S}_E$  is a deviator. However, there is no test possible to identify the constitutive law for this deviator.

An even more absurd example is given by the constraint suggested by Krawietz (personal communication, 2015),

$$\gamma(\mathbf{C}) = (\mathbf{C} - \mathbf{I}) \cdot (\mathbf{C} - \mathbf{I}) = 0 \implies \mathbf{S}_R = \mathbf{0}, \quad (45)$$

which also describes rigidity. In this case, the reactive stresses are zero, since  $\partial_{\mathbf{C}} \gamma(\mathbf{C})$  vanishes here.

So the format of [Truesdell and Noll 1965] needs some specification of the class of admitted constraint equations to avoid such unphysical results. One has to make sure that they really define a constraint manifold (which is not the case for (45))

<sup>3</sup>Bell uses the square root of  $\mathbf{C}$ , which however does not make much difference here.

<sup>4</sup>See, e.g., [McMeeking 1982; Sellers and Douglas 1990].

<sup>5</sup>See also [Vianello 2014] for the geometry of the constraint manifold.

and, in the case of multiple constraints, that the tangent spaces of these manifolds are independent of each other.

Another remedy would be to introduce the constraints in a natural way, as we show in the next section.

## 7. Introduction of internal constraints in a natural way

In [Bertram 1980; 1982] another approach to establish a theory of internal constraints has been suggested, claiming to be *in a natural way*. Here, only solids are considered for which the stresses are (at least partly) caused by elastic deformations. The idea there is, roughly speaking, to consider constraint material behavior as a limit of hyperelastic behavior with increasing stiffness for certain deformation modes. If one starts with hyperelastic behavior, one can consider a tangential stiffness tensor with 6 (not necessary different) eigenvalues, called principal stiffnesses in the case of classical (nongradient) materials. If one produces a series of such materials by incrementing one of these eigenvalues to infinity and keeping all others finite, one produces in the limit a material behavior that is constrained in such a way that the deformation mode belonging to this eigenvalue tends to zero if only finite stresses are applied. It has been shown there that for an isotropic or anisotropic hyperelastic material, this construction exactly leads to Assumptions 1a and 2a.

The method to produce internal constraints in a natural way can also be applied to gradient hyperelastic materials. We can linearize the elastic laws (11) and (12) by taking their incremental forms

$$d\mathbf{S} = 2\rho_0(\partial_{\mathbf{C}\mathbf{C}}w(\mathbf{C}, \mathbf{K})[d\mathbf{C}] + \partial_{\mathbf{C}\mathbf{K}}w(\mathbf{C}, \mathbf{K})[d\mathbf{K}]), \quad (46)$$

$$d\mathbf{G} = \rho_0(\partial_{\mathbf{K}\mathbf{C}}w(\mathbf{C}, \mathbf{K})[d\mathbf{C}] + \partial_{\mathbf{K}\mathbf{K}}w(\mathbf{C}, \mathbf{K})[d\mathbf{K}]), \quad (47)$$

with a

- fourth-order symmetric stiffness tensor (tetradic)  $\mathbf{C}^{(4)} := 2\rho_0 \partial_{\mathbf{C}\mathbf{C}}w(\mathbf{C}, \mathbf{K})$ ,
- sixth-order symmetric stiffness tensor (hexadic)  $\mathbf{C}^{(6)} := \rho_0 \partial_{\mathbf{K}\mathbf{K}}w(\mathbf{C}, \mathbf{K})$ ,
- fifth-order stiffness tensor  $\mathbf{C}^{(5)} := 2\rho_0 \partial_{\mathbf{K}\mathbf{C}}w(\mathbf{C}, \mathbf{K})$ .

For the linear gradient theory, see [Bertram and Forest 2014]. If we would restrict our attention to central symmetric behavior, the fifth-order stiffness vanishes and the hexadic is known from [Mindlin and Eshel 1968] while the tetradic is the usual one from classical elasticity.

Interesting for us is the stiffness hexadic  $\mathbf{C}^{(6)}$ , since it does not exist for classical materials. We can bring the hexadic into a spectral form,

$$\mathbf{C}^{(6)} = \sum_{i=1}^{18} \lambda_i \mathbf{P}_i^{(6)}, \quad (48)$$

with 18 (not necessarily distinct) eigenvalues  $\gamma_i$  and the same number of eigenspace projectors of sixth-order,  $\mathbf{P}_i^{(6)}$ . These are related to the third-order normalized and orthogonal eigentensors  $\mathbf{E}_i$  of the stiffness hexadic by the sum

$$\mathbf{P}_i^{(6)} = \sum_{j=1}^{M_i} \mathbf{E}_j \otimes \mathbf{E}_j \quad (49)$$

over the multiplicity  $M_i$  of the particular eigenvalue. The construction of internal constraints in a natural way consists of taking finite values for all of these eigenvalues except for one, say  $\lambda_1$ .

Let us first consider an eigenvalue of multiplicity one. In the limit, one would not be able to deform the material in the corresponding mode by applying finite stresses. Thus, we obtain the constraint equation (in this case independent of  $\mathbf{C}$ )

$$\Gamma(\mathbf{K}) := \mathbf{E}_1 \cdot \mathbf{K} = 0 \quad (50)$$

and expect the reaction hyperstresses after (30) to be

$$\mathbf{H}_R = \alpha \partial_{\mathbf{K}} \Gamma(\mathbf{K}) = \alpha \mathbf{E}_1 \quad (51)$$

with some scalar field  $\alpha$ .

Such a constraint would be considered as isotropic if it were invariant under arbitrary rotations

$$\Gamma(\mathbf{K}) = \Gamma(\mathbf{Q} * \mathbf{K}) \quad (52)$$

for all orthogonal tensors  $\mathbf{Q}$ . Clearly, this is the case if and only if

$$\mathbf{E}_1 = \mathbf{Q} * \mathbf{E}_1, \quad (53)$$

i.e., for isotropic tensors. But this would be a rather drastic restriction, which should not be made in general.

We can also superimpose the  $M$  constraints of a multiple eigenvalue in one equation:

$$\Gamma(\mathbf{K}) := \mathbf{K} \cdots \mathbf{P}_1^{(6)}[\mathbf{K}] = 0. \quad (54)$$

Examples and numerical computations of gradient materials with internal constraints will be given in a forthcoming paper by the same authors.

An alternative approach to create internal constraints was suggested by [Casey 1995; Baesu and Casey 2000] in a mechanical setting and [Casey and Krishnaswamy 1998; Casey 2011] in a thermomechanical setting, where the constrained material is identified as an equivalence class of unconstrained ones.

## 8. Example

We consider the constraint

$$\text{Grad } J = \mathbf{0}. \quad (55)$$

Here, the density of the body can be altered, but only in a homogeneous way in the three spatial directions. One may consider this as a vectorial internal constraint. This may as well be equivalently expressed by a scalar constraint equation by demanding the norm of the vector  $\text{Grad } J$  be zero.

After the chain rule, we can write  $\text{Grad } J$  in terms of  $\mathbf{K}$  as

$$\text{Grad } J = J K_{ijk} \mathbf{e}_k \quad (56)$$

with respect to some orthogonal vector basis  $\{\mathbf{e}_k\}$ . The constraint (55) can be reformulated as

$$\text{Grad } J \cdot \text{Grad } J / J^2 = K_{ikk} K_{llk} = \mathbf{K} \cdot \cdot \cdot \mathbf{P}^{(6)}[\mathbf{K}] = 0 \quad (57)$$

with

$$\mathbf{P}^{(6)} = \mathbf{I} \otimes \mathbf{e}_j \otimes \mathbf{I} \otimes \mathbf{e}_j. \quad (58)$$

By taking into account the right subsymmetry of the triadic  $\mathbf{K}$ , we can impose the index symmetries between the index pairs  $j, k$  and  $m, n$  in  $P_{ijklmn}$ .

The natural way to introduce this constraint is to add the projector with a penalty parameter to the elasticity hexadic  $\mathbf{C}^{(6)}$ , or likewise to add a penalized term  $K_{ikk} K_{llk}$  to the elastic energy. The latter corresponds to the parameter  $a_2$  in Mindlin and Eshel's representation [1968, Equation 2.4].

**Note added in proof.** One can find more results on gradient materials in the *Compendium on gradient materials* available at [http://www.ifme.ovgu.de/ifme\\_media/CompendiumGradientMaterialsJan2016.pdf](http://www.ifme.ovgu.de/ifme_media/CompendiumGradientMaterialsJan2016.pdf).

## Acknowledgment

The authors want to thank James Casey and Arnold Krawietz for helpful comments on our manuscript.

## References

- [Andreussi and Podio Guidugli 1973] F. Andreussi and P. Podio Guidugli, "Thermomechanical constraints in simple materials", *Bull. Acad. Pol. Sci. Tech.* **21** (1973), 199–205.
- [Antman and Marlow 1991] S. S. Antman and R. S. Marlow, "Material constraints, Lagrange multipliers, and compatibility. Applications to rod and shell theories", *Arch. Ration. Mech. Anal.* **116**:3 (1991), 257–299.
- [Baesu and Casey 2000] E. Baesu and J. Casey, "A treatment of internally constrained elastic-plastic materials", *Int. J. Eng. Sci.* **38**:15 (2000), 1677–1698.

- [Beatty and Hayes 1992] M. F. Beatty and M. A. Hayes, “Deformations of an elastic, internally constrained material, I: Homogeneous deformations”, *J. Elasticity* **29**:1 (1992), 1–84.
- [Bell 1973] J. F. Bell, *The experimental foundations of solid mechanics*, Handbuch der Physik **VIa**/1, Springer, Berlin, 1973. Reprinted in *Mechanics of Solids I*, Springer, Berlin, 1984.
- [Bell 1985] J. F. Bell, “Contemporary perspectives in finite strain plasticity”, *Int. J. Plast.* **1**:1 (1985), 3–27.
- [Bell 1996] J. F. Bell, “The decrease of volume during loading to finite plastic strain”, *Meccanica (Milano)* **31**:5 (1996), 461–472.
- [Bertram 1980] A. Bertram, “An introduction of internal constraints in a natural way”, *Z. Angew. Math. Mech.* **60**:6 (1980), T100–T101.
- [Bertram 1982] A. Bertram, “Material systems—a framework for the description of material behavior”, *Arch. Ration. Mech. Anal.* **80**:2 (1982), 99–133.
- [Bertram 2005] A. Bertram, *Elasticity and plasticity of large deformations: an introduction*, Springer, Heidelberg, 2005.
- [Bertram 2013] A. Bertram, “The mechanics and thermodynamics of finite gradient elasticity and plasticity”, preprint, Otto-von-Guericke-Universität Magdeburg, 2013, available at [http://ovgu.de/ifme/l-festigkeit/pdf/1/Preprint\\_Gradientenplastik\\_finite\\_16.10.12.pdf](http://ovgu.de/ifme/l-festigkeit/pdf/1/Preprint_Gradientenplastik_finite_16.10.12.pdf).
- [Bertram 2015a] A. Bertram, “Finite gradient elasticity and plasticity: a constitutive mechanical framework”, *Contin. Mech. Therm.* **27**:6 (2015), 1039–1058.
- [Bertram 2015b] A. Bertram, “Finite gradient elasticity and plasticity: a constitutive thermodynamical framework”, *Contin. Mech. Therm.* (online publication March 2015).
- [Bertram and Forest 2014] A. Bertram and S. Forest, “The thermodynamics of gradient elastoplasticity”, *Contin. Mech. Therm.* **26**:3 (2014), 269–286.
- [Bertram and Haupt 1976] A. Bertram and P. Haupt, “A note on Andreussi–Guidugli’s theory of thermomechanical constraints in simple materials”, *Bull. Acad. Pol. Sci. Tech.* **24**:1 (1976), 47–51.
- [Bertram and Svendsen 2001] A. Bertram and B. Svendsen, “On material objectivity and reduced constitutive equations”, *Arch. Mech. Stos.* **53**:6 (2001), 653–675.
- [Carlson and Tortorelli 1996] D. E. Carlson and D. A. Tortorelli, “On hyperelasticity with internal constraints”, *J. Elasticity* **42**:1 (1996), 91–98.
- [Carlson et al. 2003] D. E. Carlson, E. Fried, and D. A. Tortorelli, “Geometrically-based consequences of internal constraints”, *J. Elasticity* **70**:1-3 (2003), 101–109.
- [Casey 1995] J. Casey, “Treatment of internally constrained materials”, *J. Appl. Mech. (ASME)* **62**:2 (1995), 542–544.
- [Casey 2004] J. Casey, “Pseudo-rigid continua: basic theory and a geometrical derivation of Lagrange’s equations”, *Proc. R. Soc. Lond. A* **460**:2047 (2004), 2021–2049.
- [Casey 2006] J. Casey, “The ideal pseudo-rigid continuum”, *Proc. R. Soc. Lond. A* **462**:2074 (2006), 3185–3195.
- [Casey 2007] J. Casey, “A new definition of a pseudo-rigid continuum”, *Note Mat.* **27**:2 (2007), 43–53.
- [Casey 2011] J. Casey, “Nonlinear thermoelastic materials with viscosity, and subject to internal constraints: a classical continuum thermodynamics approach”, *J. Elasticity* **104**:1–2 (2011), 91–104.
- [Casey and Krishnaswamy 1998] J. Casey and S. Krishnaswamy, “A characterization of internally constrained thermoelastic materials”, *Math. Mech. Solids* **3**:1 (1998), 71–89.



- [Cohen 1981] H. Cohen, “Pseudorigid bodies”, *Utilitas Math.* **20** (1981), 221–247.
- [Cohen and MacSithigh 1989] H. Cohen and G. P. MacSithigh, “Plane motions of elastic pseudo-rigid bodies”, *J. Elasticity* **21**:2 (1989), 193–226.
- [Cohen and Muncaster 1984] H. Cohen and R. G. Muncaster, “The dynamics of pseudorigid bodies: general structure and exact solutions”, *J. Elasticity* **14**:2 (1984), 127–154.
- [Ericksen 1955] J. L. Ericksen, “Deformations possible in every compressible, isotropic, perfectly elastic material”, *J. Math. Phys. (MIT)* **34** (1955), 126–128.
- [Forest and Sievert 2003] S. Forest and R. Sievert, “Elastoviscoplastic constitutive frameworks for generalized continua”, *Acta Mech.* **160**:1-2 (2003), 71–111.
- [Green et al. 1970] A. E. Green, P. M. Naghdi, and J. A. Trapp, “Thermodynamics of a continuum with internal constraints”, *Int. J. Eng. Sci.* **8**:11 (1970), 891–908.
- [Gurtin and Podio Guidugli 1973] M. E. Gurtin and P. Podio Guidugli, “The thermodynamics of constrained materials”, *Arch. Ration. Mech. Anal.* **51** (1973), 192–208.
- [McMeeking 1982] R. M. McMeeking, “Finite strain tension torsion test of a thin-walled tube of elastic-plastic material”, *Int. J. Solids Struct.* **18**:3 (1982), 199–204.
- [Mindlin and Eshel 1968] R. D. Mindlin and N. N. Eshel, “On first strain-gradient theories in linear elasticity”, *Int. J. Solids Struct.* **4**:1 (1968), 109–124.
- [Podio Guidugli 1990] P. Podio Guidugli, “Constrained elasticity”, *Rend. Accad. Naz. Lincei* **1**:4 (1990), 341–350.
- [Sellers and Douglas 1990] H. S. Sellers and A. S. Douglas, “Physical theory of finite plasticity from a theoretical perspective”, *Int. J. Plast.* **6**:3 (1990), 329–351.
- [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), Article ID #012018.
- [Sławianowski 1974] J. J. Sławianowski, “Analytical mechanics of finite homogeneous strains”, *Arch. Mech. Stos.* **26** (1974), 569–587.
- [Sławianowski 1975] J. J. Sławianowski, “Newtonian dynamics of homogeneous strains”, *Arch. Mech. Stos.* **27**:1 (1975), 93–102.
- [Steigmann 2006] D. J. Steigmann, “On pseudo-rigid bodies”, *Proc. R. Soc. Lond. A* **462**:2066 (2006), 559–565.
- [Trapp 1971] J. A. Trapp, “Reinforced materials with thermo-mechanical constraints”, *Int. J. Eng. Sci.* **9**:8 (1971), 757–773.
- [Truesdell and Noll 1965] C. A. Truesdell and W. Noll, *The non-linear field theories of mechanics*, edited by S. Flügge, Handbuch der Physik **III**/3, Springer, Berlin, 1965. 2nd ed. published in 1992, 3rd ed. edited by S. Antman published in 2004.
- [Vianello 2014] M. Vianello, “Internal constraints in finite elasticity: manifolds or not”, *J. Elasticity* **114**:2 (2014), 197–211.

Received 16 Jul 2015. Revised 26 Sep 2015. Accepted 8 Nov 2015.

ALBRECHT BERTRAM: [albrecht.bertram@ovgu.de](mailto:albrecht.bertram@ovgu.de)

Otto-von-Guericke-Universität Magdeburg, Universitätsplatz 2, D-39106 Magdeburg, Germany

RAINER GLÜGE: [gluege@ovgu.de](mailto:gluege@ovgu.de)

Otto-von-Guericke-Universität Magdeburg, University of Magdeburg, Universitätsplatz 2, D-39106 Magdeburg, Germany







# UNIFIED GEOMETRIC FORMULATION OF MATERIAL UNIFORMITY AND EVOLUTION

MARCELO EPSTEIN AND MANUEL DE LEÓN

*Dedicated to Gérard Maugin, in gratitude for many years of inspiration and friendship*

The differential-geometric underpinnings of a unified theory of material uniformity and evolution are exposed in terms of the language of groupoids subordinate to geometric distributions. Both the standard theory of material uniformity and the extended theory of functionally graded materials are included in the formulation as well as their temporal counterparts in anelastic and aging processes.

## 1. Introduction

The delightful classical survey article by Alan Weinstein [1996] brings home the idea that groups, considered as the main carriers of information about the symmetries of a physical system, are actually not sufficient to convey the generality of the intuitive concept of symmetry. His example of the regular rectangular tiling of a bathroom floor, as opposed to the tiling of the whole plane, is very suggestive. Indeed, in passing from the infinite to the finite extent, while the symmetries of the individual tiles are preserved, the translational symmetries are lost. Nevertheless, any observer of the bathroom floor will agree that it still has a remnant of this distant kind of symmetry. In continuum mechanics, if we replace the bathroom floor with a material body  $\mathcal{B}$ , each tile can be identified with the tangent space  $T_X\mathcal{B}$  at each point  $X \in \mathcal{B}$ . Any given constitutive equation at  $X$  will enjoy some material symmetries, which are encoded in the material symmetry group  $\mathcal{G}_X$  at  $X$ . But consider now another point  $Y \in \mathcal{B}$ . What would the meaning be of a distant symmetry between  $X$  and  $Y$ ? The clearest answer to this question is that  $X$  and  $Y$  are thus related if they are made of the same material. In the terminology of Walter Noll [1967/68], the points are *materially isomorphic*. A body all of whose points are materially isomorphic is said to be *materially uniform*.

---

**Communicated by Raffaele Esposito.**

The financial support of the Natural Sciences and Engineering Research Council of Canada and of the Consejo Superior de Investigaciones Científicas of Spain is gratefully acknowledged.

*MSC2010:* 74A20.

*Keywords:* groupoids, distributions, material uniformity, material evolution, body-time manifold, groupoid actions, aging, morphogenesis.

In a philosophical sense, however, the notion of symmetry is, literally, in the eyes of the beholder. We may choose to declare the presence of a distant symmetry between two points if they are made of possibly different materials while enjoying the same *symmetry type*,<sup>1</sup> a much weaker condition. In this case, the symmetry groups  $\mathcal{G}_X$  and  $\mathcal{G}_Y$  need only be conjugate. We call the equivalence relation associated with this property *unisymmetry* [Epstein and de León 2000]. Unisymmetric bodies are known in engineering applications as *functionally graded materials*, usually made of two components with a spatially varying composition.

Fixing attention on a particular material point and following its material response in time, we obtain the notion of *material evolution* [Epstein and Elżanowski 2007]. If, as time goes on, the material response remains materially isomorphic to its initial response, we have a case of *pure remodeling* or *anelastic evolution* [Wang and Bloom 1973/74]. The material in an infinitesimal neighbourhood undergoes a process of reaccommodation, but does not experience any other essential changes. This is clearly another (timewise) manifestation of the concept of distant symmetry. Any other change of material behaviour can be considered as a process of *material aging*. A particular case is obtained when, while the material ages, the symmetry type is preserved. In other words, the material evolves unisymmetrically. Finally, any change in the symmetry type gives rise to a process of *morphogenesis* [Turing 1952].

When the spatial and temporal distant symmetries are combined, it is reasonable to expect that the corresponding geometrical descriptor will be a properly defined *material groupoid* based on the consideration of a body-time manifold, which is the main object of consideration in the present article. Some basic definitions pertaining to the theory of groupoids and their actions on sets are reviewed in Section 2 and extended to distinguish classes of groupoids associated with distributions in manifolds. These ideas are applied in Section 3 to define and interpret various cases of material groupoids.

## 2. Groupoids and distributions

**2.1. Groupoids.** Recall that a *groupoid* consists of a total set  $\mathcal{Z}$ , a base set  $\mathcal{M}$ , two (*projection*) surjective maps

$$\alpha : \mathcal{Z} \rightarrow \mathcal{M} \quad \text{and} \quad \beta : \mathcal{Z} \rightarrow \mathcal{M} \tag{1}$$

called, respectively, the *source* and the *target* maps, and a binary operation (*composition*) defined only for those ordered pairs  $(y, z) \in \mathcal{Z} \times \mathcal{Z}$  such that

$$\alpha(z) = \beta(y). \tag{2}$$

---

<sup>1</sup>By symmetry type we mean properties such as isotropy, transverse isotropy or orthotropy.

This operation, indicated by the reverse apposition of the operands, must be *associative*, that is,  $(xy)z = x(yz)$ , whenever the products are defined. Moreover, at each point  $m \in \mathcal{M}$ , there exists an *identity*  $\text{id}_m$  such that  $z \text{id}_m = z$  whenever  $\alpha(z) = m$ , and  $\text{id}_m z = z$  whenever  $\beta(z) = m$ . Finally, for each  $z \in \mathcal{Z}$  there exists a (unique) *inverse*  $z^{-1}$  such that  $zz^{-1} = \text{id}_{\beta(z)}$  and  $z^{-1}z = \text{id}_{\alpha(z)}$ .

It follows from this definition that to each ordered pair  $(a, b)$  of elements of  $\mathcal{M}$  one can associate a definite subset  $\mathcal{Z}_{ab}$  of  $\mathcal{Z}$ , namely the subset  $\mathcal{Z}_{ab} = \{z \in \mathcal{Z} \mid \beta(z) = b, \alpha(z) = a\}$ . It is clear that these subsets (some of which may be empty) are disjoint and that their union is equal to  $\mathcal{Z}$ . It is also clear that the various identities are elements of subsets of the form  $\mathcal{Z}_{bb}$ . It is not difficult to show that each set of the form  $\mathcal{Z}_{bb}$  is actually a group. A useful way to think of a groupoid is as a collection of symbols  $(a, b, c, \dots \in \mathcal{M})$  and arrows  $(x, y, z, \dots \in \mathcal{Z})$  connecting some of them.

One can prove that if  $\mathcal{Z}_{ab} \neq \emptyset$ , then the groups  $\mathcal{Z}_{aa}$  and  $\mathcal{Z}_{bb}$  are conjugate, and the conjugation between them is achieved by any element  $z$  of  $\mathcal{Z}_{ab}$ , namely,

$$\mathcal{Z}_{bb} = z \mathcal{Z}_{aa} z^{-1}. \quad (3)$$

Analogously, the set  $\mathcal{Z}_{ab}$  is spanned completely by composing any one of its elements with  $\mathcal{Z}_{aa}$  or with  $\mathcal{Z}_{bb}$ , that is,

$$\mathcal{Z}_{ab} = z \mathcal{Z}_{aa} = \mathcal{Z}_{bb} z. \quad (4)$$

A groupoid is said to be *transitive* if for each pair of points  $a, b \in \mathcal{M}$  there exists at least one element of the total set with  $a$  and  $b$  as the source and target points, respectively. In other words, a groupoid is transitive if, and only if,  $\mathcal{Z}_{ab} \neq \emptyset \forall (a, b) \in \mathcal{M} \times \mathcal{M}$ . In a transitive groupoid all the local groups  $\mathcal{Z}_{bb}$  are mutually conjugate. In this case, we can consider any of the local groups as the *typical group* of the transitive groupoid.

A groupoid is a *topological groupoid* if the total set  $\mathcal{Z}$  and the base set  $\mathcal{M}$  are topological manifolds and the projections  $\alpha$  and  $\beta$  are continuous, as are the operations of composition and of inverse. It follows from the definition that each of the sets  $\mathcal{Z}_{bb}$  is a topological group. If  $\mathcal{Z}$  and  $\mathcal{M}$  are smooth manifolds and if both projections are surjective submersions and all operations are smooth, we obtain a *Lie groupoid*.<sup>2</sup>

**2.2. Groupoids subordinate to a distribution.** Let  $\mathcal{C}$  be an  $n$ -dimensional manifold and let  $0 < k < n$  be an integer. A  $k$ -dimensional *distribution*  $\mathcal{D}$  in  $\mathcal{C}$  is a smooth assignment of a  $k$ -dimensional subspace  $D_c$  of the tangent space  $T_c\mathcal{C}$  to each point  $c \in \mathcal{C}$ . That the assignment is smooth means that each point  $c \in \mathcal{C}$  has

<sup>2</sup>For a thorough treatment of Lie groupoids see [Mackenzie 1987] or [Mackenzie 2005].

a neighbourhood within which there exist  $k$  smooth linearly independent vector fields that span the subspaces of the distribution. In fact, a common way to specify a  $k$ -dimensional distribution is by providing  $k$  linearly independent smooth vector fields on  $\mathcal{C}$ . A vector field  $V$  belongs to the distribution if, and only if, it is a linear combination of the vector fields defining it.

A groupoid  $\alpha, \beta : \mathcal{Z} \rightarrow \mathcal{C}$  is said to be *subordinate to a distribution*  $\mathcal{D}$  in  $\mathcal{C}$  if every element  $z \in \mathcal{Z}$  is a nonsingular linear map

$$z : D_{c_1} \rightarrow D_{c_2}, \quad (5)$$

with  $c_1, c_2 \in \mathcal{C}$ .

Given a distribution  $\mathcal{D}$  in  $\mathcal{C}$ , all possible groupoids subordinate to the distribution are subsets (subgroupoids) of the transitive Lie groupoid  $\mathcal{Z}_{\mathcal{D}}$  obtained by considering *all* the possible nonsingular linear maps between *all* subspaces  $D_c$  in the distribution.

A case of particular importance arises when the base manifold  $\mathcal{C}$  is a fibre bundle  $\pi : \mathcal{C} \rightarrow \mathcal{B}$  over an  $m$ -dimensional *base manifold*  $\mathcal{B}$ , with  $m < n$ . At each point  $c \in \mathcal{C}$  the tangent space  $T_c\mathcal{C}$  has a canonically defined *vertical subspace*  $V_c$ , which can be identified with the tangent space  $T_c\mathcal{C}_{\pi(c)}$  to the fibre over  $b = \pi(c)$  (that is, the set  $\pi^{-1}(\{b\})$ ,  $b \in \mathcal{B}$ ). A vector in  $T_c\mathcal{C}$  belongs to the vertical subspace  $V_c$  (or: is *vertical*) if, and only if, its projection by  $\pi_*$  is the zero vector of  $T_{\pi(c)}\mathcal{B}$ . The existence of these vertical subspaces allows us to define the canonical  $(n - m)$ -dimensional *vertical distribution*  $\mathcal{V}$  in the bundle  $\mathcal{C}$ . Accordingly, a groupoid  $\mathcal{Z}$  subordinated to  $\mathcal{V}$  will be called a *vertical groupoid*.

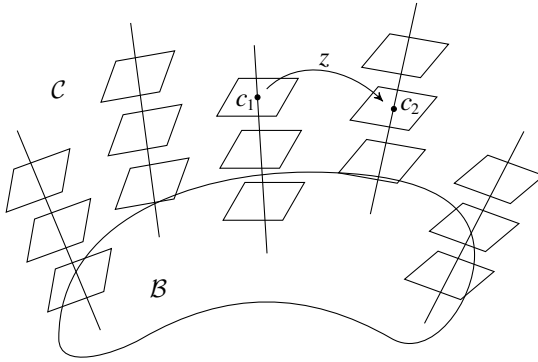
A vertical groupoid on a fibre bundle may or may not be transitive. An intermediate situation is worthy of consideration. We say that a vertical groupoid is *fibrewise transitive* if

$$\pi(c_1) = \pi(c_2) \implies \mathcal{Z}_{c_1c_2} \neq \emptyset. \quad (6)$$

A fibrewise transitive vertical groupoid is transitive if, and only if, for every pair  $a, b \in \mathcal{B}$  there exists a pair  $c_a, c_b \in \mathcal{C}$ , with  $\pi(c_a) = a$  and  $\pi(c_b) = b$ , such that  $\mathcal{Z}_{c_ac_b} \neq \emptyset$ . The truth of this assertion follows from the associative property of the groupoid composition.

**2.3. Groupoids subordinate to an Ehresmann connection.** In a fibre bundle  $\pi : \mathcal{C} \rightarrow \mathcal{B}$ , if a vector in  $T_c\mathcal{C}$  is not vertical, there is no canonical way to assign to it a vertical component. An Ehresmann connection provides this assignment.

Formally, an Ehresmann connection  $\mathcal{H}$  consists of a smooth *horizontal distribution* in  $\mathcal{C}$ . This is a smooth assignment to each point  $c \in \mathcal{C}$  of a subspace  $H_c \subset T_c\mathcal{C}$  (called the *horizontal subspace at c*), of the same dimension as the base manifold  $\mathcal{B}$ ,



**Figure 1.** An element (arrow) in a horizontal groupoid.

such that

$$T_c \mathcal{C} = H_c \oplus V_c. \quad (7)$$

In this equation,  $\oplus$  denotes the direct sum of vector spaces. Each tangent vector  $v \in T_c \mathcal{C}$  is, accordingly, uniquely decomposable as the sum of a horizontal part  $h(v)$  and a vertical part  $v(v)$ . A vector is *horizontal* if its vertical part vanishes. The only vector that is simultaneously horizontal and vertical is the zero vector.

Given a fibre bundle  $\pi : \mathcal{C} \rightarrow \mathcal{B}$  endowed with an Ehresmann connection  $\mathcal{H}$ , we define a *horizontal groupoid*  $\alpha, \beta : \mathcal{Z} \rightarrow \mathcal{C}$  as a groupoid whose elements  $z \in \mathcal{Z}$  are nonsingular linear maps

$$z : H_{c_1} \rightarrow H_{c_2}. \quad (8)$$

Notice that no a priori restriction is imposed on  $c_1, c_2 \in \mathcal{C}$ , so that, in general, we may have  $\pi(c_1) \neq \pi(c_2)$ . [Figure 1](#) illustrates this idea.

A horizontal groupoid is *fibrewise transitive* if

$$\pi(c_1) = \pi(c_2) \implies \mathcal{Z}_{c_1 c_2} \neq \emptyset. \quad (9)$$

As in the case of a vertical groupoid, a fibrewise transitive horizontal groupoid is transitive if, and only if, for every pair  $a, b \in \mathcal{B}$  there exists a pair  $c_a, c_b \in \mathcal{C}$ , with  $\pi(c_a) = a$  and  $\pi(c_b) = b$ , such that  $\mathcal{Z}_{c_a c_b} \neq \emptyset$ .

**Note.** Given an Ehresmann connection, it is possible to define a particular isomorphism between all the horizontal spaces at points lying on one and the same fibre. These *Christoffel isomorphisms*, however, may or may not belong to a given horizontal groupoid defined on the bundle.

In a product bundle  $\mathcal{C} = \mathcal{B} \times \mathcal{F}$  we can always define a canonical Ehresmann connection in an obvious way. Moreover, as dictated by convenience in particular applications, the roles of the base manifold and the typical fibre can be interchanged.

### 3. Uniformity and evolution

**3.1. *The body-time manifold.*** The concepts introduced in [Section 2](#) can, in principle, be exploited for various applications in continuum mechanics. Material bodies with internal structure, for example, are themselves modeled as fibre bundles. Classical space-time can be regarded as an affine bundle over the real line. General relativity provides an opportunity to regard a history as a world tube with identifiable world lines. For our purposes, however, which aim at a unified picture of material uniformity and evolution, we will content ourselves with a *body-time manifold* consisting of the product  $\mathcal{C} = \mathbb{R} \times \mathcal{B}$  of the time line  $\mathbb{R}$  with an ordinary body  $\mathcal{B}$ . Moreover, adopting a fixed inertial observer, we identify space-time with the product bundle  $\mathcal{S} = \mathbb{R} \times \mathbb{R}^3$ . In this simple setting, both time and space are admittedly absolute, but the essence of the unified geometric picture is not greatly altered.

A *history* of the body can be regarded as a fibre-bundle morphism

$$K : \mathcal{C} \rightarrow \mathcal{S}$$

such that the map between the base manifolds is the identity. Thus, for every instant of time  $t \in \mathbb{R}$ , our morphism provides us with a map  $\kappa_t : \mathcal{B} \rightarrow \mathbb{R}^3$ , assumed to be an embedding, whose derivative at  $X \in \mathcal{B}$  is a linear map

$$F = F(X, t) : T_X \mathcal{B} \rightarrow T_{\kappa_t(X)} \mathbb{R}^3 \cong \mathbb{R}^3,$$

called the *deformation gradient* at  $X \in \mathcal{B}$  at time  $t$ .

**3.2. *Constitutive considerations.*** The geometrical features of a material body described so far arise from the very nature of the body and of the physical space as continuous entities as well as from the *kinematic* manifestations of the former within the latter. An important feature of continuum mechanics is that the *constitutive* aspects of the material medium result in additional geometric structures. This *material geometry* arises essentially from a comparison of the material responses at different body points and at different instants of time. In view of the generality of the notion of groupoid and its ability to embrace both local and distant symmetries under the umbrella of a single mathematical entity, it is not unreasonable to expect that the material geometry alluded to above can be completely encapsulated within the compass of a single *material groupoid*.

We have just shown how any history of the body gives rise, by differentiation, to a collection of linear maps  $F$  between the vertical subspaces of the body manifold and  $\mathbb{R}^3$ . In a *simple* or *first-grade* body, we assume that the constitutive response at each body point and at each instant of time can be completely encoded in one or more functions of  $F$ . These functions may be scalar-valued (such as the energy density of a solid) or take values in a space of tensors (such as the stress). Without much loss of generality, we will assume the case of a single scalar-valued



function  $\psi = \psi_c(F)$  for each point  $c = (t, X) \in \mathcal{C}$ .<sup>3</sup> We will assume all admissible constitutive functions  $\psi_c(F)$  belong to a prespecified function space  $\Psi$ , such as  $\Psi = C^\infty(\text{GL}(3; \mathbb{R}))$ .

A useful way to look at the constitutive response of a specific body-time manifold consists of considering the Cartesian product  $\mathcal{C}_\Psi = \mathcal{C} \times \Psi$ . The constitutive response is then a *section*  $\sigma$  of  $\mathcal{C}_\Psi$ , namely a map

$$\sigma : \mathcal{C} \rightarrow \mathcal{C}_\Psi, \quad (10)$$

such that  $\text{pr}_1 \circ \sigma = \text{id}_\mathcal{C}$ , where  $\text{pr}_1$  denotes the first projection map in a product. In practice, this cross section is determined after choosing a particular *reference configuration*, that is, an identification of  $\mathcal{B}$  with a domain in  $\mathbb{R}^3$ . We will soon see the influence of this choice on the resulting geometric entities. Before proceeding, however, we need to review the notion of *action of a groupoid on a set*.

### 3.3. Action of a groupoid on a set.

**3.3.1. Group actions.** For the sake of clarity, we will first review the idea of right (or left) action of a *group* on a set, as it is widely used in physical applications. Since a groupoid is, in some sense, a generalization of a group, it should not be surprising that the extended idea of groupoid action on a set can be conceived. This extension, however, is far from trivial.

If  $\mathcal{G}$  is a group and  $A$  is a set, we say that  $\mathcal{G}$  *acts on the right* on  $A$  if for each  $g \in \mathcal{G}$  there is a map  $R_g : A \rightarrow A$  such that (i)  $R_e(a) = a$  for all  $a \in A$ , where  $e$  is the group identity; (ii)  $R_g \circ R_h = R_{hg}$  for all  $g, h \in \mathcal{G}$ . The order of the composition is the essential difference between a right and a left action. When there is no room for confusion, we also use the notation  $ag$  for  $R_g(a)$ . With this notation, property (ii) neatly reads  $a(hg) = (ah)g$ .

It is not difficult to show that each of the maps  $R_g$  is necessarily bijective. For this reason, these maps are also called *transformations*. Moreover, the inverse transformation is obtained as  $(R_g)^{-1} = R_{g^{-1}}$ . If we select a point  $a$  in  $A$  and follow its image  $ag = R_g(a)$  as  $g$  varies within  $\mathcal{G}$ , we obtain a subset of  $A$  called the *orbit* through  $a$ , denoted by  $a\mathcal{G}$ . Orbits are disjoint subsets. The relation of “belonging to the same orbit” is an equivalence relation. The orbits themselves are the equivalence classes. An often useful concept is the *quotient set*, which is the set whose elements are the orbits of  $A$ .

The action of  $\mathcal{G}$  on  $A$  is said to be *effective* if the condition  $R_g(a) = a$  for every  $a \in A$  implies  $g = e$ . The action is *free* if  $R_g(a) = a$  for *some*  $a \in A$  implies  $g = e$ . Finally, the action is *transitive* if for every  $a, b \in A$  there exists  $g \in \mathcal{G}$  such that  $R_g(a) = b$ .

---

<sup>3</sup>The time dependence is often mediated by other variables which may obey additional evolution (constitutive) equations.

As a pertinent example of a right group action, consider the following right action of the general linear group  $\mathcal{G} = \text{GL}(3; \mathbb{R})$  on the set  $\Psi$  described above. An element of  $\Psi$  is a possible constitutive equation  $\psi = \psi(F)$  for a material point. Let  $P \in \text{GL}(3; \mathbb{R})$ . We define

$$R_P(\psi) = \psi(FP). \quad (11)$$

Thus, the right action assigns to each constitutive equation another constitutive equation that, in the parlance of continuum mechanics, differs from the original one by the adoption of a different local reference. The orbit  $\psi\mathcal{G}$ , therefore, represents all the constitutive equations related in this way. From the physical standpoint, every orbit represents a different material, while the points in one orbit are different manifestations of the same constitutive law in different reference configurations. The *symmetry group*  $\mathcal{G}_\psi$  of a constitutive equation  $\psi$  is defined as the largest subgroup of  $\mathcal{G}$  that leaves  $\psi$  invariant. In other words,

$$G \in \mathcal{G}_\psi \iff R_G(\psi) = \psi. \quad (12)$$

**3.3.2. Groupoid actions.** A (*left*) action of a groupoid  $\alpha, \beta : \mathcal{Z} \rightarrow \mathcal{B}$  on a set  $A$  consists of two maps. The first map,

$$\rho : A \rightarrow \mathcal{B}, \quad (13)$$

is known as an *anchor map*, assumed to be surjective. To introduce the second map, representing the action itself, we consider first the subset  $\mathcal{Z} * A$  of the Cartesian product  $\mathcal{Z} \times A$  defined as

$$\mathcal{Z} * A = \{(z, a) \in \mathcal{Z} \times A \mid \rho(a) = \alpha(z)\}. \quad (14)$$

An *action map*  $U$ , given by

$$U : \mathcal{Z} * A \rightarrow A, \quad (15)$$

must satisfy the following (rather expected) conditions:

(1) Consistency:

$$\rho(U(z, a)) = \beta(z) \quad \forall (z, a) \in \mathcal{Z} * A. \quad (16)$$

(2) Composition:

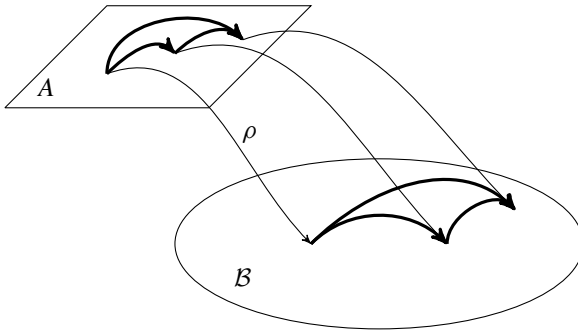
$$U(yz, a) = U(y, U(z, a)) \quad (17)$$

whenever the operations are defined.

(3) Unit:

$$U(\text{id}_{\rho(a)}, a) = a \quad \forall a \in A. \quad (18)$$

These properties are schematically represented in [Figure 2](#).



**Figure 2.** Left action of a groupoid  $\alpha, \beta : \mathcal{Z} \rightarrow \mathcal{B}$  on a set  $A$ .

### 3.4. The body-time material groupoids.

**3.4.1. Action of  $\mathcal{Z}_{\mathcal{V}}$  on  $\mathcal{C}_{\Psi}$ .** In Section 2.2 we introduced the canonical groupoid  $\mathcal{Z}_{\mathcal{D}}$  subordinate to a distribution  $\mathcal{D}$  defined as the collection of all possible nonsingular linear maps between all the subspaces in the distribution. We apply this notion to the case in which the manifold  $\mathcal{C}$  is the body-time manifold and the distribution is identified with the vertical distribution  $\mathcal{V}$ . The groupoid  $\mathcal{Z}_{\mathcal{V}}$  has a natural action on the set  $\mathcal{C}_{\Psi}$  defined in Section 3.2. The anchor map is simply the trivial bundle projection in  $\mathcal{C}_{\Psi}$ . The action  $U(z, (c, \psi))$  of an element  $z \in \mathcal{Z}_{\mathcal{V}}$  on the pair  $(c, \psi) \in \mathcal{C}_{\Psi}$  such that  $\alpha(z) = c$  is defined as

$$U(z, (c, \psi)) = (\beta(z), R_{P(z)}(\psi)), \quad (19)$$

where  $R$  is the right group action defined in (11) and  $P(z)$  is the matrix associated with the element  $z$  in the vertical groupoid  $\mathcal{Z}_{\mathcal{V}}$ . Note the apparent disagreement between the *right* character of the group action and the *left* action of the groupoid.

**3.4.2. The material groupoid.** A particular material response, as we have seen, is a particular cross section  $\sigma$  of the product bundle  $\mathcal{C}_{\Psi}$ . Consider a subgroupoid  $\mathcal{W}$  of  $\mathcal{Z}_{\mathcal{V}}$ . We can certainly restrict the action  $U$  to  $\mathcal{W}$ . The constitutive section  $\sigma$  may or may not be invariant under the action  $U$  of  $\mathcal{W}$ . We define the *material groupoid* of a given body-time manifold with constitutive response  $\sigma$  as the largest subgroupoid  $\mathcal{W}$  of  $\mathcal{Z}_{\mathcal{V}}$  that leaves  $\sigma$  invariant under the action  $U$ . This definition makes sense because, on the one hand, the disjoint subgroupoid of the unit maps is always available (so that there is no danger that the material groupoid will be empty) and, on the other hand, given two subgroupoids with the desired property, the subgroupoid generated by their union also enjoys that property.

A graphical way to visualize the material groupoid consists of drawing an arrow between each pair of points  $c_1$  and  $c_2$  in the body-time manifold  $\mathcal{C}$  for every material isomorphism  $P$  between  $c_1$  and  $c_2$ . The collection of all arrows thus drawn constitutes the material groupoid. The groupoid is transitive if all point pairs are

arrowwise connected. The typical group of a transitive material groupoid is the symmetry group of any one of the body points.

**3.4.3. The material-type groupoid.** Within the set  $\Psi$  we can introduce an equivalence relation  $\sim$  as follows.

$$\psi_1 \sim \psi_2 \iff \exists H \in \text{GL}(3; \mathbb{R}) \mid \mathcal{G}_{\psi_1} = H \mathcal{G}_{\psi_2} H^{-1}. \quad (20)$$

In other words, two constitutive equations are  $\sim$ -equivalent if, and only if, their respective symmetry groups are mutually conjugate. Define the quotient set

$$\Psi^\sim = \Psi / \sim. \quad (21)$$

The elements of the quotient set are equivalence classes  $\psi^\sim$ . Physically, two constitutive equations are in the same equivalence class if they represent the same type of material (e.g., isotropic, transversely isotropic, orthotropic).

The groupoid  $\mathcal{Z}_\gamma$  acts on  $\mathcal{C} \times \Psi^\sim$  in an obvious way. Moreover, the constitutive section  $\sigma$  induces uniquely a section  $\sigma^\sim : \mathcal{C} \rightarrow \mathcal{C} \times \Psi^\sim$  via the map  $\sim : \psi \rightarrow \psi^\sim$ . We define the *material-type groupoid* as the largest subgroupoid of  $\mathcal{Z}_\gamma$  that leaves the section  $\sigma^\sim$  invariant.

As we have suggested for the case of the material groupoid, we may visualize the material-type groupoid by drawing an arrow between each pair of points  $c_1$  and  $c_2$  in the body-time manifold  $\mathcal{C}$  for every conjugation  $H$  between the symmetry groups  $\mathcal{G}_{c_1}$  and  $\mathcal{G}_{c_2}$ . The collection of all arrows thus drawn constitutes the material-type groupoid. Clearly, the material groupoid is a subgroupoid of the material-type groupoid, since materially isomorphic points have conjugate symmetry groups. The typical group of a transitive material-type groupoid is given by the *normalizer* of the symmetry group of any of the body points.

### 3.5. Physical interpretation.

**3.5.1. The body-time material groupoid.** The terminology used to describe various kinds of distant symmetries, relevant to, among other areas, the theories of continuous distributions of defects and the theories of biological growth and remodeling, is not completely standardized. We thus start by fixing a terminological scheme to interpret the differential-geometric picture in physically meaningful terms.

In the introduction, we referred to the concepts of material isomorphism and uniformity as they pertain to the purely spatial (as opposed to temporal) component of the body-time description. Two points  $a$  and  $b$  of a material body  $\mathcal{B}$  are *materially isomorphic* if there exists a nonsingular linear map  $P : T_a\mathcal{B} \rightarrow T_b\mathcal{B}$  such that their respective constitutive functions,  $\psi_a(F)$  and  $\psi_b(F)$ , are related by the identity

$$\psi_b(F) = \psi_a(FP) \quad (22)$$

for all deformation gradients  $F$ . Points  $a$  and  $b$  are, therefore, made of the same material. Within this context, a *local material symmetry*  $G$  at point  $a \in \mathcal{B}$  can be regarded as a *material endomorphism*  $G : T_a\mathcal{B} \rightarrow T_a\mathcal{B}$ . The collection of symmetries at  $a$  forms a group  $\mathcal{G}_a$ . If  $a$  and  $b$  are materially isomorphic, their respective symmetry groups are *conjugate*. More specifically, if  $P$  is a material isomorphism, then  $\mathcal{G}_b = P\mathcal{G}_aP^{-1}$ . Conversely, if  $a$  and  $b$  are materially isomorphic, the cardinality of the set  $\mathcal{P}$  of all possible material isomorphisms between  $a$  and  $b$  is the same as the cardinality of the respective (conjugate) symmetry groups. In fact,  $\mathcal{P}$  can be generated from a given  $P$  according to the formula

$$\mathcal{P} = P\mathcal{G}_a = \mathcal{G}_bP = \mathcal{G}_bP\mathcal{G}_a.$$

A body is called *materially uniform* if all of its points are materially isomorphic. Since material isomorphism is an equivalence relation, a body is materially uniform if and only if all its points are materially isomorphic to a fixed reference body point.

The temporal counterpart of uniformity is a special kind of material evolution, whereby a material point remains materially isomorphic to a reference material point with the passage of time. This special kind of material evolution is common in the realm of biological tissues, with their natural tendency to adapt to their changing environments. A classical example is Wolff's law of trabecular bone, whose trabeculae are thought to change their orientations to follow the principal directions of stress. The fact that the material remains materially isomorphic to its initial state does not preclude the possibility of *growth* and *resorption*, whereby material of the same kind is added or removed volumetrically to the material neighbourhood. But material isomorphism does preclude the transformation of the underlying material in terms of variation of its intrinsic material properties and chemical composition. In a simplified model, we may imagine a constitutive response idealized as an elastic spring with a characteristic rest length and a given stiffness constant. Material isomorphism would imply that, while the rest length may change in time, the stiffness constant must remain unchanged. We call this special type of material evolution *remodeling*. Any other kind of evolution we call *aging*. We will later identify a particular kind of aging as worthy of further attention.

If we consider a *transitive* body-time material groupoid, its physical meaning is a body that is initially materially uniform and that evolves by pure remodeling (with no aging). In particular, it remains always materially uniform. Classical plastic evolution belongs to this material class and so does the model of tissue growth pioneered in [Rodriguez et al. 1994].

Assume, on the other hand, that the material groupoid is only *fibrewise transitive*, a concept introduced in Section 2.2. The meaning of this situation is that, while the body is materially uniform at all times, there is a process of aging taking place at all points at the same pace. In constitutive terms, this will be the case if the

dependence of the constitutive equation  $\psi = \psi(F; t, X)$  is expressed explicitly in terms of the observer time  $t$ , rather than mediated by some internal variable subject to evolution conditions. In other words, the material degradation takes place by the mere passage of time, without any coupling to other observable phenomena (such as the state of stress).

Since the body-time manifold has been defined as a Cartesian product  $\mathbb{R} \times \mathcal{B}$ , we may exchange the role of the base and the typical fibre and consider instead the product manifold  $\mathcal{B} \times \mathbb{R}$ . In this product bundle, we have a natural horizontal distribution (Ehresmann connection) induced by the constant sections. The material groupoid can now be regarded, in the terminology of Section 2.3, as a *horizontal groupoid*. If this new material groupoid happens to be fibrewise transitive, we obtain the representation of a body that, without necessarily being materially uniform, evolves by pure remodeling, without aging. Notice that under these conditions if the body is initially nonuniform, it will never attain uniformity.

**3.5.2. The body-time material-type groupoid.** We want to explore now the physical meaning of the material-type groupoid. Recall that the “arrows” of this groupoid represent only conjugation maps between the material symmetry groups of the source and target points. Put differently, the elements of this groupoid are not sensitive to any constitutive property, except the symmetries of the constitutive law. If the material-type groupoid is transitive, all that this implies is that the different material points are of the same symmetry type (isotropic, say). For all we know, part of the body may be made of cement and the rest of rubber. Moreover, as time goes on, the cement may be undergoing a process of curing and change its elastic properties. A *functionally graded material* obtained by varying the relative concentration of the components of a mixture belongs to the same category, even if the components undergo chemical reactions, as long as the symmetry type is preserved.

If the material-type groupoid is only fibrewise transitive, we have a phenomenon of *morphogenesis* or *symmetry breaking*. In this instance, all the body points undergo a change of symmetry type simultaneously. In solid materials, where the collection of all possible symmetry groups is countable, this transition is, of necessity, abrupt. In some cases of phase transition these changes may be directly observable with the naked eye and be manifest as pattern formations. Reversing the role of the base and the fibre manifolds, a fibrewise transitive material-type groupoid represents a body with initially demarcated portions obeying different symmetry types and remaining so with the passage of time.

## References

- [Epstein and de León 2000] M. Epstein and M. de León, “Homogeneity without uniformity: towards a mathematical theory of functionally graded materials”, *Internat. J. Solids Structures* **37**:51 (2000), 7577–7591.

- [Epstein and Elzanowski 2007] M. Epstein and M. Elzanowski, *Material inhomogeneities and their evolution: a geometric approach*, Springer, Berlin, 2007.
- [Mackenzie 1987] K. Mackenzie, *Lie groupoids and Lie algebroids in differential geometry*, London Mathematical Society Lecture Note Series **124**, Cambridge University Press, 1987.
- [Mackenzie 2005] K. C. H. Mackenzie, *General theory of Lie groupoids and Lie algebroids*, London Mathematical Society Lecture Note Series **213**, Cambridge University Press, 2005.
- [Noll 1967/68] W. Noll, “Materially uniform simple bodies with inhomogeneities”, *Arch. Rational Mech. Anal.* **27** (1967/68), 1–32.
- [Rodriguez et al. 1994] E. K. Rodriguez, A. Hoger, and A. D. McCulloch, “Stress-dependent finite growth in soft elastic tissues”, *J. Biomech.* **27**:4 (1994), 455–467.
- [Turing 1952] A. M. Turing, “The chemical basis of morphogenesis”, *Philos. Trans. Roy. Soc. London Ser. B* **237**:641 (1952), 37–72.
- [Wang and Bloom 1973/74] C. C. Wang and F. Bloom, “Material uniformity and inhomogeneity in anelastic bodies”, *Arch. Rational Mech. Anal.* **53** (1973/74), 246–276.
- [Weinstein 1996] A. Weinstein, “Groupoids: unifying internal and external symmetry. A tour through some examples”, *Notices Amer. Math. Soc.* **43**:7 (1996), 744–752.

Received 25 Jul 2015. Revised 9 Dec 2015. Accepted 6 Feb 2016.

MARCELO EPSTEIN: [mepstein@ucalgary.ca](mailto:mepstein@ucalgary.ca)

Department of Mechanical and Manufacturing Engineering, University of Calgary,  
2500 University Drive NW, Calgary, AB T2N 1N4, Canada

MANUEL DE LEÓN: [mdeleon@icmat.es](mailto:mdeleon@icmat.es)

Instituto de Ciencias Matemáticas, CSIC-UAM-UC3M-UCM, Nicolás Cabrera, 13-15,  
28049 Madrid, Spain









# ELECTROMECHANICS OF POLARIZED LIPID BILAYERS

DAVID J. STEIGMANN AND ASHUTOSH AGRAWAL

A model for the electromechanics of lipid bilayers, accounting for flexoelectricity, is obtained as the thin-film limit of the continuum electrostatics of nematic liquid crystals. A priori restrictions on the polarization field consistent with minimum energy considerations effectively decouple the leading-order membrane problem from the computation of the self field, yielding a substantial simplification vis a vis the three-dimensional theory. Examples illustrate the strong interplay between the electric field and membrane geometry.

## 1. Introduction

The idea that lipid bilayers can be regarded as thin liquid crystal films apparently originated in the work of Helfrich [1973]. This point of view gave rise to an associated body of work that has been thoroughly documented in [Ou-Yang et al. 1999]. The liquid-crystal framework provides a clear conceptual foundation for extensions of the basic purely mechanical theory to coupled-field problems. In the present work, we use this foundation to develop an electromechanical theory of lipid bilayers. This framework may be used to gain physical insight into various phenomena. For example, lipid vesicles have been shown to deform in the presence of applied electric fields [Winterhalter and Helfrich 1988; Kummrow and Helfrich 1991; Dimova et al. 2007; 2009; Vlahovska 2010]. The creation of nanopores in lipid membranes by external electric fields is a standard technique — known as *electroporation* — to deliver genes into cells [Neumann et al. 1982; Aihara and Miyazaki 1998; Weaver 2000] and in some cancer treatments [Davalos et al. 2005; Rubinsky et al. 2007]. The role of coupled electromechanical interactions is well recognized in the context of cochlear outer hair cells [Brownell et al. 1985; Raphael et al. 2000; Harland et al. 2015]. Electromechanical interactions also play a fundamental role in electrically active cells such as neurons. Further, experimental studies have revealed that voltage-gated ion channels exhibit sensitivity to both electrostatic and mechanical forces [Schmidt et al. 2012].

---

**Communicated by Francesco dell’Isola.**

*MSC2010:* 74P10.

*Keywords:* electromechanics, lipid membranes.

Here we adapt the three-dimensional liquid crystal theory advanced in [de Gennes and Prost 1992; Ericksen 1961; 1962; 1976; Virga 1994] to derive a two-dimensional model for the response of electrically polarized lipid bilayers to applied electric fields generated by a remote source. In this respect our approach differs substantially from recent efforts directed at modeling electromechanical interactions in lipid membranes [Gao et al. 2008; Mohammadi et al. 2014]. For definiteness and for the sake of simplicity, we base our model on the general theory for nematics [Virga 1994], incorporating modifications associated with the so-called flexoelectric effect [Meyer 1969; de Gennes and Prost 1992].

In Section 2 we summarize those aspects of the basic three-dimensional theory that are required for our purpose. This is based on an extension to liquid-crystal theory of an expression for the potential energy of a polarized material, subject to a remotely generated applied electric field regarded as an assigned function of position in the ambient space [Toupin 1956; Truesdell and Toupin 1960; Bustamante et al. 2009; Dorfmann and Ogden 2014]. Insofar as electrical interactions are concerned, we confine attention in this preliminary work to the effects of polarization and assume free charges to be absent. The relevant three-dimensional energy is used, in Section 3, to derive the leading-order-in-thickness expression for the energy of the two-dimensional model. The operative equilibrium equations and edge conditions are derived from this via a variational procedure in Sections 4 and 5, respectively, and the theory is illustrated through numerical solution of several examples involving axisymmetry in Section 6. We freely use the standard notation of the classical differential geometry of surfaces. The text by Sokolnikoff [1964] is recommended for mechanicians seeking a comprehensive treatment.

## 2. Energetics of three-dimensional liquid crystals in the presence of a stationary applied field

Numerous variational formulations of electromechanical interactions in deformable media are available in the literature. These have been extensively examined and correlated in [Bustamante et al. 2009; Dorfmann and Ogden 2014], to which the interested reader is referred for fuller expositions. There it is shown that Maxwell's equations and the equilibrium equations for a polarized medium in the presence of an applied electric field that is fixed in space, in the absence of applied loads or free electric charges, render stationary the energy functional

$$\mathcal{E} = \int_R \left( U - \frac{1}{2} \mathbf{e}_s \cdot \mathbf{p} - \mathbf{e}_a \cdot \mathbf{p} \right) dv, \quad (1)$$

where  $R$  is the volume currently occupied by the material in three-space,  $U$  is the relevant energy density,  $\mathbf{p}$  is the polarization per unit volume,  $\mathbf{e}_s$  is the electric self field generated by the polarized material, and  $\mathbf{e}_a$  is the applied electric field,

assumed to be assigned as a smooth function in all of three-space, including  $R$ . The net electric field is

$$\mathbf{e} = \mathbf{e}_s + \mathbf{e}_a. \quad (2)$$

Further, the applied field is a fixed function of position  $\mathbf{y}$  in the enveloping three-space. Its variational derivative, associated with a fixed material point (a fixed lipid molecule in the present context), is thus purely convective; i.e.,

$$\dot{\mathbf{e}}_a = (\text{grad } \mathbf{e}_a) \dot{\mathbf{y}}, \quad (3)$$

where grad is the (spatial) gradient with respect to  $\mathbf{y}$ ; and, here and henceforth, superposed dots are used to denote variational — or Gateaux — derivatives. The field  $\mathbf{e}_a$  is curl-free; accordingly, its gradient is symmetric:  $\text{grad } \mathbf{e}_a = (\text{grad } \mathbf{e}_a)^t$ . The self field is also curl-free; it is obtained from

$$\mathbf{e}_s = -\text{grad } V_s, \quad (4)$$

where the self field potential  $V_s$  is given by [Kovetz 2000]

$$4\pi \varepsilon_0 V_s(\mathbf{y}) = \int_{\partial R} \frac{\mathbf{p}' \cdot \mathbf{n}'}{|\mathbf{y} - \mathbf{y}'|} da - \int_R \frac{\text{div}' \mathbf{p}'}{|\mathbf{y} - \mathbf{y}'|} dv, \quad (5)$$

in which  $\varepsilon_0$  is the free-space permittivity;  $\mathbf{n}'$  is the exterior unit normal to  $\partial R$ , expressed as a function of the integration variable  $\mathbf{y}'$ ;  $\mathbf{p}'$  is likewise the polarization in terms of  $\mathbf{y}'$ ; and  $\text{div}'$  is the divergence with respect to  $\mathbf{y}'$ . This is defined by  $\text{div } \mathbf{p} = \text{tr}(\text{grad } \mathbf{p})$ , where  $\text{tr}(\cdot)$  is the trace.

The energy density  $U$  is a function of the polarization and appropriate deformation variables. In the conventional theory of electroelasticity the relevant deformation variable is the deformation gradient, the gradient of  $\mathbf{y} = \boldsymbol{\chi}(\mathbf{x})$  with respect to position  $\mathbf{x}$  in some fixed reference configuration. Here  $\boldsymbol{\chi}(\cdot)$  is a field describing the deformation of material points. In the present application to liquid crystals, the relevant variables are a director field  $\mathbf{d}(\mathbf{y})$  — describing the orientation of the liquid crystal molecules — and its spatial gradient

$$\mathbf{D} = \text{grad } \mathbf{d}. \quad (6)$$

We follow the conventional theory and impose  $|\mathbf{d}(\mathbf{y})| = 1$ .

The electric field is given in terms of the polarization by the partial derivative [Toupin 1956; Bustamante et al. 2009]

$$\mathbf{e} = U_p(\mathbf{d}, \mathbf{D}, \mathbf{p}). \quad (7)$$

In applications  $U$  is typically assumed to be a quadratic function of  $\mathbf{D}$ . This reflects the notion that the length scale for spatial variations of the director is typically much larger than the local length scale: the molecular length; the dimensionless

gradient is then sufficiently small to justify the termination of the Taylor expansion of  $U(\mathbf{d}, \cdot, \mathbf{p})$  at second order. Thus,

$$U = l(\mathbf{d}, \mathbf{p}) + \mathbf{L}(\mathbf{d}, \mathbf{p}) \cdot \mathbf{D} + \frac{1}{2} \mathbf{D} \cdot \mathcal{L}(\mathbf{d}, \mathbf{p})[\mathbf{D}], \quad (8)$$

in which  $l$ ,  $\mathbf{L}$  and  $\mathcal{L}$  are scalar, second-order tensor and fourth-order tensor valued functions, respectively, with  $\mathcal{L} = \mathcal{L}^t$ .

Guided by [Virga 1994], we adopt the specific forms

$$l(\mathbf{d}, \mathbf{p}) = \frac{1}{2} \chi_{\perp} |\mathbf{p}|^2 + \frac{1}{2} \chi_a (\mathbf{p} \cdot \mathbf{d})^2, \quad (9)$$

where  $\chi_{\perp}$  and  $\chi_a$  are the anisotropic dielectric constants, and

$$\begin{aligned} \mathbf{D} \cdot \mathcal{L}(\mathbf{d}, \mathbf{p})[\mathbf{D}] \\ = k_1 (\operatorname{div} \mathbf{d})^2 + k_2 (\mathbf{d} \cdot \operatorname{curl} \mathbf{d})^2 + k_3 |\mathbf{D}\mathbf{d}|^2 + (k_2 + k_4) [\operatorname{tr}(\mathbf{D}^2) - (\operatorname{div} \mathbf{d})^2], \end{aligned} \quad (10)$$

in which the latter is independent of  $\mathbf{p}$ , and  $k_1 - k_4$  are constants with  $2k_1 \geq k_2 + k_4$ ,  $k_2 \geq |k_4|$  and  $k_3 \geq 0$ , in accordance with the presumed positive-definiteness of  $\mathcal{L}$  [Virga 1994]. The second expression is the Frank energy for nematic liquid crystals.

To model the flexoelectric effect, we adopt Meyer's proposal [1969] in a form similar to that used by Ou-Yang et al. [1999]. Thus,

$$\mathbf{L}(\mathbf{d}, \mathbf{p}) \cdot \mathbf{D} = -\mathbf{p} \cdot \mathbf{f}(\mathbf{d}, \mathbf{D}), \quad \text{with } \mathbf{f}(\mathbf{d}, \mathbf{D}) = c_1 (\operatorname{div} \mathbf{d}) \mathbf{d} + c_2 \operatorname{curl} \mathbf{d} \times \mathbf{d}, \quad (11)$$

where  $c_1$  and  $c_2$  are the flexoelectric constants. The relationship (7) then furnishes an expression for the electric field

$$\mathbf{e} = \chi_{\perp} \mathbf{p} + \chi_a (\mathbf{p} \cdot \mathbf{d}) \mathbf{d} - \mathbf{f}(\mathbf{d}, \mathbf{D}). \quad (12)$$

This coincides, in the specialization  $\chi_a = \chi_{\perp}$ , with Equation (2.153) in [Ou-Yang et al. 1999] in the case when the electric field vanishes.

In the absence of polarization (8) reduces precisely to the conventional liquid-crystal energy. The stationarity of this energy is equivalent, under appropriate regularity conditions, to the well-known equilibrium equations and natural boundary conditions of liquid-crystal theory [de Gennes and Prost 1992; Ericksen 1976; Virga 1994; Steigmann 2013].

Our objective in the present work is to derive the leading-order small thickness limit of the energy (1). This limit is taken to be the energy of a polarized lipid membrane. Stationarity conditions for the limit energy are then identified with the equilibrium equations of a polarized lipid membrane in the presence of an applied field generated by a remote source.

We have in mind a lipid bilayer constituting a membrane structure in a typical animal cell. Because such a membrane is only one or two molecules across, its thickness is on the order of the local length scale embodied in the constitutive

response of the liquid crystal. Accordingly, this is the only relevant length scale in the dimension reduction procedure. In contrast, in recent work on thin films of nematic elastomers [Cesana et al. 2015], the local length scale is always much smaller than the values of film thickness deemed to be relevant. The local scale therefore vanishes with thickness, leading to a reduced theory in which the director gradient ultimately plays no role. Of course it may be argued that reliance on three-dimensional liquid crystal theory to effect a dimension reduction procedure may not be justified if the thickness is comparable to molecular dimensions. In this instance our procedure nevertheless furnishes guidance for the construction of a direct two-dimensional field theory.

### 3. Liquid crystal films

In the purely mechanical theory of thin liquid crystal films the leading-order strain energy density  $W$  is associated with the limit [Steigmann 2013]

$$\lim_{t \rightarrow 0} t^{-1} \int_R U \, dv = \int_\omega W \, da, \quad (13)$$

where  $\omega$  is the interior midsurface of the film,  $t$  is the (uniform) thickness of the film, and

$$W = U|_\omega \quad (14)$$

is the leading-order energy density on  $\omega$ . This follows by using the volume measure  $dv = \mu \, d\zeta \, da$  [Naghdi 1972], where  $\zeta$  is a linear coordinate in the direction of the unit surface normal  $\mathbf{n}$ , regarded as the restriction of  $\mathbf{d}$  to  $\omega$ , and  $\mu = 1 - 2\zeta H + \zeta^2 K$ , where  $H$  and  $K$  respectively are the mean and Gaussian curvatures of  $\omega$ . In effect, then, we suppress misalignment of the lipid molecules with the surface normal—the so-called lipid tilt—as in the classical Canham–Helfrich theory. This is appropriate if the surface density of the lipids is sufficiently high. Generalizations to accommodate tilt are described in [Steigmann 2013]. We have [Steigmann 2013]

$$\mathbf{n} = \mathbf{d}|_\omega \quad \text{and} \quad \mathbf{D}|_\omega = \nabla \mathbf{n} + \boldsymbol{\eta} \otimes \mathbf{n}, \quad (15)$$

where  $\nabla(\cdot)$  is the (two-dimensional) surface gradient on  $\omega$  and  $\boldsymbol{\eta}$  is the restriction to  $\omega$  of the derivative of  $\mathbf{d}$  in the direction of  $\mathbf{d}$ . Accordingly,

$$\mathbf{n} \cdot \boldsymbol{\eta} = 0 \quad \text{on } \omega. \quad (16)$$

The extension to the case of polarizable films in the presence of an applied field is immediate. Thus,

$$\lim_{t \rightarrow 0} t^{-1} \mathcal{E} = E, \quad (17)$$

with

$$E = \int_\omega W \, da, \quad (18)$$

where  $W$  is now given by

$$W = \left( U - \frac{1}{2} \mathbf{e}_s \cdot \mathbf{p} - \mathbf{e}_a \cdot \mathbf{p} \right)_{|\omega}. \quad (19)$$

**Remark 1.** Quantum mechanical considerations and molecular dynamics simulations in respect of polarized lipid membranes [Seelig 1978; Frischleder and Peinel 1982; Warshaviak et al. 2011] indicate that the polarization vector is essentially tangential to the film surface. In this case an estimate based on (5)—derived in the mathematically identical context of magnetostatics [Barham et al. 2012]—indicates that the magnitude  $|\mathbf{e}_s|$  of the self field is of order  $O(t \ln t)$ ; for small  $t$  this is negligible compared to unity. It follows that the *leading-order* energy, i.e., the limit of  $\mathcal{E}/t$  as  $t \rightarrow 0$ , is given by (18) but with  $W$  simplified to

$$W = U_{|\omega} - \mathbf{e}_a(\mathbf{r}) \cdot \boldsymbol{\pi}, \quad (20)$$

where  $\boldsymbol{\pi} = \mathbf{p}_{|\omega}$ ,  $\mathbf{r} = \mathbf{y}_{|\omega}$  is the position field on  $\omega$ , and

$$\mathbf{n} \cdot \boldsymbol{\pi} = 0 \quad \text{on } \omega. \quad (21)$$

**Remark 2.** The estimate on the self field effectively decouples its computation from the problem of rendering  $E$  stationary, implying, in particular, that it may be evaluated a posteriori. This feature affords a major simplification of the theory for thin films vis a vis that for bulk continua. Further, in the analogous magnetostatic setting, the condition (21), with polarization replaced by magnetization, is known to furnish energetically optimal states of magnetization in thin films [Gioia and James 1997]. Thus our approach via dimension reduction provides justification for the suppression of the self field, in the *leading order* two-dimensional model, under conditions in which the polarization field is tangential to the membrane. In contrast, in [Gao et al. 2008; Mohammadi et al. 2014] no analysis is offered to justify the suppression of the self field.

The self field at points in space remote from the membrane may be evaluated a posteriori by applying the divergence theorem to (5), for points  $\mathbf{y}$  not in the closure of  $R$ . This furnishes

$$4\pi \varepsilon_0 V_s(\mathbf{y}) = \int_R \mathbf{p}' \cdot \frac{\mathbf{y} - \mathbf{y}'}{|\mathbf{y} - \mathbf{y}'|^3} dv = t \left[ \int_\omega \boldsymbol{\pi} \cdot \frac{\mathbf{y} - \mathbf{r}}{|\mathbf{y} - \mathbf{r}|^3} da + o(t)/t \right], \quad (22)$$

in which  $\mathbf{r}$  is the membrane position field. The self field then follows by computing the gradient with respect to  $\mathbf{y}$  (cf. (4)), yielding

$$4\pi \varepsilon_0 \lim_{t \rightarrow 0} (t^{-1} \mathbf{e}_s) = \int_\omega \mathbf{G} \boldsymbol{\pi} da, \quad (23)$$

where

$$\mathbf{G} = \frac{3}{|\mathbf{y} - \mathbf{r}|^5} (\mathbf{y} - \mathbf{r}) \otimes (\mathbf{y} - \mathbf{r}) - \frac{1}{|\mathbf{y} - \mathbf{r}|^3} \mathbf{I}. \quad (24)$$

The leading-order self field in space is thus delivered by a quadrature over  $\omega$  after membrane shape has been determined.

In the expression (20) we have

$$U|_{\omega} = U(\mathbf{n}, -\mathbf{b} + \boldsymbol{\eta} \otimes \mathbf{n}, \boldsymbol{\pi}), \quad (25)$$

where

$$\mathbf{b} = -\nabla \mathbf{n} \quad (26)$$

is the (symmetric) curvature tensor on  $\omega$ ; at any particular point of  $\omega$  this maps the tangent plane  $T_{\omega}$  to itself.

The explicit form used here follows from (8)–(11). For example, the restriction to  $\omega$  of the function  $f(\mathbf{d}, \mathbf{D})$  in (11) is given by

$$f|_{\omega} = c_1(\mathbf{n} \cdot \boldsymbol{\eta} - 2H)\mathbf{n} + c_2\boldsymbol{\eta}, \quad (27)$$

where

$$H = \frac{1}{2} \operatorname{tr} \mathbf{b} \quad (28)$$

is the mean curvature of  $\omega$ . This expression may be simplified by imposing (16) but we refrain from doing so for reasons to be discussed later. Here use has been made of (15) and the formula  $\operatorname{curl} \mathbf{d} \times \mathbf{d} = \mathbf{D}\mathbf{d}$ , which follows from the fact that  $\mathbf{d}$  is a field of unit vectors [Virga 1994].

To reduce (10) we first introduce a coordinate parametrization  $\mathbf{r}(\theta^{\alpha})$  of  $\omega$ . This induces the natural tangent basis  $\mathbf{a}_{\alpha} = \mathbf{r}_{,\alpha} \in T_{\omega}$  and associated dual basis  $\mathbf{a}^{\alpha} \in T_{\omega}$ , where  $(\cdot)_{,\alpha} = \partial(\cdot)/\partial\theta^{\alpha}$ . Then, the restriction of  $\operatorname{curl} \mathbf{d}$  to  $\omega$  is [Steigmann 2013]

$$(\operatorname{curl} \mathbf{d})|_{\omega} = \mathbf{a}^{\alpha} \times \mathbf{n}_{,\alpha} + \mathbf{n} \times \boldsymbol{\eta}, \quad (29)$$

where  $\mathbf{a}^{\alpha} \times \mathbf{n}_{,\alpha} = -b_{\alpha\beta} \mathbf{a}^{\alpha} \times \mathbf{a}^{\beta}$ , with  $b_{\alpha\beta} = \mathbf{a}_{\alpha} \cdot \mathbf{b}\mathbf{a}_{\beta}$ , vanishes by virtue of the symmetry of  $\mathbf{b}$ ; accordingly,  $(\mathbf{d} \cdot \operatorname{curl} \mathbf{d})|_{\omega} = 0$ .

Using (26) with  $\mathbf{b}\mathbf{n} = \mathbf{0}$  we also derive

$$\operatorname{tr}(\mathbf{D}^2)|_{\omega} = \operatorname{tr}(\mathbf{b}^2) + (\mathbf{n} \cdot \boldsymbol{\eta})^2. \quad (30)$$

Applying the Cayley–Hamilton formula

$$\mathbf{b}^2 = 2H\mathbf{b} + K\mathbf{1}, \quad (31)$$

where

$$K = \det \mathbf{b} \quad (32)$$

is the Gaussian curvature of  $\omega$  and  $\mathbf{1}$  is the identity transformation on  $T_{\omega}$ , we then obtain

$$[\operatorname{tr}(\mathbf{D}^2) - (\operatorname{div} \mathbf{d})^2]|_{\omega} = 2H\mathbf{n} \cdot \boldsymbol{\eta} - 2K, \quad (33)$$

which again may be simplified by imposing (16).

**Remark 3.** It is well known [Virga 1994] that the combination  $\text{tr}(\mathbf{D}^2) - (\text{div } \mathbf{d})^2$  is a null Lagrangian in the three-dimensional theory. It is also well known that the Gaussian curvature is a null Lagrangian in the surface theory; in particular, the total curvature of a closed surface is fixed by its genus and thus contributes only a disposable constant to the energy if the latter is invariant, as we assume in the present work. Accordingly, (16) implies that the same combination of terms also furnishes a null Lagrangian in the two-dimensional theory.

Altogether, the surface energy reduces to (cf. (8))

$$U_{|\omega} = \frac{1}{2}\chi_{\perp}|\boldsymbol{\pi}|^2 + \frac{1}{2}\chi_a(\boldsymbol{\pi} \cdot \mathbf{n})^2 - c_1(\mathbf{n} \cdot \boldsymbol{\eta} - 2H)\mathbf{n} \cdot \boldsymbol{\pi} - c_2\boldsymbol{\eta} \cdot \boldsymbol{\pi} \\ + \frac{1}{2}k_1(\mathbf{n} \cdot \boldsymbol{\eta} - 2H)^2 + \frac{1}{2}k_3|\boldsymbol{\eta}|^2 + (k_2 + k_4)(H\mathbf{n} \cdot \boldsymbol{\eta} - K), \quad (34)$$

yielding the net energy density in the form

$$W = kH^2 + \bar{k}K + \frac{1}{2}k_3|\boldsymbol{\eta}|^2 + \frac{1}{2}\chi_{\perp}|\boldsymbol{\pi}|^2 - c_2\boldsymbol{\eta} \cdot \boldsymbol{\pi} + \tilde{\varphi}\mathbf{n} \cdot \boldsymbol{\pi} + \tilde{\psi}\mathbf{n} \cdot \boldsymbol{\eta} - \mathbf{e}_a(\mathbf{r}) \cdot \boldsymbol{\pi}, \quad (35)$$

where

$$k = 2k_1, \quad \bar{k} = -(k_2 + k_4) \quad (36)$$

and  $\tilde{\varphi}, \tilde{\psi}$  are certain scalars which will prove to be irrelevant. Accordingly,  $W$  may be regarded as a function of the list

$$\{H, K, \mathbf{r}, \mathbf{n}, \boldsymbol{\eta}, \boldsymbol{\pi}\}, \quad (37)$$

subject to the constraints (16) and (21), in which it is understood that  $H, K$  and  $\mathbf{n}$  are determined by the parametrization  $\mathbf{r}(\theta^\alpha)$ . Henceforth we require the doublet  $\{\theta^\alpha\}$  to maintain a fixed correspondence with a material point; i.e., a lipid molecule. Thus the coordinates are convected with the lipids in the course of any configurational variation.

#### 4. Variational problem and equilibrium equations

It is convenient to adopt an extended variational formulation in which the constraints are relaxed. In this formulation we do not impose (16) or (21), but instead consider the auxiliary energy

$$E^* = \int_{\Omega} [JW + \lambda(J - 1) + \bar{\varphi}\mathbf{n} \cdot \boldsymbol{\pi} + \bar{\psi}\mathbf{n} \cdot \boldsymbol{\eta}] dA, \quad (38)$$

where  $W$  is given by (35), with  $\mathbf{r}, \boldsymbol{\eta}$  and  $\boldsymbol{\pi}$  regarded as independent fields, and where  $\bar{\varphi}$  and  $\bar{\psi}$  are Lagrange-multiplier fields associated with the constraints (16) and (21). Here  $\Omega$  is the preimage of  $\omega$  in a fixed reference placement, with  $da = J dA$ . In terms of the convected-coordinate surface parametrization we have  $J = \sqrt{a/A}$ , where  $a = \det(a_{\alpha\beta})$ ,  $a_{\alpha\beta} = \mathbf{a}_\alpha \cdot \mathbf{a}_\beta$  is the surface metric, and  $A$  is the value of  $a$  on  $\Omega$ . Further,  $\lambda$  is a Lagrange-multiplier field associated with the constraint that



the map from any configuration to another preserves local surface area; and, hence, that  $J = 1$ . This restriction is appropriate in the absence of lipid distension, as in the classical Canham–Helfrich theory; the bulk incompressibility of the liquid crystal then implies that area is preserved locally [Steigmann 2013]. Generalizations to account for distension in a purely mechanical setting are discussed in [Steigmann 2013; Kim and Steigmann 2015].

We observe that  $\bar{\varphi}$  and  $\tilde{\psi}$  in (35) may be absorbed into the Lagrange multipliers and conclude that no generality is lost if (35) is replaced by

$$W = U - \mathbf{e}_a(\mathbf{r}) \cdot \boldsymbol{\pi} \quad (39)$$

in (38), where  $U$  is now given by

$$U = kH^2 + \bar{k}K + \frac{1}{2}k_3|\boldsymbol{\eta}|^2 + \frac{1}{2}\chi_\perp|\boldsymbol{\pi}|^2 - c_2\boldsymbol{\eta} \cdot \boldsymbol{\pi}. \quad (40)$$

We note that the quadratic form involving  $\boldsymbol{\eta}$  and  $\boldsymbol{\pi}$  is positive definite if and only if  $k_3 > 0$ ,  $\chi_\perp > 0$  and  $c_2^2 < k_3\chi_\perp$ .

The expression (38) reduces to the actual energy when the constraints (16) and (21) are operative, and is well defined when they are not; it therefore furnishes an extension of the actual energy to arbitrary (unconstrained) states. Stationarity with respect to the multipliers simply returns the constraints as the relevant Euler–Lagrange equations. Moreover, stationarity of  $E^*$  with respect to arbitrary variations implies stationarity with respect to constrained variations in particular, and hence stationarity of the actual energy  $E$ . We use this observation to derive equilibrium equations for the actual constrained system. We note that while the replacement of  $E$  by  $E^*$  is permissible for the purpose of extracting stationarity (i.e., equilibrium) conditions, it may not be used to study energy *minimizers*. This is a consequence of the fact that  $\inf E^* \leq \inf E$ , this following trivially from constraint relaxation.

The variational derivative of the extended energy, modulo the variations of the Lagrange multipliers, is

$$\dot{E}^* = \int_\omega [\dot{W} + (W + \lambda)\dot{J}/J + \varphi(\dot{\mathbf{n}} \cdot \boldsymbol{\pi} + \mathbf{n} \cdot \dot{\boldsymbol{\pi}}) + \psi(\dot{\mathbf{n}} \cdot \boldsymbol{\eta} + \mathbf{n} \cdot \dot{\boldsymbol{\eta}})] da, \quad (41)$$

where  $\varphi = \bar{\varphi}/J$  and  $\psi = \bar{\psi}/J$ , and it is understood, having suppressed the variations of the multipliers, that all terms in this expression are to be evaluated, post facto, at states satisfying the constraints (16) and (21). In the presence of a net lateral pressure  $p$  in the direction of the surface normal  $\mathbf{n}$ , the virtual-work statement is given by Agrawal and Steigmann [2009] as

$$\dot{E}^* = \int_\omega p\mathbf{n} \cdot \dot{\mathbf{r}} da + \int_{\partial\omega} \chi ds, \quad (42)$$

where  $\chi$  is the density of edge power, the form of which will be made explicit below. We remark that because of the definition (14), the energy in this expression is actually the energy divided by the thickness  $t$ ; the dimensions of  $p$  and  $\chi$  are affected accordingly. Thus, for example, the actual pressure is  $tp$ , and with  $p = O(1)$  this is of order  $t$ .

We consider the consequences of (42) with respect to variations of each variable in turn. The simplest are those associated with the variations  $\dot{\boldsymbol{\pi}}$  and  $\dot{\boldsymbol{\eta}}$ . They are given respectively by

$$\mathbf{e}_a = U_{\boldsymbol{\pi}} + \varphi \mathbf{n} \quad (43)$$

and

$$U_{\boldsymbol{\eta}} + \psi \mathbf{n} = \mathbf{0}, \quad (44)$$

(cf. (7) and (15)), with

$$U_{\boldsymbol{\pi}} = \chi_{\perp} \boldsymbol{\pi} - c_2 \boldsymbol{\eta} \quad \text{and} \quad U_{\boldsymbol{\eta}} = k_3 \boldsymbol{\eta} - c_2 \boldsymbol{\pi}. \quad (45)$$

Accordingly, with the constraints (16) and (21) in effect it follows that

$$\psi = 0 \quad \text{and} \quad \boldsymbol{\eta} = (c_2/k_3) \boldsymbol{\pi}; \quad (46)$$

and that

$$\varphi = \mathbf{n} \cdot \mathbf{e}_a \quad \text{and} \quad U_{\boldsymbol{\pi}} = \mathbb{P} \mathbf{e}_a, \quad (47)$$

where  $\mathbb{P} = \mathbf{I} - \mathbf{n} \otimes \mathbf{n}$  is the projection onto  $T_{\omega}$ , with  $\mathbf{I}$  the identity for 3-space. We note that  $\mathbb{P} = \mathbf{1}$ , the identity on  $T_{\omega}$ . Then, from (45)<sub>1</sub>,

$$D \boldsymbol{\pi} = \mathbb{P} \mathbf{e}_a(\mathbf{r}), \quad \text{where } D = \chi_{\perp} - c_2^2/k_3, \quad (48)$$

which furnishes the polarization uniquely in terms of the surface parametrization, provided that  $D \neq 0$  and the applied field is assigned as a function in space. When  $\chi_{\perp} > 0$ , the sign of  $D$  is controlled by the strength of the flexoelectric effect; thus  $D$  is positive or negative according as  $|c_2|$  is small or large, respectively. These alternatives correspond to the relevant quadratic form in the energy being positive definite or indefinite, respectively.

These results imply that the *equilibrium* value of the energy (40) may be regarded as a function of  $H$ ,  $K$  and  $\boldsymbol{\pi}$ ; on combining (46) and (48), the explicit expression is found to be

$$U = kH^2 + \bar{k}K + \frac{1}{2}D|\boldsymbol{\pi}|^2. \quad (49)$$

Necessary conditions for minimum energy states in the absence of polarization, derived in [Agrawal and Steigmann 2008], require  $k > 0$  but do not impose any restriction on  $\bar{k}$ .

With the foregoing in effect, (41) and (42) furnish the residual virtual-work statement

$$\int_{\omega} [\dot{W} + (W + \lambda)\dot{J}/J + \varphi \boldsymbol{\pi} \cdot \dot{\mathbf{n}}] da = \int_{\omega} p \mathbf{n} \cdot \dot{\mathbf{r}} da + \int_{\partial\omega} \chi ds, \quad (50)$$

in which all variations are induced by the virtual velocity

$$\mathbf{u} = \dot{\mathbf{r}} \quad (51)$$

with  $\dot{\boldsymbol{\pi}} = \mathbf{0}$ . In particular [Steigmann 2013],

$$\dot{J}/J = \mathbf{a}^{\alpha} \cdot \mathbf{u}_{,\alpha} \quad (52)$$

and

$$\dot{\mathbf{n}} = \varepsilon^{\beta\alpha} \mathbf{a}_{\beta} \times \mathbf{u}_{,\alpha} - (\dot{J}/J)\mathbf{n}, \quad (53)$$

where  $\varepsilon^{\beta\alpha}$  is the contravariant Levi-Civita permutation tensor ( $\varepsilon^{\beta\alpha} = e^{\beta\alpha}/\sqrt{a}$ , with  $e^{12} = -e^{21} = 1$  and  $e^{11} = e^{22} = 0$ ); and

$$\dot{W} = \dot{U} - \dot{\mathbf{e}}_a(\mathbf{r}) \cdot \boldsymbol{\pi}, \quad (54)$$

with

$$\dot{\mathbf{e}}_a = (\text{grad } \mathbf{e}_a)|_{\omega} \mathbf{u} \quad \text{and} \quad \dot{U} = U_H \dot{H} + U_K \dot{K}, \quad (55)$$

in which we have invoked  $U_{\eta} = \mathbf{0}$  (cf. (44) and (46)<sub>1</sub>); and, from (49),

$$U_H = 2kH \quad \text{and} \quad U_K = \bar{k}. \quad (56)$$

Expressions for the variations  $\dot{H}$  and  $\dot{K}$  are known [Steigmann et al. 2003] and will be recalled in the next subsection. To facilitate their representation we use the decomposition

$$\mathbf{u} = u^{\alpha} \mathbf{a}_{\alpha} + w \mathbf{n}, \quad (57)$$

where  $u^{\alpha}$  and  $w$  respectively are the *tangential* and *normal* variations of the position field.

**4.1. Tangential variations.** For tangential variations we have  $w = 0$  and

$$\dot{J}/J = u^{\alpha}_{,\alpha}, \quad \dot{H} = u^{\alpha} H_{,\alpha} \quad \text{and} \quad \dot{K} = u^{\alpha} K_{,\alpha} \quad (58)$$

[Steigmann et al. 2003]. Thus,

$$(W + \lambda)\dot{J}/J = [(W + \lambda)u^{\alpha}]_{,\alpha} - u^{\alpha}(W + \lambda)_{,\alpha}, \quad (59)$$

where

$$W_{,\alpha} = U_{,\alpha} - \mathbf{a}_{\alpha} \cdot (\text{grad } \mathbf{e}_a)|_{\omega} \boldsymbol{\pi} - \mathbf{e}_a \cdot \boldsymbol{\pi}_{,\alpha}; \quad (60)$$

whereas, with  $\dot{\boldsymbol{\pi}} = \mathbf{0}$ ,

$$\dot{W} = u^{\alpha} [U_H H_{,\alpha} + U_K K_{,\alpha} - \mathbf{a}_{\alpha} \cdot (\text{grad } \mathbf{e}_a)|_{\omega} \boldsymbol{\pi}]. \quad (61)$$

We thus reach

$$\begin{aligned} \dot{W} + (W + \lambda)\dot{J}/J \\ = [(W + \lambda)u^\alpha]_{;\alpha} + u^\alpha (U_H H_{,\alpha} + U_K K_{,\alpha} - U_{,\alpha} - \lambda_{,\alpha} + \mathbf{e}_a \cdot \boldsymbol{\pi}_{,\alpha}). \end{aligned} \quad (62)$$

Here we use the fact that  $U_\eta$  vanishes in equilibrium, together with (48) and the symmetry of  $\mathbb{P}$ , to derive

$$U_{,\alpha} = U_H H_{,\alpha} + U_K K_{,\alpha} + \mathbf{e}_a \cdot \mathbb{P}\boldsymbol{\pi}_{,\alpha}, \quad (63)$$

which furnishes

$$\dot{W} + (W + \lambda)\dot{J}/J = [(W + \lambda)u^\alpha]_{;\alpha} + u^\alpha [(\mathbf{e}_a \cdot \mathbf{n})(\mathbf{n} \cdot \boldsymbol{\pi}_{,\alpha}) - \lambda_{,\alpha}]. \quad (64)$$

To reduce the term in (50) involving  $\varphi$ , we use (53) to obtain [Steigmann 2013]

$$\dot{\mathbf{n}} = \varepsilon^{\beta\alpha} b_{\lambda\alpha} u^\lambda \mathbf{a}_\beta \times \mathbf{n}, \quad (65)$$

where  $\varepsilon_{\beta\lambda}$  is the covariant permutation tensor, together with  $\varepsilon^{\beta\alpha} \varepsilon_{\beta\lambda} = \delta_\lambda^\alpha$  (the Kronecker delta). This and

$$\mathbf{n} \times \mathbf{a}_\beta = \varepsilon_{\beta\gamma} \mathbf{a}^\gamma \quad (66)$$

yield

$$\dot{\mathbf{n}} = -b_{\lambda\alpha} u^\lambda \mathbf{a}^\alpha, \quad (67)$$

which combines with (47) to deliver

$$\varphi \boldsymbol{\pi} \cdot \dot{\mathbf{n}} = -(\mathbf{n} \cdot \mathbf{e}_a) b_{\alpha\beta} \pi^\beta u^\alpha. \quad (68)$$

On the other hand, the constraint (21) implies that

$$\boldsymbol{\pi}_{,\alpha} = \pi_{;\alpha}^\beta \mathbf{a}_\beta + b_{\alpha\beta} \pi^\beta \mathbf{n}, \quad (69)$$

where  $(\cdot)_{;\alpha}$  is the covariant derivative on  $\omega$ . Accordingly,

$$\dot{W} + (W + \lambda)\dot{J}/J + \varphi \boldsymbol{\pi} \cdot \dot{\mathbf{n}} = [(W + \lambda)u^\alpha]_{;\alpha} - u^\alpha \lambda_{,\alpha}. \quad (70)$$

Using Stokes' theorem, the surface integral over  $\omega$  of the first term on the right-hand side may be represented as an integral over the edge  $\partial\omega$ . Remarkably, the Euler equations emerging from (42) under tangential variations are then given simply by  $\lambda_{,\alpha} = 0$ ; i.e.,

$$\lambda \text{ is constant on } \omega, \quad (71)$$

as in the classical Canham–Helfrich theory for uniform lipid bilayers in the absence of electromagnetic effects [Steigmann et al. 2003; Dharmavaram and Healey 2015]. Edge conditions are discussed below.

**4.2. Normal variations.** In this case  $\mathbf{u} = w\mathbf{n}$ , yielding [Steigmann et al. 2003]

$$\begin{aligned} \dot{J}/J &= -2Hw, \\ 2\dot{H} &= \Delta w + w(4H^2 - 2K), \\ \dot{K} &= 2KHw + (\tilde{b}^{\alpha\beta} w_{,\alpha})_{;\beta}, \end{aligned} \quad (72)$$

where, for any scalar field  $\xi$ ,

$$\Delta\xi = \frac{1}{\sqrt{a}}(\sqrt{a}a^{\alpha\beta}\xi_{,\beta})_{,\alpha} \quad (73)$$

is the surface Laplacian in which  $a^{\alpha\beta}$  is the dual metric.

Recalling that  $U_\eta$  vanishes at equilibrium and noting that  $\boldsymbol{\pi}$  is fixed in the present class of variations, after a lengthy calculation presented explicitly in [Steigmann et al. 2003] we reach

$$\begin{aligned} \dot{W} + (W + \lambda)\dot{J}/J &= w\left[\Delta\left(\frac{1}{2}U_H\right) + (U_K)_{;\alpha\beta}\tilde{b}^{\alpha\beta} + U_H(2H^2 - K)\right. \\ &\quad \left.+ 2H(KU_K - W) - 2H\lambda - \mathbf{n} \cdot (\text{grad } \mathbf{e}_a)|_\omega \boldsymbol{\pi}\right] \\ &\quad + \left[\left(\frac{1}{2}U_H a^{\alpha\beta} + U_K \tilde{b}^{\alpha\beta}\right)w_{,\alpha}\right]_{;\beta} \\ &\quad - \left\{\left[(U_H)_{,\beta}a^{\alpha\beta} + (U_K)_{,\beta}\tilde{b}^{\alpha\beta}\right]w\right\}_{;\alpha}. \end{aligned} \quad (74)$$

Here  $(\cdot)_{;\alpha\beta}$  is the second covariant derivative on  $\omega$  and

$$\tilde{\mathbf{b}} = 2H\mathbf{1} - \mathbf{b} \quad (75)$$

is the cofactor of the curvature tensor.

Further, (53) now gives [Steigmann 2013]

$$\dot{\mathbf{n}} = -\mathbf{a}^\alpha w_{,\alpha}, \quad (76)$$

yielding

$$\varphi\boldsymbol{\pi} \cdot \dot{\mathbf{n}} = w(\varphi\pi^\alpha)_{;\alpha} - (\varphi\pi^\alpha w)_{;\alpha}. \quad (77)$$

Combining this with (41) and writing the integrals of the divergences as boundary integrals, from (42) the relevant Euler–Lagrange equation is found to be

$$\begin{aligned} \Delta\left(\frac{1}{2}U_H\right) + (U_K)_{;\alpha\beta}\tilde{b}^{\alpha\beta} + U_H(2H^2 - K) \\ + 2H(\boldsymbol{\pi} \cdot \mathbf{e}_a + KU_K - U) - 2H\lambda + (\varphi\pi^\alpha)_{;\alpha} \\ = p + \mathbf{n} \cdot (\text{grad } \mathbf{e}_a)|_\omega \boldsymbol{\pi}, \end{aligned} \quad (78)$$

where

$$(\varphi\pi^\alpha)_{;\alpha} = \frac{1}{\sqrt{a}}(\sqrt{a}\varphi\pi^\alpha)_{,\alpha}. \quad (79)$$

This generalizes the well-known *shape equation* of the conventional theory [Ouyang et al. 1999; Agrawal and Steigmann 2009]. For the particular energy given

by (49) it reduces to

$$k[\Delta H + 2H(H^2 - K)] - D|\boldsymbol{\pi}|^2 H + 2H(\boldsymbol{\pi} \cdot \mathbf{e}_a - \lambda) + (\varphi \pi^\alpha)_{;\alpha} \\ = p + \mathbf{n} \cdot (\text{grad } \mathbf{e}_a)|_{\omega} \boldsymbol{\pi}, \quad (80)$$

with  $\varphi = \mathbf{n} \cdot \mathbf{e}_a$ . This may be simplified by using (47) and (48) to reach

$$D\boldsymbol{\pi} \cdot \mathbf{e}_a = \mathbb{P}\mathbf{e}_a \cdot \mathbf{e}_a = |\mathbb{P}\mathbf{e}_a|^2 = |\mathbf{e}_a|^2 - \varphi^2. \quad (81)$$

Then,

$$2H\boldsymbol{\pi} \cdot \mathbf{e}_a - D|\boldsymbol{\pi}|^2 H = D^{-1}H(|\mathbf{e}_a|^2 - \varphi^2). \quad (82)$$

## 5. Edge conditions

With (71) and (78) satisfied the residual virtual-work statement (42) is

$$\int_{\partial\omega} \chi \, ds = B_t + B_n, \quad (83)$$

with

$$B_t = \int_{\partial\omega} (W + \lambda) u^\alpha v_\alpha \, ds \quad (84)$$

and

$$B_n = \int_{\partial\omega} \left( \frac{1}{2} U_H a^{\alpha\beta} + U_K \tilde{b}^{\alpha\beta} \right) v_\beta w_{,\alpha} \, ds \\ - \int_{\partial\omega} \left[ \frac{1}{2} (U_H)_{,\beta} a^{\alpha\beta} + (U_K)_{,\beta} \tilde{b}^{\alpha\beta} + \varphi \pi^\alpha \right] v_\alpha w \, ds. \quad (85)$$

The part  $B_n$  of the boundary working given above is exactly as in [Agrawal and Steigmann 2009]. The part  $B_t$  may be reduced to a more convenient form by using  $u^\alpha = \mathbf{u} \cdot \mathbf{a}^\alpha$  with  $\mathbf{a}^\alpha = v^\alpha \mathbf{v} + \tau^\alpha \boldsymbol{\tau}$ , where  $\boldsymbol{\tau} = \mathbf{n} \times \mathbf{v}$  is the unit tangent to  $\partial\omega$ . This satisfies  $\boldsymbol{\tau} = d\mathbf{r}(\theta^\alpha(s))/ds = \tau^\alpha \mathbf{a}_\alpha$ , where  $\tau^\alpha = d\theta^\alpha/ds$ . Thus,

$$u^\alpha = v^\alpha \mathbf{u} \cdot \mathbf{v} + \tau^\alpha \mathbf{u} \cdot \boldsymbol{\tau}, \quad (86)$$

yielding

$$B_t = \int_{\partial\omega} (W + \lambda) \mathbf{v} \cdot \mathbf{u} \, ds. \quad (87)$$

For smooth edges the foregoing may be combined with  $B_n$  to reduce the edge-power density to the compact form [Agrawal and Steigmann 2009]

$$\chi = \mathbf{f} \cdot \mathbf{u} - M \boldsymbol{\tau} \cdot \boldsymbol{\omega}, \quad (88)$$

where  $\boldsymbol{\omega}$  is the variation of the surface orientation ( $\dot{\mathbf{n}} = \boldsymbol{\omega} \times \mathbf{n}$ ),

$$M = \frac{1}{2} U_H + \kappa_\tau U_K \quad (89)$$

is the bending couple per unit length on  $\partial\omega$  (divided by  $t$  in accordance with the remarks following (42)),

$$\mathbf{f} = F_\nu \mathbf{v} + F_\tau \boldsymbol{\tau} + F_n \mathbf{n} \quad (90)$$

and is the edge traction (force per unit length, divided by  $t$ ) on  $\partial\omega$ , with

$$\begin{aligned} F_\nu &= W + \lambda - \kappa_\nu M, & F_\tau &= -\tau M \quad \text{and} \\ F_n &= (\tau U_K)_{,s} - \left(\frac{1}{2}U_H\right)_{,\nu} - (U_K)_{,\beta} \tilde{b}^{\alpha\beta} \nu_\alpha - \varphi \pi^\alpha \nu_\alpha. \end{aligned} \quad (91)$$

Here  $(\cdot)_{,\nu} = \nu^\alpha (\cdot)_{,\alpha}$  and  $(\cdot)_{,s} = \tau^\alpha (\cdot)_{,\alpha} = d(\cdot)/ds$  are the normal and tangential derivatives on the boundary.

Further,

$$\tau = b^{\alpha\beta} \tau_\alpha \nu_\beta \quad (92)$$

is the *twist* of the surface  $\omega$  on the  $\mathbf{v}$ ,  $\boldsymbol{\tau}$ -axes, whereas

$$\kappa_\nu = b^{\alpha\beta} \nu_\alpha \nu_\beta \quad \text{and} \quad \kappa_\tau = b^{\alpha\beta} \tau_\alpha \tau_\beta, \quad (93)$$

respectively, are the *normal* curvatures of  $\omega$  in the directions of  $\mathbf{v}$  and  $\boldsymbol{\tau}$ .

For the energy defined by (39) and (49) the bending moment and edge forces are

$$M = kH + \bar{k}\tau \quad (94)$$

and

$$\begin{aligned} F_\nu &= kH^2 + \bar{k}K + \frac{1}{2}D|\boldsymbol{\pi}|^2 - \mathbf{e}_a \cdot \boldsymbol{\pi} + \lambda - \kappa_\nu M, \\ F_\tau &= -\tau M \quad \text{and} \quad F_n = \bar{k}\tau_{,s} - kH_{,\nu} - \varphi \pi^\alpha \nu_\alpha. \end{aligned} \quad (95)$$

The force and moment are assigned on parts of the boundary that are complementary with respect to  $\partial\omega$ , respectively, to those parts where position  $\mathbf{r}$  and surface orientation  $\mathbf{n}$  are assigned.

## 6. Examples: axisymmetric states

**6.1. Uniform applied field.** We assume the length scale for the spatial variation of the applied field to be much larger than that of the overall dimensions of the lipid membrane, so that the applied field in the vicinity of the membrane is sensibly constant; we take

$$\mathbf{e}_a = E\mathbf{k}, \quad (96)$$

with  $E$  constant and  $\mathbf{k}$  a fixed unit vector.

We seek an axisymmetric solution in the class of closed surfaces of revolution with the axis of symmetry parallel to the applied field. This is parametrized

by meridional arclength  $s$ , measured from the north pole, and azimuthal angle  $\theta \in [0, 2\pi)$ . Thus,

$$\mathbf{r}(s, \theta) = r(s)\mathbf{e}_r(\theta) + z(s)\mathbf{k}, \quad (97)$$

where  $r(s)$  is the radius from the axis of symmetry,  $z(s)$  is the elevation above a base plane and  $\{\mathbf{e}_r, \mathbf{e}_\theta, \mathbf{k}\}$ , with  $\mathbf{e}_\theta = \mathbf{k} \times \mathbf{e}_r$ , is the usual polar orthonormal basis. Meridians and parallels of latitude are the curves on which  $\theta$  and  $s$ , respectively, are constant. Because  $s$  measures arclength along meridians, we have

$$(r')^2 + (z')^2 = 1, \quad (98)$$

where  $(\cdot)' = d(\cdot)/ds$ ; therefore there is  $\psi(s)$  such that

$$r'(s) = \cos \psi \quad \text{and} \quad z'(s) = \sin \psi. \quad (99)$$

Proceeding as in [Agrawal and Steigmann 2009] we have

$$\mathbf{v} = -\cos \psi \mathbf{e}_r - \sin \psi \mathbf{k}, \quad \boldsymbol{\tau} = -\mathbf{e}_\theta, \quad \mathbf{n} = \cos \psi \mathbf{k} - \sin \psi \mathbf{e}_r, \quad (100)$$

and

$$\kappa_\nu = \psi', \quad \kappa_\tau = r^{-1} \sin \psi, \quad \tau = 0. \quad (101)$$

The sum of the normal curvatures is twice the mean curvature  $H(s)$ ; hence the differential equation

$$r\psi' = 2rH - \sin \psi. \quad (102)$$

Their product yields the Gaussian curvature  $K(s)$ ; thus,

$$K = H^2 - (H - r^{-1} \sin \psi)^2. \quad (103)$$

Following the procedure outlined in Section 4.1 of [Agrawal and Steigmann 2009] and noting that the spatial gradient of  $\mathbf{e}_a$  vanishes, with some labor we reduce the shape equation (80) to

$$L' = r \left\{ p/k + (2\lambda/k)H - (E^2/Dk)[H(1 + \cos^2 \psi) - \psi' \sin^2 \psi] - 2H(H - r^{-1} \sin \psi)^2 \right\}, \quad (104)$$

with

$$H' = r^{-1}L. \quad (105)$$

We omit the details of the straightforward but lengthy derivation. The system to be solved thus consists of (99), (102), (104) and (105), for the functions  $r$ ,  $z$ ,  $\psi$ ,  $H$  and  $L$ . To render the number of differential equations consistent with the total number of side conditions, we append (cf. (71))

$$\lambda' = 0. \quad (106)$$



Consider the equilibrium of a subsurface  $\tilde{\omega} \subset \omega$  containing the pole, bounded by a parallel of latitude defined by  $s = \tilde{s}$ , where  $s$  measures meridional arclength from the pole. The force balance (modulo the multiplicative factor  $t^{-1}$ ) is

$$\int_{\tilde{\omega}} [p\mathbf{n} + (\text{grad } \mathbf{e}_a)|_{\omega} \boldsymbol{\pi}] da + \int_{\partial\tilde{\omega}} \mathbf{f} du + F\mathbf{k} = \mathbf{0}, \quad (107)$$

where  $u = r(\tilde{s})\theta$  measures arclength around the perimeter of the parallel,  $\mathbf{f}$  is the force per unit length exerted on  $\tilde{\omega}$  by the part  $\omega \setminus \tilde{\omega}$  of the membrane, and  $F$  is a point load acting at the pole and directed along the symmetry axis. Because parallels of latitude are lines of curvature on the membrane, the twist  $\tau$  vanishes on  $\partial\tilde{\omega}$  and (cf. (90) and (101)<sub>3</sub>)

$$\mathbf{f} = F_v \tilde{\mathbf{v}} + F_n \mathbf{n}, \quad (108)$$

where  $\tilde{\mathbf{v}}$  is the exterior unit normal to  $\tilde{\omega}$  (the opposite of  $\mathbf{v}$  in (100)<sub>1</sub>) and  $F_v, F_n$  are defined in (91). Thus,

$$\mathbf{f} = (F_v \cos \psi - F_n \sin \psi) \mathbf{e}_r + (F_v \sin \psi + F_n \cos \psi) \mathbf{k}, \quad (109)$$

and the periodicity of  $\mathbf{e}_r(\theta)$  yields

$$\int_{\partial\tilde{\omega}} \mathbf{f} du = 2\pi r(\tilde{s})(F_v \sin \psi + F_n \cos \psi) \mathbf{k}. \quad (110)$$

The presumed boundedness of the integrand of the first term in (107), together with the conditions

$$r(0) = 0, \quad \psi(0) = 0 \quad (111)$$

at the pole, imply that

$$F/2\pi + \lim_{\tilde{s} \rightarrow 0} (r F_n) = 0. \quad (112)$$

According to (48), the polarization at the pole is proportional to the projection of the applied field onto the plane with unit normal  $\mathbf{k}$ . This vanishes by virtue of (96), and (112) reduces to

$$F/2\pi + k \lim_{\tilde{s} \rightarrow 0} (r H') = 0, \quad \text{yielding} \quad L(0) = -F/2\pi k. \quad (113)$$

A similar condition applies at the opposite pole of the membrane.

To solve the equations it is convenient to convert the independent variable from the meridional arclength  $s$  to the surface area  $a(s)$  of the sector  $[0, s]$ , defined by

$$a(s) = 2\pi \int_0^s r(t) dt. \quad (114)$$

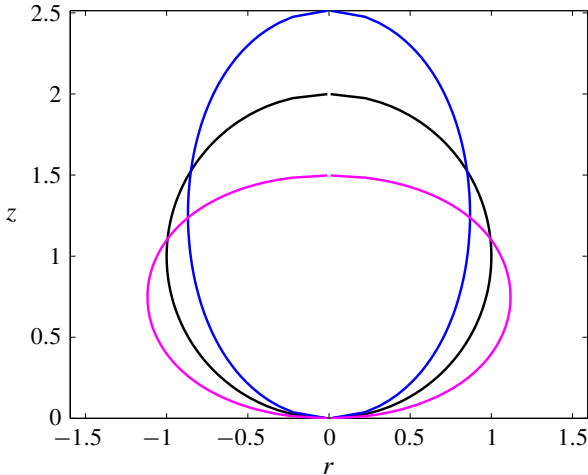
We have  $a'(s) = 2\pi r(s)$ , which is positive on the domain, implying that  $a$  and

$s$  are in one-to-one correspondence and hence that the former may replace the latter as the independent variable. The constraint on surface area is then enforced simply by integrating the equations over the domain  $[0, A]$ , where  $A$  is the assigned membrane area. The conversion of the equations is discussed in detail in [Agrawal and Steigmann 2009] and thus not presented here. To treat the equations numerically we adopt an assigned length scale  $R$ , say, which we take to be the radius of a spherical reference vesicle. This is equilibrated in the absence of pressure and applied electric field, provided that the associated value of  $\lambda$  vanishes. This radius and the modulus  $k$  are the parameters used to nondimensionalize the equations. The electric field enters the resulting system via the combination

$$\bar{E} = E^2 R^2 / Dk, \quad (115)$$

which is positive if the flexoelectric effect is weak, and negative if it is strong.

Figure 1 depicts membrane shapes obtained using the MATLAB boundary value problem solver applied to the foregoing differential equations. The boundary conditions are: zero radius and angle  $\psi$ , and vanishing point load, at the north pole; and zero radius and elevation, with  $\psi = -\pi$ , at the south pole. The reaction force at the south pole is given a posteriori by an appropriate adjustment to (120), if desired, and we impose zero lateral pressure. The effect of flexoelectricity at a given field strength manifests itself as a vertical elongation or compression of the membrane along the field direction, corresponding respectively to weak or strong flexoelectricity.



**Figure 1.** Vesicle subjected to uniform electric field. Black curve: original spherical vesicle. Blue curve: vesicle subjected to  $\bar{E} = 2.5$ . Magenta solid curve: vesicle subjected to  $\bar{E} = -2.5$ .

**6.2. Remote point charge.** We model the response of the membrane to a remote point charge located on the axis of symmetry, situated at the position

$$\mathbf{y}_c = z_c \mathbf{k}. \quad (116)$$

The associated electric potential and field are

$$V_a(\mathbf{y}) = \frac{E}{4\pi\epsilon_0|\mathbf{y}-\mathbf{y}_c|} \quad \text{and} \quad \mathbf{e}_a(\mathbf{y}) = -\text{grad } V_a = \frac{E}{4\pi\epsilon_0|\mathbf{y}-\mathbf{y}_c|^2} \mathbf{u}(\mathbf{y}), \quad (117)$$

respectively, where  $E$  is the charge,  $\epsilon_0$  is the free-space permittivity, and

$$\mathbf{u}(\mathbf{y}) = (\mathbf{y} - \mathbf{y}_c)/|\mathbf{y} - \mathbf{y}_c|. \quad (118)$$

The induced polarization is given by

$$D\boldsymbol{\pi} = \mathbf{e}_a - \varphi \mathbf{n}, \quad (119)$$

where

$$\varphi = \mathbf{n} \cdot \mathbf{e}_a = \frac{E}{4\pi\epsilon_0} [(z - z_c) \cos \psi - r \sin \psi] / [r^2 + (z - z_c)^2]^{3/2}. \quad (120)$$

Thus,

$$\boldsymbol{\pi} = \delta (\cos \psi \mathbf{e}_r + \sin \psi \mathbf{k}), \quad (121)$$

with

$$\delta = \frac{E}{4\pi\epsilon_0 D} [r \cos \psi + (z - z_c) \sin \psi] / [r^2 + (z - z_c)^2]^{3/2}. \quad (122)$$

This is used in (80), in the combination

$$(\varphi \pi^\alpha)_{;\alpha} = \frac{1}{r} (r \varphi \delta)'. \quad (123)$$

From (116)<sub>2</sub> we derive

$$\frac{4\pi\epsilon_0}{E} \text{grad } \mathbf{e}_{a|\omega} = \frac{1}{|\mathbf{r}-\mathbf{y}_c|^3} \mathbf{I} - \frac{3}{|\mathbf{r}-\mathbf{y}_c|^5} (\mathbf{r}-\mathbf{y}_c) \otimes (\mathbf{r}-\mathbf{y}_c), \quad (124)$$

and with a bit of effort we then obtain the loading term

$$\mathbf{n} \cdot (\text{grad } \mathbf{e}_{a|\omega}) \boldsymbol{\pi} = \frac{-3E}{4\pi\epsilon_0 |\mathbf{r}-\mathbf{y}_c|^5} [\mathbf{n} \cdot (\mathbf{r}-\mathbf{y}_c)] [\boldsymbol{\pi} \cdot (\mathbf{r}-\mathbf{y}_c)], \quad (125)$$

with

$$\begin{aligned} \mathbf{n} \cdot (\mathbf{r}-\mathbf{y}_c) &= (z - z_c) \cos \psi - r \sin \psi, \\ \boldsymbol{\pi} \cdot (\mathbf{r}-\mathbf{y}_c) &= \delta [r \cos \psi + (z - z_c) \sin \psi]. \end{aligned} \quad (126)$$

We substitute these results into the shape equation (80), obtaining

$$\begin{aligned} L' = r \{ [p + \mathbf{n} \cdot (\text{grad } \mathbf{e}_a)_{|\omega} \boldsymbol{\pi} - \frac{1}{r} (r \varphi \delta)'] / k \\ + (2\lambda/k) H - D^{-1} H (|\mathbf{e}_a|^2 - \varphi^2) - 2H (H - r^{-1} \sin \psi)^2 \}. \end{aligned} \quad (127)$$

Further, the condition (113)<sub>2</sub>, connecting  $L$  to the point load at a pole, remains valid in the present circumstances. The equations are nondimensionalized as before.

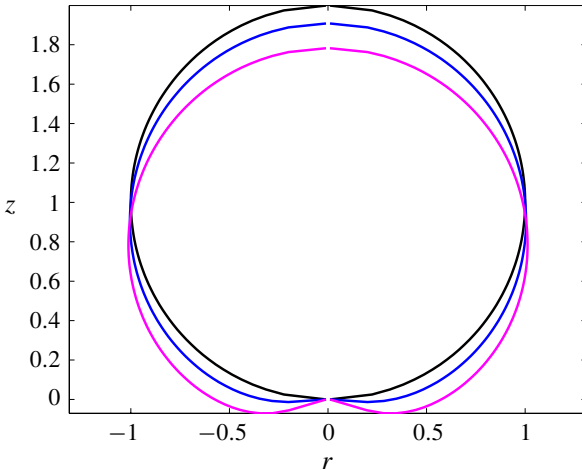
In this case the charge intensity enters the resulting system in the combination

$$\bar{E} = \left( \frac{E}{4\pi\epsilon_0} \right)^2 / Dk, \quad (128)$$

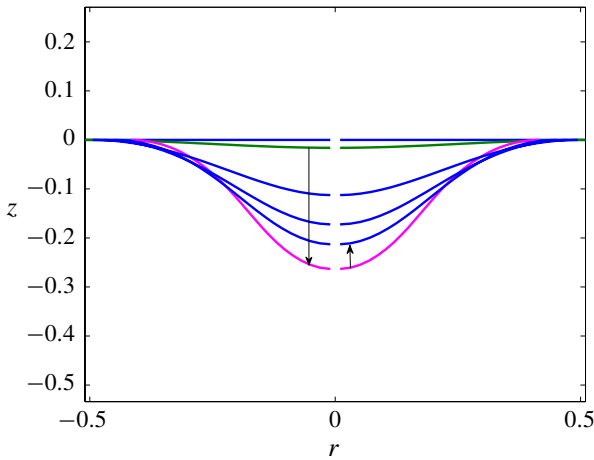
the sign of which depends on the strength of the flexoelectric effect, as before.

We present two examples. In the first the spherical vesicle of the previous example is subjected to a point charge located at the dimensionless position  $\bar{z}_c = -0.75$  below the membrane. The boundary conditions are unchanged, the pressure vanishes, and we consider only the case of weak flexoelectricity ( $\bar{E} > 0$ ). The point charge is seen to attract the membrane against the support reaction at the south pole. [Figure 2](#) depicts the associated membrane equilibria.

The second example concerns a flat disc acted upon by a point charge at position  $\bar{z}_c = -0.50$ , again situated below the membrane. This is equilibrated at vanishing charge and pressure. The areal incompressibility constraint is still operative, but here, simply for the sake of illustration, we solve the problem on the dimensionless arclength interval  $s/R \in [0, 0.5]$ , with  $R$  as in the previous examples. Accordingly, as the membrane deforms, the model accommodates areal incompressibility implicitly via recruitment of lipids through the boundary, so that the computational domain does not correspond in this instance to a fixed set of lipids. The boundary conditions are: zero radius, angle and point load at the pole (i.e., at  $s/R = 0$ ); and zero elevation and angle at the remote edge. We impose zero pressure and take the Lagrange multiplier  $\lambda$  to be zero everywhere. The latter condition renders (106) redundant, and gives rise to a state-dependent force  $F_v$  and moment  $M$  (cf. (94) and (95)<sub>1</sub>) at the edge of the (nonmaterial) domain.

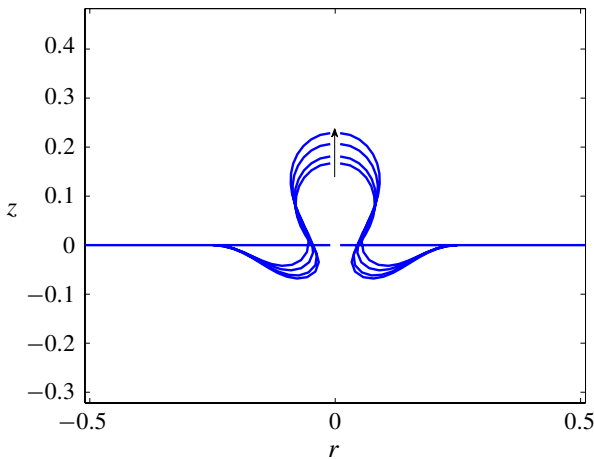


**Figure 2.** Effect of a point charge on the shape of a vesicle. Black curve: original spherical vesicle. Blue curve: vesicle subjected to  $\bar{E} = 1.75$ . Magenta curve: vesicle subjected to  $\bar{E} = 2.08$ .



**Figure 3.** A flat patch of membrane with a nearby point charge. At  $\bar{E} = 0.5$  there is an instability. The patch goes from the green curve to the magenta curve. On increasing  $\bar{E}$  further, the patch begins to flatten out. The maximum  $\bar{E}$  used is 0.83.

Figures 3 and 4 depict a sequence of predicted shapes arising in response to an increasing sequence of charge intensities in the case of weak flexoelectricity ( $\bar{E} > 0$ ). At low values of the charge intensity the membrane is attracted to the charge, as in the previous example. We observe a sudden transition, at  $\bar{E} = 0.50$ , from a slightly curved membrane to a strongly curved one. This trend, as indicated



**Figure 4.** A flat patch of membrane with a nearby point charge. The arrow indicates the direction of increasing  $\bar{E}$ . The innermost shape corresponds to  $\bar{E} = 2.0$  and the outermost shape corresponds to  $\bar{E} = 2.5$ .

by the arrows in [Figure 3](#), is reversed upon a further increase of the charge intensity, yielding ever flatter membranes, up to the value  $\bar{E} = 0.83$  ([Figure 3](#)).

We are unable to find equilibria for  $\bar{E} \in (0.83, 2.0)$  and thus offer the conjecture that no axisymmetric equilibria exist for charges in this interval. On further increase of the charge, however, we find the equilibria displayed in [Figure 4](#), in which the arrow again indicates the trend under increasing charge intensity. These results support a conjecture to the effect that equilibria associated with the disjoint intervals of charge intensity are connected by a dynamic transition. On the second branch of equilibria, the membranes have formed a sequence of buds with ever-narrowing necks, situated, remarkably, above the initial disc shape, of a kind reminiscent of those observed in the process of endocytosis. This suggests that important biological processes such as endocytosis may be controlled, to some degree, by the action of suitable electric fields. We find the shapes displayed in these figures to be quite robust under shape perturbations in the setting of MATLAB, but we have not analyzed their stability. To study stability one should work with simulations carried out on a fixed material domain; i.e., on a patch of fixed area.

## 7. Conclusions

In this work, we formulate a generalized electromechanical theory of lipid membranes systematically from the three-dimensional liquid crystal theory. We derive the Euler–Lagrange equations and edge conditions required to solve boundary value problems in a coupled electromechanical setting. In contrast to earlier studies, we find that the lipid dipoles are primarily oriented in the tangent plane, as in the analogous magnetostatic setting. This has the effect of eliminating the self field from the leading-order two-dimensional model, and yields the important simplification that the self field can be computed a posteriori. Further, it provides justification for the widespread practice of suppressing the self field on an ad hoc basis. We also find that a spatially varying electric field does not lead to a spatial variation of the Lagrange multiplier associated with areal incompressibility. Numerical examples highlight the strong interplay between electric fields and membrane geometry.

## References

- [Agrawal and Steigmann 2008] A. Agrawal and D. J. Steigmann, “Coexistent fluid-phase equilibria in biomembranes with bending elasticity”, *J. Elasticity* **93**:1 (2008), 63–80.
- [Agrawal and Steigmann 2009] A. Agrawal and D. J. Steigmann, “Boundary-value problems in the theory of lipid membranes”, *Contin. Mech. Therm.* **21**:1 (2009), 57–82.
- [Aihara and Miyazaki 1998] H. Aihara and J.-I. Miyazaki, “Gene transfer into muscle by electroporation in vivo”, *Nat. Biotechnol.* **16**:9 (1998), 867–870.
- [Barham et al. 2012] M. Barham, D. J. Steigmann, and D. White, “Magnetoelasticity of highly deformable thin films: theory and simulation”, *Int. J. Non-Linear Mech.* **47**:2 (2012), 185–196.

- [Brownell et al. 1985] W. E. Brownell, C. R. Bader, D. Bertrand, and Y. De Ribaupierre, “Evoked mechanical responses of isolated cochlear outer hair cells”, *Science* **227**:4683 (1985), 194–196.
- [Bustamante et al. 2009] R. Bustamante, A. L. Dorfmann, and R. W. Ogden, “Nonlinear electroelastostatics: a variational framework”, *Z. Angew. Math. Phys.* **60**:1 (2009), 154–177.
- [Cesana et al. 2015] P. Cesana, P. Plucinsky, and K. Bhattacharya, “Effective behavior of nematic elastomer membranes”, *Arch. Ration. Mech. Anal.* **218**:2 (2015), 863–905.
- [Davalos et al. 2005] R. V. Davalos, L. M. Mir, and B. Rubinsky, “Tissue ablation with irreversible electroporation”, *Ann. Biomed. Eng.* **33**:2 (2005), 223–231.
- [Dharmavaram and Healey 2015] S. Dharmavaram and T. J. Healey, “On the equivalence of local and global area-constraint formulations for lipid bilayer vesicles”, *Z. Angew. Math. Phys.* **66**:5 (2015), 2843–2854.
- [Dimova et al. 2007] R. Dimova, K. A. Riske, S. Aranda, N. Bezlyepkina, R. L. Knorr, and R. Lipowsky, “Giant vesicles in electric fields”, *Soft Matter* **3**:7 (2007), 817–827.
- [Dimova et al. 2009] R. Dimova, N. Bezlyepkina, M. D. Jordo, R. L. Knorr, K. A. Riske, M. Staykova, P. M. Vlahovska, T. Yamamoto, P. Yang, and R. Lipowsky, “Vesicles in electric fields: some novel aspects of membrane behavior”, *Soft Matter* **5**:17 (2009), 3201–3212.
- [Dorfmann and Ogden 2014] A. L. Dorfmann and R. W. Ogden, *Nonlinear theory of electroelastic and magnetoelastic interactions*, Springer, New York, NY, 2014.
- [Ericksen 1961] J. L. Ericksen, “Conservation laws for liquid crystals”, *Trans. Soc. Rheol.* **5** (1961), 23–34.
- [Ericksen 1962] J. L. Ericksen, “Hydrostatic theory of liquid crystals”, *Arch. Ration. Mech. Anal.* **9** (1962), 371–378.
- [Ericksen 1976] J. L. Ericksen, “Equilibrium theory of liquid crystals”, pp. 233–298 *Advances in Liquid Crystals* **2**, Academic Press, New York, NY, 1976.
- [Frischleder and Peinel 1982] H. Frischleder and G. Peinel, “Quantum-chemical and statistical calculations on phospholipids”, *Chem. Phys. Lipids* **30**:2–3 (1982), 121–158.
- [Gao et al. 2008] L.-T. Gao, X.-Q. Feng, Y.-J. Yin, and H. Gao, “An electromechanical liquid crystal model of vesicles”, *J. Mech. Phys. Solids* **56**:9 (2008), 2844–2862.
- [de Gennes and Prost 1992] P. G. de Gennes and J. Prost, *The physics of liquid crystals*, Oxford University Press, 1992.
- [Gioia and James 1997] G. Gioia and R. D. James, “Micromagnetics of very thin films”, *Proc. R. Soc. Lond. A* **453**:1956 (1997), 213–223.
- [Harland et al. 2015] B. Harland, W.-H. Lee, W. E. Brownell, S. X. Sun, and A. A. Spector, “The potential and electric field in the cochlear outer hair cell membrane”, *Med. Biol. Eng. Comput.* **53**:5 (2015), 405–413.
- [Helfrich 1973] W. Helfrich, “Elastic properties of lipid bilayers: theory and possible experiments”, *Z. Naturforsch.* **28** (1973), 693–703.
- [Kim and Steigmann 2015] C.-I. Kim and D. J. Steigmann, “Distension-induced gradient capillarity in lipid membranes”, *Contin. Mech. Therm.* **27**:4–5 (2015), 609–621.
- [Kovetz 2000] A. Kovetz, *Electromagnetic theory*, Oxford University Press, 2000.
- [Kummrow and Helfrich 1991] M. Kummrow and W. Helfrich, “Deformation of giant lipid vesicles by electric fields”, *Phys. Rev. A* **44**:12 (1991), 8356–8360.
- [Meyer 1969] R. B. Meyer, “Piezoelectric effects in liquid crystals”, *Phys. Rev. Lett.* **22**:18 (1969), 918–921.
- [Mohammadi et al. 2014] P. Mohammadi, L. P. Liu, and P. Sharma, “A theory of flexoelectric membranes and effective properties of heterogeneous membranes”, *J. Appl. Mech. (ASME)* **81**:1 (2014), Article ID #011007.

- [Naghdi 1972] P. M. Naghdi, “The theory of shells and plates”, pp. 425–640 in *Linear theories of elasticity and thermoelasticity*, edited by C. Truesdell, Handbuch der Physik **VIa/2**, Springer, Berlin, 1972. [Reprinted in 1984.](#)
- [Neumann et al. 1982] E. Neumann, M. Schaefer-Ridder, Y. Wang, and P. H. Hofschneider, “[Gene transfer into mouse lymphoma cells by electroporation in high electric fields](#)”, *Euro. Mol. Biol. Organ. J.* **1:7** (1982), 841–845.
- [Ou-Yang et al. 1999] Z.-C. Ou-Yang, J.-X. Liu, and Y.-Z. Xie, *Geometric methods in the elastic theory of membranes in liquid crystal phases*, Advanced Series on Theoretical Physical Science **2**, World Scientific, Singapore, 1999.
- [Raphael et al. 2000] R. M. Raphael, A. S. Popel, and W. E. Brownell, “[A membrane bending model of outer hair cell electromotility](#)”, *Biophys. J.* **78:6** (2000), 2844–2862.
- [Rubinsky et al. 2007] B. Rubinsky, G. Onik, and P. Mikus, “[Irreversible electroporation: a new ablation modality — clinical implications](#)”, *Technol. Cancer Res. Treat.* **6:1** (2007), 37–48.
- [Schmidt et al. 2012] D. Schmidt, J. del Mármol, and R. MacKinnon, “[Mechanistic basis for low threshold mechanosensitivity in voltage-dependent  \$K^+\$  channels](#)”, *Proc. Nat. Acad. Sci. USA* **109:26** (2012), 10352–10357.
- [Seelig 1978] J. Seelig, “ [\$^{31}\text{P}\$  nuclear magnetic resonance and the head group structure of phospholipids in membranes](#)”, *Biochim. Biophys. Acta Biomembr.* **515** (1978), 105–140.
- [Sokolnikoff 1964] I. S. Sokolnikoff, *Tensor analysis: theory and applications to geometry and mechanics of continua*, Wiley, New York, NY, 1964.
- [Steigmann 2013] D. J. Steigmann, “[A model for lipid membranes with tilt and distension based on three-dimensional liquid crystal theory](#)”, *Int. J. Non-Linear Mech.* **56** (2013), 61–70.
- [Steigmann et al. 2003] D. J. Steigmann, E. Baesu, R. E. Rudd, J. Belak, and M. McElfresh, “[On the variational theory of cell-membrane equilibria](#)”, *Interfaces Free Bound.* **5:4** (2003), 357–366.
- [Toupin 1956] R. A. Toupin, “[The elastic dielectric](#)”, *J. Rational Mech. Anal.* **5:6** (1956), 849–915.
- [Truesdell and Toupin 1960] C. Truesdell and R. A. Toupin, “[The classical field theories](#)”, pp. 226–793 in *Principles of classical mechanics and field theory*, edited by S. Flügge, Handbuch der Physik **III/1**, Springer, Berlin, 1960.
- [Virga 1994] E. G. Virga, *Variational theories for liquid crystals*, Applied Mathematics and Mathematical Computation **8**, Chapman & Hall, London, 1994.
- [Vlahovska 2010] P. M. Vlahovska, “[Nonequilibrium dynamics of lipid membranes: deformation and stability in electric fields](#)”, pp. 101–146 *Advances in Planar Lipid Bilayers and Liposomes* **12**, Academic Press, Oxford, 2010.
- [Warshaviak et al. 2011] D. T. Warshaviak, M. J. Mueller, and M. Chachisvilis, “[Effect of membrane tension on the electric field and dipole potential of lipid bilayer membrane](#)”, *Biochim. Biophys. Acta Biomembr.* **1808:10** (2011), 2608–2617.
- [Weaver 2000] J. C. Weaver, “[Electroporation of cells and tissues](#)”, *IEEE Trans. Plasma Sci.* **28:1** (2000), 24–33.
- [Winterhalter and Helfrich 1988] M. Winterhalter and W. Helfrich, “[Deformation of spherical vesicles by electric fields](#)”, *J. Colloid Interface Sci.* **122:2** (1988), 583–586.

Received 19 Jun 2015. Revised 13 Nov 2015. Accepted 14 Dec 2015.

DAVID J. STEIGMANN: [dsteigmann@berkeley.edu](mailto:dsteigmann@berkeley.edu)

Department of Mechanical Engineering, University of California, 6189 Etchevery Hall No. 1740, Berkeley, CA 94720-1740, United States

ASHUTOSH AGRAWAL: [ashutosh@uh.edu](mailto:ashutosh@uh.edu)

Department of Mechanical Engineering, University of Houston, Houston, TX 77204, United States







# ORTHOGONAL POLYNOMIALS AND RIESZ BASES APPLIED TO THE SOLUTION OF LOVE'S EQUATION

PIERLUIGI VELLUCCI AND ALBERTO MARIA BERSANI

In this paper we reinvestigate the structure of the solution of a well-known Love's problem, related to the electrostatic field generated by two circular coaxial conducting disks, in terms of orthogonal polynomial expansions, enlightening the role of the recently introduced class of the Lucas–Lehmer polynomials. Moreover we show that the solution can be expanded more conveniently with respect to a Riesz basis obtained starting from Chebyshev polynomials.

## 1. Introduction

In 1949, E. R. Love [1949] considered the electrostatic field generated by two identical circular coaxial conducting disks at equal, and at equal and opposite, potentials, the potential at infinity being taken equal to zero. He established a celebrated expression for the potential, involving the solution of an integral equation of well-known type, much simpler than that considered by other authors in previous works.

Love's integral equation is a Fredholm equation of the second kind. It has found applications in several applied physics fields such as polymer structures, aerodynamics, fracture mechanics, hydrodynamics, and elasticity engineering. Recently, a polynomial expansion scheme was proposed by M. Agida and A. S. Kumar [2010] as an analytical method for solving Love's integral equation in the case of a rational kernel. Their study is concerned with the calculation of the normalized field created conjointly by two similar plates of radius  $R$ , separated by a distance  $kR$ , where  $k$  is a positive real parameter, and at equal or opposite potential, with zero potential at infinity; the solution of this problem solves a Love's second kind integral equation; see also [Love 1990; Ren et al. 1999].

We propose two different approaches to this problem. In Section 2, starting from a classical technique, based on the expansion of the solution in orthogonal polynomials, we employ a class of polynomials introduced in [Vellucci and Bersani

---

**Communicated by Antonio Carcaterra.**

*MSC2010:* 00A69, 33C45.

*Keywords:* integral equations, numerical approximation and analysis, Love equation, Chebyshev polynomials, Lucas–Lehmer primality test, exponential bases, Riesz bases.

2016], in order to solve a modified version of the original Love's equation. In Section 3, we recall a work by M. Norgren and B. L. G. Jonsson [2009], and we show that their results are still valid, expanding the solution of Love's integral equation with respect to a nonharmonic Fourier cosine series, which is a particular case of a Riesz basis [Sun and Zhou 1999].

For literature related to the numerical solutions of singular integral equations of the deterministic type, we refer to the fundamental book by L. Fox and I. B. Parker [1968], where different analytical methods for the solution of random integral equations were investigated.

## 2. Chebyshev polynomials approach

**2.1. Preliminaries.** In the following, we introduce the mathematical tools to employ an analytical method for solving Love's integral equation in the case of a rational kernel. Afterwards, we recall a short summary on Love's original problem.

**2.1.1. Chebyshev and Lucas–Lehmer polynomials.** The Chebyshev polynomials of the first kind [Chebyshev 1858; Chebyshev 1875; Erdélyi et al. 1953; Gatteschi 1973; Rivlin 1990] satisfy the recurrence relation

$$\begin{cases} T_0(x) = 1, \\ T_1(x) = x, \\ T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x), \quad n \geq 2. \end{cases}$$

The polynomials  $T_n(x)$  are orthogonal with respect to the weight function  $1/\sqrt{1-x^2}$  defined on  $x \in (-1, 1)$ . In a previous paper [Vellucci and Bersani 2016], we studied a class of polynomials  $L_n(x) = L_{n-1}^2(x) - 2$ , created by means of the same iterative formula used to build the well-known Lucas–Lehmer sequence, employed in primality tests [Lucas 1878; Lehmer 1930; Ribenboim 1988; Bressoud 1989; Koshy 2001]. It is clearly crucial to choose the first term of the polynomial sequence. In [Vellucci and Bersani 2016], we showed that the Lucas–Lehmer polynomials are orthogonal in the interval  $(-2, 2)$  with respect to the weight  $w(x) = 1/(4\sqrt{4-x^2})$  and such that their zeros belong to the interval  $(-2, 2)$ , that is to their orthogonality interval. From now on we will consider  $L_0 = x$ .

Let us first recall some important properties of these polynomials.

**Proposition 1.** *For each  $n \geq 1$ ,*

$$L_n(x) = 2T_{2n-1}\left(\frac{1}{2}x^2 - 1\right) \tag{1}$$

**Proposition 2.** *The polynomials  $L_n(x)$  are orthogonal with respect to the weight function  $1/(4\sqrt{4-x^2})$  defined on  $x \in (-2, 2)$ .*

**Corollary 3.** *With  $x = 2 \cos \theta$ , the polynomials  $L_n(x)$  admit the representation*

$$L_n(2 \cos \theta) = 2 \cos(2^n \theta) \quad (2)$$

For  $|x| \leq 2$  we can show another formula for  $L_n$ :

$$L_n(x) = \left( \frac{1}{2}x^2 - 1 - \sqrt{\left(\frac{1}{2}x^2 - 1\right)^2 - 1} \right)^{2^{n-1}} + \left( \frac{1}{2}x^2 - 1 + \sqrt{\left(\frac{1}{2}x^2 - 1\right)^2 - 1} \right)^{2^{n-1}}. \quad (3)$$

We change the sign inside the radical, factoring out the imaginary unit:

$$L_n(x) = \left( \frac{1}{2}x^2 - 1 - \iota \sqrt{1 - \left(\frac{1}{2}x^2 - 1\right)^2} \right)^{2^{n-1}} + \left( \frac{1}{2}x^2 - 1 + \iota \sqrt{1 - \left(\frac{1}{2}x^2 - 1\right)^2} \right)^{2^{n-1}}. \quad (4)$$

We then calculate the powers of the complex conjugate numbers  $L_n^+$  and  $L_n^-$ , depending on the variable  $x$ . Let

$$\begin{aligned} L_n(x) &= \left( \frac{1}{2}x^2 - 1 + \sqrt{\left(\frac{1}{2}x^2 - 1\right)^2 - 1} \right)^{2^{n-1}} + \left( \frac{1}{2}x^2 - 1 - \sqrt{\left(\frac{1}{2}x^2 - 1\right)^2 - 1} \right)^{2^{n-1}} \\ &= L_n^+(x) + L_n^-(x). \end{aligned} \quad (5)$$

The absolute value of both complex numbers is unitary, since

$$\begin{aligned} |L_n^-| &= \left| \frac{1}{2}x^2 - 1 - \iota \sqrt{1 - \left(\frac{1}{2}x^2 - 1\right)^2} \right|^{2^{n-1}}, \\ |L_n^+| &= \left| \frac{1}{2}x^2 - 1 + \iota \sqrt{1 - \left(\frac{1}{2}x^2 - 1\right)^2} \right|^{2^{n-1}}, \end{aligned}$$

and

$$|L_n^+| = |L_n^-| = \sqrt{\left(\frac{1}{2}x^2 - 1\right)^2 + 1 - \left(\frac{1}{2}x^2 - 1\right)^2} = 1. \quad (6)$$

Moreover, since  $L_1(\pm\sqrt{2}) = 0$ ,  $L_2(\pm\sqrt{2}) = -2$ , and  $L_n(\pm\sqrt{2}) = 2$  for all  $n \geq 3$ , the argument of  $L_n(\pm\sqrt{2})$  is 0 for every  $n \geq 3$ . In the other cases, since we can write  $x = 2 \cos \vartheta$  when  $|x| \leq 2$ , it follows that  $\frac{1}{2}x^2 - 1 = \cos 2\vartheta$ . Thus for  $|x| \neq \sqrt{2}$ , we can also put

$$\vartheta(x) = \frac{1}{2} \arctan \frac{\sqrt{1 - \left(\frac{1}{2}x^2 - 1\right)^2}}{\frac{1}{2}x^2 - 1} + b\pi \quad (7)$$

where  $b$  is a binary digit. Finally, using (2), we obtain  $L_n(x) = 2 \cos(2^n \vartheta(x))$ .

By further setting

$$\theta(x) = \frac{1}{2} \arctan \frac{\sqrt{1 - \left(\frac{1}{2}x^2 - 1\right)^2}}{\frac{1}{2}x^2 - 1}, \quad (8)$$

we can write

$$L_n(x) = 2 \cos(2^n \theta(x) + 2^n b\pi) = 2 \cos(2^n \theta(x)). \quad (9)$$

On the other hand, the Chebyshev polynomials of the first kind can be defined as the unique polynomials satisfying

$$T_n(t) = \cos(n \arccos t)$$

or, in other words, as the unique polynomials satisfying

$$T_n(\cos(\vartheta)) = \cos(n\vartheta)$$

for  $n = 0, 1, 2, 3, \dots$ . Therefore, by [Proposition 1](#),

$$L_n(x) = 2T_{2^{n-1}}\left(\frac{1}{2}x^2 - 1\right) = 2 \cos\left(2^{n-1} \arccos\left(\frac{1}{2}x^2 - 1\right)\right).$$

**2.1.2. Love's problem.** Two leading cases of the problem are considered here: to specify the field generated by two identical circular coaxial conducting disks (a) at equal potentials and (b) at equal and opposite potentials, the potential at infinity being taken as zero. The results established by Love are as follows: the upper sign referring to the case of equally charged disks and the lower to that of oppositely charged disks. For [Theorem 4](#) we refer to [[Love 1949](#), Figures 1 and 2].

**Theorem 4** [[Love 1949](#)]. *In the two leading cases described above, the potential at any point  $(p, \zeta, \zeta')$ , specified by its distance  $r = pa$  from the axis of the disks and its axial distances  $z = \zeta a$  and  $z' = \zeta' a$  from their planes, is*

$$\frac{V_0}{\pi} \int_{-1}^1 \left( \frac{1}{\sqrt{\rho^2 + (\zeta + it)^2}} \pm \frac{1}{\sqrt{\rho^2 + (\zeta' + it)^2}} \right) f(t) dt, \quad (10)$$

where  $V_0$  is potential of the disks,  $a$  is the radius of the disks, each square root has positive real part, and  $f(t)$  is the solution of the integral equation

$$f(x) \pm \frac{1}{\pi} \int_{-1}^1 \frac{k}{k^2 + (x-t)^2} f(t) dt = 1, \quad |x| \leq 1 \quad (11)$$

where  $k$  is the spacing parameter.

**Theorem 5** [[Love 1949](#)]. *For every positive  $k$ , (11) has a continuous solution, and no other solution: it is real and even, and is specifiable by the Neumann series*

$$f(x) = 1 + \sum_{n=1}^{\infty} (\mp 1)^n \int_{-1}^1 K_n(x, t) dt, \quad (12)$$

where the iterated kernels  $K_n(x, t)$ , for  $n \in \mathbb{N}$ ,  $n > 1$ , are given by

$$K_1(x, t) = \frac{1}{\pi} \frac{k}{k^2 + (x-t)^2} \quad \text{and} \quad K_n(x, t) = \int_{-1}^1 K_{n-1}(x, s) K_1(s, t) ds.$$

**Theorem 6** [Love 1949]. *The capacitance of each disk in the two cases is*

$$\frac{a}{\pi} \int_{-1}^1 f(t) dt,$$

and the components of the field at all points not on the disks are given by the appropriate formal differentiations of (10).

**2.2. The classical approach to the problem in terms of orthogonal polynomials.**

For the solution of the problems we will refer to [Fox and Parker 1968]. When the upper and lower disks are at potentials  $V_0$  and  $\pm V_0$ , the potential  $V$  at any point whose spheroidal coordinates are  $(\mu, \eta)$  with respect to the upper disk and  $(\mu', \eta')$  with respect to the lower one is expressed in terms of Legendre functions. The upper disk, specified in cylindrical polar coordinates  $(r, \theta, z)$  by  $r \leq a$  and  $z = 0$ , is taken as “focal disk”  $\eta = 0$  of spheroidal coordinates  $(\mu, \eta)$ ; in actual study these are such that  $-2 \leq \mu \leq 2, \eta \geq 0$ .

Then (10) can be rewritten in the form

$$\frac{V_0}{2\pi} \int_{-2}^2 \left( \frac{1}{\sqrt{\rho^2 + (\zeta + it/2)^2}} \pm \frac{1}{\sqrt{\rho^2 + (\zeta' + it/2)^2}} \right) f(t/2) dt, \quad (13)$$

where each square root has positive real part, and  $f(t)$  is the solution of the integral equation

$$f(x) \pm \frac{1}{2\pi} \int_{-2}^2 \frac{k}{k^2 + (x - t/2)^2} f(t/2) dt = 1, \quad |x| \leq 2. \quad (14)$$

By the linear transformation  $t = 2y$ , both equations can be reduced to Love’s original form. In (14) we put  $k = 1$  and consider positive sign, so

$$f(x) + \frac{1}{2\pi} \int_{-2}^2 \frac{1}{1 + (x - t/2)^2} f(t/2) dt = 1, \quad |x| \leq 2. \quad (15)$$

We make the substitution  $x \mapsto \frac{1}{2}x^2 - 1$  in (15), yielding

$$f\left(\frac{1}{2}x^2 - 1\right) + \frac{1}{2\pi} \int_{-2}^2 \frac{1}{1 + \left(\frac{1}{2}(x^2 - t) - 1\right)^2} f\left(\frac{1}{2}t\right) dt = 1.$$

We can find a Chebyshev series solution as follows: write

$$f(x) = \sum_{r=0}^{\infty} a_r T_r(x),$$

substitute it into (15), interchange the order of integration and summation in the first term. Then we arrive at the equation

$$\sum_{r=0}^{\infty} a_r T_r\left(\frac{1}{2}x^2 - 1\right) + \frac{1}{2\pi} \sum_{s=0}^{\infty} a_s \int_{-2}^2 \frac{T_s\left(\frac{1}{2}t\right)}{1 + \left(\frac{1}{2}(x^2 - t) - 1\right)^2} dt = 1 \quad (16)$$

for  $|x| \leq 2$ . If we can now determine the expansion

$$\frac{1}{2} \int_{-2}^2 \frac{T_s\left(\frac{1}{2}t\right)}{1 + \left(\frac{1}{2}(x^2 - t) - 1\right)^2} dt = \sum_{r=0}^{\infty} b_{sr} T_r\left(\frac{1}{2}x^2 - 1\right),$$

we can equate the corresponding coefficients of each  $T_r(x)$  on both sides of (15), which is legitimate since the Chebyshev polynomials form a complete set of independent functions, to produce an infinite set of algebraic equations for the required coefficients  $a_r$ , given by

$$a_r + \sum_{s=0}^{\infty} a_s b_{sr} = 0, \quad r = 1, 2, \dots \quad (17)$$

and, for  $r = 0$ ,

$$a_0 + \sum_{s=0}^{\infty} a_s b_{s,0} = 1.$$

The  $a_r$  will decrease rapidly for sufficiently large  $r$ , so that in a convenient method of solving (17) we select the first  $n + 1$  rows and columns, perform Gaussian elimination and back-substitution for the last few coefficients —  $a_n, a_{n-1}, a_{n-2}$ , say — decide by inspection whether convergence is sufficiently rapid for the required precision with this selected value of  $n$ , and if necessary add some extra rows and columns with only a small additional amount of work.

Let's go back to (16). Let

$$J = \{1, 2, 4, \dots\} = \{2^{r-1} \mid r \in \mathbb{N}\},$$

and rewrite (16) in this way:

$$\sum_{r=0}^{\infty} a_r T_r\left(\frac{1}{2}x^2 - 1\right) + \sum_{s=0}^{\infty} a_s \sum_{r=0}^{\infty} c_{sr} T_r\left(\frac{1}{2}x^2 - 1\right) = 1,$$

where  $c_{sr} = b_{sr}/\pi$ . Then

$$\begin{aligned} \sum_{r \in J} a_r T_r\left(\frac{1}{2}x^2 - 1\right) + \sum_{r \notin J} a_r T_r\left(\frac{1}{2}x^2 - 1\right) \\ + \sum_{s=0}^{\infty} a_s \left( \sum_{r \in J} c_{sr} T_r\left(\frac{1}{2}x^2 - 1\right) + \sum_{r \notin J} c_{sr} T_r\left(\frac{1}{2}x^2 - 1\right) \right) = 1. \end{aligned}$$

By (1),

$$\begin{aligned} \frac{1}{2} \sum_{r=1}^{\infty} a_r L_r(x) + \sum_{r \notin J} a_r T_r\left(\frac{1}{2}x^2 - 1\right) \\ + \frac{1}{2} \sum_{s=0}^{\infty} a_s \sum_{r=1}^{\infty} c_{sr} L_r(x) + \sum_{s=0}^{\infty} a_s \sum_{r \notin J} c_{sr} T_r\left(\frac{1}{2}x^2 - 1\right) = 1 \end{aligned}$$

By Proposition 1, we note that solving (17), a subset of first  $n + 1$  rows and columns selected to perform Gaussian elimination, is due to Lucas–Lehmer polynomials. They not only cannot by themselves guarantee the convergence to the solution, but also their contributions can be neglected. In fact, by above reasoning, since

$$f(x) = \sum_{r \notin J} a_r T_r\left(\frac{1}{2}x^2 - 1\right) + \frac{1}{2} \sum_{r \in J} a_r L_r(x),$$

we have

$$\left| f(x) - \sum_{r \notin J} a_r T_r\left(\frac{1}{2}x^2 - 1\right) \right| \leq \frac{1}{2} \sum_{r \in J} |a_r| = \frac{1}{2} \sum_{r=1}^{\infty} |a_{2r-1}|.$$

Accordingly, when the term on the right hand side can be considered “small” with respect to other contributions, a convenient method of solving (17) should be to select the first  $n + 1$  rows and columns, perform Gaussian elimination and back-substitution for the last few coefficients —  $a_n, a_{n-1}, a_{n-2}$ , say — and delete terms due to Lucas–Lehmer polynomials.

### 3. An alternative approach: nonharmonic Fourier series

The capacitance of a circular parallel plate capacitor can be calculated by expanding the solution of the Love’s integral equation in terms of a Fourier cosine series. In previous literature, this kind of expansion was carried out numerically, leading to accuracy problems at small plate separations. Norgren and Jonsson [2009] calculated analytically all expansion integrals in terms of the sine and cosine integrals. Hence, they approximated the kernel using considerably large matrices, resulting in improved numerical accuracy for the capacitance. Previously, G. T. Carlson and B. L. Illman [1994], solved the Love’s equation through an expansion of the kernel into a Fourier cosine series. To calculate the expansion coefficients of the kernel, they use numerical integration. Hence, as noted in [Norgren and Jonsson 2009], their method is limited by a combination of the accuracy of the integration and the large number of terms needed. The accumulated errors effectively limit the expansion to about 100 terms, which is insufficient for the convergence at very small separations. Let us observe that both the methods recalled here make use of orthogonal expansions.

In this section we will use some basic facts about nonharmonic Fourier series, and we recall them below.

It is well known that the family of exponentials  $\{e^{int}\}_{n \in \mathbb{Z}}$  forms an orthonormal basis in  $L^2(-\pi, \pi)$ . The natural question that arises is: what happens if we replace it by a classical system of exponentials  $\{e^{i\lambda_n t}\}_{n \in \mathbb{Z}}$ ; these bases are very useful for the study of the so-called almost periodic functions. See, for example, [Andres et al. 2006; Besicovitch 1932].

The celebrated work of Paley and Wiener [1934] kicked off studies on classical systems of exponentials  $\{e^{i\lambda_n t}\}_{n \in \mathbb{Z}}$  in  $L^2(0, T)$ , where  $T > 0$ . They proved that if  $\lambda_n \in \mathbb{R}$ ,  $n \in \mathbb{Z}$ , and

$$|\lambda_n - n| \leq L < \pi^{-2}, \quad n \in \mathbb{Z}$$

then the system  $\{e^{i\lambda_n t}\}_{n \in \mathbb{Z}}$  forms a Riesz basis in  $L^2[-\pi, \pi]$ , i.e., a family of the form  $\{Ue_k\}_{k=1}^\infty$ , where  $\{e_k\}_{k=1}^\infty$  is an orthonormal basis for a separable infinite-dimensional Hilbert space  $\mathcal{H}$  and  $U: \mathcal{H} \rightarrow \mathcal{H}$  is a bounded bijective operator.

M. I. Kadec [1964] extended this result to the case  $L < \frac{1}{4}$ . This is the so-called *Kadec- $\frac{1}{4}$  theorem*, which over the following 50 years has been extensively generalized; see, for example [Avdonin 1974; Bailey 2010; Pavlov 1979; Sedletskii 2003; Sun and Zhou 1999; Vellucci 2015]. Let us recall Kadec's original result:

**Theorem 7.** *If  $\{\lambda_n\}_{n \in \mathbb{Z}}$  is a sequence of real numbers for which*

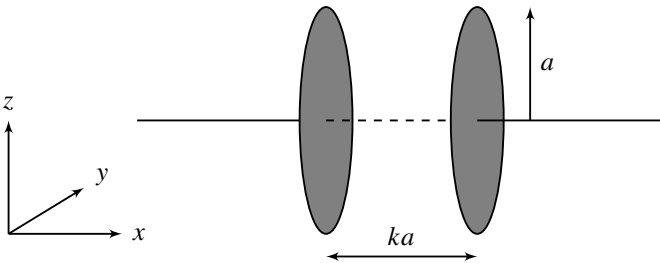
$$\sup_n |\lambda_n - n| < \frac{1}{4}, \quad n = 0, \pm 1, \dots$$

*then the system  $\{e^{i\lambda_n t}\}_{n \in \mathbb{Z}}$  is a Riesz basis for  $L^2[-\pi, \pi]$ .*

Therefore, if  $L = \sup_n |\lambda_n - n| < \frac{1}{4}$ , then the sine system  $\{\sin \lambda_n t\}_1^\infty$  as well as the cosine system  $1 \cup \{\cos \lambda_n t\}_1^\infty$  is a Riesz basis in  $L^2(0, \pi)$ .

We now approach the problem described in [Norgren and Jonsson 2009]. The circular parallel plate capacitor is depicted in Figure 1.

The distance between the circular plates is here put equal to their common radius. Accordingly, the normalized separation between the plates, a constant  $k$ , is set for



**Figure 1.** A circular parallel plate capacitor can be viewed as a cylindrical volume whose bases are the capacitor's plates.



the sake of simplicity equal to 1. The model is idealized in the sense that the plates have zero thickness.

The capacitance of the parallel plate capacitor is [Carlos and Illman 1994]

$$C = 4\epsilon_0 a \int_0^1 f(s) ds, \tag{18}$$

where  $a$  is the radius of the circular plate and the function  $f(s)$  is the solution of the modified Love's integral equation

$$f(s) - \int_0^1 K(s, t) f(t) dt = 1, \quad 0 \leq s \leq 1, \tag{19}$$

with kernel

$$K(s, t) = \frac{1}{\pi} \left( \frac{1}{1 + (s - t)^2} + \frac{1}{1 + (s + t)^2} \right). \tag{20}$$

To solve (19) numerically, we follow the approach in [loc. cit.] and expand the kernel and the unknown function into the (nonharmonic) Fourier cosine expansion in terms of the functions

$$\tilde{\psi}_n(s) = \sqrt{2 - \delta_{n,0}} \cos(\lambda_n s), \quad n = 0, 1, \dots$$

which in our study have been orthonormalized to fulfill the orthogonality relation

$$\int_0^\pi \psi_n(s) \psi_m(s) ds = \delta_{m,n}$$

and satisfy Kadec's assumption on  $L = \sup_n |\lambda_n - n| < \frac{1}{4}$ . Here,  $\delta_{m,n}$  denotes the Kronecker delta function.

This orthonormalization process is shown in the following

**Theorem 8** (orthonormalization process). *Consider  $L^2(-\pi, \pi)$  and a sequence  $\{\lambda_n\}_{n \in \mathbb{Z}} \subset \mathbb{R}$  which satisfies Kadec's assumption. Let  $P = (I - S)^{-1} = \sum_{m=0}^\infty S^m$ , where*

$$S(f)(x) = \sum_{n=-\infty}^\infty \hat{f}(n) (e^{inx} - e^{i\lambda_n x})$$

and  $\{\hat{f}(n)\}$  are the Fourier coefficients of  $f$ . Then  $P(e^{i\lambda_n x}) = e^{inx}$  for each  $n \in \mathbb{Z}$ .

*Proof.* By Kadec's theorem, we have that  $\|S\| < 1$ . Hence,  $P = (I - S)^{-1} = \sum_{m=0}^\infty S^m$ . To show that  $P(e^{i\lambda_n x}) = e^{inx}$ , we write

$$e^{i\lambda_n x} = (I - S)e^{inx} = e^{inx} - \sum_k c_k (e^{ikx} - e^{i\lambda_k x})$$

where  $c_k = \langle e^{inx}, e^{ikx} \rangle$ . Thus

$$e^{inx} - e^{i\lambda_n x} = \sum_k \delta_{n,k} (e^{ikx} - e^{i\lambda_k x}). \quad \square$$

In this way we have orthonormalized the Riesz basis  $\{e^{i\lambda_n x}\}$ , in an easy way. Further results on the orthonormalization of more complex Riesz bases, such as  $\{\phi(t - n)\}_{n \in \mathbb{Z}}$ , applied for example to the study of a “digital filter,” can be found in [Meyer 1989]. For our purposes it is sufficient to consider the basis introduced at the beginning of Section 3 and used in Theorem 7.

Carrying out the expansions of  $f(s)$  and  $K(s, t)$  in terms of  $\{\psi_n\}$ , we obtain

$$f(s) = \sum_{m=0}^{\infty} f_m \psi_m(s), \quad \text{where } f_m = \int_0^{\pi} f(s) \psi_m(s) ds \quad (21)$$

and

$$K(s, t) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} K_{mn} \psi_m(s) \psi_n(t), \quad (22)$$

where

$$K_{mn} = \int_0^{\pi} \int_0^{\pi} K(s, t) \psi_n(t) \psi_m(s) ds dt.$$

These equations yield the infinite linear system of equations for the coefficients  $\{f_n\}_{n=0}^{\infty}$ :

$$\sum_{n=0}^{\infty} (\delta_{m,n} - K_{mn}) f_n = \delta_{m,0}, \quad m = 0, 1, \dots \quad (23)$$

From (18), (21), and from the orthonormalization process described in Theorem 8 and guaranteed by Kadec’s assumption, which allows to us expand the kernel and the unknown function into the (nonharmonic) Fourier cosine expansion in terms of the functions  $\{\cos(\lambda_n s)\}$ , the capacitance reduces to  $C = 4\varepsilon_0 a f_0$ , where  $f_0$  is simply the  $(0, 0)$ -element in the inverse of the matrix with elements  $\delta_{m,n} - K_{mn}$ , as obtained in [Norgren and Jonsson 2009].

Furthermore, Norgren and Jonsson derive analytical expressions for the expansion of the kernel  $K(s, t)$ . Proceeding as in [loc. cit.], it is easy to prove that, in the general case when  $m \neq n$  and  $m, n > 0$ ,

$$K_{mn} = \frac{2}{\pi} \tilde{I}_3(n\pi, m\pi), \quad (24)$$

where  $\tilde{I}_3(n\pi, m\pi) = P I_3(\lambda_n \pi, \lambda_m \pi)$ , with  $P$  as in Theorem 8 and  $I_3$  as defined in [loc. cit.]. The application of the operator  $P$  denotes here the orthonormalization process performed on the set of functions  $\{\cos(\lambda_n s)\}_{n \in \mathbb{Z}}$ .

We have extended the results of [Carlos and Illman 1994; Norgren and Jonsson 2009] to a (nonharmonic) Fourier cosine expansion in terms of the set of functions  $\{\cos(\lambda_n s)\}_{n \in \mathbb{Z}}$ , employing a simple procedure, due to Theorem 8, to orthonormalize the Riesz basis  $\{e^{i\lambda_n x}\}$  under Kadec’s assumption. Therefore, we have found a further expansion of the solution that it is not in terms of orthogonal polynomials, but in terms of nonharmonic functions  $\cos(\lambda_n s)$ ,  $s \in \mathbb{R}$ .

#### 4. Conclusion and perspectives

Orthogonal functions, other classes of polynomials, and Riesz bases have shown to be very powerful for the search of solutions of several problems in disparate fields, from physics to engineering, from economics to biology, and so on. In this paper we applied a new class of orthogonal polynomials, called Lucas–Lehmer polynomials [Vellucci and Bersani 2016] and the tool of Riesz bases [Paley and Wiener 1934] in order to reinvestigate a classical problem, due to Love [1949], obtaining a further expansion of the solution that it is not in terms of orthogonal polynomials, but in terms of nonharmonic functions  $\cos(\lambda_n s)$ ,  $s \in \mathbb{R}$ , suitably orthonormalized, thanks to Theorem 8 which uses the celebrated result due to Kadec [1964]. Many other applications can be investigated in the future, mainly in the field of mechanics. In particular, great attention has been recently paid to peridynamics and fracture mechanics, which are approached in terms of integral equations. In this framework our researches will be addressed to apply our techniques to the study of some integral equations of the type introduced by Piola in 1848, recently rediscovered in the framework of peridynamics and fracture dynamics, and reported in the paper [dell’Isola et al. 2015].

#### References

- [Agida and Kumar 2010] M. Agida and A. S. Kumar, “A Boubaker polynomials expansion scheme solution to random Love’s equation in the case of a rational kernel”, *Electron. J. Theor. Phys.* **7**:24 (2010), 319–326.
- [Andres et al. 2006] J. Andres, A. M. Bersani, and R. F. Grande, “Hierarchy of almost-periodic function spaces”, *Rend. Mat. Appl. (7)* **26**:2 (2006), 121–188.
- [Avdonin 1974] S. A. Avdonin, “On the question of Riesz bases of exponential functions in  $L^2$ ”, *Vestnik Leningrad. Univ.* **13** (1974), 5–12. In Russian.
- [Bailey 2010] B. Bailey, “Sampling and recovery of multidimensional bandlimited functions via frames”, *J. Math. Anal. Appl.* **367**:2 (2010), 374–388.
- [Besicovitch 1932] A. S. Besicovitch, *Almost periodic functions*, Cambridge University Press, 1932.
- [Bressoud 1989] D. M. Bressoud, *Factorization and primality testing*, Springer, New York, 1989.
- [Carlos and Illman 1994] G. T. Carlos and B. L. Illman, “The circular disk parallel plate capacitor”, *Amer. J. Phys.* **62**:12 (1994), 1099–1105.
- [Chebyshev 1858] P. L. Chebyshev, “Sur une nouvelle série”, *Bull. Acad. sci. St. Pétersbourg* **17** (1858). Reprinted as pp. 379–384 in *Œuvres de P. L. Tchebychef*, vol. I, St. Petersburg, 1899.
- [Chebyshev 1875] P. L. Chebyshev, “Объ интерполировании величинъ равноотстоящихъ”, *Zap. Imp. Akad. Nauk (St. Petersburg)* **25**:5 (1875). Translated as “Sur l’interpolation des valeurs équidistantes”, on pp. 219–242 in *Œuvres de P. L. Tchebychef*, vol. II, St. Petersburg, 1899.
- [dell’Isola et al. 2015] 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.
- [Erdélyi et al. 1953] A. Erdélyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi, *Higher transcendental functions*, vol. II, McGraw-Hill, New York-Toronto-London, 1953.

- [Fox and Parker 1968] L. Fox and I. B. Parker, *Chebyshev polynomials in numerical analysis*, Oxford University Press, London-New York-Toronto, 1968.
- [Gatteschi 1973] L. Gatteschi, *Funzioni speciali*, Unione Tipografico-Editrice Torinese, 1973.
- [Kadec 1964] M. Ī. Kadec, “The exact value of the Paley–Wiener constant”, *Dokl. Akad. Nauk SSSR* **155** (1964), 1253–1254. In Russian; translated in *Sov. Math.* **5** (1964), 559–561.
- [Koshy 2001] T. Koshy, *Fibonacci and Lucas numbers with applications*, Wiley, New York, 2001.
- [Lehmer 1930] D. H. Lehmer, “An extended theory of Lucas’ functions”, *Ann. of Math.* (2) **31**:3 (1930), 419–448.
- [Love 1949] E. R. Love, “The electrostatic field of two equal circular co-axial conducting disks”, *Quart. J. Mech. Appl. Math.* **2** (1949), 428–451.
- [Love 1990] E. R. Love, “The potential due to a circular parallel plate condenser”, *Mathematika* **37**:2 (1990), 217–231.
- [Lucas 1878] E. Lucas, “Theorie des Fonctions Numeriques Simplement Periodiques”, *Amer. J. Math.* **1**:4 (1878), 289–321.
- [Meyer 1989] Y. Meyer, “Wavelets and operators”, pp. 256–365 in *Analysis at Urbana* (Urbana, IL, 1986–87), vol. 1, edited by E. Berkson and T. Peck, London Math. Soc. Lecture Note Ser. **137**, Cambridge Univ. Press, 1989.
- [Norgren and Jonsson 2009] M. Norgren and B. L. G. Jonsson, “The capacitance of the circular parallel plate capacitor obtained by solving the Love integral equation using an analytic expansion of the kernel”, *Progr. Electromag. Res.* **97** (2009), 357–372.
- [Paley and Wiener 1934] R. E. A. C. Paley and N. Wiener, *Fourier transforms in the complex domain*, Amer. Math. Soc. Colloq. Publ. **19**, American Mathematical Society, New York, 1934.
- [Pavlov 1979] B. S. Pavlov, “The basis property of a system of exponentials and the condition of Muckenhoupt”, *Dokl. Akad. Nauk SSSR* **247**:1 (1979), 37–40. In Russian; translated in *Sov. Math.* **20** (1979), 655–659.
- [Ren et al. 1999] Y. Ren, B. Zhang, and H. Qiao, “A simple Taylor-series expansion method for a class of second kind integral equations”, *J. Comput. Appl. Math.* **110**:1 (1999), 15–24.
- [Ribenboim 1988] P. Ribenboim, *The book of prime number records*, Springer, New York, 1988.
- [Rivlin 1990] T. J. Rivlin, *Chebyshev polynomials*, 2nd ed., Wiley, New York, 1990.
- [Sedletskii 2003] A. M. Sedletskii, “Nonharmonic analysis”, *J. Math. Sci. (New York)* **116**:5 (2003), 3551–3619.
- [Sun and Zhou 1999] W. Sun and X. Zhou, “On the stability of multivariate trigonometric systems”, *J. Math. Anal. Appl.* **235**:1 (1999), 159–167.
- [Vellucci 2015] P. Vellucci, “A simple pointview for Kadec-1/4 theorem in the complex case”, *Ric. Mat.* **64**:1 (2015), 87–92.
- [Vellucci and Bersani 2016] P. Vellucci and A. M. Bersani, “The class of Lucas–Lehmer polynomials”, (2016). To appear in *Rend. Mat. Appl.* (7). arXiv 1603.01989

Received 15 Oct 2015. Revised 8 Dec 2015. Accepted 6 Feb 2016.

PIERLUIGI VELLUCCI: [pierluigi.vellucci@sbai.uniroma1.it](mailto:pierluigi.vellucci@sbai.uniroma1.it)

Dipartimento di Scienze di Base e Applicate per l’Ingegneria, Sapienza Università di Roma,  
Via Antonio Scarpa 16, I-00161 Rome, Italy

ALBERTO MARIA BERSANI: [alberto.bersani@sbai.uniroma1.it](mailto:alberto.bersani@sbai.uniroma1.it)

Dipartimento di Scienze di Base e Applicate per l’Ingegneria, Sapienza Università di Roma,  
Via Antonio Scarpa 16, I-00161 Rome, Italy





## MODELING CAPILLARY HYSTERESIS IN UNSATURATED POROUS MEDIA

GÉRARD GAGNEUX AND OLIVIER MILLET

This paper deals with the modeling of cyclic hysteresis phenomena for flows in unsaturated porous media, using a dynamic regularization process of Sobolev type. The addition of a kinematic regularizing term of third-order partial derivatives, depending on a strictly positive, small real parameter, enables us to capture the missing information of the ill-posed hysteresis phenomena via Rankine–Hugoniot and “entropy” inequalities. When this parameter tends to zero, an oriented hysteresis loop, corresponding to the realistic problem modeled, emerges from the flow of an associated auxiliary ordinary differential equation.

### 1. Introduction

The modeling of moisture transport in partially saturated porous media is of major importance for civil engineering, soil physics, and pharmaceutical applications. The hysteresis effects, often neglected in the modeling as they are difficult to be taken into account, play a central role in the imbibition and drying process.

In this paper, we propose an original modeling of cyclic hysteresis phenomena in partially saturated porous media, in the simplified case of water–air flows. The approach used is based on the artificial introduction of an unstable spinodal interval and on Sobolev’s method of dynamic regularization, inspired by the works of P. I. Plotnikov [1996; 1994], publicized by L. C. Evans and M. Portilheiro [2004; Evans 2004]. The hysteresis graph is replaced by Cartesian curves and an artificial spinodal interval generating instabilities, with associated attractive–repulsive dynamics.

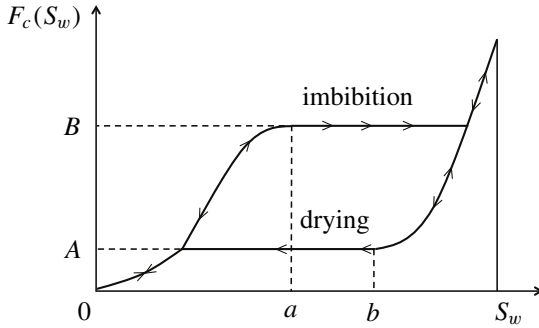
The additional information to describe the hysteresis effects is introduced on the form of entropy-type inequalities. This way, the asymptotic limit of viscous approximate solutions generates effects of irreversibility and enables us to recover the expected hysteresis loop.

---

**Communicated by Francesco dell’Isola.**

*MSC2010:* primary 35K65; secondary 47J40, 76S05.

*Keywords:* capillary hysteresis loop, unsaturated porous media.



**Figure 1.** Graph of  $F_c$  representing the hysteresis loop during an imbibition-drying process.

## 2. The physical problem

**2.1. Richards equation.** The flow of two fluid phases (water–air flow), isothermal and immiscible, in an unsaturated porous medium is considered. To focus on the study of the hysteresis effects, in particular irreversibility, the gravity is neglected and the porous medium is assumed to be homogeneous and isotropic.<sup>1</sup> Moreover, the water vapor in the gas phase is neglected and the air pressure is assumed to be constant and equal to the atmospheric pressure.<sup>2</sup>

The water saturation  $S_w$  is classically governed by a Richards equation,

$$\varphi \frac{\partial S_w}{\partial t} - \Delta \Theta_c(S_w) = 0, \quad (1)$$

where  $\varphi$  denotes the porosity of the porous medium considered. We assume that the residual saturation of each fluid is equal to zero.

Even if the mathematical analysis of this equation is now well stated [Gagneux and Madaune-Tort 1995; Lions 1969], it ignores the hysteresis and dynamic effects that play a major role in the behavior of unsaturated porous media.

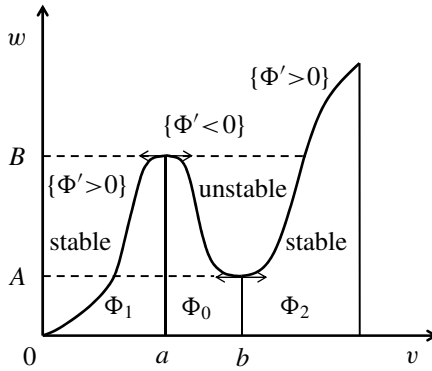
**2.2. Hysteresis modeling.** The capillary hysteresis effects can be modeled with a multivalued operator  $F_c$  whose oriented graph  $F_c$  of  $\mathbb{R}^2$  is represented in Figure 1. The circulation sense depends both on the values of  $S_w$  and on the sign of  $\frac{\partial S_w}{\partial t}$ . It characterizes the imbibition and drying phases, through the differential inclusion<sup>3</sup>

$$0 \in \left\{ \varphi \frac{\partial S_w}{\partial t} - \Delta F_c \left( S_w, \text{sign} \left( \frac{\partial S_w}{\partial t} \right) \right) \right\}. \quad (\mathcal{P}_{\text{hyst}})$$

<sup>1</sup>The analysis holds also for the anisotropic case.

<sup>2</sup>It is equivalent to assume that the gas phase moves fast and is connected to outside.

<sup>3</sup>In the sense of [Aubin and Cellina 1984].



**Figure 2.** The spinodal interval  $[a, b]$  and the unstable part of the graph of  $\Phi$ .

It is equivalent to search a pair  $(S_w, q_w)$  with  $q_w \in \{F_c(S_w, \text{sign}(\frac{\partial S_w}{\partial t}))\}$  which satisfies

$$\varphi \frac{\partial S_w}{\partial t} - \Delta q_w = 0 \quad (2)$$

associated to a Cauchy initial condition  $S_{w(0)}$  and Neumann homogeneous boundary conditions. To simplify the problem without loss of generality, we will consider in the following a normalized porosity  $\varphi = 1$  (this is always possible using a homothetical time scaling, when  $\varphi$  is constant).

In a first step, using a mathematical artifice, we replace the part of the graph  $F_c$  representing the loop by a cubic Cartesian curve  $\Phi$  (Figure 2). Then, to approach problem  $(\mathcal{P}_{\text{hyst}})$ , a nonlinear monotone diffusion equation with an ad hoc “spinodal” interval  $]a, b[$  is introduced. The graph  $F_c$  is replaced on  $]a, b[$  by a cubic spline function, denoted  $\Phi_0$ , whose slope is strictly negative everywhere on  $]a, b[$  and assuring a  $\mathcal{C}^1$  continuity at  $(a, B)$  and  $(b, A)$  with the preserved part.

The graph  $F_c$  is decomposed by splitting its domain of definition into three distinct parts. This leads us to introduce three injective functions,  $\Phi_0$ ,  $\Phi_1$  and  $\Phi_2$ , defined on  $[a, b]$ ,  $[0, a]$  and  $[b, 1]$ , respectively. We denote by  $\beta_0$ ,  $\beta_1$  and  $\beta_2$  their respective inverse functions and by  $\Phi$  the numerical function of class  $\mathcal{C}^1$  on  $[0, 1]$  whose graph is the joining of the graphs of  $\Phi_0$ ,  $\Phi_1$  and  $\Phi_2$ .

This substitution enables us to give sense to the initial formal problem in a suitable mathematical functional framework, via the initial–boundary value system

$$\begin{cases} \frac{\partial v}{\partial t} - \Delta \Phi(v) = 0 & \text{in } Q = ]0, T[ \times \Omega, \\ \frac{\partial \Phi(v)}{\partial n} = 0 & \text{on } \Sigma = ]0, T[ \times \Gamma, \\ v(0) = S_{w(0)} & \text{in } \Omega, \end{cases} \quad (P_\Phi)$$

where  $\Omega$  denotes a bounded domain of  $\mathbb{R}^d$ ,  $d \geq 1$ , with a Lipschitz boundary  $\Gamma$  and an associated external unit normal vector  $n$ . The new forward–backward problem  $(P_\Phi)$  with variable parabolicity direction is ill posed without any supplementary information, because of the nonmonotonic function  $\Phi$ . The dynamic regularization process that follows will enable us to regularize the problem. Note that a problem similar to  $(P_\Phi)$  has been studied in [Smarrazzo and Tesei 2010; 2012; Smarrazzo 2008].

### 3. Dynamic regularization process of Sobolev type

The classical notations that follow are introduced. Let  $\Omega$  be a bounded domain of  $\mathbb{R}^d$  whose boundary  $\Gamma$  is a Lipschitz manifold of dimension  $d - 1$ . For  $T > 0$ , we write  $Q = ]0, T[ \times \Omega$ ,  $\Sigma = ]0, T[ \times \Gamma$  and let  $\Delta$  be the Laplacian operator<sup>4</sup> of  $\mathbb{R}^d$ . We denote by  $H^s(\Omega)$ ,  $s \in \mathbb{R}$ , the classical Hilbert spaces [Lions and Magenes 1968]. For all  $\varepsilon > 0$ , the embedding of  $H^s(\Omega)$  into  $H^{s-\varepsilon}(\Omega)$  is compact. Moreover, we identify  $L^2(\Omega) = H^0(\Omega)$  to its dual, so that the dual of  $H^1(\Omega)$ , denoted  $H^1(\Omega)'$ , can be identified to an superspace of  $L^2(\Omega)$  with  $H^1(\Omega) \hookrightarrow L^2(\Omega) \hookrightarrow H^1(\Omega)'$ , the embeddings being dense and continuous. In addition, an initial state  $S_{w(0)} \in L^\infty(\Omega)$  satisfying  $0 \leq S_{w(0)} \leq 1$  a.e. in  $\Omega$  is given.

The dynamical regularization process of Sobolev type used is based on “artificial viscosity”. A parameter  $\lambda > 0$  is introduced in the initial ill-posed problem  $(P_\Phi)$ , which is transformed into the third order boundary problem

$$\begin{cases} \frac{\partial v_\lambda}{\partial t} - \Delta \Phi(v_\lambda) - \lambda \Delta \frac{\partial v_\lambda}{\partial t} = 0 & \text{in } Q = ]0, T[ \times \Omega, \\ \frac{\partial}{\partial n} \left( \Phi(v_\lambda) + \lambda \frac{\partial v_\lambda}{\partial t} \right) = 0 & \text{on } \Sigma = ]0, T[ \times \Gamma, \\ v_\lambda(0) = S_{w(0)} & \text{in } \Omega. \end{cases} \quad (P_\Phi)_\lambda$$

We then introduce the auxiliary dynamical unknown  $w_\lambda$  defined by

$$w_\lambda = \Phi(v_\lambda) + \lambda \frac{\partial v_\lambda}{\partial t}, \quad (3)$$

or equivalently

$$\frac{\partial v_\lambda}{\partial t} = \frac{w_\lambda - \Phi(v_\lambda)}{\lambda}, \quad t > 0. \quad (4)$$

Note that the dynamics created by (4) drive the system onto stable parts of the graph, as we will see in the sequel.

<sup>4</sup>In the sense of distributions.



Problem  $(P_\Phi)_\lambda$  may be rewritten as

$$\begin{cases} \frac{\partial v_\lambda}{\partial t} - \Delta w_\lambda = 0 & \text{in } Q = ]0, T[ \times \Omega, \\ \frac{\partial w_\lambda}{\partial n} = 0 & \text{on } \Sigma = ]0, T[ \times \Gamma, \\ v_\lambda(0) = S_{w(0)} & \text{in } \Omega. \end{cases} \quad (5)$$

Equivalently, for nearly all  $t$ ,  $w_\lambda$  is a solution of the elliptic problem parametrized in time<sup>5</sup>

$$\begin{cases} w_\lambda - \lambda \Delta w_\lambda = \Phi(v_\lambda) & \text{in } \Omega, t > 0, \\ \frac{\partial w_\lambda}{\partial n} = 0 & \text{on } \Gamma, t > 0. \end{cases} \quad (6)$$

The “ $-$  Laplacian” operator (denoted again  $-\Delta$ ) specifically associated with homogeneous Neumann boundary conditions on  $\Gamma$  is obviously a nonbounded operator from  $L^2(\Omega)$  into  $L^2(\Omega)$  whose domain is  $H^2(\Omega)$ . We introduce the Yosida regularization  $-\Delta_\lambda := -\Delta(J_\lambda) = (I - J_\lambda)/\lambda$  and its resolvent  $J_\lambda = (I - \lambda \Delta)^{-1}$ . According to (6), it follows that

$$w_\lambda = J_\lambda \Phi(v_\lambda). \quad (7)$$

Problem  $(P_\Phi)_\lambda$  can be formulated again in a well-posed form<sup>6</sup> for the operator  $-\Delta_\lambda \Phi$  in  $L^2(\Omega)$ :

$$\begin{cases} \frac{\partial v_\lambda}{\partial t} - \Delta_\lambda \Phi(v_\lambda) = 0, & t \in ]0, T[, \\ v_\lambda(0) = S_{w(0)} & \text{in } L^2(\Omega) \text{ and a.e. in } \Omega. \end{cases} \quad (P_\lambda)$$

The following proposition summarizes the properties of the solutions of  $(P_\lambda)$ :

**Proposition 1.** *Let us denote by  $g : \mathbb{R} \rightarrow \mathbb{R}$  a Lipschitz nondecreasing function, a so-called “entropy function”, and let*

$$G(r) = \int_0^r g(s) ds, \quad G_\Phi(r) = \int_0^r g(\Phi(s)) ds, \quad g_{1/2}(r) = \int_0^r \sqrt{g'(s)} ds, \quad r \in \mathbb{R}. \quad (8)$$

From the Rademacher theorem [Evans and Gariepy 1992],  $g'$  is a bounded Borelian representative of the derivative of  $g$  in its class.

For all  $\lambda > 0$ , the solution  $v_\lambda$  of  $(P_\lambda)$  associated to  $w_\lambda$  has the following properties:

(i) *Estimations using entropy inequations (for each entropy function  $g$ ):*

$$\frac{\partial}{\partial t} G_\Phi(v_\lambda) \leq \operatorname{div}(g(w_\lambda) \nabla w_\lambda) - g'(w_\lambda) |\nabla w_\lambda|^2 \quad \text{in } Q. \quad (9)$$

<sup>5</sup>According to an observation of [Evans and Portilheiro 2004].

<sup>6</sup>Thanks to the Cauchy–Lipschitz–Picard theorem via a first-order differential equation with given initial condition.

Using the function  $G$  defined in (8), the last inequality can be rewritten as

$$\frac{\partial}{\partial t} G_{\Phi}(v_{\lambda}) \leq \Delta(G(w_{\lambda})) - G''(w_{\lambda})|\nabla w_{\lambda}|^2 \quad \text{in } Q.$$

The following inequalities hold:

$$\int_{\Omega} G_{\Phi}(v_{\lambda}(t, x)) dx \leq \int_{\Omega} G_{\Phi}(v_{\lambda}(s, x)) dx \leq \int_{\Omega} G_{\Phi}(S_{w(0)}) dx, \quad t > s > 0,$$

$$\|g_{\frac{1}{2}}(w_{\lambda})\|_{L^2([0, T]; H^1(\Omega))}^2 \stackrel{\text{def}}{=} \int_Q g'(w_{\lambda})|\nabla w_{\lambda}|^2 dx dt \leq C_g, \quad C_g = C(g).$$

(ii) We have the following uniform a priori estimates:

$$\|v_{\lambda}\|_{L^{\infty}(Q)} + \|w_{\lambda}\|_{L^{\infty}(Q)} \leq C_1,$$

$$\|w_{\lambda}\|_{L^2([0, T]; H^1(\Omega))} + \sqrt{\lambda} \left\| \frac{\partial v_{\lambda}}{\partial t} \right\|_{L^2(Q)} \leq C_2,$$

$$\left\| \frac{\partial v_{\lambda}}{\partial t} \right\|_{L^2([0, T]; (H^1(\Omega))')} \leq C_3.$$

The frame constants depend on the extremum values of  $\Phi$  and  $S_{w(0)}$ .

*Proof.* The general principle of the proof of this proposition may be found in [Evans 2004, p. 427]. This classical computation is somewhat akin to an entropy flux calculation for a hyperbolic conservation law, through choices of nondecreasing functions  $g$  (see also [Gagneux and Millet 2015] for more details). We note that the inequality (9) is straightforward from the following relation, for any function  $\Phi$ :

$$\frac{\partial}{\partial t} G_{\Phi}(v_{\lambda}) - \operatorname{div}(g(w_{\lambda})\nabla w_{\lambda}) = -g'(w_{\lambda})|\nabla w_{\lambda}|^2 - (g(w_{\lambda}) - g(\Phi(v_{\lambda}))) \frac{w_{\lambda} - \Phi(v_{\lambda})}{\lambda} \quad (10)$$

stated in [Evans and Portilheiro 2004; Evans 2004; Plotnikov 1996].  $\square$

## 4. Study of capillary effects

**4.1. Generalized “entropic” solutions.** It follows from the uniform estimates of Proposition 1 that we can find subsequences  $\{v_{\lambda_k}\}$  and  $\{w_{\lambda_k}\}$  and a pair<sup>7</sup>  $(v, w)$  such that, as  $\lambda_k \rightarrow 0$ ,

$$v_{\lambda_k} \rightharpoonup v \quad \text{in } L^{\infty}(Q) \text{ weakly-}^*, \quad (11)$$

$$\frac{\partial v_{\lambda_k}}{\partial t} \rightharpoonup \frac{\partial v_{\lambda}}{\partial t} \quad \text{in } L^2([0, T]; (H^1(\Omega))') \text{ weakly,} \quad (12)$$

$$w_{\lambda_k} \rightharpoonup w \quad \text{in } L^{\infty}(Q) \text{ weakly-}^* \text{ and in } L^2([0, T]; H^1(\Omega)) \text{ weakly,} \quad (13)$$

<sup>7</sup>A vanishing viscosity limit.

and

$$w_{\lambda_k} - \Phi(v_{\lambda_k}) \rightarrow 0 \quad \text{in } L^p(Q) \text{ strongly, for any finite } p.$$

Furthermore, we can assume that<sup>8</sup>

$$v_{\lambda_k} \rightarrow v \quad \text{in } C^0([0, T]; H^1(\Omega)') \text{ strongly.} \quad (14)$$

The associated Cauchy condition is given by

$$\begin{cases} v(0, \cdot) = S_{w(0)} & \text{in } H^1(\Omega)', \text{ a.e. in } \Omega, \\ 0 \leq v(t, \cdot) \leq 1 & \text{in } \Omega, t > 0, \\ \int_{\Omega} v(t, x) dx = \int_{\Omega} S_{w(0)}(x) dx, & t > 0. \end{cases}$$

In addition, the pair<sup>9</sup>  $(v, w)$  belongs to the functional frame

$$\begin{cases} v \in L^\infty(Q) \cap C^0([0, T]; H^1(\Omega)'), \\ \frac{\partial v}{\partial t} \in L^2([0, T]; H^1(\Omega)'), \\ w \in L^\infty(Q) \cap L^2([0, T]; H^1(\Omega)), \end{cases}$$

and is a solution of the boundary value problem

$$\begin{cases} \frac{\partial v}{\partial t} - \Delta w = 0 & \text{in } \mathcal{D}'(Q) \text{ and } L^2([0, T]; H^1(\Omega)'), \\ \frac{\partial w}{\partial n} = 0 & \text{on } \Sigma, \\ v(0, \cdot) = S_{w(0)} & \text{a.e. in } \Omega. \end{cases}$$

Because of the nonmonotonicity of  $\Phi$ , the information (11)–(14) is not sufficient to conclude that  $w = \Phi(v)$ , as we will see in what follows.

Let us denote by  $\Xi$  the complementary of the set of points of  $\mathcal{L}^{d+1}$ -approximate continuity of  $v$  according to the rigorous definition of the shock wave [Evans and Gariépy 1992; Gagneux and Madaune-Tort 1995]. The set  $\Xi$  is Borelian and  $\mathcal{L}^{d+1}$ -negligible because  $v$  is in  $L^\infty(Q)$ .

Let us assume that  $\Xi$  is a countable union of smooth hypersurfaces  $\Xi^i$  of  $\mathbb{R}^{d+1}$  which admit a unit normal vector  $v^i = (v_1^i, \dots, v_d^i, v_{d+1}^i) = (\tilde{v}^i, v_{d+1}^i)$ .

Using the usual notations for jumps in hyperbolic scalar laws and for the Hausdorff measure  $\mathcal{H}^d$ , very informally, the pair  $(v, w)$  satisfies the Rankine–Hugoniot and entropy conditions for all  $i$ , integrating by parts locally in a vicinity of a given transition interface via appropriate smooth functions with compact support:

$$\begin{aligned} v_{d+1}^i[v] &= \tilde{v}^i \cdot [\nabla w] \quad \text{and} \quad [w] = 0 \quad \mathcal{H}^d\text{-a.e. on } \Xi^i, \\ v_{d+1}^i[G_\Phi(v)] - \tilde{v}^i \cdot [\nabla w]g(w) &\leq 0 \quad \mathcal{H}^d\text{-a.e. on } \Xi^i. \end{aligned} \quad (15)$$

<sup>8</sup>From a classical compactness result of J. A. Dubinskii [Lions 1969, pp. 141–142].

<sup>9</sup>The pair  $(v, w)$  is called a *generalized solution* of the problem  $(P_\Phi)$ .

Relation (15) may be written in the form

$$v_{d+1}^i([G_\Phi(v)] - g(w)[v]) \leq 0 \quad \mathcal{H}^d\text{-a.e. on } \Xi^i \quad (16)$$

with the notations (8) for the definition of  $g$  (entropy function) and  $G_\Phi$ . In this form, relation (16) will be very useful to highlight the further developments.

**4.2. Associated hysteresis effects.** The analysis of hysteresis effects relies on the following proposition:

**Proposition 2.** *There exist three  $\mathcal{L}^{d+1}$ -measurable and bounded functions,  $\Lambda_0$ ,  $\Lambda_1$  and  $\Lambda_2$ , that are representative of the respective influence of the three branches of the graph of  $\Phi^{-1}$  (in the sense of the set theory) through the functions  $\beta_0$ ,  $\beta_1$  and  $\beta_2$ . Moreover, we have*

$$0 \leq \Lambda_i \leq 1 \quad \text{and} \quad \sum_{i=0}^2 \Lambda_i = 1 \quad \mathcal{L}^{d+1}\text{-a.e. in } Q.$$

In addition, when  $\lambda_k$  tends to  $0^+$ ,

$$\mu(v_{\lambda_k}) \text{ converges to } \sum_{i=0}^2 \Lambda_i \mu(\beta_i(w)) \text{ in } L^\infty(Q) \text{ weakly-}^*$$

for any numerical continuous function  $\mu$ . In addition, we have the following strong convergences:

$$w_{\lambda_k} \text{ and } \Phi(v_{\lambda_k}) \text{ converge to } w \text{ in } L^p(Q) \text{ strongly for any finite } p.$$

Finally, for any Lipschitz nondecreasing function  $g$ , we have the entropy relation in the sense of the measures in  $Q$

$$\frac{\partial}{\partial t} \left( \sum_{i=0}^2 \Lambda_i G_\Phi(\beta_i(w)) \right) \leq \operatorname{div}(g(w)\nabla w) - g'(w)|\nabla w|^2 \quad \text{in } Q. \quad (17)$$

*Proof.* The difficult technical proof of this proposition is not detailed here and can be found in [Evans and Portilheiro 2004; Evans 2004; Plotnikov 1996; 1994] with some adjustments.  $\square$

When  $\Lambda_0$  is equal to zero everywhere,<sup>10</sup> the method provides a response corresponding to the initial problem, thanks to the information contained in the complementary entropy relation (17). That is the main goal of the following proposition, based on the complementary information on the entropy given by (15)–(16), which enables us to determine the sense of circulation of the hysteresis loop.

<sup>10</sup>That corresponds in the final result to the neutralization of the decreasing part of the cubic introduced artificially to create a repulsive region.

**Proposition 3.** *In the framework of Proposition 2, we assume that*

$$\Lambda_0 = 0 \text{ in } Q, \quad \Lambda_1 = 1 \text{ in } Q_1 \quad \text{and} \quad \Lambda_2 = 1 \text{ in } Q_2,$$

where  $Q_1$  and  $Q_2$  are two open subsets of the cylinder  $Q$ , with a Lipschitz interface  $\Sigma_{1,2} = \overline{Q_1} \cap \overline{Q_2}$  admitting a unit normal vector  $v = (v_1, \dots, v_d, v_{d+1})$  of  $\mathbb{R}^d \times \mathbb{R}$ , oriented into  $Q_1$ .

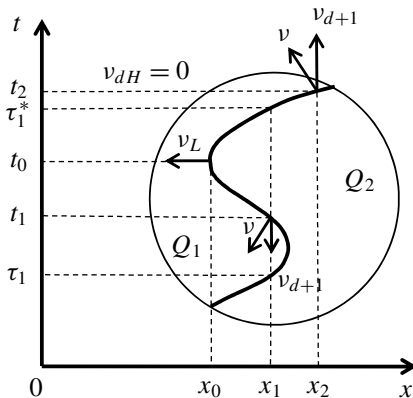
Using the notations of Figure 2, the problem can be written more precisely as a problem with free surface:

$$\begin{cases} v = \beta_1(w) \text{ and } \frac{\partial v}{\partial t} - \Delta \Phi_1(v) = 0 & \text{in } Q_1, \\ v = \beta_2(w) \text{ and } \frac{\partial v}{\partial t} - \Delta \Phi_2(v) = 0 & \text{in } Q_2. \end{cases} \quad (18)$$

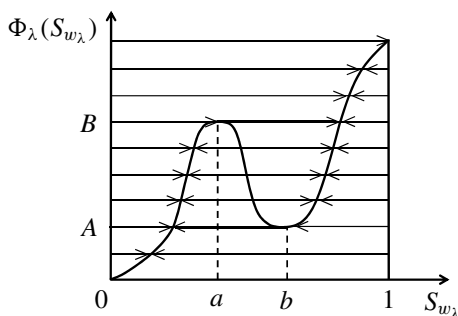
As a consequence of the information contained in the Rankine–Hugoniot and entropy relations (15)–(16), which are justified here along the shock wave  $\Sigma_{1,2}$ , the sign of the component  $v_{d+1}$  of the normal vector  $v$ , i.e., its orientation during the time, is specified by the relations

$$\begin{cases} v_{d+1} = 0 & \text{if } A < w < B, \\ v_{d+1} \geq 0 & \text{if } w = A, \\ v_{d+1} \leq 0 & \text{if } w = B. \end{cases} \quad (19)$$

**4.3. Interpretation of the results.** According to Proposition 3, the expected hysteresis effect is well described by the pair  $(v, w)$ , the generalized solution of  $(P_\Phi)$ . The change of the expression of the state law, which governs the diffusion process according to the values of the reduced saturation, is given by (18). Moreover, relation (19) reveals that the interface  $\Sigma_{1,2}$  evolves only if  $w$  takes the value  $A$  or  $B$  (see Figure 1).



**Figure 3.** More general case illustrating Proposition 3.



**Figure 4.** Flow of the auxiliary ordinary differential equation and emergence of the hysteresis effects when  $\lambda \rightarrow 0$ .

To illustrate the resulting hysteretic behavior, let us consider the generic example of [Figure 3](#) on the previous page, zooming on a time interval representing three possible states of the point  $x_1$  at three different times. We focus on the possible states corresponding to the abscissa  $x_1$ .

At the point  $(t_1, x_1)$ , we have a jump from  $Q_1$  to  $Q_2$ ,  $w = B$ , and we are in the imbibition phase (see also [Figure 1](#)). On the contrary, at the point  $(t_2, x_2)$ , we have a jump from  $Q_2$  to  $Q_1$ ,  $w = A$ , and we are in the drainage phase.

Therefore, the entropy method linked to the Sobolev regularization leads to a hysteresis loop similar to that obtained for Stefan’s supercooling problem [[Evans 2004](#)]. The flow of the auxiliary ordinary differential equation (3) leads to hysteresis effects when  $\lambda \rightarrow 0$  ([Figure 4](#)).

## 5. Conclusion

The hysteresis phenomena of flows in unsaturated porous media has been modeled with success, using the artificial introduction of an unstable spinodal interval and on a dynamic regularization process of Sobolev type.

## References

- [Aubin and Cellina 1984] J.-P. Aubin and A. Cellina, *Differential inclusions: set-valued maps and viability theory*, Grundlehren der Math. Wissenschaften **264**, Springer, Berlin, 1984.
- [Evans 2004] L. C. Evans, “A survey of entropy methods for partial differential equations”, *Bull. Amer. Math. Soc. (N.S.)* **41**:4 (2004), 409–438.
- [Evans and Gariépy 1992] L. C. Evans and R. F. Gariépy, *Measure theory and fine properties of functions*, CRC Press, Boca Raton, FL, 1992.
- [Evans and Portilheiro 2004] L. C. Evans and M. Portilheiro, “Irreversibility and hysteresis for a forward–backward diffusion equation”, *Math. Models Methods Appl. Sci.* **14**:11 (2004), 1599–1620.
- [Gagneux and Madaune-Tort 1995] G. Gagneux and M. Madaune-Tort, *Analyse mathématique de modèles non linéaires de l’ingénierie pétrolière*, Mathématiques & Applications **22**, Springer, Berlin, 1995.

- [Gagneux and Millet 2015] G. Gagneux and O. Millet, “A model for capillary hysteresis loops in unsaturated porous media theory”, 2015. Submitted to *Math. Mech. Solids*.
- [Lions 1969] J.-L. Lions, *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod, Paris, 1969.
- [Lions and Magenes 1968] J.-L. Lions and E. Magenes, *Problèmes aux limites non homogènes et applications, I*, Travaux et Recherches Mathématiques **17**, Dunod, Paris, 1968.
- [Plotnikov 1994] P. I. Plotnikov, “Passage to the limit with respect to viscosity in an equation with a variable direction of parabolicity”, *Differentsial'nye Uravneniya* **30**:4 (1994), 665–674, 734. In Russian; translated in *Differential Equations* **30**:4 (1994), 614–622.
- [Plotnikov 1996] P. I. Plotnikov, “Forward–backward parabolic equations and hysteresis”, pp. 183–209, 257–258 in *Boundary-value problems of mathematical physics and related problems of function theory, XXVII*, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) **233**, 1996. Reprinted in *J. Math. Sci.* **93**:5 (1999), 747–766.
- [Smarrazzo 2008] F. Smarrazzo, “On a class of equations with variable parabolicity direction”, *Discrete Contin. Dyn. Syst.* **22**:3 (2008), 729–758.
- [Smarrazzo and Tesei 2010] F. Smarrazzo and A. Tesei, “Long-time behavior of solutions to a class of forward–backward parabolic equations”, *SIAM J. Math. Anal.* **42**:3 (2010), 1046–1093.
- [Smarrazzo and Tesei 2012] F. Smarrazzo and A. Tesei, “Degenerate regularization of forward–backward parabolic equations: the regularized problem”, *Arch. Ration. Mech. Anal.* **204**:1 (2012), 85–139.

Received 10 Sep 2015. Revised 3 Dec 2015. Accepted 15 Jan 2016.

GÉRARD GAGNEUX: [gerard.maryse.gagneux@gmail.com](mailto:gerard.maryse.gagneux@gmail.com)

Laboratoire des Sciences de l'Ingénieur pour l'Environnement, UMR-CNRS 7356,  
Université de La Rochelle, avenue Michel Crépeau, 17042 La Rochelle Cedex 1, France

OLIVIER MILLET: [olivier.millet@univ-lr.fr](mailto:olivier.millet@univ-lr.fr)

Laboratoire des Sciences de l'Ingénieur pour l'Environnement, UMR-CNRS 7356,  
Université de La Rochelle, avenue Michel Crépeau, 17042 La Rochelle Cedex 1, France









# DISCRETE DOUBLE-POROSITY MODELS FOR SPIN SYSTEMS

ANDREA BRAIDES, VALERIA CHIADÒ PIAT AND MARGHERITA SOLCI

We consider spin systems between a finite number  $N$  of “species” or “phases” partitioning a cubic lattice  $\mathbb{Z}^d$ . We suppose that interactions between points of the same phase are coercive while those between points of different phases (or possibly between points of an additional “weak phase”) are of lower order. Following a discrete-to-continuum approach, we characterize the limit as a continuum energy defined on  $N$ -tuples of sets (corresponding to the  $N$  strong phases) composed of a surface part, taking into account homogenization at the interface of each strong phase, and a bulk part that describes the combined effect of lower-order terms, weak interactions between phases, and possible oscillations in the weak phase.

## 1. Introduction

In this paper, we consider lattice spin energies mixing strong ferromagnetic interactions and weak (possibly antiferromagnetic) pair interactions. The geometry that we have in mind is a periodic system of interactions such as that whose periodicity cell is represented in [Figure 1](#). In that picture, the strong interactions between nodes of the lattice (circles) are represented by solid lines and weak ones by dashed lines. In this particular case, we have two three-periodic systems of “strong sites”, i.e., sites connected by strong interactions, and isolated “weak sites” (pictured as white circles). Note that we may also have one or more infinite systems of connected weak interactions as in [Figure 2](#). In a discrete environment, the topological requirements governing the interactions between the strong and weak phases characteristic of continuum high-contrast models are substituted with assumptions on long-range interactions. In particular, contrary to the continuum case, for discrete systems with second-neighbor (or longer-range) interactions, we may have a limit multiphase system even in dimension 1 (see the examples in [Section 6](#)).

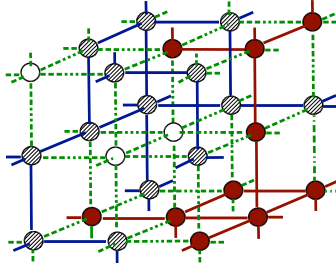
This paper is part of a general study of spin systems by means of variational techniques through the computation of continuum approximate energies, for which

**Communicated by Raffaele Esposito.**

The authors are members of INdAM-GNAMPA.

*MSC2010:* 49M25, 39A12, 39A70, 35Q82.

*Keywords:* spin systems, lattice energies, double porosity,  $\Gamma$ -convergence, homogenization, discrete to continuum, high contrast, interfacial energies, multiphase materials.



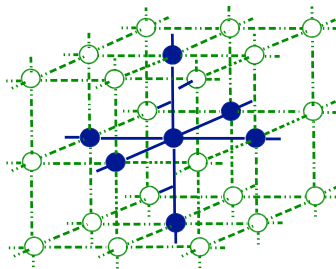
**Figure 1.** Picture of a double-porosity system.

homogenization results have been proved in the ferromagnetic case (i.e, when all interactions are strong) [Caffarelli and de la Llave 2005; Braides and Piatnitski 2013], and a general discrete-to-continuum theory of representation and optimization has been developed (see the survey [Braides 2014a]). In particular, a discrete-to-continuum compactness result and an integral representation of the limit by means of surface energies defined on sets of finite perimeter have been proved [Alicandro and Gelli 2016]. In that result, the coercivity of energies is obtained by assuming that nearest neighbors are always connected through a chain of strong interactions. Double-porosity systems can be interpreted as energies for which this condition does not hold but is satisfied separately on (finitely many) infinite connected components.

We are going to consider energies defined on functions parametrized on the cubic lattice  $\mathbb{Z}^d$  of the form

$$F_\varepsilon(u) = \sum_{(\alpha, \beta) \in \varepsilon \mathcal{N}_1 \cap (\Omega \times \Omega)} \varepsilon^{d-1} a_{\alpha\beta}^\varepsilon (u_\alpha - u_\beta)^2 + \sum_{(\alpha, \beta) \in \varepsilon \mathcal{N}_0 \cap (\Omega \times \Omega)} \varepsilon^d a_{\alpha\beta}^\varepsilon (u_\alpha - u_\beta)^2 + \sum_{\alpha \in \Omega \cap \varepsilon \mathbb{Z}^d} \varepsilon^d g(u_\alpha), \quad (1)$$

where  $\Omega$  is a regular open subset of  $\mathbb{R}^d$  and  $u_\alpha \in \{-1, +1\}$  denote the values of a spin function. For explanatory purposes, in this formula and the rest of the



**Figure 2.** A double-porosity system with an infinite connected weak component.

introduction, we use a simplified notation with respect to the rest of the paper, defining  $u = \{u_\alpha\}$  on the nodes of  $\Omega \cap \varepsilon \mathbb{Z}^d$  (instead of, equivalently, on the nodes of  $(1/\varepsilon)\Omega \cap \mathbb{Z}^d$ ). We denote by  $\mathcal{N}_1$  the set of pairs of nodes in  $\mathbb{Z}^d \times \mathbb{Z}^d$  between which we have strong interactions and by  $\mathcal{N}_0$  the set of pairs in  $\mathbb{Z}^d \times \mathbb{Z}^d$  between which we have weak interactions. The difference between these two types of interactions in the energy is the scaling factor:  $\varepsilon^{d-1}$  for strong interactions and  $\varepsilon^d$  for weak interaction. We suppose that all coefficients are obtained by scaling fixed coefficients on  $\mathbb{Z}^d$ , i.e.,

$$a_{\alpha\beta}^\varepsilon = a_{\alpha/\varepsilon, \beta/\varepsilon} \quad \text{if } \alpha, \beta \in \varepsilon \mathbb{Z}^d, \quad (2)$$

and  $a_{jk}$  are periodic of some integer period  $T$ . Moreover, we assume that the coefficients of the strong interactions are strictly positive, i.e.,  $a_{jk} > 0$  if  $(j, k) \in \mathcal{N}_1$ . The “forcing” term containing  $g$  and depending only on the point values  $u_\alpha$  is of lower order with respect to strong interactions but of the same order as the weak interactions.

We suppose that there are  $N$  infinite connected components of the graph of points linked by strong interactions, which we denote by  $C_1, \dots, C_N$ . Note that weak interactions in  $\mathcal{N}_0$  are due either to the existence of “weak sites” or to weak bonds between different “strong components” and, if we have more than one strong graph, the interactions in  $\mathcal{N}_0$  are present also in the absence of a weak component. We will describe the asymptotic behavior of energies (1) using the notation and techniques of  $\Gamma$ -convergence (see, e.g., [Braides 2002; 2006]).

If we consider only the strong interactions restricted to each strong connected component  $C_j$ , we obtain energies

$$F_\varepsilon^j(u) = \sum_{(\alpha, \beta) \in \varepsilon \mathcal{N}_1^j \cap (\Omega \times \Omega)} \varepsilon^{d-1} a_{\alpha\beta}^\varepsilon (u_\alpha - u_\beta)^2, \quad (3)$$

where  $\mathcal{N}_1^j$  is the restriction to  $C_j \times C_j$  of the set  $\mathcal{N}_1$ . This is a discrete analog of an energy on a *perforated domain*, the perforation being  $\mathbb{Z}^d \setminus C_j$ .

We prove an extension lemma that allows us to define for each  $j \in \{1, \dots, N\}$  a *discrete-to-continuum convergence* of (the restriction to  $C_j$  of) a sequence of functions  $u^\varepsilon$  to a function  $u^j \in \text{BV}(\Omega; \{\pm 1\})$ , which is compact under an equi-boundedness assumption for the energies  $F_\varepsilon^j(u^\varepsilon)$ . Thanks to this lemma, such energies behave as ferromagnetic energies with positive coefficients on the whole of  $\mathbb{Z}^d$ , which can be *homogenized* thanks to [Braides and Piatnitski 2013]; i.e., their  $\Gamma$ -limit with respect to the convergence  $u^\varepsilon \rightarrow u^j$  exists and is of the form

$$F^j(u^j) = \int_{S(u^j) \cap \Omega} f_{\text{hom}}^j(v_{u^j}) d\mathcal{H}^{d-1} \quad (4)$$

where  $S(u^j)$  is the set of jump points of  $u^j$ , which can also be interpreted as the interface between  $\{u^j = 1\}$  and  $\{u^j = -1\}$ .

Taking into account separately the restrictions of  $u^\varepsilon$  to all of the components  $C_j$ , we define a *vector-valued* limit function  $u = (u^1, \dots, u^N)$  and a convergence  $u^\varepsilon \rightarrow u$  and consider the  $\Gamma$ -limit of the whole energy with respect to that convergence. The combination of the weak interactions and the forcing term gives rise to a term of the form

$$\int_{\Omega} \varphi(u) dx$$

depending on the values of all components of  $u$ . In the case that  $\bigcup_{j=1}^N C_j$  is all of  $\mathbb{Z}^d$ , the function  $\varphi(z^1, \dots, z^N)$  is simply computed as the average of the  $T$ -periodic function

$$i \mapsto \sum_{k \in \mathbb{Z}^d} a_{ik} (u_i - u_k)^2 + g(u_i)$$

where  $u$  takes the value  $z^j$  on  $C_j$ . Note that with this condition only (weak) interactions between different  $C_j$  are taken into account. Note moreover that the restriction of the last term  $g$  to  $\varepsilon C_j$  is continuously converging to

$$K_j \int_{\Omega} g(u^j) dx,$$

where  $K_j = T^{-d} \#\{i \in C_j : i \in \{0, \dots, T\}^d\}$  is the percentage of sites in  $C_j$ . In general,  $\varphi$  is obtained by optimizing the combined effect of weak pair interactions and  $g$  on the free sites in the complement of all  $C_j$ .

Such different interactions can be summed up to describe the  $\Gamma$ -limit of  $F_\varepsilon$  that finally takes the form

$$F_{\text{hom}}(u) = \int_{S(u) \cap \Omega} f_{\text{hom}}(u^+, u^-, v_u) d\mathcal{H}^{d-1} + \int_{\Omega} \varphi(u) dx, \quad (5)$$

where  $f_{\text{hom}}(u^+, u^-, v) = \frac{1}{2} \sum_{j=1}^N f_{\text{hom}}^j(v) |u_j^+ - u_j^-|$ .

We note that the presence of two terms of different dimensions in the limit highlights the combination of bulk homogenization effects due to periodic oscillations besides the optimization of the interfacial structure. The effect of those oscillations on the variational motions of such systems (in the sense of [Ambrosio et al. 2008; Braides 2014b]) is addressed in [Braides and Solci 2015]. With respect to [Braides et al. 2015], we remark that the case of spin systems allows a very easy proof of an extension lemma from connected discrete sets and at the same time permits us to highlight the possibility to include a weak phase with antiferromagnetic interactions, optimized by microscopic oscillations.

Discrete problems modeling high-contrast media in the case of elastic energies have recently been considered in [Braides et al. 2015], but double-porosity homogenization in a continuum framework is a long-standing issue. The interest in double-porosity systems came at first from geophysics. The notion of double porosity, or double permeability, is borne from studies carried out on naturally fractured porous rocks such as oil fields. The benefits of describing oil flow and stock capacity in these kinds of soils justified theoretical studies undertaken during the 1960s. The double-porosity model was first introduced by [Barenblatt et al. 1960], and it has been used since in a wide range of engineering specialties. The first rigorous mathematical result on the subject was obtained in [Arbogast et al. 1990], where a linear parabolic equation with asymptotically degenerating coefficients was considered. This result was subsequently generalized in [Panassenko 1991; Bourgeat et al. 1996; 1998; 1999; Sandrakov 1999a; 1999b; Pankratov and Piatnitski 2002; Marchenko and Khruslov 2006] also for nonperiodic domains and various rates of contrast. On the physical level of rigor, double-porosity models were studied in [Panfilov 2000]. Linear double-porosity models with thin fissures were considered in [Pankratov and Rybalko 2003; Amaziane et al. 2009b]. The singular double-porosity model was considered in [Bourgeat et al. 2003]. The works [Bourgeat et al. 1999; Marchenko and Khruslov 2006; Pankratov and Rybalko 2003; Amaziane et al. 2009b] are carried out in the framework of Khruslov's mesoscopic energy characteristic methods. In addition, note that the double-porosity model was also obtained using the two-scale convergence method in [Hornung 1997]. Elliptic and parabolic nonlinear double-porosity models, including homogenization in variable Sobolev spaces, were also obtained in [Pankratov et al. 2003; Amaziane et al. 2006; 2009a; Choquet and Pankratov 2010]. Finally, the double-porosity models of multiphase flows, including the nonequilibrium ones, were also obtained in [Choquet 2004; Yeh 2006; Amaziane and Pankratov 2016; Konyukhov and Pankratov 2015] (see also [Hornung 1997] and the references therein). A reformulation in terms of  $\Gamma$ -convergence can be found in [Braides et al. 2004] with related results for nonconvex integrands. An approach using  $\Gamma$ -convergence and a two-scale formulation at the same time is given in [Cherdantsev and Cherednichenko 2012]. Double-porosity models for interfacial energies on the continuum were previously examined in [Solci 2009; 2012; Braides and Solci 2013].

The results in the present paper may be regarded as a geometrically simplified model of continuum ones (but with more freedom in the lattice interactions), but the same framework may also be useful for other discrete models actually developed in mechanics. Among them are pantographic systems made of beams and used for modeling of some metamaterials [Seppecher et al. 2011] and investigations of two- and three-dimensional lattices in order to develop models used in nano- and micromechanics.

The plan of the paper is the following. In [Section 2](#), we introduce the geometric setting, identifying the “strong” and possibly “weak” phases of the lattice network, and define the microscopic energy. In [Section 3](#), we prove a compactness theorem and a homogenization result for each separate strong phase. The resulting energies will provide the interfacial energy part of the limit. In [Section 4](#), we define the interaction term between the strong phases by proving an asymptotic formula. The main convergence result is stated and proved in [Section 5](#), where the compactness theorem in [Section 3](#) applied to each strong phase is used to define a multiphase limit. Finally in [Section 6](#), some simple examples are provided, which in particular also exhibit nontrivial limits in dimensions 1 and 2.

## 2. Notation

The numbers  $d, m, T$ , and  $N$  are positive integers. We introduce a  $T$ -periodic *label function*  $J : \mathbb{Z}^d \rightarrow \{0, 1, \dots, N\}$  and the corresponding sets of sites

$$A_j = \{k \in \mathbb{Z}^d : J(k) = j\}, \quad j = 0, \dots, N.$$

Sites interact through possibly long- (but finite-)range interactions, whose range is defined through a system  $P^j = \{P_k^j\}$  of finite subsets  $P_k^j \subset \mathbb{Z}^d$  for  $j = 0, \dots, N$  and  $k \in A_j$ . We suppose

- (*T-periodicity*)  $P_{k+m}^j = P_k^j$  for all  $m \in T\mathbb{Z}^d$  and
- (*symmetry*) if  $k \in A_j$  for  $j = 1, \dots, N$  (*hard components*) and  $i \in P_k^j$ , then  $k+i \in A_j$  and  $-i \in P_{k+i}^j$ , and  $0 \in P_k^j$ .

We say that two points  $k, k' \in A_j$  are  $P^j$ -*connected* in  $A_j$  if there exists a path  $\{k_n\}_{n=0, \dots, K}$  such that  $k_n \in A_j$ ,  $k_0 = k$ ,  $k_K = k'$ , and  $k_n - k_{n-1} \in P_{k_{n-1}}^j$ .

We suppose

- (*connectedness*) there exists a unique infinite  $P^j$ -connected component of each  $A_j$  for  $j = 1, \dots, N$ , which we denote by  $C_j$ .

Clearly, the connectedness assumption is not a modeling restriction upon introducing more labeling parameters if the number of infinite connected components is finite. Note that we do not make any assumptions on  $A_0$  and  $P^0$ . In particular, if  $k \in A_j$  for  $j = 0, \dots, N$  and  $i \in P_k^0$ , then  $k+i$  may belong to any  $A_{j'}$  with  $j' \neq j$ .

We consider the following sets of bonds between sites in  $\mathbb{Z}^d$ : for  $j = 1, \dots, N$

$$N_j = \{(k, k') : k, k' \in A_j, k' - k \in P_k^j \setminus \{0\}\}$$

and for  $j = 0$

$$N_0 = \{(k, k') : k' - k \in P_k^0 \setminus \{0\}, J(k)J(k') = 0 \text{ or } J(k) \neq J(k')\}.$$

Note that the set  $N_0$  takes into account interactions not only among points of the set  $A_0$  but also among pairs of points in different  $A_j$ . More refined notation could be introduced by defining a range of interactions  $P^{ij}$  and the corresponding sets  $N_{ij}$ , in which case the sets  $N_j$  would correspond to  $N_{jj}$  for  $j = 1, \dots, N$  and  $N_0$  to the union of the remaining sets. However, for simplicity of presentation, we limit our notation to a single index.

We consider interaction energy densities associated with positive numbers  $a_{kk'}$  for  $k, k' \in \mathbb{Z}^d$  and the forcing term  $g$ . We suppose that for all  $k, k' \in \mathbb{Z}^d$

- (*coerciveness on the hard phase*) there exists  $c > 0$  such that  $a_{kk'} \geq c > 0$  if  $k \in C_j$  and  $k' - k \in P_k^j$  for  $j \geq 1$ ,
- (*T-periodicity*)  $a_{k+m, k'+m} = a_{kk'}$  for all  $m \in T\mathbb{Z}^d$ ,
- (*symmetry*)  $a_{k'k} = a_{kk'}$ , and
- (*T-periodicity of the forcing term*)  $g(k+m, 1) = g(k, 1)$  and  $g(k+m, -1) = g(k, -1)$  for all  $m \in T\mathbb{Z}^d$ .

Note that we do not suppose that the  $a_{kk'}$  are positive for weak interactions. They can be negative as well, thus favoring oscillations in the weak phase.

Given  $\Omega$ , a bounded regular open subset of  $\mathbb{R}^d$ , for  $u : (1/\varepsilon)\Omega \cap \mathbb{Z}^d \rightarrow \{+1, -1\}$ , we define the energies

$$F_\varepsilon(u) = F_\varepsilon\left(u, \frac{1}{\varepsilon}\Omega\right) = \sum_{j=1}^N \sum_{(k, k') \in \mathcal{N}_j^\varepsilon(\Omega)} \varepsilon^{d-1} a_{kk'} (u_k - u_{k'})^2 + \sum_{(k, k') \in \mathcal{N}_0^\varepsilon(\Omega)} \varepsilon^d a_{kk'} (u_k - u_{k'})^2 + \sum_{k \in \mathbb{Z}^\varepsilon(\Omega)} \varepsilon^d g(k, u_k), \quad (6)$$

where

$$\mathcal{N}_j^\varepsilon(\Omega) = N_j \cap \frac{1}{\varepsilon}(\Omega \times \Omega), \quad j = 0, \dots, N, \quad \mathbb{Z}^\varepsilon(\Omega) = \mathbb{Z}^d \cap \frac{1}{\varepsilon}\Omega. \quad (7)$$

The first sum in the energy takes into account all interactions between points in  $A_j$  (*hard phases*), which are supposed to scale differently than those between points in  $A_0$  (*soft phase*) or between points in different phases. The latter are contained in the second sum. The third sum is a zero-order term taking into account all types of phases with the same scaling.

Note that the first sum may also take into account points in  $A_j \setminus C_j$ , which form “islands” of the hard phase  $P^j$ -disconnected from the corresponding infinite component. Furthermore, in this energy, we may have sites that do not interact at all with hard phases.

**Remark 2.1** (choice of the parameter space). The energy is defined on discrete functions parametrized on  $(1/\varepsilon)\Omega \cap \mathbb{Z}^d$ . The choice of this notation, rather than interpreting  $u$  as defined on  $\Omega \cap \varepsilon\mathbb{Z}^d$ , allows a much easier notation for the coefficients, which in this way are  $\varepsilon$ -independent rather than obtained by scaling as in (2).

### 3. Homogenization of perforated discrete domains

In this section, we separately consider the interactions in each infinite connected component of the hard phases introduced above. To that end, we fix one of the indices  $j$ , with  $j > 0$ , dropping it in the notation of this section (in particular, we use the symbol  $C$  in place of  $C_j$ , etc.), and define the energies

$$\mathcal{F}_\varepsilon(u) = \mathcal{F}_\varepsilon\left(u, \frac{1}{\varepsilon}\Omega\right) = \sum_{(k,k') \in N_C^\varepsilon(\Omega)} \varepsilon^{d-1} a_{kk'} (u_k - u_{k'})^2, \quad (8)$$

where

$$N_C^\varepsilon(\Omega) = \left\{ (k, k') \in (C \times C) \cap \frac{1}{\varepsilon}(\Omega \times \Omega) : k' - k \in P_k, k \neq k' \right\}. \quad (9)$$

We also introduce the notation  $C^\varepsilon(\Omega) = C \cap (1/\varepsilon)\Omega$ .

**Definition 3.1.** We define the *piecewise-constant interpolation* of a function  $u : \mathbb{Z}^d \cap (1/\varepsilon)\Omega \rightarrow \mathbb{R}^m$ ,  $k \mapsto u_k$ , as

$$u(x) = u_{\lfloor x/\varepsilon \rfloor},$$

where  $\lfloor y \rfloor = (\lfloor y_1 \rfloor, \dots, \lfloor y_d \rfloor)$  and  $\lfloor s \rfloor$  stands for the integer part of  $s$ . The *convergence* of a sequence  $(u^\varepsilon)$  of discrete functions is understood as the  $L^1_{\text{loc}}(\Omega)$  convergence of these piecewise-constant interpolations. Note that, since we consider local convergence in  $\Omega$ , the value of  $u(x)$  close to the boundary is not involved in the convergence process.

We prove an extension and compactness lemma with respect to the convergence of piecewise-constant interpolations.

**Lemma 3.2** (extension and compactness). *Let  $C$  be a  $T$ -periodic subset of  $\mathbb{Z}^d$   $P$ -connected in the notation of the previous section, and let  $u^\varepsilon : \mathbb{Z}^d \cap (1/\varepsilon)\Omega \rightarrow \{+1, -1\}$  be a sequence such that*

$$\sup_\varepsilon \varepsilon^{d-1} \#\{(k, k') \in N_C^\varepsilon(\Omega) : u_k^\varepsilon \neq u_{k'}^\varepsilon\} < +\infty. \quad (10)$$

*Then there exists a sequence  $\tilde{u}^\varepsilon : \mathbb{Z}^d \cap (1/\varepsilon)\Omega \rightarrow \mathbb{R}^m$  such that  $\tilde{u}_k^\varepsilon = u_k^\varepsilon$  if  $k \in C^\varepsilon(\Omega)$  and  $\text{dist}(k, \partial(1/\varepsilon)\Omega) > c = c(P)$  with  $\tilde{u}^\varepsilon$  converging to some  $u \in \text{BV}_{\text{loc}}(\Omega; \{+1, -1\})$  up to subsequences.*

*Proof.* For a fixed  $M \in \mathbb{N}$  and  $j \in \mathbb{Z}^d$ , we consider the discrete cubes of side length  $M$

$$Q_M(j) := jM + \{0, M-1\}^d.$$



For each  $j$ , we also define the cube

$$Q'_{3M}(j) = \bigcup_{\|i-j\|_\infty \leq 1} Q_M(i),$$

which is a discrete cube centered at  $Q_M(j)$  and with side length  $3M$ .

For all  $\varepsilon$ , we consider the family

$$\mathcal{Q}_M^\varepsilon := \left\{ Q_M(j) : j \in \mathbb{Z}^d, Q'_{3M}(j) \subset \frac{1}{\varepsilon} \Omega \right\}.$$

We suppose that  $M$  is large enough such that, if  $k, k' \in Q_M(j) \cap C$ , then there exists a  $P$ -path connecting  $k$  and  $k'$  contained in  $Q'_{3M}(j)$ . The existence of such  $M$  follows easily from the connectedness hypotheses. Indeed, we may take  $M$  as the length of the longest shortest  $P$ -path connecting two points in  $C$  with distance not greater than  $2\sqrt{d}$  (in particular belonging to neighboring periodicity cubes) and construct such a  $P$ -path by concatenating a family of those shortest paths.

We define the set of indices

$$\mathcal{S}_\varepsilon = \{j \in \mathbb{Z}^d : Q_M(j) \in \mathcal{Q}_M^\varepsilon \text{ and } u^\varepsilon \text{ is not constant on } C \cap Q_M(j)\}.$$

By our choice of  $M$ , if  $j \in \mathcal{S}_\varepsilon$ , then there exist  $k, k' \in Q'_{3M}(j) \cap C$  with  $k' - k \in P$  such that  $u_k^\varepsilon \neq u_{k'}^\varepsilon$ . Let

$$K := \sup_\varepsilon \varepsilon^{d-1} \#\{(k, k') \in N_C^\varepsilon(\Omega) : u_k^\varepsilon \neq u_{k'}^\varepsilon\}.$$

Then we deduce that

$$\#\mathcal{S}_\varepsilon \leq 3^d K \frac{1}{\varepsilon^{d-1}} \quad (11)$$

(the factor  $3^d$  comes from the fact that  $k, k' \in Q'_{3M}(j)$  for  $3^d$  possible  $j$ ).

We define

$$\tilde{u}^\varepsilon = \begin{cases} \text{constant value of } u^\varepsilon \text{ on } Q_M(j) \cap C & \text{on } Q_M(j) \text{ if } Q_M(j) \in \mathcal{Q}_M^\varepsilon \text{ and } j \notin \mathcal{S}_\varepsilon, \\ u^\varepsilon & \text{elsewhere.} \end{cases}$$

This will be the required extension. However, we will prove the convergence of  $\tilde{u}^\varepsilon$  as a consequence of the convergence of the functions

$$v^\varepsilon = \begin{cases} \tilde{u}^\varepsilon & \text{on } Q_M(j) \text{ if } Q_M(j) \in \mathcal{Q}_M^\varepsilon \text{ and } j \notin \mathcal{S}_\varepsilon, \\ 1 & \text{elsewhere.} \end{cases}$$

By (11), we have that for fixed  $\Omega' \Subset \Omega$

$$\|v^\varepsilon - \tilde{u}^\varepsilon\|_{L^1(\Omega')} = O(\varepsilon)$$

(recall that we identify the function with its scaled interpolations in  $L^1(\Omega)$ ).

If the value of  $v^\varepsilon$  differs on two neighboring  $Q_M(j)$  and  $Q_M(j')$  with  $\|j - j'\|_1 = 1$ , then upon taking a suitable larger  $M$ , we may also suppose that there exist

$k, k' \in (Q'_{3M}(j) \cup Q'_{3M}(j)) \cap C$  with  $k - k' \in P$  and  $u_k^\varepsilon \neq u_{k'}^\varepsilon$ . Arguing as for (11), we deduce that the number of such  $j$  is  $O(\varepsilon^{1-d})$  so that

$$\mathcal{H}^{d-1}(\partial\{v^\varepsilon = 1\} \cap \Omega') = O(1),$$

which implies the compactness of the family  $(v^\varepsilon)$  in  $\text{BV}_{\text{loc}}(\Omega)$ .  $\square$

The compactness theorem above proves that the domain of the limit is functions  $u \in \text{BV}(\Omega, \{+1, -1\})$ , which can be identified with the *sets of finite perimeter*  $E = \{u = 1\}$ . In this case, the set of discontinuity points  $S(u)$  coincides, up to sets of  $\mathcal{H}^{n-1}$ -measure 0, with the *reduced boundary*  $\partial^*\{u = 1\}$ , whose *inner normal* we denote by  $\nu$  [Braides 1998].

**Theorem 3.3** (homogenization on discrete perforated domains). *The energies  $\mathcal{F}_\varepsilon$  defined in (8)  $\Gamma$ -converge with respect to the  $L^1_{\text{loc}}(\Omega)$  topology to the energy*

$$\bar{\mathcal{F}}_{\text{hom}}(u) = \int_{\Omega \cap \partial^*\{u=1\}} f_{\text{hom}}(\nu) d\mathcal{H}^{d-1}, \quad (12)$$

defined on  $u \in \text{BV}(\Omega, \{+1, -1\})$ , where the energy density  $f_{\text{hom}}$  satisfies

$$f_{\text{hom}}(\nu) = \lim_{T \rightarrow +\infty} \frac{1}{T^{d-1}} \inf \left\{ \sum_{(k,k') \in \tilde{N}_C(Q_T^\nu)} a_{kk'} (u_k - u_{k'})^2 : u_k = \text{sign}\langle k, \nu \rangle \text{ if } k \notin Q_T^\nu \right\}, \quad (13)$$

where

$$\text{sign } x = \begin{cases} 1 & \text{if } x > 0, \\ -1 & \text{if } x \leq 0, \end{cases} \quad (14)$$

$Q_T^\nu$  is a cube centered at 0 and with one side orthogonal to  $\nu$ ,  $Q_T^\nu = T Q^\nu$ , and  $\tilde{N}_C(Q_T^\nu)$  denotes all pairs in  $(k, k') \in N_C^1(\mathbb{R}^d)$  such that either  $k \in Q_T^\nu$  or  $k' \in Q_T^\nu$ .

*Proof.* In [Braides and Piatnitski 2013], this theorem is proved under the additional assumption that the energies  $\mathcal{F}_\varepsilon$  are equicoercive with respect to the weak BV-convergence. This assumption can be substituted with Lemma 3.2. Indeed, if  $u^\varepsilon$  is a sequence converging to  $u$  in  $L^1_{\text{loc}}(\Omega)$  and with equibounded energies, then by Lemma 3.2, we may find a sequence  $\tilde{u}^\varepsilon$  coinciding with  $u^\varepsilon$  on  $C^\varepsilon(\Omega')$  for every fixed  $\Omega' \Subset \Omega$  and  $\varepsilon$  sufficiently small and converging to some  $\tilde{u}$  in  $\text{BV}(\Omega; \{\pm 1\})$ . Since  $\tilde{u}^\varepsilon = u^\varepsilon$  on  $C^\varepsilon(\Omega')$ , we have that  $\tilde{u} = u$  and  $\mathcal{F}_\varepsilon(\tilde{u}^\varepsilon, (1/\varepsilon)\Omega') = \mathcal{F}_\varepsilon(u^\varepsilon, (1/\varepsilon)\Omega')$ . Then we can give a lower estimate on each  $\Omega'$  fixed using the proof of [Braides and Piatnitski 2013] and hence on  $\Omega$  by internal approximation. Note that neither the proof of the existence of the limit in (13) therein nor the construction of the recovery sequences depends on the coerciveness assumption, so the proof is complete.  $\square$

#### 4. Definition of the interaction term

The homogenization result in [Theorem 3.3](#) will describe the contribution of the hard phases to the limiting behavior of energies  $F_\varepsilon$ . We now characterize their interactions with the soft phase.

For all positive integers  $M$  and  $z_1, \dots, z_N \in \{+1, -1\}$ , we define the minimum problem

$$\varphi_M(z_1, \dots, z_N) = \frac{1}{M^d} \min \left\{ \sum_{(k,k') \in N_0(Q_M)} a_{kk'} (v_k - v_{k'})^2 + \sum_{k \in Z(Q_M)} g(k, v_k) : v \in \mathcal{V}_M \right\},$$

where

$$Q_M = \left[ -\frac{M}{2}, \frac{M}{2} \right)^d, \quad N_0(Q_M) = N_0 \cap (Q_M \times Q_M), \quad Z(Q_M) = \mathbb{Z}^d \cap Q_M \quad (15)$$

and the minimum is taken over the set  $\mathcal{V}_M = \mathcal{V}_M(z_1, \dots, z_N)$  of all  $v$  constant on each connected component of  $A_j \cap Q_M$  and  $v = z_j$  on  $C_j$  for  $j = 1, \dots, N$ .

**Proposition 4.1.** *The limit  $\varphi$  of  $\varphi_M$  as  $M \rightarrow +\infty$  exists.*

*Proof.* We first show that

$$\varphi_{KM} \geq \varphi_M \quad \text{for all } K \in \mathbb{N}. \quad (16)$$

To that end, let  $\bar{v}$  be a minimizer for  $\varphi_{KM}(z_1, \dots, z_N)$ . Then we have

$$\begin{aligned} & K^d M^d \varphi_{KM}(z_1, \dots, z_N) \\ &= \sum_{(k,k') \in N_0(Q_{KM})} a_{kk'} (\bar{v}_k - \bar{v}_{k'})^2 + \sum_{k \in Z(Q_{KM})} g(k, \bar{v}_k) \\ &= \sum_{l \in \mathbb{Z}^d \cap Q_K} \left( \sum_{(k,k') \in N_0(Q_M + lM)} a_{kk'} (\bar{v}_k - \bar{v}_{k'})^2 + \sum_{k \in Z(Q_M + lM)} g(k, \bar{v}_k) \right) \\ &\quad + \sum_{(k,k') \in N_0(Q_{KM}) \setminus \bigcup_l N_0(Q_M + lM)} a_{kk'} (\bar{v}_k - \bar{v}_{k'})^2 \\ &\geq \sum_{l \in \mathbb{Z}^d \cap Q_K} \left( \sum_{(k,k') \in N_0(Q_M + lM)} a_{kk'} (\bar{v}_k - \bar{v}_{k'})^2 + \sum_{k \in Z(Q_M + lM)} g(k, \bar{v}_k) \right). \end{aligned}$$

Let  $\bar{l} \in \mathbb{Z}^d \cap Q_K$  minimize the expression in parentheses. Then we deduce

$$K^d M^d \varphi_{KM}(z_1, \dots, z_N) \geq K^d \left( \sum_{(k,k') \in N_0(Q_M + \bar{l}M)} a_{kk'} (\bar{v}_k - \bar{v}_{k'})^2 + \sum_{k \in Z(Q_M + \bar{l}M)} g(k, \bar{v}_k) \right),$$

from which (16) follows by taking  $v_k = \bar{v}_{k - \bar{l}M}$  in the computation of  $\varphi_M(z_1, \dots, z_N)$ .

We remark that for  $L \geq L'$  we have

$$L^d \varphi_L \geq (L')^d \varphi_{L'} - \max|g|(L^d - (L')^d). \quad (17)$$

Hence, fixing  $n$ ,  $L$ , and  $M$  with  $L \geq M2^n$  and taking  $L' = \lfloor L/(M2^n) \rfloor M2^n$  in (17), we have, using (16) with  $K = \lfloor L/(M2^n) \rfloor 2^n$

$$\begin{aligned} \varphi_L &\geq \frac{1}{L^d} \left( \left\lfloor \frac{L}{M2^n} \right\rfloor M2^n \right)^d \varphi_{\lfloor L/(M2^n) \rfloor M2^n} - \max|g| \left( 1 - \left( \left\lfloor \frac{L}{M2^n} \right\rfloor \frac{M2^n}{L} \right)^d \right) \\ &\geq \left( \left\lfloor \frac{L}{M2^n} \right\rfloor \frac{M2^n}{L} \right)^d \varphi_M - \max|g| \left( 1 - \left( \left\lfloor \frac{L}{M2^n} \right\rfloor \frac{M2^n}{L} \right)^d \right). \end{aligned}$$

Letting  $L \rightarrow +\infty$ , we then obtain

$$\liminf_{L \rightarrow +\infty} \varphi_L \geq \varphi_M$$

and the conclusion follows by taking the upper limit in  $M$ .  $\square$

Let  $R$  be defined by

$$R = \max\{|k - k'| : k, k' \in A_j \setminus C_j \text{ that are } P^j\text{-connected, } j = 1, \dots, N\}, \quad (18)$$

and for all  $M$  positive integer, set

$$D_M = \bigcup_{j=1}^N \bigcup \{P^j\text{-connected components } B \text{ of } A_j \setminus C_j \text{ not intersecting } Q_{M-R}\}.$$

For all  $z_1, \dots, z_N \in \{+1, -1\}$ , we define

$$\begin{aligned} \tilde{\varphi}_M(z_1, \dots, z_N) &= \frac{1}{M^d} \min \left\{ \sum_{(k,k') \in N_0(Q_M)} a_{kk'} (v_k - v_{k'})^2 + \sum_{k \in Z(Q_M)} g(k, v_k) \right. \\ &\quad \left. : v \in \mathcal{V}_M, v_k = 1 \text{ if } k \in D_M \right\}. \quad (19) \end{aligned}$$

**Proposition 4.2.** *There is a positive constant  $c$  independent of  $M$  such that*

$$\tilde{\varphi}_M \geq \varphi_M \geq \tilde{\varphi}_M - \frac{c}{M}. \quad (20)$$

*Proof.* The first inequality is trivial. To prove the second, let  $\bar{v}$  be a minimizer for  $\varphi_M(z_1, \dots, z_N)$  and define  $v$  by

$$v_k = \begin{cases} 1 & \text{if } k \in D_M \\ \bar{v}_k & \text{otherwise.} \end{cases}$$

Using  $v$  as a test function for  $\tilde{\varphi}_M(z_1, \dots, z_N)$ , we obtain

$$\begin{aligned}
M^d \tilde{\varphi}_M(z_1, \dots, z_N) &\leq \sum_{(k,k') \in N_0(Q_M), k, k' \notin D_M} a_{kk'} (v_k - v_{k'})^2 + \sum_{k \in Z(Q_M) \setminus D_M} g(k, v_k) \\
&\quad + 2 \sum_{(k,k') \in N_0(Q_M), k \in D_M} a_{kk'} (v_k - v_{k'})^2 + \sum_{k \in Z(Q_M) \cap D_M} g(k, v_k) \\
&\leq \sum_{(k,k') \in N_0(Q_M), k, k' \notin D_M} a_{kk'} (\bar{v}_k - \bar{v}_{k'})^2 + \sum_{k \in Z(Q_M) \setminus D_M} g(k, \bar{v}_k) \\
&\quad + \sum_{(k,k') \in N_0(Q_M), k \in D_M} a_{kk'} + \sum_{k \in Z(Q_M) \cap D_M} g(k, 1) \\
&\leq M^d \varphi_M(z_1, \dots, z_N) + \#D_M \#P_0 \max a_{ij} + \#D_M 2 \max |g|.
\end{aligned}$$

As  $\#D_M \leq 2^d M^{d-1} R$ , the result follows with  $c = 2^d R(\#P_0 \max a_{ij} + 2 \max |g|)$ .  $\square$

### 5. Statement of the convergence result

We now have all the ingredients to characterize the asymptotic behavior of  $F_\varepsilon$  defined in (6).

**Definition 5.1** (multiphase discrete-to-continuum convergence). We define the *convergence*

$$u^\varepsilon \rightarrow (u^1, \dots, u^N) \quad (21)$$

as the  $L^1_{\text{loc}}(\Omega; \mathbb{R}^m)$  convergence  $\tilde{u}_j^\varepsilon \rightarrow u^j$  of the extensions of the restrictions of  $u^\varepsilon$  to  $C_j$  as in Lemma 3.2, which is a compact convergence as ensured by that lemma.

The total contribution of the hard phases will be given separately by the contribution on the infinite connected components and the finite ones. The first one is obtained by independently computing the limit relative to the energy restricted to each component

$$\mathfrak{F}_\varepsilon^j(u) = \sum_{(k,k') \in N_j^\varepsilon(\Omega)} \varepsilon^{d-1} a_{kk'} (v_k - v_{k'})^2, \quad (22)$$

where

$$N_j^\varepsilon(\Omega) = N_{C_j}^\varepsilon(\Omega) = \left\{ (k, k') \in (C_j \times C_j) \cap \frac{1}{\varepsilon}(\Omega \times \Omega) : k - k' \in P_k^j, k \neq k' \right\}, \quad (23)$$

which is characterized by Theorem 3.3 as

$$\mathfrak{F}_{\text{hom}}^j(u) = \int_{\Omega \cap \partial^* \{u=1\}} f_{\text{hom}}^j(v) d\mathcal{H}^{d-1}. \quad (24)$$

In Section 4, we introduced the energy density  $\varphi$ , which describes the interactions between the hard phases. Taking all contributions into account, we may state the following convergence result.

**Theorem 5.2** (double-porosity homogenization). *Let  $\Omega$  be a Lipschitz bounded open set, and let  $F_\varepsilon$  be defined by (6) with the notation of Section 2. Then the  $\Gamma$ -limit of  $F_\varepsilon$  with respect to the convergence (21) exists, and it equals*

$$F_{\text{hom}}(u^1, \dots, u^N) = \sum_{j=1}^N \int_{\Omega \cap \partial^* \{u^j=1\}} f_{\text{hom}}^j(v) d\mathcal{H}^{d-1} + \int_{\Omega} \varphi(u^1, \dots, u^N) dx \quad (25)$$

on functions  $u = (u^1, \dots, u^N) \in (\text{BV}(\Omega; \{1, -1\}))^N$ , where  $\varphi$  is defined in Proposition 4.1 and  $f_{\text{hom}}^j$  are defined by (24).

Note that there is no contribution of the finite connected components of  $A_j$ .

**Remark 5.3** (nonhomogeneous lower-order term). In our hypotheses, the lower-order term  $g$  depends on the fast variable  $k$ , which is integrated out in the limit. We may easily include a measurable dependence on the slow variable  $\varepsilon k$  by assuming  $g = g(x, k, z)$  is a Carathéodory function (this covers in particular the case  $g = g(x, z)$ ) and substitute the last sum in (6) by

$$\sum_{k \in Z^\varepsilon(\Omega)} \varepsilon^d g(\varepsilon k, k, u_k).$$

Correspondingly, in Theorem 5.2, the integrand in the last term in (25) must be substituted by  $\varphi(x, u^1, \dots, u^N)$ , where the definition of this last function is the same but taking  $g(x, k, z)$  in place of  $g(k, z)$  so that  $x$  simply acts as a parameter.

The proof of Theorem 5.2 will be subdivided into a lower and an upper bound.

*Proof of the lower bound.* Let  $u^\varepsilon \rightarrow (u^1, \dots, u^N)$  be such that  $F_\varepsilon(u^\varepsilon) \leq c < +\infty$ . Fixing  $M \in \mathbb{N}$ , we introduce the notation

$$\begin{aligned} J_M^\varepsilon &= \left\{ z \in \mathbb{Z}^d : Q_M + zM \subset \frac{1}{\varepsilon} \Omega \right\}, \\ R^\varepsilon &= \mathcal{N}_0^\varepsilon(\Omega) \setminus \bigcup_{z \in J_M^\varepsilon} \mathcal{N}_0^\varepsilon(Q_M + zM), \\ S^\varepsilon &= Z^\varepsilon(\Omega) \setminus \bigcup_{z \in J_M^\varepsilon} Z(Q_M + zM) \end{aligned}$$

and write

$$F_\varepsilon(u^\varepsilon) = \sum_{j=1}^N \text{I}_j^\varepsilon + \text{II}^\varepsilon + \text{III}^\varepsilon + \text{IV}^\varepsilon + \text{V}^\varepsilon,$$

where

$$\text{I}_j^\varepsilon = \mathcal{F}_\varepsilon^j(u),$$

$$\begin{aligned}
\text{II}^\varepsilon &= \sum_{j=1}^N \sum_{(k,k') \in \mathcal{N}_j^\varepsilon(\Omega) \setminus (C_j \times C_j)} \varepsilon^{d-1} a_{kk'} (v_k - v_{k'})^2, \\
\text{III}^\varepsilon &= \sum_{z \in J_M^\varepsilon} \varepsilon^d \left( \sum_{(k,k') \in \mathcal{N}_0^\varepsilon(Q_M + zM)} a_{kk'} (v_k - v_{k'})^2 + \sum_{k \in Z(Q_M + zM)} g(k, v_k) \right), \\
\text{IV}^\varepsilon &= \sum_{(k,k') \in R^\varepsilon} \varepsilon^d a_{kk'} (v_k - v_{k'})^2, \\
\text{V}^\varepsilon &= \sum_{k \in S^\varepsilon} \varepsilon^d g(k, v_k).
\end{aligned}$$

Note that

$$\begin{aligned}
\text{II}^\varepsilon &\geq 0, \\
\text{IV}^\varepsilon &\geq -c/M + o(1), \\
\text{V}^\varepsilon &\geq -\max |g| \left( \left| \Omega \setminus \varepsilon^d \bigcup_{z \in J_M^\varepsilon} (Q_M + zM) \right| + o(1) \right),
\end{aligned} \tag{26}$$

where we have taken into account that the interactions in  $\text{IV}^\varepsilon$  may be negative and

$$\liminf_{\varepsilon \rightarrow 0} \sum_{j=1}^N \text{I}_j^\varepsilon \geq \sum_{j=1}^N \liminf_{\varepsilon \rightarrow 0} \text{I}_j^\varepsilon \geq \sum_{j=1}^N \int_{\Omega \cap \partial^* \{u^j=1\}} f_{\text{hom}}^j(v) d\mathcal{H}^{d-1}. \tag{27}$$

It remains to estimate  $\text{III}^\varepsilon$ . To that end, we introduce the set of indices

$$\Lambda_M^\varepsilon = \{z \in J_M^\varepsilon : u^\varepsilon \text{ constant on every connected component of } A_j \cap (Q_{3M} + zM), \\ j = 1, \dots, N\}.$$

Note that

$$\#(J_M^\varepsilon \setminus \Lambda_M^\varepsilon) \leq \frac{cM}{\varepsilon^{d-1}}. \tag{28}$$

We then write

$$\begin{aligned}
\text{III}^\varepsilon &= \sum_{z \in \Lambda_M^\varepsilon} \varepsilon^d \left( \sum_{(k,k') \in \mathcal{N}_0^\varepsilon(Q_M + zM)} a_{kk'} (v_k - v_{k'})^2 + \sum_{k \in Z(Q_M + zM)} g(k, v_k) \right) \\
&\quad + \sum_{z \in J_M^\varepsilon \setminus \Lambda_M^\varepsilon} \varepsilon^d \left( \sum_{(k,k') \in \mathcal{N}_0^\varepsilon(Q_M + zM)} a_{kk'} (v_k - v_{k'})^2 + \sum_{k \in Z(Q_M + zM)} g(k, v_k) \right) \\
&\geq \sum_{z \in \Lambda_M^\varepsilon} \varepsilon^d M^d \varphi_M(u_1^\varepsilon, \dots, u_N^\varepsilon) - c\varepsilon^d M^d \max(|g| + |a_{kk'}|) \#(J_M^\varepsilon \setminus \Lambda_M^\varepsilon),
\end{aligned}$$

where  $u_j^\varepsilon$  is the constant value taken by  $u^\varepsilon$  on  $(Q_M + zM) \cap C_j$ . Here we suppose  $M$  is large enough so that the connected component of  $C_j$  containing  $(Q_M + zM) \cap C_j$

is connected in  $Q_{3M} + zM$ . We set

$$U^\varepsilon = \sum_{z \in \Lambda_M^\varepsilon} (u_1^\varepsilon, \dots, u_N^\varepsilon) \chi_{Q_M + zM}$$

and  $\varphi_M(0, \dots, 0) = 0$ . Note that  $U^\varepsilon \rightarrow U := (u^1, \dots, u^N)$  in  $L^1(\Omega)^N$  so that

$$\liminf_{\varepsilon \rightarrow 0} \text{III}^\varepsilon \geq \liminf_{\varepsilon \rightarrow 0} \left( \int_{\Omega} \varphi_M(U^\varepsilon) dx - \varepsilon \max |g| c_M M^d \right) = \int_{\Omega} \varphi_M(U) dx \quad (29)$$

by the Lebesgue dominated convergence theorem and the estimate (28).

Summing up the inequalities (26), (27), and (29), we get

$$\liminf_{\varepsilon \rightarrow 0} F_\varepsilon(u^\varepsilon) \geq \sum_{j=1}^N \int_{\Omega \cap \partial^* \{u^j=1\}} f_{\text{hom}}^j(v) d\mathcal{H}^{d-1} + \int_{\Omega} \varphi_M(U) dx. \quad (30)$$

The lower-bound inequality then follows by taking the limit as  $M \rightarrow +\infty$ , using Proposition 4.1 and the Lebesgue dominated convergence theorem.  $\square$

*Proof of the upper bound.* We fix  $U = (u^1, \dots, u^N) \in \text{BV}(\Omega; \{1, -1\})^N$ . For every  $j = 1, \dots, N$ , we choose  $u^{j,\varepsilon} \rightarrow u^j$  a recovery sequence for  $\mathcal{F}_{\text{hom}}^j(u^j)$ . We tacitly extend all functions defined on  $Z^\varepsilon(\Omega)$  to all of  $\mathbb{Z}^d$  with the value  $+1$  outside  $Z^\varepsilon(\Omega)$ . This does not affect the value of the energies but allows us to rigorously define some sets of indices  $z$  in the sequel.

We fix  $M \in \mathbb{N}$  large enough. As in Section 4, we introduce the sets of indices

$$\tilde{J}_M^\varepsilon = \left\{ z \in \mathbb{Z}^d : (Q_M + zM) \cap \frac{1}{\varepsilon} \Omega \neq \emptyset \right\},$$

$$\Lambda_M^{j,\varepsilon} = \{ z \in J_M^\varepsilon : u^\varepsilon \text{ constant on every connected component of } A_j \cap (Q_{3M} + zM) \}$$

and give the estimate

$$\sum_{j=1}^N \#(\tilde{J}_M^\varepsilon \setminus \Lambda_M^{j,\varepsilon}) \leq \frac{cM}{\varepsilon^{d-1}}. \quad (31)$$

Note that, if  $z \in \bigcap_{j=1}^N \Lambda_M^{j,\varepsilon}$ , then  $u^{j,\varepsilon} =: u^{j,\varepsilon,z}$  is constant on  $C_j \cap (Q_M + zM)$  for  $j = 1, \dots, N$ . Let  $v^{\varepsilon,z}$  be a minimizer for  $\tilde{\varphi}_M(u^{1,\varepsilon,z}, \dots, u^{N,\varepsilon,z})$ .

We define

$$u_k^\varepsilon = \begin{cases} u_k^{j,\varepsilon} & \text{if } k \in C_j, j = 1, \dots, N, \\ v^{\varepsilon,z}(k - zM) & \text{if } k \in Q_M + zM \text{ and } z \in \bigcap_{j=1}^N \Lambda_M^{j,\varepsilon}, \\ 1 & \text{otherwise.} \end{cases}$$



We first estimate the energy on the strong connections. By the definition of  $u^{j,\varepsilon}$ , we have for all  $j = 1, \dots, N$

$$\lim_{\varepsilon \rightarrow 0} \sum_{(k,k') \in \mathcal{N}_j^\varepsilon(\Omega) \cap (C_j \times C_j)} \varepsilon^{d-1} a_{kk'} (u_k^\varepsilon - u_{k'}^\varepsilon)^2 = \mathcal{F}_{\text{hom}}^j(u^j) \quad (32)$$

since  $u^\varepsilon = u^{j,\varepsilon}$  on  $C_j$ . On the strong connections between points not in the infinite connected components  $C_j$ ,

$$\sum_{(k,k') \in \mathcal{N}_j^\varepsilon(\Omega) \setminus (C_j \times C_j)} \varepsilon^{d-1} a_{kk'} (u_k^\varepsilon - u_{k'}^\varepsilon)^2 = 0 \quad (33)$$

since  $u^\varepsilon$  is constant on every connected component of  $A_j \setminus C_j$ . Note that here we have used the condition that  $v^{\varepsilon,z} = 1$  on  $D_M$  in the definition of  $\tilde{\varphi}_M$ .

We then examine the contribution due to the interaction between weak connections and the term  $g$ . We first look at the contributions on the cubes in the sets  $\Lambda_M^{j,\varepsilon}$ , where we can use the definition of  $\tilde{\varphi}_M$ : for every  $z \in \bigcap_{j=1}^N \Lambda_M^{j,\varepsilon}$ ,

$$\sum_{(k,k') \in \mathcal{N}_0^\varepsilon(Q_M + zM)} a_{kk'} (u_k^\varepsilon - u_{k'}^\varepsilon)^2 + \sum_{k \in Z(Q_M + zM)} g(k, u_k^\varepsilon) = \tilde{\varphi}_M(u^{1,\varepsilon,z}, \dots, u^{N,\varepsilon,z}).$$

The contributions interior to all other cubes in  $\tilde{J}_M^\varepsilon$  sum up to

$$\begin{aligned} \sum_{z \notin \bigcap_{j=1}^N \Lambda_M^{j,\varepsilon}} \varepsilon^d \left( \sum_{(k,k') \in \mathcal{N}_0^\varepsilon(Q_M + zM)} a_{kk'} (u_k^\varepsilon - u_{k'}^\varepsilon)^2 + \sum_{k \in Z(Q_M + zM)} g(k, u_k^\varepsilon) \right) \\ \leq \varepsilon^d M^d (\#P_0 \max a_{il} + \max |g|) \sum_{j=1}^N \#(J_M^\varepsilon \setminus \Lambda_M^{j,\varepsilon}) \\ \leq \varepsilon M^d c'_M + o(1) \end{aligned} \quad (34)$$

by (31) and the fact that the boundary of  $\Omega$  has zero measure. Finally, the contribution due to the weak connection across the boundary of neighboring cubes is given by

$$\begin{aligned} \sum_{z \neq z' \in \bigcap_{j=1}^N \Lambda_M^{j,\varepsilon}} \varepsilon^d \sum_{(k,k') \in \mathcal{N}_0^\varepsilon(\Omega), k \in Q_M + zM, k' \in Q_M + z'M} a_{kk'} (u_k^\varepsilon - u_{k'}^\varepsilon)^2 \\ \leq \varepsilon^d M^{d-1} \#J_M^\varepsilon \#P_0 \max a_{il} \leq \#P_0 \max a_{il} \frac{|\Omega|}{M}. \end{aligned}$$

From the inequalities above, we obtain

$$\limsup_{\varepsilon \rightarrow 0} F_\varepsilon(u^\varepsilon) \leq \sum_{j=1}^N \mathcal{F}_{\text{hom}}^j(u^j) + \int_{\Omega} \tilde{\varphi}_M(u^1, \dots, u^N) dx + \#P_0 \max a_{il} \frac{|\Omega|}{M}.$$

Letting  $M \rightarrow +\infty$  and using Propositions 4.2 and 4.1 then gives the result.  $\square$

### 6. Examples

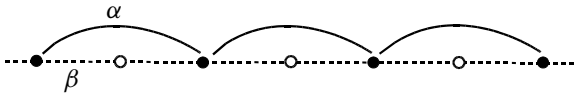
In the pictures in the following examples, weak connections are denoted by a dashed line and strong connections by a continuous line.

**6.1. One-dimensional examples.** In this section, we consider easy one-dimensional examples, highlighting the possibility of double-porosity behavior if long-range interactions are allowed, contrary to the continuum case. We use a slightly different notation than above, with the sums depending only on one index. The factor  $\frac{1}{4}$  is just a normalization since  $(u_i - u_j)^2$  is always a multiple of 4.

**Example 6.1** (weak inclusions on alternating lattice). Consider a system of weak nearest-neighbor interactions and strong next-to-nearest-neighbor interactions on the even-odd lattice (see figure below); namely,

$$F_\varepsilon(u) = \frac{\beta}{4} \sum_{i=1}^{N_\varepsilon} \varepsilon(u_i - u_{i-1})^2 + \frac{\alpha}{4} \sum_{j=1}^{N_\varepsilon/2-1} (u_{2j+1} - u_{2j-1})^2 + \sum_{i=1}^{N_\varepsilon} \varepsilon g(u_i),$$

where we assume that  $\Omega = [0, 1]$  and  $N_\varepsilon = 1/\varepsilon \in 2\mathbb{N}$ . In this case  $N = 1$ ,  $A_1 = C_1 = 1 + 2\mathbb{N}$ , and  $A_0 = 2\mathbb{N}$ .

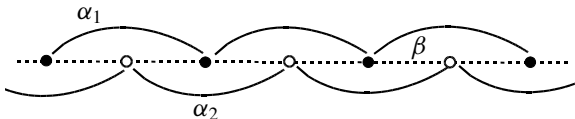


The  $\Gamma$ -limit is

$$\begin{aligned} F_{\text{hom}}(u) &= \alpha \#S(u) + \frac{1}{2} \int_0^1 g(u) \, dx + \frac{1}{2} \int_0^1 \min\{g(u), g(-u) + 2\beta\} \, dx \\ &= \alpha \#S(u) + \int_0^1 g(u) \, dx - \frac{1}{2} \int_0^1 \max\{0, g(u) - g(-u) - 2\beta\} \, dx. \end{aligned}$$

The last term favors states with the same value on  $A_0$  and  $A_1$  if the integrand is 0 and of opposite sign if the integrand is positive. Note that this is always the case if we have a strong-enough “antiferromagnetic” nearest-neighbor interaction, i.e.,  $\beta$  is negative and  $2|\beta| > |g(1) - g(-1)|$ .

**Example 6.2** (interacting sublattices). Consider a system of weak nearest-neighbor interactions and strong next-to-nearest-neighbor interactions:



Here

$$F_\varepsilon(u) = \frac{\beta}{4} \sum_{i=1}^{N_\varepsilon} \varepsilon(u_i - u_{i-1})^2 + \frac{\alpha_1}{4} \sum_{j=1}^{N_\varepsilon/2-1} (u_{2j+1} - u_{2j-1})^2 \\ + \frac{\alpha_2}{4} \sum_{j=0}^{N_\varepsilon/2-1} (u_{2j+2} - u_{2j})^2 + \sum_{i=1}^{N_\varepsilon} \varepsilon g(u_i),$$

where we assume that  $N_\varepsilon = 1/\varepsilon \in 2\mathbb{N}$ . In this case,  $N = 2$ ,  $A_1 = C_1 \stackrel{i=1}{=} 1 + 2\mathbb{N}$ ,  $A_2 = C_2 = 2\mathbb{N}$ , and  $A_0 = \emptyset$ .

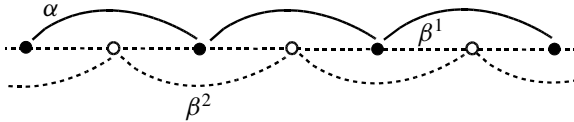
The  $\Gamma$ -limit is

$$F_{\text{hom}}(u^1, u^2) = \alpha_1 \#S(u^1) + \alpha_2 \#S(u^2) \\ + \frac{1}{2} \int_0^1 g(u^1) dx + \frac{1}{2} \int_0^1 g(u^2) dx + \frac{\beta}{4} \int_0^1 (u^2 - u^1)^2.$$

Note that, since  $A_0 = \emptyset$ , we have no optimization in the interacting term, which then is just the pointwise limit of the nearest-neighbor interactions. Note moreover that in the case  $\beta = 0$  the interactions are completely decoupled.

**Example 6.3** (interacting weak and strong sublattices). We consider the same pattern of interactions as in the previous example but with only strong connections on the odd lattice as in [Example 6.1](#) (see figure below), i.e., with

$$F_\varepsilon(u) = \frac{\beta^1}{4} \sum_{i=1}^{N_\varepsilon} \varepsilon(u_i - u_{i-1})^2 + \frac{\beta^2}{4} \sum_{j=0}^{N_\varepsilon/2-1} \varepsilon(u_{2j+2} - u_{2j})^2 \\ + \frac{\alpha}{4} \sum_{j=1}^{N_\varepsilon/2-1} (u_{2j+1} - u_{2j-1})^2 + \sum_{i=1}^{N_\varepsilon} \varepsilon g(u_i).$$



In this case, we have three possibilities:

- the minimizing values on the even lattice agree with those on the odd lattice (ferromagnetic overall behavior),
- the minimizing values on the even lattice disagree with those on the odd lattice (antiferromagnetic overall behavior), or
- the values on the even lattice alternate (antiferromagnetic behavior on the weak lattice).

The value of  $\varphi$  is obtained by optimizing over these three possibilities; i.e.,

$$\varphi(u) = \min \left\{ g(u), \frac{g(u) + g(-u)}{2} + \beta^1, \frac{3g(u) + g(-u)}{4} + \frac{\beta^1 + \beta^2}{2} \right\},$$

and we have

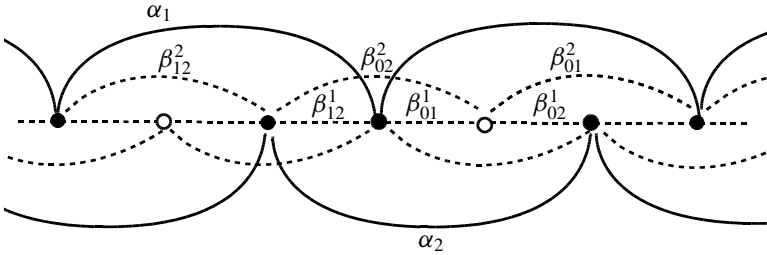
$$F_{\text{hom}}(u) = \alpha \#S(u) + \int_0^1 \varphi(u) dx.$$

**Example 6.4** (third-neighbor hard phases). In the system described in the figure below, involving strong third-neighbor interactions, we have two strong components and a  $\Gamma$ -limit obtained by minimization of the nearest and next-to-nearest neighbors. Using the same notation of the previous examples for the coefficients, we can write the limit as

$$F_{\text{hom}}(u^1, u^2) = \alpha_1 \#S(u^1) + \alpha_2 \#S(u^2) + \int_0^1 \varphi(u^1, u^2) dx,$$

and

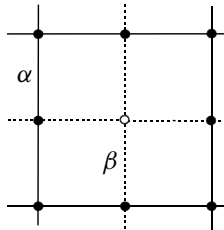
$$\begin{aligned} \varphi(u^1, u^2) = & \frac{1}{3}(g(u^1) + g(u^2)) + \frac{1}{4}\beta_{12}^2(u^2 - u^1)^2 \\ & + \frac{1}{3} \min\left\{\frac{1}{4}((\beta_{01}^1 + \beta_{01}^2)(v - u^1)^2 + (\beta_{02}^1 + \beta_{02}^2)(v - u^2)^2) + g(v) : v \in \{-1, 1\}\right\}. \end{aligned}$$



**6.2. Higher-dimensional examples.** In the following examples, we go back to the notation used in the statement of the main result. The normalization factor  $\frac{1}{8}$  takes into account that each pair of nearest neighbors is accounted for twice.

**Example 6.5** (a nearest-neighbor system with soft inclusions). Consider a nearest-neighbor system in two dimensions in which  $A_0 = 2\mathbb{Z}^2$  and strong and weak interactions are given respectively by

$$\frac{1}{8}\alpha(u_k - u_{k'})^2, \quad \frac{1}{8}\varepsilon\beta(u_k - u_{k'})^2.$$



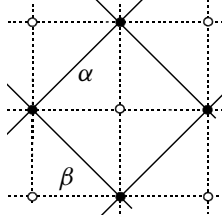
In this case,

$$F_{\text{hom}}(u) = \frac{1}{2}\alpha \int_{S(u) \cap \Omega} \|v_u\|_1 d\mathcal{H}^1 + \int_{\Omega} \varphi(u) dx,$$

where

$$\varphi(u) = \min \left\{ g(u), \frac{3g(u) + g(-u)}{4} + \beta \right\}.$$

**Example 6.6** (a lattice with weak nearest-neighbor interactions). Consider strong interactions on a lattice of next-to-nearest neighbors as in the figure:



with weak nearest-neighbor interactions on the square lattice given respectively by

$$\frac{1}{8}\alpha(u_k - u_{k'})^2, \quad \frac{1}{8}\varepsilon\beta(u_k - u_{k'})^2,$$

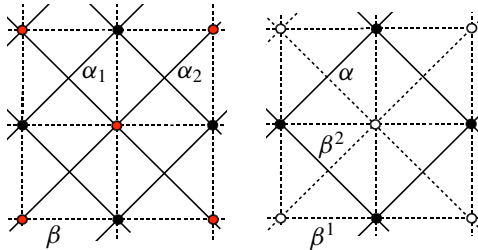
(the factor  $\frac{1}{8}$  takes into account that each pair is accounted for twice). We only have one strong component, and with this choice of coefficients,

$$F_{\text{hom}}(u) = \alpha \int_{\Omega \cap \partial\{u=1\}} \|v\|_{\infty} d\mathcal{H}^1 + \int_{\Omega} \varphi(u) dx,$$

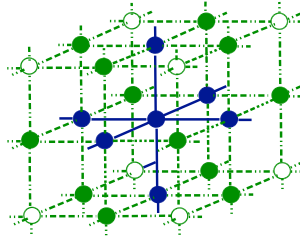
where

$$\varphi(u) = \min \{ g(u), \frac{1}{2}(g(u) + g(-u)) + \beta \}.$$

**Example 6.7.** We include just the pictorial description of two more two-dimensional systems with a limit with two parameters (below, left) and with one limit parameter but with the possibility of an oscillating behavior on the weak lattice (below, right), analogous to the one-dimensional Examples 6.2 and 6.3, respectively.



**Example 6.8.** We finally consider a three-dimensional two-periodic geometry, with one strong connected component pictured in Figure 3. Even in the absence of the



**Figure 3.** Oscillations in the infinite weak component.

forcing term  $g$ , we may have several competing microstructures in the determination of  $\varphi$ . In Figure 3, we have represented the uniform data  $u = +1$  on the strong component with solid circles and a system of ferromagnetic connections between strong and weak sites (positive coefficients) and of antiferromagnetic connections between weak sites (a negative coefficient  $\alpha$ ). Correspondingly, the minimal states have the value  $+1$  on weak sites connected with the strong component (represented by solid circles) and the value  $-1$  on the other sites (represented by white circles). Note that in this case the contribution of the weak phase is a constant.

## References

- [Alicandro and Gelli 2016] R. Alicandro and M. S. Gelli, “Local and nonlocal continuum limits of Ising-type energies for spin systems”, *SIAM J. Math. Anal.* **48**:2 (2016), 895–931.
- [Amaziane and Pankratov 2016] B. Amaziane and L. Pankratov, “Homogenization of a model for water-gas flow through double-porosity media”, *Math. Methods Appl. Sci.* **39**:3 (2016), 425–451.
- [Amaziane et al. 2006] B. Amaziane, L. Pankratov, and A. Piatnitski, “Homogenization of a class of quasilinear elliptic equations in high-contrast fissured media”, *Proc. Roy. Soc. Edinburgh (A)* **136**:6 (2006), 1131–1155.
- [Amaziane et al. 2009a] B. Amaziane, L. Pankratov, and A. Piatnitski, “Nonlinear flow through double porosity media in variable exponent Sobolev spaces”, *Nonlinear Anal. Real World Appl.* **10**:4 (2009), 2521–2530.
- [Amaziane et al. 2009b] B. Amaziane, L. Pankratov, and V. Rybalko, “On the homogenization of some double-porosity models with periodic thin structures”, *Appl. Anal.* **88**:10–11 (2009), 1469–1492.
- [Ambrosio et al. 2008] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, 2nd ed., Birkhäuser, Basel, 2008.
- [Arbogast et al. 1990] T. Arbogast, J. Douglas, Jr., and U. Hornung, “Derivation of the double porosity model of single phase flow via homogenization theory”, *SIAM J. Math. Anal.* **21**:4 (1990), 823–836.
- [Barenblatt et al. 1960] G. I. Barenblatt, Iu. P. Zheltov, and I. N. Kochina, “Basic concepts in the theory of seepage of homogeneous liquids in fissured rocks [strata]”, *J. Appl. Math. Mech.* **24**:5 (1960), 1286–1303.

- [Bourgeat et al. 1996] A. Bourgeat, S. Luckhaus, and A. Mikelić, “Convergence of the homogenization process for a double-porosity model of immiscible two-phase flow”, *SIAM J. Math. Anal.* **27**:6 (1996), 1520–1543.
- [Bourgeat et al. 1998] A. Bourgeat, A. Mikelić, and A. Piatnitski, “Modèle de double porosité aléatoire”, *C. R. Acad. (I)* **327**:1 (1998), 99–104.
- [Bourgeat et al. 1999] A. Bourgeat, M. Goncharenko, M. Panfilov, and L. Pankratov, “A general double porosity model”, *C. R. Acad. (IIB)* **327**:12 (1999), 1245–1250.
- [Bourgeat et al. 2003] A. Bourgeat, G. A. Chechkin, and A. L. Piatnitski, “Singular double porosity model”, *Appl. Anal.* **82**:2 (2003), 103–116.
- [Braides 1998] A. Braides, *Approximation of free-discontinuity problems*, Lecture Notes in Mathematics **1694**, Springer, Berlin, 1998.
- [Braides 2002] A. Braides,  *$\Gamma$ -convergence for beginners*, Oxford Lecture Series in Mathematics and its Applications **22**, Oxford University Press, Oxford, 2002.
- [Braides 2006] A. Braides, “A handbook of  $\Gamma$ -convergence”, pp. 101–213 in *Handbook of differential equations: stationary partial differential equations*, vol. 3, ch. 2, edited by M. Chipot and P. Quittner, Elsevier, Amsterdam, 2006.
- [Braides 2014a] A. Braides, “Discrete-to-continuum variational methods for Lattice systems”, pp. 997–1015 in *Proceedings of the International Congress of Mathematicians* (Seoul, 2014), vol. IV, edited by S. Y. Jang et al., Kyung Moon Sa, Seoul, 2014.
- [Braides 2014b] A. Braides, *Local minimization, variational evolution and  $\Gamma$ -convergence*, Lecture Notes in Mathematics **2094**, Springer, Cham, 2014.
- [Braides and Piatnitski 2013] A. Braides and A. Piatnitski, “Homogenization of surface and length energies for spin systems”, *J. Funct. Anal.* **264**:6 (2013), 1296–1328.
- [Braides and Solci 2013] A. Braides and M. Solci, “Multi-scale free-discontinuity problems with soft inclusions”, *Boll. Unione Mat. Ital.* (9) **6**:1 (2013), 29–51.
- [Braides and Solci 2015] A. Braides and M. Solci, “Motion of discrete interfaces through mushy layers”, preprint, 2015. To appear in *J. Nonlinear Sci.* [arXiv 1506.04062v1](https://arxiv.org/abs/1506.04062v1)
- [Braides et al. 2004] A. Braides, V. Chiadò Piat, and A. Piatnitski, “A variational approach to double-porosity problems”, *Asymptot. Anal.* **39**:3–4 (2004), 281–308.
- [Braides et al. 2015] A. Braides, V. Chiadò Piat, and A. Piatnitski, “Homogenization of discrete high-contrast energies”, *SIAM J. Math. Anal.* **47**:4 (2015), 3064–3091.
- [Caffarelli and de la Llave 2005] L. A. Caffarelli and R. de la Llave, “Interfaces of ground states in Ising models with periodic coefficients”, *J. Stat. Phys.* **118**:3–4 (2005), 687–719.
- [Cherdantsev and Cherednichenko 2012] M. Cherdantsev and K. D. Cherednichenko, “Two-scale  $\Gamma$ -convergence of integral functionals and its application to homogenisation of nonlinear high-contrast periodic composites”, *Arch. Ration. Mech. Anal.* **204**:2 (2012), 445–478.
- [Choquet 2004] C. Choquet, “Derivation of the double porosity model of a compressible miscible displacement in naturally fractured reservoirs”, *Appl. Anal.* **83**:5 (2004), 477–499.
- [Choquet and Pankratov 2010] C. Choquet and L. Pankratov, “Homogenization of a class of quasilinear elliptic equations with non-standard growth in high-contrast media”, *Proc. Roy. Soc. Edinburgh (A)* **140**:3 (2010), 495–539.
- [Hornung 1997] U. Hornung (editor), *Homogenization and porous media*, Interdisciplinary Applied Mathematics **6**, Springer, New York, 1997.
- [Konyukhov and Pankratov 2015] A. Konyukhov and L. Pankratov, “Upscaling of an immiscible non-equilibrium two-phase flow in double porosity media”, *Appl. Anal.* (online publication July 2015).

- [Marchenko and Khruslov 2006] V. A. Marchenko and E. Ya. Khruslov, *Homogenization of partial differential equations*, Progress in Mathematical Physics **46**, Birkhäuser, Boston, 2006.
- [Panassenko 1991] G. P. Panassenko, “Multicomponent homogenization of processes in strongly non-homogeneous structures”, *Sb. Math.* **69**:1 (1991), 143–153.
- [Panfilov 2000] M. Panfilov, *Macroscale models of flow through highly heterogeneous porous media*, Theory and Applications of Transport in Porous Media **16**, Springer, Dordrecht, 2000.
- [Pankratov and Piatnitski 2002] L. Pankratov and A. Piatnitski, “Nonlinear ‘double porosity’ type model”, *C. R. Math.* **334**:5 (2002), 435–440.
- [Pankratov and Rybalko 2003] L. S. Pankratov and V. A. Rybalko, “Asymptotic analysis of a double porosity model with thin fissures”, *Sb. Math.* **194**:1 (2003), 123–150.
- [Pankratov et al. 2003] L. Pankratov, A. Piatnitskii, and V. Rybalko, “Homogenized model of reaction–diffusion in a porous medium”, *C. R. Mécanique* **331**:4 (2003), 253–258.
- [Sandrakov 1999a] G. V. Sandrakov, “Homogenization of elasticity equations with contrasting coefficients”, *Sb. Math.* **190**:12 (1999), 1749–1806.
- [Sandrakov 1999b] G. V. Sandrakov, “Homogenization of parabolic equations with contrasting coefficients”, *Izv. Math.* **63**:5 (1999), 1015–1061.
- [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** (2011), 012018.
- [Solci 2009] M. Solci, “Double-porosity homogenization for perimeter functionals”, *Math. Methods Appl. Sci.* **32**:15 (2009), 1971–2002.
- [Solci 2012] M. Solci, “Multiphase double-porosity homogenization for perimeter functionals”, *Math. Methods Appl. Sci.* **35**:5 (2012), 598–620.
- [Yeh 2006] L.-M. Yeh, “Homogenization of two-phase flow in fractured media”, *Math. Models Methods Appl. Sci.* **16**:10 (2006), 1627–1651.

Received 1 Dec 2015. Revised 21 Jan 2016. Accepted 28 Feb 2016.

ANDREA BRAIDES: [braides@mat.uniroma2.it](mailto:braides@mat.uniroma2.it)

Dipartimento di Matematica, Università di Roma Tor Vergata, via della ricerca scientifica 1, I-00133 Roma, Italy

VALERIA CHIADÒ PIAT: [d001854@polito.it](mailto:d001854@polito.it)

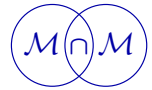
Dipartimento di Matematica, Politecnico di Torino, corso Duca degli Abruzzi 24, I-10129 Torino, Italy

MARGHERITA SOLCI: [margherita@uniss.it](mailto:margherita@uniss.it)

Dipartimento di Architettura, Design, Urbanistica, Università di Sassari, piazza Duomo 6, I-07041 Alghero, Italy







# CORRECTION TO THE ARTICLE ON THE THEORY OF DIFFUSION AND SWELLING IN FINITELY DEFORMING ELASTOMERS

GARY J. TEMPLET AND DAVID J. STEIGMANN

Volume 1:1 (2013), 105–128

We are grateful to Professor Patrizio Neff for drawing our attention to an incorrect statement in the paper, to the effect that the conditions listed in (43) are necessary and sufficient for the polyconvexity of the strain-energy function of an isotropic material. In fact, these inequalities are shown in the paper to be necessary and sufficient for the polyconvexity of the function defined by (38). However, not every polyconvex, isotropic function is expressible in the form (38), and so the conditions (43), while sufficient for polyconvexity of an isotropic strain-energy function, are not necessary.

Received 1 Mar 2016. Accepted 14 Mar 2016.

GARY J. TEMPLET: [gjtempl@berkeley.edu](mailto:gjtempl@berkeley.edu)  
Berkeley, CA, United States

DAVID J. STEIGMANN: [dsteigmann@berkeley.edu](mailto:dsteigmann@berkeley.edu)  
Department of Mechanical Engineering, University of California, Berkeley, Berkeley, CA 94720,  
United States



Communicated by Francesco dell’Isola.

MSC2010: 74FXX.

Keywords: errata.





## 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 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  $\text{BIB}\TeX$  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](mailto: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.

Gradient materials with internal constraints Albrecht Bertram and Rainer Glüge	1
Unified geometric formulation of material uniformity and evolution Marcelo Epstein and Manuel de León	17
Electromechanics of polarized lipid bilayers David J. Steigmann and Ashutosh Agrawal	31
Orthogonal polynomials and Riesz bases applied to the solution of Love's equation Pierluigi Vellucci and Alberto Maria Bersani	55
Modeling capillary hysteresis in unsaturated porous media G�rard Gagneux and Olivier Millet	67
Discrete double-porosity models for spin systems Andrea Braides, Valeria Chiad� Piat and Margherita Solci	79
Correction to "On the theory of diffusion and swelling in finitely deforming elastomers" Gary J. Templet and David J. Steigmann	103

*MEMOCS* is a journal of the International Research Center for the Mathematics and Mechanics of Complex Systems at the Universit  dell'Aquila, Italy.

