J. Mech. Phys. Solids 215 (2026) 106741

Contents lists available at ScienceDirect

Journal of the Mechanics and Physics of Solids

journal homepage: www.elsevier.com/locate/jmps

Nonlinear electroelasticity via inverse motion: A gauge-invariant formulation with explicit variational structure

Paolo Maria Mariano

DICEA, Università di Firenze, via Santa Marta 3, Firenze, I-50139, Italy

ARTICLE INFO

Keywords:
Nonlinear electromagnetism
Lagrangians
Electro-elastic interactions
Polarization
Configurational actions
Inverse motion

ABSTRACT

We develop a variational formulation of nonlinear electroelasticity based on inverse motion and gauge invariance. Motivated by Ericksen's reconsideration of elastic dielectrics from the viewpoint of energy minimization, the proposed framework describes both mechanical and electromagnetic fields directly in the current configuration, providing a unified setting for electroelastic interactions. We discuss the relation between the inverse-motion formulation and the classical theories of Toupin and Ericksen, showing that standard electroelastic models are recovered in the small-strain regime. Particular attention is devoted to flexoelectric effects, interpreted as manifestations of latent microstructure and incorporated through higher-gradient contributions. The variational implications of the theory are illustrated through explicit coercivity conditions ensuring stability in a representative quadratic model and through an existence theorem for minimizers of a broad class of nonlinear electroelastic functionals. The resulting framework provides a unified nonlinear setting for electroelasticity and flexoelectricity while making explicit the connection between electroelastic couplings, stability, and energy minimization.

1. Introduction

The description of electroelastic coupling at finite strains, grounded in the pioneering work of Toupin (1956, 1960), involves a fundamental structural issue: electromagnetic fields are naturally formulated in the current (Eulerian) representation, whereas nonlinear elasticity is typically described with respect to a reference configuration (Lagrangian setting). The topic has been largely discussed in treatises (e.g., Brown 1966, Truesdell and Toupin 1960, Dorfmann and Ogden 2014, Maugin 1988, Wang 1979) and further foundational analyses (Ericksen, 2002; Fosdick and Tang, 2007; Liu, 2014; Man, 2025; Steigmann, 2007; Dorfmann and Ogden, 2006). Existing approaches reconcile this mismatch by transforming one set of fields into the representation of the other (Bustamante et al., 2009; Davì and Mariano, 2001; Dorfmann and Ogden, 2014) or by adopting mixed formulations (Xiao and Bhattacharya, 2008).

Beyond this representational issue, however, there is a deeper question concerning the variational structure of electroelasticity. Toupin's formulation provides a general variational framework leading to the governing equations of electroelastic media. Ericksen (2007b,a) later revisited the theory from a different perspective, emphasizing that a formulation naturally adapted to energy minimization may be particularly advantageous when questions of stability arise. In this sense, the issue is not merely the derivation of equilibrium equations, but the possibility of interpreting electroelastic interactions within a framework where minimizers, stability, and variational structure directly appear.

Motivated by these considerations, we adopt here a viewpoint based on inverse motion. Instead of describing the body through the direct motion from a reference configuration to the current one, we consider the inverse map and formulate both mechanical and

E-mail address: paolomaria.mariano@unifi.it

https://doi.org/10.1016/j.jmps.2026.106741
Received 4 May 2026; Received in revised form 19 June 2026; Accepted 22 June 2026
Available online 24 June 2026

electromagnetic fields directly in the current configuration. This choice allows one to avoid pull-back operations, places all fields in a common geometric representation, and leads naturally to a gauge-invariant variational formulation.

The approach proposed here is not intended as an alternative to existing electroelastic theories. Rather, it is a reformulation aimed at clarifying their geometric structure and variational content. Its main advantage lies in the possibility of treating mechanical and electromagnetic fields within a unified setting in the current configuration, while retaining direct connections with classical electroelastic models.

To make the framework concrete, we explicitly establish its relation with classical electroelasticity and discuss its small-strain limit, where the distinction between direct and inverse descriptions becomes essentially negligible. We then consider a representative electroelastic energy inspired by recent theories of flexoelectric materials and rewrite it in terms of inverse motion. The resulting functional admits minimizers under large strains. In this way, the variational implications of the general framework can be assessed on a specific class of models of current interest.

A further point of interest concerns the influence of electroelastic couplings on material stability. In particular, flexoelectric effects, interpreted here as manifestations of latent microstructure, introduce higher-gradient contributions that may either stabilize or destabilize the material response. This aspect is analyzed within the present variational setting through explicit coercivity and minimization results.

The formulation developed here is intrinsically compatible with gauge invariance and admits a natural extension to relativistic settings. Nevertheless, the primary focus of the paper remains nonlinear electroelasticity and its variational structure. The relativistic extension is presented only as a natural continuation of the inverse-motion framework.

The paper is organized as follows. Section 2 recalls the basic elements of gauge invariance. Section 3 introduces the inverse-motion framework. Section 4 presents the Lagrangian formulation and the associated d'Alembert-Lagrange principle. Section 5 derives the governing equations. Section 6 establishes the connection with classical electroelasticity, discusses the relation with the formulations of Toupin and Ericksen, and analyzes the small-strain regime. Section 7 incorporates flexoelectric effects into the general framework. Section 8 discusses stability in the small-strain regime and proves existence of minimizers for a representative flexoelectric energy under large strains. Finally, Section 9 outlines a possible relativistic extension.

Regarding notations, for $\mathfrak{A}$ and $\mathfrak{B}$ third-rank tensors, the product $\mathfrak{A} : \mathfrak{B}$ gives a second-rank tensor with components (say) $\boldsymbol{N}_i^l = \mathfrak{A}_{ijk}\mathfrak{B}^{jkl}$, while for $\boldsymbol{A}$ and $\boldsymbol{B}$ two second-rank tensors, the product $\boldsymbol{A}\boldsymbol{B}$ gives a second-rank tensor with components (say) $\boldsymbol{C}_i^j = \boldsymbol{A}_{ik}\boldsymbol{B}^{kj}$, where summation over repeated indices is accepted here and below. An interposed dot between elements of a generic linear space and those of its dual will indicate in the sequel duality pairing (that is, saturation of all components), independently of the linear space considered; the result of this product is a scalar. A superposed dot indicates the time derivative, while D means spatial derivative.

2. Gauge invariance: Essential elements

Field theories may include in their structure some indeterminacies. For example, adding an arbitrary constant to the gravitational potential, we get the same force; two wave functions of (say) electrons that differ by the phase cannot be distinguished by measurements. These and other cases are said as endowed with a gauge freedom and can be exploited in terms of transformation, that is, a change of fields, to simplify the problems tackled. In fact, the key aspect is that as long as a gauge transformation is carried out throughout a problem in hand, the underlying physics remains unchanged.

In electromagnetism, gauge transformations involve the differentiable *vector potential* A and *scalar potential* Φ. Indeed, the *magnetic flux density* B, a divergence-free field in the free space, is given by the curl of A, that is,

$$B = \mathrm{curl} A\,. \tag{1}$$

Evidently, the transformation

$$A \longmapsto A + D\chi, \tag{2}$$

with χ any differentiable function of space and time, leaves invariant B.

The scalar potential Φ enters the definition of electric field E. Indeed, we have

$$E = -D\Phi - \dot{A}\,. \tag{3}$$

By using the transformation (2), we obtain $E = -D\Phi - \dot{A} - D\dot{\chi}$. Thus, in order that the electric field E remains invariant, we need to consider in addition the transformation

$$\Phi \longmapsto \Phi - \dot{\chi}\,. \tag{4}$$

The pair of transformations (2) and (4) is what we impose on the electromagnetic field when we require gauge invariance.

In the present framework, gauge invariance plays a structural role. It allows the electromagnetic contribution to be formulated directly in the current configuration and leads naturally to the variational construction discussed in the following sections.

3. Classical space-time: 4-vectors and 4-tensors

The inverse-motion formulation proposed here is naturally expressed in terms of space-time variables. For this reason, before introducing the variational structure of the theory, it is convenient to collect some elementary notions concerning a four-dimensional representation of space and time. At this stage, the setting remains entirely classical and no relativistic assumptions are involved.

The ambient where we place first the theory is the classical space-time $\mathbb{R}^3 \times \mathbb{R}$. Its elements are called here *4-vectors*. Any $\boldsymbol{a} \in \mathbb{R}^3 \times \mathbb{R}$ has a component of space type, here indicated by $\mathsf{a} \in \mathbb{R}^3$ and one of time type in $\mathbb{R}$, say, a, that is,

$$\boldsymbol{a} = (\mathsf{a}, a)\,.$$

A linear map from $\mathbb{R}^3 \times \mathbb{R}$ onto itself, that is, an element $\boldsymbol{N}$ of the linear space $\mathrm{Hom}(\mathbb{R}^3 \times \mathbb{R}, \mathbb{R}^3 \times \mathbb{R})$, is what we call a second-rank *4-tensor*. It has the following matrix structure:

$$[\boldsymbol{N}] = \begin{bmatrix} [\mathsf{N}] & \mathsf{v} \\ \mathsf{w}^{\top} & a \end{bmatrix},$$

where N is a second rank tensor from $\mathbb{R}^3$ onto itself, that is, $\mathsf{N} \in \mathrm{Hom}(\mathbb{R}^3, \mathbb{R}^3)$, while $\mathsf{v}, \mathsf{w} \in \mathbb{R}^3$ and $a \in \mathbb{R}$. The superscript $\top$ indicates standard transposition.

Consider $\boldsymbol{a} \in \mathbb{R}^3 \times \mathbb{R}$ as the value of a space-time differentiable field. Its four-dimensional derivative $D_4\tilde{\boldsymbol{a}} = D_4\tilde{\boldsymbol{a}}(x,t)$ has the following matrix structure:

$$[D_4\boldsymbol{a}] = \begin{bmatrix} [D\mathsf{a}] & \dot{\mathsf{a}} \\ Da^{\top} & \dot{a} \end{bmatrix}.$$

Its trace is the four-dimensional divergence, indicated by div_4 and given by

$$\mathrm{div}_4\boldsymbol{a} = \mathrm{div}\,\mathsf{a} + \dot{a}\,.$$

For χ the value of a space-time differentiable function $\tilde{\chi}$, we have

$$D_4\chi = (D\chi, \dot{\chi})\,.$$

Moreover, consider $\boldsymbol{N}$ as a space-time differentiable field. By definition, we have[1]

$$\mathrm{div}_4\boldsymbol{N} = (\mathrm{div}\,\mathsf{N} + \dot{\mathsf{v}}, \mathrm{div}\,\mathsf{w} + \dot{a})\,. \tag{5}$$

4. Inverse motion and a d'Alembert-Lagrange principle

The following formal developments provide the structural setting of the theory. Their implications for classical electroelasticity, variational formulations, and stability are made explicit in the subsequent sections.

4.1. A view on the inverse motion and polarization

Traditionally, a body is intended as an abstract manifold $\mathfrak{B}$ of finite dimension, parameterized by a reference configuration $\mathcal{B}$, a fit region. To render significant the statement that different observers (here considered to be frames of reference in the space-time) record the same reference configuration Truesdell and Noll (1965), Truesdell (1977), it is necessary to consider first two identical copies of the ambient space, say, $\mathbb{R}^3$ and $\bar{\mathbb{R}}^3$. We take $\bar{\mathbb{R}}^3$ as a reference space and select $\mathcal{B}$ by an embedding $\kappa : \mathfrak{B} \longrightarrow \bar{\mathbb{R}}^3$ of the body manifold, so that $\mathcal{B} = \kappa(\mathfrak{B})$. Then, motions with respect to $\mathcal{B}$ are defined by one-to-one differentiable maps

$$y : \mathcal{B} \times \mathbb{R} \longrightarrow \mathbb{R}^3,$$

assumed to preserve the space orientation. As usual, we indicate by F its spatial derivative. Thus, the orientation constraint formally reads $\det F > 0$. At every $t \in \mathbb{B}$, the fit region $\mathcal{B}_t := y(\mathcal{B}, t)$ represents the current configuration reached by the body.

The standard electroelastic formulation combines two geometrically different descriptions. Mechanical fields are commonly referred to a reference configuration, whereas electromagnetic fields are naturally expressed in the current configuration. The inverse motion provides a way to remove this asymmetry and to describe both classes of fields within a common geometric setting.

Here, we reverse the view. We consider a finite interval of time, say $[t_1, t_2]$, $t_1 < t_2$, and the 'tube'

$$\mathbb{B} := \bigcup_{t \in [t_1, t_2]} (t, \mathcal{B}_t),$$

where $\mathcal{B}_t$ is not necessarily identified by a motion $\tilde{y}$. Then, we consider a map

$$u : \mathbb{B} \longrightarrow \mathfrak{B}\,.$$

The value $u = u(x,t)$ indicates what material element $u \in \mathfrak{B}$ occupies the place $x \in \mathbb{R}^3$ at the instant $t \in \mathbb{R}$; in other words, if we would think in terms of the relativity theory, we should say that u shows what material element is involved in the event (x,t). We assume that u is differentiable and such that $\det Du > 0$ at every event (x,t). Differentiability of u is intended in terms of charts over $\mathfrak{B}$, as it is customary in defining differentiable maps between manifolds, because differentiability is a local property.

Polarization in the matter is described by a differentiable field

$$p : \mathbb{B} \longrightarrow B_p,$$

where B_p is a ball in $\mathbb{R}^3$ with radius $|p_{\max}|$, centered at zero. Precisely,

$$B_p := \{p \in \mathbb{R}^3 : \ |p| \leq |p_{\max}|\},$$

The field p represents the polarization $p = p(x,t)$ that pertains to the event $\bar{e} = (x,t)$; $|p_{\max}|$ indicates the maximum polarization density.

[1] A similar formalism has been adopted in Gurtin (1997).

4.2. The lagrangian

Let $\Omega \in \mathbb{R}^3$ be an open compact set. We consider a space-time tube $\mathbb{T} = \Omega \times [t_1, t_2]$ so wide to strictly include $\mathbb{B}$, that is, $\mathbb{B} \subset \mathbb{T}$. Then, we take three differentiable 4-vector fields

$$\boldsymbol{a} : \mathbb{R}^3 \times \mathbb{R} \longrightarrow \mathbb{R}^3 \times \mathbb{R}, \qquad \boldsymbol{u} : \mathbb{B} \longrightarrow \mathfrak{B} \times \{0\}, \qquad \boldsymbol{p} : \mathbb{B} \longrightarrow B_p \times \{0\},$$

so that

$$\boldsymbol{a} = \boldsymbol{a}(x,t) = (A, -\Phi), \qquad \boldsymbol{u} = \boldsymbol{u}(x,t) = (u, 0), \qquad \boldsymbol{p} = \boldsymbol{p}(x,t) = (p, 0),$$

where, as above, A and Φ are the vector and scalar electromagnetic potentials, respectively.

These fields enter the following general Lagrangian $\mathsf{L}(\mathbb{T})$:

$$\mathsf{L}(\mathbb{T}) := \int_{\mathbb{T}} \mathcal{L}_e(\boldsymbol{a}, D_4\boldsymbol{a}, \boldsymbol{p})\, \mathrm{d}\mu + \int_{\mathbb{B}} \mathcal{L}_m(x, \boldsymbol{u}, D_4\boldsymbol{u}, \boldsymbol{p}, D_4\boldsymbol{p})\, \mathrm{d}\mu,$$

where $\mathrm{d}\mu$ is the volume measure in $\mathbb{R}^3 \times \mathbb{R}$, that is, $\mathrm{d}\mu = \mathrm{d}x^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3 \wedge \mathrm{d}t$. The Lagrangian density $\mathcal{L}_e$ is assumed to be such that

$$\mathcal{L}_e(\boldsymbol{a}, D_4\boldsymbol{a}, \boldsymbol{p}) = \mathcal{L}_v(\boldsymbol{a}, D_4\boldsymbol{a})(1 - \mathbf{1}_{\mathbb{B}}) + \mathcal{L}_{em}(\boldsymbol{a}, D_4\boldsymbol{a}, \boldsymbol{p})\mathbf{1}_{\mathbb{B}}, \tag{6}$$

where $\mathbf{1}_{\mathbb{B}}$ is the indicator function of $\mathbb{B}$. $\mathcal{L}_v$ is the Lagrangian in the free space (possibly with isolated point charges) while $\mathcal{L}_{em}$ is the Lagrangian referring to electrodynamics in the matter. The latter depends on the polarization, which generates a local perturbation of the electric field.

The decomposition of $\mathcal{L}_e$ reflects the distinction between electromagnetic fields in the vacuum and electromagnetic interactions within the material body. It is not essential for the formal developments below, but it clarifies the physical origin of the various contributions and facilitates comparison with classical electroelastic formulations.

4.3. Gauge invariance of $\mathcal{L}_e$

We require gauge invariance for $\mathcal{L}_e$, which means

$$\mathcal{L}_e(\boldsymbol{a}, D_4\boldsymbol{a}, \boldsymbol{p}) = \mathcal{L}_e(\boldsymbol{a} + D_4\chi, D_4\boldsymbol{a} + D_4^2\chi, \boldsymbol{p}), \tag{7}$$

for any arbitrary choice of a twice differentiable space-time scalar function χ. The arbitrariness of χ implies the possibility of choosing $D_4\chi = -\boldsymbol{a}$, which also means $D_4^2\chi = -\mathrm{sym}D_4\boldsymbol{a}$, where sym extracts the symmetric part of its argument. Thus, the gauge condition (7) implies

$$\mathcal{L}_e(\boldsymbol{a}, D_4\boldsymbol{a}, \boldsymbol{p}) = \mathcal{L}_e(\mathrm{skew}D_4\boldsymbol{a}, \boldsymbol{p}),$$

where skew extracts the skew-symmetric part of its argument.

Set

$$\mathcal{F} := 2\mathrm{skew}D_4\boldsymbol{a}\,.$$

A common choice for $\mathcal{L}_v$ is a quadratic form proportional to $|\mathcal{F}|^2$. Other more intricate choices can be made for $\mathcal{L}_v$ as, for example, the Euler-Heisenberg structure, which describes the production of electron-positron pairs from a background electromagnetic field. For this reason, to leave open possibly different specific choices, we maintain $\mathcal{L}_e$ as general as possible in the following developments.

Note that

$$\mathrm{skew}D_4\boldsymbol{a} := \frac{1}{2}\begin{bmatrix} [DA] - [DA]^\top & \dot{A} + D\Phi \\ -(D\Phi + \dot{A})^\top & 0 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} [B\times] & -E \\ E^\top & 0 \end{bmatrix}. \tag{8}$$

Thus, there exist (1) a local differentiable function ℓ that depends on the triple (B, E, p) and (2) two other differentiable functions ℓ_v and ℓ_m such that

$$\mathcal{L}_e(\mathrm{skew}D_4\boldsymbol{a}, \boldsymbol{p}) = -\ell(B, E, p) = -\ell_v(B, E)(1 - \mathbf{1}_{\mathbb{B}}) - \ell_m(B, E, p)\mathbf{1}_{\mathbb{B}}\,.$$

Define the 4-tensor $\boldsymbol{T}$ as

$$\boldsymbol{T} := -\frac{\partial \mathcal{L}_e}{\partial \mathrm{skew}D_4\boldsymbol{a}}\,.$$

Evidently, $\boldsymbol{T}$ is a skew-symmetric tensor. Thus, in agreement with the structure of the last term in (8), we can say that there are vectors, say H and D such that

$$\boldsymbol{T} := \begin{bmatrix} [\mathsf{H}\times] & \mathsf{D} \\ -\mathsf{D}^\top & 0 \end{bmatrix}. \tag{9}$$

In the free space, at every t we have

$$\dot{\mathcal{L}}_v(\mathrm{skew}D_4\boldsymbol{a}) = -\dot{\ell}_v(B, E),$$

which reads

$$\boldsymbol{T}\cdot\overset{\cdot}{\overline{\operatorname{skew}D_4\boldsymbol{a}}}=\frac{\partial\mathscr{E}_v}{\partial\boldsymbol{B}}\cdot\dot{\boldsymbol{B}}+\frac{\partial\mathscr{E}_v}{\partial\boldsymbol{E}}\cdot\dot{\boldsymbol{E}}\,. \tag{10}$$

Since $(\mathbf{v}\times)\cdot(\mathbf{w}\times)=2$ for any pair of vectors $(\mathbf{v},\mathbf{w})$, the Eq. (10) becomes

$$\mathsf{H}\cdot\dot{\boldsymbol{B}}-\mathsf{D}\cdot\dot{\boldsymbol{E}}=\frac{\partial\mathscr{E}_v}{\partial\boldsymbol{B}}\cdot\dot{\boldsymbol{B}}+\frac{\partial\mathscr{E}_v}{\partial\boldsymbol{E}}\cdot\dot{\boldsymbol{E}}$$

and is valid for any choice of the time rates involved. Thus, the following identifications hold:

$$\mathsf{H}=\frac{\partial\mathscr{E}_v}{\partial\boldsymbol{B}},\qquad \mathsf{D}=-\frac{\partial\mathscr{E}_v}{\partial\boldsymbol{E}}\,.$$

In this sense, H and D are *dual* to $\boldsymbol{B}$ and $\boldsymbol{E}$. The vector D is commonly called an *electric displacement*. In contrast, there is no universal consensus on the nomenclature attributed to the magnetic dual H, that is a *magnetic field*.[2]

The 4-tensor field with values $\boldsymbol{T}$ suffers a bounded discontinuity across the surface

$$\Gamma:=\bigcup_{t\in[t_1,t_2]}(t,\partial\mathcal{B}_t),$$

while for the rest it is continuous and continuously differentiable. The field $\tilde{\boldsymbol{a}}$, above defined, is instead considered continuous everywhere.

The variational structure considered below includes processes that may involve charge transport and electric currents. For this reason, the electromagnetic field alone is not sufficient to describe the interaction mechanism and an additional source field must be introduced. We therefore consider a 4-current density field whose role is analogous to that played by external actions in the classical d'Alembert-Lagrange principle.

A 4-vector field $\tilde{\boldsymbol{j}}:\mathfrak{E}\longrightarrow\mathfrak{E}$ with values $\boldsymbol{j}$ is defined by

$$\boldsymbol{j}=(\mathrm{j},\mathfrak{c}),$$

where $\mathrm{j}\in\mathbb{R}^3$ is the vector of *electric currents*, while $\mathfrak{c}\in\mathbb{R}$ is the scalar density of charges that induce polarization. In dielectrics, the case treated by Ericksen in Ericksen (2007b), $\mathrm{j}=0$ and there are only bound charges. In conducting ferroelectrics, a case encompassed in the present formalism, j may be different from zero. In any case, outside the body considered, we can possibly have currents in the free space and charges. In the free space, we consider $\tilde{\boldsymbol{j}}$ continuously differentiable.

4.4. A d'Alembert-Lagrange principle

We assume that physically realizable paths in the chosen state space, that is, the space of triples $(\tilde{\boldsymbol{a}},\tilde{\boldsymbol{u}},\tilde{\boldsymbol{p}})$, satisfy the following d'Alembert-Lagrange principle:

$$\delta\mathsf{L}(\mathbb{T})+\int_{\mathbb{T}}\boldsymbol{j}\cdot\delta(\boldsymbol{a}+D_4\Phi)\,\mathrm{d}\mu=0\,. \tag{11}$$

The choice of a d'Alembert-Lagrange principle, rather than a purely Hamiltonian stationarity condition, is motivated by the possibility of including non-conservative current-related effects while preserving a variational structure.

Variations are considered as

$$\boldsymbol{a}\longmapsto\boldsymbol{a}+\delta\boldsymbol{a},\quad \boldsymbol{u}\longmapsto\boldsymbol{u}+\delta\boldsymbol{u},\quad \boldsymbol{p}\longmapsto\boldsymbol{p}+\delta\boldsymbol{p},\quad \Phi\longmapsto\Phi+\delta\Phi,$$

where $\delta\boldsymbol{a}$ and $\delta\Phi$ are assumed to vanish on $\partial\mathbb{T}$, while $\boldsymbol{u}+\delta\boldsymbol{u}$ and $\boldsymbol{p}+\delta\boldsymbol{p}$ are required to coincide with values $\bar{\boldsymbol{u}}$ and $\bar{\boldsymbol{p}}$ on a portion $\partial_{pre}\mathbb{B}$ of the boundary $\partial_{pre}\mathbb{B}$, where, in principle, the spatial components u and p are prescribed. In particular, we consider $\delta\boldsymbol{u}=(\delta u,0)$ and $\delta\boldsymbol{p}=(\delta p,0)$.

Remark 1. If we would skip the last integral in (11), the first variation of the Lagrangian alone would not obtain the continuity law for the electric current (see below), as it occurs in classical variational derivation of the Maxwell equations. A d'Alembert-Lagrange structure allows one to include non-conservative effects.

5. The first variation

The purpose of this section is to derive the governing equations in a representation where both electromagnetic and mechanical fields are described in the current configuration. In this way, the geometric advantage of the inverse-motion formulation becomes explicit already at the level of the variational principle.

[2] Some authors call H the *magnetic field*, saying that $\boldsymbol{B}$ is the *magnetic induction*. Others claim that the induction is properly a phenomenon and call $\boldsymbol{B}$ a *magnetic field*, denoting H as a *magnetic intensity*. We do not take part in the diatribe, accepting the adopted compromise.

5.1. Dirichlet boundary conditions along $\partial\mathbb{B}$

For convenience, first we separate the various addenda of $\mathsf{L}(\mathbb{T})$. The technique that we use in this section is a standard matter in the calculus of variation.

For the integral

$$\int_{\mathbb{T}} \mathcal{L}_e(\operatorname{skew} D_4\boldsymbol{a}, \boldsymbol{p})\,\mathrm{d}\mu$$

we have

$$\begin{aligned}\delta\int_{\mathbb{T}} \mathcal{L}_e\,\mathrm{d}\mu &= \int_{\mathbb{T}} \frac{\partial \mathcal{L}_e}{\partial \operatorname{skew} D_4\boldsymbol{a}}\cdot \operatorname{skew} D_4\delta\boldsymbol{a}\,\mathrm{d}\mu + \int_{\mathbb{B}} \frac{\partial \mathcal{L}_e}{\partial \boldsymbol{p}}\cdot\delta\boldsymbol{p}\,\mathrm{d}\mu = \int_{\mathbb{T}} \frac{\partial \mathcal{L}_e}{\partial \operatorname{skew} D_4\boldsymbol{a}}\cdot D_4\delta\boldsymbol{a}\,\mathrm{d}\mu + \int_{\mathbb{B}} \frac{\partial \mathcal{L}_e}{\partial \boldsymbol{p}}\cdot\delta\boldsymbol{p}\,\mathrm{d}\mu\\ &= \int_{\mathbb{T}} \left(\operatorname{div}_4\left(\left(\frac{\partial \mathcal{L}_e}{\partial \operatorname{skew} D_4\boldsymbol{a}}\right)^{\top}\delta\boldsymbol{a}\right) - \delta\boldsymbol{a}\cdot\operatorname{div}_4\frac{\partial \mathcal{L}_e}{\partial \operatorname{skew} D_4\boldsymbol{a}}\right)\mathrm{d}\mu + \int_{\mathbb{B}} \frac{\partial \mathcal{L}_e}{\partial \boldsymbol{p}}\cdot\delta\boldsymbol{p}\,\mathrm{d}\mu\\ &= \int_{\mathbb{T}} \delta\boldsymbol{a}\cdot\operatorname{div}_4\boldsymbol{T}\,\mathrm{d}\mu + \int_{\Gamma} [\![\,\boldsymbol{T}\,]\!]\boldsymbol{n}_{space}\cdot\delta\boldsymbol{a}\,\mathrm{d}\mathcal{H}^3 + \int_{\mathbb{B}} \frac{\partial \mathcal{L}_e}{\partial \boldsymbol{p}}\cdot\delta\boldsymbol{p}\,\mathrm{d}\mu,\end{aligned} \tag{12}$$

where $\mathrm{d}\mathcal{H}^3 = \mathrm{d}\mathcal{H}^2\wedge\mathrm{dt}$, with $\mathrm{d}\mathcal{H}^2$ the two-dimensional Hausdorff measure over $\partial\mathcal{B}_t$ at every instant t, the 4-vector $\boldsymbol{n}_{space}$ is given by $\boldsymbol{n}_{space} = (\mathbf{n}, 0)$, with $\mathbf{n}$ the normal to $\partial\mathcal{B}_t$ at every instant t and in every point where it is well defined, and $[\![\,\boldsymbol{T}\,]\!]$ indicates the jump of $\boldsymbol{T}$ across Γ, that is, $[\![\,\boldsymbol{T}\,]\!] = \boldsymbol{T}^{+} - \boldsymbol{T}^{-}$, where $\boldsymbol{T}^{\pm}$ are, respectively, the outer and inner limits of $\boldsymbol{T}$ to $\partial\mathcal{B}_t$ at every instant t, with respect to the normal $\mathbf{n}$.

Assume for simplicity that

$$\partial_{pre}\mathbb{B} = \partial\mathbb{B}\,. \tag{13}$$

In other words, Dirichlet-type boundary conditions apply for the fileds $\bar{\boldsymbol{u}}$ abd $\bar{\boldsymbol{p}}$

For the integral

$$\int_{\mathbb{B}} \mathcal{L}_m(x, \boldsymbol{u}, D_4\boldsymbol{u}, \boldsymbol{p}, D_4\boldsymbol{p})\,\mathrm{d}\mu,$$

due to the condition (13), since the variations are chosen to match the boundary conditions, we have

$$\begin{aligned}\delta\int_{\mathbb{B}} \mathcal{L}_m\,\mathrm{d}\mu &= \int_{\mathbb{B}} \left(\frac{\partial \mathcal{L}_m}{\partial \boldsymbol{u}}\cdot\delta\boldsymbol{u} + \frac{\partial \mathcal{L}_m}{\partial D_4\boldsymbol{u}}\cdot\delta D_4\boldsymbol{u} + \frac{\partial \mathcal{L}_m}{\partial \boldsymbol{p}}\cdot\delta\boldsymbol{p} + \frac{\partial \mathcal{L}_m}{\partial D_4\boldsymbol{p}}\cdot\delta D_4\boldsymbol{p}\right)\mathrm{d}\mu\\ &= \int_{\mathbb{B}} \left(\delta\boldsymbol{u}\cdot\left(\frac{\partial \mathcal{L}_m}{\partial \boldsymbol{u}} - \operatorname{div}_4\frac{\partial \mathcal{L}_m}{\partial D_4\boldsymbol{u}}\right) + \delta\boldsymbol{p}\cdot\left(\frac{\partial \mathcal{L}_m}{\partial \boldsymbol{p}} - \operatorname{div}_4\frac{\partial \mathcal{L}_m}{\partial D_4\boldsymbol{p}}\right)\right)\mathrm{d}\mu\\ &= \int_{\mathbb{B}} \left(\delta\boldsymbol{u}\cdot\left(\frac{\partial \mathcal{L}_m}{\partial \boldsymbol{u}} + \operatorname{div}_4\boldsymbol{P}\right) + \delta\boldsymbol{p}\cdot\left(\frac{\partial \mathcal{L}_m}{\partial \boldsymbol{p}} + \operatorname{div}_4\boldsymbol{K}\right)\right)\mathrm{d}\mu,\end{aligned} \tag{14}$$

where

$$\boldsymbol{P} := -\frac{\partial \mathcal{L}_m}{\partial D_4\boldsymbol{u}} \qquad \text{and} \qquad \boldsymbol{K} := -\frac{\partial \mathcal{L}_m}{\partial D_4\boldsymbol{p}}\,.$$

The resulting Euler–Lagrange equations separate naturally into three groups: electromagnetic balances, mechanical balances associated with inverse motion, and balances of microstructural interactions associated with polarization.

Therefore, the variational principle (11) becomes

$$\begin{aligned}&\int_{\mathbb{T}} \delta\Phi\cdot\operatorname{div}_4\boldsymbol{j}\,\mathrm{d}\mu + \int_{\mathbb{T}} \delta\boldsymbol{a}\cdot(\operatorname{div}_4\boldsymbol{T} + \boldsymbol{j})\,\mathrm{d}\mu + \int_{\Gamma} [\![\,\boldsymbol{T}\,]\!]\boldsymbol{n}_{space}\cdot\delta\boldsymbol{a}\,\mathrm{d}\mathcal{H}^3\\ &\qquad + \int_{\mathbb{B}} \left(\delta\boldsymbol{u}\cdot\left(\frac{\partial \mathcal{L}_m}{\partial \boldsymbol{u}} + \operatorname{div}_4\boldsymbol{P}\right) + \delta\boldsymbol{p}\cdot\left(\frac{\partial \mathcal{L}_m}{\partial \boldsymbol{p}} + \frac{\partial \mathcal{L}_e}{\partial \boldsymbol{p}} + \operatorname{div}_4\boldsymbol{K}\right)\right)\mathrm{d}\mu = 0.\end{aligned}$$

The arbitrariness of $\mathbb{T}$ and the test field considered imply in the free space

$$\operatorname{div}_4\boldsymbol{T} + \boldsymbol{j} = 0 \tag{15}$$

and

$$\operatorname{div}_4\boldsymbol{j} = 0\,; \tag{16}$$

in $\mathbb{B}$ we obtain

$$\frac{\partial \mathcal{L}_m}{\partial \boldsymbol{u}} + \operatorname{div}_4\boldsymbol{P} = 0 \tag{17}$$

and

$$\frac{\partial \mathcal{L}_m}{\partial \boldsymbol{p}} + \frac{\partial \mathcal{L}_e}{\partial \boldsymbol{p}} + \operatorname{div}_4\boldsymbol{K} = 0, \tag{18}$$

while, along Γ, we get

$$[\![\, \boldsymbol{T} \,]\!]\boldsymbol{n}_{space} = 0\,. \tag{19}$$

When sufficient regularity is granted, on the basis of the previous local balance equations, we compute the following identities:

$$\mathrm{div}_4((\mathrm{skew}D_4\boldsymbol{a})^\top\boldsymbol{T}) = -(\mathrm{skew}D_4\boldsymbol{a})^\top\boldsymbol{j} + \boldsymbol{T}D_4\mathrm{skew}D_4\boldsymbol{a},$$

$$\mathrm{div}_4((D_4\boldsymbol{u})^\top\boldsymbol{P}) = -(D_4\boldsymbol{u}^\top)\frac{\partial\mathcal{L}_m}{\partial\boldsymbol{u}} + \boldsymbol{P}D_4^2\boldsymbol{u},$$

$$\mathrm{div}_4((D_4\boldsymbol{p})^\top\boldsymbol{K}) = -(D_4\boldsymbol{p})^\top\left(\frac{\partial\mathcal{L}_m}{\partial\boldsymbol{p}} + \frac{\partial\mathcal{L}_e}{\partial\boldsymbol{p}}\right) + \boldsymbol{K}D_4^2\boldsymbol{p}\,.$$

Since, for $\mathcal{L} = \mathcal{L}_e + \mathcal{L}_m$, we have in $\mathcal{B}$

$$\begin{aligned} D_4\mathcal{L} &= \frac{\delta\mathcal{L}}{\partial\mathrm{skew}D_4\boldsymbol{a}}D_4\mathrm{skew}D_4\boldsymbol{a} + \frac{\delta\mathcal{L}}{\partial\boldsymbol{u}}D_4\boldsymbol{u} + \frac{\delta\mathcal{L}}{\partial D_4\boldsymbol{u}}D_4^2\boldsymbol{u} + \frac{\delta\mathcal{L}}{\partial\boldsymbol{p}}D_4\boldsymbol{p} + \frac{\delta\mathcal{L}}{\partial D_4\boldsymbol{p}}D_4^2\boldsymbol{p} + \frac{\delta\mathcal{L}}{\partial x} \\ &= -\boldsymbol{T}D_4\mathrm{skew}D_4\boldsymbol{a} + \frac{\delta\mathcal{L}}{\partial\boldsymbol{u}}D_4\boldsymbol{u} - \boldsymbol{P}D_4D_4\boldsymbol{u} + \frac{\delta\mathcal{L}}{\partial\boldsymbol{p}}D_4\boldsymbol{p} - \boldsymbol{K}D_4^2\boldsymbol{p} + \frac{\delta\mathcal{L}}{\partial x}, \end{aligned}$$

defining

$$\boldsymbol{M} := \mathcal{L}\boldsymbol{I}_4 + (\mathrm{skew}D_4\boldsymbol{a})^\top\boldsymbol{T} + (D_4\boldsymbol{u})^\top\boldsymbol{P} + (D_4\boldsymbol{p})^\top\boldsymbol{K},$$

with $\boldsymbol{I}_4$ the 4×4 identity tensor and summing up the previous identities, we obtain

$$\mathrm{div}_4\boldsymbol{M} - \frac{\delta\mathcal{L}}{\partial x} = -(\mathrm{skew}D_4\boldsymbol{a})^\top\boldsymbol{j}. \tag{20}$$

When there are no currents or localized charges beyond the dipoles in $\mathbb{B}$, and the body is homogeneous, the previous balance reduces to the conservation law

$$\mathrm{div}_4\boldsymbol{M} = 0. \tag{21}$$

The tensor $\boldsymbol{M}$ is gauge invariant, due to the presence of $\mathrm{skew}D_4\boldsymbol{a}$.

Remark 2. If we would start above from considering the identity

$$\mathrm{div}_4((D_4\boldsymbol{a})^\top\boldsymbol{T}) = -(D_4\boldsymbol{a})^\top\boldsymbol{j} + \boldsymbol{T}D_4^2\boldsymbol{a},$$

instead of (20), we would obtain

$$\mathrm{div}_4(\mathcal{L}\boldsymbol{I}_4 + (D_4\boldsymbol{a})^\top\boldsymbol{T} + (D_4\boldsymbol{u})^\top\boldsymbol{P} + (D_4\boldsymbol{p})^\top\boldsymbol{K}) - \frac{\delta\mathcal{L}}{\partial x} = -(D_4\boldsymbol{a})^\top\boldsymbol{j},$$

which is not gauge invariant. Then, imposing gauge invariance, that is, requiring that the equation would remain invariant when we substitute $D_4\boldsymbol{a}$ with $D_4\boldsymbol{a} + D_4\chi$, we would obtain once again (20).

5.2. Decomposition into space and time components

On the basis of (5) and taking into account (9), since $\mathrm{div}(\mathsf{v}\times) = -\mathrm{curl}\mathsf{v}$, for any $\mathsf{v}\in\mathbb{R}^3$, the Eq. (15) gives

$$\mathrm{curl}\mathsf{H} = \dot{\mathsf{D}} + \mathsf{j} \tag{22}$$

and

$$\mathrm{div}\mathsf{D} = \mathfrak{c}. \tag{23}$$

Also, from the definitions (1) and (3), we obtain

$$\mathrm{div}\boldsymbol{B} = 0 \tag{24}$$

and

$$\mathrm{curl}E = -\dot{B}, \tag{25}$$

while Eq. (16) explicitly reads

$$\mathrm{div}\mathsf{j} = -\dot{\mathfrak{c}}, \tag{26}$$

which complete the system of Maxwell equations. On the basis of (19), along Γ we also have at every instant t

$$[\mathsf{H}]\times\mathbf{n} = 0 \qquad \text{and} \qquad [\mathsf{D}]\cdot\mathbf{n} = 0 \qquad \text{along}\ \partial\mathcal{B}_t. \tag{27}$$

The Eq. (17) reads

$$\mathrm{div}P + b = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial\mathcal{L}}{\partial\dot{u}}, \tag{28}$$

where $P := -\frac{\partial \mathcal{L}}{\partial Du}$ is a stress associated with the gradient of inverse motion, that is, F^{-1}, and $b := \frac{\partial \mathcal{L}}{\partial u}$ is the (co)vector of non-inertial body forces in the present representation.

Indeed, since $F^{-1} = \mathrm{adj}F(\det F)^{-1}$, distinguishing between space and time, we can intend $\mathcal{L}$ as a function of the type

$$\mathcal{L}(Du, \cdots) = f(\mathrm{adj}F, (\det F)^{-1}, \cdots),$$

in which $\cdots$ represent the other state variables. Thus, we have

$$P = \frac{\partial \mathcal{L}}{\partial Du} = \frac{1}{(\det F)^2}\frac{\partial f}{\partial (\det F)^{-1}}\mathrm{cof}F - \frac{\partial f}{\partial \mathrm{adj}F}\frac{\mathrm{d}(\mathrm{adj}F)}{\mathrm{d}F}.$$

Finally, the Eq. (18) of microstructural interactions is, indeed,

$$\mathrm{div}S - \mathbf{z} + \boldsymbol{\beta} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{p}}, \tag{29}$$

where $S := -\frac{\partial \mathcal{L}}{\partial Dp}$, $\mathbf{z}$ is a dipole self-action while $\boldsymbol{\beta}$ is an external action on the polarization; they are such that

$$\boldsymbol{\beta} = \frac{\partial \mathcal{L}_e}{\partial p} \qquad \text{and} \qquad \mathbf{z} = -\frac{\partial \mathcal{L}_m}{\partial p}.$$

Remark 3. The balance of microstructural interactions could include Coriolis type inertial effects associated with the spin of $\boldsymbol{p}$, not explored here.

5.3. Mixed boundary conditions along Γ

Assume that along a set $\bar{\Gamma} \subset \Gamma$ no data are assigned. In this case, the d'Alembert-Lagrange principle gives

$$\int_{\mathbb{T}} \delta\Phi \cdot \mathrm{div}_4 \boldsymbol{j}\,\mathrm{d}\mu + \int_{\mathbb{T}} \delta\boldsymbol{a}\cdot(\mathrm{div}_4\boldsymbol{T} + \boldsymbol{j})\,\mathrm{d}\mu + \int_{\Gamma} [\![\boldsymbol{T}]\!]\boldsymbol{n}_{space}\cdot\delta\boldsymbol{a}\,\mathrm{d}\mathcal{H}^3 + \int_{\mathbb{B}}\left(\delta\boldsymbol{u}\cdot\left(\frac{\partial \mathcal{L}_m}{\partial \boldsymbol{u}} + \mathrm{div}_4\boldsymbol{P}\right) + \delta\boldsymbol{p}\cdot\left(\frac{\partial \mathcal{L}_m}{\partial \boldsymbol{p}} + \frac{\partial \mathcal{L}_e}{\partial \boldsymbol{p}} + \mathrm{div}_4\boldsymbol{K}\right)\right)\mathrm{d}\mu$$
$$+ \int_{\bar{\Gamma}}\left(\delta\boldsymbol{u}\cdot\frac{\partial \mathcal{L}_m}{\partial D_4\boldsymbol{u}}\boldsymbol{n} + \delta\boldsymbol{p}\cdot\frac{\partial \mathcal{L}_m}{\partial D_4\boldsymbol{p}}\boldsymbol{n}\right)\mathrm{d}\mathcal{H}^3 = 0.$$

Along $\bar{\Gamma}$ the test fields $\delta\boldsymbol{u}$ and $\delta\boldsymbol{p}$ can be arbitrary. However, due to their structure, this arbitrariness is only in the spatial component. Thus, in addition to the already determined equations, we should have two additional natural conditions:

$$\frac{\partial \mathcal{L}_m}{\partial D_4\boldsymbol{u}}\boldsymbol{n}_{space} = 0 \qquad \text{and} \qquad \frac{\partial \mathcal{L}_m}{\partial D_4\boldsymbol{p}}\boldsymbol{n}_{space} = 0 \qquad \text{along}\,\bar{\Gamma}\,. \tag{30}$$

Indeed, we can have along $\bar{\Gamma}$ a distributed dead load with density $\boldsymbol{f}$. In this case, we can consider a modified d'Alembert-Lagrange principle, that is,

$$\delta\mathsf{L}(\mathbb{T}) + \int_{\mathbb{T}} \boldsymbol{j}\cdot\delta(\boldsymbol{a} + D_4\Phi)\,\mathrm{d}\mu - \int_{\bar{\Gamma}} \boldsymbol{f}\cdot\delta\boldsymbol{u}\,\mathrm{d}\mathcal{H}^3 = 0. \tag{31}$$

Then, the natural boundary condition $(30)_1$ becomes

$$\frac{\partial \mathcal{L}_m}{\partial D_4\boldsymbol{u}}\boldsymbol{n}_{space} = \boldsymbol{f}.$$

We do not suggest an analogous treatment for $(30)_2$ because we do not know any loading device able to act directly on the polarization by contact on the body boundary.

We could have along $\bar{\Gamma}$ also live loads. In other words, we could prescribe, for example that Du, the spatial part of $D_4\boldsymbol{u}$ be a certain function $\boldsymbol{h}(u)$ or even that an analogous condition is prescribed for Dp, provided that this last condition could physically be realizable. In these cases, the addition to $\mathcal{L}$ of an appropriate null Lagrangian would accommodate the variational principle in order to fit the prescribed boundary conditions (techniques in Olver (2024) are appropriate to represent and analyze this case). It is a full divergence of an appropriate function. In terms of Du, the null Lagrangian should be a linear combination of all minors of Du. When we would also involve Dp, we should consider a mixed combination of all minors of the two gradients (see, e.g., Sharma and Basak 2021).

What has been developed so far contains Ericksen's theory of elastic dielectrics as a special case, obtained by suppressing currents, restricting attention to static processes, and adopting the specific energy structure proposed in Ericksen (2007b). In this sense, the present framework may be viewed as a nonlinear and dynamically consistent extension of Ericksen's formulation. At the same time, because Ericksen's theory was originally introduced as a reformulation of Toupin's electroelasticity Toupin (1956, 1960) from the viewpoint of energy minimization, the present construction establishes a direct bridge between Toupin's variational framework and modern electroelastic models involving flexoelectric and microstructural effects. The next section makes this connection explicit.

6. Comparison with classical electroelasticity

6.1. Comparison with large strain electrostatics

The purpose of this section is twofold. First, we clarify the relation between the present inverse-motion formulation and the classical theories of Toupin (1956, 1960) and Ericksen (2007b). Second, we show that the proposed framework recovers the standard electroelastic models in the small-strain regime.

Indicate by X the point associated with the coordinate chart around a generic material element in $\mathcal{B} = \kappa(\mathfrak{B})$.

The standard approach to elastic dielectrics is formulated in terms of the direct motion $y = y(X, t)$ and is based on a variational principle leading to the governing field equations and boundary conditions (Toupin, 1956, 1960).

A natural observation made by Ericksen (2007b) is that the stationarity conditions obtained in the classical formulation do not necessarily identify minimizers of the total energy. Rather, they may correspond to more general critical points, including saddle points. This distinction becomes important whenever stability issues are considered.

Along the same path indicated by Ericksen (2007b), in the present framework the deformation gradient is expressed as $F = (Du)^{-1}$, rather than $F = Dy$.

Consider a static problem with energy written in the reference configuration,

$$\mathcal{E}[y, p] = \int_{\mathcal{B}_0} W(F, p, Dp)\, dX.$$

Under the change of variables

$$x = y(X, t)$$

and the identification

$$u(x) = X,$$

we obtain the equivalent representation

$$\mathcal{E}[u, p] = \int_{\mathcal{B}} \widetilde{W}(F^{-1}, p, Dp)\, dx.$$

Thus, the inverse-motion formulation does not alter the physical content of the theory. It simply expresses it directly in the current configuration.

In Toupin's formulation the electrostatic potential enters as an independent variable whose variation generates Gauss' law, whereas in Ericksen's treatment the electromagnetic contribution is incorporated directly into an energy functional whose minimizers represent stable equilibrium states. Consequently, the inverse-motion framework retains the general variational structure of Toupin's theory while allowing, in appropriate settings, a direct interpretation in terms of energy minimization as advocated by Ericksen. In particular, this perspective clarifies the role of electroelastic couplings in stability analysis, where the distinction between general stationary points and energy minimizers becomes crucial.

Indeed, as already observed by Toupin (1956) and further remarked by Ericksen (2007b), the stationarity conditions obtained in the classical formulation do not necessarily correspond to a minimum of the total energy. In particular, the coupling between mechanical and electromagnetic fields may lead to a saddle-point structure rather than to a pure minimization problem. In Toupin's formulation Toupin (1956), the variational principle involves variables with different energetic roles, so that the resulting equilibrium conditions are naturally of min–max type. In static settings, one may either retain the full saddle-point structure or, under appropriate assumptions, reduce the problem to a minimization principle in the sense advocated by Ericksen.

Formally, the energy may be viewed as a functional $\mathcal{E}(u, p, \phi)$ that is convex in some variables and concave in others, leading naturally to a saddle-point problem.

The distinction between the present approach and classical electroelastic theories becomes immaterial in the small-strain regime, where the two representations effectively coincide.

6.2. Small-strain setting and recovery of classical models

The present formulation reduces to the standard theory of electroelasticity in the small-strain regime.

To see this, consider the inverse motion written as

$$u(x) = x - v(x),$$

where v is a displacement field.

Then, $Du = I - Dv$ and, for small strains, namely when $|Dv| \ll 1$, we obtain

$$F = (Du)^{-1} \approx I + Dv.$$

The linearized strain tensor is

$$\varepsilon = \frac{1}{2}\left(Dv + (Dv)^T\right),$$

so that the kinematics coincides with the classical infinitesimal theory.

Consider an energy density of the form $W(F, p, Dp)$. Expanding around the reference configuration yields

$$W \approx W_0 + \frac{1}{2}(\mathbb{C}\varepsilon)\cdot\varepsilon + \frac{\alpha}{2}|p|^2 + \frac{\beta}{2}|Dp|^2 + \gamma\, p\cdot D(\operatorname{tr}\varepsilon),$$

which corresponds to a standard electroelastic-type energy. The associated Euler–Lagrange equations are

$$\operatorname{div}(\mathbb{C}\varepsilon) + \gamma D(\operatorname{div} p) = 0,$$

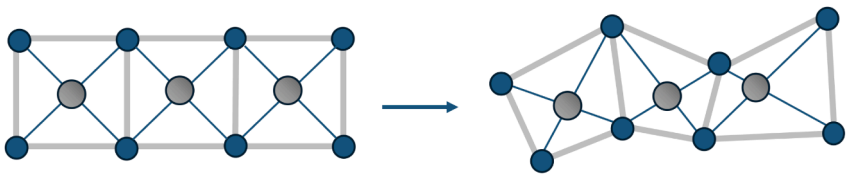

Fig. 1. Breaking symmetry of a centrosymmetric square lattice due to nonuniform strain.

and

$$-\beta\Delta p + \alpha p + \gamma D(\operatorname{tr}\varepsilon) = 0.$$

These equations coincide with the standard linearized electroelastic system obtained from classical formulations. Thus, in the small-strain limit, the inverse-motion formulation yields the same governing equations as the classical theory. This shows that the present framework is a consistent nonlinear extension of standard models rather than an alternative constitutive theory.

Consequently, the inverse-motion representation does not introduce a new electroelasticity. It provides instead a reformulation in which mechanical and electromagnetic fields share the same geometric setting while preserving the classical electroelastic limit.

Gauge invariance plays here a role analogous to the one of objectivity and covariance, commonly adopted in continuum mechanics. In the present formulation, it implies that the electromagnetic contribution depends on the physical fields rather than on the specific choice of potentials. This restriction is reflected directly in the admissible constitutive structure.

7. Considering flexoelectricity

In 1880, the brothers Pierre and Paul-Jacques Curie discovered that uniform strain may induce polarization in certain classes of materials: the phenomenon is what we call piezoelectricity. In its simplest occurrence there is a linear link between the strain (second-rank) tensor and polarization vector. The coefficient of proportionality is thus a third-rank tensor. Constitutive tensors of even rank are admitted only when material elements are not centrosymmetric. Crystallographic group theory is discriminant regarding this aspect. Thus, centrosymmetric materials do not exhibit piezoelectric effects under uniform strain.

However, nonuniform strain may break mirror symmetry and determine polarization through the contribution of second-neighbor interactions measured by the dependence of the energy on strain gradients (Fig. 1 provides a pictorial representation of the way a nonuniform strain may induce polarization by breaking symmetry). This effect is called *flexoelectricity* (for a clean overview of the relevant phenomenology and technological perspectives, see Krichen and Sharma 2016, Wang et al. 2019).

Flexoelectricity provides a particularly convenient testing ground for the inverse-motion formulation. Indeed, unlike standard dielectric theories, it naturally involves higher gradients and latent microstructural effects. Consequently, it offers a concrete setting in which the variational advantages of the present framework become visible.

In the common description of flexoelectricity, the energy is assumed to depend on the second derivative of the deformation. In other words, the resulting scheme is that of a continuum endowed with two microstructures: one is the polarization, described by $\boldsymbol{p}$, while the other is *latent*. By this we mean that the phase field describing this additional (virtual) microstructure depends on the deformation gradient (in other words, there is an internal constraint (Capriz, 1985)). Second-neighbor interactions become relevant (Mariano, 2017).

To account for flexoelectricity within the general framework discussed above, it is therefore necessary to include the second spatial derivative of $\boldsymbol{u}$. A few formal variations are consequently required.

7.1. The basic case

Consider the spatial projector Π^s with matrix structure

$$[\Pi^s] = \begin{bmatrix} [I] & \mathbf{0} \\ \mathbf{0}^\top & 0 \end{bmatrix},$$

where I is the 3×3 identity and $\mathbf{0}$ is the null vector in $\mathbb{R}^3$.

The 4-tensor

$$D_4\boldsymbol{u}^s := D_4\boldsymbol{u}\,\Pi^s$$

has the matrix structure

$$[D_4\boldsymbol{u}^s] = \begin{bmatrix} [D\mathfrak{a}] & \mathbf{0} \\ \mathbf{0}^\top & 0 \end{bmatrix}.$$

In the inverse-motion representation, a natural extension of the Lagrangian includes $D_4^2\boldsymbol{u}^s$.

To include the flexoelectric effect, the Lagrangian $\mathsf{L}(\mathbb{T})$ changes into

$$\mathsf{L}(\mathbb{T}) := \int_{\mathbb{T}} \mathcal{L}_e(\boldsymbol{a}, D_4\boldsymbol{a}, \boldsymbol{p})\,\mathrm{d}\mu + \int_{\mathbb{B}} \mathcal{L}_m(x, \boldsymbol{u}, D_4\boldsymbol{u}, D_4^2\boldsymbol{u}^s, \boldsymbol{p}, D_4\boldsymbol{p})\,\mathrm{d}\mu, \tag{32}$$

where

$$D_4^2\boldsymbol{u}^s = D_4(D_4\boldsymbol{u}^s).$$

To compute the first variation of $\mathsf{L}(\mathbb{T})$, we can proceed as above. As a result, we derive equations structurally identical to (15), (16), and (18), while (17) changes into

$$\frac{\partial \mathcal{L}_m}{\partial \boldsymbol{u}} + \mathrm{div}_4\big(\boldsymbol{P} - \mathrm{div}_4\mathfrak{P}\big) = 0. \tag{33}$$

Here, $\mathfrak{P}$, the so-called hyperstress, is a third-rank 4-tensor with structure $\mathfrak{P} = (\mathbf{S}, 0)$, given by

$$\mathfrak{P} := -\frac{\partial \mathcal{L}}{\partial D_4^2\boldsymbol{u}^s}.$$

Distinguishing space and time, (33) reduces to

$$\mathrm{div}(P - \mathrm{div}\mathbf{S}) + b = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{u}}. \tag{34}$$

The appearance of the hyperstress tensor $\mathfrak{P}$ is the direct signature of second-neighbor interactions. In this sense, flexoelectricity is represented as a higher-gradient effect associated with a latent microstructure whose descriptor is constrained to depend on the deformation gradient.

The natural boundary conditions remain the same. However, the computation of the first variation imposes an additional natural condition:

$$\frac{\partial \mathcal{L}_m}{\partial D_4^2\boldsymbol{u}}\boldsymbol{n}_{\text{space}} = \mathbf{0} \qquad \text{along } \bar{\Gamma}. \tag{35}$$

The Eq. (20) becomes

$$\mathrm{div}_4\boldsymbol{Z} - \frac{\delta \mathcal{L}}{\partial x} = -(\mathrm{skew}D_4\boldsymbol{a})^\top \boldsymbol{j}, \tag{36}$$

with

$$\boldsymbol{Z} := \mathcal{L}\,\boldsymbol{I}_4 + (\mathrm{skew}D_4\boldsymbol{a})^\top\boldsymbol{T} + (D_4\boldsymbol{u})^\top(\boldsymbol{P} - \mathrm{div}_4\mathfrak{P}) + (D_4\boldsymbol{p})^\top\boldsymbol{K}.$$

Remark 4. Since $\boldsymbol{u}$ represents the inverse motion, Eq. (17) (see also (28)) has configurational nature. Indeed, it results from variations involving $\mathfrak{B}$, which is associated with a reference configuration once embedded in a linear space, commonly $\mathbb{R}^3$. The same interpretation applies to (33).

In contrast, Eq. (18) is the balance of microstructural interactions associated with polarization and expressed in Eulerian representation (see also Davì and Mariano 2001, Mariano 2024). Thus, $\boldsymbol{M}$ and $\boldsymbol{K}$ are stresses associated with forces.

When linear constitutive structures are accepted in the small-strain regime, odd constitutive tensors appear. In Hong and Vanderbilt (2013), Stengel (2013a) (see also Stengel 2014, 2013b), a consistent and nearly complete framework for deriving them from first principles has been proposed. The derivation procedure involves only lists of atoms, presumed placements in a unit supercell, and the corresponding atomic numbers.

The present formulation is therefore compatible with modern first-principles descriptions of flexoelectricity, where the relevant higher-order constitutive tensors are derived from atomistic information rather than postulated phenomenologically (Hong and Vanderbilt, 2013; Stengel, 2013a, 2014).

7.2. Flexoelectricity in nematics

In liquid crystals, "the combination of molecular asymmetry and broken orientational symmetry produces a nonzero electric polarization that couples to either bend or splay" (Harden et al., 2010) through the flexoelectric moduli. This phenomenon is particularly evident in bent-core nematics. They are thermotropic, banana-shaped mesogenes with polar and chiral supramolecular structure, despite being achiral.

Nematic flexoelectricity provides a useful example of the flexibility of the inverse-motion framework. In this case, polarization is no longer the only microstructural descriptor. A second field describing orientational order must be introduced and coupled to the electromechanical response.

With $\mathbb{S}^2$ the unit sphere in $\mathbb{R}^3$, a new differentiable phase field

$$m : \mathbb{B} \longrightarrow \mathbb{S}^2,$$

needs to be considered. When we account for the hand-to-tail symmetry of nematic molecules, as a target space, the unit sphere must be substituted by the projective space obtained by identifying the antipodes in $\mathbb{S}^2$. The field m describes the (averaged) orientation m of stick molecules within the material element that occupies x in the instant t. It is considered as the space-like component of a 4-vector field

$$\boldsymbol{m} : \mathbb{B} \longrightarrow \mathbb{S}^2 \times \{0\},$$

with values given by

$$\boldsymbol{m} = \tilde{\boldsymbol{m}}(x,t) = (m,0)\,.$$

The field m plays the role of an order parameter describing the local orientational state of the nematic phase. In this sense, the theory becomes a multi-field electroelastic model involving both polarization and orientational microstructure. It enters the Lagrangian, which is now

$$\mathsf{L}(\mathbb{T}) := \int_{\mathbb{T}} \mathcal{L}_e(\boldsymbol{a}, D_4\boldsymbol{a}, \boldsymbol{p})\,\mathrm{d}\mu + \int_{\mathbb{B}} \mathcal{L}_m(x, \boldsymbol{u}, D_4\boldsymbol{u}, \boldsymbol{p}, D_4\boldsymbol{p}, \boldsymbol{m}, D_4\boldsymbol{m})\,\mathrm{d}\mu\,.$$

The d'Alembert-Langrange principle (11) implies the results mentioned above and an additional balance equation given by

$$\frac{\partial \mathcal{L}_m}{\partial \boldsymbol{m}} + \mathrm{div}_4 \overline{\boldsymbol{K}} = 0, \tag{37}$$

where

$$\overline{\boldsymbol{K}} := -\frac{\partial \mathcal{L}_m}{\partial D_4 \boldsymbol{m}},$$

and a natural boundary condition:

$$\frac{\partial \mathcal{L}_m}{\partial D_4 \boldsymbol{m}} \boldsymbol{n}_{space} = 0 \qquad \text{along } \Gamma\,.$$

The Eq. (36) changes into

$$\mathrm{div}_4 \bar{\boldsymbol{Z}} - \frac{\delta \mathcal{L}}{\partial x} = -(\mathrm{skew} D_4\boldsymbol{a})^{\top}\boldsymbol{j}. \tag{38}$$

where, now,

$$\bar{\boldsymbol{Z}} := \mathcal{L}\boldsymbol{I}_4 + (\mathrm{skew} D_4\boldsymbol{a})^{\top}\boldsymbol{T} + (D_4\boldsymbol{u})^{\top}(\boldsymbol{P} - \mathrm{div}_4 \mathfrak{P}) + (D_4\boldsymbol{p})^{\top}\boldsymbol{K} + D_4\boldsymbol{m}^{\top}\bar{\boldsymbol{K}}.$$

Once again, when $\mathcal{L}$ does not depend on x, in the absence of currents, we have

$$\mathrm{div}_4 \bar{\boldsymbol{Z}} = 0.$$

Distinguishing space and time components, (37) becomes

$$\mathrm{div}\bar{\mathrm{S}} - \bar{\mathrm{z}} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \mathcal{L}}{\partial \dot{m}}, \tag{39}$$

where

$$\bar{\mathrm{S}} = \frac{\partial \mathcal{L}_m}{\partial Dm} \qquad \text{and} \qquad \bar{\mathbf{z}} = -\frac{\partial \mathcal{L}_m}{\partial m}.$$

We do not include here an external bulk force that act directly on the nematic order.

The sequence

classical nematic flexoelectricity ⟶ inverse motion ⟶ variational formulation

is summarized in Fig. 2.

8. Electroelastic coupling: Energy minimization and material stability

The previous sections have shown that the inverse-motion formulation provides a variational framework for nonlinear electroelasticity and its flexoelectric extensions. A natural question, already emphasized by Ericksen in his re-examination of elastic dielectrics, is whether such a framework can be used to characterize stable equilibrium states through energy minimization. The purpose of this section is to show that this is indeed the case for a representative class of electroelastic energies.

As already noted, the governing equations of electroelasticity arise as stationarity conditions of an energy functional. However, stationarity does not in general imply stability: depending on the structure of the functional, equilibrium configurations may correspond to minima, maxima, or saddle points. In particular, the coupling between mechanical and electromagnetic fields leads, in general, to a saddle-point structure of the variational problem, as emphasized in classical analyses.

Let (u, p) be an equilibrium configuration and consider admissible variations $(\delta u, \delta p)$. Stability may be analyzed through the second variation of the energy, that is, $\delta^2\mathcal{E}[u,p](\delta u, \delta p)$. A sufficient condition for stability is the non-negativity of the second variation for all admissible perturbations. In the present framework, electroelastic interactions introduce additional contributions to $\delta^2\mathcal{E}$, which may be grouped as follows:

- standard elastic terms, associated with Du;
- polarization terms, associated with p and Dp;
- coupling terms, involving both deformation and polarization gradients;

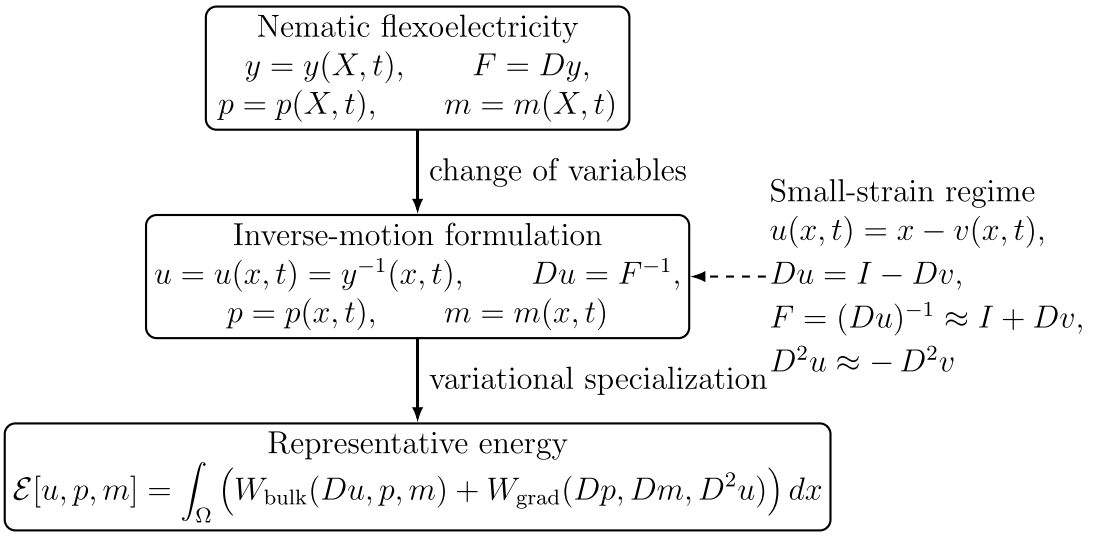

Fig. 2. Relation between the classical description of nematic flexoelectricity and the inverse-motion formulation. In the small-strain regime, the distinction between direct and inverse descriptions becomes "essentially" negligible, while the inverse formulation keeps all fields in the current configuration and preserves a direct variational structure.

- higher-order terms, involving D^2u in the case of flexoelectric effects.

The higher-order terms provide a regularizing contribution, introducing an intrinsic length scale, while the coupling terms may either enhance or reduce the effective stiffness of the material, as shown in Mariano et al. (2025).

In the linearized regime, stability may be characterized either through the positivity properties of the acoustic tensor or, equivalently, by the coercivity of the associated quadratic energy functional.

These observations are consistent with general results on materials with microstructure, where additional fields and higher-order gradients may both stabilize and destabilize the response, depending on the coupling parameters. In the present setting, flexoelectricity can be interpreted as a manifestation of latent microstructure, whose influence on stability is captured directly at the level of the energy functional.

This perspective clarifies that the inverse-motion formulation is not only a geometrically consistent description of electroelasticity, but also a natural framework for analyzing the stability of coupled electromechanical systems. In particular, stability requires that $\delta^2\mathcal{E}$ be non-negative on the admissible subspace compatible with the constraints imposed by the electromagnetic fields.

8.1. A representative quadratic energy and small-strain stability

We illustrate the stability implications of electroelastic couplings through a representative quadratic energy in the infinitesimal regime. Let v be the displacement field already mentioned in Section 6.2, and $\varepsilon(v) = \frac{1}{2}(Dv + (Dv)^{\top})$. On $\mathcal{H} := H^2_0(\Omega;\mathbb{R}^3) \times H^1_0(\Omega;\mathbb{R}^3)$, define

$$\mathcal{Q}(v,p) = \int_\Omega \left(\frac{1}{2}\,(\mathbb{C}\varepsilon)\cdot\varepsilon + \frac{\alpha}{2}|p|^2 + \frac{\beta}{2}|Dp|^2 + \gamma\, p\cdot D(\operatorname{tr}\varepsilon) + \frac{\eta}{2}|D^2v|^2 \right) dx, \tag{40}$$

with $\alpha, \beta, \eta > 0$ and $\mathbb{C}$ satisfying strong ellipticity on symmetric tensors: $(\mathbb{C}A)\cdot A \geq c_{\rm el}|A|^2$ for any second-rank tensor A, with $c_{\rm el}$ a positive constant.

Proposition 1. *If*

$$\alpha\eta > 3\gamma^2, \tag{41}$$

then $\mathcal{Q}$ is coercive on $\mathcal{H}$, i.e., there exists $c_* > 0$ such that

$$\mathcal{Q}(v,p) \geq c_* \left(\|\varepsilon(v)\|^2_{L^2(\Omega)} + \|D^2v\|^2_{L^2(\Omega)} + \|p\|^2_{L^2(\Omega)} + \|Dp\|^2_{L^2(\Omega)} \right) \tag{42}$$

for all $(v,p) \in \mathcal{H}$. Thus, $(v,p) = (0,0)$ is a strict minimizer of $\mathcal{Q}$. Equivalently, the second variation of the energy is positive definite on $\mathcal{H}$.

The functional (40) may be regarded as the quadratic counterpart of the flexoelectric energies discussed in the previous section. It contains the four basic ingredients of the theory: elasticity, polarization, electromechanical coupling, and higher-gradient regularization.

Proof. Since $\operatorname{tr}\varepsilon(v) = \operatorname{div} v$, for each $i = 1,2,3$, we have $\partial_i(\operatorname{div} v) = \partial_{ij}v_j$, where summation over repeated indices is adopted. By the Cauchy–Schwarz inequality, $\left|\partial_i(\operatorname{div} v)\right|^2 \le 3|\partial_{ij}v_j|^2$. Summing over i gives $|D(\operatorname{div} v)|^2 \le 3\,|D^2 v|^2$ a.e. in $\mathcal{B}$, so that

$$\|D(\operatorname{div} v)\|_{L^2}^2 \le 3\|D^2 v\|_{L^2}^2.$$

Hence, by the Cauchy–Schwarz inequality, we have

$$\left|\int_\Omega \gamma\, p\cdot D(\operatorname{tr}\varepsilon(v))\,dx\right| \le \sqrt{3}\,|\gamma|\,\|p\|_{L^2}\,\|D^2 v\|_{L^2}.$$

Therefore,

$$\begin{aligned}\frac{\alpha}{2}\|p\|_{L^2}^2 + \frac{\eta}{2}\|D^2 v\|_{L^2}^2 + \gamma\int_{\mathcal{B}} p\cdot D(\operatorname{tr}\varepsilon(v))\,dx \\ \ge \frac{\alpha}{2}\|p\|_{L^2}^2 + \frac{\eta}{2}\|D^2 v\|_{L^2}^2 - \sqrt{3}\,|\gamma|\,\|p\|_{L^2}\,\|D^2 v\|_{L^2}.\end{aligned}$$

The right-hand side can be written as the 2×2 block

$$\frac{1}{2}\begin{pmatrix}\|p\|_{L^2}\\ \|D^2 v\|_{L^2}\end{pmatrix}^{\top}\begin{pmatrix}\alpha & -\sqrt{3}|\gamma|\\ -\sqrt{3}|\gamma| & \eta\end{pmatrix}\begin{pmatrix}\|p\|_{L^2}\\ \|D^2 v\|_{L^2}\end{pmatrix}.$$

Under (41), this matrix is positive definite. Let $\lambda_{\min}(M) > 0$ denote its smallest eigenvalue. Then,

$$\frac{\alpha}{2}\|p\|_{L^2}^2 + \frac{\eta}{2}\|D^2 v\|_{L^2}^2 + \gamma\int_\Omega p\cdot D(\operatorname{tr}\varepsilon(v))\,dx \ge \frac{\lambda_{\min}(M)}{2}\Big(\|p\|_{L^2}^2 + \|D^2 v\|_{L^2}^2\Big). \tag{43}$$

The remaining terms are nonnegative: $\frac{1}{2}\int \mathbb{C}\varepsilon(v) : \varepsilon(v) \ge \frac{c_{\mathrm{el}}}{2}\|\varepsilon(v)\|_{L^2}^2$ and $\frac{\beta}{2}\|Dp\|_{L^2}^2$. Combining the previous estimates with the definition of $\mathcal{Q}$, we obtain

$$\mathcal{Q}(v,p) \ge \frac{c_{\mathrm{el}}}{2}\,\|\varepsilon(v)\|_{L^2}^2 + \frac{\beta}{2}\,\|Dp\|_{L^2}^2 + \frac{\lambda_{\min}(M)}{2}\Big(\|p\|_{L^2}^2 + \|D^2 v\|_{L^2}^2\Big).$$

Hence (42) holds with

$$c_* = \frac{1}{2}\min\{c_{\mathrm{el}}, \beta, \lambda_{\min}(M)\} > 0.$$

This proves the coercivity estimate.

Since $\mathcal{Q}$ is a quadratic functional, the second variation at the reference state is simply

$$\delta^2\mathcal{Q}(0,0)[(v,p),(v,p)] = 2\,\mathcal{Q}(v,p),$$

so (42) implies that the second variation is positive definite on $\mathcal{H}$. Therefore, $(0,0)$ is a strict minimizer of $\mathcal{Q}$. □

Proposition 1 shows explicitly how the competition between flexoelectric coupling and higher-gradient regularization determines the stability region. In this sense, the variational structure introduced through inverse motion is not merely a geometric reformulation. It provides direct information on the stability of electroelastic equilibria.

Remark 5. Condition (41) makes explicit the competition between stabilizing quadratic terms and the flexoelectric coupling γ: increasing $|\gamma|$ shrinks the stability region. This provides a concrete small-strain counterpart of the general picture discussed above.

Remark 6. The factor 3 in (41) stems from $\|D(\operatorname{div} v)\|_{L^2}^2 \le 3\|D^2 v\|_{L^2}^2$ in three dimensions. If the higher-order term is written as $\frac{\eta}{2}|D(\operatorname{tr}\varepsilon(v))|^2$, the sharper condition $\alpha\eta > \gamma^2$ is obtained.

8.2. *Explicit energy minimization for electroelastic coupling described through inverse motion*

To exhibit in a concrete way the variational content of the inverse-motion formulation, we consider here a static version of the free energy density introduced in Rahmati et al. (2025) for flexoelectricity with nematic order and optical effects. In the version here, we avoid considering an optical term in the original formulation and the nematic order m, here considered as a frozen direction at every x. Thus, the energy considered is

$$\mathcal{F}(u,p) := \int_\Omega \Big(\widehat{W}_{\mathrm{el}}(x,u,Du) + \widehat{W}_{\mathrm{pol}}(x,u,p) + \widehat{W}_{\mathrm{flex}}(x,u,Du,p) + \frac{\beta}{2}|Dp|^2\Big)\,dx, \tag{44}$$

where $\beta > 0$ and

$$\widehat{W}_{\mathrm{el}}(x,u,Du) := \Phi\big(x,u,M(Du)\big), \tag{45}$$

$$\widehat{W}_{\mathrm{pol}}(x,u,p) := \Psi_{\mathrm{pol}}(x,u,p), \tag{46}$$

$$\widehat{W}_{\mathrm{flex}}(x,u,Du,p) := \Psi_{\mathrm{flex}}(x,u,Du,p). \tag{47}$$

Here, $M(Du)$ denotes the 3-vector constituted by the minors of Du (see Giaquinta et al. 1998),

$$M(Du) = (Du, \operatorname{cof} Du, \det Du),$$

and $\Psi_{\rm pol}$ is a measurable polarization energy density, while $\Psi_{\rm flex}$ is a measurable flexoelectric contribution. The inverse-motion rewriting presented above of the energy in Rahmati et al. (2025) is immediate: every occurrence of F^{-1} in the direct-motion representation is replaced by Du, while every occurrence of F is replaced by $(Du)^{-1}$.

This example plays a role analogous to that of the quadratic model considered in the previous subsection. There, the emphasis was on stability. Here, the emphasis is on existence. Together, the two examples show how the inverse-motion framework accommodates both stability and minimization issues in electroelasticity.

A possible expression of the energy is a special case of one discussed in Rahmati et al. (2025):

$$\Psi_{\rm pol}(x,u,p)=\frac{1}{2}\,(\det Du)\,p\cdot\mathbb{A}(m(x))\,p,$$

with $\mathbb{A}(m(x))$ uniformly positive definite, and

$$\Psi_{\rm flex}(x,u,Du,p)=J_u\,(m(x)\otimes p)\,:\,\mathbb{F}\big((Dm(x))(Du)^{-1}\big).$$

The theorem below is however stated in a more general form.

We consider Dirichlet-type data u_0 and p_0, respectively over portions $\partial\Omega_u$ and $\partial\Omega_p$ of $\partial\Omega$. We will tacitly intend them in the sense of traces, when appropriate. In particular, we need to consider $p_0\in W^{1/2,2}(\partial\Omega_p;B_p)$. To discuss the existence of minimizers, we need to define an *admissible functional class* $\mathcal{A}$ of (u,p) pairs. Thus, we define

$$\mathcal{A}:=\Big\{(u,p)\;\Big|\;u\in{\rm dif}^{\,r,1}(\Omega,\mathcal{B}),\;\;p\in W^{1,2}(\Omega;B_p),\;\;u=u_0\;\text{on}\;\partial\Omega_u,\;\;p=p_0\;\text{on}\;\partial\Omega_p\Big\},$$

where $W^{1,2}(\Omega;B_p)$ means that $p\in W^{1,2}(\Omega;\mathbb{R}^3)$ and $p(x)\in B_p$ for a.e. $x\in\Omega$, while ${\rm dif}^{\,r,1}$ denotes the class of weak diffeomorphisms with summable minors. Specifically, ${\rm dif}^{\,1,1}$ is a subclass of $W^{1,1}(\Omega;\mathbb{R}^3)$. It contains orientation-preserving deformations u, which satisfy an integral condition that avoids self-penetration, are such that $|M(Du)|\in L^1(\Omega;\mathbb{R}^3)$ and do not show discontinuities in the sense that their graphs have no 'holes' or 'fractures'; the class ${\rm dif}^{\,1,1}$ is complete (weakly sequentially closed) (Giaquinta et al., 1998).[3] Then, ${\rm dif}^{\,r,1}$ is defined as the class of those maps in ${\rm dif}^{\,1,1}$ with $|M(Du)|\in L^r(\Omega;\mathbb{R}^3)$.[4]

We assume what follows:

(H1) $\Phi(x,u,\cdot)$ is polyconvex with respect to Du, namely it is convex as a function of $M(Du)=(Du,{\rm cof}\,Du,\det Du)$;

(H2) Φ is measurable in x, lower semicontinuous in $(u,M(Du))$, and there exist constants $c_1>0$, $c_2\geq 0$, exponents $r>1$, and a convex function $\vartheta:(0,+\infty)\to[0,+\infty]$ with $\vartheta(t)\to+\infty$ as $t\to 0^+$, such that

$$\Phi(x,u,M(Du))\geq c_1|M(Du)|^r+\vartheta(\det Du)-c_2; \tag{48}$$

(H3) $\Psi_{\rm pol}:\Omega\times\mathcal{B}\times B_p\to\mathbb{R}$ is measurable in x, lower semicontinuous in (u,p), and bounded from below, namely there exists $c_{\rm pol}\geq 0$ such that

$$\Psi_{\rm pol}(x,u,p)\geq -c_{\rm pol}\qquad\text{for all admissible }(x,u,p);$$

(H4) $\Psi_{\rm flex}$ is measurable in x and lower semicontinuous in $(u,M(Du),p)$; moreover, it is bounded from below by an L^1-function uniformly on the admissible class, namely there exists $\zeta\in L^1(\Omega)$ such that

$$\Psi_{\rm flex}(x,u,Du,p)\geq\zeta(x)\qquad\text{for all admissible }(u,p);$$

(H5) there exists one pair $(u_*,p_*)\in\mathcal{A}$ such that $\mathcal{F}(u_*,p_*)<+\infty$.

The assumptions below are standard in the direct method of the calculus of variations. Their role is to isolate the structural features of the inverse-motion formulation from technical constitutive choices.

Theorem 1. *Under assumptions* (H1)–(H5)*, the functional $\mathcal{F}$ attains its minimum on $\mathcal{A}$.*

Proof. By (H5), the infimum $\mu:=\inf_{(u,p)\in\mathcal{A}}\mathcal{F}(u,p)$ is finite. Let $\{(u_j,p_j)\}\subset\mathcal{A}$ be a minimizing sequence, so that $\mathcal{F}(u_j,p_j)\to\mu$. Since each p_j is admissible, we have $p_j(x)\in B_p$ for a.e. $x\in\Omega$. Therefore $\{p_j\}$ is bounded in $L^\infty(\Omega;\mathbb{R}^3)$, and in particular in $L^2(\Omega;\mathbb{R}^3)$.

Moreover, since $\beta>0$ and the minimizing sequence has bounded energy, the term

$$\int_\Omega\frac{\beta}{2}|Dp_j|^2\,dx$$

[3] Indeed, in defining ${\rm dif}^{\,1,1}$, we could directly refer to the body manifold $\mathfrak{B}$ because the Sobolev space $W^{1,1}(\Omega,\mathfrak{B})$ can be intrinsically defined for maps that take values on a finite-dimensional differentiable Riemannian manifold, provided that the manifold itself is geodesically complete (see, for example, Mariano and Mucci 2025 and relevant references quoted therein).

[4] In the standard approach of the direct method, minimizers are in $W^{1,p}$, with $p\geq 3$. For $p<3$ one cannot exclude that minimizers admit holes. Referring to Cartesian currents, as done here, although left unexpressed for reducing the analytical weight of the reading, allows one to define a subclass of $W^{1,1}$, with the triple $(F,{\rm cof}\,F,\det F)\in L^r$. This is a way to select a subspace of $W^{1,p}$, with $p>1$, of maps that avoid the non-required presence of holes, as the very concept of elasticity imposes (otherwise we were referring to elastic-brittle materials). These maps have the minimal analytical regularity needed to formulate the mechanical setting. Indeed, the choice of a functional class has per se constitutive character.

is uniformly bounded. It follows that $\{p_j\}$ is bounded in $W^{1,2}(\Omega;\mathbb{R}^3)$. On the mechanical side, from (48) and the boundedness of $\mathcal{F}(u_j,p_j)$ we deduce

$$\int_\Omega |M(Du_j)|^r\,dx \le C, \qquad \int_\Omega \vartheta(\det Du_j)\,dx \le C,$$

for some constant $C>0$ independent of j. Hence $\{u_j\}$ is bounded in the class $\mathrm{dif}^{r,1}(\Omega,\mathcal{B})$.

By the compactness and closure properties of weak diffeomorphisms (Giaquinta et al., 1998), up to extraction of a (not relabeled) subsequence, there exists $u\in \mathrm{dif}^{r,1}(\Omega,\mathcal{B})$ such that $u_j \rightharpoonup u$ and $M(Du_j)\rightharpoonup M(Du)$, in the relevant weak topologies, and u satisfies the prescribed trace condition.

Likewise, from the boundedness of $\{p_j\}$ in $W^{1,2}(\Omega;\mathbb{R}^3)$, there exists $p\in W^{1,2}(\Omega;\mathbb{R}^3)$ such that, up to extraction of a (possibly not relabeled) subsequence, we get

$$p_j \rightharpoonup p \quad \text{in } W^{1,2}(\Omega;\mathbb{R}^3), \qquad p_j \to p \quad \text{in } L^2(\Omega;\mathbb{R}^3).$$

Passing to a further subsequence if necessary, we may also assume that $p_j(x)\to p(x)$ for a.e. $x\in\Omega$. Since B_p is closed and $p_j(x)\in B_p$ for a.e. x, it follows that

$$p(x)\in B_p \qquad \text{for a.e. } x\in\Omega.$$

Hence, $p\in W^{1,2}(\Omega;B_p)$ and the trace condition on $\partial\Omega_p$ is preserved. Therefore, $(u,p)\in\mathcal{A}$.

We now pass to the limit in the energy terms.

By (H1)–(H2), the map $(u,Du)\mapsto \Phi(x,u,M(Du))$ is polyconvex with respect to Du, measurable in x, and lower semicontinuous. Therefore, by standard semicontinuity results (Giaquinta et al., 1998),

$$\int_\Omega \Phi(x,u,M(Du))\,dx \le \liminf_{j\to\infty}\int_\Omega \Phi(x,u_j,M(Du_j))\,dx.$$

Since Ψ_{pol} is lower semicontinuous in (u,p) and bounded from below, while $p_j\to p$ strongly in $L^2(\Omega;\mathbb{R}^3)$ and u_j converges in the admissible class, we obtain

$$\int_\Omega \Psi_{\mathrm{pol}}(x,u,p)\,dx \le \liminf_{j\to\infty}\int_\Omega \Psi_{\mathrm{pol}}(x,u_j,p_j)\,dx.$$

For the flexoelectric term, assumption (H4) gives directly

$$\int_\Omega \Psi_{\mathrm{flex}}(x,u,Du,p)\,dx \le \liminf_{j\to\infty}\int_\Omega \Psi_{\mathrm{flex}}(x,u_j,Du_j,p_j)\,dx.$$

Finally, the Dirichlet term in Dp is convex and weakly lower semicontinuous in $W^{1,2}(\Omega;\mathbb{R}^3)$, so

$$\int_\Omega \frac{\beta}{2}|Dp|^2\,dx \le \liminf_{j\to\infty}\int_\Omega \frac{\beta}{2}|Dp_j|^2\,dx.$$

Collecting the previous inequalities yields

$$\mathcal{F}(u,p)\le \liminf_{j\to\infty}\mathcal{F}(u_j,p_j)=\mu.$$

Since $(u,p)\in\mathcal{A}$, the reverse inequality $\mu\le\mathcal{F}(u,p)$ holds by definition of infimum. Therefore, we have $\mathcal{F}(u,p)=\mu$ and (u,p) is a minimizer. □

Theorem 1 shows that the inverse-motion formulation is compatible with the standard existence theory of nonlinear elasticity. In particular, the passage from direct to inverse motion does not alter the variational structure required by the direct method, while naturally accommodating electroelastic and flexoelectric couplings.

Remark 7. The pointwise constraint $p(x)\in B_p$ is physically natural, since the polarization intensity cannot exceed a material-dependent saturation value. From the variational viewpoint, the constraint is also convenient because the admissible class is closed under almost-everywhere convergence, so that it is preserved in the direct method.

Remark 8. The theorem above does not claim to cover the whole functional adopted in Rahmati et al. (2025). Its role is different: it shows that, after reducing the full photo-flexoelectric theory to a representative static electroelastic sector and rewriting it in inverse-motion variables, one obtains a genuine minimization problem with a standard existence theory. The quadratic polarization term used in Rahmati et al. (2025) is recovered as a particular constitutive choice within the more general class covered here.

Remark 9. Theorem 1 places the present electroelastic model within the broader variational theory of complex bodies developed in Mariano and Modica (2009). The novelty here lies in the simultaneous presence of polarization effects, flexoelectric couplings, and regularization through the inverse-motion description.

8.3. An explicit flexoelectric variant

Consider a variant of the previous energy where $\Psi_{\rm flex}(x,u,Du,p)$ is substituted by

$$\Psi_{\rm flex}(x,u,Du,D^2u,p)=\Psi_{\rm flex}(x,u,M(Du),D(\text{cof }Du),D(\det Du),p),$$

which is what we refer to as gradient polyconvexity. By maintaining previous assumption (H1)-(H3) and (H5), here we assume that there are constants $q>2$, $s>0$, $c_i>0$, $i=1,2$, such that

$$\Psi_{\rm flex}(x,u,M(Du),\Delta_1,\Delta_2,p)\geq c_1(|M(Du)|^{\bar r}+|\Delta_1|^q-|\Delta_2|^s)+\vartheta(\det Du)-c_2,$$

where θ is as in (H1). In this context, minimizers u belong to $\text{dif}^{r,2,p}$, that is, u is in a subset of the Sobolev space $W^{2,p}$, with $p>2$, such that cof Du is in $W^{1,q}$, $q\geq p/(p-1)$ and det Du is in L^r. The functional choices of the polarization field remain invariant. The proof is based on tools discussed in Kruzik et al. (2022) with reference to a different setting.

9. A relativistic extension

The inverse-motion formulation developed above admits a natural covariant extension. The purpose of this section is not to develop a complete relativistic theory of electroelasticity, but rather to show that the geometric structure introduced in the previous sections remains compatible with a relativistic setting.

Relativistic formulations of elasticity have been discussed by several authors (Carter and Quintana, 1972; Maugin, 1977; Kijowski and Magli, 1992; Beig and Schmidt, 2003), including media with microstructure and Cosserat-type effects (Maugin, 1978; Maitra and Tromp, 2024). Here we restrict attention to the structural aspects needed to extend the inverse-motion framework.

Instead of the derivatives D and D_4, we consider a four-dimensional spacetime manifold $\mathfrak{E}$ endowed with a Lorentzian metric $\boldsymbol{g}$ and associated Levi–Civita connection. The corresponding covariant derivative is denoted by ∇.[5]

The inverse motion is now interpreted as a submersion

$$\boldsymbol{u}:\mathfrak{E}\longrightarrow\mathfrak{B},$$

whose fibres represent the world-lines of material particles. The electromagnetic potential remains a spacetime field $\boldsymbol{a}$, while polarization is described by a spacetime field $\boldsymbol{p}$.

The classical derivatives appearing in the electroelastic Lagrangian are replaced by their covariant counterparts,

$$D_4\boldsymbol{u}\longrightarrow\nabla\boldsymbol{u},\qquad D_4\boldsymbol{p}\longrightarrow\nabla\boldsymbol{p},$$

while the flexoelectric contribution involves the spatial restriction of the second covariant derivative,

$$D_4^2\boldsymbol{u}^s\longrightarrow\nabla^2\boldsymbol{u}^s.$$

The relativistic counterpart of the electroelastic Lagrangian is then

$$\mathsf{L}(\mathbb{T}):=k\int_{\mathbb{T}}R\sqrt{-g}\,\mathrm{d}\mu_{\boldsymbol{g}}+\int_{\mathbb{T}}\mathcal{L}\sqrt{-g}\,\mathrm{d}\mu_{\boldsymbol{g}},\tag{49}$$

where R is the scalar curvature associated with the metric $\boldsymbol{g}$, $k>0$ contains the gravitational constant, $g=\det\boldsymbol{g}$, and $\mathcal{L}$ denotes the relativistic electroelastic density.

The electromagnetic contribution may be written in the standard covariant form

$$\mathcal{L}_v=-\frac{1}{4}\mathcal{F}^\flat\cdot\mathcal{F},\qquad\mathcal{F}:=2\,\text{skew}\,\nabla\boldsymbol{a},$$

where $\mathcal{F}^\flat=\boldsymbol{g}\mathcal{F}\boldsymbol{g}$.

The d'Alembert–Lagrange principle becomes

$$\delta\mathsf{L}(\mathbb{T})+\int_{\mathbb{T}}\boldsymbol{j}\cdot\delta(\boldsymbol{a}+D_4\Phi)\sqrt{-g}\,\mathrm{d}\mu_{\boldsymbol{g}}=0.\tag{50}$$

By taking variations with respect to the metric, the electromagnetic potential, the inverse motion and the polarization field, one obtains the Einstein equations

$$\boldsymbol{R}-\frac{1}{2}\boldsymbol{g}\,R=\kappa^{-1}\mathbb{P},$$

where

$$\mathbb{P}=\frac{1}{2}\boldsymbol{g}\,\mathcal{L}-\frac{\partial\mathcal{L}}{\partial\boldsymbol{g}}$$

is the energy–momentum tensor of the coupled system.

At the same time, the electroelastic balances retain the same structure as in the classical theory, namely

$$\text{div}_4\hat{\boldsymbol{T}}+\boldsymbol{j}=0,\tag{51}$$

[5] For simplicity we adopt units in which the speed of light is equal to one.

$$\mathrm{div}_4 \boldsymbol{j} = 0, \tag{52}$$

$$\frac{\partial \mathcal{L}_m}{\partial \boldsymbol{u}} + \mathrm{div}_4\left(\hat{\boldsymbol{P}} - \mathrm{div}_4 \hat{\mathfrak{P}}\right) = 0, \tag{53}$$

and

$$\frac{\partial \mathcal{L}_m}{\partial \boldsymbol{p}} + \frac{\partial \mathcal{L}_e}{\partial \boldsymbol{p}} + \mathrm{div}_4 \hat{\boldsymbol{K}} = 0. \tag{54}$$

Therefore, the inverse-motion formulation extends naturally to a covariant setting in which electroelastic, flexoelectric and microstructural effects coexist with gravitational interactions.

A detailed analysis of the resulting relativistic theory, including its conservation laws and geometric properties, lies beyond the scope of the present work and will be discussed elsewhere.

10. Concluding remarks

The formulation proposed here originates from a question already highlighted by Ericksen in his re-examination of elastic dielectrics: how should electroelastic interactions be represented when stability and energy minimization are regarded as central issues rather than mere consequences of stationarity conditions? To address this question, we have adopted an inverse-motion representation in which mechanical and electromagnetic fields are described directly in the current configuration. The resulting framework is gauge invariant, variationally consistent, and naturally suited to the treatment of electroelastic couplings involving polarization, flexoelectricity, and latent microstructure.

The proposed formulation recovers the classical electroelastic theory in the small-strain regime, while providing a unified setting for nonlinear problems. In particular, it allows one to connect electroelastic interactions with questions of stability and energy minimization. This aspect has been illustrated through explicit coercivity results for a representative quadratic flexoelectric energy and through an existence theorem for a broader class of nonlinear electroelastic functionals. The analysis therefore extends Ericksen's proposal for elastic dielectrics to nonlinear electroelastic media including ferroelectric and flexoelectric effects, also in the presence of currents. At the same time, it establishes a direct link between inverse motion, gauge invariance, and the variational structure of electroelasticity.

Finally, the geometric nature of the construction makes possible a covariant extension to relativistic settings. Although only briefly outlined here, this extension suggests further developments at the interface between continuum mechanics, electromagnetism, and field theory.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work falls within programs of the research group on "Theoretical Mechanics" in the "Centro di Ricerca Matematica Ennio De Giorgi" of the Scuola Normale Superiore, Pisa. GNFM-INDAM is acknowledged.

Data availability

No data was used for the research described in the article.

References

Beig, R., Schmidt, B.G., 2003. Relativistic elasticity. Class. Quantum Gravity 20, 889–904.
Brown, W.F., 1966. Magnetoelastic Interactions. Springer Verlag, New York.
Bustamante, R., Dorfmann, A., Ogden, R.W., 2009. Nonlinear electroelastostatics: a variational framework. Z. Angew. Math. Phys. 60, 154–177.
Capriz, G., 1985. Continua with latent microstructure. Arch. Rational Mech. Anal. 90, 43–56.
Carter, B., Quintana, H., 1972. Foundations of general relativistic high-pressure elasticity theory. Proc. Royal Soc. A 331, 57–83.
Davì, F., Mariano, P.M., 2001. Evolution of domain walls in ferroelectric solids. J. Mech. Phys. Solids 49, 1701–1726.
Dorfmann, L., Ogden, R.W., 2006. Nonlinear electroelastic deformations. J. Elasticity 82, 99–127.
Dorfmann, L., Ogden, R.W., 2014. Nonlinear Theory of Electroelastic and Magnetoelastic Interactions. Springer Verlag, New York.
Ericksen, J.L., 2002. Electromagnetic effects in thermoelastic materials. Math. Mech. Solids 7, 165–189.
Ericksen, J.L., 2007a. On formulating and assessing continuum theories of electromagnetic fields in elastic materials. J. Elasticity 87, 95–108.
Ericksen, J.L., 2007b. Theory of elastic dielectric revisited. Arch. Rational Mech. Anal. 183, 299–313.
Fosdick, R., Tang, H., 2007. Electrodynamics and thermomechanics of material bodies. J. Elasticity 88, 255–297.
Giaquinta, M., Modica, G., Souček, J., 1998. Cartesian Currents in the Calculus of Variations. Vol. I and II. Springer-Verlag, Berlin.
Gurtin, M.E., 1997. Dynamical theories of electromagnetism and superconductivity based on gauge invariance and energy. Arch. Ration. Mech. Anal. 137, 49–97.
Harden, J., Chambers, M., Verduzco, R., Luchette, P., Gleeson, J.T., Sprunt, S., Jákli, A., 2010. Giant flexoelectricity in bent-core nematic liquid crystal elastomers. Appl. Phys. Lett. 96, art. n. 102907.
Hong, J., Vanderbilt, D., 2013. First-principles theory and calculation of flexoelectricity. Phys. Rev. B. 88, art. n. 174107.
Kijowski, J., Magli, G., 1992. Relativistic elastomechanics as a lagrangian field theory. J. Geom. Phys. 9, 207–223.

Krichen, S., Sharma, P., 2016. Flexoelectricity: a perspective on an unusual electromechanical coupling. J. Appl. Mech. 83, art. n. 030801.
Kruzik, M., Mariano, P.M., Mucci, D., 2022. Crack occurrence in bodies with gradient polyconvex energies. J. Nonlinear Sci. 32, art. n. 16.
Liu, L., 2014. An energy formulation of continuum magneto-electro-elasticity with applications. J. Mech. Phys. Solids 63, 451–480.
Maitra, M., Tromp, J., 2024. Cosserat elasticity as the weak-field limit of einstein-cartan relativity. Phys. Rev. D 109, 104052.
Man, C.-S., 2025. Fusion of nonlinear elasticity with galilean electromagnetism. J. Elasticity 157, art. n. 32.
Mariano, P.M., 2017. Second-neighbor interactions in classical field theories: invariance of the relative power and covariance. Math. Meth. Appl. Sci. 40, 1316–1332.
Mariano, P.M., 2024. Topological features in continuum mechanics: from simple to complex bodies. Proc. Royal Soc. A 480, 1–30. art. n. 20230919
Mariano, P.M., Markenscoff, X., Zeka, D., 2025. In continua with 'active' microstructure stability competes with multi-field destabilizing effects. Proc. Royal Soc. A 481, 20250623.
Mariano, P.M., Modica, G., 2009. Ground states in complex bodies. ESAIM - Contr. Optim. Calc. Var. 15, 377–402.
Mariano, P.M., Mucci, D., 2025. Rod deformations with possibly irrecoverable corners: conditions for global injectivity. J. Diff. Eq. 438, art. n. 113381.
Maugin, G.A., 1977. Infinitesimal discontinuities in initially stressed relativistic elastic solids. Comm. Math. Phys. 53, 233–256.
Maugin, G., 1978. On the covariant equations of the relativistic electrodynamics of continua. IV. media with spin. J. Math. Phys. 19, 1220–1226.
Maugin, G.A., 1988. Continuum Mechanics of Electromagnetic Solids. North Holland, Amsterdam.
Olver, P.J., 2024. Boundary conditions and null lagrangians in the calculus of variations and elasticity. J. Elast. 155, 75–108.
Rahmati, A.H., Mozaffari, K., Liu, L., Sharma, P., 2025. A static and dynamic theory for photo-flexoelectric liquid crystal elastomers and the coupling of light, deformation and electricity. J. Mech. Phys. Solids 195, art. n. 105949.
Sharma, B.L., Basak, N., 2021. Null lagrangians in cosserat elasticity. J. Elast. 143, 337–358.
Steigmann, D.J., 2007. On the formulation of balance laws for electromagnetic continua. Math. Mech. Solids 14, 390–402.
Stengel, M., 2013a. Flexoelectricity from density-functional perturbation theory. Phys. Rev. B. 88, 174106.
Stengel, M., 2013b. Microscopic response to inhomogeneous deformations in curvilinear coordinates. Nat. Commun. 4, art. n. 2693.
Stengel, M., 2014. Surface control of flexoelectricity. Phys. Rev. B. 90, art. n. 201112(R).
Toupin, R.A., 1956. The elastic dielectric. J. Rational Mech. Anal. 5, 849–915.
Toupin, R.A., 1960. Stress tensors in elastic dielectrics. Arch. Ration. Mech. Anal. 5, 440–552.
Truesdell, C.A., 1977. A first course in rational continuum mechanics. Academic Press New York.
Truesdell, C.A., Noll, W., 1965. The non-linear field theories of mechanics. Vol. III/3. Springer, Berlin.
Truesdell, C.A., Toupin, R.A., 1960. Classical field theories of mechanics. In: Flügge S., (Eds.), Handbuch Der Physics. Springer, Berlin. Vol. III/1, pp. 226–793.
Wang, C.C., 1979. Mathematical principles of mechanics and electromagnetism. Plenum Press New York.
Wang, B., Gua, Y., Zhange, S., Chena, L.-Q., 2019. Flexoelectricity in solids: progress, challenges, and perspectives. Prog. Material Sciences 106, art. n. 100570.
Xiao, Y., Bhattacharya, K., 2008. A continuum theory of deformable, semiconducting ferroelectrics. Arch. Ration. Mech. Anal. 189, 59–95.