Preprint, Institute of Mathematics, AS CR, Prague. 2008-10-22

Equilibrium of phases with interfacial energy: a variational approach

M. Šilhavý
Institute of Mathematics of the AV ČR
Žitná 25
115 67 Prague 1
Czech Republic
e-mail: silhavy@math.cas.cz

Abstract The paper proves the existence of equilibrium two phase states with elastic solid bulk phases and deformation dependent interfacial energy. The states are pairs $(\mathbf{y}, E)$ consisting of the deformation $\mathbf{y}$ on the body and the region E occupied by one of the phases in the reference configuration. The bulk energies of the two phases are polyconvex functions representing two wells of the substance. The interfacial energy is interface polyconvex. The last notion is introduced and discussed below, together with the interface quasiconvexity and interface null lagarangians. The constitutive theory and equilibrium theory of the interface are discussed in detail under appropriate smoothness hypotheses. Various forms of the interfacial stress relations for the standard and configurational (Eshelby) interfacial stresses are established. The equilibrium equations are derived by a variational argument emphasizing the roles of outer and inner variations.

Keywords Phase transitions, phase interface, interface polyconvexity, interface null lagrangians, interfacial stress, interfacial Eshelby tensor

MSC 2000 classification 74A50, 49Q20

Contents

1 Introduction

Consider a deformed body in a state with two coexistent phases separated by a phase interface. If a bounded open set $\Omega \subset \mathbb{R}^3$ is the reference region, then the state is described by a deformation function $\mathbf{y} : \Omega \to \mathbb{R}^3$ and by an open subset E of Ω occupied by the first phase; the region occupied by the second phase is the complement of E in Ω. The phase interface $\mathcal{S}$ is the part of the boundary of E that

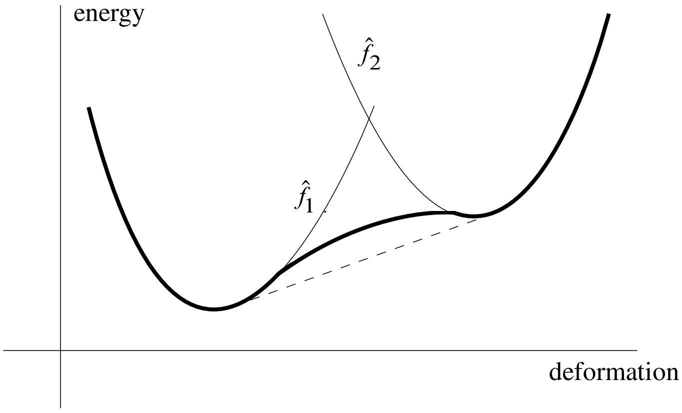

Fig. 1.1.

is contained in Ω (the rest of the boundary of E being a subset of the boundary of Ω, possibly empty). The deformation function $\mathbf{y}$ gives the actual position $\mathbf{y}(\mathbf{x})$ of the material point $\mathbf{x} \in \Omega$, in particular for $\mathbf{x} \in \mathcal{S}$ the value $\mathbb{y}(\mathbf{x}) = \mathbf{y}(\mathbf{x})$ gives the actual position of the interface points. We assume that the energy of each of the two bulk phases $\alpha = 1, 2$ depends on the deformation gradient

$$\mathbf{F} = \nabla \mathbf{y}$$

via the response functions $\hat f_\alpha$, $\alpha = 1, 2$, as

$$f_\alpha(\mathbf{x}) = \hat f_\alpha(\mathbf{F}(\mathbf{x})), \quad \mathbf{x} \in E_\alpha, \quad E_1 := E, \quad E_2 := \Omega \sim \operatorname{cl} E$$

where throughout, cl and bd denote the closure and boundary. The energy of the interface depends on the normal $\mathbb{m}$ to the interface and on the surface deformation gradient [19, 16]

$$\mathbb{F} = \nabla \mathbb{y}, \quad \mathbb{F}\mathbb{m} = \mathbf{0},$$

where ∇ denotes the surface gradient (see Section 7), via the response function $\hat{\mathbb{f}}$,

$$\mathbb{f}(\mathbf{x}) = \hat{\mathbb{f}}(\mathbb{F}(\mathbf{x}), \mathbb{m}(\mathbf{x})), \quad \mathbf{x} \in \mathcal{S}.$$

The total energy of the state $(\mathbf{y}, E)$ is

$$\mathsf{E}(\mathbf{y}, E) = \mathsf{E}_{\mathrm{b}}(\mathbf{y}, E) + \mathsf{E}_{\mathrm{if}}(\mathbf{y}, E)$$

where

$$\mathsf{E}_{\mathrm{b}}(\mathbf{y}, E) = \int_E \hat f_1(\mathbf{F})\, d\mathcal{L}^3 + \int_{\Omega \sim E} \hat f_2(\mathbf{F})\, d\mathcal{L}^3,$$

$$\mathsf{E}_{\mathrm{if}}(\mathbf{y}, E) = \int_{\mathcal{S}} \hat{\mathbb{f}}(\mathbb{F}, \mathbb{m})\, d\mathcal{H}^2 \tag{1.1}$$

are the bulk and interface energies, with $d\mathcal{L}^3$ the referential volume element and $d\mathcal{H}^2$ the referential area element. Equilibrium states of the system correspond to minimum energy under the constraints imposed by the environment of the system. Here the region E is unknown, its determination is a part of the solution of the problem.

The above format pertains to fluid/fluid interface, solid/melt interface, solid/solid interface, or grain boundaries. The two bulk energies represent two energy wells of the substance, see Figure 1.1.

In case of fluid phases, the problem of energy minimizing states was solved by Gurtin [14–15]. In this case the bulk energies depend on $\mathbf{F}$ via the specific volume $v = \det \mathbf{F}$,

$$\hat{f}_a(\mathbf{F}) = \tilde{f}_a(\det \mathbf{F}),$$

and it is realistic to assume (and also follows from symmetry considerations) that the interfacial energy reduces to surface tension, i.e., is proportional to the area of the deformed surface,

$$\hat{\mathbb{f}}(\mathbb{F}, \mathbb{m}) = \sigma |\operatorname{cof} \mathbb{F} \mathbb{m}|,$$

where $\sigma > 0$ is the surface tension coefficient and $\operatorname{cof} \mathbb{F}$ is the cofactor tensor of $\mathbb{F}$, see Section 9. Under the external conditions of “canister” type, viz., for states occupying in the actual configuration a prescribed region $\Sigma \subset \mathrm{R}^3$, the minimum energy states are either homogeneous, i.e., single phase states of constant specific volume, or pairwise homogeneous, i.e., states of two coexistent phases, each of constant specific volume, separated by an interface minimizing the spatial area of the interface under fixed volumes of the phases. The minimum area is a function of one of the two volumes and the problem reduces to the minimization of a function of one variable.

The present paper addresses the question of the existence of minimum energy states $(\mathbf{y}, E)$ in case of solid bulk phases under the prescribed boundary displacement. Then $\hat{f}_a$ are the energy wells of the two solid phases (e.g., corresponding to the austenite and martensite phases, or two martensite phases), and $\hat{\mathbb{f}}$ is a function of $\mathbb{F}$ and $\mathbb{m}$ that represents the symmetry of the bulk phases. If $\mathbf{z}_0 : \operatorname{bd} \Omega \to \mathrm{R}^3$ is the prescribed boundary displacement we consider states from the set

$$\mathcal{A}(\mathbf{z}_0) = \{ (\mathbf{z}, F) \in \mathcal{G} : \mathbf{z} = \mathbf{z}_0 \ \text{ on } \ \operatorname{bd} \Omega \}$$

where $\mathcal{G}$ denotes the collection of all states $(\mathbf{y}, E)$. The problem is to find a state $(\mathbf{y}, E) \in \mathcal{A}(\mathbf{z}_0)$ such that

$$\mathsf{E}(\mathbf{y}, E) \le \mathsf{E}(\mathbf{z}, E) \quad \text{for all} \quad (\mathbf{z}, F) \in \mathcal{A}(\mathbf{z}_0). \tag{1.2}$$

It is well known that in the absence of interfacial energy ($\hat{\mathbb{f}} \equiv 0$) the problem (1.2) generally does not have a solution, since in the process of approaching the least energy the body exhibits states $(\mathbf{y}^i, E^i)$ of complicated patterns of coexistent phases with finer and finer microstructure and with the area of the interface approaching infinity. As the theory does not have any length scale, there is no limit on the fineness of the microstructure, i.e., it is infinitely fine in the limit. The Young measure minimizers represent the idealized limiting states. The least energy is given by the quasiconvex envelope $\hat{f}^{\mathrm{qc}}$ (see [6; Section 6.3] for the definition) of the minimum energy

$$\hat{f}(\mathbf{F}) = \min \{ \hat{f}_1(\mathbf{F}), \hat{f}_2(\mathbf{F}) \}.$$

In particular, under the affine boundary conditions

$$\mathbf{y}(\mathbf{x}) = \mathbf{A}\mathbf{x}, \quad \mathbf{x} \in \operatorname{bd} \Omega,$$

where $\mathbf{A}$ is a prescribed constant affine deformation gradient, one has

$$\inf \{ \mathsf{E}(\mathbf{y}, E) \in \mathcal{G} : \mathbf{y}(\mathbf{x}) = \mathbf{A}\mathbf{x} \ \text{ on } \ \operatorname{bd} \Omega \} = \hat{f}^{\mathrm{qc}}(\mathbf{A})$$

where we assume the referential volume of Ω equal to 1 for simplicity. The quasiconvex envelope is schematically represented in Figure 1.1 by the dashed line, although it must be understood that $\hat{f}^{\mathrm{qc}}$ is not the convex envelope, as the figure might suggest.

To obtain the existence of the solutions to the problem (1.2), we assume that the bulk phases are described by two polyconvex wells. The mutual relation of the wells is arbitrary, so that the geometric incompatibility induced by symmetry can occur. Thus we assume that

$$\hat{f}_\alpha(\mathbf{F}) = \Phi_\alpha(\mathbf{F}, \operatorname{cof} \mathbf{F}, \det \mathbf{F}) \tag{1.3}$$

for all $i = 1, 2$ and all values of deformation gradient $\mathbf{F}$, where $\Phi_\alpha : \mathrm{W} \to \mathrm{R}$ are convex functions on $\mathrm{W} = \mathrm{Lin} \times \mathrm{Lin} \times \mathrm{R}$ with Lin the set of all second order tensors. The existence theory for a single phase minimizer with polyconvex energy is well understood [3, 12, 26, 13] and the corresponding part of the present theory merely transfers those results.

The paper proposes and discusses convexity notions of the interfacial energy $\hat{\mathbb{f}}$, viz., the interface quasiconvexity, interface null lagrangians, and interface polyconvexity. The interface quasiconvexity ensures the stability of a planar homogeneously deformed interface $\mathcal{T}$ against curved inhomogeneously deformed interfaces $\mathcal{S}$ with the same boundary data. Mathematically it is equivalent to the sequential lowersemicontinuity of the surface energy term (1.1) under an appropriate convergence, as will be shown elsewhere [31]. An interface null lagrangian is an interfacial energy $\hat{\mathbb{f}}$ such that $\hat{\mathbb{f}}$ and $-\hat{\mathbb{f}}$ are interface quasiconvex. An explicit form is given below [(1.5)]. An interface polyconvex surface energy is a convex, positively 1 homogeneous function of interface null lagrangians; it is automatically interface quasiconvex, and our existence result is based on the interface polyconvexity. We note that Fonseca [10] establishes two particular cases of the present notion of interface quasiconvexity as necessary conditions for metastable minima (see below in this section for a discussion of the relationship).

In detail, $\hat{\mathbb{f}}$ is said to be interface quasiconvex if

$$\int_{\mathcal{S}} \hat{\mathbb{f}}(\nabla \mathbb{y}, \mathbb{m})\, d\mathcal{H}^2 \geq \mathcal{H}^2(\mathcal{T}) \hat{\mathbb{f}}(\mathbb{G}, \mathbb{m}) \tag{1.4}$$

for every surface deformation gradient $\mathbb{G}$, every unit vector $\mathbb{m}$ with $\mathbb{G}\mathbb{m} = \mathbf{0}$, every planar 2 dimensional region $\mathcal{T}$ of normal $\mathbb{m}$, every (curved) surface $\mathcal{S}$ of normal $\mathbb{m}$ and every smooth map $\mathbb{y} : \mathcal{S} \to \mathrm{R}^3$ such that

$$\operatorname{bd} \mathcal{S} = \operatorname{bd} \mathcal{T}, \quad \mathbb{y}(\mathbf{x}) = \mathbb{G}\mathbf{x}, \quad \mathbf{x} \in \operatorname{bd} \mathcal{T};$$

see Figure 1.2. Here $\operatorname{bd} \mathcal{S}$ and $\operatorname{bd} \mathcal{T}$ denote the (relative) boundaries of the 2 dimensional surfaces $\mathcal{S}$ and $\mathcal{T}$ in R^3. We emphasize that the surface $\mathcal{S}$ is not the deformed interface $\mathcal{T}$ but instead a *different interface* consisting of material points different from those of $\mathcal{T}$. Thus testing (1.4) involves implicitly a change of the interface, mathematically reflected by the variation of the integration domain, from $\mathcal{T}$ to $\mathcal{S}$, with $\mathcal{H}^2(\mathcal{S}) \geq \mathcal{H}^2(\mathcal{T})$. The last fact, which has no counterpart in the standard bulk quasiconvexity notion, has strong consequences which we shall mention below. Here we note that while the constant bulk energies are trivially quasiconvex, a constant interfacial energy $\hat{\mathbb{f}}$ is interface quasiconvex if and only if the constant value of $\hat{\mathbb{f}}$ is nonnegative. The interface quasiconvexity rules out surface wrinkling and prefers homogeneous surface deformations over the inhomogeneous ones.

Working in a different format of the interfacial energy than (1.1), Fonseca [10] established two related but weaker quasiconvexity properties of the interfacial energy.

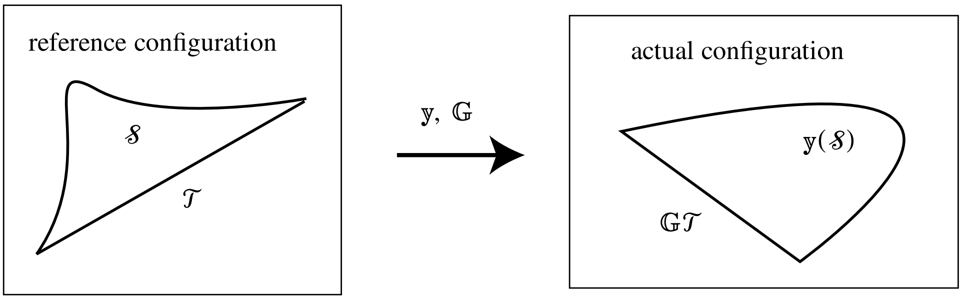

Fig. 1.2.

For energies of the format (1.1) they correspond to particular cases of the above definition. Namely, rephrasing slightly [10], the first condition [10; Proposition 4.3(i)] amounts to testing (1.4) in the particular case $\mathcal{S} = \mathcal{T}$ with $\mathbb{y}$ as above while the second condition [10; Proposition 4.9(i)] to $\mathcal{S}$ and $\mathcal{T}$ as above but $\mathbb{y}(\mathbf{x}) = \mathbf{F}\mathbf{x}$ for all $\mathbf{x} \in \mathcal{S}$ where $\mathbf{F} \in \mathrm{Lin}$ is constant and such that $\mathbf{F}\mathbf{x} = \mathbb{G}\mathbf{x}$ for all $\mathbf{x} \in \mathcal{T}$. Individually, each of these conditions is weaker (less restrictive) than the interface quasiconvexity; whether the combination of these two is equivalent to the interface quasiconvexity is an open question.

A surface energy is said to be an interface null lagrangian if it satisfies (1.4) with the equality sign for all objects listed above. It will be proved [31] that $\hat{\mathbb{f}}$ is an interface null lagrangian if and only if

$$\hat{\mathbb{f}}(\mathbb{F}, \mathbb{m}) = \mathbf{c} \cdot \mathbb{m} + \mathbf{\Omega} \cdot (\mathbb{F} \times \mathbb{m}) + \mathbf{a} \cdot \operatorname{cof} \mathbb{F} \mathbb{m} \tag{1.5}$$

for each $\mathbb{F}$ and $\mathbb{m}$ with $\mathbb{F}\mathbb{m} = \mathbf{0}$, where $\mathbf{c}$ and $\mathbf{a}$ are constant vectors and $\mathbf{\Omega}$ a constant second order tensor. Here $\mathbb{F} \times \mathbb{m}$ is a second order tensor defined by

$$(\mathbb{F} \times \mathbb{m})\mathbf{t} = \mathbb{F}(\mathbb{m} \times \mathbf{t})$$

for any vector $\mathbf{t}$; in components,

$$(\mathbb{F} \times \mathbb{m})_{iA} = \varepsilon_{ABC} \mathbb{F}_{iB} \mathbb{m}_C$$

where ε_{ABC} is the permutation symbol, summation convention applies, and $\mathbb{F}_{iB}$, $\mathbb{m}_C$ are the components of $\mathbb{F}$ and $\mathbb{m}$ with $i = 1, 2, 3$ the spatial indices and $A, B, C = 1, 2, 3$ the referential indices. Since $\mathbb{F}\mathbb{m} = \mathbf{0}$ we have

$$\mathbb{F} = -(\mathbb{F} \times \mathbb{m}) \times \mathbb{m};$$

thus $\mathbb{F} \times \mathbb{m}$ carries the same information as $\mathbb{F}$; however, it is $\mathbb{F} \times \mathbb{m}$, and not $\mathbb{F}$, that enters the interface null lagrangians. The list

$$\mathbb{m}, \quad \mathbb{F} \times \mathbb{m}, \quad \operatorname{cof} \mathbb{F} \mathbb{m} \tag{1.6}$$

of basic interface null lagrangians consists of the unit interface normal vector $\mathbb{m}$, which need not be commented, of $\mathbb{F} \times \mathbb{m}$, which, like $\mathbb{F}$, describes the elongation of lines in the interface, and of the vector $\operatorname{cof} \mathbb{F} \mathbb{m}$ whose direction is the normal to the interface in the deformed configuration, and whose magnitude $|\operatorname{cof} \mathbb{F} \mathbb{m}| \equiv |\operatorname{cof} \mathbb{F}|$ is the ratio of the areas of the interface in the deformed and reference configurations. In dimension 3 there are 15 independent (scalar) interface null lagrangians; we recall that there are 20 standard null lagrangians (counting the constant null lagrangians in them). We note that $\hat{\mathbb{f}}$ = constant $\neq 0$ is *not* an interface null lagrangian, this being another manifestation of the variation of the integration domain in the interface quasiconvexity.

The surface energy $\hat{\mathbb{f}}$ is said to be interface polyconvex if $\hat{\mathbb{f}}$ is the supremum of some family of interface null lagrangians. The interface polyconvexity is a sufficient condition for the interface quasiconvexity, as in case of the bulk counterparts of these notions. The energy $\hat{\mathbb{f}}$ is interface polyconvex if and only if it is of the form

$$\hat{\mathbb{f}}(\mathbb{F}, \mathbb{m}) = \Phi(\mathbb{m}, \mathbb{F} \times \mathbb{m}, \operatorname{cof} \mathbb{F} \mathbb{m}) \tag{1.7}$$

where $\Phi : \mathrm{X} \to \mathrm{R}$ is a positively 1 homogenous convex function on $\mathrm{X} = \mathrm{R}^3 \times \mathrm{Lin} \times \mathrm{R}^3$, where the positive 1 homogeneity of Φ means

$$\Phi(t\mathrm{A}) = t\,\Phi(\mathrm{A})$$

for each $t \geq 0$ and each argument $\mathrm{A} \in \mathrm{X}$. The positive 1 homogeneity requirement does not occur in the definition of the standard (bulk) polyconvexity (1.3); the convexity of Φ_a suffices. This is again related to the increase of the area of the competitor interface $\mathcal{S}$ as the use of Jensen's inequality similar to that in case of bulk polyconvexity involves an integration with respect to two measures of which one has a bigger total mass.

For the existence result, we define the state space as follows.

Definition 1.1. Let $\Omega \subset \mathrm{R}^3$ be a bounded open set with Lipschitz boundary, and let $2 \leq p < \infty$, $3/2 \leq q < \infty$. We denote by $\mathcal{G}^{p,q}(\Omega, \mathrm{R}^3)$ the set of all pairs $(\mathbf{y}, E)$ where

(i) $\mathbf{y} \in W^{1,p}(\Omega, \mathrm{R}^3)$, $\operatorname{cof} \nabla \mathbf{y} \in L^q(\Omega, \mathrm{Lin})$,

(ii) E is a subset of Ω of finite perimeter,

(iii) there exist measures $\mathbb{D}$ and $\mathbb{p}$ on Ω with values in Lin, and R^3, respectively, such that

$$-\int_E \nabla \mathbf{y} \operatorname{curl} \mathbf{v}\, d\mathcal{L}^3 = \int_\Omega d\mathbb{D}\mathbf{v}, \quad \int_E \operatorname{cof} \nabla \mathbf{y} \cdot \nabla \mathbf{v}\, d\mathcal{L}^3 = \int_\Omega \mathbf{v} \cdot d\mathbb{p}$$

for every infinitely differentiable testvectorfield $\mathbf{v}$ with support in Ω.

We call the elements $(\mathbf{y}, E)$ of $\mathcal{G}^{p,q}(\Omega, \mathrm{R}^3)$ states. If $\mathbf{y}$ is smooth then the integration by parts and the identities

$$\operatorname{curl} \nabla \mathbf{y} = \mathbf{0}, \quad \operatorname{div}(\operatorname{cof} \nabla \mathbf{y}) = \mathbf{0}$$

show that the measures $\mathbb{D}$ and $\mathbb{p}$ automatically exist and are given by

$$\mathbb{D} = \mathbb{F} \times \mathbb{m}\, \mathcal{H}^2 \llcorner \mathcal{S}, \quad \mathbb{p} = \operatorname{cof} \mathbb{F} \mathbb{m}\, \mathcal{H}^2 \llcorner \mathcal{S} \tag{1.8}$$

where $\mathcal{H}^2 \llcorner \mathcal{S}$ is the area measure restricted to the interface $\mathcal{S} := \Omega \cap \mathrm{bd}_* E$ with $\mathrm{bd}_* E$ the measure theoretic boundary of E. Hence the measures

$$\mathbb{b} = \mathbb{m}\, \mathcal{H}^2 \llcorner \mathcal{S}, \quad \mathbb{D}, \quad \mathbb{p} \tag{1.9}$$

provide measure theoretic generalizations of the basic interface null lagrangians (1.6). We then define the interfacial energy as the Φ function of the triplet (1.9) of measures, i.e., we put

$$\mathsf{E}_{\mathrm{if}}(\mathbf{y},E) = \int_\Omega \Phi(\mathbb{A})\, d|\mathbb{J}| \tag{1.10}$$

where we interpret $\mathbb{J} := (\mathbb{b}, \mathbb{D}, \mathbb{p})$ as a measure with values in X, the symbol $|\mathbb{J}|$ denotes the total variation measure of $\mathbb{J}$ and $\mathbb{A} : \Omega \to \mathrm{X}$ is a vectorfield such that we have the polar decomposition identity $\mathbb{J} = \mathbb{A}\,|\mathbb{J}|$; cf. [1; Corollary 1.29 and Section 2.6] for the discussion of these notions in a general context. The definition (1.10) reduces to (1.1) in case of a state $(\mathbf{y}, E)$ with $\mathbf{y}$ smooth by (1.8) and (1.7). The requirement (i) in the above definition comes from the refinement of Ball's existence theory given in [26].

Theorem 1.2. *Let* $2 \le p < \infty$, $3/2 \le q < \infty$ *and assume that*

(i) $\hat f_a$, $a = 1, 2$, *are polyconvex in the sense of* (1.3) *where* $\Phi_a : W \to [0, \infty]$ *are continuous convex functions,* $\hat{\mathbb{f}}$ *is interface polyconvex in the sense of* (1.7) *where* $\Phi : Z \to [0, \infty)$ *is a positively* 1 *homogeneous convex function,*

(ii) *for all* $a = 1, 2$, *all* $\mathbf{F} \in \mathrm{Lin}$, *all* $\mathrm{A} \in \mathrm{X}$, *some* $c > 0$ *and some* $d \in \mathrm{R}$ *we have*

$$\hat f_a(\mathbf{F}) \ge c(|\mathbf{F}|^p + |\operatorname{cof} \mathbf{F}|^q) + d, \quad \Phi(\mathrm{A}) \ge c|\mathrm{A}|, \tag{1.11}$$

(iii) $\hat f_a(\mathbf{F}) = \infty$ *if* $\det \mathbf{F} \le 0$.

If $\mathbf{z}_0 \in W^{1,p}(\Omega, \mathrm{R}^m)$ *and* E *is finite for some element of the set*

$$\mathcal{A}(\mathbf{z}_0) = \{(\mathbf{z}, F) \in \mathcal{G}^{p,q}(\Omega, \mathrm{R}^3) : \mathbf{z} = \mathbf{z}_0 \text{ on } \operatorname{bd}\Omega\}$$

then the problem (1.2) *has a solution, i.e., there exists an* $(\mathbf{y}, E) \in \mathcal{A}(\mathbf{z}_0)$ *such that*

$$\mathsf{E}(\mathbf{y}, E) \le \mathsf{E}(\mathbf{z}, F)$$

for all $(\mathbf{z}, F) \in \mathcal{A}(\mathbf{z}_0)$; *each solution* $(\mathbf{y}, E)$ *of the problem satisfies*

$$\det \nabla \mathbf{y} > 0 \text{ for } \mathcal{L}^3 \text{ a.e. point of } \Omega. \tag{1.12}$$

We allow $\hat f_a$ to take the value ∞ not only to incorporate Condition (iii), which leads to the orientation preserving property (1.12), but also to allow the effective domains

$$\operatorname{eff\,dom} \hat f_a = \{\mathbf{F} \in \mathrm{Lin} : \hat f_a(\mathbf{F}) < \infty\}$$

be proper subsets of the set Lin_+ of deformation gradients of positive determinant. Thus one may assume that the effective domains are disjoint, and/or exclude states with deformation gradient in the spinodal region.

The solvability of (1.2) rules out the minimizers with infinitely fine microstructure occurring in the absence of interfacial energy. Indeed, the surface energy $\hat{\mathbb{f}}$ introduces a length scale

$$l \simeq \mathbb{f}/f$$

where $\mathbb{f}$ and f are the typical values of $\hat{\mathbb{f}}$ and $\hat f_a$ of the problem. The infimum energy

$$\mathsf{E}_{\mathrm{eff}}(\mathbf{A}) = \inf\{\mathsf{E}(\mathbf{y}, E) \in \mathcal{G} : \mathbf{y}(\mathbf{x}) = \mathbf{A}\mathbf{x} \text{ on } \operatorname{bd}\Omega\}$$

is a minimum under the conditions of the existence theorem. The function $\mathsf{E}_{\mathrm{eff}}$ is schematically shown in Figure 1.1 by the bold line; thus the quasiconvex envelope of $\hat f$ is now only a lower bound for $\mathsf{E}_{\mathrm{eff}}$. Moreover, the function $\mathsf{E}_{\mathrm{eff}}$ is different for different shapes of Ω, presumably approaching $\hat f^{\mathrm{qc}}$ for $\hat{\mathbb{f}}$ approaching 0 or for the size of the specimen Ω approaching ∞.

The coercivity condition $(1.11)_2$ excludes surface energies $\hat{\mathbb{f}}$ that depend only on the interface normal $\mathbb{m}$; Theorem 6.6 (below) gives an existence result in this situation.

2 Constitutive theory. States

We denote by $\mathrm{Lin}(V, W)$ the set of all linear transformations from a vectorspace V into a vectorspace W. Throughout the paper, let m, n be positive integers and write $\mathrm{Lin} := \mathrm{Lin}(\mathrm{R}^n, \mathrm{R}^m)$ unless stipulated otherwise. Denote by S^{n-1} the unit sphere in R^n. We model the reference configuration of the body by an open bounded subset Ω of R^n and consider deformations $\mathbf{y} : \Omega \to \mathrm{R}^m$. In applications, $m = n = 3$; however, the considerations presented below for general m and n do not simplify if $m = n = 3$ although they occasionally simplify if $m = n$.

In the treatment of the constitutive theory for the interface it is necessary to take into account that the surface energy $\hat{\mathbb{f}} = \hat{\mathbb{f}}(\mathbb{F}, \mathbb{m})$ is defined on pairs satisfying $\mathbb{F}\mathbb{m} = \mathbf{0}$, this set G forms a *submanifold* of the space $\mathrm{Lin} \times \mathrm{R}^n$. Thus the derivatives of $\hat{\mathbb{f}}$ belong to the tangent space of G and hence the "partial derivatives" with respect to $\mathbb{F}$ and $\mathbb{m}$ are not independent. The derivative of a map on a manifold is defined in Section 7; Remark 2.2 (below) discusses other possible consistent choices of the constitutive theory for the interface.

Definitions 2.1 (Constitutive information and response functions).

(i) The two bulk phases are indexed by $\alpha = 1, 2$, each phase is described by the ***bulk energy*** $\hat{f}_\alpha : U_\alpha \to \mathrm{R}$ where $U_\alpha \subset \mathrm{Lin}$ is an open set and $\hat{f}_\alpha$ are class 2 functions. We define the response functions for the standard and configurational stresses $\hat{\mathbf{S}}_\alpha : U_\alpha \to \mathrm{Lin}(\mathrm{R}^n, \mathrm{R}^m)$, $\hat{\mathbf{C}}_\alpha : U_\alpha \to \mathrm{Lin}(\mathrm{R}^n, \mathrm{R}^n)$ by

$$\hat{\mathbf{S}}_\alpha = \mathrm{D}\hat{f}_\alpha, \quad \hat{\mathbf{C}}_\alpha = \hat{f}_\alpha \mathbf{1} - \mathbf{F}^{\mathrm{T}}\,\mathrm{D}\hat{f}_\alpha \tag{2.1}$$

for each $\mathbf{F} \in U_\alpha$, where $\hat{\mathbf{S}}_\alpha$, $\hat{\mathbf{C}}_\alpha$, $\hat{f}_\alpha$ and its derivatives are evaluated at $\mathbf{F}$.

(ii) The interface is desctibed by the ***interfacial energy*** $\hat{\mathbb{f}} : \mathbb{U} \to \mathrm{R}$ where $\mathbb{U}$ is a (relatively) open subset of the class ∞ manifold

$$\mathrm{G} = \{ (\mathbb{F}, \mathbb{m}) \in \mathrm{Lin} \times \mathrm{S}^{n-1} : \mathbb{F}\mathbb{m} = \mathbf{0} \}$$

(see Proposition 8.1, below) and $\hat{\mathbb{f}}$ is a class 2 function. The derivative of $\hat{\mathbb{f}}$ at $(\mathbb{F}, \mathbb{m}) \in \mathrm{G}$ is an element of the tangent space $\mathrm{Tan}(\mathrm{G}, (\mathbb{F}, \mathbb{m}))$ of G at $(\mathbb{F}, \mathbb{m})$ given by (8.1) (below); we write $\mathrm{D}\hat{\mathbb{f}} = (\mathrm{D}_1\hat{\mathbb{f}}, \mathrm{D}_2\hat{\mathbb{f}})$ for its components in Lin and R^n, respectively. We define the ***response functions*** for the ***standard*** and ***configurational stresses*** $\hat{\mathbb{S}} : \mathrm{G} \to \mathrm{Lin}(\mathrm{R}^n, \mathrm{R}^m)$ and $\hat{\mathbb{C}} : \mathrm{G} \to \mathrm{Lin}(\mathrm{R}^n, \mathrm{R}^n)$ by

$$\hat{\mathbb{S}} = \mathrm{D}_1\hat{\mathbb{f}}\,\mathbb{P}, \tag{2.2}$$

$$\hat{\mathbb{C}} = \hat{\mathbb{f}}\,\mathbb{P} - \mathbb{F}^{\mathrm{T}}\,\mathrm{D}_1\hat{\mathbb{f}}\,\mathbb{P} + \mathbb{m} \otimes (\mathbb{F}^{\mathrm{T}}\,\mathrm{D}_1\hat{\mathbb{f}}\,\mathbb{m} - \mathrm{D}_2\hat{\mathbb{f}}) \tag{2.3}$$

for every $(\mathbb{F}, \mathbb{m}) \in \mathrm{G}$ where $\mathbb{P} = \mathbf{1} - \mathbb{m} \otimes \mathbb{m}$ and $\hat{\mathbb{S}}$, $\hat{\mathbb{C}}$, $\hat{\mathbb{f}}$ and its derivatives are evaluated at $(\mathbb{F}, \mathbb{m})$.

The form of the stress relations (2.2) and (2.3) is motivated by the variations formulas for the total energy, (3.1) and (3.2), by the correponding balance equations (3.4) and (3.5), and by the fact that with the above definitions $\hat{\mathbb{S}}$ and $\hat{\mathbb{C}}$ neatly exchange their roles under the exchange of the actual and reference configurations, Section 4 (below). Some authors call the configurational stress only the part

$$\hat{\mathbb{f}}\,\mathbb{P} - \mathbb{F}^{\mathrm{T}}\,\mathrm{D}_1\,\hat{\mathbb{f}}\,\mathbb{P}$$

and the remaining part

$$\mathbb{m} \otimes (\mathbb{F}^{\mathrm{T}}\,\mathrm{D}_1\,\hat{\mathbb{f}}\,\mathbb{m} - \mathrm{D}_2\,\hat{\mathbb{f}})$$

the configurational shear. The 'partial derivatives' $\mathrm{D}_1\,\hat{\mathbb{f}}$ and $\mathrm{D}_2\,\hat{\mathbb{f}}$ satisfy

$$\mathrm{D}_1\,\hat{\mathbb{f}}(\mathbb{F},\mathbb{m})\mathbb{m} + \mathbb{F}\,\mathrm{D}_2\,\hat{\mathbb{f}}(\mathbb{F},\mathbb{m}) = \mathbf{0}, \quad \mathbb{m}\cdot\mathrm{D}_2\,\hat{\mathbb{f}}(\mathbb{F},\mathbb{m}) = 0 \tag{2.4}$$

by (8.1) (below).

Remarks 2.2. We discuss other equivalent and consistent forms of the stress relations with different partial derivatives employed in the literature (often tacitly); the reader is referred to Section 8 for proofs.

(i) Assume that the domain $\mathbb{U}$ of $\hat{\mathbb{f}}$ is G. The response function $\hat{\mathbb{f}}$ has a 'canonical' extension to a function $\tilde{\mathbb{f}} : \mathrm{H} \to \mathrm{R}$ where $\mathrm{H} := \mathrm{Lin}\times\mathrm{S}^{n-1} \to \mathrm{R}$ given by

$$\tilde{\mathbb{f}}(\mathbf{F},\mathbb{m}) = \hat{\mathbb{f}}(\mathbf{F}\mathbb{P},\mathbb{m}), \quad \mathbb{P} := \mathbf{1} - \mathbb{m}\otimes\mathbb{m} \tag{2.5}$$

for any $(\mathbf{F},\mathbb{m}) \in \mathrm{H}$. The variables $\mathbf{F} \in \mathrm{Lin}$ and $\mathbb{m} \in \mathrm{S}^{n-1}$ are now independent. The set H is a class ∞ manifold in $\mathrm{Lin}\times\mathrm{R}^n$ of dimension $(m+1)n-1$ with the tangent space given by (8.2) (below). We write $\mathrm{D}\,\tilde{\mathbb{f}} = (\mathrm{D}_1\,\tilde{\mathbb{f}}, \mathrm{D}_2\,\tilde{\mathbb{f}})$ for the components of the derivative of $\tilde{\mathbb{f}}$ in Lin and R^n, respectively. The stress relations (2.2) and (2.3) then read

$$\hat{\mathbb{S}} = \mathrm{D}_1\,\tilde{\mathbb{f}}, \tag{2.6}$$

$$\hat{\mathbb{C}} = \tilde{\mathbb{f}}\,\mathbb{P} - \mathbf{F}^{\mathrm{T}}\,\mathrm{D}_1\,\tilde{\mathbb{f}} - \mathbb{m}\otimes\mathrm{D}_2\,\tilde{\mathbb{f}} \tag{2.7}$$

where $\hat{\mathbb{S}}$, $\hat{\mathbb{C}}$ are evaluated at $(\mathbb{F},\mathbb{m}) \in \mathrm{G}$ and $\tilde{\mathbb{f}}$ and its derivatives at *any* $(\mathbf{F},\mathbb{m}) \in \mathrm{H}$ satisfying $\mathbb{F} = \mathbf{F}\mathbb{P}$. For a given $\mathbb{F}$, there are infinitely many $\mathbf{F}$ satisfying the last relation. The four most natural choices of $\mathbf{F}$ are: the tensor $\mathbb{F}$ itself, the two limiting values $\mathbf{F}_\alpha$, $\alpha = 1, 2$, of the bulk deformation gradient on the interface and their mean value $\mathbf{E} := \langle\mathbf{F}\rangle := \frac{1}{2}(\mathbf{F}_1 + \mathbf{F}_2)$ (see below).

(ii) The possibility discussed in (i) has the interface normal $\mathbb{m}$ restriced by $|\mathbb{m}| = 1$. It is further possible to extend the function $\tilde{\mathbb{f}} : \mathrm{Lin}\times\mathrm{S}^{n-1} \to \mathrm{R}$ to a function $\bar{\mathbb{f}} : \mathrm{Lin}\times(\mathrm{R}^n \sim \{\mathbf{0}\}) \to \mathrm{R}$ by setting

$$\bar{\mathbb{f}}(\mathbf{F}, t\,\mathbb{m}) = \tilde{\mathbb{f}}(\mathbf{F},\mathbb{m})$$

for all $\mathbf{F} \in \mathrm{Lin}$ and $\mathbb{m} \in \mathrm{S}^{n-1}$ and all $t > 0$. If we write $\mathrm{D}\,\bar{\mathbb{f}} = (\mathrm{D}_1\,\bar{\mathbb{f}}, \mathrm{D}_2\,\bar{\mathbb{f}})$ for the components of the derivative of $\bar{\mathbb{f}}$ in Lin and R^n, respectively, then the stress relations (2.2) and (2.3) read

$$\hat{\mathbb{S}} = \mathrm{D}_1\,\bar{\mathbb{f}},$$

$$\hat{\mathbb{C}} = \bar{\mathbb{f}}\,\mathbb{P} - \mathbf{F}^{\mathrm{T}}\,\mathrm{D}_1\,\bar{\mathbb{f}} - \mathbb{m}\otimes\mathrm{D}_2\,\bar{\mathbb{f}}$$

where $\hat{\mathbb{S}}$, $\hat{\mathbb{C}}$ are evaluated at $(\mathbb{F},\mathbb{m}) \in \mathrm{G}$ and $\bar{\mathbb{f}}$ and its derivatives at any $(\mathbf{F},\mathbb{m})$ such that $\mathbf{F}\mathbb{P} = \mathbb{F}$.

(iii) Let $m = n$ and assume that the domain $\mathbb{U}$ of $\hat{\mathbb{f}}$ is

$$\mathrm{G}_0 = \{(\mathbb{F},\mathbb{m}) \in \mathrm{G} : \mathrm{cof}\,\mathbb{F} \neq \mathbf{0}\}.$$

Then the normal $\mathbb{m}$ in the pair $(\mathbb{F},\mathbb{m})\in \mathrm{G}_0$ is locally a function of $\mathbb{F}$ [see Proposition 8.3 (below)]. This in turn implies that there exists a function $\mathbb{f}_\circ$ defined on a relatively open subset of

$$\mathrm{Lin}_{n-1} := \{\mathbb{F}\in\mathrm{Lin} : \det\mathbb{F}=0, \operatorname{cof}\mathbb{F}\neq \mathbf{0}\}$$

such that

$$\mathbb{f}_\circ(\mathbb{F}) = \hat{\mathbb{f}}(\mathbb{F},\mathbb{m}) \tag{2.8}$$

for each $(\mathbb{F},\mathbb{m})$ from some neighborhood of an arbitrarily chosen point of G_0. (The last relation is global if $\hat{\mathbb{f}}$ satisfies the mild constraint $\hat{\mathbb{f}}(\mathbb{F},\mathbb{m}) = \hat{\mathbb{f}}(\mathbb{F},-\mathbb{m})$ for each $(\mathbb{F},\mathbb{m})\in\mathrm{G}_0$.) The derivative of $\mathbb{f}_\circ$ is an element of the tangent space of Lin_{n-1} [see (8.3) (below)]. The interfacial stress relations (2.2) and (2.3) read

$$\hat{\mathbb{S}} = \mathrm{D}\,\mathbb{f}_\circ\,\mathbb{P}, \tag{2.9}$$

$$\hat{\mathbb{C}} = \mathbb{f}_\circ\,\mathbb{P} - \mathbb{F}^{\mathrm{T}}\,\mathrm{D}\,\mathbb{f}_\circ\,\mathbb{P} + \mathbb{m}\otimes\mathbb{F}^{\mathrm{T}}\,\mathrm{D}\,\mathbb{f}_\circ\,\mathbb{m} \tag{2.10}$$

for arguments related as above.

Definition 2.3 (States). Given the constitutive information from Definitions 2.1, we say that $(\mathbf{y},E)$ is a ***state*** if $\mathbf{y}:\Omega\to\mathrm{R}^m$ is a continuous map and E is an open subset of Ω such that

(i) $\mathcal{S} := \Omega\cap\operatorname{bd}E$ is a class 2 surface of dimension $n-1$ of normal $\mathbb{m}:\mathcal{S}\to\mathrm{S}^{n-1}$;

(ii) with the notation

$$E_1 := E,\quad E_2 = \Omega\sim\operatorname{cl}E$$

the maps $\mathbf{y}_\alpha := \mathbf{y}\,|\,E_\alpha$, $\alpha=1,2$, and $\mathbb{y} := \mathbf{y}\,|\,\mathcal{S}$ are of class 2 with their gradients $\nabla\mathbf{y}_\alpha$, and $\mathbb{\nabla}\,\mathbb{y}$ having continuous extensions $\mathbf{F}_\alpha$ and $\mathbb{F}$ to the closure of their respective domains;

(iii) we have $\operatorname{ran}\mathbf{F}_\alpha\subset U_\alpha$, $\alpha=1,2$, and $\operatorname{ran}\mathbb{F}\subset\mathbb{U}$.

Here $f\,|\,M$ denotes the restriction of a map f to a subset M of its domain of definition $\operatorname{dom}f$ and $\operatorname{ran}f=\{f(x):x\in\operatorname{dom}f\}$ denotes the range of f. One has

$$\mathbf{F}_1=\nabla\mathbf{y}\quad\text{in } E,\qquad \mathbf{F}_2=\nabla\mathbf{y}\quad\text{in } \Omega\sim\operatorname{cl}E, \tag{2.11}$$

and the values of $\mathbf{F}_\alpha$ on $\operatorname{cl}E_\alpha\sim E_\alpha$ are the limits of the gradients in (2.11). In particular, $\mathbf{F}_\alpha$ are well defined on $\mathcal{S}$ and we denote by $[\mathbf{F}] := \mathbf{F}_1|\mathcal{S}-\mathbf{F}_2|\mathcal{S}$ the jump of the deformation gradient across the interface. However, let us emphasize that $\mathbf{y}$ is continuous. Also,

$$\mathbb{F} = \mathbb{\nabla}\,\mathbb{y} = \mathbb{\nabla}\,\mathbf{y}\quad\text{on } \mathcal{S} \tag{2.12}$$

and $\mathbb{F}:\operatorname{cl}\mathcal{S}\to\mathrm{Lin}$ is the continuous extension of the surface gradient in (2.12).

Definition 2.4 (Energy and stresses associated with states). Let $(\mathbf{y},E)$ be a state. We define

(i) the ***energy*** $\mathsf{E}(\mathbf{y},E)$ of the state by

$$\mathsf{E}(\mathbf{y},E) = \mathsf{E}_{\mathrm{b}}(\mathbf{y},E) + \mathsf{E}_{\mathrm{if}}(\mathbf{y},E) \tag{2.13}$$

where

$$\mathsf{E}_{\mathrm{b}}(\mathbf{y},E) = \int_E \hat{f}_1(\nabla\mathbf{y})\,d\mathcal{L}^n + \int_{\Omega\sim E}\hat{f}_2(\nabla\mathbf{y})\,d\mathcal{L}^n, \tag{2.14}$$

$$\mathsf{E}_{\mathrm{if}}(\mathbf{y}, E) = \int_{\mathcal{S}} \hat{\mathbb{f}}(\nabla \mathbb{y}, \mathbb{m})\, d\mathcal{H}^{n-1}$$

are the ***bulk*** and ***interfacial energies***, respectively;

(ii) the ***bulk standard stress*** $\mathbf{S}$, the ***bulk configurational stress*** $\mathbf{C}$ and the ***bulk energy density*** f on $\Omega \sim \mathcal{S}$ by

$$\mathbf{S} = \begin{cases} \hat{\mathbf{S}}_1 \circ \mathbf{F}_1 & \text{on} \quad E, \\ \hat{\mathbf{S}}_2 \circ \mathbf{F}_2 & \text{on} \quad \Omega \sim \operatorname{cl} E, \end{cases}$$

and similarly for $\mathbf{C}$ and f; here and below $\hat{\mathbf{S}}_1 \circ \mathbf{F}_1$ denotes the composition of the maps $\hat{\mathbf{S}}_1$ and $\mathbf{F}_1$, i.e., $(\hat{\mathbf{S}}_1 \circ \mathbf{F}_1)(\mathbf{x}) = \hat{\mathbf{S}}_1(\mathbf{F}_1(\mathbf{x}))$ for each $\mathbf{x} \in \operatorname{dom} \mathbf{F}_1 = \operatorname{cl} E$ and similarly for compositions of general maps;

(iii) the ***jumps*** $[\mathbf{S}]$, $[\mathbf{C}]$ $[f]$ of the bulk stresses on $\mathcal{S}$ and of the bulk energy on $\mathcal{S}$ by

$$[\mathbf{S}] = \hat{\mathbf{S}}_1 \circ \mathbf{F}_1 | \mathcal{S} - \hat{\mathbf{S}}_2 \circ \mathbf{F}_2 | \mathcal{S}$$

and similarly for $[\mathbf{C}]$, $[f]$ and $[\mathbf{S}\mathbb{m} \cdot \mathbf{F}\mathbb{m}]$;

(iv) the ***interfacial standard stress*** $\mathbb{S}$, ***interfacial configurational stress*** $\mathbb{C}$, and the ***interfacial energy density*** $\mathbb{f}$ on $\mathcal{S}$ by

$$\mathbb{S} = \hat{\mathbb{S}} \circ (\mathbb{F}, \mathbb{m})$$

and similarly for $\mathbb{C}$, and $\mathbb{f}$, where we use the notation of Definition 2.3.

As in case of the jump of $\mathbf{F}$, the jumps defined in (iii) are the differences of the limits of the corresponding bulk fields from the two sides of the interface. Note that $\mathbb{S}$ and $\mathbb{C}$ are superficial tensors, i.e., $\mathbb{S}\mathbb{m} = \mathbb{C}\mathbb{m} = \mathbf{0}$.

3 Smooth minimizers of energy and equilibrium equations

The condition of minimum energy leads to equilibrium equations in the bulk phases and on the interface. Below I present a derivation within the framework of the present constitutive theory emphasizing the role of outer and inner variations to obtain the balances of standard forces and configurational forces, respectively. In the absence of interfacial energy, the outer and inner variations (to be defined below) are known to lead to the Euler Lagrange and to the Noether equations, respectively, [11; Chaper 3], and indeed the bulk equations $(3.4)_{1,2}$ are the respective instances of these.

Definition 3.1 (Local perturbations and minima).

(i) A state $(\mathbf{z}, F)$ is said to be a ***local perturbation*** of the state $(\mathbf{y}, E)$ if there exists a compact subset K of Ω with

$$\mathbf{z} | (\Omega \sim K) = \mathbf{z} | (\Omega \sim K), \quad F \cap (\Omega \sim K) = E \cap (\Omega \sim K).$$

(ii) The state $(\mathbf{y}, E)$ is said to be a ***local minimizer of energy*** if $\mathsf{E}(\mathbf{y}, E) \le \mathsf{E}(\mathbf{z}, F)$ for each local perturbation $(\mathbf{z}, F)$ of $(\mathbf{y}, E)$.

Thus a local perturbation $(\mathbf{z}, F)$ is identical with the state $(\mathbf{y}, E)$ near the boundary of Ω and (ii) considers the minima of total energy in this class of states. For the considerations below, and in particular for the validity of the interfacial configuration force balance, it is crucial that the interface in the state $(\mathbf{z}, F)$ can be different from that of $(\mathbf{y}, E)$ (apart from the mentioned coincidence near the boundary of Ω). Thus in passing from $(\mathbf{y}, E)$ to $(\mathbf{z}, F)$, part of the phase 1 is transformed into the phase 2 and/or conversely. A stronger notion of minimum is considered in the existence theorems in Section 6. The reader is referred to [27] and [10] for different but related notions of minima.

Lemma 3.2 (Outer and inner variations). *Let $(\mathbf{y}, E)$ be a state. With the notation of Definitions* 2.3 *and* 2.4 *we have the following statements, in which $t \in \mathrm{R}$ is a parameter and $\delta > 0$ a number with $|t|$, δ sufficiently small:*

(i) *Let $\boldsymbol{\alpha} \in C_0^\infty(\Omega, \mathrm{R}^m)$ and let $\mathbf{y}_t : \Omega \to \mathrm{R}^m$ be defined by*

$$\mathbf{y}_t = \mathbf{y} + t\boldsymbol{\alpha}.$$

Then $(\mathbf{y}_t, E)$ is a state that is a local perturbation of $(\mathbf{y}, E)$, the function $t \mapsto \mathsf{E}(\mathbf{y}_t, E)$ is continuously differentiable and

$$\left.\frac{d\mathsf{E}(\mathbf{y}_t, E)}{dt}\right|_{t=0} = \int_{\Omega \sim \mathcal{S}} \mathbf{S} \cdot \nabla \boldsymbol{\alpha} \, d\mathcal{L}^n + \int_{\mathcal{S}} \mathbb{S} \cdot \mathbb{\nabla} \boldsymbol{\alpha} \, d\mathcal{H}^{n-1}. \tag{3.1}$$

*The family $\{(\mathbf{y}_t, E) : |t| < \delta\}$ is said to be an **outer variation** of $(\mathbf{y}, E)$.*

(ii) *Let $\boldsymbol{\beta} \in C_0^\infty(\Omega, \mathrm{R}^n)$ and let $\boldsymbol{\phi}_t : \Omega \to \mathrm{R}^n$ be defined by*

$$\boldsymbol{\phi}_t(\mathbf{x}) = \mathbf{x} + t\boldsymbol{\beta}(\mathbf{x}),$$

$\mathbf{x} \in \Omega$. *Then $\boldsymbol{\phi}_t$ maps Ω bijectively onto Ω; if we define*

$$\mathbf{y}_t = \mathbf{y} \circ \boldsymbol{\phi}_t^{-1}, \quad E_t = \boldsymbol{\phi}_t(E)$$

then $(\mathbf{y}_t, E_t)$ is a state that is a local perturbation of $(\mathbf{y}, E)$, the function $t \mapsto \mathsf{E}(\mathbf{y}_t, E_t)$ is continuously differentiable and

$$\left.\frac{d\mathsf{E}(\mathbf{y}_t, E_t)}{dt}\right|_{t=0} = \int_{\Omega \sim \mathcal{S}} \mathbf{C} \cdot \nabla \boldsymbol{\beta} \, d\mathcal{L}^n + \int_{\mathcal{S}} \mathbb{C} \cdot \mathbb{\nabla} \boldsymbol{\beta} \, d\mathcal{H}^{n-1}. \tag{3.2}$$

*The family $\{(\mathbf{y}_t, E_t) : |t| < \delta\}$ is said to be an **inner variation** of $(\mathbf{y}, E)$.*

The outer variations do not change the referential position of the interface, whereas the inner variations do. The role of standard and configurational stresses was analyzed from different standpoints in many works (apart form the pioneering work of Eshelby [7–8], the reader is referred to [17, 22, 28–29] for bulk quantities and [20, 19, 16–17, 32] for the interface quantities). The derivation below justifies the particular forms of the interfacial stress relations postulated above. Note also that the stress relations continue to hold also in dynamical situations, although it is well known that variational arguments do not suffice [17, 29].

Proof Throughout the proof, a superimposed dot denotes the derivative of a quantity parametrized by t at $t = 0$ and $\mathbf{F}$, $\mathbb{F}$ and $\mathbb{m}$ refer to the deformation gradients and the normal of the state $(\mathbf{y}, E)$ with the interface $\mathcal{S}$. Continuity considerations show that for sufficiently small values of $|t|$ the pairs $(\mathbf{y}_t, E)$ and $(\mathbf{y}_t, E_t)$ of (i) and (ii) satisfy the requirements of Definition 2.3 and thus are states. We omit the derivations of the bulk contributions in (3.1) and (3.2).

(i): The surface deformation gradient $\mathbb{F}_t$ corresponding to $(\mathbf{y}_t, E)$ is

$$\mathbb{F}_t = \mathbb{F} + t\nabla \boldsymbol{\alpha};$$

while the interface normal is unchanged; hence

$$\mathsf{E}_{\mathrm{if}}(\mathbf{y}_t, E) = \int_{\mathcal{S}} \hat{\mathbb{f}}(\mathbb{F} + t\nabla \boldsymbol{\alpha}, \mathbb{m})\, d\mathcal{H}^{n-1};$$

a differentiation and the interfacial stress relation provide

$$\frac{d\mathsf{E}_{\mathrm{if}}(\mathbf{y}_t, E)}{dt}\bigg|_{t=0} = \int_{\mathcal{S}} \mathrm{D}_1 \hat{\mathbb{f}} \cdot \nabla \boldsymbol{\alpha}\, d\mathcal{H}^{n-1} = \int_{\mathcal{S}} \mathbb{S} \cdot \nabla \boldsymbol{\alpha}\, d\mathcal{H}^{n-1}$$

where we have employed $\nabla \boldsymbol{\alpha} = \nabla \boldsymbol{\alpha}\mathbb{P}$. Thus (3.1) follows.

(ii): (Cf. [10; Proof of Proposition 3.2] for a similar derivation within a different constitutive framework.) The interface corresponding to the state $(\mathbf{y}_t, E_t)$ is $\mathcal{S}_t = \boldsymbol{\phi}_t(\mathcal{S})$. The normal $\mathbb{m}_t$ to $\mathcal{S}_t$ is calculated using the standard formula; the composition of $\mathbb{m}_t$ with $\boldsymbol{\phi}_t$ is then given by

$$\bar{\mathbb{m}}_t := \mathbb{m}_t \circ \boldsymbol{\phi}_t = \nabla\boldsymbol{\phi}_t^{-\mathrm{T}}\mathbb{m}/|\nabla\boldsymbol{\phi}_t^{-\mathrm{T}}\mathbb{m}|.$$

The surface deformation gradient $\mathbb{F}_t$ of $\mathbb{y}_t := \mathbf{y}_t|\mathcal{S}_t = \mathbf{y} \circ \boldsymbol{\phi}_t^{-1}|\mathcal{S}_t$ is calculated by the chain rule for the surface gradients as the product of the surface deformation gradient $\mathbb{F}$ of $\mathbb{y} = \mathbf{y}|\mathcal{S}$ and the surface deformation gradient $\nabla\boldsymbol{\phi}_t^{-1}(\mathbf{1} - \bar{\mathbb{m}}_t \otimes \bar{\mathbb{m}}_t)$ of $\boldsymbol{\phi}_t^{-1}|\mathcal{S}_t$. The composite map $\mathbb{F}_t \circ \boldsymbol{\phi}_t$ is given by

$$\bar{\mathbb{F}}_t := \mathbb{F}_t \circ \boldsymbol{\phi}_t = \mathbb{F}\,\boldsymbol{\phi}_t^{-1}(\mathbf{1} - \bar{\mathbb{m}}_t \otimes \bar{\mathbb{m}}_t).$$

Then

$$\mathsf{E}_{\mathrm{if}}(\mathbf{y}_t, E_t) = \int_{\mathcal{S}_t} \hat{\mathbb{f}}(\mathbb{F}_t, \mathbb{m}_t)\, d\mathcal{H}^{n-1}.$$

By the change of variables formula,

$$\mathsf{E}_{\mathrm{if}}(\mathbf{y}_t, E_t) = \int_{\mathcal{S}} \bar{\mathbb{f}}_t \mathbb{J}_t\, d\mathcal{H}^{n-1} \tag{3.3}$$

where $\mathbb{J}_t = |\operatorname{cof} \nabla\boldsymbol{\phi}_t\, \mathbb{m}|$ and $\bar{\mathbb{f}}_t := \hat{\mathbb{f}}(\bar{\mathbb{F}}_t, \bar{\mathbb{m}}_t)$. To calculate the time derivative of the integrand, first note that

$$\dot{\bar{\mathbb{m}}}_t = -\mathbb{P}\,\nabla\dot{\boldsymbol{\phi}}_t^{\mathrm{T}}\mathbb{m},$$

$$\dot{\bar{\mathbb{F}}}_t = -\mathbb{F}\,\nabla\dot{\boldsymbol{\phi}}_t\,\mathbb{P} + \mathbb{F}\,\nabla\dot{\boldsymbol{\phi}}_t^{\mathrm{T}}\mathbb{m} \otimes \mathbb{m}.$$

The chain rule gives

$$\dot{\bar{\mathbb{f}}}_t = -\mathrm{D}_1\hat{\mathbb{f}} \cdot \mathbb{F}\,\nabla\dot{\boldsymbol{\phi}}_t\,\mathbb{P} + \mathrm{D}_1\hat{\mathbb{f}} \cdot \mathbb{F}\,\nabla\dot{\boldsymbol{\phi}}_t^{\mathrm{T}}\mathbb{m} \otimes \mathbb{m} - \mathrm{D}_2\hat{\mathbb{f}} \cdot \mathbb{P}\,\nabla\dot{\boldsymbol{\phi}}_t^{\mathrm{T}}\mathbb{m}$$

where the derivatives of $\hat{\mathbb{f}}$ are evaluated at $(\mathbb{F}, \mathbb{m})$, which can be written as

$$\dot{\bar{\mathbb{f}}}_t = \left(-\mathbb{F}^{\mathrm{T}}\,\mathrm{D}_1\hat{\mathbb{f}}\,\mathbb{P} + \mathbb{m} \otimes (\mathbb{F}^{\mathrm{T}}\,\mathrm{D}_1\hat{\mathbb{f}}\,\mathbb{m} - \mathrm{D}_2\hat{\mathbb{f}})\right) \cdot \nabla\dot{\boldsymbol{\phi}}_t$$

where $(2.4)_2$ has been employed. Thus the differentiation of (3.3) using $\dot{\mathbb{J}}_t = \mathbb{P} \cdot \nabla\dot{\boldsymbol{\phi}}_t$ gives

$$\frac{d\mathsf{E}_{\mathrm{if}}(\mathbf{y}_t, E_t)}{dt}\bigg|_{t=0} = \int_{\mathcal{S}} \mathbb{C} \cdot \boldsymbol{\beta}\, d\mathcal{H}^{n-1}.$$

Thus (3.2) follows. $\square$

Proposition 3.3. *If* $(\mathbf{y}, E)$ *is a local minimzer of energy then*

$$\operatorname{div}\mathbf{S} = \mathbf{0}, \qquad \operatorname{div}\mathbf{C} = \mathbf{0} \qquad \text{in } \Omega \sim \mathcal{S}, \tag{3.4}$$

$$\mathbb{div}\,\mathbb{S} + [\mathbf{S}]\mathbb{m} = \mathbf{0}, \quad \mathbb{div}\,\mathbb{C} + [\mathbf{C}]\mathbb{m} = \mathbf{0} \qquad \text{on } \mathcal{S}, \tag{3.5}$$

where we use the notation of Definitions 2.3 *and* 2.4. *Equation* $(3.4)_2$ *and the tangential component of* $(3.5)_2$ *is a consequence of* $(3.4)_1$ *and* $(3.5)_1$. *Granted* $(3.5)_1$, *the normal component of* $(3.5)_2$ *is equivalent to*

$$[f - \mathbf{S}\mathbb{m} \cdot \mathbf{F}\mathbb{m}] - (\mathbb{f}\mathbb{P} - \mathbb{F}^{\mathrm{T}}\mathbb{S}) \cdot \mathbb{L} + \mathbb{div}\,\mathbb{t} = 0 \qquad \text{on } \mathcal{S}, \tag{3.6}$$

where $\mathbb{L} = \nabla\mathbb{m}$ *is the curvature tensor and* $\mathbb{t} : \mathcal{S} \to \mathrm{R}^n$ *is given by*

$$\mathbb{t} = \mathbb{F}^{\mathrm{T}}\,\mathrm{D}_1\hat{\mathbb{f}} \circ (\mathbb{F}, \mathbb{m})\,\mathbb{m} - \mathrm{D}_2\hat{\mathbb{f}} \circ (\mathbb{F}, \mathbb{m}).$$

These are the balances of standard and configurational forces. Here $(3.4)_1$ is standard, $(3.4)_2$ is derived in [7–8], $(3.5)_1$ in [18], $(3.5)_2$ in [20, 19, 16–17], and (3.6) in [20, 19, 16–17]. See also [17] for further references.

Proof If $(\mathbf{y}_t, E)$ and $(\mathbf{y}_t, E_t)$ are the variations of $(\mathbf{y}, E)$ as in Lemma 3.2(i), (ii), one obtains from the local minimum condition that

$$\int_{\Omega\sim\mathcal{S}} \mathbf{S} \cdot \nabla\boldsymbol{\alpha}\, d\mathcal{L}^n + \int_{\mathcal{S}} \mathbb{S} \cdot \nabla\boldsymbol{\alpha}\, d\mathcal{H}^{n-1} = 0, \tag{3.7}$$

$$\int_{\Omega\sim\mathcal{S}} \mathbf{C} \cdot \nabla\boldsymbol{\beta}\, d\mathcal{L}^n + \int_{\mathcal{S}} \mathbb{C} \cdot \nabla\boldsymbol{\beta}\, d\mathcal{H}^{n-1} = 0 \tag{3.8}$$

for any $\boldsymbol{\alpha} \in C_0^\infty(\Omega, \mathrm{R}^m)$ and $\boldsymbol{\beta} \in C_0^\infty(\Omega, \mathrm{R}^n)$. Employing the divergence theorem on E and $\Omega \sim \operatorname{cl} E$ to the volume integral and the surface divergence theorem (Section 7) to the surface integrals one obtains

$$\int_{\Omega\sim\mathcal{S}} \boldsymbol{\alpha} \cdot \operatorname{div}\mathbf{S}\, d\mathcal{L}^n + \int_{\mathcal{S}} \boldsymbol{\alpha} \cdot \big([\mathbf{S}]\mathbb{m} + \mathbb{div}\,\mathbb{S}\big)\, d\mathcal{H}^{n-1} = 0,$$

$$\int_{\Omega\sim\mathcal{S}} \boldsymbol{\beta} \cdot \operatorname{div}\mathbf{C}\, d\mathcal{L}^n + \int_{\mathcal{S}} \boldsymbol{\beta} \cdot \big([\mathbf{C}]\mathbb{m} + \mathbb{div}\,\mathbb{C}\big)\, d\mathcal{H}^{n-1} = 0;$$

the arbitrariness of $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ gives the balance equations.

To obtain $(3.4)_2$ and the tangential component of $(3.5)_2$ as a consequence of $(3.4)_1$ and $(3.5)_1$, we note that the last two relations are equivalent to (3.7) for each $\boldsymbol{\alpha} \in C_0^\infty(\Omega, \mathrm{R}^m)$. Let $\boldsymbol{\beta} \in C_0^\infty(\Omega, \mathrm{R}^n)$ be such that

$$\boldsymbol{\beta} \cdot \mathbb{m} = 0 \qquad \text{on } \mathcal{S}, \tag{3.9}$$

and apply (3.7) with $\boldsymbol{\alpha}$ given by

$$\boldsymbol{\alpha} = \begin{cases} \mathbf{F}\boldsymbol{\beta} & \text{on } \Omega \sim \mathcal{S}, \\ \mathbb{F}\boldsymbol{\beta} & \text{on } \mathcal{S}. \end{cases}$$

By (3.9), $\boldsymbol{\alpha}$ is a continuous vectorfield that is continuously differentiable on $\Omega \sim \mathcal{S}$ and on $\mathcal{S}$. Using the stress relation $(2.1)_1$ and the identity

$$\nabla(\mathbf{F}\boldsymbol{\beta}) = \nabla_{\boldsymbol{\beta}}\mathbf{F} + \mathbf{F}\nabla\boldsymbol{\beta},$$

in which $\nabla_{\boldsymbol{\beta}}\mathbf{F}$ is the directional derivative of $\mathbf{F}$ in the direction $\boldsymbol{\beta}$, one obtains

$$\mathbf{S}\cdot\nabla(\mathbf{F}\boldsymbol{\beta})=\boldsymbol{\beta}\cdot\nabla f+\mathbf{F}^{\mathrm{T}}\mathbf{S}\cdot\nabla\boldsymbol{\beta}$$

where f is the field of the free energy corresponding to the state $(\mathbf{y},E)$. Integrating by parts the term $\boldsymbol{\beta}\cdot\nabla f$ and using (3.9) one obtains

$$\int_{\Omega\sim\mathcal{S}}\mathbf{S}\cdot\nabla\boldsymbol{\alpha}\,d\mathcal{L}^n=-\int_{\Omega\sim\mathcal{S}}\mathbf{C}\cdot\nabla\boldsymbol{\beta}\,d\mathcal{L}^n.$$

Similarly, we use the stress relation (2.2), the chain rule, the commutation formula

$$\mathbb{\nabla}(\mathbb{F}\boldsymbol{\beta})=\mathbb{\nabla}_{\boldsymbol{\beta}}\mathbb{F}+\mathbb{F}\mathbb{L}\boldsymbol{\beta}\otimes\mathbb{m}+\mathbb{F}\mathbb{\nabla}\boldsymbol{\beta}$$

and [see (3.9)]

$$\mathbb{L}\boldsymbol{\beta}\equiv\mathbb{\nabla}_{\boldsymbol{\beta}}\mathbb{m}=-\mathbb{\nabla}\boldsymbol{\beta}^{\mathrm{T}}\mathbb{m}$$

to obtain

$$\mathbb{S}\cdot\mathbb{\nabla}(\mathbb{F}\boldsymbol{\beta})=\boldsymbol{\beta}\cdot\mathbb{\nabla}\mathbb{f}+(\mathrm{D}_2\hat{\mathbb{f}}-\mathbb{F}^{\mathrm{T}}\mathrm{D}_1\hat{\mathbb{f}}\,\mathbb{m})\cdot\mathbb{\nabla}\boldsymbol{\beta}^{\mathrm{T}}\mathbb{m}+\mathbb{F}^{\mathrm{T}}\mathrm{D}_1\hat{\mathbb{f}}\cdot\mathbb{\nabla}\boldsymbol{\beta}.$$

Integrating by parts the term $\boldsymbol{\beta}\cdot\mathbb{\nabla}\mathbb{f}$ one obtains

$$\int_{\mathcal{S}}\mathbb{S}\cdot\mathbb{\nabla}\boldsymbol{\alpha}\,d\mathcal{H}^{n-1}=-\int_{\mathcal{S}}\mathbb{C}\cdot\mathbb{\nabla}\boldsymbol{\beta}\,d\mathcal{H}^{n-1}.$$

Hence (3.7) implies (3.8) for all $\boldsymbol{\beta}\in C_0^\infty(\Omega,\mathrm{R}^n)$ satisfying (3.9). The arbitrarines of $\boldsymbol{\beta}$ then gives $(3.4)_2$ and the tangential component of $(3.5)_2$.

Finally, we multiply $(3.5)_2$ by $\mathbb{m}$ to obtain

$$\mathbb{m}\cdot\mathbb{div}\,\mathbb{C}+\mathbb{m}\cdot[\mathbf{C}]\mathbb{m}=0$$

and use the formulas

$$\mathbb{div}(\mathbb{C}^{\mathrm{T}}\mathbb{m})=\mathbb{m}\cdot\mathbb{div}\,\mathbb{C}+\mathbb{C}\cdot\mathbb{L},$$

$$\mathbb{C}\cdot\mathbb{L}=(\mathbb{f}\mathbb{P}-\mathbb{F}^{\mathrm{T}}\mathbb{S})\cdot\mathbb{L},\quad \mathbb{m}\cdot[\mathbf{C}]\mathbb{m}=[f-\mathbf{S}\mathbb{m}\cdot\mathbf{F}\mathbb{m}],\quad \mathbb{C}^{\mathrm{T}}\mathbb{m}=\mathbb{t}$$

to obtain (3.6). □

4 The exchange of the actual and reference configurations

This section we discusses the exchange of the roles of the standard and configurational stresses under the exchange of the actual and reference configurations. We consider the format the energy of Section 2 with $m=n$.

Given a state $(\mathbf{y},E)$ with $\mathbf{y}$ injective, the actual configuration of the body is $\bar{\Omega}:=\mathbf{y}(\Omega)$, the actual configuration of the interface is $\bar{\mathcal{S}}:=\mathbb{y}(\mathcal{S})$, and the spatial interface normal $\bar{\mathbb{m}}\circ\mathbb{y}=\operatorname{cof}\mathbb{F}\mathbb{m}/|\operatorname{cof}\mathbb{F}\mathbb{m}|=\operatorname{cof}\mathbb{F}\mathbb{m}/|\operatorname{cof}\mathbb{F}|$. The fields of referential energy densities f, $\mathbb{f}$ and referential stresses $\mathbf{S}$, $\mathbf{C}$, $\mathbb{S}$, $\mathbb{C}$ associated with the state $(\mathbf{y},E)$ can be transformed to the actual configuration of the body via the formulas

$$\bar{f}\circ\mathbf{y}=f/J,\quad \bar{\mathbf{S}}\circ\mathbf{y}=\mathbf{S}\mathbf{F}^{\mathrm{T}}/J,\quad \bar{\mathbf{C}}\circ\mathbf{y}=\mathbf{C}\mathbf{F}^{\mathrm{T}}/J,$$

$$\bar{\mathbb{f}}\circ\mathbb{y}=\mathbb{f}/\mathbb{J},\quad \bar{\mathbb{S}}\circ\mathbb{y}=\mathbb{S}\mathbb{F}^{\mathrm{T}}/\mathbb{J},\quad \bar{\mathbb{C}}\circ\mathbb{y}=\mathbb{C}\mathbb{F}^{\mathrm{T}}/\mathbb{J},$$

where $J=|\det\mathbf{F}|$ and $\mathbb{J}=|\operatorname{cof}\mathbb{F}|$ are the bulk and interface jacobians, measuring the change of the volumes and areas under the passage to the actual configuration. The above transformation formulas are dictated by the geometry alone. The equilibrium equations take the forms

$$\begin{aligned} &\operatorname{Div} \bar{\mathbf{S}} = \mathbf{0}, && \operatorname{Div} \bar{\mathbf{C}} = \mathbf{0} && \text{in } \bar{\Omega} \sim \bar{\mathcal{S}}, \\ &\mathbb{D}\mathrm{iv}\, \bar{\mathbb{S}} + [\bar{\mathbf{S}}]\bar{\mathbb{m}} = \mathbf{0}, && \mathbb{D}\mathrm{iv}\, \bar{\mathbb{C}} + [\bar{\mathbf{C}}]\bar{\mathbb{m}} = \mathbf{0} && \text{on } \bar{\mathcal{S}}, \end{aligned}$$

where $\operatorname{Div}, \mathbb{D}\mathrm{iv}$ denote the spatial bulk and surface divergences.

Let the symbol $\hat{f}$ stand for any of the energy functions $\hat{f}_a$, $a = 1, 2$. Assume that the domain of definition of $\hat{f}$ is

$$U = \operatorname{Lin}_+ := \{ \mathbf{F} \in \operatorname{Lin} : \det \mathbf{F} > 0 \},$$

and that the domain of definition of $\hat{\mathbb{f}}$ is

$$\mathbb{U} = \mathrm{G}_0 = \{ (\mathbb{F}, \mathbb{m}) \in \operatorname{Lin} \times \mathrm{S}^{n-1} : \mathbb{F}\mathbb{m} = \mathbf{0},\ \operatorname{cof} \mathbb{F} \neq \mathbf{0} \}.$$

Under the exchange of the actual and reference configurations, the deformation $\mathbf{y}$ is replaced by its inverse $\mathbf{y}^{-1}$ and hence the bulk deformation gradient $\mathbf{F}$ is replaced by the inverse $\mathbf{F}^{-1}$, the surface deformation gradient $\mathbb{F}$ by the pseudoinverse $\mathbb{F}^{-1}$ (see Section 7) and the referential interface normal $\mathbb{m}$ by the spatial normal $\bar{\mathbb{m}}$.

Using the above transformation formulas, one finds that under the exchange of the actual and reference configurations the response functions $\hat{f}$ and $\hat{\mathbb{f}}$ change to the response functions $\hat{f}^{\#} : U \to \mathrm{R}$ and $\hat{\mathbb{f}}^{\#} : \mathbb{U} \to \mathrm{R}$ given by

$$\begin{aligned} \hat{f}^{\#}(\mathbf{F}) &= \det \mathbf{F}\, \hat{f}(\mathbf{F}^{-1}), \\ \hat{\mathbb{f}}^{\#}(\mathbb{F}, \mathbb{m}) &= |\operatorname{cof} \mathbb{F}|\, \hat{\mathbb{f}}(\mathbb{F}^{-1}, \operatorname{cof} \mathbb{F}\mathbb{m} / |\operatorname{cof} \mathbb{F}|) \end{aligned} \tag{4.1}$$

for each $\mathbf{F} \in U$ and each $(\mathbb{F}, \mathbb{m}) \in \mathbb{U}$. In these definitions, we have denoted by $\mathbf{F}$ and $(\mathbb{F}, \mathbb{m})$ the natural variables of $\hat{f}^{\#}$ and $\hat{\mathbb{f}}^{\#}$, i.e. the variables previously denoted by $\mathbf{F}^{-1}$ and $(\mathbb{F}^{-1}, \bar{\mathbb{m}})$. We let $\hat{\mathbf{S}}$, $\hat{\mathbf{C}}$, $\hat{\mathbb{S}}$ and $\hat{\mathbb{C}}$ denote the response functions for the stresses calculated from $\hat{f}$ and $\hat{\mathbb{f}}$ and the same letters with the superscript $^{\#}$ denote the response functions for the stresses calculated from $\hat{f}^{\#}$ and $\hat{\mathbb{f}}^{\#}$ according to Definition 2.1.

Proposition 4.1. *Under the passage from the response functions from $\hat{f}$ and $\hat{\mathbb{f}}$ to $\hat{f}^{\#}$ and $\hat{\mathbb{f}}^{\#}$ the standard and configurational stresses echange their roles according to*

$$\hat{\mathbf{S}}^{\#}(\mathbf{F}) = \det \mathbf{F}\, \hat{\mathbf{C}}(\mathbf{F}^{-1})\mathbf{F}^{-\mathrm{T}}, \quad \hat{\mathbf{C}}^{\#}(\mathbf{F}) = \det \mathbf{F}\, \hat{\mathbf{S}}(\mathbf{F}^{-1})\mathbf{F}^{-\mathrm{T}},$$

for each $\mathbf{F} \in U$ *and*

$$\hat{\mathbb{S}}^{\#}(\mathbb{F}, \mathbb{m}) = |\operatorname{cof} \mathbb{F}|\, \hat{\mathbb{C}}(\mathbb{F}^{-1}, \bar{\mathbb{m}})\mathbb{F}^{-\mathrm{T}}, \quad \hat{\mathbb{C}}^{\#}(\mathbb{F}, \mathbb{m}) = |\operatorname{cof} \mathbb{F}|\, \hat{\mathbb{S}}(\mathbb{F}^{-1}, \bar{\mathbb{m}})\mathbb{F}^{-\mathrm{T}} \tag{4.2}$$

for each $(\mathbb{F}, \mathbb{m}) \in \mathbb{U}$ *where we write* $\bar{\mathbb{m}} = \operatorname{cof} \mathbb{F}\mathbb{m} / |\operatorname{cof} \mathbb{F}|$ *for brevity.*

The assertion about the bulk response functions is a result of [5], see also [30]; the reader is referred to the proofs there. Assertions (4.2) are the surface counterparts of these; see also [32]. The factor $|\operatorname{cof} \mathbb{F}|$ and of $\mathbb{F}^{-\mathrm{T}}$ in (4.8) indicates that the tensors $\mathbb{S}^{\#}, \mathbb{C}^{\#}$ are transformed from the actual configuration for the reference configuration. The exchange of the roles of $\mathbb{S}$ and $\mathbb{C}$ suggests the interpretation that the Eshelby tensor is a stress tensor of *configurational forces* associated with deformations and defects in the reference configuration [22, 17].

Proof By Remark 2.2(iii) on some neighborhood N of an arbitrarily chosen point of $\mathbb{U}$ there exist functions $\mathbb{f}_\circ$, $\mathbb{S}_\circ$, $\mathbb{C}_\circ$ such that

$$\hat{\mathbb{f}}(\mathbb{G}, \mathbb{m}) = \mathbb{f}_\circ(\mathbb{G}), \quad \hat{\mathbb{S}}(\mathbb{G}, \mathbb{m}) = \mathbb{S}_\circ(\mathbb{G}), \quad \hat{\mathbb{C}}(\mathbb{G}, \mathbb{m}) = \mathbb{C}_\circ(\mathbb{G})$$

for each $(\mathbb{G}, \mathbb{m}) \in N$. Similarly on some neighborhood N' of an arbitrarily chosen point of $\mathbb{U}$ there exist functions $\mathbb{f}_\circ^\#$, $\mathbb{S}_\circ^\#$, $\mathbb{C}_\circ^\#$ such that

$$\hat{\mathbb{f}}^\#(\mathbb{F}, \mathbb{m}) = \mathbb{f}_\circ^\#(\mathbb{F}), \quad \hat{\mathbb{S}}^\#(\mathbb{F}, \mathbb{m}) = \mathbb{S}_\circ^\#(\mathbb{F}), \quad \hat{\mathbb{C}}^\#(\mathbb{F}, \mathbb{m}) = \mathbb{C}_\circ(\mathbb{F})$$

for each $(\mathbb{F}, \mathbb{m}) \in N'$. By (4.1) we have

$$\mathbb{f}_\circ^\#(\mathbb{F}) = |\operatorname{cof}\mathbb{F}|\mathbb{f}_\circ(\mathbb{F}^{-1}) \tag{4.3}$$

for each point $\mathbb{F}$ from some neighborhood of an arbitrarily chosen point of Lin_{n-1}. Evaluating the stress relations (2.9) and (2.10) for the response functions $\mathbb{f}_\circ$, $\mathbb{S}_\circ$, $\mathbb{C}_\circ$ at $\mathbb{F}^{-1}$ one obtains

$$\mathbb{S}_\circ = \mathrm{D}\,\mathbb{f}_\circ\bar{\mathbb{P}}, \tag{4.4}$$

$$\mathbb{C}_\circ = \mathbb{f}_\circ\bar{\mathbb{P}} - \mathbb{F}^{-\mathrm{T}}\,\mathrm{D}\,\mathbb{f}_\circ\bar{\mathbb{P}} + \bar{\mathbb{m}} \otimes \mathbb{F}^{-\mathrm{T}}\,\mathrm{D}\,\mathbb{f}_\circ\bar{\mathbb{m}} \tag{4.5}$$

with the response functions evaluated at $\mathbb{F}^{-1}$ and where $\bar{\mathbb{m}}$ is a unit vector related to $\mathbb{F}^{-1}$ by $\mathbb{F}^{-1}\bar{\mathbb{m}} = \mathbf{0}$ and $\bar{\mathbb{P}} = \mathbf{1} - \bar{\mathbb{m}} \otimes \bar{\mathbb{m}}$. Evaluating the stress relations (2.9) and (2.10) for the response functions $\mathbb{f}_\circ^\#$, $\mathbb{S}_\circ^\#$, $\mathbb{C}_\circ^\#$ at $\mathbb{F}$ one obtains

$$\mathbb{S}_\circ^\# = \mathrm{D}\,\mathbb{f}_\circ^\#\,\mathbb{P}, \tag{4.6}$$

$$\mathbb{C}_\circ^\# = \mathbb{f}_\circ^\#\,\mathbb{P} - \mathbb{F}^{\mathrm{T}}\,\mathrm{D}\,\mathbb{f}_\circ^\#\,\mathbb{P} + \mathbb{m} \otimes \mathbb{F}^{\mathrm{T}}\,\mathrm{D}\,\mathbb{f}_\circ^\#\,\mathbb{m} \tag{4.7}$$

with the response functions evaluated at $\mathbb{F}$ and where $\mathbb{m}$ is a unit vector related to $\mathbb{F}$ by $\mathbb{F}\mathbb{m} = \mathbf{0}$ and $\mathbb{P} = \mathbf{1} - \mathbb{m} \otimes \mathbb{m}$.

If $\mathbb{H}$ and $\hat{\mathbb{J}}$ are maps on Lin_{n-1} defined by

$$\hat{\mathbb{H}}(\mathbb{F}) = \mathbb{F}^{-1}, \quad \hat{\mathbb{J}}(\mathbb{F}) = |\operatorname{cof}\mathbb{F}|$$

then

$$\mathrm{D}\,\hat{\mathbb{H}}(\mathbb{F})\mathbf{A} = -\mathbb{F}^{-1}\mathbf{A}\mathbb{F}^{-1} + \mathbb{m} \otimes \mathbb{F}^{-\mathrm{T}}\mathbb{F}^{-1}\mathbf{A}\mathbb{m} + \mathbb{F}^{-1}\mathbb{F}^{-\mathrm{T}}\mathbf{A}^{\mathrm{T}}\bar{\mathbb{m}} \otimes \bar{\mathbb{m}},$$
$$\mathrm{D}\,\hat{\mathbb{J}}(\mathbb{F}) = |\operatorname{cof}\mathbb{F}|\mathbb{F}^{-\mathrm{T}}$$

for any $\mathbb{F} \in \mathrm{Lin}_{n-1}$ and any $\mathbf{A} \in \mathrm{Lin}$ where $\mathbb{m}, \bar{\mathbb{m}}$ are any of the two pairs of unit vectors such that $\mathbb{F}\mathbb{m} = \mathbb{F}^{-1}\bar{\mathbb{m}} = \mathbf{0}$. Differentiating (4.3) with the help of these formulas one obtains the relationship between the derivatives of $\mathbb{f}_\circ$ and $\mathbb{f}_\circ^\#$ as follows:

$$\mathrm{D}\,\mathbb{f}_\circ^\# = |\operatorname{cof}\mathbb{F}|\big(\mathbb{f}_\circ\mathbb{F}^{-\mathrm{T}} - \mathbb{F}^{-\mathrm{T}}\,\mathrm{D}\,\mathbb{f}_\circ\mathbb{F}^{-\mathrm{T}} + \mathbb{F}^{-\mathrm{T}}\mathbb{F}^{-1}\,\mathrm{D}\,\mathbb{f}_\circ^{\mathrm{T}}\mathbb{m} \otimes \mathbb{m} + \bar{\mathbb{m}} \otimes \mathbb{F}^{-1}\mathbb{F}^{-\mathrm{T}}\,\mathrm{D}\,\mathbb{f}_\circ\bar{\mathbb{m}}\big)$$

where $\mathrm{D}\,\mathbb{f}_\circ^\#$ is evaluated at $\mathbb{F}$ and $\mathbb{f}_\circ$ and its derivatives at $\mathbb{F}^{-1}$ and $\mathbb{m}$, $\bar{\mathbb{m}}$ are as above. Combining this formula with (4.6), (4.7) one obtains

$$\mathbb{S}_\circ^\# = |\operatorname{cof}\mathbb{F}|\big(\mathbb{f}_\circ\bar{\mathbb{P}} - \mathbb{F}^{-\mathrm{T}}\,\mathrm{D}\,\mathbb{f}_\circ\bar{\mathbb{P}} + \bar{\mathbb{m}} \otimes \mathbb{F}^{-\mathrm{T}}\,\mathrm{D}\,\mathbb{f}_\circ\bar{\mathbb{m}}\big)\mathbb{F}^{-\mathrm{T}}$$
$$\mathbb{C}_\circ^\# = |\operatorname{cof}\mathbb{F}|\,\mathrm{D}\,\mathbb{f}_\circ\mathbb{F}^{-\mathrm{T}}$$

which by (4.4) and (4.5) gives

$$\mathbb{S}_\circ^\#(\mathbb{F}) = |\operatorname{cof}\mathbb{F}|\mathbb{C}_\circ(\mathbb{F}^{-1})\mathbb{F}^{-\mathrm{T}}, \quad \mathbb{C}_\circ^\#(\mathbb{F}) = |\operatorname{cof}\mathbb{F}|\mathbb{S}_\circ(\mathbb{F}^{-1})\mathbb{F}^{-\mathrm{T}}, \tag{4.8}$$

i.e., (4.2). □

5 Interface quasiconvexity, null lagrangians and polyconvexity

Let $n \geq 2$ and put

$$s := \min\{m, n\}, \quad t := \min\{m, n-1\}.$$

For the purpose of the following definition, by an oriented surface $\mathcal{S}$ of normal $\mathbb{m}$ we mean a bounded class ∞ surface in R^n of dimension $n-1$ for which $\mathbb{m}$ is a continuous field of unit normal, such that the boundary $\operatorname{bd}\mathcal{S} := \operatorname{cl}\mathcal{S} \sim \mathcal{S}$ is a class ∞ surface of dimension $n-2$, with the orientation of $\operatorname{bd}\mathcal{S}$ dictated by the Stokes theorem. We say that $\mathcal{T}$ is a planar surface of normal $\mathbb{mm}$ if $\mathcal{T}$ is a subset of some $n-1$ dimensional hyperplane in R^n.

Definitions 5.1. Let $\hat{\mathbb{f}} : \mathrm{G} \to \mathrm{R} \cup \{\infty\}$ be a continuous function. We say that $\hat{\mathbb{f}}$ is

(i) ***interface quasiconvex*** if

$$\int_{\mathcal{S}} \hat{\mathbb{f}}(\nabla \mathbb{y}, \mathbb{m})\, d\mathcal{H}^{n-1} \geq \mathcal{H}^{n-1}(\mathcal{T})\hat{\mathbb{f}}(\mathbb{G}, \mathbb{mm}) \tag{5.1}$$

for every $(\mathbb{G}, \mathbb{mm}) \in \mathrm{G}$, every planar surface $\mathcal{T}$ of normal $\mathbb{mm}$, every orented surface $\mathcal{S}$ of normal $\mathbb{m}$ and every continuous map $\mathbb{y} : \operatorname{cl}\mathcal{S} \to \mathrm{R}^m$ that is class 1 on $\mathcal{S}$ such that

$$\operatorname{bd}\mathcal{S} = \operatorname{bd}\mathcal{T} \quad \text{and} \quad \mathbb{y}(\mathbf{x}) = \mathbb{G}\mathbf{x} \ \text{ if } \ \mathbf{x} \in \operatorname{bd}\mathcal{T};$$

(ii) an ***interface null lagrangian*** if $\hat{\mathbb{f}}$ is finite valued and $\pm\hat{\mathbb{f}}$ are interface quasiconvex [in other words, (5.1) holds with the equality sign for each collection of objects listed in (i)];

(iii) ***interface polyconvex*** if $\hat{\mathbb{f}}$ is the supremum of some family of interface null lagrangians.

Remarks 5.2.

(i) Recall from the introduction that the interface quasiconvexity involves a change of the *referential* position of the interface, and note that this is in accordance with inner variations of Section 3.

(ii) It is easily seen that if $\hat{\mathbb{f}}$ is the supremum of some family of interface quasiconvex functions then $\hat{\mathbb{f}}$ is interface quasiconvex; in particular any interface polyconvex function is interface quasiconvex. Any standard (bulk) polyconvex function is also the supremum of some family of standard (bulk) null lagrangians.

(iii) The technical details of the definition of interface quasiconvexity will be discussed in [31]. Thus one may consider the function $\hat{\mathbb{f}}$ occurring above to be a Borel function with locally bounded negative part, one may require that the surface $\mathcal{S}$ be a part of the boundary of an open n dimensional set in R^n, replace the class ∞ quality of $\mathcal{S}$ by its Lipchitz character or even require that $\mathcal{S}$ is an integral $n-1$ dimensional current. Similarly, the deformation $\mathbb{y}$ of $\mathcal{S}$ may be required to be Lipschitz continuous with $\nabla\mathbb{y}$ interpreted as the approximate surface derivative of $\mathbb{y}$.

(iv) When lifted to the graphs of $\mathbb{y}$ and $\mathbb{G}$ in $\mathrm{R}^{m+n} \equiv \mathrm{R}^n \times \mathrm{R}^m$, the interface quasiconvexity is closely related to the semiellipticity of parametric integrands of degree $n-1$ by Almgren and Federer [9; Subsection 5.1.2] as follows. The function $\hat{\mathbb{f}}$ defines [31] a parametric integrand Φ of degree $n-1$ in R^{m+n} defined on simple nonvertical $n-1$ vectors (those with the projection onto R^n different from $\mathbf{0}$). Then modulo the existence of a continuous extension of Φ to the set of all simple $n-1$ vectors, the interface quasiconvexity of $\hat{\mathbb{f}}$ becomes equivalent to the semiellipticity of Φ. It is interesting to note that under this identification, the interface polyconvexity translates to the *convexity* of the parametric integrand. Recall that the graphs of the bulk deformation (interpreted as n dimensional currents in R^{m+n}) are basic to the approach to nonliner elasticity of bulk phases by Giaquinta, Modica & Souček [12–13].

(v) The main motivation of the interface quasiconvexity comes from two related sources. (a): the necessity of the interface quasiconvexity along $\mathbb{y}$ corresponding to a minimizer of the total energy and (b): the necessity and sufficiency of the lowersemicontinuity of the surface energy E_{if} with respect to a suitable convergence of states $(\mathbf{y}, E)$ with migrating interface. These matters are counterparts of the corresponding "bulk" assertions [23–24, 3–4]. Assertions (a) (b) depend on the definition of minimizer, definition of the interface quasiconvexity, and on the convergence in the space of states, as will be discussed in detail in [31]; here we mention only the following versions.

(vi) Roughly, (a) reads: *if $(\mathbf{y}, E)$ is a sufficiently smooth minimizer of the total energy E on the space of states $(\mathbf{y}, E)$ consisting of a Lipschitzian deformation $\mathbf{y}$ and a subset E of Ω of finite perimeter, then $\hat{\mathbb{f}}$ is interface quasiconvex at every $(\mathbb{G}, \mathbb{m}) = (\nabla\mathbb{y}(\mathbf{x}), \mathbb{m}(\mathbf{x}))$ corresponding to $\mathbf{x} \in \mathcal{S}$.* Here a slightly modified version of the interface quasiconvexity is needed, along the lines discussed in (ii). We recall from the introduction that Fonseca [10; Propositions 4.3(i) and 4.9(i)] established quasiconvexity properties of $\hat{\mathbb{f}}$ crresponding to particular choices of $\mathcal{S}$ and $\mathbb{y}$ in Definition 5.1(i) as a consequence of minima defined in [10].

(vii) For the property (b), we mention the following $W^{1,\infty}$ version of lowersemicontinuity: *If $\hat{\mathbb{f}}$ is a finite valued, continuous and interface polyconvex energy and $(\mathbf{y}^i, E^i)$, $i = 1, \ldots$, and $(\mathbf{y}, E)$ are states (consisting of a Lipschitz deformation and of a set of finite perimeter) such that*

$$\mathbf{y}^i \to \mathbf{y} \quad \text{uniformly on } \Omega, \quad \sup\{\mathrm{Lip}(\mathbf{y}^i) : i = 1, \ldots\} < \infty,$$
$$\nabla 1_{E^i} \rightharpoonup^* \nabla 1_E \quad \text{in } \mathcal{M}(\Omega, \mathrm{R}^n)$$

then

$$\liminf_{i\to\infty} \mathsf{E}_{\mathrm{if}}(\mathbf{y}^i, E^i) \geq \mathsf{E}_{\mathrm{if}}(\mathbf{y}, E). \tag{5.2}$$

Here $\mathrm{Lip}(f)$ denotes the Lipschitz constant of the map f, 1_M is the characteristic function of the set M, the gradients of 1_{E^i} and 1_E are interpreted as measures in the space $\mathcal{M}(\Omega, \mathrm{R}^n)$ of vector valued measures in Ω and $\rightharpoonup^*$ denotes the weak* convergence of measures. Equation (5.2) can be reconstructed from the proof of Theorem 6.4 (below).

(viii) If $m = n$, the interface quasiconvex functions, interface null lagrangians, and interface polyconvex functions are preserved unter the exchange of the actual and reference configurations discussed in Section 4, i.e., if $\hat{\mathbb{f}}$ has any of these 3 properties then $\hat{\mathbb{f}}^{\#}$ defined by (4.1) has the same property. This is analogous to the bulk assertions [2].

We now give a complete description of interface null lagrangians as linear combinations, with constant tensorial coefficients, of the members of the list

$$\wedge_k \mathbb{F} \wedge \mathbb{m}, \quad k = 0, \dots, t \tag{5.3}$$

for each $(\mathbb{F}, \mathbb{m}) \in \mathrm{G}$. The reader is referred to Section 9 (below) for the notation. In particular,

$$\wedge_0 \mathbb{F} \wedge \mathbb{m} = \mathbb{m},$$

and if $m = n$ then

$$\wedge_{n-1} \mathbb{F} \wedge \mathbb{m} = *(\operatorname{cof} \mathbb{F}\mathbb{m})$$

For $m = n = 3$ the list (5.3) is equivalent to the list (1.6) described in the introduction. In dimension 3 there are 15 independet (scalar) interface null lagrangians; we recall that there are 20 standard null lagrangians (counting the constant).

Theorem 5.3. *A function* $\hat{\mathbb{f}} : \mathrm{G} \to \mathrm{R}$ *is an interface null lagrangian if and only if it is of the form*

$$\hat{\mathbb{f}}(\mathbb{F}, \mathbb{m}) = \sum_{k=0}^{t} \mathbf{\Omega}_k \cdot (\wedge_k \mathbb{F} \wedge \mathbb{m})$$

for all $(\mathbb{F}, \mathbb{m}) \in \mathrm{G}$ *where*

$$\mathbf{\Omega}_k \in \operatorname{Lin}(\wedge_{k+1} \mathrm{R}^n, \wedge_k \mathrm{R}^m)$$

are constants for all $k = 0, \dots, t$. *If* $m = n = 3$ *then a general form of an interface null lagrangian is*

$$\hat{\mathbb{f}}(\mathbb{F}, \mathbb{m}) = \mathbf{c} \cdot \mathbb{m} + \mathbf{\Omega} \cdot (\mathbb{F} \times \mathbb{m}) + \mathbf{a} \cdot \operatorname{cof} \mathbb{F}\mathbb{m}$$

for each $(\mathbb{F}, \mathbb{m}) \in \mathrm{G}$ *where* $\mathbf{c}, \mathbf{a} \in \mathrm{R}^3$ *and* $\mathbf{\Omega} \in \operatorname{Lin}(\mathrm{R}^3, \mathrm{R}^3)$ *are constants.*

The reader is referred to [31] for a proof and an interpretation of this proposition as well as for the proofs of the assertions below in this section.

Theorem 5.4. *A function* $\hat{\mathbb{f}} : \mathrm{G} \to \mathrm{R} \cup \{\infty\}$ *is interface polyconvex if and only if there exists a positively* 1 *homogeneous function* $\Psi : \mathrm{Y} \to \mathrm{R} \cup \{\infty\}$ *defined on*

$$\mathrm{Y} := \prod_{k=0}^{t} \operatorname{Lin}(\wedge_{k+1} \mathrm{R}^n, \wedge_k \mathrm{R}^m)$$

such that

$$\hat{\mathbb{f}}(\mathbb{F}, \mathbb{m}) = \Psi(\wedge_0 \mathbb{F} \wedge \mathbb{m}, \wedge_1 \mathbb{F} \wedge \mathbb{m}, \dots, \wedge_t \mathbb{F} \wedge \mathbb{m})$$

for each $(\mathbb{F}, \mathbb{m}) \in \mathrm{G}$. *If* $m = n = 3$ *then* $\hat{\mathbb{f}}$ *is interface polyconvex if and only if there exists a a positively* 1 *homogeneous convex function* Φ *on the space* X *such that*

$$\hat{\mathbb{f}}(\mathbb{F}, \mathbb{m}) = \Phi(\mathbb{m}, \mathbb{F} \times \mathbb{m}, \operatorname{cof} \mathbb{F}\mathbb{m}) \tag{5.4}$$

for each $(\mathbb{F}, \mathbb{m}) \in \mathrm{G}$.

Assume that $m = n = 3$ and discuss some interface polyconvex energies. The principle of objectivity requires that realistic energies satisfy

$$\hat{\mathbb{f}}(\mathbf{R}\mathbb{F}, \mathbb{m}) = \hat{\mathbb{f}}(\mathbb{F}, \mathbb{m})$$

for all $(\mathbb{F}, \mathbb{m}) \in \mathrm{G}$ and all rotations $\mathbf{R} \in SO(3)$. In addition, the symmetry of the substance imposes restrictions of the type

$$\hat{\mathbb{f}}(\mathbb{F}\mathbf{R}, \mathbf{R}^{\mathrm{T}}\mathbb{m}) = \hat{\mathbb{f}}(\mathbb{F}, \mathbb{m}) \tag{5.5}$$

for all $(\mathbb{F}, \mathbb{m}) \in \mathrm{G}$ and all rotations $\mathbf{R}$ from the symmetry group of the substance; these will not be discussed here.

We first give a sufficient confition for the interface polyconvexity of "isotropic interfaces" in analogy with the sufficient condition for polyconvexity of isotropic materials by Ball [3].

Proposition 5.5. *Let $n = 3$ and let $g : [0, \infty)^3 \times \mathrm{R}^3 \to \mathrm{R}$ be a positively 1 homogeneous convex function such that*

(i) *for each $s > 0$, $\mathbf{p} \in \mathrm{R}^3$ the function $g(\cdot, \cdot, s, \mathbf{p})$ is symmetric under the exchange of its two arguments,*

(ii) *for each $\mathbf{p} \in \mathrm{R}^3$ the function $g(\cdot, \cdot, \cdot, \mathbf{p})$ is nondecreasing.*

Let $\hat{\mathbb{f}} : \mathrm{G} \to \mathrm{R}$ be defined by

$$\hat{\mathbb{f}}(\mathbb{F}, \mathbb{m}) = g(\lambda_1, \lambda_2, \lambda_1\lambda_2, \mathbb{m})$$

for each $(\mathbb{F}, \mathbb{m}) \in \mathrm{G}$ where $\lambda_1, \lambda_2, 0$ are the singular numbers of $\mathbb{F}$. Then $\hat{\mathbb{f}}$ is interface polyconvex.

The following three examples consider polyconvex energies corresponding to the three varibles on the right hand side of (5.4). Of these examples, the first and the third have direct mechanical relevances.

Example 5.6 (Wulff energy). Let $\hat{\mathbb{f}} : \mathrm{G} \to \mathrm{R}$ be given by

$$\hat{\mathbb{f}}(\mathbb{F}, \mathbb{m}) = \varphi(\mathbb{m}) \tag{5.6}$$

for each $(\mathbb{F}, \mathbb{m}) \in \mathrm{G}$ and $\varphi : \mathrm{S}^{n-1} \to \mathrm{R}$ is an even function; we call $\hat{\mathbb{f}}$ the Wulff energy. It is used to model the growth of crystals, with φ restricted by the symmetry of the lattice. One finds that the stress response functions are

$$\hat{\mathbb{S}} = \mathbf{0}, \quad \hat{\mathbb{C}} = \varphi(\mathbb{m})\mathbb{P} - \mathbb{m} \otimes \mathrm{D}\,\varphi(\mathbb{m})$$

where $\mathbb{P} = \mathbf{1} - \mathbb{m} \otimes \mathbb{m}$. The energy (5.6) produces no standard stress, the interface equilibrium is governed solely by the configurational stress. Already the special case

$$\hat{\mathbb{f}}(\mathbb{F}, \mathbb{m}) = \tau = \mathrm{const} > 0 \tag{5.7}$$

leads to nontrivial phenomena with $\hat{\mathbb{C}} = \tau\mathbb{P}$. [The assumption (5.7) is sometimes confused with the surface tension (see Example 5.8 below).] *The energy $\hat{\mathbb{f}}$ is interface polyconvex if and only if the positively 1 homogeneous extension of φ is convex on R^n.* It is known that there are substances for which the positively 1 homogeneous extension of φ is not convex.

Example 5.7 ("Self–dual" energy). Let $\hat{\mathbb{f}} : \mathrm{G} \to \mathrm{R}$ be given by

$$\hat{\mathbb{f}}(\mathbb{F}, \mathbb{m}) = \Psi(\mathbb{F} \times \mathbb{m}). \tag{5.8}$$

for each $(\mathbb{F}, \mathbb{m}) \in \mathrm{G}$ where $\Psi : \mathrm{Lin} \to \mathrm{R}$. One has

$$\hat{\mathbb{S}} = -\mathrm{D}\,\Psi \times \mathbb{m},$$

$$\hat{\mathbb{C}} = \Psi \mathbb{P} + \mathbb{F}^{\mathrm{T}}\,\mathrm{D}\,\Psi \times \mathbb{m} + \mathbb{m} \otimes \mathbb{P}(\mathbb{F}^{\mathrm{T}}\,\mathrm{D}\,\Psi)^{\times}$$

where for any $\boldsymbol{M} \in \mathrm{Lin}$ we define the axial vector $\boldsymbol{M}^{\times} \in \mathrm{R}^3$ of a tensor $\boldsymbol{M} \in \mathrm{Lin}$ by $\boldsymbol{M}^{\times} \cdot \mathbf{a} = \mathrm{tr}(\boldsymbol{M} \times \mathbf{a})$ for any $\mathbf{a} \in \mathrm{R}^3$. *The energy $\hat{\mathbb{f}}$ is interface polyconvex if and only if Ψ is a positively* 1 *homogeneous convex function.* One can show that if Ψ is positively 1 homogeneous function then the function $\hat{\mathbb{f}}^{\#}$ corresponding to the exchange of the actual and reference configurations (Section 4) is of the same format as $\hat{\mathbb{f}}$ in (5.8), viz.,

$$\hat{\mathbb{f}}^{\#}(\mathbb{F}, \mathbb{m}) = \Psi^{\mathrm{T}}(\mathbb{F} \times \mathbb{m})$$

for each $(\mathbb{F}, \mathbb{m}) \in \mathrm{G}$, where $\Psi^{\mathrm{T}} : \mathrm{Lin} \to \mathrm{R}$ is given by

$$\Psi^{\mathrm{T}}(\mathbb{A}) = \Psi(\mathbb{A}^{\mathrm{T}})$$

for each $\mathbb{A} \in \mathrm{Lin}$. In this sense $\hat{\mathbb{f}}$ is self–dual. The energy

$$\hat{\mathbb{f}}(\mathbb{F}, \mathbb{m}) = c|\mathbb{F}|$$

for each $(\mathbb{F}, \mathbb{m}) \in \mathrm{G}$ where c is a nonnegative constant is a particular case with $\hat{\mathbb{f}}$ polyconvex since $|\mathbb{F}| = |\mathbb{F} \times \mathbb{m}|$.

Example 5.8 (Surface tension). Let $\hat{\mathbb{f}} : \mathrm{G} \to \mathrm{R}$ be given by

$$\hat{\mathbb{f}}(\mathbb{F}, \mathbb{m}) = \psi(\mathrm{cof}\,\mathbb{F}\mathbb{m}), \tag{5.9}$$

for each $(\mathbb{F}, \mathbb{m}) \in \mathrm{G}$, where $\psi : \mathrm{R}^n \to \mathrm{R}$ is a given function. One obtains

$$\hat{\mathbb{S}} = \big((\mathbf{v} \cdot \mathrm{D}\,\psi)\bar{\mathbb{P}} - \mathbf{v} \otimes \mathrm{D}\,\psi\big)\mathbb{F}^{-\mathrm{T}},$$

$$\hat{\mathbb{C}} = \big(\psi - (\mathbf{v} \cdot \mathrm{D}\,\psi)\big)\mathbb{P}$$

where $\mathbf{v} = \mathrm{cof}\,\mathbb{F}\mathbb{m}$ and ψ and its derivative are evaluated at $\mathbf{v}$. *The energy $\hat{\mathbb{f}}$ is interface polyconvex if and only if ψ is a positively* 1 *homogeneous convex function.* One finds that if $\hat{\mathbb{f}}$ is of the form (5.9) with ψ positively 1 homogeneous then $\hat{\mathbb{f}}^{\#}$ is of the form considered in Example 5.6, with $\varphi = \psi$, i.e.,

$$\hat{\mathbb{f}}^{\#}(\mathbb{F}, \mathbb{m}) = \psi(\mathbb{m})$$

for each $(\mathbb{F}, \mathbb{m}) \in \mathrm{G}$. In this sense the present example and Example 5.6 are dual to each other. However, unlike Example 5.6, in the present example for a realistic model the function ψ in (5.9) cannot be arbitrary. Namely, the objectivity (5.5) requires $\psi(\mathbf{R}\mathbf{v}) = \psi(\mathbf{v})$ for each $\mathbf{v} \in \mathrm{R}^3$, $\mathbf{R} \in SO(3)$ which implies that ψ is a multiple of the euclidean norm,

$$\psi(\mathbf{v}) = \sigma|\mathbf{v}|$$

for each $\mathbf{v} \in \mathrm{R}^3$ where σ is a constant, nonnegative if $\hat{\mathbb{f}}$ is polyconvex, i.e.,

$$\hat{\mathbb{f}}(\mathbb{F},\mathbb{m}) = \sigma|\operatorname{cof}\mathbb{F}| \tag{5.10}$$

for each $(\mathbb{F},\mathbb{m}) \in \mathrm{G}$. The stresses corresponding to this particular case are

$$\hat{\mathbb{S}} = \sigma|\operatorname{cof}\mathbb{F}|\mathbb{F}^{-\mathrm{T}}, \quad \hat{\mathbb{C}} = \mathbf{0}. \tag{5.11}$$

Equation (5.10) and (5.11) are the constitutive equations of the surface tension. We note that the spatial standard stress reduces to

$$\bar{\mathbb{S}} = \sigma\bar{\mathbb{P}}$$

where $\bar{\mathbb{P}} = \mathbf{1} - \bar{\mathbb{m}} \otimes \bar{\mathbb{m}}$ with $\bar{\mathbb{m}} = \operatorname{cof}\mathbb{F}\,\mathbb{m}/|\operatorname{cof}\mathbb{F}|$ the normal to the actual configuration of the interface. *The surface tension produces no configurational stress.*

6 The existence of energy minimizing states

This section outlines the existence theory for the minimizers of energy. We first enlarge the state space in Definition 6.1. The reader is referred to Section 9 for the necessary notions of multilinear algebra; furthermore, $\mathcal{M}(\Omega, V)$ denotes the set of all measures on an open set $\Omega \subset \mathrm{R}^n$ with values in a finite dimensional vectorspace V. Recall

$$s := \min\{m,n\}, \quad t := \min\{m,n-1\}.$$

Definitions 6.1 (State spaces for the existence theory). Let $\Omega \subset \mathrm{R}^n$ be a bounded open set with Lipschitz boundary and $t \le p \le \infty$, $1 \le q \le \infty$.

(i) We denote by $\mathcal{G}^p(\Omega,\mathrm{R}^m)$ the set of all pairs $(\mathbf{y},E)$ with $\mathbf{y} \in W^{1,p}(\Omega,\mathrm{R}^m)$ and $E \subset \Omega$ a $\mathcal{L}^n$ measurable set such that for each k with $0 \le k \le t$ there exists a measure $\mathbb{B}_k \in \mathcal{M}(\Omega, \mathrm{Lin}(\wedge_{k+1}\mathrm{R}^n, \wedge_k\mathrm{R}^m))$ satisfying

$$\int_E \wedge_k \nabla\mathbf{y}\, \partial\boldsymbol{\xi}\, d\mathcal{L}^n = (-1)^{k+1}\int_\Omega d\mathbb{B}_k\,\boldsymbol{\xi} \tag{6.1}$$

for each $\boldsymbol{\xi} \in \mathcal{D}_{k+1}(\Omega)$ [in the integral on the right hand side of (6.1) the integration measure $\mathbb{B}_k$ precedes the integrand $\boldsymbol{\xi}$ for algebraic reasons].

For the definitions in (ii) and (iii) we assume that $m = n$.

(ii) We denote by $\mathcal{G}^{p,q}(\Omega,\mathrm{R}^n)$ the set of all $(\mathbf{y},E) \in \mathcal{G}^p(\Omega,\mathrm{R}^n)$ with

$$\wedge_{n-1}\nabla\mathbf{y} \in L^q(\Omega, \mathrm{Lin}(\wedge_{n-1}\mathrm{R}^n, \wedge_{n-1}\mathrm{R}^n)). \tag{6.2}$$

(iii) We denote by $\mathcal{G}_0^{p,q}(\Omega,\mathrm{R}^n)$ the set of all pairs $(\mathbf{y},E)$ where $\mathbf{y} \in W^{1,p}(\Omega,\mathrm{R}^n)$, (6.2) holds, and E is a subset of Ω of finite perimeter.

We call the elements $(\mathbf{y},E)$ of the sets introduced in (i)–(iii) ***states.*** In (i), we call the measure $\mathbb{B}_k$ the interface null lagrangian of order k corresponding to $(\mathbf{y},E)$. We write $\mathbb{B}_k = \mathbb{B}_k(\mathbf{y},E)$ to indicate the dependence on $(\mathbf{y},E)$; we abbreviate $\mathbb{J}(\mathbf{y},E) = (\mathbb{B}_0,\dots,\mathbb{B}_t)$.

The set $\mathcal{G}^p(\Omega,\mathrm{R}^m)$ with $p > s = \min\{m,n\}$ will be employed in Theorem 6.5 (below) with m and n arbitrary. If $m = n$ and only states with deformation gradient of positive determinant can have finite energy, one may work in larger spaces $\mathcal{G}^{p,q}(\Omega, \mathrm{R}^n)$ where $p \geq n-1$ and $q \geq n/(n-1)$ by employing the results of [26], see Theorem 6.5 (below). Finally, the set $\mathcal{G}_0^{p,q}(\Omega, \mathrm{R}^n)$ is employed in case $m = n$ with the interfacial energy depending only on the normal to the interface, Theorem 6.6 (below).

Remarks 6.2.

(i) Underlying the definition of the interface null lagrangians is the vanishing of the exterior derivative of the bulk jacobian minors $\wedge_k \nabla \mathbf{y}$. Namely, if $\mathbf{y} \in W^{1,p}(\Omega, \mathrm{R}^m)$ with $p \geq k$ then

$$\int_\Omega \wedge_k \nabla \mathbf{y} \, \partial \boldsymbol{\xi} \, d\mathcal{L}^n = \mathbf{0} \tag{6.3}$$

for each $\boldsymbol{\xi} \in \mathcal{D}_{k+1}(\Omega)$; we here recall that the interior derivative ∂ is dual (formal adjoint) of the exterior derivative. The reader is referred to [13; Corollary 2, Subsection 3.2.3] for a coordinate version of (6.3). Since (6.1) involves the bulk integral over E, one expects that the integration by parts will result in an object $\mathbb{B}_k$ concentrated on the boundary of E. This is indeed the case, as we shall show now (the reader is referred to [31] for the proofs of (ii)–(iv)).

(ii) *If* $(\mathbf{y}, E) \in \mathcal{G}^p(\Omega, \mathrm{R}^m)$ *with* $p \geq t$ *then* E *is a set of finite perimeter and*

$$\mathbb{B}_0 = \mathbb{m} \, \mathcal{H}^{n-1} \llcorner \operatorname{bd}_*(E, \Omega)$$

where $\operatorname{bd}_*(E, \Omega) := \Omega \cap \operatorname{bd}_* E$. *Moreover, if* $1 \leq k \leq t$ *then*

$$\operatorname{spt} \mathbb{B}_k \subset \operatorname{cl} \operatorname{bd}_*(E, \Omega).$$

(iii) We note that $\mathcal{G}^\infty(\Omega, \mathrm{R}^m) \subset \mathcal{G}^p(\Omega, \mathrm{R}^m)$ for all $p \geq t$. *The set* $\mathcal{G}^\infty(\Omega, \mathrm{R}^m)$ *is the set of all pairs* $(\mathbf{y}, E)$ *where* $\mathbf{y} : \Omega \to \mathrm{R}^m$ *is Lipschitz and* $E \subset \Omega$ *is a set of finite perimeter.*

(iv) *If* $(\mathbf{y}, E) \in \mathcal{G}^\infty(\Omega, \mathrm{R}^m)$ *then the measures* $\mathbb{B}_k$ *are given by*

$$\mathbb{B}_k = \wedge_k \nabla \mathbb{y} \wedge \mathbb{m} \, \mathcal{H}^{n-1} \llcorner \operatorname{bd}_*(E, \Omega)$$

were $\mathbb{y} = \mathbf{y} | \operatorname{bd}_*(E, \Omega)$ *and* $\nabla \mathbb{y}$ *is the approximate surface gradient of the Lipschitz map* $\mathbb{y}$ *on the* $\mathcal{H}^{n-1}$ *rectifiable set* $\operatorname{bd}_*(E, \Omega)$. Thus $\mathbb{B}_k$ are the measure theoretic generalizations of the interface null lagrangians.

Definition 6.3 (Energy functionals for the existence theory). Let $\hat{f}_\alpha : \mathrm{Lin} \to [0, \infty]$, $\alpha = 1, 2$, be functions of the forms

$$\hat{f}_\alpha(\mathbf{F}) = \Phi_\alpha(\wedge_1 \mathbf{F}, \dots, \wedge_s \mathbf{F}) \tag{6.4}$$

for all $\alpha = 1, 2$ and all $\mathbf{F} \in \mathrm{Lin}$, where $\Phi_\alpha : \mathrm{Z} \to [0, \infty]$ are continuous convex functions on

$$\mathrm{Z} = \prod_{k=1}^{s} \mathrm{Lin}(\wedge_k \mathrm{R}^n, \wedge_k \mathrm{R}^m).$$

(i) Let $\hat{\mathbb{f}} : \mathrm{G} \to [0, \infty)$ be a function of the form

$$\hat{\mathbb{f}}(\mathbb{F}, \mathbb{m}) = \Phi(\wedge_0 \mathbb{F} \wedge \mathbb{m}, \dots, \wedge_t \mathbb{F} \wedge \mathbb{m}) \tag{6.5}$$

for each $(\mathbb{F}, \mathbb{m}) \in \mathrm{G}$ where $\Phi : \mathrm{Y} \to [0, \infty)$ is a positively 1 homogeneous convex funtion. If $t \leq p \leq \infty$, we define the ***total energy*** $\mathsf{E} : \mathcal{G}^p(\Omega, \mathrm{R}^m) \to [0, \infty]$ by (2.13) for each $(\mathbf{y}, E) \in \mathcal{G}^p(\Omega, \mathrm{R}^m)$ where E_b is given by (2.14) and

$$\mathsf{E}_{\mathrm{if}}(\mathbf{y},E)=\int_{\mathrm{R}^n}\Phi(\mathbb{A})\,d\,|\mathbb{J}(\mathbf{y},E)| \tag{6.6}$$

where $|\mathbb{J}(\mathbf{y},E)|$ is the total variation of $\mathbb{J}(\mathbf{y},E)$ and $\mathbb{A}:\Omega\to\mathrm{Y}$ satisfies $\mathbb{J}(\mathbf{y},E)=\mathbb{A}\,|\mathbb{J}(\mathbf{y},E)|$; cf. [1; Corollary 1.29 and Section 2.6].

(ii) Let $m=n$ and let $\varphi:\mathrm{R}^n\to[0,\infty)$ be a positively 1 homogeneous convex function. We define the ***total energy*** $\mathsf{E}:\mathcal{G}_0^{p,q}(\Omega,\mathrm{R}^n)\to[0,\infty]$ by (2.13) for each $(\mathbf{y},E)\in\mathcal{G}_0^{p,q}(\Omega,\mathrm{R}^n)$ where E_{b} is given by (2.14) and

$$\mathsf{E}_{\mathrm{if}}(\mathbf{y},E)=\int_{\mathrm{bd}_*(E,\Omega)}\varphi(\mathbb{m})\,d\mathcal{H}^{n-1} \tag{6.7}$$

where $\mathbb{m}$ is the measure theoretic normal to E.

The definition (6.6) reduces to

$$\mathsf{E}_{\mathrm{if}}(\mathbf{y},E)=\int_{\mathrm{bd}_*(E,\Omega)}\hat{\mathbb{f}}(\nabla\mathbf{y},\mathbb{m})\,d\mathcal{H}^{n-1}$$

if $(\mathbf{y},E)\in\mathcal{G}^\infty(\Omega,\mathrm{R}^m)$ by Remark 6.2(iv). The definition (6.7) corresponds to $\hat{\mathbb{f}}:\mathrm{G}\to[0,\infty)$ given by $\hat{\mathbb{f}}(\mathbb{F},\mathbb{m})=\varphi(\mathbb{m})$. The definition of E in Item (i) will be used when $\hat{\mathbb{f}}$ is coercive with respect to its variables. The Wulff energy of Example 5.6 does not satisfy these coercivity requirements; nevertheless, one can prove the existence also in this case; then the definition of E in Item (ii) will be used.

We first consider the case of general m and n.

Theorem 6.4. *Let $s<p<\infty$ and assume that*

(i) $\hat{f}_a$, $a=1,2$, *are polyconvex in the sense of* (6.4) *where Φ_a are continuous convex $[0,\infty]$ valued functions, $\hat{\mathbb{f}}$ is interface polyconvex in the sense of* (6.5) *where Φ is a positively* 1 *homogeneous convex $[0,\infty)$ valued function,*

(ii) *for all $a=1,2$, all $\mathbf{F}\in\mathrm{Lin}$, all $\mathrm{A}\in\mathrm{Y}$, some $c>0$ and some $d\in\mathrm{R}$ we have*

$$\hat{f}_a(\mathbf{F})\ge c|\mathbf{F}|^p+d,\quad \Phi(\mathrm{A})\ge c|\mathrm{A}|.$$

Given $\mathbf{z}_0\in W^{1,p}(\Omega,\mathrm{R}^m)$, consider the Dirichlet class

$$\mathcal{A}(\mathbf{z}_0)=\{(\mathbf{y},E)\in\mathcal{G}^p(\Omega,\mathrm{R}^m):\mathbf{y}=\mathbf{z}_0\ \text{ on }\ \mathrm{bd}\,\Omega\}$$

and let E be given by Definition 6.3(i). *If E is finite on some element of $\mathcal{A}(\mathbf{z}_0)$ then there exists an $(\mathbf{y},E)\in\mathcal{A}(\mathbf{z}_0)$ such that*

$$\mathsf{E}(\mathbf{y},E)\le\mathsf{E}(\mathbf{z},F)$$

for all $(\mathbf{z},F)\in\mathcal{A}(\mathbf{z}_0)$.

Proof Let $\mathbf{M}(\mu)$ denote the mass of the measure $\mu\in\mathcal{M}(\Omega,V)$, i.e., $\mathbf{M}(\mu)=|\mu|(\Omega)$ where $|\mu|$ denotes the total variation of μ. Let $(\mathbf{y}^i,E^i)\in\mathcal{A}(\mathbf{z}_0)$ be a minimizing sequence. By the coercivity assumptions on $\hat{f}_a$ and Φ the sequences $|\nabla\mathbf{y}^i|_{L^p}$ and $\mathbf{M}(\mathbb{J}(\mathbf{y}^i,E^i))$ are bounded. Combining the boundedness of $|\nabla\mathbf{y}^i|_{L^p}$ with the Dirichlet boundary data, one obtains the boundedness of $|\mathbf{y}^i|_{W^{1,p}}$. Standard compactness theorems for Sobolev space and for the spaces of measures give that for some subsequence of $(\mathbf{y}^i,E^i)$, denoted again $(\mathbf{y}^i,E^i)$, and some $\mathbf{y}\in W^{1,p}(\Omega,\mathrm{R}^m)$, $\boldsymbol{\Delta}\in\mathcal{M}(\Omega,\mathrm{Y})$ we have

$$\mathbf{y}^i \rightharpoonup \mathbf{y} \quad \text{in } W^{1,p}(\Omega, \mathrm{R}^m), \tag{6.8}$$

$$\mathbb{J}(\mathbf{y}^i, E^i) \rightharpoonup^* \boldsymbol{\Delta} \quad \text{in } \mathcal{M}(\Omega, \mathrm{Y}). \tag{6.9}$$

From $\mathbb{B}_0(\mathbf{y}^i, E^i) = \mathbb{m}^{E^i} \mathcal{H}^{n-1} \llcorner \operatorname{bd}_*(E^i, \Omega)$ we deduce that the sequence $\mathrm{D}1_{E^i}$ is bounded in $\mathcal{M}(\Omega, \mathrm{R}^n)$. The imbedding theorem from BV functions (e.g., [1; Corollary 3.49, Chapter 3] implies

$$1_{E^i} \to 1_E \quad \text{in } L^1(\Omega). \tag{6.10}$$

for some set $E \subset \Omega$ of finite perimeter, i.e.,

$$\mathcal{L}^n(\Delta(E^i, E)) \to 0, \tag{6.11}$$

where $\Delta(E^i, E)$ is the symmetric difference of E^i and E. The inequality $s < p$, Equation (6.8) and the weak sequential continuity of minors (e.g., [25; Theorem 2.3(ii)]) gives

$$\wedge_k \nabla \mathbf{y}^i \rightharpoonup \wedge_k \nabla \mathbf{y} \quad \text{in } L^{p/s}(\Omega, \mathrm{Lin}(\wedge_k \mathrm{R}^n, \wedge_k \mathrm{R}^m)), \quad 0 \le k \le s. \tag{6.12}$$

The equiintegrability of the sequence $\wedge_k \nabla \mathbf{y}^i$ and (6.11) yield

$$1_{E^i} \wedge_k \nabla \mathbf{y}^i \rightharpoonup 1_E \wedge_k \nabla \mathbf{y} \quad \text{in } L^1(\Omega, \mathrm{Lin}(\wedge_k \mathrm{R}^n, \wedge_k \mathrm{R}^m)), \quad 0 \le k \le s$$

and in particular,

$$\int_{E^i} \wedge_k \nabla \mathbf{y}^i \, \partial \boldsymbol{\xi} \, d\mathcal{L}^n \to \int_E \wedge_k \nabla \mathbf{y} \, \partial \boldsymbol{\xi} \, d\mathcal{L}^n$$

for each $\boldsymbol{\xi} \in \mathcal{D}_{k+1}(\mathrm{R}^n)$, which can be rewritten as

$$\int_\Omega d\mathbb{B}_k(\mathbf{y}^i, E^i) \boldsymbol{\xi} \to \int_E \wedge_k \nabla \mathbf{y} \, \partial \boldsymbol{\xi} \, d\mathcal{L}^n. \tag{6.13}$$

Hence (6.9) yields

$$\int_E \wedge_k \nabla \mathbf{y} \, \partial \boldsymbol{\xi} \, d\mathcal{L}^n = \int_\Omega d\boldsymbol{\Delta}_k \, \boldsymbol{\xi}$$

where we write $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_0, \dots, \boldsymbol{\Delta}_s)$ for the components of $\boldsymbol{\Delta}$. Thus $(\mathbf{y}, E) \in \mathcal{G}^p(\Omega, \mathrm{R}^m)$ and $\mathbb{B}_k(\mathbf{y}, E) = \boldsymbol{\Delta}_k$. Equation (6.13) reduces to

$$\mathbb{J}(\mathbf{y}^i, E^i) \rightharpoonup^* \mathbb{J}(\mathbf{y}, E) \quad \text{in } \mathcal{M}(\Omega, \mathrm{Y}). \tag{6.14}$$

Let $\Xi : \mathrm{R} \times \mathrm{Z} \to [0, \infty]$ be defined by

$$\Xi(\tau, \mathrm{M}) = |\tau| \Phi_1(\mathrm{M})$$

for each $\tau \in \mathrm{R}$ and $\mathrm{M} \in \mathrm{Y}$; note that the function $\Xi(\tau, \cdot)$ is convex for each $\tau \in \mathrm{R}$. We have

$$\int_{E^i} \hat{f}_1(\nabla \mathbf{y}^i) \, d\mathcal{L}^n = \int_\Omega \Xi(1_{E^i}, \wedge_1 \nabla \mathbf{y}^i, \dots, \wedge_s \nabla \mathbf{y}^i) \, d\mathcal{L}^n.$$

The Ioffe lowersemicontinuity theorem [1; Theorem 5.8, Chapter 5] and (6.10) and (6.12) then give

$$\liminf_{i \to \infty} \int_\Omega \Xi(1_{E^i}, \wedge_0 \nabla \mathbf{y}^i, \dots, \wedge_s \nabla \mathbf{y}^i) \, d\mathcal{L}^n \ge \int_\Omega \Xi(1_E, \wedge_0 \nabla \mathbf{y}, \dots, \wedge_s \nabla \mathbf{y}) \, d\mathcal{L}^n.$$

Thus

$$\liminf_{i \to \infty} \int_{E^i} \hat{f}_1(\nabla \mathbf{y}^i) \, d\mathcal{L}^n \ge \int_E \hat{f}_1(\nabla \mathbf{y}) \, d\mathcal{L}^n, \tag{6.15}$$

$$\liminf_{i\to\infty} \int_{\Omega\sim E^i} \hat{f}_2(\nabla \mathbf{y}^i)\, d\mathcal{L}^n \ge \int_{\Omega\sim E} \hat{f}_2(\nabla \mathbf{y})\, d\mathcal{L}^n \tag{6.16}$$

where the last relation is obtained analogously. Using (6.14) and the Reshetnyak lowersemicontinuity theorem (e.g., [1; Theorem 2.38, Chapter 2]), one obtains

$$\liminf_{i\to\infty} \mathsf{E}_{\mathrm{if}}(\mathbf{y}^i, E^i) \ge \mathsf{E}_{\mathrm{if}}(\mathbf{y}, E). \tag{6.17}$$

Thus (6.15), (6.16) and (6.17) provide

$$\liminf_{i\to\infty} \mathsf{E}(\mathbf{y}^i, E^i) \ge \mathsf{E}(\mathbf{y}, E).$$

Clearly, $(\mathbf{y}, E) \in \mathcal{A}(\mathbf{z}_0)$. □

If $m = n$ and the bulk energies $\hat{f}_\alpha$ satisfy Ball's constraint $\hat{f}_\alpha(\mathbf{F}) \to \infty$ as $\det \mathbf{F} \to 0$, one can exclude states $(\mathbf{y}, E)$ with $\det \mathbf{F} \le 0$ and improve the exponent p in the definition of the state space as follows.

Theorem 6.5. *Let $m = n$, $p \ge n-1$, $q \ge n/(n-1)$ and assume that*

(i) $\hat{f}_\alpha$, $\alpha = 1, 2$, *are polyconvex in the sense of* (6.4) *where Φ_α are continuous convex $[0,\infty]$ valued functions, $\hat{\mathbb{f}}$ is interface polyconvex in the sense of* (6.5) *where Φ is a positively* 1 *homogeneous convex $[0,\infty)$ valued function,*

(ii) *for all $\alpha = 1, 2$, all $\mathbf{F} \in \mathrm{Lin}$, all $\mathrm{A} \in \mathrm{Y}$, some $c > 0$ and some $d \in \mathrm{R}$ we have*

$$\hat{f}_\alpha(\mathbf{F}) \ge c(|\mathbf{F}|^p + |\operatorname{cof} \mathbf{F}|^q) + d, \quad \Phi(\mathrm{A}) \ge c|\mathrm{A}|,$$

(iii) $\hat{f}_\alpha(\mathbf{F}) = \infty$ *if* $\det \mathbf{F} \le 0$.

Given $\mathbf{z}_0 \in W^{1,p}(\Omega, \mathrm{R}^m)$, consider the Dirichlet class

$$\mathcal{A}(\mathbf{z}_0) = \{(\mathbf{y}, E) \in \mathcal{G}^{p,q}(\Omega, \mathrm{R}^n) : \mathbf{y} = \mathbf{z}_0 \text{ on } \operatorname{bd}\Omega\}$$

and let E be given by Definition 6.3(i). *If E is finite at some element of $\mathcal{A}(\mathbf{z}_0)$ then there exists an $(\mathbf{y}, E) \in \mathcal{A}(\mathbf{z}_0)$ such that*

$$\mathsf{E}(\mathbf{y}, E) \le \mathsf{E}(\mathbf{z}, F)$$

for all $(\mathbf{z}, F) \in \mathcal{A}(\mathbf{z}_0)$. Each solution $(\mathbf{y}, E)$ of the problem satisfies

$$\det \nabla \mathbf{y} > 0 \text{ for } \mathcal{L}^n \text{ a.e. point of } \Omega. \tag{6.18}$$

The proof follows the lines of the proof of Theorem 6.4; the details will be given in [31]. Theorem 1.2 is a particular case.

Finally, we consider the case of the Wulff energy.

Theorem 6.6. *Let $m = n$, $p \ge n-1$, $q \ge n/(n-1)$, and assume that*

(i) $\hat{f}_\alpha$, $\alpha = 1, 2$, *are polyconvex in the sense of* (6.4) *where Φ_α are continuous convex $[0,\infty]$ valued functions, $\varphi : \mathrm{R}^n \to [0,\infty)$ is a positively* 1 *homogeneous convex function;*

(ii) *for all $\alpha = 1, 2$, all $\mathbf{F} \in \mathrm{Lin}$, all $\mathbf{p} \in \mathrm{R}^n$, some $c > 0$ and some $d \in \mathrm{R}$ we have*

$$\hat{f}_\alpha(\mathbf{F}) \ge c(|\mathbf{F}|^p + |\operatorname{cof} \mathbf{F}|^q) + d, \quad \varphi(\mathbf{p}) \ge c|\mathbf{p}|,$$

(iii) $\hat{f}_\alpha(\mathbf{F}) = \infty$ *if* $\det \mathbf{F} \le 0$.

Given $\mathbf{z}_0 \in W^{1,p}(\Omega, \mathrm{R}^m)$*, consider the Dirichlet class*

$$\mathcal{A}(\mathbf{z}_0) = \{ (\mathbf{y}, E) \in \mathcal{G}_0^{p,q}(\Omega, \mathrm{R}^n) : \mathbf{y} = \mathbf{z}_0 \ \text{ on } \ \mathrm{bd}\,\Omega \}$$

and let E *be given by Definition* 6.3(ii)*. If* E *is finite at some element of* $\mathcal{A}(\mathbf{z}_0)$ *then there exists an* $(\mathbf{y}, E) \in \mathcal{A}(\mathbf{z}_0)$ *such that*

$$\mathsf{E}(\mathbf{y}, E) \le \mathsf{E}(\mathbf{z}, F)$$

for all $(\mathbf{z}, F) \in \mathcal{A}(\mathbf{z}_0)$*. Each solution* $(\mathbf{y}, E)$ *of the problem satisfies* (6.18)*.*

We refer to [31] for the proof.

7 Appendix A. Differentiation on manifolds

We deal with manifolds of (at least) class 1 embedded in finite dimensional vectorspaces [9; Subsections 3.1.19–3.1.20], which we call simply manifolds or synonymously surfaces. In this section we define derivatives (gradients) of maps on manifolds, which we call surface derivatives or surface gradients. We need the surface derivatives in two different ways: (a) for fields defined on the phase interface $\mathcal{S} \subset \mathrm{R}^n$, (b) for the response function $\hat{\mathfrak{f}}$, which is a function defined on the manifold G (see Section 8).

Throughout the section, let V, W be finite dimensional inner product spaces, and f is a map with the domain $\mathrm{dom}\, f$ a relatively open subset of a manifold $\mathcal{M}$ in V with the range $\mathrm{ran}\, f$ in W. If $x \in M$, we denote by $\mathrm{Tan}(\mathcal{M}, x)$ the tangent space to $\mathcal{M}$ at x, a k dimensional subspace of V where k is the dimension of $\mathcal{M}$.

We say that f is differentiable at $x \in \mathrm{dom}\, f$ if there exists a $\mathrm{D} f(x) \in \mathrm{Lin}(V, W)$, called the derivative of f at x, such that

$$\mathrm{D} f(x) P = \mathrm{D} f(x) \tag{7.1}$$

where P is the orthogonal projection onto $\mathrm{Tan}(\mathrm{dom}\, f, x)$ and

$$\lim_{\substack{y \to x \\ y \in \mathrm{dom}\, f,\, y \neq x}} |f(y) - f(x) - \mathrm{D} f(x)(y - x)| / |y - x| = 0.$$

The map $\mathrm{D} f(x)$ is uniquely determined. We note that $\mathrm{D} f(x)$ is a linear transformation defined on the entire space V and not just on the tangent space; however, it vanishes on the orthogonal complement of the tangent space by (7.1). This convetion renders the derivatives of f at different points of $\mathcal{M}$ belong to the same linear space $\mathrm{Lin}(V, W)$. Other authors (e.g., [9; Subsection 3.1.22]) mean by the derivative the restriction of the derivative in the present sense to the tangent space at the given point. If f is differentiable at $x \in \mathcal{M}$ and $a \in V$ we denote by $\mathrm{D}_a f(x) := \mathrm{D} f(x) a$ the directional derivative of f at x in the direction a. If $V = \mathrm{R}^n$ and $\boldsymbol{e}_i$ is the standard basis in R^n, $1 \le i \le n$, we denote by $\mathrm{D}_i f(x) = \mathrm{D} f(x) \boldsymbol{e}_i \equiv \mathrm{D}_{\boldsymbol{e}_i} f(x)$ is the i th partial derivative of f at x. If the range W of f is R, we identify $\mathrm{D} f(x) \in \mathrm{Lin}(V, \mathrm{R})$ with an equally denoted vector in V, such that

$$\mathrm{D} f(x) a = a \cdot \mathrm{D} f(x)$$

for each $a \in V$; then $\mathrm{D} f(x) \in \mathrm{Tan}(\mathcal{M}, x)$.

Observe that if $I : \mathcal{M} \to V$ is the identity map $I(x) = x$ for every $x \in \mathcal{M}$ then

$$\mathrm{D} I(x) = P(x) \tag{7.2}$$

for every $x \in \mathcal{M}$ where $P(x)$ is the orthogonal projection onto $\mathrm{Tan}(\mathcal{M}, x)$.

If $T \in \mathrm{Lin}(V, W)$ is an injective map, we define the pseudoinverse $T^{-1} \in \mathrm{Lin}(W, V)$ as the unique linear map such that

$$T^{-1}T = P, \quad T T^{-1} = Q$$

where P and Q are the orthogonal projections onto $(\ker T)^{\perp}$ and $\operatorname{ran} T$. One has

$$(T^{-1})^{-1} = T.$$

If T maps V bijectively onto W then the pseudoinverse coincides with the usual inverse.

We retain the symbol D for the derivative of the response functions $\hat{f}_\alpha$, $\alpha = 1, 2$, and $\hat{\mathbb{f}}$. However, if the variable x in the definition above has the meaning of the referential position $\mathbf{x}$ of a material point of Ω, we write ∇ for D in case of a map f defined on an open subset of Ω and $\mathbb{\nabla}$ for D in case f is a map defined on the phase interface $\mathcal{S}$ in Ω. If $\boldsymbol{\phi}$ is a local parametrization of $\mathcal{S}$ then f is differentiable at $\mathbf{x} \in \mathcal{S}$ if and only if $f \circ \boldsymbol{\phi}$ is differentiable at $\mathbf{p} := \boldsymbol{\phi}^{-1}(\mathbf{x})$ and then

$$\mathbb{\nabla} f(\mathbf{x}) = \nabla(f \circ \boldsymbol{\phi})(\mathbf{p}) \nabla\boldsymbol{\phi}(\mathbf{p})^{-1}.$$

If g is a extension of f to a neighborhood of $\mathbf{x}$ in R^n that is differentiable at $\mathbf{x}$ then

$$\mathbb{\nabla} f(\mathbf{x}) = \nabla g(\mathbf{x}) \mathbb{P}(\mathbf{x})$$

where $\mathbb{P}(\mathbf{x})$ is the orthogonal projection onto the tangent space of $\mathcal{M}$ at $\mathbf{x}$. If f is a map defined in a neighborhood of $\mathcal{S}$ we use the notation

$$\mathbb{\nabla} f := \mathbb{\nabla}(f | \mathcal{S}).$$

If f is a map on a relatively open subset of a manifold we say that f is a class 0 map if f is continuous. We say that f is a class 1 map if f is differentiable at every point of its domain and the derivative is a continuous function from $\operatorname{dom} f$ into $\mathrm{Lin}(V, W)$; we put $\mathrm{D}^1 f := \mathrm{D} f$. Proceeding inductively, we say that f is a class $s \geq 2$ map if $\operatorname{dom} f$ is a class s manifold and f is class $s - 1$, $\mathrm{D}^{s-1} f$ is differentiable at every point of $\operatorname{dom} f$ and the derivative $\mathrm{D}^s f = \mathrm{D}(\mathrm{D}^{s-1} f)$ is continuous.

If $\mathbb{Q} : \mathcal{S} \to \mathrm{Lin}(\mathrm{R}^n, V)$ is a class 1 map on a surface $\mathcal{S}$ of dimension k, we define the surface divergence $\mathbb{div}\, \mathbb{Q} : \mathcal{S} \to V$ by

$$\mathbb{a} \cdot \mathbb{div}\, \mathbb{Q} = \operatorname{tr}\big(\mathbb{\nabla}(\mathbb{Q}^{\mathrm{T}} \mathbb{a})\big)$$

for each $\mathbb{a} \in V$ where the transpose is defined by $\mathbf{b} \cdot \mathbb{Q}^{\mathrm{T}} \mathbb{a} = \mathbb{Q}\mathbf{b} \cdot \mathbb{a}$ for each $\mathbf{b} \in \mathrm{R}^n$ and $\mathbb{a} \in V$. We say that $\mathbb{Q}$ is superficial if $\mathbb{Q} = \mathbb{Q}\mathbb{P}$.

We define the relative boundary of a surface $\mathcal{S} \subset \mathrm{R}^n$ by $\operatorname{bd} \mathcal{S} = \operatorname{cl} \mathcal{S} \sim \mathcal{S}$. The surface $\mathcal{S}$ is said to be a k dimensional oriented surface with Lipschitz boundary provided the following two conditions hold. (a) The closure of $\mathcal{S}$ is a subset of some k dimensional oriented surface $\mathcal{U}$ and (b) for each point $\mathbf{x}$ of $\operatorname{bd} \mathcal{S}$ there is an orientation preserving diffeomorphism from an open set in R^k onto some (relative) neighborhood N of $\mathbf{x}$ in $\mathcal{U}$ whose inverse maps $N \cap \mathcal{S}$ into some Lipschitz region in

R^k. There exists an $\mathcal{H}^{k-1}$ almost unique vectorfield $\mathbb{m} : \operatorname{bd} \mathcal{S} \to \mathrm{S}^{n-1}$ such that for each diffeomorphism of the type just described we have

$$\mathbb{m} \circ \boldsymbol{\phi} = \nabla \boldsymbol{\phi}^{-\mathrm{T}} \mathbf{v} / |\nabla \boldsymbol{\phi}^{-\mathrm{T}} \mathbf{v}|$$

where $\mathbf{v} : \operatorname{bd} \boldsymbol{\phi}^{-1}(\mathcal{S}) \to \mathrm{S}^{k-1}$ is the Lipschitz normal to the Lipschitz region $\boldsymbol{\phi}^{-1}(\mathcal{S})$ and $\nabla \boldsymbol{\phi}^{-1}$ is the pseudoinverse of the injective map $\nabla \boldsymbol{\phi}$; one has $\mathbb{m}(\mathbf{x}) \in \operatorname{Tan}(\mathcal{U}, \mathbf{x})$ for $\mathcal{H}^{k-1}$ a.e. $\mathbf{x} \in \operatorname{bd} \mathcal{S}$. We call $\mathbb{m}$ the relative normal of $\mathcal{S}$.

Theorem 7.1 (Surface divergence theorem). *If $\mathcal{S}$ is a bounded oriented surface (of arbitrary dimension) of class 2 with Lipschitz boundary and if $\mathbb{Q} : \mathcal{S} \to \operatorname{Lin}(\mathrm{R}^n, V)$ is a superficial field of class 1 with bounded derivative and $\mathbb{Q}$ can be continuously extended to $\operatorname{cl} \mathcal{S}$ then*

$$\int_{\mathcal{S}} \operatorname{\mathbb{div}} \mathbb{Q} \, d\mathcal{H}^k = \int_{\operatorname{bd} \mathcal{S}} \mathbb{Q} \mathbb{m} \, d\mathcal{H}^{k-1}$$

where on the right hand side $\mathbb{Q}$ denotes the continuous extension just mentioned.

8 Appendix B. Derivatives of the interfacial energy

We discuss alternative equivalent forms of the interfacial stress relations (2.2) and (2.3).

Proposition 8.1. *The set* G *is a class ∞ manifold in $\operatorname{Lin} \times \mathrm{R}^n$ of dimension $(m+1)(n-1)$ with the tangent space*

$$\operatorname{Tan}(\mathrm{G}, (\mathbb{F}, \mathbb{n})) = \{ (\mathbb{G}, \mathbb{m}) \in \operatorname{Lin} \times \mathrm{R}^n : \mathbb{G}\mathbb{n} + \mathbb{F}\mathbb{m} = \mathbf{0},\ \mathbb{m} \cdot \mathbb{n} = 0 \} \tag{8.1}$$

for every $(\mathbb{F}, \mathbb{n}) \in \mathrm{G}$.

Proof Let $\Lambda : \operatorname{Lin} \times \mathrm{R}^n \to \mathrm{R}^m \times \mathrm{R}$ be defined by $\Lambda(\mathbb{F}, \mathbb{n}) = (\mathbb{F}\mathbb{n}, |\mathbb{n}|^2 - 1)$, $(\mathbb{F}, \mathbb{n}) \in \operatorname{Lin} \times \mathrm{R}^n$, so that $\mathrm{G} = \Lambda^{-1}(0)$. The map Λ is class ∞ on $\operatorname{Lin} \times \mathrm{R}^n$; its derivative is evaluated straightforwardly; from its form one finds that $\operatorname{ran} \mathrm{D}\, \Lambda(\mathbb{F}, \mathbb{n}) = \mathrm{R}^m \times \mathrm{R}$ whenever $\mathbb{n} \neq \mathbf{0}$. Thus G is a submanifold of $\operatorname{Lin} \times \mathrm{R}^n$ of class ∞ of dimension $\dim(\operatorname{Lin} \times \mathrm{R}^n) - \dim(\mathrm{R}^m \times \mathrm{R}) = (m+1)(n-1)$ and

$$\operatorname{Tan}(\mathrm{G}, (\mathbb{F}, \mathbb{n})) = \ker \mathrm{D}\, \Lambda(\mathbb{F}, \mathbb{n})$$

which gives (8.1). □

Proposition 8.2. *The set* H *is a class ∞ manifold in $\operatorname{Lin} \times \mathrm{R}^n$ of dimension $(m+1)n-1$ with the tangent space*

$$\operatorname{Tan}(\mathrm{H}, (\mathbf{F}, \mathbb{n})) = \{ (\boldsymbol{G}, \mathbb{m}) \in \mathrm{H} : \mathbb{m} \cdot \mathbb{n} = 0 \} \tag{8.2}$$

for every $(\mathbf{F}, \mathbb{n}) \in \mathrm{H}$. If $\tilde{\mathbb{f}} : \mathrm{H} \to \mathrm{R}$ is defined by (2.5) then $\mathbb{n} \cdot \mathrm{D}_2 \tilde{\mathbb{f}} = 0$ and

$$\mathrm{D}_1 \tilde{\mathbb{f}} = \mathrm{D}_1 \hat{\mathbb{f}}\, \mathbb{P},$$
$$\mathrm{D}_2 \tilde{\mathbb{f}} = \mathrm{D}_2 \hat{\mathbb{f}} - \mathbb{P}\mathbf{F}^{\mathrm{T}} \mathrm{D}_1 \hat{\mathbb{f}}\, \mathbb{n} - \mathbb{P}\, \mathrm{D}_1 \hat{\mathbb{f}}^{\mathrm{T}} \mathbf{F}\, \mathbb{n}$$

where the derivatives of $\tilde{\mathbb{f}}$ are evaluated at $(\mathbf{F}, \mathbb{n}) \in \mathrm{H}$ and the derivatives of $\hat{\mathbb{f}}$ at $(\mathbb{F}, \mathbb{n}) \in \mathrm{G}$ related to $(\mathbf{F}, \mathbb{n})$ by $\mathbb{F} = \mathbf{F}\mathbb{P}$; Equations (2.6) and (2.7) hold.

Proof The manifold character of H and its tangents space are immediate. The asserted relations follow by differentiating (2.5) respecting the fact that the derivative must be an element of the tangent space. □

Proposition 8.3. *If $m = n$ then the set* Lin_{n-1} *is a class* ∞ *manifold in* Lin *of dimension* $n^2 - 1$ *with a continuous unit normal to* Lin_{n-1} *at* $\mathbb{F} \in \mathrm{Lin}_{n-1}$ *given by*

$$\mathbb{N}(\mathbb{F}) = |\operatorname{cof} \mathbb{F}|^{-1} \operatorname{cof} \mathbb{F}. \tag{8.3}$$

Each point of Lin_{n-1} *has a neighborhood* N *and a map* $\mathbb{m} : N \to \mathrm{S}^{n-1}$ *such that* $\mathbb{F}\mathbb{m}_\circ(\mathbb{F}) = \mathbf{0}$ *for each* $\mathbb{F} \in N$; *the map* $\Theta : N \to \mathrm{G}$ *given by* $\Theta(\mathbb{F}) = (\mathbb{F}, \mathbb{m}_\circ(\mathbb{F}))$ *for each* $\mathbb{F} \in N$ *is a class* ∞ *diffeomorphism from* N *onto some relatively open subset of* G. *The map* $\mathbb{m}_\circ$ *is locally uniquely determined to within the change of sign; the derivatives of* Θ *and* $\mathbb{m}_\circ$ *at* $\mathbb{F} \in N$ *are given by*

$$\mathrm{D}\Theta\,\mathbf{A} = \big(\mathbf{A} - (\mathbf{A}\cdot\mathbb{N})\mathbb{N}, -\mathbb{F}^{-1}\mathbf{A}\mathbb{m}\big), \tag{8.4}$$

$$\mathrm{D}\mathbb{m}_\circ\mathbf{A} = -\mathbb{F}^{-1}\mathbf{A}\,\mathbb{m}_\circ \tag{8.5}$$

for each $\mathbf{A} \in \mathrm{Lin}$. *If* $\mathbb{f}_\circ$ *is a function on a relatively open subset of* Lin_{n-1} *such that* (2.8) *holds then*

$$\mathrm{D}\,\mathbb{f}_\circ = \mathrm{D}_1\hat{\mathbb{f}} - \mathbb{F}^{-\mathrm{T}}\mathrm{D}_2\hat{\mathbb{f}} \otimes \mathbb{m} - (\mathrm{D}_1\hat{\mathbb{f}}\cdot\mathbb{N})\mathbb{N} \tag{8.6}$$

where the derivative of $\mathbb{f}_\circ$ *and* $\mathbb{N}$ *are evaluated at* $\mathbb{F}$ *and the derivatives of* $\hat{\mathbb{f}}$ *at* $(\mathbb{F}, \mathbb{m})$. *Equations* (2.9) *and* (2.10) *hold.*

Proof If $g : U \to \mathrm{R}$ is given by $U = \{\mathbf{F} \in \mathrm{Lin} : \operatorname{cof}\mathbf{F} \neq \mathbf{0}\}$ and $g(\mathbf{F}) = \det\mathbf{F}$ for each $\mathbf{F} \in U$ then f is class ∞ and $\mathrm{D}g(\mathbf{F}) = \operatorname{cof}\mathbf{F} \neq \mathbf{0}$ for every point of the domain U. Hence $\mathrm{Lin}_{n-1} \equiv g^{-1}(0)$ is a class ∞ manifold of the indicated dimension and $\mathbb{N}$ is the unit normal to Lin_{n-1}. The existence of the map $\mathbb{m}_\circ$ follows by solving $\mathbb{F}\mathbb{m} = \mathbf{0}$ by the implicit function theorem. Differentiating the relation $\mathbb{F}\mathbb{m}_\circ(\mathbb{F}) = \mathbf{0}$ and using the requirement that the derivative satisfies $\mathrm{D}\mathbb{m}_\circ\mathbb{N} = \mathbf{0}$ as a consequence of (7.1) in the present case one obtains (8.5). Equation (8.4) is then a consequence of (7.2) and (8.5) if one observes that $\mathrm{id} - \mathbb{N}\otimes\mathbb{N}$ is the projection onto the tangent space of Lin_{n-1}. To prove the assertions about $\mathbb{f}_\circ$, we note that $\mathbb{f}_\circ = \hat{\mathbb{f}} \circ \Theta$. The derivative $\mathrm{D}\,\mathbb{f}_\circ$ of $\mathbb{f}_\circ$ at $\mathbb{F}$ is an element of the tangent space to Lin_{n-1} at $\mathbb{F}$ which means by (8.3) that

$$\mathrm{D}\,\mathbb{f}_\circ \cdot \operatorname{cof}\mathbb{F} = 0.$$

Evaluating the derivative by the chain rule $\mathrm{D}\,\mathbb{f}_\circ = \mathrm{D}\hat{\mathbb{f}}\,\mathrm{D}\Theta$ and combining with (8.4) one obtains (8.6). Equations (2.9) and (2.10) are proved by noting that $\mathrm{D}\,\mathbb{f}_\circ\mathbb{P} = \mathrm{D}_1\hat{\mathbb{f}}\,\mathbb{P}$ and that $\mathbb{F}^{\mathrm{T}}\mathrm{D}\,\mathbb{f}_\circ\mathbb{m} = \mathbb{F}^{\mathrm{T}}\mathrm{D}_1\hat{\mathbb{f}}\,\mathbb{m} - \mathrm{D}_2\hat{\mathbb{f}}$ by using $\mathbb{F}^{\mathrm{T}}\operatorname{cof}\mathbb{F} = \mathbf{0}$ since $\det\mathbb{F} = 0$. □

9 Appendix C. Multilinear algebra

We use the conventions from [9; Chapter One]; see also [21; Section 1.7].

If r is an integer with $0 \le r \le n$, we denote by $\wedge_r\mathrm{R}^n$ the inner product space of all r vectors in R^n, i.e., the set of all r linear completely antisymmetric maps $\boldsymbol{\alpha}$ from the dual space of R^n into R. We put $\wedge_0\mathrm{R}^n = \mathrm{R}$, note that $\wedge_1\mathrm{R}^n$ is canonically isomorphic with R^n and recall that $\wedge_n\mathrm{R}^n$ is unidimensional. We also put $\wedge_r\mathrm{R}^n = \{\mathbf{0}\}$ if r is an integer with $r < 0$ or $r > n$. We denote by $\boldsymbol{\alpha}\wedge\boldsymbol{\beta}$ the wedge product of an r vector $\boldsymbol{\alpha}$ and an s vector $\boldsymbol{\beta}$; if $\mathbf{a}_1, \dots, \mathbf{a}_r$ are vectors in R^n we denote by $\mathbf{a}_1 \wedge \cdots \wedge \mathbf{a}_r \equiv \prod_{i=1}^r \mathbf{a}_i$ the wedge product of $\mathbf{a}_1, \dots, \mathbf{a}_r$, an element of $\wedge_r\mathrm{R}^n$. If $\boldsymbol{\alpha}$ is an r vector in R^n and $\boldsymbol{\beta}$

and s vector with $s \leq r$ we define a contraction $\boldsymbol{\alpha} \llcorner \boldsymbol{\beta}$ of $\boldsymbol{\alpha}$ by $\boldsymbol{\beta}$ to be an $r-s$ vector in R^n such that

$$(\boldsymbol{\alpha} \llcorner \boldsymbol{\beta}) \cdot \boldsymbol{\gamma} = \boldsymbol{\alpha} \cdot (\boldsymbol{\gamma} \wedge \boldsymbol{\beta})$$

for each $r-s$ vector $\boldsymbol{\gamma}$.

If $\mathbf{A} \in \mathrm{Lin}(\mathrm{R}^n, \mathrm{R}^m)$ where m is a positive integer, we denote by $\wedge_r \mathbf{A}$ the r th exterior power of $\mathbf{A}$, a unique element of $\mathrm{Lin}(\wedge_r \mathrm{R}^n, \wedge_r \mathrm{R}^m)$ having the property

$$\wedge_r \mathbf{A}(\prod_{i=1}^{r} \mathbf{a}_i) = \prod_{i=1}^{r} (\mathbf{A}\, \mathbf{a}_i)$$

for each $\mathbf{a}_1, \ldots, \mathbf{a}_r \in \mathrm{R}^n$. The matrix elements of $\wedge_r \mathbf{A}$ relative to the bases in $\wedge_r \mathrm{R}^n$ and $\wedge_r \mathrm{R}^m$ consisting of the r fold wedge products of the standard basis vectors in R^n and R^m are minors of order r of the matrix of $\mathbf{A}$. In particular if $m = n$ then $\wedge_n \mathbf{A}\, \boldsymbol{\beta} = \det \mathbf{A}\, \boldsymbol{\beta}$ for each n vector $\boldsymbol{\beta}$; in the same situation, $\wedge_{n-1} \mathbf{A} = * \operatorname{cof} \mathbf{A} *$ where $*$ is the Hodge operator mapping $\wedge_r \mathrm{R}^n$ isometrically onto $\wedge_{n-r} \mathrm{R}^n$. We put $\wedge_0 \mathbf{A} = 1$ in all situations. Clearly, $\wedge_r \mathbf{A} = \mathbf{0}$ if $r > \min\{m, n\}$.

If $\mathbf{A} \in \mathrm{Lin}(\mathrm{R}^n, \mathrm{R}^m)$ and $\mathbf{a} \in \mathrm{R}^n$ we define $\wedge_r \mathbf{A} \wedge \mathbf{a} \in \mathrm{Lin}(\wedge_{r+1} \mathrm{R}^n, \wedge_r \mathrm{R}^m)$ by

$$(\wedge_r \mathbf{A} \wedge \mathbf{a})\boldsymbol{\beta} = \wedge_r \mathbf{A}(\boldsymbol{\beta} \llcorner \mathbf{a})$$

for each r vector $\boldsymbol{\beta}$ in R^n. We have $\wedge_0 \mathbf{A}\mathbf{a} = \mathbf{a}$ and, if $m = n$, $\wedge_{n-1} \mathbf{A} \wedge \mathbf{a} = * \operatorname{cof} \mathbf{A}\, \mathbf{a}$. If $m = n = 3$ we can identify $\wedge_1 \mathbf{A} \wedge \mathbf{a} = \mathbf{A} \wedge \mathbf{a} \in \mathrm{Lin}(\wedge_2 \mathrm{R}^3, \mathrm{R}^3)$ with a polar second order tensor $\mathbf{A} \times \mathbf{a}$.

If $\Omega \subset \mathrm{R}^n$ is open, we denote by $\mathcal{D}_r(\Omega)$ the set of all infinitely differentiable r vectorfields $\boldsymbol{\xi} : \mathrm{R}^n \to \wedge_r \mathrm{R}^n$ whose support is compact and contained in Ω. We define the ***interior derivative*** $\partial \boldsymbol{\xi}$ of $\boldsymbol{\xi}$ as an element of $\mathcal{D}_{r-1}(\Omega)$ given by

$$\partial \boldsymbol{\xi} = (-1)^r \sum_{i=1}^{n} \mathrm{D}_i \, \boldsymbol{\xi} \llcorner \boldsymbol{e}^i,$$

where D_i denote the partial derivatives and $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ is the standard basis in R^n. If $\mathbf{a}$ is a 1 vectorfield then $\partial \mathbf{a} = -\operatorname{div} \mathbf{a}$. The factor $(-1)^r$ is chosen so as to render valid the integration by parts formula

$$\int_{\mathrm{R}^n} \partial \boldsymbol{\xi} \cdot \boldsymbol{\omega} \, d\mathcal{L}^n = \int_{\mathrm{R}^n} \boldsymbol{\xi} \cdot \mathrm{D} \boldsymbol{\omega} \, d\mathcal{L}^n$$

for every smooth $r-1$ form $\boldsymbol{\omega}$ on R^n where $\mathrm{D}\,\boldsymbol{\omega}$ denotes the exterior derivative.

References

1 Ambrosio, L.; Fusco, N.; Pallara, D.: *Functions of bounded variation and free discontinuity problems* Oxford, Clarendon Press (2000)

2 Ball, J. M.: *Constitutive inequalities and existence theorems in nonlinear elastostatics* In *Nonlinear Analysis and Mechanics, Heriot-Watt Symposium* R.J. Knops (ed.), Pitman, London 1977

3 Ball, J. M.: *Convexity conditions and existence theorems in nonlinear elasticity* Arch. Rational Mech. Anal. **63** (1977) 337–403

4 Ball, J. M.; Murat, F.: *$W^{1,p}$-quasiconvexity and variational problems for multiple integrals* J. Funct. Anal. **58** (1984) 225–253

5 Chadwick, P.: *Applications of an energy-momentum tensor in non-linear elastostatics* J. Elasticity **5** (1975) 249–258

6 Dacorogna, B.: *Direct methods in the calculus of variations. Second Edition* Berlin, Springer (2008)

7 Eshelby, J. D.: *The force on an elastic singularity* Phil. Trans. Royal Soc. London **A244** (1951) 87–112

8 Eshelby, J. D.: *Continuum theory of defects* In *Solid State Physics, vol. 3* D. Turnbull (ed.), pp. 79–144, Academic Press, New York 1956

9 Federer, H.: *Geometric measure theory* New York, Springer (1969)

10 Fonseca, I.: *Interfacial energy and the Maxwell rule* Arch. Rational Mech. Anal. **106** (1989) 63–95

11 Giaquinta, M.; Hildebrandt, S.: *Calculus of variations I* Berlin, Springer (1996)

12 Giaquinta, M.; Modica, G.; Souček, J.: *Cartesian currents, weak diffeomorphisms and existence theorems in non-linear elasticity* Arch. Rational Mech. Anal. **106** (1989) 97–160

13 Giaquinta, M.; Modica, G.; Souček, J.: *Cartesian Currents in the Calculus of Variations I, II* Berlin, Springer (1998)

14 Gurtin, M. E.: *On a theory of phase transitions with interfacial energy* Arch. Rational Mech. Anal. **87** (1985) 187–212

15 Gurtin, M. E.: *On the two-phase Stefan problem with interfacial energy and entropy* Arch. Rational Mech. Anal. **96** (1986) 199–241

16 Gurtin, M. E.: *The nature of configurational forces* Arch. Rational Mech. Anal. **131** (1995) 67–100

17 Gurtin, M. E.: *Configurational forces as basic concepts of continuum physics* New York, Springer (2000)

18 Gurtin, M. E.; Murdoch, I.: *A continuum theory of elastic material surfaces* Arch. Rational Mech. Anal. **57** (1975) 291–323

19 Gurtin, M. E.; Struthers, A.: *Multiphase thermomechanics with interfacial structure, 3. Evolving phase boundaries in the presence of bulk deformation* Arch. Rational Mech. Anal. **112** (1990) 97–160

20 Leo, P. H.; Sekerka, R. F.: *The effect of surface stress on crystal-melt and crystal-crystal equilibrium* Acta Metall. **37** (1989) 3119–3138

21 Marsden, J. E.; Hughes, T. J. R.: *Mathematical foundations of elasticity* Englewood Cliffs, Prentice-Hall (1983)

22 Maugin, G.: *Material inhomogeneities in elasticity* London, Chapman & Hall (1993)

23 Morrey, Jr, C. B.: *Quasi-convexity and the lower semicontinuity of multiple integrals* Pacific J. Math. **2** (1952) 25–53

24 Morrey, Jr, C. B.: *Multiple integrals in the calculus of variations* New York, Springer (1966)

25 Müller, S.: *Variational Models for Microstructure and Phase Transitions* In *Calculus of variations and geometric evolution problems (Cetraro, 1996) Lecture notes in Math. 1713* pp. 85–210, Springer, Berlin 1999

26 Müller, S., Tang, Q.; Yan, B. S.: *On a new class of elastic deformations not allowing for cavitation* Ann. Inst. H. Poincaré, Analyse non linéaire **11** (1994) 217–243

27 Parry, G. P.: *On shear bands in unloaded crystals* J. Mech. Phys. Solids **35** (1987) 367–382

28 Podio-Guidugli, P.: *Configurational balances via variational arguments* Interfaces and Free Boundaries **3** (2001) 223–232

29 Podio-Guidugli, P.: *Configurational forces: are they needed?* Mechanics Research Communications **29** (2002) 513—519

30 Šilhavý, M.: *The mechanics and thermodynamics of continuous media* Berlin, Springer (1997)

31 Šilhavý, M.: *Phase transitions with interfacial energy: convexity conditions and the existence of minimizers* (2008) In preparation.

32 Steinmann, P.: *On boundary potential energies in deformational and configurational mechanics* Journal of the Mechanics and Physics of Solids **56** (2008) 772–800