Linear Γ-limits of multiwell energies in nonlinear elasticity theory*

Bernd Schmidt
Zentrum Mathematik, Technische Universität München
Boltzmannstr. 3, 85747 Garching, Germany
schmidt@ma.tum.de

Abstract

We derive linearized theories from nonlinear elasticity theory for multiwell energies. Under natural assumptions on the nonlinear stored energy densities, the properly rescaled nonlinear energy functionals are shown to Γ-converge to the relaxation of a corresponding linearized model. Minimizing sequences of problems with displacement boundary conditions and body forces are investigated and found to correspond to minimizing sequences of the linearized problems. As applications of our results we discuss the validity and failure of a formula that is widely used to model multiwell energies in the regime of linear elasticity. Applying our convergence results to the special case of single well densities, we also obtain a new strong convergence result for the sequence of minimizers of the nonlinear problem.

1 Introduction and overview

Consider an elastic body occupying a reference configuration $\Omega \subset \mathbb{R}^d$, $d \in \mathbb{N}$, subject to some deformation $y : \Omega \to \mathbb{R}^d$. Assuming the body is hyperelastic, the stored energy of such a deformation can be written in terms of a stored energy function W:

$$\text{energy of } y = \int_\Omega W(x, \nabla y(x))\, dx. \tag{1}$$

If $y(x) = x + \varepsilon u(x)$ is given in terms of a small displacement εu and Ω is such that the energy of y is minimized for $y(x) = x$, then Taylor-expanding (1) and rescaling by ε^{-2} we find the energy formula of linear elasticity

$$\text{linearized energy of } u = \frac{1}{2} \int_\Omega \nabla_2^2 W(x, \mathrm{Id})(\nabla u(x))\, dx$$

in the limit $\varepsilon \to 0$. Frame indifference of W implies that the quadratic form $\nabla_2^2 W(x, \mathrm{Id})$ in fact only depends on the linear strain $e(u) = \frac{1}{2}((\nabla u)^T + \nabla u)$ rather than the full gradient ∇u.

Although standard, this relation between nonlinear (finite) elasticity theory and its linear (infinitesimal) counterpart has been given a precise meaning only recently by Dal Maso, Negri and Percivale (cf. [11]). Assuming in particular that W is minimized precisely at $SO(d)$, they prove that the functional of

*Preprint. The original publication appeared in Continuum Mech. Thermodyn. **20** (2008) pp. 375–396, DOI 10.1007/s00161-008-0087-8

linear elasticity arises as a Γ-limit of the properly rescaled nonlinear functionals and that minimizers of boundary value problems converge to minimizers of the linearized problem.

In this paper we will more generally allow for a family W_ε of stored energy densities with multiple energy wells, i.e., whose set of minimizers is of the form

$$SO(d)U_1(\varepsilon) \cup \ldots \cup SO(d)U_N(\varepsilon).$$

(For the sake of simplicity we will concentrate on homogeneous energy densities, i.e., W_ε not explicitly depending on x, with wells of equal height. However, we will indicate all the necessary steps to treat the more general case.) Those energies are important when modeling materials with different 'variants', i.e., preferred strains represented by the wells $SO(d)U_i(\varepsilon)$, that occur, e.g., in the martensitic phase of shape memory alloys (see, e.g., [6] for more details). Note that in this case the stored energy is not only non-convex, but not even quasiconvex in general. Consequently such materials tend to build microstructures in order to assume energetically favorable configurations. We wish to understand the limiting behavior of (minimizers of) the properly rescaled energy functionals

$$\varepsilon^{-2} \int_\Omega W_\varepsilon(\mathrm{Id} + \varepsilon \nabla u).$$

If $U_i(\varepsilon) = \mathrm{Id} + \varepsilon U_i$, $U_i^T = U_i$, and $\nabla^2 W_\varepsilon(U_i(\varepsilon)) \to a_i$, a formal linearization at the energy wells leads to the commonly used energy functional of the form

$$\frac{1}{2} \int_\Omega \min_i \langle a_i(e(u) - U_i), (e(u) - U_i) \rangle . \tag{2}$$

This 'well minima formula' is usually referred to as the energy functional of the KRS-theory, named after Khachaturyan, Roitburd and Shatalov (cf. [15, 16, 21, 22, 17]). Note that the integrand is non-quadratic for more than one well, so the limiting theory actually is only geometrically linear. Kohn shows in [18] that (2) formally arises from nonlinear elasticity by Taylor-expansion and proper rescaling of an energy density of the form

$$W_\varepsilon(F) = \min_i W_\varepsilon^{(i)}(F), \tag{3}$$

$W_\varepsilon^{(i)}$ minimized at $\mathrm{Id} + \varepsilon U_i$, where it has constant Hessian a_i. In general – as noted by Ball and James [5] – a functional as in (2) is expected to describe deformations that are close not only to the identity mapping but also to the different energy wells. Bhattacharya compares the nonlinear and linear theories in [7]. We also refer to [7] and the references therein and to the more recent article [3] by Ball for a detailed account of the material science and mathematical literature for multiwell energies.

We will adopt the point of view that has proved very successful recently in a variety of problems in elasticity theory involving the derivation of effective theories, which consists of an ansatz free study by Γ-convergence of various scalings of the energy functional. This is particularly suited for multiwell problems where minimal energies may not be attained and fine mixtures of phases have to be considered instead (see [4]). Assuming quadratic growth of the energy densities at our energy wells, however, the appropriate energy regime for deformations as considered by Ball and James in [5] would be

$$\int_\Omega W_\varepsilon(\mathrm{Id} + \varepsilon \nabla u) \ll \varepsilon^2.$$

This leads to compatibility restrictions on the resulting linear displacement gradients as we will show in Corollary 2.8.

In order to obtain a linearized theory without these restrictions (i.e., which corresponds to the classical energy scaling of linear elasticity), we will consider energies that scale quadratically in ε. In this regime, however, formula (2) can be justified only under additional assumptions on the limiting behavior of the nonlinear energy densities W_ε.

Let us briefly comment on the choice of the various small parameters that enter the energy functional. In order that the system be described by a (geometrically) linear model, the energy wells have to be sufficiently close to each other. Otherwise a linearized displacement field cannot properly connect strains belonging to different wells. This introduces a small parameter ε, in terms of which the typical distance between energy wells is measured. Now the physically interesting regime is when the displacements scale with the same parameter ε. This is precisely the regime where the limiting linear strain 'feels' the effect of the different energy wells. (If $\nabla y - \mathrm{Id}$ converges to 0 more slowly than ε, then different wells could not be resolved in the linearized regime anymore: in (2), e.g., this would lead to $U_1 = \ldots = U_N = \mathrm{Id}$. Also from the point of view of the applications, a linearized theory is not expected to apply on scales much larger than the distance between the energy wells. If, on the other hand, $|\nabla y - \mathrm{Id}|$ is much smaller than ε, then one would effectively only try to linearize at one particular well. Moreover, mathematically this would immediately lead to a loss of compactness, spoiling the convergence of the variational minimum problems.) The third small parameter is given by the energy scaling of the problem, which also determines the scaling of the loading term. Assuming quadratic growth of the energy density away from the wells, the appropriate scaling is given by ε^2. In fact, due to the multiwell structure of the energy densities, also energies of order $\ll \varepsilon^2$ appear to be interesting. As mentioned above, we briefly discuss these scalings in Corollary 2.8.

More generally than in (1) we will consider the energy functionals

$$\int_\Omega W_\varepsilon(\mathrm{Id} + \varepsilon \nabla u) + \int_\Omega \varepsilon^2 l u, \quad u = g \text{ on } \partial\Omega_*,$$

including some loading term scaling with ε^2, for displacements subject to boundary conditions on the Dirichlet boundary $\partial\Omega_* \subset \partial\Omega$. We state our main results in Section 2. Assuming that the rescaled nonlinear stored energy densities converge to some function V on the symmetric $d \times d$ matrices in a suitable sense, we prove in Theorem 2.1 and Proposition 2.2 that the equicoercive family of rescaled nonlinear energies $\mathcal{E}_\varepsilon$ Γ-converges L^2-strongly and H^1-weakly to the relaxation $\mathcal{E}^{\mathrm{rel}}$ of the linearized functional $\mathcal{E}$ with density V. (See Section A in the appendix for a review of the definitions and some properties of quasiconvexity, quasiconvexity on linear strains, Γ-convergence and relaxation.) More precisely, the Γ-limit yields a geometrically linearized theory in terms of an integral functional of the linear strain. The limiting energy density V, however, will in general not be quadratic, i.e., the limiting energy functional is still nonlinear.

This in particular implies that the minimal energy values of $\mathcal{E}_\varepsilon$ converge to the minimal energy of $\mathcal{E}$. For multiwell energies with $\mathcal{E}^{\text{rel}} \neq \mathcal{E}$, however, a limiting deformation, i.e., a limit of a low energy sequence, will in general only be a weak H^1-limit. In addition it will not be a minimizer of the linear problem $\mathcal{E}$ but only of the relaxed functional $\mathcal{E}^{\text{rel}}$ as in the presence of multiple phases the material is expected to develop microstructure so as to reduce energy. In order to capture this effect we will also study the recovery sequences, i.e., those approximating sequences that have the correct energetic limit, and prove in Theorem 2.3 that they are also recovery sequences for the relaxation of $\mathcal{E}$. A Young measure interpretation is given in Corollary 2.5. As a consequence we obtain a complete picture of the convergence of minimum problems including the convergence of minimizing energy deformations in Theorem 2.4.

We conclude this section with two little corollaries on the convergence of the relaxed energy densities and on the previously mentioned restrictions at lower energy scalings. Furthermore we discuss extensions to more general situations, in particular to wells that are only locally minimizing and to non-homogeneous energy densities.

Before proving all this in Section 4, we will focus on two applications in Section 3. Firstly we will revisit the single well case considered in [11]. Here we have $\mathcal{E}^{\text{rel}} = \mathcal{E}$ and our methods give in fact a stronger convergence result for low energy sequences under less restrictive growth assumptions on the energy density as compared to [11]. Secondly we will investigate in detail two examples to which our converge scheme applies. We reconsider Kohn's well minima example and show that (2) correctly describes the corresponding linearized regime. The second example, however, in spite of being a natural candidate for a nonlinear multiwell energy which – compared to the first example – even has the advantage of being smooth, does not fit into the framework of KRS-theory. We will show that even the relaxation of its linearized limit can be different from the the relaxed KRS-functional.

2 Main results

Suppose that $F \mapsto W_\varepsilon(F)$, $F \in \mathbb{R}^{d\times d}$, is a family of frame indifferent multiwell energies with wells at $SO(d)U_1(\varepsilon), \ldots, SO(d)U_N(\varepsilon)$ for some $U_i(\varepsilon) = U_i^T(\varepsilon) > 0$. Here $U_i(\varepsilon)$ are the local minimizers of W_ε among positive symmetric matrices with energies $W_\varepsilon(U_i(\varepsilon)) = 0$, which by assumption shall satisfy $U_i(\varepsilon) = \text{Id} + \varepsilon U_i + o(\varepsilon)$ for some $U_1, \ldots, U_N$. We also suppose that W_ε is measurable, continuous in an ε-independent neighborhood of Id, and scales quadratically at the energy wells in the direction perpendicular to infinitesimal rotations, more precisely, for $\mathcal{U}_\varepsilon := SO(d)U_1(\varepsilon) \cup \ldots \cup SO(d)U_N(\varepsilon)$ we have

$$W_\varepsilon(F) \geq c \operatorname{dist}^2(F, \mathcal{U}_\varepsilon) \tag{4}$$

for all $F \in \mathbb{R}^{d\times d}$ and some $c > 0$ independent of ε. In addition we will require that the W_ε satisfy the orientation preserving condition

$$W_\varepsilon(F) = \infty \text{ whenever } \det(F) \leq 0. \tag{5}$$

This assumption is physically reasonable and mathematically convenient in the proofs of Section 4 but can in fact be dropped as we will see at the end of that section.

Example. By frame indifference we can express $W_\varepsilon(F) = \tilde{W}_\varepsilon(\frac{1}{2}(F^T F - \mathrm{Id}))$ in terms of the Green–St. Venant tensor $\frac{1}{2}(F^T F - \mathrm{Id})$. An admissible family of energy densities W_ε arises, e.g., by a suitable rescaling of some fixed $\tilde{W}$: Suppose $\tilde{W} \geq 0$ is a continuous function on the symmetric matrices, minimized at A_i with $\tilde{W}(A_i) = 0$, $i = 1, \ldots, N$, and such that $\tilde{W}(F) \geq c|F - A_i|^2$ for some $c > 0$ if F is close to A_i (which holds, e.g., if the second derivative $\nabla^2 \tilde{W}(A_i)$ of $\tilde{W}$ at A_i exists and is positive definite). In addition assume that $\liminf_{|F|\to\infty} |F|^{-1}\tilde{W}(F) > 0$. Set $\tilde{W}_\varepsilon(A) := \varepsilon^2 \tilde{W}(\frac{1}{\varepsilon}A)$ and define W_ε by $W_\varepsilon(F) = \tilde{W}_\varepsilon(\frac{1}{2}(F^T F - \mathrm{Id}))$.

Then W_ε satisfies all our requirements: $U_i(\varepsilon)$ is given by $\varepsilon A_i = \frac{1}{2}(U_i^2(\varepsilon) - \mathrm{Id})$, whence $U_i(\varepsilon) = \mathrm{Id} + \varepsilon A_i + O(\varepsilon^2)$ and so $U_i = A_i$. Also, for $|F| \leq C$,

$$\begin{aligned} W_\varepsilon(F) = \tilde{W}_\varepsilon\left(\frac{1}{2}(F^T F - \mathrm{Id})\right) &\geq c \min_i \varepsilon^2 \left|\frac{1}{2\varepsilon}(F^T F - \mathrm{Id}) - A_i\right|^2 \\ &= \frac{c}{4} \min_i \left|F^T F - U_i^2(\varepsilon)\right|^2 \geq c' \min_i \left|\sqrt{F^T F} - U_i(\varepsilon)\right|^2 \\ &\geq c' \min_i \operatorname{dist}^2(F, SO(d)U_i(\varepsilon)) \end{aligned}$$

for suitable $c, c' > 0$. Here we have made use of the inequality $|\sqrt{G} - \sqrt{H}| \leq C|G - H|$ for all positive $G, H \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$ whenever H lies in a sufficiently small neighborhood of Id, which follows elementary from the growth and local Lipschitz properties of $G \mapsto \sqrt{G}$. The claim now follows from $\tilde{W} > 0$ on $\{|F| \geq C\}$ for suitable C and the growth assumption on $\tilde{W}$ at ∞.

Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain. In order to derive a linearized theory we will in particular investigate minimum problems for the energy functionals

$$\mathcal{E}_\varepsilon(u) := \varepsilon^{-2} \int_\Omega W_\varepsilon(\mathrm{Id} + \varepsilon \nabla u) - \int_\Omega lu, \tag{6}$$

subject to some boundary data $u = g$ on the Dirichlet boundary $\partial\Omega_* \subset \partial\Omega$ and some loading l, in their limit as $\varepsilon \to 0$. Our main aim is to derive a good condition that guarantees convergence of the minimum energy and of (almost) minimizers of $\mathcal{E}_\varepsilon$. (Note that the infimum of $\mathcal{E}_\varepsilon$ might not be attained.) In fact we would like to identify a corresponding linear problem of the form

$$\mathcal{E}(u) := \int_\Omega V(e(u)) - \int_\Omega lu, \quad u = g \text{ on } \partial\Omega_* \tag{7}$$

with $V : \mathbb{R}^{d\times d}_{\rm sym} \to \mathbb{R}$ to be determined from W_ε.

For given W_ε let $V_\varepsilon : \mathbb{R}^{d\times d}_{\rm sym} \to \mathbb{R}$ be the rescaled energy

$$V_\varepsilon(F) := \varepsilon^{-2} W_\varepsilon(\mathrm{Id} + \varepsilon F).$$

We will see that under suitable assumptions on the convergence of V_ε – which are, in particular, satisfied in Kohn's model – a linear limit exists that is naturally linked to a corresponding linear model.

Our main Γ-convergence result is the following. Let $g \in W^{1,\infty}(\Omega, \mathbb{R}^d)$ be some boundary data and $l \in L^2(\Omega, \mathbb{R}^d)$ some loading. For a closed subset $\partial\Omega_*$ of $\partial\Omega$ with $\mathcal{H}^{d-1}(\partial\Omega_*) > 0$ consider the functionals $\mathcal{E}, \mathcal{E}_\varepsilon : H^1_{g,\partial\Omega_*} \to [0, \infty]$ defined in (7) and (6), where $H^1_{g,\partial\Omega_*}$ denotes the closure of $\{v \in W^{1,\infty}(\Omega, \mathbb{R}^d) : v = g \text{ on } \partial\Omega_*\}$ in $H^1(\Omega, \mathbb{R}^d)$.

Theorem 2.1 *Suppose $V_\varepsilon \to V$ uniformly on compacta and V satisfies the growth condition $V(F) \le \alpha(1 + |F|^2)$ for some constant $\alpha \in \mathbb{R}$. Then $\mathcal{E}_\varepsilon$ Γ-converges to the relaxation $\mathcal{E}^{\rm rel} : H^1_{g,\partial\Omega_*} \to [0, \infty)$ of $\mathcal{E}$ given by*

$$\mathcal{E}^{\rm rel}(u) := \int_\Omega Q^e V(e(u)) - \int_\Omega lu$$

with respect to both the strong L^2- and the weak H^1-topology on $H^1_{g,\partial\Omega_*}$.

Here $Q^e V$ denotes the quasiconvexification on linear strains of V. See Section A in the appendix for a definition of quasiconvexity on linear strains and Γ-convergence. Also note that by fame indifference, the convergence assumption on V_ε is of course equivalent to asking that $\varepsilon^{-2} W_\varepsilon(\mathrm{Id} + \varepsilon\cdot)$ converges to $V(\frac{(\cdot)^T + (\cdot)}{2})$ uniformly on compacta of $\mathbb{R}^{d\times d}$.

Theorem 2.1 is complemented by the following compactness result:

Proposition 2.2 *The functionals $\mathcal{E}_\varepsilon$ are equicoercive with respect to the strong L^2-topology and the weak H^1-topology.*

In general, minimizers of $\mathcal{E}_\varepsilon$ and $\mathcal{E}$ do not exist, which leads us to the study of low energy sequences. Also, for multiwell energies V, the minimizers of the relaxed functional $\mathcal{E}^{\rm rel}$ may not be unique. So in terms of the convergence of the displacements u, we can only hope for subsequential convergence. Especially for incompatible wells for which $Q^e V$ is non-convex, we cannot expect strong convergence of low energy sequences as energy minimization requires fine phase mixtures, i.e., the formation of microstructure, in general. In this case, the relevant information on the linearized displacements is encoded in the recovery sequences, i.e., those approximations that give the correct energy in the linear limit. The following theorem shows that recovery sequences are indeed 'relaxing' sequences of the corresponding linear problem.

Theorem 2.3 *Suppose the assumptions of Theorem 2.1 are satisfied. If u_ε is a recovery sequence for u, i.e., $u_\varepsilon \to u$ in L^2 and $\mathcal{E}_\varepsilon(u_\varepsilon) \to \mathcal{E}^{\rm rel}(u)$, then*

$$\lim_{\varepsilon\to 0} \mathcal{E}(u_\varepsilon) = \mathcal{E}^{\rm rel}(u).$$

As a consequence of Theorem 2.1, Proposition 2.2 and Theorem 2.3 we obtain the following convergence result for our nonlinear-to-linear variational limit:

Theorem 2.4 *Suppose the assumptions of Theorem 2.1 are satisfied. Then the minimal energy converges, i.e.,*

$$\lim_{\varepsilon\to 0}\inf_{v\in H^1_{g,\partial\Omega_*}}\mathcal{E}_\varepsilon(v)=\inf_{v\in H^1_{g,\partial\Omega_*}}\mathcal{E}(v)=\min_{v\in H^1_{g,\partial\Omega_*}}\mathcal{E}^{\mathrm{rel}}(v).$$

Moreover, if u_ε *is a low energy sequence of* $\mathcal{E}_\varepsilon$*, i.e.,*

$$\mathcal{E}_\varepsilon(u_\varepsilon)=\inf_{v\in H^1_{g,\partial\Omega_*}}\mathcal{E}_\varepsilon(v)+o(1),$$

then u_ε *is also a minimizing sequence for* $\mathcal{E}$*, i.e.,*

$$\lim_{\varepsilon\to 0}\mathcal{E}_\varepsilon(u_\varepsilon)=\lim_{\varepsilon\to 0}\mathcal{E}(u_\varepsilon)=\inf_{v\in H^1_{g,\partial\Omega_*}}\mathcal{E}(v).$$

Furthermore, there exists a subsequence that converges to some $u\in H^1_{g,\partial\Omega_*}$ *(weakly in* H^1 *and hence strongly in* L^2*), and* u *is a minimizer of* $\mathcal{E}^{\mathrm{rel}}$*.*

Note that this sheds some new light also on the single well case. This will be detailed in Paragraph 3.1.

The formation of microstructure is often quantified in terms of the Young measures induced by the gradients of low energy sequences. (See, e.g., [20] for basic results on (gradient) Young measures.)

Corollary 2.5 *Suppose the assumptions of Theorem 2.1 are satisfied. Let* (ν_x) *be a* ($W^{1,2}$-) *gradient Young measure induced by a recovery sequence* u_ε *for some* u *in* $H^1_{g,\partial\Omega_*}$*. If* (ν^e_x) *denotes the image measure under the transformation* $\mathbb{R}^{d\times d}\to\mathbb{R}^{d\times d}_{\mathrm{sym}}$, $F\mapsto\frac{1}{2}(F^T+F)$*, then*

$$\mathcal{E}^{\mathrm{rel}}(u)=\int_\Omega\int_{\mathbb{R}^{d\times d}_{\mathrm{sym}}}V(F)\,d\nu^e_x(F)\,dx-\int_\Omega lu\,dx.$$

In particular, if u_ε *is a low energy sequence for* $\mathcal{E}_\varepsilon$*, this equals* $\inf_{v\in H^1_{g,\partial\Omega_*}}\mathcal{E}(v)$*.*

So indeed low energy, resp., recovery, sequences induce gradient Young measures that represent the correct microstructure to relax the corresponding linearized problem.

In the simplest setting with affine boundary data and zero loading this implies convergence of the relaxed energy densities.

Corollary 2.6 *Suppose the assumptions of Theorem 2.1 are satisfied. Then for all* $F\in\mathbb{R}^{d\times d}$*,*

$$\lim_{\varepsilon\to 0}\varepsilon^{-2}QW_\varepsilon(\mathrm{Id}+\varepsilon F)=Q^eV\left(\frac{F^T+F}{2}\right).$$

Remark 2.7 Assuming a uniform standard 2-growth assumption for W_ε, it is not hard to modify the proof of Theorem 2.1 to see that $\mathcal{E}_\varepsilon$ Γ-converges to $\mathcal{E}^{\mathrm{rel}}$ whenever $\varepsilon^{-2}QW_\varepsilon(\mathrm{Id}+\varepsilon\cdot) \to Q^eV\left(\frac{(\cdot)^T+(\cdot)}{2}\right)$ pointwise. However, our assumptions in Theorem 2.1 in terms of the convergence properties of V_ε itself will be much easier to check in applications (see Section 3) and do not exclude physically interesting cases in which W_ε equals infinity on some subset of $\mathbb{R}^{d\times d}$ away from $SO(d)$, e.g., for gradients F that do not satisfy $\det(F) > 0$.

The following Corollary 2.8 implies that at energy scales $\ll \varepsilon^2$ the limiting strain is constrained to a proper subset of $\mathbb{R}^{d\times d}_{\mathrm{sym}}$. In particular it shows that a strain $\mathrm{Id}+\varepsilon F$ can only be compatible with the energy wells at $SO(d)U_i(\varepsilon)$, i.e. yield zero energy, in the limit $\varepsilon \to 0$ if the linear strain $\frac{F^T+F}{2}$ is compatible with $\{U_1,\dots,U_N\}$.

Corollary 2.8 *Suppose there is a sequence u_ε with $u_\varepsilon(x) = Fx$ for $x \in \partial\Omega$ such that, for $l = 0$,*

$$\liminf_{\varepsilon\to 0} \mathcal{E}_\varepsilon(u_\varepsilon) = 0.$$

Then $\frac{F^T+F}{2} \in Q^e_p\{U_1,\dots,U_N\} = \mathbf{Q}^e_p\{U_1,\dots,U_N\}$ for all $p \in [1,\infty)$.

Before we prove these results in the next section let us revisit the example discussed at the beginning of this section.

Example. Assume W_ε arises from some fixed function $\tilde{W}$ as described in the example on page 5. It is not hard to see that $V_\varepsilon \to \tilde{W}$ uniformly on compact subsets of $\mathbb{R}^{d\times d}_{\mathrm{sym}}$. So if in addition $\tilde{W}$ is quadratically bounded from above, then the assumptions of Theorem 2.1 are satisfied. In this case our results show that $\tilde{W}$ gives an energy density of the linearized problem.

Remark 2.9 The results of this section can be extended in several ways:

(i) Except for Corollary 2.8, the condition that W_ε be minimal precisely at $U_i(\varepsilon)$ can be relaxed. In fact, one only needs that

$$W_\varepsilon(F) \geq c\,\mathrm{dist}^2(F, SO(d)) - C\varepsilon^2 \tag{8}$$

for all $F \in \mathbb{R}^{d\times d}$ and some $c, C > 0$ independent of ε. In particular, energy wells may have different heights $\varepsilon^2 w_i$ scaling with ε^2.

(ii) More generally, the main results remain true for x-dependent families of energies. Suppose $W_\varepsilon : \Omega \times \mathbb{R}^{d\times d} \to (-\infty,\infty]$ is a frame indifferent family of measurable functions such that (8) with $W_\varepsilon(F)$ replaced by $W_\varepsilon(x,F)$ holds for a.e. x and all F. If $V_\varepsilon \to V$ uniformly on $\Omega \times K$ for all compact subsets K of $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ and V is continuous and satisfies a 2-growth condition $V(x,F) \leq C(1+|F|^2)$, then Theorems 2.1, 2.3, 2.4, Proposition 2.2 and

Corollary 2.5 and, if (4) holds, Corollary 2.8 are true for

$$\mathcal{E}_\varepsilon(u) = \varepsilon^{-2} \int_\Omega W_\varepsilon(x, \mathrm{Id} + \varepsilon \nabla u(x))\, dx - \int_\Omega l(x) u(x)\, dx,$$
$$\mathcal{E}(u) = \int_\Omega V(x, e(u)(x))\, dx - \int_\Omega l(x) u(x)\, dx,$$
$$\mathcal{E}^{\mathrm{rel}}(u) = \int_\Omega Q^e V(x, e(u)(x))\, dx - \int_\Omega l(x) u(x)\, dx,$$

where $Q^e V(x, \cdot)$ is the quasiconvexification on linear strains of $V(x, \cdot)$ for fixed x.

(iii) The orientation preserving assumption (5) is not necessary.

3 Applications

In this section we will discuss applications of our main convergence results. Motivated by our results in the previous paragraphs we will call a family of energy densities W_ε *admissible* if the conditions specified at the beginning of Section 2 are satisfied. If $V_\varepsilon \to V$ uniformly on compact subsets of $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ and V satisfies a 2-growth assumption from above, then we will say that V is their *linear limit.* In this case Theorems 2.1, 2.3, 2.4, Proposition 2.2 and Corollaries 2.5, 2.6, 2.8 apply, so in fact V is an energy density for the geometrically linearized theory.

We first consider the special case of a limiting single well energy. This case has been studied previously by Dal Maso, Negri and Percivale in [11]. However, we obtain a stronger convergence result for the deformations under weaker growth assumptions than in [11].

Next we investigate two examples which are natural models for multiwell energies and to which our general theorems apply. In both cases we obtain a formula for the linearized energy, in particular, we identify the linear limits $V^{(1)}$ and $V^{(2)}$, respectively. Only one of them (the nonlinear energy considered by Kohn in [18]) is in agreement with the well minima formula (2). The second linear limit $V^{(2)}$ thus gives a non-pathological counterexample to (2). We will discuss a two-well example showing that although $V^{(1)}$ and $V^{(2)}$ might agree at the energy wells up to the second order, in general $Q^e V^{(1)} \neq Q^e V^{(2)}$. So in fact also the corresponding relaxed functionals $(\mathcal{E}^{(1)})^{\mathrm{rel}}$ and $(\mathcal{E}^{(2)})^{\mathrm{rel}}$ are different. In particular, the minimal energies of the limiting functional are model-dependent.

3.1 The single well

Suppose $W_\varepsilon = W$ is independent of ε, where $W \geq 0$ is a frame indifferent single well energy such that

$$W(F) \geq c\, \mathrm{dist}^2(F, SO(d)) \text{ for all } F \in \mathbb{R}^{d\times d}, \quad W(F) = 0 \text{ iff } F \in SO(d) \qquad (9)$$

and such that W is C^2 in a neighborhood of $SO(d)$. Then clearly $V_\varepsilon = \varepsilon^{-2} W_\varepsilon(\mathrm{Id} + \varepsilon \cdot)$ converges to $V = \nabla^2 W(\mathrm{Id})$ uniformly on compacta. Note that by our assumptions the quadratic form $\nabla^2 W(\mathrm{Id})$ vanishes on antisymmetric matrices and is positive definite on $\mathbb{R}^{d\times d}_{\mathrm{sym}}$. In particular, V is uniformly strictly convex on $\mathbb{R}^{d\times d}_{\mathrm{sym}}$. But then $Q^e V = V$ and hence, by strict convexity on linear strains and Korn's inequality, it follows that the minimizer of the limiting functional $\mathcal{E} = \mathcal{E}^{\mathrm{rel}}$ is unique (cf. Kirchoff's uniqueness theorem).

More generally assume that W_ε is an admissible family of energy densities with linear limit V. In addition suppose that $V : \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}$ is uniformly strictly quasiconvex on linear strains, i.e., there exists $\gamma > 0$ such that

$$\int_\Omega V(F + e(v)) \geq \int_\Omega (V(F) + \gamma |e(v)|^2) \quad \forall F \in \mathbb{R}^{d\times d}_{\mathrm{sym}},\ v \in C_c^\infty(\Omega, \mathbb{R}^d). \tag{10}$$

In addition to the results of Section 2 we have the following

Theorem 3.1 *Suppose the assumptions of Theorem 2.1 are satisfied and V satisfies* (10)*.*

(i) Let u_ε be a low energy sequence of $\mathcal{E}_\varepsilon$. Then – up to subsequences – u_ε converges strongly in H^1 to a minimizer u of $\mathcal{E} = \mathcal{E}^{\mathrm{rel}}$. So if the minimizer u is unique, the whole sequence u_ε converges strongly.

(ii) If u_ε is a recovery sequence for $u \in H^1_{g,\partial\Omega_*}$, then $u_\varepsilon \to u$ strongly in H^1.

Remark 3.2 As noted above, uniqueness of the minimizer of $\mathcal{E}$ follows, e.g., if V is strictly convex. Another instance is given when $g = F\cdot$ is affine on $\partial\Omega_* = \partial\Omega$ and $l = 0$. Since by density (10) holds for all $v \in H^1_0$, Korn's inequality immediately implies that the unique minimizer is the affine function $x \mapsto Fx$ itself.

The proof is a typical convexity argument (compare Section 3.C in [12]).

Proof. By quasiconvexity of V we have $\mathcal{E} = \mathcal{E}^{\mathrm{rel}}$. If u_ε is a low energy sequence, then, by Theorem 2.4, $u_\varepsilon \rightharpoonup u$ in H^1 for a subsequence where u is a minimizer of $\mathcal{E}$ and $\mathcal{E}(u_\varepsilon) \to \mathcal{E}(u)$. On the other hand, a recovery sequence for an element $u \in H^1_{g,\partial\Omega_*}$ has bounded energy and satisfies $\mathcal{E}(u_\varepsilon) \to \mathcal{E}(u)$ by Theorem 2.3. So in order to prove (i) and (ii), it suffices to show that $u_\varepsilon \rightharpoonup u$ in H^1 and $\int_\Omega V(e(u_\varepsilon)) \to \int_\Omega V(e(u))$ imply $u_\varepsilon \to u$ in H^1.

If $0 \leq \tilde\gamma \leq \gamma$, then $\tilde V$ with $\tilde V(F) = V(F) - \tilde\gamma |F|^2$ is quasiconvex on linear strains. Choosing $\tilde\gamma > 0$ so small that $\tilde V$ satisfies a growth condition $-C \leq \tilde V(F) \leq C(1 + |F|^2)$, we deduce from lower semicontinuity

$$\begin{aligned}\int_\Omega V(e(u)) - \tilde\gamma |e(u)|^2 &= \int_\Omega \tilde V(e(u)) \leq \liminf_{\varepsilon\to 0} \int_\Omega \tilde V(e(u_\varepsilon)) \\ &= \lim_{\varepsilon\to 0} \int_\Omega V(e(u_\varepsilon)) - \tilde\gamma \limsup_{\varepsilon\to 0} \int_\Omega |e(u_\varepsilon)|^2.\end{aligned}$$

So $\limsup_{\varepsilon\to 0} \|e(u_\varepsilon)\|^2_{L^2} \leq \|e(u)\|^2_{L^2}$ and therefore weak convergence $e(u_\varepsilon) \rightharpoonup e(u)$ in L^2 improves to strong convergence $e(u_\varepsilon) \to e(u)$ in L^2. By Korn's and Poincaré's inequalities we obtain $u_\varepsilon \to u$ in H^1. □

The adaption to x-dependent W_ε is straightforward if the constants c in (9) and γ in (10) can be chosen independently of x.

3.2 Validity and failure of the well minima formula

If W_ε is given as the minimum over quadratic single well energies as in (3) (also cf. [18]), then the results of Section 2 lead in fact to a justification of the linear functional in the well minima formula (2). Let $W_{i,\varepsilon}$, $i = 1, \dots, N$, be admissible single well energies minimized, respectively, at $SO(d)U_i(\varepsilon)$ with $\nabla^2 W_{i,\varepsilon}(U_i(\varepsilon)) = a_i$ on symmetric matrices and $W_{i,\varepsilon}(U_i(\varepsilon)) = \varepsilon^2 w_i$ (see Remark 2.9). So

$$\begin{aligned} W_{i,\varepsilon}(F) = \frac{1}{2} \left\langle a_i \left(\sqrt{F^T F} - U_i(\varepsilon) \right), \sqrt{F^T F} - U_i(\varepsilon) \right\rangle + \varepsilon^2 w_i \\ + o\left(\left| \sqrt{F^T F} - U_i(\varepsilon) \right|^2 \right) \end{aligned}$$

and a_i is positive definite on $\mathbb{R}^{d\times d}_{\mathrm{sym}}$. Assume that the last summand on the right-hand side comprising the higher order contributions to $W_{i,\varepsilon}$ divided by $\left|\sqrt{F^T F} - U_i(\varepsilon)\right|^2$ converges to 0 uniformly in ε as $\sqrt{F^T F} - U_i(\varepsilon) \to 0$. Define $W_\varepsilon^{(1)}$ by

$$W_\varepsilon^{(1)}(F) = \min_i W_{i,\varepsilon}(F).$$

Proposition 3.3 *$W_\varepsilon^{(1)}$ is admissible and has a linear limit $V^{(1)} : \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}$ given by*

$$V^{(1)}(F) = \min_i \left[\frac{1}{2} \left\langle a_i \left(F - U_i \right), F - U_i \right\rangle + w_i \right].$$

Proof. $W_\varepsilon^{(1)}$ is an admissible family of energy functions since

$$\begin{aligned} \mathrm{dist}^2(F, SO(d)) &\le \mathrm{dist}^2(F, \mathcal{U}_\varepsilon) + C\varepsilon^2 = \min_i \mathrm{dist}^2 \left(\sqrt{F^T F}, SO(d)U_i(\varepsilon) \right) + C\varepsilon^2 \\ &\le \min_i \left| \sqrt{F^T F} - U_i(\varepsilon) \right|^2 + C\varepsilon^2 \le C W_\varepsilon^{(1)}(F) + C\varepsilon^2. \end{aligned}$$

Then, since

$$\lim_{\varepsilon\to 0} \varepsilon^{-2} W_{i,\varepsilon}(\mathrm{Id} + \varepsilon F) = \frac{1}{2} \left\langle a_i \left(\frac{F^T + F}{2} - U_i \right), \frac{F^T + F}{2} - U_i \right\rangle + w_i$$

uniformly on compact subsets of $\mathbb{R}^{d\times d}$, also $V_\varepsilon^{(1)}$ converges uniformly on compact subsets of $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ to $V^{(1)}$. Clearly $V^{(1)}$ satisfies the growth condition $V^{(1)}(F) \le \alpha(1 + |F|^2)$ for a suitable constant $\alpha \in \mathbb{R}$. $\square$

This proves that in fact

$$u \mapsto \int_\Omega V^{(1)}(e(u)) = \int_\Omega \min_i \left[\frac{1}{2} \left\langle a_i \left(e(u) - U_i \right), e(u) - U_i \right\rangle + w_i \right]$$

is the functional of linear elasticity induced by the nonlinear functional $u \mapsto \int_\Omega W_\varepsilon(\nabla u)$. More generally than in the preceding example, we see that the well minima formula (2) is justified, whenever $V_\varepsilon \to \min_i \left[\frac{1}{2}\langle a_i(\cdot - U_i), \cdot - U_i\rangle + w_i\right]$ uniformly on compacta, since the latter function clearly has quadratic growth.

Our next example deals with smooth energy densities. Suppose $U_i(\varepsilon) = \mathrm{Id} + \varepsilon U_i$ and $W_\varepsilon^{(2)}$ is given by

$$W_\varepsilon^{(2)}(F) = \frac{\varepsilon^2}{2} \prod_i \rho_i \left(\varepsilon^{-1}\left(\sqrt{F^T F} - U_i(\varepsilon)\right)\right)$$

if $\det(F) \geq 0$, where the $\rho_i : \mathbb{R}^{d\times d}_{\mathrm{sym}} \to [0, \infty)$ are smooth functions such that, for some constants $c_2 > c_1 > 0$ and positive definite a_i, $\rho_i(F) = \langle a_i F, F\rangle$ if $|F| \leq c_1$, $\rho_i(F) > 0$ if $|F| \geq c_1$ and $\alpha|F|^{\frac{2}{N}} \leq \rho_i(F) \leq \beta|F|^{\frac{2}{N}}$ for $|F| \geq c_2$ for some $\alpha, \beta > 0$. If $\det(F) < 0$ let $W_\varepsilon^{(2)}(F) = \infty$. Note that $\nabla^2 W_\varepsilon^{(2)}(U_i(\varepsilon)) = \prod_{j\neq i} \rho_j(U_i - U_j) a_i$ on symmetric matrices.

Proposition 3.4 *$W_\varepsilon^{(2)}$ is admissible. It has a linear limit $V^{(2)} : \mathbb{R}^{d\times d}_{\mathrm{sym}} \to \mathbb{R}$ given by*

$$V^{(2)}(F) = \frac{1}{2} \prod_i \rho_i \left(F - U_i\right).$$

Proof. Since the energy wells are separated by a distance $\geq c\,\varepsilon$ for some $c > 0$, it is not hard to see that we have

$$\mathrm{dist}^2(F, \mathcal{U}_\varepsilon) \leq C W_\varepsilon(F).$$

V_ε is given by

$$\begin{aligned} \varepsilon^{-2} W_\varepsilon(\mathrm{Id} + \varepsilon F) &= \frac{1}{2} \prod_i \rho_i \left(\varepsilon^{-1}\left(\sqrt{(\mathrm{Id} + \varepsilon F)^T(\mathrm{Id} + \varepsilon F)} - U_i(\varepsilon)\right)\right) \\ &= \frac{1}{2} \prod_i \rho_i \left(\frac{F^T + F}{2} - U_i\right) + O(\varepsilon) \end{aligned}$$

converging uniformly on compact subsets of $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ to $V^{(2)}$ which obviously has quadratic growth. ☐

In the following example will show that even in the simplest case where W_ε is an isotropic double well potential with two incompatible wells, (2) might yield the wrong (relaxed) limiting energy.

If $V^{(1)}$, derived from $W_\varepsilon^{(1)}$ as in Proposition 3.3, is given by

$$V^{(1)}(F) = \min\left\{\frac{1}{2}\langle a(F - U_1), (F - U_1)\rangle, \langle a(F - U_2), (F - U_2)\rangle\right\},$$

where a is isotropic with bulk modulus κ and shear modulus μ, i.e.,

$$aF = \kappa\,\mathrm{tr}(F)\mathrm{Id} + 2\mu\left(F - \frac{1}{d}\,\mathrm{tr}(F)\mathrm{Id}\right),$$

$U_1 - U_2 = \eta\,\mathrm{Id}$ for some $\eta > 0$ and $d \geq 2$, then we infer from Proposition 4.4 and Theorem 3.1 in [18] that the quasiconvexification on linear strains of $V^{(1)}$ satisfies

$$Q^e V^{(1)}(F) = \min_{\theta\in[0,1]} \Big\{ \frac{\theta}{2}\langle a(F-U_1),(F-U_1)\rangle + \frac{1-\theta}{2}\langle a(F-U_2),(F-U_2)\rangle - \frac{\theta(1-\theta)c\eta^2}{2} \Big\}$$

where $c = \frac{\kappa^2 d^3}{\kappa d + 2(d-1)\mu}$. In particular, if $\mu = \frac{1}{2}$ and $\kappa = \frac{2\mu}{d} = \frac{1}{d}$, then a is the identity tensor, $c = 1$ and it is elementary to see that, for $\eta = 1$,

$$Q^e V^{(1)}(tU_1 + (1-t)U_2) = \begin{cases} \frac{dt^2}{2}, & \text{if } t \leq \frac{1}{2} - \frac{1}{2d}, \\ \frac{d(d-1)}{2}t(1-t) - \frac{(d-1)^2}{8}, & \text{if } \frac{1}{2} - \frac{1}{2d} \leq t \leq \frac{1}{2} + \frac{1}{2d}, \\ \frac{d(1-t)^2}{2}, & \text{if } t \geq \frac{1}{2} + \frac{1}{2d}. \end{cases}$$

Note that $Q^e V^{(1)}$ still has double-well structure and, in particular,

$$Q^e V^{(1)}\left(\frac{U_1+U_2}{2}\right) = \frac{d-1}{8}. \tag{11}$$

Now suppose $W_\varepsilon^{(2)}$ is given by

$$W_\varepsilon^{(2)}(F) = \frac{\varepsilon^2}{2}\rho\left(\varepsilon^{-1}\left|\sqrt{F^TF} - U_1(\varepsilon)\right|\right)\rho\left(\varepsilon^{-1}\left|\sqrt{F^TF} - U_2(\varepsilon)\right|\right),$$

where $U_i(\varepsilon) = \mathrm{Id} + \varepsilon U_i$ and the smooth function $\rho(t)$ equals t^2 for $|t| \leq \frac{1}{4}$, is positive for $t \neq 0$ and scales linearly near ∞. Assume moreover that $\rho(\sqrt{d}) = 1$ and $\rho(\frac{\sqrt{d}}{2}) < \frac{1}{2}$. Note that, for $V_\varepsilon^{(2)} = \varepsilon^{-2}W_\varepsilon^{(2)}(\mathrm{Id} + \varepsilon\cdot)$, $\nabla^2 V_\varepsilon^{(2)}(U_i)$ is the identity tensor on symmetric matrices as for $V^{(1)}$ above. But by Proposition 3.4 the linear limit is given by

$$V^{(2)}(F) = \frac{1}{2}\rho\left(|F-U_1|\right)\rho\left(|F-U_2|\right)$$

and

$$Q^e V^{(2)}\left(\frac{U_1+U_2}{2}\right) \leq V^{(2)}\left(\frac{U_1+U_2}{2}\right) = \frac{1}{2}\rho^2\left(\left|\frac{U_2-U_1}{2}\right|\right) < \frac{1}{8}.$$

Comparing with (11), we find that indeed $Q^e V^{(1)} \neq Q^e V^{(2)}$.

4 Proofs of the main results

4.1 Compactness

With the help of the geometric rigidity result of Friesecke, James and Müller (cf. [13]), Dal Maso, Negri and Percivale prove compactness of finite energy sequences for a single-well energy in [11]. The adaption to our multiwell set-up is straightforward, and we refer to [11] for the largest part of the proof. In the following, C denotes a generic constant, which is independent of ε but whose value might change from line to line.

Proposition 4.1 *Suppose* (u_ε) *is a sequence in* $H^1_{g,\partial\Omega_*}(\Omega,\mathbb{R}^d)$*. Then*

$$\|\nabla u_\varepsilon\|_{L^2}^2 \le C\mathcal{E}_\varepsilon(u_\varepsilon) + C\int_{\partial\Omega_*} |g|^2\, d\mathcal{H}^{d-1} + C\|l\|_{L^2}\|u_\varepsilon\|_{L^2} + C.$$

Proof. Define $\mathcal{I}_\varepsilon : L^2(\Omega,\mathbb{R}^d) \to [0,\infty]$ by

$$\mathcal{I}_\varepsilon(v) := \begin{cases} \varepsilon^{-2}\int_\Omega W_\varepsilon(\mathrm{Id} + \varepsilon\nabla v), & \text{if } v \in H^1_{g,\partial\Omega_*}, \\ \infty, & \text{otherwise.} \end{cases} \tag{12}$$

By (4), there exist rotations $R_\varepsilon(x)$ and indices $i(x)$ such that

$$\int_\Omega |\mathrm{Id} + \varepsilon\nabla u_\varepsilon(x) - R_\varepsilon(x)U_{i(x)}(\varepsilon)|^2\, dx \le C\varepsilon^2\mathcal{I}_\varepsilon(u_\varepsilon).$$

Since by assumption $|U_i(\varepsilon) - \mathrm{Id}| \le C\varepsilon$ for all i, it follows from $|\mathrm{Id} + \varepsilon\nabla u_\varepsilon(x) - R_\varepsilon(x)U_{i(x)}(\varepsilon)|^2 \le 2|\mathrm{Id} + \varepsilon\nabla u_\varepsilon(x) - R_\varepsilon(x)|^2 + 2|R_\varepsilon(x) - R_\varepsilon(x)U_{i(x)}(\varepsilon)|^2$ that

$$\int_\Omega |\mathrm{Id} + \varepsilon\nabla u_\varepsilon(x) - R_\varepsilon(x)|^2\, dx \le C\varepsilon^2(1 + \mathcal{I}_\varepsilon(u_\varepsilon)).$$

Now the geometric rigidity result in [13] yields constant rotations R_ε such that $\int_\Omega |\mathrm{Id} + \varepsilon\nabla u_\varepsilon(x) - R_\varepsilon|^2\, dx \le C\int_\Omega |\mathrm{Id} + \varepsilon\nabla u_\varepsilon(x) - R_\varepsilon(x)|^2\, dx$ and hence

$$\int_\Omega |\mathrm{Id} + \varepsilon\nabla u_\varepsilon - R_\varepsilon|^2 \le C\varepsilon^2(1 + \mathcal{I}_\varepsilon(u_\varepsilon)). \tag{13}$$

As shown in the proof of Proposition 3.4 in [11], this implies that

$$\int_\Omega |\nabla u_\varepsilon|^2 \le C(1 + \mathcal{I}_\varepsilon(u_\varepsilon)) + C\int_{\partial\Omega_*} |g|^2\, d\mathcal{H}^{d-1},$$

which proves the claim. □

Proof of Proposition 2.2. Immediate from Proposition 4.1 and Poincaré's inequality. □

For later use we also state the following variant of Proposition 4.1, which can be proved using a refinement of the main geometric rigidity result (see Proposition 5 in [14]).

Lemma 4.2 *Suppose* (u_ε) *is a sequence in* $H^1_{g,\partial\Omega_*}(\Omega,\mathbb{R}^d)$ *such that* $\varepsilon^{-2}\mathrm{dist}^2(\mathrm{Id}+\varepsilon\nabla u_\varepsilon,\mathcal{U}_\varepsilon)$ *is equiintegrable. Then also* $|\nabla u_\varepsilon|^2$ *is equiintegrable.*

Proof. As above note that

$$\mathrm{dist}^2(\mathrm{Id} + \varepsilon\nabla u_\varepsilon, SO(d)) \le C(\varepsilon^2 + \mathrm{dist}^2(\mathrm{Id} + \varepsilon\nabla u_\varepsilon,\mathcal{U}_\varepsilon)),$$

whence $\varepsilon^{-2}\mathrm{dist}^2(\mathrm{Id} + \varepsilon\nabla u_\varepsilon, SO(d))$ is equiintegrable. The refined geometric rigidity in [14] yields constant rotations R_ε such that also $\varepsilon^{-2}|\mathrm{Id} + \varepsilon\nabla u_\varepsilon - R_\varepsilon|^2$ is equiintegrable. As in the proof of Proposition 3.4 in [11] we see that

$$\int_\Omega |\mathrm{Id} - R_\varepsilon|^2 \le C\varepsilon^2\int_\Omega (1 + \varepsilon^{-2}\mathrm{dist}^2(\mathrm{Id} + \varepsilon\nabla u_\varepsilon,\mathcal{U}_\varepsilon)) + C\varepsilon^2\int_{\partial\Omega_*} |g|^2 d\mathcal{H}^{d-1}$$

and thus $|\mathrm{Id} - R_\varepsilon| \le C\varepsilon$. It follows that $|\nabla u_\varepsilon|^2$ is equiintegrable. □

4.2 Γ-convergence

Theorem 2.1 follows from a combination of the compactness result in Proposition 4.1 and the approximation result for quasiconvex functions in Lemma A.3.

Proof of Theorem 2.1. Define $\mathcal{I}_\varepsilon$ as in (12). Since $v \mapsto \int_\Omega lv$ is a continuous functional on L^2 and, by Proposition 4.1, bounded energy sequences are bounded in H^1, it suffices to show that the L^2-Γ-limit of $\mathcal{I}_\varepsilon$ is given by

$$\mathcal{I}^{\mathrm{rel}}(v) := \int_\Omega Q^e V(e(v)).$$

Suppose $u_\varepsilon \to u$ in $L^2(\Omega, \mathbb{R}^d)$. To prove the lower bound we may without loss of generality assume that u_ε is a sequence such that $\mathcal{I}_\varepsilon(u_\varepsilon)$ is bounded and hence, thanks to Proposition 4.1, $\|\nabla u_\varepsilon\|_{L^2} \le C$ for some suitable constant C, whence $u_\varepsilon \rightharpoonup u$ in $H^1_{g,\partial\Omega_*}$. By frame indifference,

$$\begin{aligned}\varepsilon^{-2}\int_\Omega W_\varepsilon(\mathrm{Id}+\varepsilon\nabla u_\varepsilon) &= \varepsilon^{-2}\int_\Omega W_\varepsilon\left(\sqrt{(\mathrm{Id}+\varepsilon\nabla u_\varepsilon)^T(\mathrm{Id}+\varepsilon\nabla u_\varepsilon)}\right)\\ = \varepsilon^{-2}\int_\Omega W_\varepsilon\left(\mathrm{Id}+\varepsilon e(u_\varepsilon)+f(\varepsilon\nabla u_\varepsilon)\right) &= \int_\Omega V_\varepsilon\left(e(u_\varepsilon)+\varepsilon^{-1}f(\varepsilon\nabla u_\varepsilon)\right),\end{aligned}$$

where the error term $f(F) = \sqrt{(\mathrm{Id}+F)^T(\mathrm{Id}+F)} - \mathrm{Id} - \frac{F^T+F}{2}$ can be bounded by

$$|f(F)| \le C\min\{|F|, |F|^2\}. \tag{14}$$

Noting that, for a matrix $A \in \mathbb{R}^{d\times d}$, $\mathrm{dist}(A, SO(d)) \ge |\sqrt{A^TA} - \mathrm{Id}|$ (with equality if $\det(A) > 0$), due to our assumptions on W_ε we have

$$\begin{aligned}&V_\varepsilon\left(\frac{F^T+F}{2}+\varepsilon^{-1}f(\varepsilon F)\right) = \varepsilon^{-2}W_\varepsilon(\mathrm{Id}+\varepsilon F)\\ &\ge c\,\varepsilon^{-2}\mathrm{dist}^2(\mathrm{Id}+\varepsilon F, SO(d)) - C \ge c\,\varepsilon^{-2}\left|\sqrt{(\mathrm{Id}+\varepsilon F)^T(\mathrm{Id}+\varepsilon F)} - \mathrm{Id}\right|^2 - C\\ &= c\left|\frac{F^T+F}{2}+\varepsilon^{-1}f(\varepsilon F)\right|^2 - C.\end{aligned}$$

Let $\delta > 0$ and choose ψ_k approximating the quasiconvex envelope QU_δ of U_δ : $F \mapsto V(\frac{F^T+F}{2}) + \delta|F|^2$ as in Lemma A.3. Since $V_\varepsilon \to V$ uniformly on compact subsets of the symmetric matrices and ψ_k grows linearly at ∞, we can therefore find $\varepsilon(k) > 0$ such that, for all $\varepsilon \le \varepsilon(k)$,

$$\begin{aligned}&V_\varepsilon\left(\frac{F^T+F}{2}+\varepsilon^{-1}f(\varepsilon F)\right) + \delta\left|F+\varepsilon^{-1}f(\varepsilon F)\right|^2\\ &\ge \psi_k\left(F+\varepsilon^{-1}f(\varepsilon F)\right) - \frac{1}{k}.\end{aligned} \tag{15}$$

It follows that

$$\begin{aligned}
&\int_\Omega V_\varepsilon\left(e(u_\varepsilon)+\varepsilon^{-1}f(\varepsilon\nabla u_\varepsilon)\right)\\
&\geq \int_\Omega \psi_k\left(\nabla u_\varepsilon+\varepsilon^{-1}f(\varepsilon\nabla u_\varepsilon)\right)-\int_\Omega \delta\left|\nabla u_\varepsilon+\varepsilon^{-1}f(\varepsilon\nabla u_\varepsilon)\right|^2-\frac{1}{k}\\
&\geq \int_\Omega \psi_k\left(\nabla u_\varepsilon\right)-\int_\Omega\left|\psi_k\left(\nabla u_\varepsilon\right)-\psi_k\left(\nabla u_\varepsilon+\varepsilon^{-1}f(\varepsilon\nabla u_\varepsilon)\right)\right|-C\delta\|\nabla u_\varepsilon\|_{L^2}^2-\frac{1}{k}
\end{aligned}$$

by (14).

For any $M>0$, the second term on the right-hand side can be estimated by splitting the integrand into two parts according to $|\nabla u_\varepsilon|>M$ or $|\nabla u_\varepsilon|\leq M$:

$$\begin{aligned}
&\int_\Omega\left|\psi_k\left(\nabla u_\varepsilon\right)-\psi_k\left(\nabla u_\varepsilon+\varepsilon^{-1}f(\varepsilon\nabla u_\varepsilon)\right)\right|\\
&\leq \int_\Omega \chi_{\{|\nabla u_\varepsilon|>M\}}\left(\left|\psi_k\left(\nabla u_\varepsilon\right)\right|+\left|\psi_k\left(\nabla u_\varepsilon+\varepsilon^{-1}f(\varepsilon\nabla u_\varepsilon)\right)\right|\right) \qquad (16)\\
&\qquad+\int_\Omega \chi_{\{|\nabla u_\varepsilon|\leq M\}}\left|\psi_k\left(\nabla u_\varepsilon\right)-\psi_k\left(\nabla u_\varepsilon+\varepsilon^{-1}f(\varepsilon\nabla u_\varepsilon)\right)\right|
\end{aligned}$$

Now let $\eta>0$ be arbitrary. Since for suitable constants $\tilde{a}_k,\tilde{b}_k$ we have $\psi_k(F)\leq\tilde{a}_k|F|+\tilde{b}_k$, using (14) and the fact that (∇u_ε) is bounded in L^2 and hence equiintegrable we find that for all ε the first integral on the right-hand side of (16) is bounded from above by

$$\begin{aligned}
&\int_\Omega \chi_{\{|\nabla u_\varepsilon|>M\}}\left(2\tilde{a}_k\left|\nabla u_\varepsilon\right|+\tilde{a}_k\left|\varepsilon^{-1}f(\varepsilon\nabla u_\varepsilon)\right|+2\tilde{b}_k\right)\\
&\leq \int_\Omega \chi_{\{|\nabla u_\varepsilon|>M\}}\left(C\tilde{a}_k\left|\nabla u_\varepsilon\right|+2\tilde{b}_k\right)<\eta,
\end{aligned}$$

whenever $M=M(k,\eta)$ is chosen large enough. Now choosing ε sufficiently small we can bound the second term on the right-hand side of (16) by η, too, since on $\{|\nabla u_\varepsilon|\leq M\}$ we have $|\varepsilon^{-1}f(\varepsilon\nabla u_\varepsilon)|\leq C\varepsilon M^2$ by (14) and ψ_k, being quasiconvex and finite valued, is continuous.

Summarizing, we have shown that for each k

$$\liminf_{\varepsilon\to 0}\varepsilon^{-2}\int_\Omega W_\varepsilon(\mathrm{Id}+\varepsilon\nabla u_\varepsilon)\geq\liminf_{\varepsilon\to 0}\int_\Omega\psi_k\left(\nabla u_\varepsilon\right)-C\delta-\frac{1}{k}. \qquad (17)$$

Since by quasiconvexity of ψ_k the functional $v\mapsto\int_\Omega\psi_k(\nabla v)$ is weakly lower semicontinuous on $W^{1,1}$ and since $\psi_k(F)\to QU_\delta(F)$ as $k\to\infty$, it follows that

$$\liminf_{\varepsilon\to 0}\mathcal{I}_\varepsilon(u_\varepsilon)\geq\int_\Omega QU_\delta(\nabla u)-C\delta$$

by monotone convergence. Now note that the quasiconvex hull of $F\mapsto V(\frac{F^T+F}{2})$ clearly satisfies $QV(F)\leq V(\frac{F^T+F}{2})\leq U_\delta(F)$, whence $QU_\delta(F)\geq QV(F)$. So finally sending $\delta\to 0$, we obtain that indeed

$$\liminf_{\varepsilon\to 0}\mathcal{I}_\varepsilon(u_\varepsilon)\geq\int_\Omega QV(\nabla u)=\int_\Omega Q^eV(e(u)).$$

It remains to provide a recovery sequence for given $u \in H^1_{g,\partial\Omega_*}$. As the limiting functional is continuous with respect to the strong H^1-topology, by a general density argument in the theory of Γ-convergence we may assume that $u \in W^{1,\infty}_{g,\partial\Omega_*}$. Suppose that, for given $\delta > 0$, we can choose $u^\delta \in W^{1,\infty}(\Omega;\mathbb{R}^d)$ with $u^\delta - u \in H^1_0$ and $\|u^\delta - u\|_{L^2} \leq \delta$ such that

$$\int_\Omega V(e(u^\delta)) \leq \int_\Omega Q^e V\left(e(u)\right) + \delta. \tag{18}$$

Then, setting $u_\varepsilon = u^\delta$ for all ε, as before we find

$$\varepsilon^{-2} \int_\Omega W_\varepsilon(\mathrm{Id} + \varepsilon\nabla u_\varepsilon) = \int_\Omega V_\varepsilon(e(u_\varepsilon) + \varepsilon^{-1} f(\varepsilon\nabla u_\varepsilon)).$$

Since ∇u_ε takes values in a compact subset of matrices, $V_\varepsilon \to V$ uniformly on compacta and thus V is continuous, we find by (14) that this converges to

$$\lim_{\varepsilon\to 0} \mathcal{I}_\varepsilon(u_\varepsilon) = \int_\Omega V(e(u^\delta)) \leq \mathcal{I}^{\mathrm{rel}}(u) + \delta.$$

Now choosing a diagonal sequence as $\delta \to 0$ we obtain a recovery sequence for u, which finishes the proof.

It remains to prove (18). Let $\gamma > 0$. Since V satisfies a 2-growth assumption from above, the constant sequence of functionals

$$v \mapsto \begin{cases} \int_\Omega V(e(v)) + \gamma|\nabla v|^2, & \text{if } v - u \in H^1_0, \\ \infty, & \text{otherwise,} \end{cases}$$

L^2-Γ-converges to their lower semicontinuous envelope

$$v \mapsto \begin{cases} \int_\Omega Q\left[V\left(\frac{(\cdot)^T + (\cdot)}{2}\right) + \gamma|\cdot|^2\right](\nabla v), & \text{if } v - u \in H^1_0, \\ \infty, & \text{otherwise,} \end{cases}$$

by Theorem A.4. In particular, there exists $u^{\gamma,\delta}$ with $u^{\gamma,\delta} - u \in H^1_0$ and $\|u^{\gamma,\delta} - u\|_{L^2} \leq \delta$ such that

$$\int_\Omega V(e(u^{\gamma,\delta})) + \gamma|\nabla u^{\gamma,\delta}|^2 \leq \int_\Omega Q\left[V\left(\frac{(\cdot)^T + (\cdot)}{2}\right) + \gamma|\cdot|^2\right](\nabla u) + \frac{\delta}{2}. \tag{19}$$

Let $F \in \mathbb{R}^{d\times d}$. Since $V(A) \geq c|A|^2 - C$ for all $A \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$, it follows from Korn's inequality (see, e.g., [23]) that

$$\int_\Omega |\nabla v|^2 \leq C \int_\Omega |e(v)|^2 + C|F|^2 \leq C \int_\Omega V(e(v)) + C|F|^2$$

for all $v \in H^1$ with $v(x) = Fx$ on $\partial\Omega$. But then

$$\int_\Omega V(e(v)) + \gamma|\nabla v|^2 \leq (1 + C\gamma) \int_\Omega V(e(v)) + C\gamma|F|^2,$$

so taking the infimum over all v such that $x \mapsto v(x) - Fx$ lies in C_c^∞ leads to

$$Q\left[V\left(\frac{(\cdot)^T + (\cdot)}{2}\right) + \gamma|\cdot|^2\right](F) \le (1 + C\gamma)Q^e V\left(\frac{F^T + F}{2}\right) + C\gamma|F|^2. \quad (20)$$

By (19) we therefore have

$$\int_\Omega V(e(u^{\gamma,\delta})) \le \int_\Omega (1 + C\gamma)Q^e V(e(u)) + C\gamma + \frac{\delta}{2}.$$

Now choosing γ sufficiently small and approximating $u^{\gamma,\delta}$ in H^1 with a $W^{1,\infty}$ function that satisfies the same boundary values finally gives us u^δ satisfying (18). □

Remark 4.3 The above proof of Equation (18) is already half the proof of the version of Theorem A.4 for linear strains. Indeed, if V satisfies a standard 2-growth condition on $\mathbb{R}^{d\times d}_{\rm sym}$, then the constant sequence of functionals

$$v \mapsto \begin{cases} \int_\Omega V(e(v)), & \text{if } v - u \in H^1_0, \\ \infty, & \text{otherwise,} \end{cases}$$

L^2-Γ-converges to their lower semicontinuous envelope

$$v \mapsto \begin{cases} \int_\Omega Q^e V(e(v)), & \text{if } v - u \in H^1_0, \\ \infty, & \text{otherwise.} \end{cases}$$

The remaining part of the proof of this relaxation result is straightforward: Just note that, in order to prove the Γ-liminf inequality, one may assume that $\int_\Omega V(e(v_n)) < C$ for a sequence $v_n \to v$ in L^2 and hence that v_n is bounded in H^1 by Korn's inequality. Now use Theorem A.4 similarly as before to conclude the proof.

4.3 Convergence of almost minimizers

Proof of Theorem 2.3. Let u_ε be a recovery sequence for u. We have to show that

$$\lim_{\varepsilon\to 0} \mathcal{E}(u_\varepsilon) = \mathcal{E}^{\rm rel}(u), \quad (21)$$

equivalently, that each sequence $\varepsilon = \varepsilon_k$ converging to 0 has a subsequence such that (21) holds. By continuity of the loading term we may replace $\mathcal{E}$ by $\mathcal{I}$ as in the previous proofs. Furthermore, we may assume that $\mathcal{I}_\varepsilon(u_\varepsilon)$ is bounded and $u_\varepsilon \rightharpoonup u$ in $H^1_{g,\partial\Omega_*}$.

We first claim that, for $A_\varepsilon := e(u_\varepsilon) + \varepsilon^{-1} f(\varepsilon\nabla u_\varepsilon)$, $|A_\varepsilon|^2$ is equiintegrable. Suppose this were not the case. Then

$$\limsup_{\varepsilon\to 0} \int_{|A_\varepsilon|\ge M} |A_\varepsilon|^2 \ge \gamma > 0$$

for all M and a suitable γ and – upon passing to a subsequence – we can replace the $\limsup_{\varepsilon\to 0}$ by $\liminf_{\varepsilon\to 0}$.

Let ψ_k as in the proof of Theorem 2.1 and define $\tilde{\psi}_k$ by

$$\tilde{\psi}_k(F) = \psi_k(F) + \frac{c}{2}\chi_{\{|F^T+F|\geq 2R_k\}}\left|\frac{F^T+F}{2}\right|^2,$$

where c is the constant from (4). If R_k is sufficiently large, we may replace ψ_k by $\tilde{\psi}_k$ in (15) in the proof of Theorem 2.1, which (as in the derivation of Equation (17)) yields

$$\begin{aligned}
&\liminf_{\varepsilon\to 0} \varepsilon^{-2}\int_\Omega W_\varepsilon(\mathrm{Id}+\varepsilon\nabla u_\varepsilon) \\
&\geq \liminf_{\varepsilon\to 0}\left(\int_\Omega \psi_k(\nabla u_\varepsilon) - C\delta - \frac{1}{k} + \int_\Omega \frac{c}{2}\chi_{\{|A_\varepsilon|\geq R_k\}}|A_\varepsilon|^2\right) \\
&\geq \int_\Omega \psi_k(\nabla u) - C\delta - \frac{1}{k} + \liminf_{\varepsilon\to 0}\frac{c}{2}\int_{\{|A_\varepsilon|\geq R_k\}}|A_\varepsilon|^2.
\end{aligned}$$

Continuing as in the proof of Theorem 2.1, we see that

$$\liminf_{\varepsilon\to 0}\varepsilon^{-2}\int_\Omega W_\varepsilon(\mathrm{Id}+\varepsilon\nabla u_\varepsilon) \geq \int_\Omega Q^e V(e(u)) + \frac{c\gamma}{2}.$$

But u_ε is a recovery sequence, hence $\gamma = 0$ in contradiction to our assumption.

In particular, this proves that

$$\varepsilon^{-2}\mathrm{dist}^2(\mathrm{Id}+\varepsilon\nabla u_\varepsilon, SO(d)) = \varepsilon^{-2}|\mathrm{Id}+\varepsilon e(u_\varepsilon) + f(\varepsilon\nabla u_\varepsilon) - \mathrm{Id}|^2 = |A_\varepsilon|^2$$

is equiintegrable. By Lemma 4.2 this in turn implies that also $|\nabla(u_\varepsilon)|^2$ is equiintegrable and we can conclude the proof as follows.

For each $M > 0$,

$$\begin{aligned}
\int_\Omega Q^e V(e(u)) &= \lim_{\varepsilon\to 0}\varepsilon^{-2}\int_\Omega W_\varepsilon(\mathrm{Id}+\varepsilon\nabla u_\varepsilon) = \lim_{\varepsilon\to 0}\int_\Omega V_\varepsilon(A_\varepsilon) \\
&\geq \limsup_{\varepsilon\to 0}\int_{\{|\nabla u_\varepsilon|\leq M\}} V_\varepsilon(A_\varepsilon) = \limsup_{\varepsilon\to 0}\int_{\{|\nabla u_\varepsilon|\leq M\}} V(e(u_\varepsilon))
\end{aligned}$$

since $V_\varepsilon \to V$ uniformly on compacta and $A_\varepsilon - e(u_\varepsilon) \to 0$ uniformly on $\{|\nabla u_\varepsilon| \leq M\}$ by (14). Since M was arbitrary, by the 2-growth assumption on V and the equiintegrability of $|e(u_\varepsilon)|^2$ we therefore obtain that

$$\int_\Omega Q^e V(e(u)) \geq \limsup_{\varepsilon\to 0}\int_\Omega V(e(u_\varepsilon)).$$

Since by lower semicontinuity also $\int_\Omega Q^e V(e(u)) \leq \liminf_{\varepsilon\to 0}\int_\Omega V(e(u_\varepsilon))$, we finally get that

$$\int_\Omega Q^e V(e(u)) = \lim_{\varepsilon\to 0}\int_\Omega V(e(u_\varepsilon)).$$

□

For later use we note that the proof shows that in fact $|\nabla u_{\varepsilon_k}|^2$ is equiintegrable for every sequence $\varepsilon_k \to 0$.

Proof of Theorem 2.4. Suppose that u_ε is a low energy sequence. By Theorem 2.1 and Proposition 2.2 we have

$$\lim_{\varepsilon\to 0} \mathcal{E}_\varepsilon(u_\varepsilon) = \min_{v\in H^1_{g,\partial\Omega_*}} \mathcal{E}^{\rm rel}(v) \le \inf_{v\in H^1_{g,\partial\Omega_*}} \mathcal{E}(v)$$

and $u_\varepsilon \rightharpoonup u$ in $H^1_{g,\partial\Omega_*}$ up to subsequences, where u is a minimizer of $\mathcal{E}^{\rm rel}$. It remains to show that

$$\lim_{\varepsilon\to 0} \mathcal{E}_\varepsilon(u_\varepsilon) = \lim_{\varepsilon\to 0} \mathcal{E}(u_\varepsilon).$$

Passing to a suitable subsequence of an arbitrary subsequence, we may assume that $u_\varepsilon \rightharpoonup u$ in H^1, where u is a minimizer of $\mathcal{E}^{\rm rel}$. By the first part of the proof we have $\mathcal{E}_\varepsilon(u_\varepsilon) \to \mathcal{E}^{\rm rel}(u)$, i.e., u_ε is a recovery sequence for u. By Theorem 2.3 we therefore obtain that indeed $\mathcal{E}(u_\varepsilon) \to \mathcal{E}^{\rm rel}(u)$. □

Proof of Corollary 2.5. Suppose (ν_x) is induced by the subsequence ∇u_{ε_k}, $\varepsilon_k \to 0$. By Theorem 2.3 we have

$$\mathcal{E}^{\rm rel}(u) = \lim_{\varepsilon_k\to 0} \mathcal{E}(u_{\varepsilon_k}) = \lim_{\varepsilon_k\to 0} \int_\Omega V(e(u_{\varepsilon_k})) - \int_\Omega lu.$$

As shown in the proof of Theorem 2.3, $V(e(u_{\varepsilon_k}))$ is equiintegrable, whence $V(e(u_{\varepsilon_k})) \rightharpoonup \bar{V}$ in $L^1(\Omega)$, where $\bar{V}(x) = \int_{\mathbb{R}^{d\times d}_{\rm sym}} V(F)\, d\nu^e_x(F)$. So

$$\mathcal{E}^{\rm rel}(u) = \int_\Omega \bar{V}(x) - \int_\Omega lu.$$

Note that if u_ε is a low energy sequence, then u is a minimizer of $\mathcal{E}^{\rm rel}$. □

4.4 Relaxed densities and compatibility

Proof of Corollary 2.6. Let $F \in \mathbb{R}^{d\times d}$, $\partial\Omega_* = \partial\Omega$ and $g(x) = Fx$, $l = 0$. On the one hand, we clearly have that

$$\inf_{v\in H^1_{g,\partial\Omega_*}} \mathcal{E}_\varepsilon(v) \le \inf_{v-F\cdot\in C^\infty_c} \mathcal{E}_\varepsilon(v) = \varepsilon^{-2} QW_\varepsilon(\mathrm{Id} + \varepsilon F)|\Omega|,$$

and so

$$\liminf_{\varepsilon\to 0} \varepsilon^{-2} QW_\varepsilon(\mathrm{Id} + \varepsilon F)|\Omega| \ge \inf_{v\in H^1_{g,\partial\Omega_*}} \mathcal{E}(v)$$

by Theorem 2.4. Since V satisfies a 2-growth condition from above, the latter expression is equal to $Q^e V(\frac{F^T+F}{2})|\Omega|$.

On the other hand, choosing u with $u - F\cdot \in C^\infty_c$ such that, for given $\eta > 0$,

$$\int_\Omega V(e(u)) \le Q^e V\left(\frac{F^T + F}{2}\right)|\Omega| + \eta|\Omega|,$$

we see as in the proof of Theorem 2.1 that the constant sequence $u_\varepsilon = u$ satisfies

$$\int_\Omega \varepsilon^{-2} W_\varepsilon(\mathrm{Id} + \varepsilon\nabla u_\varepsilon) \to \int_\Omega V\left(e(u)\right).$$

It follows that

$$\limsup_{\varepsilon\to 0} \varepsilon^{-2} QW_\varepsilon(\mathrm{Id} + \varepsilon F)|\Omega| \le \limsup_{\varepsilon\to 0} \int_\Omega \varepsilon^{-2} W_\varepsilon(\mathrm{Id} + \varepsilon \nabla u_\varepsilon)$$
$$\le \int_\Omega V\left(e(u)\right) \le Q^e V\left(\frac{F^T + F}{2}\right)|\Omega| + \eta|\Omega|.$$

Since $\eta > 0$ was arbitrary, the claim is proved. □

Proof of Corollary 2.8. Define $\tilde{W}_\varepsilon$ by $\tilde{W}_\varepsilon(F) := c\,\mathrm{dist}^2(F, \mathcal{U}_\varepsilon)$ (resp., $+\infty$ if $\det(F) \le 0$). If $c > 0$ is sufficiently small, then $0 \le \tilde{W}_\varepsilon \le W_\varepsilon$ by (4) and so

$$\liminf_{\varepsilon\to 0} \varepsilon^{-2} \int_\Omega \tilde{W}_\varepsilon\left(\mathrm{Id} + \varepsilon \nabla u_\varepsilon\right) = 0. \tag{22}$$

On the other hand, if $A \in \mathbb{R}^{d\times d}_{\mathrm{sym}}$, then

$$\tilde{V}_\varepsilon(A) := \varepsilon^{-2} \tilde{W}_\varepsilon(\mathrm{Id} + \varepsilon A) = c\,\varepsilon^{-2} \mathrm{dist}^2\left(\mathrm{Id} + \varepsilon A, \mathcal{U}_\varepsilon\right)$$
$$= c\,\min_i |A - U_i|^2 + O(\varepsilon)$$

with an error term $O(\varepsilon)$, which can be chosen uniformly on compact subsets of $\mathbb{R}^{d\times d}_{\mathrm{sym}}$. By Theorem 2.4 and (22) we therefore get

$$0 = \lim_{\varepsilon\to 0} \varepsilon^{-2} \int_\Omega \tilde{W}_\varepsilon\left(\mathrm{Id} + \varepsilon \nabla u_\varepsilon\right) \ge \inf_{u - F\cdot \in H^1_0} \int_\Omega Q^e \tilde{V}(u) = Q^e \tilde{V}\left(\frac{F^T + F}{2}\right) \ge 0 \tag{23}$$

for $\tilde{V} : \mathbb{R}^{d\times d}_{\mathrm{sym}} \to [0,\infty)$, $\tilde{V}(F) := c\,\min_i |F - U_i|^2 = c\,\mathrm{dist}^2(F, \{U_1, \ldots, U_N\})$ (satisfying a 2-growth condition from above). The claim now follows from Theorem A.1. □

4.5 Extensions

It is easy to see that – except for Corollary 2.8 – assumption (4) may be replaced by the weaker condition (8). The extension to x-dependent W_ε, however, is not so straightforward. In the following we provide the necessary modifications of the previous proofs.

Propositions 4.1 and 2.2 are clear. To prove Theorem 2.1 note first that Lemma A.3 can be replaced by

Lemma 4.4 *Suppose $\psi : \Omega \times \mathbb{R}^{m\times n} \to \mathbb{R}$ is a Carathéodory function such that $\psi(x,\cdot)$ is quasiconvex for a.e. x satisfying the growth condition*

$$\alpha|F|^p - \beta \le \psi(x, F) \le \beta|F|^p + \beta \quad \text{for a.e. } x \in \Omega \text{ and all } F \in \mathbb{R}^{m\times n}$$

for some $p > 1$ and constants $\alpha, \beta > 0$. Then there are Carathéodory functions $\psi_j : \Omega \times \mathbb{R}^{m\times n} \to \mathbb{R}$, quasiconvex in the second variable for a.e. x, such that, for a.e. x, $\psi_j(x,\cdot) \le \psi_{j+1}(x,\cdot)$, $\psi_j(x,\cdot) \to \psi(x,\cdot)$ pointwise as $j \to \infty$ and there exist $a_j, b_j \in \mathbb{R}$ independent of x such that

$$|\psi_j(x, F)| \le a_j|F| + b_j.$$

Proof. This follows by fixing the first variable and quasiconvexifying in the second (as in Theorem A.4): We can apply the proof of Lemma 4.1 in [19], where it is shown that for fixed x one can take

$$\psi_j(x,F) := Q[\min\{j(1+|\cdot|),\psi(x,\cdot)\}](F).$$

□

Now (17) can be seen as before. In order to estimate $\int_\Omega \chi_{|\nabla u_\varepsilon|\le M}|\psi_k(x,\nabla u_\varepsilon)-\psi_k(x,\nabla u_\varepsilon+\varepsilon^{-1}f(\varepsilon\nabla u_\varepsilon))|$ one uses that the integrand converges to 0 as $\varepsilon\to 0$ boundedly in x.

When constructing the recovery sequence, in particular when proving (20), we can start from $V(x_0,A)\ge c|A|^2-C$ for fixed $x_0\in\Omega$ and follow the same arguments to see that

$$Q\left[V\left(x_0,\frac{(\cdot)^T+(\cdot)}{2}\right)+\gamma|\cdot|^2\right](F)\le(1+C\gamma)Q^eV\left(x_0,\frac{F^T+F}{2}\right)+C\gamma|F|^2$$

for all $x_0\in\Omega$. The remaining parts of the proof do not need to be modified.

In the proof of Theorem 2.3 we find as before that

$$\int_\Omega Q^eV(x,e(u))\ge\limsup_{\varepsilon\to 0}\int_{\{|\nabla u_\varepsilon|\le M\}}V_\varepsilon(x,A_\varepsilon).$$

By the uniform convergence $V_\varepsilon\to V$ on $\Omega\times K$ for compact K we may replace V_ε by V. Now, using the Scorza-Dragoni Theorem, for each $\eta>0$ we find a compact subset K_η of Ω such that V is uniformly continuous on $\{(x,F):x\in K,\ |F|\le M+1\}$ and $|\Omega\setminus K_\eta|\le\eta$. It follows that

$$\int_\Omega Q^eV(x,e(u))\ge\limsup_{\varepsilon\to 0}\int_{K_\eta\cap\{|\nabla u_\varepsilon|\le M\}}V(x,e(u_\varepsilon)).$$

Again by the equiintegrability of $e(u_\varepsilon)$ we can deduce that

$$\int_\Omega Q^eV(x,e(u))\ge\limsup_{\varepsilon\to 0}\int_\Omega V(x,e(u_\varepsilon)).$$

The reverse inequality follows as before.

The proof of Theorem 2.4 remains unchanged. For Corollary 2.5 it suffices to note that $V(\cdot,e(u_{\varepsilon_k}))$ is equiintegrable and so $V(\cdot,e(u_{\varepsilon_k}))\rightharpoonup\bar V$ in $L^1(\Omega)$, where $\bar V(x)=\int_{\mathbb{R}^{d\times d}_{\rm sym}}V(x,F)\,d\nu^e_x(F)$, since V is a Carathéodory function (see e.g. [2]). Also the extension to x-dependent W_ε for Corollary 2.8 is clear if (4) holds.

We finally briefly describe what needs to be changed in the preceding proofs if condition (5) is dropped. In the proof of Theorem 2.1 let $D_\varepsilon:=\{x\in\Omega:\det(\mathrm{Id}+\varepsilon\nabla u_\varepsilon)\le 0\}$. Noting that on D_ε we have $\mathrm{dist}(\cdot,SO(d))\ge\gamma$ for some $\gamma>0$, instead of (17) we first arrive at

$$\begin{aligned}&\liminf_{\varepsilon\to 0}\varepsilon^{-2}\int_\Omega W_\varepsilon(\mathrm{Id}+\varepsilon\nabla u_\varepsilon)\\&\ge\liminf_{\varepsilon\to 0}\left(\gamma\,\varepsilon^{-2}|D_\varepsilon|+\int_{\Omega\setminus D_\varepsilon}\psi_k\left(\nabla u_\varepsilon\right)\right)-C\delta-\frac{1}{k}.\end{aligned}$$

$\mathcal{I}_\varepsilon(u_\varepsilon)$ being bounded, this implies that $|D_\varepsilon| \to 0$ as $\varepsilon \to 0$ and this in turn yields (17) by the boundedness of u_ε in H^1 and the linear growth of ψ_k at ∞. For the existence of a recovery sequence one can use Theorem 2.1 for $\bar{W}_\varepsilon$ with $\bar{W}_\varepsilon(F) = W_\varepsilon(F)$ if $\det(F) > 0$ and $\bar{W}_\varepsilon(F) = \infty$ otherwise.

Similarly, in the proof of Theorem 2.3 we obtain

$$\begin{aligned}
&\liminf_{\varepsilon\to 0} \varepsilon^{-2} \int_\Omega W_\varepsilon(\mathrm{Id} + \varepsilon\nabla u_\varepsilon) \\
&\geq \liminf_{\varepsilon\to 0} \Big(\gamma \int_{D_\varepsilon} \varepsilon^{-2} \mathrm{dist}^2(\mathrm{Id} + \varepsilon\nabla u_\varepsilon, SO(d)) + \int_{\Omega\setminus D_\varepsilon} \psi_k\left(\nabla u_\varepsilon\right) \\
&- C\delta - \frac{1}{k} + \int_{\Omega\setminus D_\varepsilon} \frac{c}{2} \chi_{\{|A_\varepsilon \geq R_k\}} |A_\varepsilon|^2 \Big).
\end{aligned}$$

Now using that u_ε is a recovery sequence for u and that

$$\lim_{k\to\infty} \liminf_{\varepsilon\to 0} \int_{\Omega\setminus D_\varepsilon} \psi_k\left(\nabla u_\varepsilon\right) - \frac{1}{k} \geq \mathcal{I}^{\mathrm{rel}}(u)$$

as shown above, we deduce that $|A_\varepsilon|^2 \chi_{\Omega\setminus D_\varepsilon}$ is equiintegrable. For the same reason we have that

$$\liminf_{\varepsilon\to 0} \int_{D_\varepsilon} \varepsilon^{-2} \mathrm{dist}^2(\mathrm{Id} + \varepsilon\nabla u_\varepsilon, SO(d)) = 0.$$

Passing to a suitable subsequence we therefore get that

$$\begin{aligned}
&\varepsilon^{-2}\mathrm{dist}^2(\mathrm{Id} + \varepsilon\nabla u_\varepsilon, SO(d)) \\
&= \chi_{D_\varepsilon} \mathrm{dist}^2(\mathrm{Id} + \varepsilon\nabla u_\varepsilon, SO(d)) + \chi_{\Omega\setminus D_\varepsilon} \varepsilon^{-2} |\mathrm{Id} + \varepsilon e(u_\varepsilon) + f(\varepsilon\nabla u_\varepsilon) - \mathrm{Id}|^2 \\
&= \chi_{D_\varepsilon} \mathrm{dist}^2(\mathrm{Id} + \varepsilon\nabla u_\varepsilon, SO(d)) + \chi_{\Omega\setminus D_\varepsilon} |A_\varepsilon|^2
\end{aligned}$$

is equiintegrable. Similarly as before we deduce that

$$\int_\Omega Q^e V(e(u)) \geq \limsup_{\varepsilon\to 0} \int_{\{|\nabla u_\varepsilon| \leq M\}\setminus D_\varepsilon} V(e(u_\varepsilon))$$

for each M. So the claim follows as before since $|D_\varepsilon| \to 0$ as $\varepsilon \to 0$.

A Appendix: Quasiconvexity and Γ-convergence

In this appendix we collect some notation used throughout the rest of the paper and we recall some results on quasiconvexity and Γ-convergence that are needed in the main part of this paper.

A.1 Quasiconvexity

Recall that a function $f : \mathbb{R}^{m\times n} \to \mathbb{R}$ on the $m\times n$ matrices is called quasiconvex, if

$$f(F) \leq \fint_U f(F + \nabla u)$$

for all $u \in C_c^\infty(U;\mathbb{R}^m)$ (or, equivalently, $u \in W_0^{1,\infty}(U;\mathbb{R}^m)$), where U is a domain in $\mathbb{R}^n$ and $⨍_U$ denotes the averaged integral $\frac{1}{|U|}\int_U$. The quasiconvexification (or quasiconvex hull) of a function $f : \mathbb{R}^{m\times n} \to \mathbb{R}$ is given by

$$Qf := \sup\{g : g \le f, g \text{ is quasiconvex}\}.$$

It turns out that

$$Qf(F) = \inf_{u\in C_c^\infty(U;\mathbb{R}^d)} ⨍_U f(F + \nabla u)$$

(see, e.g., [9]).

If $m = n = d$ and f is only defined on the space $\mathbb{R}^{d\times d}_{\text{sym}}$ of symmetric matrices, following [24] we call f quasiconvex on linear strains if, for each $F \in \mathbb{R}^{d\times d}_{\text{sym}}$,

$$f(F) \le ⨍_U f\left(F + e(u)\right)$$

for all $u \in C_c^\infty(U;\mathbb{R}^d)$ (resp., $u \in W_0^{1,\infty}(U;\mathbb{R}^d)$), where $e(u) = \frac{1}{2}((\nabla u)^T + \nabla u)$ denotes the linear strain. (I.e., if its extension $F \mapsto f(\frac{F^T+F}{2})$ to all of $\mathbb{R}^{d\times d}$ is quasiconvex in the usual sense.) Correspondingly, the quasiconvex hull on linear strains of f is

$$\begin{aligned} Q^e f(F) &= \sup\{g(F) : g \le f, g \text{ is quasiconvex on linear strains}\} \\ &= \inf_{u\in C_c^\infty(U;\mathbb{R}^d)} ⨍_U f\left(F + e(u)\right). \end{aligned}$$

The quasiconvex hull of a closed subset K of $\mathbb{R}^{m\times n}$ is defined as

$$QK = \{F \in \mathbb{R}^{m\times n} : f(F) \le \sup_{G\in K} f(G) \text{ for all quasiconvex } f\}.$$

If $\mathcal{Q}_p$, $1 \le p < \infty$, denotes the set of non-negative quasiconvex functions f satisfying a p-growth estimate $0 \le f(F) \le C(1 + |F|^p)$ for some $C = C(f)$, then the strong p-quasiconvex hull of K is defined to be

$$\mathbf{Q}_p K = \{F \in \mathbb{R}^{m\times n} : f(F) \le \sup_{G\in K} f(G) \text{ for all } f \in \mathcal{Q}_p\}.$$

The weak p-quasiconvex hull of K is defined by

$$Q_p K = \{F \in \mathbb{R}^{m\times n} : Q\text{dist}^p(F, K) = 0\},$$

where $F \mapsto Q\text{dist}^p(F, K)$ is the quasiconvexification of the p-distance function $F \mapsto \text{dist}^p(F, K)$ (see, e.g., [26]).

Analogously we define the spaces $\mathcal{Q}_p^e$ and the hulls $\mathbf{Q}_p^e K$ and $Q_p^e K$ by replacing $\mathbb{R}^{m\times n}$ by $\mathbb{R}^{d\times d}_{\text{sym}}$, 'quasiconvex' by 'quasiconvex on linear strains' and $Q\text{dist}^p$ by $Q^e\text{dist}^p$ (cf. [24]).

Obviously we have the inclusions

$$QK \subset \mathbf{Q}_p K \subset Q_p K \quad \text{and} \quad Q^e K \subset \mathbf{Q}_p^e K \subset Q_p^e K.$$

The following result is due to Zhang (see Prop. 2.5 in [25] and Theorem 4 in [24]).

Theorem A.1 *Suppose $K \subset \mathbb{R}^{m\times n}$, resp., $\mathbb{R}^{d\times d}_{\rm sym}$, is compact, then*

$$QK = \mathbf{Q}_p K = Q_p K, \quad \textit{resp.}, \quad \mathbf{Q}^e_p K = Q^e_p K = Q^e_1 K$$

for all $1 \le p < \infty$.

Remark A.2 It appears to be unknown if, for compact $K \subset \mathbb{R}^{d\times d}_{\rm sym}$, the weak and strong p-quasiconvex envelopes on linear strains are in fact equal to $Q^e K$.

We close this paragraph with the following approximation result for quasiconvex functions due to Kristensen (cf. [19]).

Lemma A.3 *Suppose $\psi : \mathbb{R}^{m\times n} \to \mathbb{R}$ is a quasiconvex function satisfying the growth condition*

$$\alpha|F|^p - \beta \le \psi(F) \le \beta|F|^p + \beta \quad \forall\, F \in \mathbb{R}^{m\times n}$$

for some $p > 1$ and constants $\alpha, \beta > 0$. Then there are quasiconvex $\psi_j : \mathbb{R}^{m\times n} \to \mathbb{R}$ such that $\psi_j \le \psi_{j+1}$, $\psi_j \to \psi$ pointwise as $j \to \infty$ and there exist $a_j, r_j > 0$, $b_j \in \mathbb{R}$ such that

$$\psi_j(F) = \psi_j^{**}(F) = a_j|F| + b_j \quad \textit{if} \quad |F| \ge r_j,$$

where ψ_j^{**} denotes the convex envelope of ψ_j.

A.2 Γ-convergence and relaxation

A sequence of functionals $\mathcal{F}_\varepsilon : M \to [-\infty, \infty]$ on a metric space M is said to Γ-converge to some functional $\mathcal{F} : M \to [-\infty, \infty]$ if the following conditions are satisfied:

(i) ('lim inf-inequality') If $u_\varepsilon \to u$ in M, then

$$\liminf \mathcal{F}_\varepsilon(u_\varepsilon) \ge \mathcal{F}(u).$$

(ii) ('recovery sequence') For every $u \in M$ there exists a sequence $u_\varepsilon \to u$ in M such that

$$\lim \mathcal{F}_\varepsilon(u_\varepsilon) = \mathcal{F}(u).$$

(See [10] for general information about Γ-convergence).

We recall some standard results on the relaxation of integral functionals (see, e.g., [8, 1]) in a form most suitable for our purposes. Let Ω be an open subset of $\mathbb{R}^n$ and suppose $f : \Omega \times \mathbb{R}^{m\times n} \to \mathbb{R}$ is a Carathéodory function, i.e., measurable in the first variable and continuous in the second, satisfying a standard p-growth estimate

$$\alpha|F|^p \le f(x, F) \le \beta(1 + |F|^p) \qquad \forall x \in \Omega, F \in \mathbb{R}^{m\times n}$$

with $p > 1$ for some $\alpha, \beta > 0$. Define the functional $\mathcal{F} : W^{1,p}(\Omega, \mathbb{R}^m) \to [0, \infty)$ by

$$\mathcal{F}(u) = \int_\Omega f(x, \nabla u).$$

If $\varphi \in W^{1,p}(\Omega, \mathbb{R}^m)$ is some boundary data, we define the functional $\mathcal{F}^\varphi$: $W^{1,p}(\Omega, \mathbb{R}^m) \to [0, \infty]$ by

$$\mathcal{F}^\varphi(u) = \begin{cases} \mathcal{F}(u), & \text{if } u - \varphi \in W_0^{1,p}(\Omega, \mathbb{R}^m), \\ \infty, & \text{otherwise.} \end{cases}$$

Theorem A.4 *The sequentially weakly lower semicontinuous envelopes of $\mathcal{F}$ and $\mathcal{F}^\varphi$ on $W^{1,p}(\Omega, \mathbb{R}^m)$ are given, respectively, by*

$$\mathcal{F}^{\mathrm{rel}}(u) = \int_\Omega Qf(x, \nabla u) \quad \text{and} \quad (\mathcal{F}^\varphi)^{\mathrm{rel}}(u) = \begin{cases} \mathcal{F}^{\mathrm{rel}}(u), & \text{if } u - \varphi \in W_0^{1,p}, \\ \infty, & \text{otherwise.} \end{cases}$$

Here $Qf(x, \cdot)$ is a Carathéodory function equal to the quasiconvex envelope of $f(x, \cdot)$ for a.e. x.

In fact, the constant sequences $\mathcal{F}_\varepsilon \equiv \mathcal{F}$ and $\mathcal{F}_\varepsilon^\varphi \equiv \mathcal{F}^\varphi$ Γ-converge to $\mathcal{F}^{\mathrm{rel}}$ respectively $(\mathcal{F}^\varphi)^{\mathrm{rel}}$ with respect to the strong L^p-topology and the weak $W^{1,p}$-topology.

Acknowledgments

I would like to thank Kaushik Bhattacharya for proposing me the problem and for interesting discussions on the subject. I am also grateful to an unknown referee for helpful comments on the first version of this paper.

References

[1] E. Acerbi, N. Fusco. *Semicontinuity problems in the calculus of variations.* Archive Ration. Mech. Anal. **86** (1984), 125–145.

[2] J. Ball. *A version of the fundamental theorem for Young measures.* In: *PDEs and continuum models of phase transition,* Lecture Notes in Phys. **344** Springer Verlag, Berlin · Heidelberg 1989.

[3] J. Ball. *Mathematical models of martensitic microstructure.* Materials Science Eng. A **378** (2004), 61–69.

[4] J. Ball, R. D. James. *Fine phase mixtures as minimizers of energy.* Archive Ration. Mech. Anal. **100** (19987), 13–52.

[5] J. Ball, R. D. James. *Proposed experimental tests of a theory of fine microstructure and the two-well problem.* Phil. Trans. R. Soc. Lond. A **139** (1992), 389–450.

[6] K. Bhattacharya. *Microstructure of Martensite: Why it forms and how it gives rise to the shape memory effect.* Oxford University Press, Oxford 2003.

[7] K. Bhattacharya. *Comparison of the geometrically nonlinear and linear theories of martensitic transformation.* Continuum Mech. Thermodyn. **5** (1993), 205–242.

[8] A. Braides, A. Defrancheschi. *Homogenization of Multiple Integrals.* Clarendon Press, Oxford 1998.

[9] B. Dacorogna. *Direct Methods in the Calculus of Variations.* Springer-Verlag, Berlin · Heidelberg · New York 1989.

[10] G. Dal Maso. *An Introduction to* Γ*-convergence.* Birkhäuser, Boston · Basel · Berlin 1993.

[11] G. Dal Maso, M. Negri, D. Percivale. *Linearized elasticity as* Γ*-limit of finite elasticity.* Set-valued Anal. **10** (2002), 165–183.

[12] L. C. Evans. *Weak Convergence Methods for Nonlinear Partial Differential Equations.* CBMS Series in Mathematics **74**, AMS, Providence, RI 1990.

[13] G. Friesecke, R. D. James, S Müller. *A theorem on geometric rigidity and the derivation of nonlinear plate theory from three-dimensional elasticity.* Comm. Pure Appl. Math. **55** (2002), 1461–1506.

[14] G. Friesecke, R. D. James, S Müller. *A hierarchy of plate models derived from nonlinear elasticity by* Γ*-convergence.* Archive Ration. Mech. Anal. **180** (2006), 183–236.

[15] A. G. Khachaturyan. *Some questions concerning the theory of phase transformations in solids.* Sov. Phys. Solid State **8** (1967), 2163–2168.

[16] A. G. Khachaturyan. *Theory of structural transformations in solids.* John Wiley and sons, New York 1983.

[17] A. G. Khachaturyan, G. A. Shatalov. *Theory of macroscopic periodicity for a phase transition in the solid state.* Soviet Phys. JETP **29** (1969), 557–561.

[18] R. V. Kohn. *The relaxation of a double-well energy.* Continuum Mech. Thermodyn. **3** (1991), 193–236.

[19] J. Kristensen. *Lower semicontinuity in spaces of weakly differentiable functions.* Math. Ann. **313** (1999), 653–710.

[20] S. Müller. *Variational models for Microstructure and Phase Transitions.* In: *Calculus of Variations and Geometric Evolution Problems*, Springer Verlag, Berlin · Heidelberg · New York 1999.

[21] A. L. Roitburd. Kristallografiya **12** (1967), 567 ff. In Russian.

[22] A. L. Roitburd. *Martensitic transformation as a typical phase transformation in solids.* In: *Solid state physics* **33** (1978), 317–390.

[23] R. Temam. *Mathematical Problems in Plasticity.* Gauthier-Villars, Paris 1985.

[24] K. Zhang. *An approximation theorem for sequences of linear strains and its applications.* ESAIM Control Optim. Calc. Var. **10** (2004), 224–242.

[25] K. Zhang. *On the structure of quasiconvex hulls.* Annales de l'I.H.P. – Anal. non lin. **15** (1998), 663–686.

[26] K. Zhang. *Rank-one connections at infinity and quasiconvex hulls.* J. Convex Anal. **7** (2000), 19–45.