COMPUTATION OF THE LAVRENTIEV PHENOMENON*

CARSTEN CARSTENSEN AND CHRISTOPH ORTNER

Abstract. Amongst the more exciting phenomena in the field of nonlinear partial differential equations is the Lavrentiev phenomenon which occurs in the calculus of variations. We prove that a conforming finite element method fails if, and only if, the Lavrentiev phenomenon is present. Consequently, nonstandard finite element methods have to be designed for the detection of the Lavrentiev phenomenon in the computational calculus of variations.

1. Introduction

The calculus of variations is concerned with the minimisation problem

$$\inf E(\mathcal{A}_1) := \inf_{v\in\mathcal{A}_1} E(v), \tag{1.1}$$

where $E : \mathcal{A}_1 \to \mathbb{R} \cup \{+\infty\}$, and where $\mathcal{A}_1$ (or more generally $\mathcal{A}_p$) is the first-order Sobolev space

$$\mathcal{A}_p := W_0^{1,p}(\Omega;\mathbb{R}^m) = \{v \in W^{1,p}(\Omega)^m : v|_{\partial\Omega} = 0\},$$

based on a bounded Lipschitz domain $\Omega \subset \mathbb{R}^n$ with piecewise hyperplanar boundary $\partial\Omega$.

We shall assume throughout that E is *proper* on $\mathcal{A}_\infty$, i.e., there exists $v \in \mathcal{A}_\infty$ so that $E(v) < +\infty$. In particular, $\mathcal{A}_\infty \subset \mathcal{A}_1$ always implies

$$-\infty \le \inf E(\mathcal{A}_1) \le \inf E(\mathcal{A}_\infty) < +\infty.$$

The *Lavrentiev phenomenon* is the surprising property of E that, in some cases,

$$\inf E(\mathcal{A}_1) \neq \inf E(\mathcal{A}_\infty), \tag{L}$$

named after its first occurence in the literature [17]. Other well-known examples are the one-dimensional examples of Mania [21] and of Ball and Mizel [5, 4], or the convex example of Foss, Hrusa and Mizel [15].

Date: March 2, 2009.

1991 *Mathematics Subject Classification.* 65N12,65N25,65N30,65N50.

Key words and phrases. Lavrentiev phenomenon, finite element method, computational calculus of variations.

*Supported by the DFG Research Center MATHEON "Mathematics for key technologies" in Berlin, the Hausdorff Institute of Mathematics at Bonn, Germany, and EPSRC Critical Mass programme 'New Frontiers in the Mathematics of Solids' (OxPDE), Oxford, UK.

In nonlinear elasticity, the Lavrentiev phenomenon is closely related to the occurence of cavitation [2].

For the conforming finite element discretization of (1.1) assume we are given a family of finite element spaces $(V_\ell)_{\ell\in\mathbb{N}_0}$ to solve

$$\inf E(V_\ell) := \inf_{v_\ell\in V_\ell} E(v_\ell). \tag{1.2}$$

Since conforming Finite element function are always Lipschitz continuous, any finite element space V_ℓ is contained in $\mathcal{A}_\infty$, and hence standard finite element methods cannot compute singular minimisers,

$$\inf E(\mathcal{A}_1) < \inf E(\mathcal{A}_\infty) \le \inf E(V_\ell).$$

A finite element method is given by sequence of finite element spaces

$$V_0, V_1, V_2, \dots \subset \overline{\cup_{\ell=0}^{\infty} V_\ell} \subseteq \mathcal{A}_\infty.$$

The respective infimal energy densities are possibly convergent towards some limit

$$\inf E(\mathcal{A}_\infty) \le \liminf_{\ell\to\infty} \inf E(V_\ell).$$

We say that the finite element method (FEM) is convergent if E and the sequence of discrete subspaces $V_0, V_1, V_2, \dots$ allow for

$$\inf E(\mathcal{A}_1) = \lim_{\ell\to\infty} \inf E(V_\ell). \tag{C}$$

Therein, the convergence of the entire sequence (not merely of *some* subsequence but for all subsequences) is part of the statement as well as the equality of that limit to $\inf E(\mathcal{A}_1)$.

It is obvious that (C) implies that (L) is false. Section 2 below provides a general framework that allows for the converse and establishes

$$\text{(C)} \quad \Longleftrightarrow \quad \text{NOT (L)}$$

under natural assumptions on the energy density.

A consequence of this equivalence is that conforming finite element methods are inappropriate tools for detecting the singular minimisers associated to the Lavrentiev phenomenon (L).

Several classes of numerical schemes have been introduced in the literature to allow for a numerical detection of (L), including the penalty method of Ball and Knowles [3, 16] and its extension to polyconvex integrands by Negron–Marrero [22], the element-removal method of Li [18, 19], and the truncation method of Bai and Li [1].

Section 3 introduces a penalty method, which is related to the methods of Ball & Knowles [3] and of Negron–Marrero [22] but, as we believe, is easier to use and more efficient in practise. We establish general convergence results in Sections 4 and 5. In Section 6 we discuss some connections of our results with the theory of Γ-convergence. We conclude with the description of a practical implementation and several computational examples.

2. Finite Element Failure is Equivalent to the Lavrentiev Phenomenon

This section is devoted to the equivalence of (C) and NOT (L) in a general setting which is entirely free of growth conditions and notions of convexity. However, there are standard conditions on the uniform convergence of the mesh-size towards zero in the finite element methods and global continuity of an energy density.

Suppose that $\mathcal{T}_1, \mathcal{T}_2, \mathcal{T}_3, \dots$ denotes a sequence of regular triangulations into simplices of a Lipschitz domain $\Omega \subset \mathbb{R}^n$ with piecewise flat boundary $\partial\Omega$ which is perfectly matched by the triangulations. Suppose that the triangulation is shape regular in the sense that the largest n dimensional ball inside each simplex T and the smallest ball outside have uniformly bounded ratios: There exists a universal positive constant $C_{\text{shaperegular}}$ which does not depend on T or ℓ such that one finds midpoints m_T and M_T, and radii r_T and R_T, satisfying

$$B(m_T, r_T) \subset T \subset B(M_T, R_T) \quad \text{and} \quad R_T/r_T \leq C_{\text{shaperegular}}.$$

Convergence of the mesh-size, written $h_\ell \to 0$, is understood in the sense that

$$\lim_{\ell\to\infty} \max_{T\in\mathcal{T}_\ell} R_T = 0.$$

The finite-dimensional space V_ℓ of piecewise affine finite element functions (piecewise with respect to the triangulation $\mathcal{T}_\ell$)

$$V_\ell := \{v_\ell \in C_0(\Omega; \mathbb{R}^m) : \forall T \in \mathcal{T}_\ell,\ v_\ell|_T \text{ affine } \}$$

belongs to $\mathcal{A}_\infty$.

Let the energy density $W : \overline{\Omega} \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}$ be continuous and define the energy

$$E(v) := \int_\Omega W(x, v(x), Dv(x))dx$$

for all $v \in \mathcal{A}_\infty$. In fact, if v is Lipschitz continuous, then the set of triples $\{(x, v(x), Dv(x)) : x \in \bar{\Omega}\}$ as well as the set $\{W(x, v(x), Dv(x)) : x \in \bar{\Omega}\}$ are contained in compact sets. Consequently, $E(v) \in \mathbb{R}$ and $E : \mathcal{A}_\infty \to \mathbb{R}$ is well defined. For an arbitrary function $v \in \mathcal{A}_1$, this is no longer clear. Throughout this section, we simply assume that

$$E : \mathcal{A}_1 \to \mathbb{R} \cup \{+\infty\}$$

is some extension of $E|_{\mathcal{A}_\infty}$. In applications, this may be guaranteed by growth control from below and we refer to the literature (e.g., [11]) for this well-understood argument in the direct method of the calculus of variations. The question of attainment of a global or discrete minimum is irrelevant here and bypassed by a consequent discussion of infima instead of minima, e.g., for any $\ell = 0, 1, 2, \dots$,

$$E_\ell := \inf E(V_\ell) := \inf_{v_\ell \in V_\ell} E(v_\ell) \in \mathbb{R} \cup \{\pm\infty\}.$$

We emphasize that there is no nestedness assumption on the finite element spaces and so the convergence of the infimal energies E_ℓ does *not* follow automatically. In fact, it is stated in the theorem as a conclusion.

Theorem 2.1 (Finite Element Failure $\Leftrightarrow$ Lavrentiev Phenomenon)**.** *If $W : \bar\Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}$ is continuous, then $\lim_{\ell=\infty} E_\ell = \inf E(\mathcal{A}_\infty)$, and in particular,*

$$\lim_{\ell\to\infty} E_\ell = \inf E(\mathcal{A}_1) \quad \Longleftrightarrow \quad \inf E(\mathcal{A}_1) = \inf E(\mathcal{A}_\infty).$$

The direction $\Longrightarrow$ in the theorem's assertion is obvious from the introduction and $V_\ell \subset \mathcal{A}_\infty$:

$$\inf E(\mathcal{A}_\infty) \le \liminf_{\ell\to\infty} E_\ell = \inf E(\mathcal{A}_1) \le \inf E(\mathcal{A}_\infty).$$

The converse $\Longleftarrow$ requires a density argument stated in terms of the nodal interpolation operator. Given a continuous function $v : \overline{\Omega} \to \mathbb{R}^m$ and a triangulation $\mathcal{T}_\ell$, the nodal interpolation $v_\ell := I_\ell v$ of v is defined on each simplex $T \in \mathcal{T}_\ell$ with vertices $z_1, \dots, z_{n+1}$ through linear interpolation of the values $v(z_j)$ at the $n+1$ vertices z_j.

Lemma 2.2. *There exists a constant C which depends only $C_{\rm shaperegular}$, such that, for any $v \in W^{1,\infty}(\Omega; \mathbb{R}^m)$, the piecewise affine function $v_\ell = I_\ell v$ satisfies*

$$\|v_\ell\|_{W^{1,\infty}(\Omega)} \le C \|v\|_{W^{1,\infty}(\Omega)} \quad \textit{for all } \ell = 0, 1, 2, \dots.$$

Moreover, $v_\ell \to v$ in $L^\infty(\Omega; \mathbb{R}^m)$, and $Dv_\ell \to Dv$ pointwise a.e. in Ω, as $\ell \to \infty$.

Proof. The stability of the nodal interpolation operator, as well as the convergence in the L^∞-norm, are standard results and can, for example, be found in [9].

The theorem of Rademacher implies that, for almost all x in some simplex T, that $Dv(x)$ exists in the sense of a Fréchet derivative, i.e.,

$$Dv(x)(y-x) = v(y) - v(x) + o(|x-y|),$$

for some function $y \mapsto o(|x-y|)$ with

$$\lim_{y\to x} o(|x-y|)/|x-y| = 0.$$

Fix some $x \in \Omega$ so that, for any $\ell \in \mathbb{N}_0$, x lies the interior of an element $T \in \mathcal{T}_\ell$, then

$$\begin{aligned} &Dv(x)(z_j - z_k) = v_\ell(z_j) - v_\ell(z_k) + o(|x-z_j|) + o(|x-z_k|) \\ &= Dv_\ell(x)(z_j - z_k) + o(|x-z_j|) + o(|x-z_k|) \quad \text{for all } j,k = 1,\dots,n+1. \end{aligned}$$

Since the tangential vectors are linearly independent and the interior angles do not deteriorate, we have

$$\sup_{j,k=1,\dots,n+1} (Dv(x) - Dv_\ell(x))(z_j - z_k) \ge c|Dv(x) - Dv_\ell(x)| r_T,$$

where c depends only on $C_{\text{shaperegular}}$. It now follows easily that

$$\lim_{\ell\to\infty} |Dv(x) - Dv_\ell(x)| = 0. \qquad \square$$

Proof of Theorem 2.1. Given $v \in \mathcal{A}_\infty$ and its nodal interpolant $v_\ell := I_\ell v$ for all $\ell \in \mathbb{N}_0$, the previous lemma shows that

$$\lim_{\ell\to\infty} (v_\ell(x), Dv_\ell(x)) = (v(x), Dv(x)) \in \mathbb{R}^m \times \mathbb{R}^{m\times n} \quad \text{for a.e. } x \in \Omega.$$

Since W is continuous, this yields pointwise convergence of the energy density

$$\lim_{\ell\to\infty} W(x, v_\ell(x), Dv_\ell(x)) = W(x, v(x), Dv(x)) \quad \text{for a.e. } x \in \Omega.$$

Furthermore, the boundedness of v_ℓ in $W^{1,\infty}(\Omega)$ and the assumption that W is continuous implies that $W(x, v_\ell(x), Dv_\ell(x))$ is bounded uniformly in x and ℓ. Consequently, Lebesgue's dominated convergence theorem shows

$$\lim_{\ell\to\infty} \int_\Omega W(x, v_\ell(x), Dv_\ell(x))dx = \int_\Omega W(x, v(x), Dv(x))dx = E(v).$$

Therefore,

$$\inf E(\mathcal{A}_\infty) \le \liminf_{\ell\to\infty} E_\ell \le \limsup_{\ell\to\infty} E_\ell \le \lim_{\ell\to\infty} E(v_\ell) = E(v).$$

Since v was an arbitrary element in $\mathcal{A}_\infty$, we deduce

$$\liminf_{\ell\to\infty} E_\ell = \limsup_{\ell\to\infty} E_\ell = \inf E(\mathcal{A}_\infty).$$

In particular, we can conclude that $\lim_\ell E_\ell = \inf E(\mathcal{A}_\infty)$ exists. From this, the assertion of Theorem 2.1 follows immediately. $\square$

3. Penalisation and Discrete Scheme

In many examples there exists a *coupling function*

$$\gamma : \Omega \times \mathbb{R}^m \times R^{m\times n} \to \mathbb{M}$$

where $\mathbb{M} := \mathbb{R}^\mu \equiv \mathbb{R}^{M\times N}$ is a space of matrices, and an *extended energy density*

$$\phi : \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \times \mathbb{M} \to \mathbb{R}$$

such that the energy density W is given by

$$W(x, v, F) := \phi(x, v, F, \gamma(x, v, F))$$

for all $x \in \Omega$, $v \in \mathbb{R}^m$, $F \in \mathbb{R}^{m\times n}$. In this case, we also define

$$\Phi(v, \eta) := \int_\Omega \phi(x, v(x), Dv(x), \eta(x))dx \quad \text{for } (v, \eta) \in \mathcal{A}_1 \times L^1(\Omega; \mathbb{M})$$

and, with the abbreviation $\gamma(\cdot, v, Dv)(x) := \gamma(x, v(x), Dv(x))$ for $x \in \Omega$, we observe that

$$E(v) = \Phi(v, \gamma(\cdot, v, Dv)). \tag{3.1}$$

Example 3.1 (Singular Minimiser). Given $W : \Omega \times \mathbb{R}^m \times \mathbb{R}^{m\times n} \to \mathbb{R}$ let $\mathbb{M} = \mathbb{R}^{m\times n}$ and consider

$$\phi(x, v, F, \eta) := W(x, v, \eta) \quad \text{and} \quad \gamma(x, v, F) := F.$$

Example 3.2 (Polyconvex Materials). By definition, at almost all material points $x \in \Omega$ and all $v \in \mathbb{R}^m$, a polyconvex energy density $W(x, v, \cdot) : \mathbb{R}^{m\times n} \to \mathbb{R}$ can be written in the form

$$W(x, v, F) = \phi(x, v, \gamma(F)),$$

where ϕ is convex in its third component (with x, v fixed), and $\gamma : \mathbb{R}^{m\times n} \to \mathbb{M}$ maps a deformation gradient F to the vector of minors (sub-determinants) of F and $\mathbb{M}$ is the space of all those minors (e.g. $\mathbb{M} = \mathbb{R}^{19}$ for $m = 3 = n$).

On the continuous level, this looks as a trivial complication of the formulation but the point is that the discretisation relaxes the condition

$$\eta = \gamma(x, v, F) \quad \text{in} \quad W(x, v, F) = \phi(x, v, F, \eta).$$

Since the immediate substitution cannot detect singular minimisers with a Lavrentiev phenomenon the 'coupling' $\eta = \gamma(x, v(x), Dv(x))$ will be weakened by introducing a penalty functional,

$$\Psi_\ell : L^1(\Omega; \mathbb{M}) \times L^1(\Omega; \mathbb{M}) \to \mathbb{R} \cup \{+\infty\},$$

which is written, via some density $\psi_\ell : \Omega \times \mathbb{M} \times \mathbb{M} \to [0, \infty]$,

$$\Psi_\ell(\eta, \zeta) := \int_\Omega \psi_\ell(x, \eta(x), \zeta(x))dx \qquad \text{for } \eta, \zeta \in L^1(\Omega; \mathbb{M}).$$

The proposed discrete minimisation problem reads: Minimise the discrete energy

$$E_\ell(v, \eta) := \Phi(v, \eta) + \Psi_\ell(\eta, \gamma(\cdot, v, Dv))$$

over $(v, \eta) \in V_\ell \times Y_\ell$ where V_ℓ and Y_ℓ are suitable finite element spaces.

Example 3.3 (Penalisation). A typical class of distance functionals is given for $1 \le p < \infty$ and positive parameters ϵ_ℓ which possibly depends on the position x in the spatial domain (e.g., piecewise constant with respect to the triangulation $\mathcal{T}_\ell$) and

$$\psi_\ell(x, \eta, \zeta) := \epsilon_\ell^{-1} |\eta - \zeta|^p$$

for all $x \in \Omega$ and $\eta, \zeta \in \mathbb{M}$.

4. Polyconvex Energy Densities

An important class of energy functionals, especially in the field of nonlinear elasticity, are those where the stored energy density is polyconvex. As a prototypical model problem, we consider the stored energy density

$$W(x, u, F) = \phi(x, F, \det F) - f(x) \cdot u, \tag{4.1}$$

where $f \in L^q(\Omega)^n$ for some $q > 1$, and $\phi : \Omega \times \mathbb{R}^{n\times n} \times \mathbb{R} \to [0, +\infty]$. We assume throughout this section that ϕ satisfies

$$
\begin{aligned}
&|F|^n + \Gamma(\eta) \lesssim \phi(x, F, \eta) \lesssim 1 + |F|^n + \Gamma(\eta), \text{ and} \\
&\phi(x, \cdot, \cdot) \text{ is convex and l.s.c. in } \mathbb{R}^{n\times n} \times \mathbb{R} \quad \text{for a.a. } x \in \Omega,
\end{aligned}
\tag{4.2}
$$

where $\Gamma : \mathbb{R} \to [0, +\infty]$ is convex and has superlinear growth, i.e., $\liminf_{|s|\to\infty} \Gamma(s)/s = +\infty$ [6, 10]. We remark, that the growth condition $|F|^n + \Gamma(g)$ may be replaced by $|F|^p$ for some $p > n$. In fact, the latter implies the former.

The space of admissible functions is defined as

$$V = u_D + W_0^{1,n}(\Omega)^n,$$

where $u_D \in W^{1,n}(\Omega)^n$ and $E(u_D) < +\infty$. Under these conditions the minimization problem

$$u \in \operatorname{argmin} E(V) \tag{4.3}$$

has at least one solution [11, Theorem 2.10].

To discretize the problem, we fix a sequence $u_{D,\ell} \in \mathrm{P}^1(\mathcal{T}_\ell)^n$ such that $u_{D,\ell} \to u_D$ strongly in $W^{1,n}(\Omega)^n$, and we discretize V and $L^1(\Omega)$, respectively, by

$$V_\ell = u_{D,\ell} + \mathrm{P}^1(\mathcal{T}_\ell)^n, \qquad \text{and} \quad Y_\ell = \mathrm{P}^0(\mathcal{T}_\ell).$$

Further, we assume that we have a *penalty functional* $\Psi : L^1(\Omega)^2 \to [0, +\infty]$ such that, for all sequences (η_ℓ) and (ζ_ℓ),

$$\Psi(\eta_\ell, \zeta_\ell) \to 0 \quad \Leftrightarrow \quad \|\eta_\ell - \zeta_\ell\|_{L^1} \to 0. \tag{4.4}$$

Given a sequence $\varepsilon_\ell \searrow 0$, we discretize (4.3) by

$$(u_\ell, \eta_\ell) \in \operatorname{argmin} E_\ell(V_\ell, Y_\ell),$$

where

$$
\begin{aligned}
E_\ell(v_\ell, \eta_\ell) &= \Phi(v_\ell, \eta_\ell) + \varepsilon_\ell^{-1}\Psi(\det Dv_\ell, \eta_\ell) \\
&= \int_\Omega \big(\phi(x, Dv_\ell, \eta_\ell) - f \cdot v_\ell\big) dx + \varepsilon_\ell^{-1}\Psi(\det Dv_\ell, \eta_\ell).
\end{aligned}
$$

Theorem 4.1. *Assume that* (4.1), (4.2), *and* (4.4) *hold. Then there exists a sequence* $\varepsilon_\ell \searrow 0$ *such that, for any sequence* $(u_\ell, \xi_\ell) \in \operatorname{argmin} E_\ell(V_\ell, Y_\ell)$, *we have*

$$\Phi(u_\ell, \xi_\ell) \to \inf E(V) \text{ and } \varepsilon_\ell^{-1}\Psi(\det Du_\ell, \xi_\ell) \to 0.$$

Moreover, the family $\{u_\ell; \ell \in \mathbb{N}\}$ *is precompact in the weak topology of* $W^{1,n}(\Omega)^n$, *and each accumulation point* u *is a minimizer of* E *in* V. *In particular, there exists a subsequence* $\ell_k \nearrow \infty$ *such that*

$$
\begin{aligned}
u_{\ell_k} &\rightharpoonup u \qquad \text{weakly in } W^{1,n}(\Omega)^n, \\
\xi_{\ell_k} &\rightharpoonup \det Du \qquad \text{weakly in } L^1(\Omega),
\end{aligned}
$$

where u *solves* (4.3).

The proof of Theorem 4.1 is contained in the following three Lemmas.

Lemma 4.2. *Assume that* (4.1), (4.2), *and* (4.4) *hold. For every* $u \in V$ *there exists a sequence* $(u_\ell, \xi_\ell) \in V_\ell \times Y_\ell$ *such that*

$$u_\ell \to u \qquad \text{strongly in } W^{1,n}(\Omega)^n, \tag{4.5}$$

$$\lim_{\ell\to\infty} \Psi(\det Du_\ell, \xi_\ell) = 0, \ \text{and} \tag{4.6}$$

$$\lim_{\ell\to\infty} \Phi(u_\ell, \xi_\ell) = \Phi(u, \det Du) = E(u). \tag{4.7}$$

Proof. We assume without loss of generality that $E(u) < \infty$. We take an arbitrary sequence $u_\ell \in V_\ell$ such that $u_\ell \to u$ strongly in $W^{1,n}(\Omega)^n$ which also implies $\det Du_\ell \to \det Du$ strongly in $L^1(\Omega)$. The variable $\xi_\ell \in Y_\ell$ is defined as

$$\xi_\ell(x) = |T|^{-1} \int_T \det Du \, dx \qquad x \in T, T \in \mathcal{T}_\ell.$$

It follows that $\xi_\ell \to \det Du$ strongly in $L^1(\Omega)$, and in particular that $\Psi(\det Du_\ell, \xi_\ell) \to 0$. Thus, we have shown (4.5) and (4.6).

To prove (4.7) we first use Jensen's inequality to estimate, for $x \in T$,

$$\Gamma(\xi_\ell(x)) = \Gamma\Big(|T|^{-1} \int_T \det Du \, dx\Big) \le |T|^{-1} \int_T \Gamma(\det Du) dx =: \Gamma_\ell(x),$$

i.e., Γ_ℓ is a majorant for $\Gamma(g_\ell)$. From its definition, and since $\Gamma(\det Du) \in L^1(\Omega)$, it follows immediately that $\Gamma_\ell \to \Gamma(\det Du)$ strongly in $L^1(\Omega)$.

Hence, we obtain that

$$\big|\phi(x, Du_\ell, \xi_\ell)\big| \lesssim 1 + |Du_\ell|^n + \Gamma_\ell =: a_\ell,$$

where a_ℓ is strongly convergent in $L^1(\Omega)$. For any subsequence we can extract a further subsequence such that $(Du_\ell, \xi_\ell) \to (Du, \xi)$ pointwise, and hence we can use a variant of Lebesgue's dominated convergence theorem [14, Sec. 1.3, Th. 4] to deduce (4.7). □

Lemma 4.3. *Assume that* (4.1), (4.2), *and* (4.4) *hold. There exists a sequence* $\varepsilon_\ell \searrow 0$ *such that*

$$\limsup_{\ell\to\infty} \min E_\ell(V_\ell, Y_\ell) \le \inf E(V). \tag{4.8}$$

Proof. Let $u \in \operatorname{argmin} E(V)$, and let (u_ℓ, ξ_ℓ) be the sequence constructed in Lemma 4.2. Then

$$\Psi(\det Du_\ell, \xi_\ell) \to 0,$$

and chosing $\varepsilon_\ell' = \Psi(\det Du_\ell, \xi_\ell)^{1/2}$, and $\varepsilon_\ell = \sqrt{\varepsilon_\ell'}$, we obtain

$$\limsup_{\ell\to\infty} \inf E_\ell(V_\ell, Y_\ell) \le \limsup_{\ell\to\infty} E_\ell(u_\ell, \xi_\ell) = E(u). \qquad \square$$

In the previous Lemma, we showed that it is possible to choose a sequence ε_ℓ such that the upper bound (4.8) holds. It remains to show that the limit is in fact equal.

Lemma 4.4. *Assume that* (4.1), (4.2), *and* (4.4) *hold. Suppose that a sequence* $\varepsilon_\ell \searrow 0$ *is fixed. Suppose furthermore that* $u_\ell \in V_\ell, \xi_\ell \in Y_\ell$ *such that*

$$\limsup_{\ell\to\infty} E_\ell(u_\ell, \xi_\ell) \le \inf E(V), \tag{4.9}$$

then there exists a subsequence $\ell_k \uparrow \infty$ *and* $u \in \operatorname{argmin} E(V)$ *such that*

$$\begin{aligned} u_{\ell_k} &\rightharpoonup u \qquad \textit{weakly in } W^{1,n}(\Omega)^n \\ \xi_{\ell_k} &\rightharpoonup \det Du \qquad \textit{weakly in } L^1(\Omega), \\ \Phi(u_\ell, \xi_\ell) &\to E(u), \quad \textit{and } \varepsilon_\ell^{-1}\Psi(\det Du_\ell, \xi_\ell) \to 0. \end{aligned}$$

Proof. It follows from (4.9) that $E_\ell(u_\ell, \xi_\ell)$ is bounded by some constant M. Using (4.2), we obtain

$$M \ge E_\ell(u_\ell, \xi_\ell) \gtrsim \|\nabla u_\ell\|^n_{L^n} - C\|u_\ell\|_{L^{q'}} + \int_\Omega \Gamma(\xi_\ell)dx + \varepsilon_\ell^{-1}\Psi(\det Du_\ell, \xi_\ell),$$

and since $L^{q'}(\Omega)^n$ is continuously embedded in $W^{1,n}(\Omega)^n$, there exists $M' \in \mathbb{R}$ such that

$$\|u_\ell\|^n_{W^{1,n}} + \int_\Omega \Gamma(\eta_\ell)dx + \varepsilon_\ell^{-1}\Psi(\det Du_\ell, \eta_\ell) \le M'.$$

We can therefore deduce the existence of a subsequence $\ell_k \nearrow \infty$, and of functions $u \in W^{1,n}(\Omega)^n$ and $\xi \in L^1(\Omega)$ such that

$$u_\ell \rightharpoonup u \text{ weakly in } W^{1,n}(\Omega)^n \text{ and } \xi_\ell \rightharpoonup \xi \text{ weakly in } L^1(\Omega).$$

(We note that the superlinear bound implies equi-integrability of the sequence (ξ_ℓ) which implies its precompactness in the weak tolopogy of $L^1(\Omega)$ [13, Cor. IV.8.11].)

Since $\det Du_\ell \rightharpoonup' \det Du$ in the sense of distributions [11, Sec. 4.2, Th. 2.6, (5)], and using (4.4), it follows that $\xi = \det Du$. Using sequential weak lower semi-continuity of energies with convex integrands [11, Sec. 3.3, Th. 3.4], we can estimate

$$\begin{aligned} E(u) \le\ & \liminf_{k\to\infty} \int_\Omega \Big(\phi(x, u_{\ell_k}, \xi_{\ell_k}) - f\cdot u_{\ell_k}\Big)dx \\ \le\ & \liminf_{k\to\infty} \int_\Omega \Big(\phi(x, u_{\ell_k}, \xi_{\ell_k}) - f\cdot u_{\ell_k}\Big)dx \\ & \qquad + \limsup_{k\to\infty} \varepsilon_{\ell_k}^{-1}\Gamma(\det Du_{\ell_k}, \xi_{\ell_k}) \\ \le\ & \limsup_{\ell\to\infty} E_\ell(u_\ell, \xi_\ell) \ \le\ \inf E(V). \end{aligned}$$

It follows therefore that $E(u) = \inf E(V)$. Moreover, this implies that all inequalities in the above chain of estimates must be equalities, and hence,

$$\limsup_{k\to\infty} \varepsilon_{\ell_k}^{-1}\Psi(\det Du_{\ell_k}, \xi_{\ell_k}) = 0.$$

Since the proof applies also if we begin with an arbitrary subsequence, it follows that the energy of the entire sequence converges in this sense. □

Remarks 4.5. 1. In practise, the condition that Ψ is continuous in the strong topology of $L^1(\Omega;\mathbb{R}^n)$ requires that

$$\Psi(\eta,\zeta) = \int_\Omega \psi(|\eta-\zeta|)dx,$$

where ψ has 1-growth at infinity. Typical penalty densities ψ are $\psi(t) = |t|$, or, if one prefers a smooth functional, $\psi(t) = (t^2+1)^{1/2} - 1$.

2. If ϕ satisfies a stronger growth condition, for example $\phi(x,F,g) \gtrsim |F|^p$ for some $p > n$ then this additional integrability allows us to use a penalty functional which is only continuous in $L^{p/n}(\Omega;\mathbb{R}^n)$.

3. We have only shown the existence of some sequence ε_ℓ for which we obtain convergence of the penalty method. We will show in Section 7 below, how this sequence can be constructed in practise.

4. More general polyconvex material models, where ϕ depends on all minors of the gradient, can be easily incorporated in our analysis. One would then have to decouple all minors which appear in the definition of the functional. Similar convergence can then be obtained whenever the growth conditions from above and below are the same, and are sufficiently strong so that the direct method can be applied.

5. Examples with Lavrentiev phenomenon

In many problems, decoupling the gradient is sufficient, and it is the goal of this section to make this precise. This is possible whenever W is convex in the third component, but it is also a useful approach if it is unclear which variable should be relaxed. We begin again with a more general discussion which we then make precise at two classes of problems, general one-dimensional functionals with continuous integrands, and higher-dimensional examples with mild v-dependence of the integrand.

We assume throughout that $W = \phi : \Omega\times\mathbb{R}^m\times\mathbb{R}^{m\times n} \to (-\infty,+\infty]$ is lower semi-continuous in all three variables, continuous at every point (x,v,η) where $\phi(x,v,\eta) < \infty$, and that it satisfies the lower bound

$$\phi(x,v,\eta) \gtrsim -1 - |v|^q, \tag{5.1}$$

where $1 \le q < n/(n-1)$ if $n \ge 2$ and $1 \le q < \infty$ if $n = 1$. This implies in particular that, for $v \in W^{1,1}(\Omega)^m$ and $\eta \in L^1(\Omega)^{m\times n}$, the functionals

$$\Phi(v,\eta) = \int_\Omega W(x,v,\eta)dx, \text{ and } E(v) = \Phi(v,Dv)$$

are well-defined in $(-\infty,+\infty]$. Let $u_D \in W^{1,1}(\Omega)^m$ such that $E(u_D) < \infty$ and define $V = u_D + W_0^{1,1}(\Omega)^m$.

We will assume that the penalty has 1-growth, namely, that there $\psi : \mathbb{R}^{m\times n} \to [0,\infty)$ such that functional of the form

$$\begin{aligned} &\Psi(\eta,\zeta) = \int_\Omega \psi(\eta-\zeta)dx, &&\text{for all } \eta,\zeta \in L^1(\Omega)^{m\times n}, \text{ and} \\ &|\eta| - 1 \lesssim \psi(\eta) \lesssim |\eta| + 1 &&\text{for all } \eta \in \mathbb{R}^{m\times n}. \end{aligned} \tag{5.2}$$

To discretize the problem of minimizing E over V, we take $u_{D,\ell} \in \mathrm{P}^1(\mathcal{T}_\ell)$, such that $u_{D,\ell} \to u_D$ strongly in $W^{1,1}(\Omega)^m$, and define

$$V_\ell = u_{D,\ell} + \mathrm{P}^1_0(\mathcal{T}_\ell)^m \text{ and } Y_\ell = \mathrm{P}^0(\Omega)^{m\times n}$$

to discretize, respectively, the variables u and η. We approximate Φ using the midpoint rule: For $v_\ell \in \mathrm{P}^1(\mathcal{T}_\ell)$, we set $\bar{v}_\ell(x) = (v_\ell)_T$ for $x \in T \in \mathcal{T}_\ell$, and for $v_\ell \in V_\ell$ and $\eta_\ell \in Y_\ell$, we define

$$\Phi_\ell(v_\ell,\eta_\ell) = \int_\Omega \phi(\bar{x}_\ell,\bar{v}_\ell,\eta_\ell)dx = \sum_{T\in\mathcal{T}_\ell} |T|\phi\big((x)_T,(v_\ell)_T,\eta_\ell|_T\big).$$

The functional $\Phi_\ell$is extended in an obvious way to $V_\ell \times L^1(\Omega)^{m\times n}$.

Remark 5.1. We could have included a quadrature approximation in our analysis in Section 4 as well. For the sake of simplicity, we decided not to do so. In the present case, we are in fact unable to prove convergence of the penalty method *without* the quadrature approximation.

Our first aim is an approximation result akin to Lemma 4.2. In Lemma 5.4 below we reduce this task to the following general condition which can be quite easily checked for different problems,

$$\begin{aligned} &\forall u \in V \ \exists\, u_\ell \in V_\ell, \zeta \in L^1(\Omega)^{m\times n} \text{ such that} \\ &\qquad \text{(i)} \quad \phi(x,u,\zeta) \in L^1(\Omega), \\ &\qquad \text{(ii)} \quad u_\ell \to u \text{ strongly in } W^{1,1}(\Omega)^m, \text{ and} \\ &\qquad \text{(iii)} \quad \limsup_{\ell\to\infty} \Phi_\ell(u_\ell,\zeta) \le \Phi(u,\zeta). \end{aligned} \tag{5.3}$$

Example 5.2 (1D Examples). Suppose that $n = 1$, that $\phi : \Omega \times \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ is globally continuous, and assume that $u_{D,\ell} = u_D$ for all ℓ; then (5.3) holds. This class includes in particular problems of Maniá type [5, 21].

Let $u \in V$, and let u_ℓ be its piecewise affine nodal interpolant. Then $u_\ell \to u$ strongly in $W^{1,1}(\Omega)^m$, $(\bar{x}_\ell, \bar{u}_\ell(x)) \to (x, u(x))$ uniformly in Ω, and in particular, since ϕ is globally continuous,

$$\Phi_\ell(u_\ell,\zeta) \to \Phi(u,\zeta),$$

for any fixed $\zeta \in \mathbb{R}^m$. □

Example 5.3 (Weak coupling of u and Du). Suppose that, in addition to (5.1),

$$\phi(x,v,\eta) \lesssim |v|^q + \Gamma(\eta), \tag{5.4}$$

where $\Gamma : \mathbb{R}^{m\times n} \to [0, +\infty]$ is proper, then (5.3) holds. We note that this class includes in particular the example of Foss, Hrusa, and Mizel [15] and Ball's example of cavitation [2].

Let $u \in V$ and take $u_\ell \in V_\ell$ converging strongly in $W^{1,1}(\Omega)^m$ to u. In particular, we have $\bar{u}_\ell \to u$ strongly in L^q. Further, let $\zeta \in \mathbb{R}^{m\times n}$ such that $\Gamma(\zeta) < +\infty$. In view of the growth condition imposed in (5.4), we obtain $\phi(x, u(x), \zeta) \in L^1(\Omega)$. Let $\ell_j \nearrow \infty$ be a subsequence such that

$$\limsup_{\ell\to\infty} \Phi_\ell(u_\ell, \zeta) = \lim_{j\to\infty} \Phi_{\ell_j}(u_{\ell_j}, \zeta).$$

Upon extracting a further subsequence, we may assume that $(\bar{x}_{\ell_j}, \bar{u}_{\ell_j}) \to (x, u)$ pointwise a.e. in Ω. Since $\phi(x, u(x), \zeta) \in L^1(\Omega)$, it is finite for a.a. $x \in \Omega$, and hence continuous at those points. We therefore obtain

$$\lim_{j\to\infty} \phi(\bar{x}_j, \bar{u}_j, \zeta) = \phi(x, u, \zeta) \quad \text{pointwise a.e. in } \Omega.$$

The majorant

$$\phi(\bar{x}_{\ell_j}, \bar{u}_{\ell_j}, \zeta) \le |\bar{u}_{\ell_j}|^q + \Gamma(\zeta)$$

is strongly convergent in $L^1(\Omega)$ and hence we can use Fatou's Lemma to obtain (5.3) (iii). □

Having shown that (5.3) indeed holds for several interesting problem classes, we establish the basic approximation result which it implies.

Lemma 5.4. *Fix $\varepsilon > 0$, and suppose that (5.2) and (5.3) hold, then, for every $u \in V$ there exists a sequence $(u_\ell, \xi_\ell) \in V_\ell \times Y_\ell$ such that*

$$\limsup_{\ell\to\infty} \big[\Phi(u_\ell, \xi_\ell) + \varepsilon^{-1}\Psi(Du_\ell, \xi_\ell)\big] \le E(u).$$

(The sequence u_ℓ can be chosen independent of the value of ε.)

Proof. We take the sequence u_ℓ specified in (5.3). For every $T \in \mathcal{T}_\ell$ and $x \in T$, we define

$$\bar{\phi}_\ell(x) = \inf_{\xi\in\mathbb{R}^{m\times n}} \big[\phi(\bar{x}_\ell(x), \bar{u}_\ell(x), \xi) + \varepsilon^{-1}\psi(\xi, Du_\ell(x))\big].$$

Since $\bar{\phi}_\ell$ can be constructed elementwise it is piecewise constant and, given the growth condition on ϕ from below, finite. In particular, it is measurable and its integral is well-defined with a value in $(-\infty, +\infty]$.

There exists a subsequence $\ell_j \nearrow \infty$ such that

$$\limsup_{\ell\to\infty} \int_\Omega \bar{\phi}_\ell \, dx = \lim_{j\to\infty} \int_\Omega \bar{\phi}_{\ell_j}, \text{ and}$$
$$(\bar{u}_{\ell_j}, Du_{\ell_j}) \to (u, Du) \text{ pointwise a.e. in } \Omega.$$

From the definition of $\bar{\phi}_\ell$, we have

$$\bar{\phi}_\ell \le \phi(\bar{x}_\ell, \bar{u}_\ell, Du) + \varepsilon^{-1}\psi(Du, Du_\ell) \text{ for a.a. } x \in \Omega,$$

and since we assumed that ϕ is continuous at every point where it is finite, and ψ is globally continuous, we obtain

$$\limsup_{j\to\infty} \bar{\phi}_{\ell_j}(x) \le \phi(x, u(x), Du(x)) \quad \text{for a.a. } x \in \Omega. \tag{5.5}$$

Again using the definition of $\bar{\phi}_\ell$ we obtain the majorant

$$\bar{\phi}_\ell \le \phi(\bar{x}_\ell, \bar{u}_\ell, \zeta) + \varepsilon^{-1}\psi(\zeta, Du_\ell) =: m_\ell,$$

where $\zeta \in L^1(\Omega)^{m\times n}$ is taken from (5.3). Since, by assumption (5.3) (i), $\phi(x, u, \zeta) \in L^1(\Omega)$, it follows that ϕ is continuous at $(x, u(x), \zeta(x))$ for a.a. $x \in \Omega$, and hence

$$m_{\ell_j}(x) \to m(x) := \phi(x, u(x), \zeta(x)) \quad \text{for a.a. } x \in \Omega.$$

Condition (5.3) (iii) translates as

$$\liminf_{j\to\infty} \int_\Omega m_{\ell_j}\, dx \le \int_\Omega m\, dx.$$

Applying Fatou's lemma to the sequence $m_\ell - \bar{\phi}_\ell$ gives

$$\int_\Omega \liminf_{j\to\infty} (m_{\ell_j} - \bar{\phi}_{\ell_j}) dx \le \liminf_{j\to\infty} \int_\Omega (m_{\ell_j} - \bar{\phi}_{\ell_j}) dx,$$

which, under the above conditions, is easily rearranged to yield

$$\limsup_{\ell\to\infty} \int_\Omega \bar{\phi}_\ell\, dx = \limsup_{j\to\infty} \int_\Omega \bar{\phi}_{\ell_j}\, dx \le \int_\Omega \limsup_{j\to\infty} \bar{\phi}_{\ell_j}\, dx \le E(u).$$

It remains to show that there exists a sequence $\xi_\ell \in Y_\ell$ such that

$$\limsup_{\ell\to\infty} \int_\Omega \bar{\phi}_\ell\, dx = \limsup_{\ell\to\infty} \Phi_\ell(u_\ell, \xi_\ell).$$

To this end we choose $\xi_\ell(x)$, for $x \in T$ and for $T \in \mathcal{T}_\ell$, piecewise constant, such that

$$\phi(\bar{x}_\ell, \bar{u}_\ell(x), \xi_\ell(x)) \le \bar{\phi}_\ell(x) + 1/\ell.$$

The existence of such functions follows from the definition of $\bar{\phi}_\ell$. □

Next, we will deduce from Lemma 5.4 the existence of a sequence $\varepsilon_\ell \searrow 0$ for which the same upper bound still holds.

Lemma 5.5. *Suppose that* (5.2) *and* (5.3) *hold, then there exists a sequence* $\varepsilon_\ell \searrow 0$ *such that*

$$\limsup_{\ell\to\infty} \inf E_\ell(V_\ell, Y_\ell) \le \inf E(V).$$

Proof. Let $v_k \in V$ such that $E(v_k) \le \inf E(V) + 1/k$. According to Lemma 5.4, for every $k \in \mathbb{N}$, there exists $\ell_k \in \mathbb{N}$ such that, for all $\ell \ge \ell_k$,

$$\inf_{(u_\ell,\xi_\ell)\in V_\ell\times Y_\ell} \Phi_\ell(u_\ell, \xi_\ell) + k\Psi(\xi_\ell, Du_\ell) \le E(v_k) + 1/k \le \inf E(V) + 2/k.$$

We may assume that $\ell_k \le \ell_{k+1}$ for all k. If we define

$$\varepsilon_\ell = 1/k \quad \text{for } \ell_k \le \ell < \ell_{k+1}, \quad k = 1, 2, \dots,$$

and $\varepsilon_\ell = 1$ for $1 \le \ell < \ell_1$, then $\varepsilon_\ell \searrow 0$ and

$$\inf E_\ell(V_\ell, Y_\ell) \le \inf E(V) + 2\varepsilon_\ell \qquad \text{for all } \ell \ge \ell_1. \qquad \square$$

We only need to prove a lower bound now. Here, we distinguish two cases, whether ϕ is convex in the third component or only quasiconvex.

Theorem 5.6 (Convex energies)**.** *Suppose that* (5.2) *and* (5.3) *hold, and assume in addition that ϕ is* convex *in its third component. Let $\varepsilon_\ell \searrow 0$ be the sequence established in Lemma 5.5, and let $(u_\ell, \xi_\ell) \in V_\ell \times Y_\ell$ be a sequence satisfying the following conditions:*

(i) (u_ℓ, ξ_ℓ) *are* approximate minimizers, *i.e.*,

$$\limsup_{\ell\to\infty} \big(E_\ell(u_\ell, \xi_\ell) - \inf E_\ell(V_\ell, Y_\ell)\big) = 0. \tag{5.6}$$

(ii) *There exists $u \in V$ such that*

$$u_\ell \rightharpoonup u \qquad \text{weakly in } W^{1,1}(\Omega)^m. \tag{5.7}$$

Then $u \in \operatorname{argmin} E(V)$,

$$\begin{aligned} &\lim_{\ell\to\infty} \Phi_\ell(u_\ell, \xi_\ell) = E(u), \\ &\lim_{\ell\to\infty} \varepsilon_\ell^{-1} \Psi(Du_\ell, \xi_\ell) = 0, \ \text{and} \\ &\xi_\ell \rightharpoonup Du \qquad \text{weakly in } L^1(\Omega)^m. \end{aligned} \tag{5.8}$$

Proof. By the construction of ε_ℓ and assumption (5.6) we have

$$\limsup_{\ell\to\infty} E_\ell(u_\ell, \xi_\ell) \le \inf E(V).$$

In particular, $\Psi(\xi_\ell, Du_\ell) \to 0$ which implies $\xi_\ell \rightharpoonup Du$ weakly in $W^{1,1}(\Omega)^{m\times n}$. We can therefore deduce that

$$E(u) \le \liminf_{\ell\to\infty} \Phi_\ell(u_\ell, \xi_\ell).$$

Using the same arguments as in the proof of Lemma 4.4, we can conclude the proof of the theorem. $\square$

Theorem 5.7 (Quasiconvex energies)**.** *Suppose that* (5.2) *and* (5.3) *hold, and assume in addition that ϕ is* quasiconvex *in its third component. Let $\varepsilon_\ell \searrow 0$ be the sequence established in Lemma 5.5, and let $(u_\ell, \xi_\ell) \in V_\ell \times Y_\ell$ be a sequence satisfying (i) and (ii) in Theorem 5.6, as well as:*

(iii) *There exists a monotone family of subsets $\Omega_k \nearrow \Omega$ such that*

$$\lim_{\ell\to\infty} \big\|\phi(\bar x_\ell, \bar u_\ell, Du_\ell) - \phi(\bar x_\ell, \bar u_\ell, \xi_\ell)\big\|_{L^1(\Omega_k)} = 0 \ \text{and} \tag{5.9}$$

$$\forall k \in \mathbb{N} \quad \sup_{\ell\ge k} \|u_\ell\|_{W^{1,\infty}(\Omega_k)} < \infty. \tag{5.10}$$

Then $u \in \operatorname{argmin} E(V)$, *and the conclusion* (5.8) *remains true as well.*

Proof. Set $\Omega_k' = \Omega \setminus \Omega_k$. In view of the bound (5.10), for fixed $k \in \mathbb{N}$, we have

$$u_\ell \overset{*}{\rightharpoonup} u \qquad \text{weakly-}* \text{ in } W^{1,\infty}(\Omega)^m.$$

Since ϕ is quasiconvex in its third component it follows from (5.9) that

$$\begin{aligned}\int_{\Omega_k} \phi(x,u,Du)dx \leq& \liminf_{\ell\to\infty} \int_{\Omega_k} \phi(\bar{x}_\ell,\bar{u}_\ell,Du_\ell)dx\\ =& \liminf_{\ell\to\infty} \int_{\Omega_k} \phi(\bar{x}_\ell,\bar{u}_\ell,\xi_\ell)dx,\end{aligned}$$

Using the lower bound (5.1) and the compactness of the embedding $W^{1,1}(\Omega)^m \subset L^q(\Omega)^m$, we obtain

$$\begin{aligned}\int_{\Omega_k} \phi(x,u,Du)dx \leq& \liminf_{\ell\to\infty} \Big(\Phi_\ell(u_\ell,\xi_\ell) - \int_{\Omega_k'} \phi(\bar{x}_\ell,\bar{u}_\ell,\xi_\ell)dx \Big)\\ \leq& \liminf_{\ell\to\infty} \Phi_\ell(u_\ell,\xi_\ell) + \limsup_{\ell\to\infty} C(|\Omega_k'| + \|u_\ell\|^q_{L^q(\Omega_k')})\\ =& \liminf_{\ell\to\infty} \Phi_\ell(u_\ell,\xi_\ell) + C(|\Omega_k'| + \|u\|^q_{L^q(\Omega_k')}).\end{aligned}$$

Setting $\delta_k = C(|\Omega_k'| + \|u\|^q_{L^q(\Omega_k')})$, we can further estimate

$$\begin{aligned}\int_{\Omega_k} \phi(x,u,Du)dx \leq& \liminf_{\ell\to\infty} \Phi_\ell(u_\ell,\xi_\ell) + \delta_k\\ \leq& \liminf_{\ell\to\infty} \Phi_\ell(u_\ell,\xi_\ell) + \limsup_{\ell\to\infty} \Psi(Du_\ell,\xi_\ell) + \delta_k\\ \leq& \limsup_{\ell\to\infty} E_\ell(u_\ell,\xi_\ell) + \delta_k\\ \leq& \inf E(V) + \delta_k \qquad \text{for all } k \in \mathbb{N}.\end{aligned} \tag{5.11}$$

Adding the term $C(1+|u|^q)$ to the integral on the left-hand side, the integrand becomes non-negative, and the bound becomes

$$\int_{\Omega_k} \big[\phi(x,u,Du) + C(1+|u|^q)\big]dx \leq \inf E(V) + \int_\Omega C(1+|u|^q)dx.$$

Taking the supremum over k on the left-hand side (employing, for example, the Beppo-Levi theorem), it follows that $\phi(x,u,Du)$ is integrable and that $u \in \operatorname{argmin} E(V)$. Furthermore, we can let $k \to \infty$ and thus $\delta_k \to \infty$ in (5.11) from which we can deduce the separate convergence of the energy contributions (compare also with the proof of Lemma 4.4). ☐

6. Connection with Γ-convergence

Our main results, Theorems 4.1, 5.6, and 5.7 can be understood as Γ-convergence (also known as epi-convergence) results. We refer to the monographs of Braides [8] and Dal Maso [12] for an introduction.

We will demonstrate this point of view at the example of the polyconvex case. To this end, suppose that (4.1), (4.2), and (4.4) hold, and define, for $v \in W^{1,n}(\Omega)^n$, $\eta \in L^1(\Omega)$ and $\varepsilon \in [0, \infty)$,

$$F(v, \eta, \varepsilon) = \begin{cases} E(v) & \text{if } v \in V, \eta = \det Dv, \varepsilon = 0, \\ \Phi(v, \eta) + \varepsilon^{-1}\Psi(\det Dv, \eta) & \text{if } v \in V, \varepsilon \in (0, \infty), \\ +\infty & \text{otherwise;} \end{cases}$$

$$F_\ell(v, \eta, \varepsilon) = \begin{cases} \Phi(v, \eta) + \varepsilon^{-1}\Psi(\det Dv, \eta) & \text{if } v \in V_\ell, \eta \in Y_\ell, \varepsilon \in (0, \infty), \\ +\infty & \text{otherwise.} \end{cases}$$

A minor modification of Lemma 4.3 shows that, for each $u \in V$, $\xi = \det Dv$, there exists a sequence $v_\ell \to v$ strongly in $W^{1,n}(\Omega)^n$, $\xi_\ell \to \xi$ strongly in $L^1(\Omega)$, and $\varepsilon_\ell \to 0$ such that

$$\limsup_{\ell\to\infty} F_\ell(u_\ell, \xi_\ell, \varepsilon_\ell) \le F(u, \xi, 0). \tag{6.1}$$

If $\xi \ne \det Du$ then $F(u, \xi, 0) = +\infty$ and hence (6.1) is trivially satisfied.

On the other hand, in Lemma 4.4, we have proven that, whenever $u_\ell \rightharpoonup u$ weakly in $W^{1,n}(\Omega)^n$, $\xi_\ell \rightharpoonup \xi$ weakly in $L^1(\Omega)$, and $\varepsilon_\ell \to 0$, then

$$F(u, \xi, 0) \le \liminf_{\ell\to\infty} F_\ell(u_\ell, \xi_\ell, \varepsilon_\ell). \tag{6.2}$$

Strictly speaking we have shown this for the case $\xi = \det Du$, but we have also shown that all accumulation points of families with bounded energy satisfy this. Hence, (6.2) is indeed correct.

In the language of Γ-convergence (6.1) and (6.2) are, respectively, called the *limsup* and *liminf conditions* (here only for $\varepsilon = 0$), and together they can be written as

$$\Gamma\text{–}\lim_{\ell\to\infty} F_\ell(v, \eta, 0) = F(v, \eta, 0) \qquad \text{for all } v \in V, \eta \in L^1(\Omega), \tag{6.3}$$

where Γ-convergence is understood with respect to the weak $W^{1,n}(\Omega)^n \times L^1(\Omega) \times [0, \infty)$-topology. In fact, it is straightforward to verify that

$$\Gamma\text{–}\lim_{\ell\to\infty} F_\ell = F$$

holds in the entire space $W^{1,n}(\Omega)^n \times L^1(\Omega) \times [0, \infty)$, however, this is less relevant for our purposes.

Thus, Theorem 4.1 can be interpreted as a Γ-convergence result in the sense of (6.3). In an obvious way, Theorems 5.6 and 5.7 can also be written in this way. We note however, that our original statements are slightly stronger in that we obtain separate convergence of the different contributions to the energy.

To conclude, we note that the statement

$$\Gamma\text{–}\lim_{\ell\to\infty} F_\ell(\cdot, \cdot, \varepsilon_\ell) = F(\cdot, \cdot, 0)$$

for a fixed sequence $\varepsilon_\ell \to 0$, is in general *false*. To see this, observe that to obtain (6.1), the choice of the sequence (ε_ℓ) may strongly depend on the limit point u which we are aiming to approximate.

7. Algorithms and Numerical Examples

In the preceding sections, we have formulated a general class of numerical methods for the solution of problems of the calculus of variations. The purpose of the present section is to demonstrate how they can be efficiently implemented and to demonstrate their practicality at several examples. We aim to give as much detail as possible so that our numerical results may be easily reproduced.

7.1. **Optimization of non-differentiable energies.** We begin by describing the implementation of the non-differentiable functionals which arise in our penalization procedure. Recall that we are aiming to minimize an energy which can be written in the form

$$\begin{aligned} E(v) &= \int_\Omega W(x, v, Dv)dx \\ &= \int_\Omega \phi\big(x, v, Dv, g(Dv)\big)dx \end{aligned}$$

over a convex and closed subset $\mathcal{A} \subset W^{1,1}(\Omega)^m$, where $\phi(x, v, F, \eta)$ and $g(F)$ are assumed to be smooth (at least twice differentiable) in v, F, and η. For the sake of simplicity we do not consider $g = g(x, v, Dv)$, but this is not a true restriction.

We shall consider general penalty functionals of the type

$$E_\varepsilon(v, \eta) = \int_\Omega \phi(x, v, Dv, \eta)dx + \varepsilon^{-1} \int_\Omega \big|g(x, v, Dv) - \eta\big|dx, \tag{7.1}$$

defined for $v \in V_\ell = u_{D,\ell} + \mathrm{P}^1(\mathcal{T}_\ell)^m$, $\eta \in Y_\ell = \mathrm{P}^0(\mathcal{T}_\ell)^k$.

By a simple variable transformation, we can replace η by $\eta + g(F)$ to obtain a new functional

$$\int_\Omega \phi\big(x, v, Dv, g(Dv) + \eta\big)dx + \int_\Omega |\eta|dx.$$

Next, we split the variable η into $\eta = \eta^+ - \eta^-$ where $\eta^+ = \max(\eta, 0)$ and $\eta^- = -\min(\eta, 0)$, and hide $g(F)$ within a newly defined energy density

$$\widetilde{\phi}(x, u, F, \eta) = \phi\big(x, u, F, g(F) + \eta\big),$$

to rewrite the functional as

$$\widetilde{E}_\varepsilon(v, \eta^+, \eta^-) = \int_\Omega \widetilde{\phi}(x, v, Dv, \eta^+ - \eta^-)dx + \varepsilon^{-1} \int_\Omega |\eta^+|_1 + |\eta^-|_1 dx, \tag{7.2}$$

where $|\cdot|_1$ denotes the ℓ^1-norm. Upon making η^+ and η^- independent variables but imposing the bound constraints $\eta^+ \geq 0$ and $\eta^- \geq 0$ we have thus turned the original non-differentiable problem to minimize (7.1) into a smooth but constrained optimization problem. In particular, we define (7.2) for all $v \in V_\ell$ and for all $\eta^+, \eta^- \in Y_\ell^+$, where

$$Y_\ell^+ = \{\eta \in Y_\ell : w \geq 0 \text{ in } \Omega\}.$$

Functionals of type (7.2) can be easily implemented with its gradient and hessian provided exactly. Our own implementation uses the trust region software TRON [20] to solve the *local* minimization problem

$$\min_{\substack{u\in V_\ell \\ \xi^\pm\in Y_\ell^+}} \widetilde{E}_\varepsilon(u,\xi^+,\xi^-). \tag{7.3}$$

7.2. **Adaptive mesh refinement for the penalty method.** At several points in the continuation algorithm for the penalty method described in the following section, we have to refine the mesh based on one of two principles: (i) either to reduce the overall energy or (ii) to reduce the contribution from the penalty term.

(i) To reduce the overall energy, we use a DWR-type idea [7]. Let (u_ℓ,ξ_ℓ) be a local minimum of $\widetilde{E}_\varepsilon$, computed using the method described above. We then define the error indicators

$$\eta_{\rm e} = \sum_{T\in\mathcal{T}_\ell} \eta_T, \quad \text{where}$$

$$\eta_T = \left| \int_T \partial_F \widetilde{\phi}(x,u_\ell,Du_\ell,\xi_\ell^+ - \xi_\ell^-) : (Du_\ell - G_\ell)dx \right|,$$

where $G_\ell \in \mathrm{P}^1(\mathcal{T}_\ell)^{m\times n}$ is a gradient recovery defined at each node z of the mesh $\mathcal{T}_\ell$ by

$$G_\ell(z) = \fint_{\cup\{T\in\mathcal{T}_\ell : z\in T\}} Du_\ell \, dx.$$

The value $\eta_{\rm e}$ gives an indication how much the "elastic" energy may be lowered by local mesh refinement. On the other hand, the value of the penalty integral

$$\eta_{\rm p} = \int_\Omega \psi_\varepsilon(v_\varepsilon^+ - v_\varepsilon^-)dx$$

indicates how much the "penalty" energy can be lowered. If $\eta_{\rm e} > C_{\rm e,p}\eta_{\rm p}$ then the mesh is refined by marking a fraction of all elements which have the largest indicators η_T for refinement. Otherwise all those elements are marked where $\xi^+ + \xi^-$ is non-zero (up to a threshhold which takes round-off errors and premature termination of the optimization into account).

(ii) To reduce the penalty energy we use the very same procedure. All those elements are marked for refinement where $\xi^+ + \xi^-$ are non-zero.

7.3. **Continuation algorithm.** A major difficulty which one encounters when solving problems involving the Lavrentiev phenomenon is the so-called repulsion property. For example, if $u_j \to x^{1/3}$ strongly in $L^1(0,1)$, but $u_j \in W^{1,\infty}(0,1)$ for all j, then

$$\int_0^1 |u_{j,x}|^6 (u_j^3 - x)^2 dx \to +\infty.$$

We can immagine this effect as a huge energy barrier that needs to be overcome (or a complicated path to be found) when moving from a Lipschitz function to the global minimum. In our computations, we see the effect in that even for sufficiently small meshes it is often difficult to find the correct minimizers and that the penalty method converges to the Galerkin solution instead. In particular, we observed that a local minimum when ε is chosen too small in relation to the current mesh since in that case the penalty method becomes in effect a Galerkin method again.

Thus the problem may be overcome by, either increasing ε, or decreasing the mesh size. The former is clearly not desirable while the latter may be prohibitively expensive. Our solution therefore was to consider a continuation with respect to the parameter ε. By initially choosing ε very large, the Galerkin solution is automatically discarded, even for coarse meshes. We then gradually decrease ε and adapt the mesh whenever there is a danger that we may "fall out" of the basin of attraction of the exact minimizer because ε has become too small for the current mesh. This may be controlled by requiring that at all times the total energy $\widetilde{E}_\varepsilon$ must be below a critical value which should be less than the energy of the Galerkin solution.

(1) Choose $\varepsilon_{\rm dec} \in (0,1)$, $E_{\rm goal} \in \mathbb{R}$, ε_0, an initial mesh $\mathcal{T}_0$, and two bounds $N^1_{\rm opt}, N^2_{\rm opt}$ (see remarks below how to coose them) for number of iterations of the optimization. Set $\ell = 0$ and $N_{\rm opt} = N^2_{\rm opt}$.
(2) Minimize $\widetilde{E}_{\varepsilon_\ell}$, allowing at most $N_{\rm opt}$ iterations.
(3) Determine next action:
   (3.1) If the optimization converged and $\widetilde{E}_{\varepsilon_\ell} \le E_{\rm goal}$ accept the step, set $\ell \leftarrow \ell + 1$, $\varepsilon_\ell = \varepsilon_{\ell-1} \cdot \varepsilon_{\rm dec}$, $\mathcal{T}_\ell = \mathcal{T}_{\ell-1}$, $N_{\rm opt} = N^1_{\rm opt}$ and continue at (2).
   (3.2) If the optimization converged but $\widetilde{E}_{\varepsilon_\ell} > E_{\rm goal}$, use refinement strategy (i) of the previous section to obtain a new mesh $\mathcal{T}_\ell$, set $N_{\rm opt} = N^2_{\rm opt}$, and redo step (2).
   (3.3) If the optimization did not converge, use refinement strategy (ii) of the previous section to obtain a new mesh $\mathcal{T}_\ell$, set $\varepsilon_\ell = \varepsilon_{\ell-1}$, $N_{\rm opt} = N^2_{\rm opt}$, and redo step (2).

Some further comments to refine the continuation algorithm are required.

- The initial parameters for step (1) have to be chosen in such a way that the first step is always succesful.
- The algorithm terminates unsuccesfully when a maximum number of elements is reached, and succesfully when a prescribed goal $\varepsilon_{\rm goal}$ for ε_ℓ is achieved.

- If the algorithm has terminated succesfully, we usually "post-process" the solution by performing a few additional mesh refinements (but fixing ε) using strategy (i) to confirm that the penalty energy and support of $\xi_\ell^+ + \xi_\ell^-$ tend to zero.
- After ε_ℓ is decreased in step (3.1) we only expect a small change in the solution. Therefore the optimization should essentially behave like Newton's method and therfore terminate in few steps. We therefore set the maximum number of iterations to a relatively small number (say $N_{\rm opt}^1 = 20$). This setting prevents us from spending many iterations on finding an entirely new equilibrium when ε_ℓ becomes too small for the current mesh and the penalty solution ceases to be a local minimizer.
- On the other hand, after the mesh is refined in either step (3.2) or (3.3) we expect a large change in the solution because the support of $\xi^+ + \xi^-$ may shrink and we therefore allow a larger number of iterations (say $N_{\rm opt}^2 = 10^6$, but we usually observe termination in far fewer iterations).

We have not addressed the question under which the algorithm is considered to have failed. When no Lavrentiev phenomenon occurs, we observe, in general, that for large ε a state satisfying the requirement $\widetilde{E}_{\varepsilon_\ell} \le E_{\rm goal}$ is found but that eventually, the algorithm will keep refining the mesh without being able to uphold this bound. We have therefore implemented a safety check which terminates the algorithm when a prescribed number of elements is reached.

As a warning, we also note that for sufficiently large ε it is sometimes possible to find *reasonably looking* solutions which indicate a Lavrentiev gap, but which may disappear as ε becomes small. It is therefore crucial to be able to drive ε as close to zero as possible.

7.4. **Maniá-type examples.** In this section we present numerical results for one-dimensional problems of the type

$$\text{(7.4)}\quad E(v) = \int_0^1 \Big(|v_x|^n (v^m - x^k)^2 + \nu |v_x|^2\Big) dx$$
$$\mathcal{A} = \big\{ v \in W^{1,1}(0,1) : v(0) = 0, v(1) = 1 \big\} = \mathrm{id} + W_0^{1,1}(0,1),$$

where $k, m, n \in \mathbb{N}$ and $\nu \ge 0$. This class includes in particular Maniá's original example [21] ($n = 6, m = 3, k = 1, \nu = 0$), and the *regular* example of Ball and Mizel [5, 4] ($n = 14, m = 3, k = 2, 0 < \nu < 2.4 \times 10^{-3}$). The idea in these examples is that, for $\nu = 0$ the infimum of the energy is always zero with exact solution $u^*(x) = x^{k/m}$, but that the power n can be chosen large to make approximation *difficult.* Moreover, if m and k are chosen such that $u^* \in H^1(0,1)$ then a perturbation of the functional with sufficiently small positive ν does not change whether E exhibits a Lavrentiev phenomenon or not.

The $x^{1/3}$ singularity for the original Maniá example is expensive (though not impossible) to resolve and so we have chosen to compute the solution for $n = 8, m = 2, k = 1, \nu = 0$ instead. We have plotted an accurate Galerkin solution, the solution of the penalty method for $\varepsilon = 10^{-1}$ and $\sharp\mathcal{T} = X$ in Figure 7.1, and the iterations of the contributions to the energy of the penalty method as well as the support of $\xi_\ell^+ + \xi_\ell^-$ in Figure 7.2.

In addition, we also computed the solution for the regular example of Ball and Mizel with $n = 14, m = 3, k = 2, \nu = 10^{-3}$, and we have plotted the solution in Figure 7.3. The evolution of the energy and of the support of the penalty variable is similar as in the previous example.

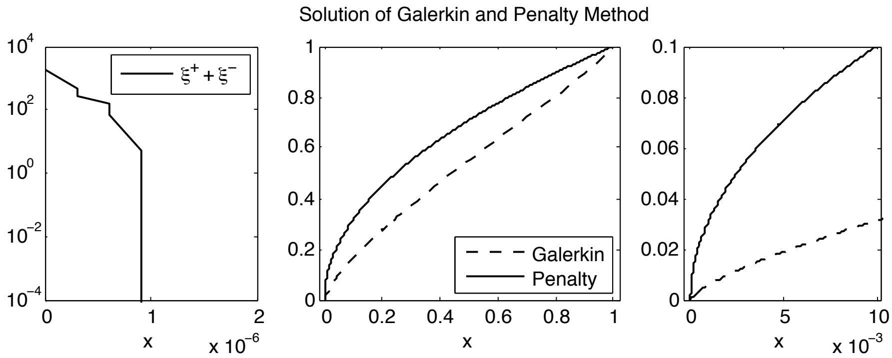

FIGURE 7.1. Final solutions of the Galerkin and the Penalty methods for the Maniá problem (7.4) with parameters $n = 8, m = 2, k = 1, \nu = 0$ *before* the reduction step. The error of the penalty solution in the L^∞-norm is $\|u_\ell - u^*\|_{L^\infty} \approx 7.83 \times 10^{-5}$.

7.5. **A convex example in 2D.** In this section, we present numerical results for a modification of the example provided by Foss, Hrusa and Mizel [15]. In their original example, a semi-circle Ω is transformed into a quarter-circle $y(\Omega)$, with stored energy

$$E(y) = \int_\Omega \Big[\big(|Dy|^2 - 2\det Dy\big)^4 + \nu\big(\frac{\kappa}{\det Dy} + 3^{2-\kappa 2}(1 + |Dy|^2)^{\kappa/2}\big)\Big]\, dx,$$

where κ and ν are parameters, creating a singularity at the original. The idea of the example is similar as in the regular examples of Ball and Mizel. For $\nu = 0$ the map $y^*(x) = r^{1/2}(\cos(\theta/2), \sin(\theta/2))$ gives zero energy but the large power makes approximation difficult and it can be shown that the problem exhibits the Lavrentiev phenomenon. Further, the deformation y^* has finite energy for $\nu > 0$ and hence, for ν sufficiently small the Lavrentiev effect remains.

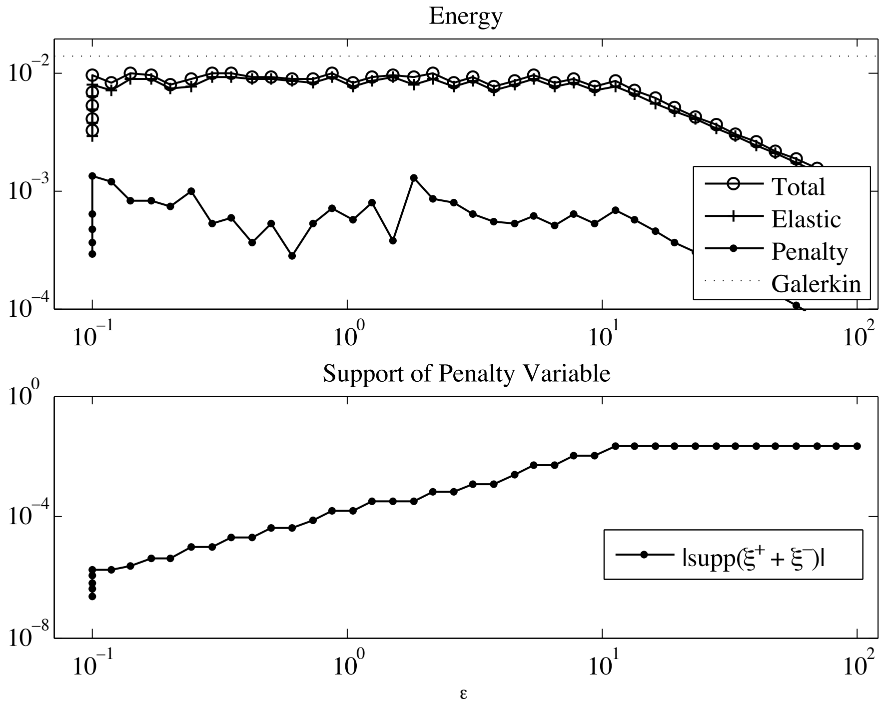

FIGURE 7.2. Evolution of the contributions to the penalty energy $\widetilde{E}_{\varepsilon_\ell}$ and of the support of the penalty variables at each step of the continuation algorithm outlined in Section 7.3. The clear convergence of $|\mathrm{supp}(\xi^+ + \xi^-)|$ to zero is a strong indicator for the convergence of the method.

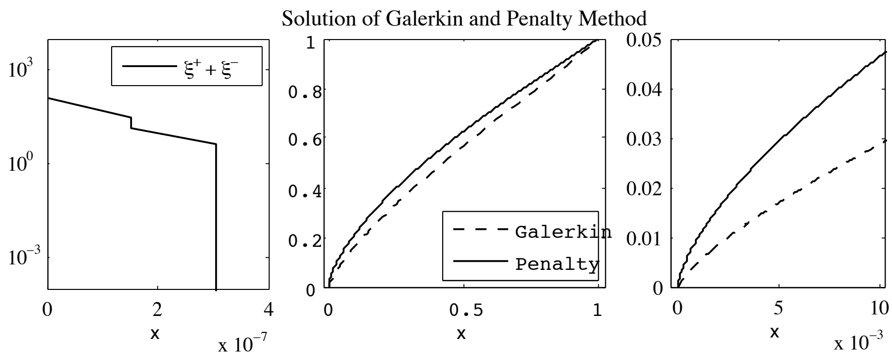

FIGURE 7.3. Final solution of the Galerkin and Penalty methods for Ball and Foss' [5] version of the Maniá problem (7.4) with parameters $n = 17, m = 3, k = 2, \nu = 10^{-3}$ *before* the reduction step. The different orders of the singularity at the origin are a clear indication for a Lavrentiev gap.

We note that the map term $F \mapsto (|F|^2 - 2\det F)$ is a non-negative quadratic form and hence the stored energy density

$$W_0(F) = \big(|F|^2 - 2\det F\big)^4$$

is convex. The polyconvex terms are fairly unimportant for the Lavrentiev effect and hence we decided to ignore them completely (though me should mention that we also performed succesful computations with the full Foss/Hrusa/Mizel example). Instead, upon noting that $y^* \in H^1(\Omega)$ we regularize W_0 by a quadratic and define

(7.5)

$$E(v) = \int_\Omega \big[W_0(Dy) + \nu|Dy|^2\big]\,dx$$

$$\mathcal{A} = \big\{v \in W^{1,1}(\Omega) : v(x) = r(\cos(\theta/2), \sin(\theta/2)) \text{ if } |x| = 1,$$
$$v_1(\{x_2 = 0, x_1 < 0\}) = \{0\} \text{ and } v_2(\{x_2 = 0, x_1 > 0\} = \{0\}\big\}$$

The solution and the evolution of the energy during optimization for the case $\nu = 0$ are plotted in Figures 7.4, 7.5 and 7.6. For the case $\nu = 10^{-3}$, we have only plotted the radial component of the solution in Figure 7.7. The evolutions of energy and support of the penalty variables during the optimization is similar as in Figure 7.6.

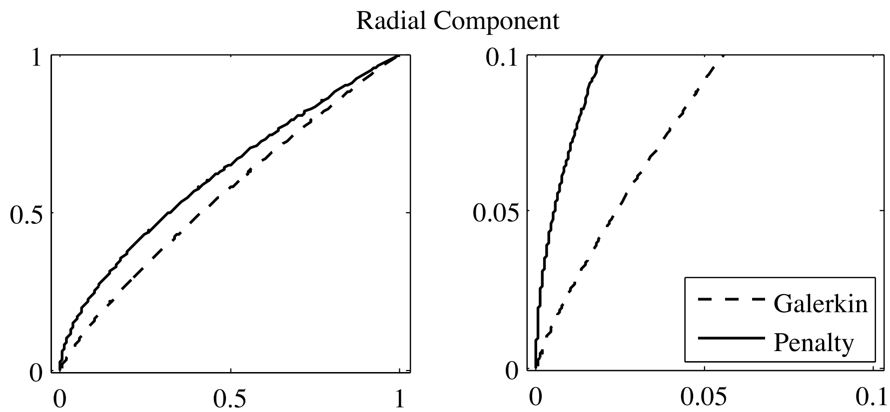

FIGURE 7.4. Radial components of the solution of the Galerkin and the Penalty methods for the modified Foss/Hrusa/Mizel problem (7.5) with $\nu = 0$ *before* the reduction step. The different orders in the singularities at the origin are a clear indivator for a Lavrentiev gap. The error of the penalty solution *before* reduction is $\|u_\ell - u^*\|_{L^\infty} \approx 1.07 \times 10^{-1}$.

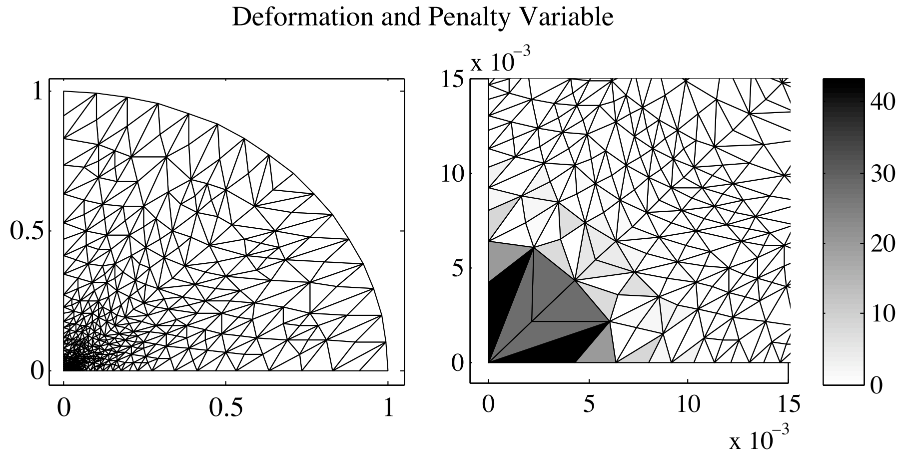

FIGURE 7.5. Plot of the deformation given by the solution of the Penalty method for the Foss/Hrusa/Mizel problem (7.5) with $\nu = 0$. The shade of the elements represents the size of the penalty variables $\xi^+ + \xi^-$.

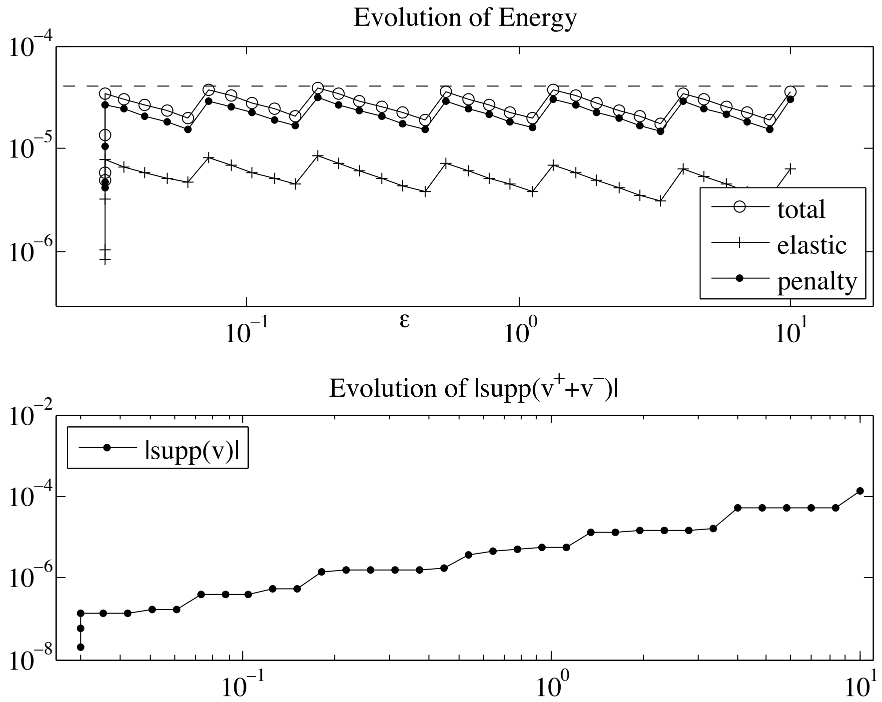

FIGURE 7.6. Evolution of the contributions to the penalty energy $\widetilde{E}_{\varepsilon_\ell}$ and of the support of the penalty variables at each step of the continuation algorithm outlined in Section 7.3. The apparent convergence of $|\mathrm{supp}(\xi^+ + \xi^-)|$ to zero is a strong indicator for the convergence of the method.

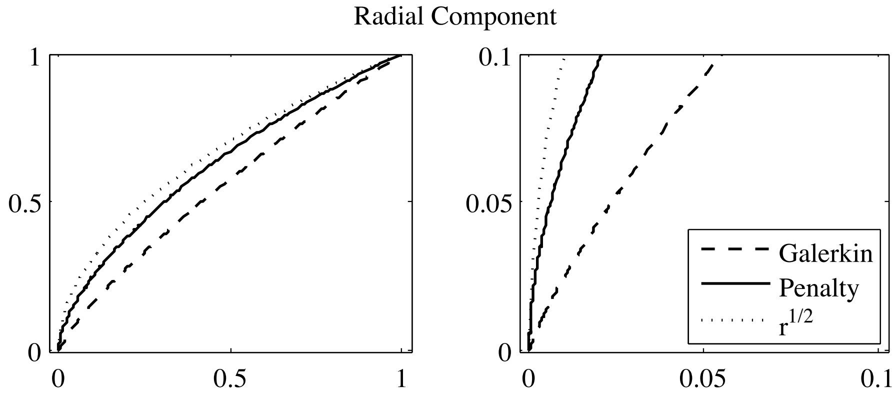

Figure 7.7. Radial components of the solution of the Galerkin and the Penalty methods for the modified Foss/Hrusa/Mizel problem (7.5) with $\nu = 0$ *before* the reduction step. The different orders in the singularities at the origin are a clear indivator for a Lavrentiev gap. For comparison, the exact solution for the case $\nu = 0$ is plotted as well.

References

[1] Y. Bai and Z.-P. Li. A truncation method for detecting singular minimizers involving the Lavrentiev phenomenon. *Math. Models Methods Appl. Sci.*, 16(6):847–867, 2006.

[2] J. M. Ball. Discontinuous equilibrium solutions and cavitation in nonlinear elasticity. *Philos. Trans. Roy. Soc. London Ser. A*, 306(1496):557–611, 1982.

[3] J. M. Ball and G. Knowles. A numerical method for detecting singular minimizers. *Numer. Math.*, 51(2):181–197, 1987.

[4] J. M. Ball and V. J. Mizel. Singular minimizers for regular one-dimensional problems in the calculus of variations. *Bull. Amer. Math. Soc. (N.S.)*, 11(1):143–146, 1984.

[5] J. M. Ball and V. J. Mizel. One-dimensional variational problems whose minimizers do not satisfy the Euler-Lagrange equation. *Arch. Rational Mech. Anal.*, 90(4):325–388, 1985.

[6] John M. Ball. Convexity conditions and existence theorems in nonlinear elasticity. *Arch. Rational Mech. Anal.*, 63(4):337–403, 1976/77.

[7] W. Bangerth and R. Rannacher. *Adaptive finite element methods for differential equations*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2003.

[8] Andrea Braides. Γ-*convergence for beginners*, volume 22 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, Oxford, 2002.

[9] Susanne C. Brenner and L. Ridgway Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer, New York, third edition, 2008.

[10] Philippe G. Ciarlet. *Mathematical elasticity. Vol. I*, volume 20 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, 1988. Three-dimensional elasticity.

[11] B. Dacorogna. *Direct methods in the calculus of variations*, volume 78 of *Applied Mathematical Sciences*. Springer-Verlag, Berlin, 1989.

[12] Gianni Dal Maso. *An introduction to* Γ*-convergence*. Progress in Nonlinear Differential Equations and their Applications, 8. Birkhäuser Boston Inc., Boston, MA, 1993.

[13] N. Dunford and J. T. Schwartz. *Linear operators. Part I*. Wiley Classics Library. John Wiley & Sons Inc., New York, 1988. General theory, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication.

[14] Lawrence C. Evans and Ronald F. Gariepy. *Measure theory and fine properties of functions*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[15] M. Foss, W. J. Hrusa, and V. J. Mizel. The Lavrentiev gap phenomenon in nonlinear elasticity. *Arch. Ration. Mech. Anal.*, 167(4):337–365, 2003.

[16] G. Knowles. Finite element approximation to singular minimizers, and applications to cavitation in nonlinear elasticity. In *Differential equations and mathematical physics (Birmingham, Ala., 1986)*, volume 1285 of *Lecture Notes in Math.*, pages 236–247. Springer, Berlin, 1987.

[17] A. Lavrentiev. Sur quelques problémes du calcul des variations. *Ann. Mat. Pura Appl.*, 41:107–124, 1926.

[18] Z.-P. Li. Element removal method for singular minimizers in variational problems involving Lavrentiev phenomenon. *Proc. Roy. Soc. London Ser. A*, 439(1905):131–137, 1992.

[19] Z.-P. Li. Element removal method for singular minimizers in problems of hyperelasticity. *Math. Models Methods Appl. Sci.*, 5(3):387–399, 1995.

[20] Chih-Jen Lin and Jorge J. Moré. Newton's method for large bound-constrained optimization problems. *SIAM J. Optim.*, 9(4):1100–1127 (electronic), 1999. Dedicated to John E. Dennis, Jr., on his 60th birthday.

[21] B. Maniá. Sopra un esempio di Lavrentieff. *Boll. Un. Mat. Ital.*, 13:147–153, 1934.

[22] P. V. Negrón-Marrero. A numerical method for detecting singular minimizers of multidimensional problems in nonlinear elasticity. *Numer. Math.*, 58(2):135–144, 1990.

Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany
E-mail address: cc@mathematik.hu-berlin.de

Mathematical Institute, 24-29 St Giles', Oxford OX1 3LB, UK
E-mail address: ortner@maths.ox.ac.uk