Analysis of the Quasicontinuum Method

Christoph Ortner

Oriel College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Michaelmas 2006

Abstract

Analysis of the Quasicontinuum Method

Christoph Ortner — Doctor of Philosophy
Oriel College — Michaelmas Term 2006

The aim of this work is to provide a mathematical and numerical analysis of the static quasicontinuum (QC) method. The QC method is, in essence, a finite element method for atomistic material models. By restricting the set of admissible deformations to linear splines with respect to a finite element mesh, the computational complexity of atomistic material models is reduced considerably.

We begin with a general review of atomistic material models and the QC method and, most importantly, a thorough discussion of the correct concept of static equilibrium. For example, it is shown that, in contrast to global energy minimization, a 'dynamic' selection procedure based on gradient flows models the physically correct behaviour.

Next, an atomistic model with long-range Lennard–Jones type interactions is analyzed in one dimension. A rigorous demonstration is given for the existence and stability of elastic as well as fractured steady states, and it is shown that they can be approximated by a QC method if the mesh is sufficiently well adapted to the exact solution; this can be measured by the interpolation error.

While the *a priori* error analysis is an important theoretical step for understanding the approximation properties of the QC method, it is in general unclear how to compute the QC deformation whose existence is guaranteed by the *a priori* analysis. An *a posteriori* analysis is therefore performed as well. It is shown that, if a computed QC deformation is stable and has a sufficiently small residual, then there exists a nearby exact solution and the error is estimated. This *a posteriori existence* idea is also analyzed in an abstract setting.

Finally, extensions of the ideas to higher dimensions are investigated in detail.

Acknowledgements

First and foremost, I thank my parents, Günther and Roswitha Ortner, who have always so generously supported me in every imaginable respect. I dedicate this work to them.

I deeply thank my partner Kimberley Brownlee for being a great role model and a wonderful companion.

It was a pleasure to work with my doctoral adviser Professor Endre Süli. He introduced me to many of the ideas I use in this thesis, including finite element methods, the quasicontinuum method, or how to use fixed point theory to considerably strengthen many results. In a similar spirit, and in no particular order, I thank Dr Bernd Kirchheim who introduced me to many advanced topics in the calculus of variations, Dr Johannes Zimmer with whom I shared interesting discussions on local minimization and who introduced me to the concept of curves of maximal slope, and Professor Nick Gould who was a fantastic source of information on numerical optimization methods. They all have been a great source of help and inspiration.

For three months, I visited the 'Istituto di Matematica Applicata e Tecnologie Informatiche' (IMATI-CNR) at Pavia, supported by the European research project *HPRN-CT-2002-00284: New Materials, Adaptive Systems and their Nonlinearities. Modelling, Control and Numerical Simulation.* A considerable part of this thesis, the most important ideas in Chapters 3–5, were developed during this visit. I thank Carlo Lovadina and Matteo Negri for their invitation.

In the first two years of my doctorate I received financial support from the University of Oxford through a *Scatcherd European Scholarship.* I am equally grateful to Oriel College and the Oxford University Computing Laboratory from which I received generous grants during the last year.

Contents

List of Figures

List of Tables

List of Symbols

General Notation:

d	space dimension
$\mathbb{R}$, $\mathbb{N}$, $\mathbb{Z}$	real numbers, non-negative integers, integers
$\mathcal{H}^k$	k-dimensional Hausdorff measure
$\lvert\cdot\rvert$	Lebesgue measure in $\mathbb{R}^d$; or Euclidean ℓ^2-norm of a vector or matrix
$\lvert\cdot\rvert_p$	ℓ^p-norm of a vector or matrix
$D(\phi)$	domain of definition of a functional ϕ; $D(\phi) = \{u : \phi(u) < \infty\}$
ϕ', ϕ''	first and second derivatives of the functional ϕ; cf. §2.2 and §4.1.1
$\lvert\partial\phi\rvert(u)$	local slope of ϕ at u; cf. §2.2.1.
$A : B$	for $A, B \in \mathbb{R}^{n\times m}$, $A : B$ denotes the double-contraction $A : B = \sum_{i=1}^{n}\sum_{j=1}^{m} A_{ij}B_{ij}$
$a \vee b$, $a \wedge b$	for real numbers a, b, $a \vee b = \max(a, b)$ and $a \wedge b = \min(a, b)$
$(u)_A$	average of the function u over the set A; $(u)_A = \lvert A\rvert^{-1}\int_A u(x)\,\mathrm{d}x$
$L(\mathscr{X}, \mathscr{Y})$	space of bounded linear functions from $\mathscr{X}$ to $\mathscr{Y}$, where $\mathscr{X}$ and $\mathscr{Y}$ are normed linear spaces

Continuum analysis:

Ω	an open connected subset of $\mathbb{R}^d$
$\lVert\cdot\rVert_{\mathrm{L}^p(\Omega)}$	Lebesgue-norms: $\lVert u\rVert_{\mathrm{L}^p(\Omega)} = \begin{cases} \left(\int_\Omega \lvert u\rvert_p^p\,\mathrm{d}x\right)^{1/p}, & \text{if } p \in [1,\infty) \\ \operatorname{ess.sup}_\Omega \lvert u\rvert_\infty, & \text{if } p = \infty. \end{cases}$
$\lvert\cdot\rvert_{\mathrm{W}^{1,p}(\Omega)}$	Sobolev-semi-norms: $\lvert u\rvert_{\mathrm{W}^{1,p}(\Omega)} = \lVert\nabla u\rVert_{\mathrm{L}^p(\Omega)}$; cf. §A.1.1
$\mathrm{BV}(\Omega)$	space of functions of bounded variation; cf. §A.1.3
$\lvert Du\rvert(\Omega)$, $\lVert\cdot\rVert_{\mathrm{BV}}$,	total variation of $u \in \mathrm{BV}(\Omega)$ and BV-norm; cf. §A.1.3
$\mathscr{T}$, $\mathscr{N}$, $\mathscr{E}$	a regular simplicial finite element mesh, its vertex set and its face set; cf. §3.3
$S^k(\mathscr{T})$, $S_0^k(\mathscr{T})$	continuous, piecewise polynomial splines of order k on the mesh $\mathscr{T}$; cf. §3.3 and §A.3
$X(a, b; Y)$	space of functions from the interval (a, b) to the metric space Y, e.g., $X = \mathrm{C}^1$ or $X = \mathrm{AC}$, the space of absolutely continuous functions from (a, b) to Y

Atomistic analysis:

J	atomistic interaction potential of Lennard–Jones type
ℓ^p_ε, $\mathrm{w}^{k,p}_\varepsilon$, $\mathrm{w}^{1,\infty}_{\varepsilon,f}$	symbols for Sobolev-type (semi-)norms in the one-dimensional discrete analysis; cf. §4.1; and §4.4 for $\mathrm{w}^{1,\infty}_{\varepsilon,f}$
$F'_n(y)$, $F''_{nm}(y)$	representation of the first and second derivatives of a one-dimensional atomistic energy; cf. §4.3.1
$\|\cdot\|_*$	dual norm; cf. §4.2.1 and §5.3.1
Ω	the atomistic domain: a finite subset of $\mathbb{R}^d$
$\bar{\Omega}$	a continuum representation of the atomistic domain, obtained by taking the union of all QC elements
E	atomistic energy functional
E_ξ	Contribution to the energy E from the atom at site ξ, cf. §1.4.4
E'_ξ	Pointwise residual; $E'_\xi(y) = \varepsilon^{d-1}(\partial E/\partial y_\xi)(y)$; cf. §4.4.2 and §6.2.
$\tilde{E}$	quasicontinuum approximation of an atomistic energy functional E; cf. §1.4 and §4.2.2
$\mathscr{T}$, $\mathscr{N}$, $\mathscr{E}$	a regular simplicial finite element mesh, its vertex set and its face set; cf. Chapter 6
$S^1(\mathscr{T})$	QC deformations obtained by continuous piecewise affine spline interpolation; cf. §1.4.2
$\tilde{\mathscr{A}}$	coarse-grained set of admissable deformations; cf. §1.4
$\bar{Y}$, $\nabla\bar{Y}_\kappa$	If Y is a QC deformation then $\bar{Y}$ is the 'classical' spline on $\mathscr{T}$ with the same nodal values and $\nabla\bar{Y}_\kappa$ its gradient in an element $\kappa \in \mathscr{T}$
ϕ_z	finite element shape function; cf. §6.1.3.

Chapter 1

Introduction

During the course of my doctoral research I have studied several mathematical models for material failure and techniques for their analysis. For example, a mathematically highly attractive model of fracture (cf. [48]) is based on the minimization of the free energy

$$E(u) = \int_\Omega \frac{1}{2}|\nabla u|^2 \,\mathrm{d}x + \kappa \mathcal{H}^{d-1}(S_u), \tag{1.1}$$

where the displacement u is a piecewise Sobolev function and $\mathcal{H}^{d-1}(S_u)$ is the surface measure of its discontinuity set, over a set of admissible displacements $\mathscr{A}$. The process of fracture is described by the balance between the bulk term $\int \frac{1}{2}|\nabla u|^2 \,\mathrm{d}x$ and the surface energy $\kappa \mathcal{H}^{d-1}(S_u)$. Using the direct method of the calculus of variations (cf. [4] for an analysis of (1.1)), it is possible to prove the existence of minimizers and the convergence of numerical discretizations. In short, if u_N is a minimizer of E (or an appropriate approximation thereof) over a suitable discrete space $\mathscr{A}_N$, $N \in \mathbb{N}$, then a subsequence of $(u_N)_{N\in\mathbb{N}}$ converges to a minimizer of E in $\mathscr{A}$ [17, 70]; cf. also §A.3.2. However, not only is (1.1) non-convex, but in every point of the configuration space it is discontinuous and has directions in which it is strictly convex and directions in which it is strictly concave. Any expert in numerical optimization would agree that it is virtually impossible to compute global minimizers of this model. In that case, we are fully justified in questioning the value of numerical approximation results based on this technique (unless of course it can be demonstrated that the discrete minimizers can be computed). As general and mathematically satisfying they might be, they bear little relationship with reality. The joint work with Negri [71] is an attempt to analyze fracture models based on different principles than that of global energy minimization; however, this was restricted to very simple model problems.

While initially still closely related to the direct method (cf. [75, 24]), as a consequence of experiences such as this, much of my research was centered around local

energy minimization. In §1.1, I try to explain the philosophy that I have developed over the past years and which I try to follow throughout this work.

Another area that I have studied with great interest were atomistic material models to which an introduction is given in §1.2. Here, the situation is a slightly different one. Although there are still some instances in the literature which study atomistic models via global energy minimization (cf. for example [15, 49]) it is usually acknowledged that global minimization leads to unphysical behaviour (cf. §1.1.2). Despite their complexities, atomistic material models have a much more accessible structure than free discontinuity problems such as (1.1). For example, gradient flows with respect to a sufficiently strong metric are well-defined and stable evolutions (in a uniform sense). This observation is used in Chapter 2 to analyze elastic equilibria of atomistic chains which are not global energy minima.

While this analysis was interesting and helped considerably to advance my analytical understanding of atomistic material models, it is not sufficient for numerical purposes. The gradient flow evolution has little or no connection to physical dynamics and it would be a waste of resources and computing power to compute static equilibria via a gradient flow evolution. It was therefore necessary to look for more *direct* methods to analyze equilibrium points.

Often, in the numerical analysis of nonlinear problems, it is assumed that an exact solution u satisfying a nonlinear equation $\mathscr{F}(u)$ exists, that $\mathscr{F}'$ is Lipschitz continuous at u and $\mathscr{F}'(u)$ is an isomorphism. If it is assumed that a numerical solution U is sufficiently close to such a u then it is easy to bound the error by its residual,

$$\|u - U\| \leq \|\mathscr{F}'(\theta)^{-1}(\mathscr{F}(u) - \mathscr{F}(U))\| \lesssim \|\mathscr{F}'(u)^{-1}\|\|\mathscr{F}(U)\|,$$

where $\theta \in \operatorname{conv}\{u, U\}$. In other words, *if a numerical solution is good then it is very good*. In my opinion, this assumption is not too different from the assumption that global minima of a discretized model can be computed. I therefore found my early results on the approximation properties of quasicontinuum methods (finite element-based coarse-graining methods for atomistic models; cf. §1.4), which were of this type, quite unsatisfactory. The crucial breakthrough was achieved after the joint work with Süli [78] in which we used a technique adapted from [63] to prove the existence of numerical DGFEM solutions near *stable* exact solutions of nonlinear elliptic or hyperbolic equations. Using a fixed point iteration, similar to Newton's method, if the same assumptions on an exact solution u are made as above, then the existence of a numerical solution which is *sufficiently close* can be shown rigorously. I have made use of this technique in Chapter 4 where I give a general *a priori* analysis of one-dimensional

atomistic models and their numerical approximation. Under natural conditions on the atomistic model, the following results are shown:

- the existence, local uniqueness and stability of elastic critical points;
- the existence, local uniqueness and stability of fractured atomistic solutions; and
- subject to natural conditions on the mesh quality, the existence and approximation properties of a quasicontinuum solution.

In particular the last point is based on ideas which have been used extensively to study the approximation properties of numerical discretizations to nonlinear operator equations [23, 40].

A reinterpretation of the fixed point argument mentioned above makes it possible to prove that

- if U is a numerical solution, $\mathscr{F}'(U)$ is an isomorphism, $\mathscr{F}'$ is locally Lipschitz continuous at U and $\mathscr{F}(U)$ is sufficiently small, then there exists a *nearby* exact solution satisfying $\|u - U\| \lesssim \|\mathscr{F}'(U)^{-1}\|\|\mathscr{F}(U)\|$.

With slightly different aims, this idea has been developed in some depth for weak as well as strong solutions to the nonlinear Laplace equation (see [81] for a fairly recent overview article). Also for dynamical systems, similar ideas can be employed (see for example [33]) [1]. Surprisingly though, this principle which I label *a posteriori existence*, seem to be unknown in the adaptive finite element community. Chapter 3 is therefore devoted entirely to its investigation in a simple setting. In addition to the abstract result outlined above, two applications, a numerical investigation of local minima of a double-well energy and the semilinear Laplace equation, are given.

In Chapter 5 this idea is applied to the quasicontinuum method. In addition to proving the existence of an exact atomistic solution *a posteriori*, an adaptive optimization method is developed which includes the adaptive mesh-refinement procedure within the optimization algorithm rather than the reverse which is common practise. This has the advantage that during the formation of a defect, which is a computationally expensive process to capture, the mesh is automatically adapted during the optimization whereas otherwise, it would be adapted after its termination and the solution computed from the beginning (cf. Chapter 5 for a more detailed discussion). The optimization method used is a proximal point algorithm which is, in essence, an implicit Euler discretization of a gradient flow. The analytical techniques studied in

[1] I thank Willy Dörfler, Mats Larson and Stig Larssen for pointing out these references to me.

Chapter 2 were therefore used to great advantage. The following results are given in Chapter 5:

- Bounds on the residual and the inf-sup constant of quasicontinuum approximations of solutions;
- If Y is a stable quasicontinuum solution with sufficiently small residual, then it is shown that there exists a nearby exact solution and the error is bounded in terms of the residual and the inf-sup constant;
- Each iteration of the proximal point optimization method is compared to a related *exact* problem and the error is estimated;
- The convergence properties of the proximal point optimization method are analyzed rigorously in the non-adaptive case and assessed through numerical examples in the adaptive case.

In order to be able to provide an entirely rigorous theory without any — possibly unjustified — assumptions, it has become a disappointing feature of this thesis that almost all results are one-dimensional. The last chapter is therefore included to generalize some aspects of the *a posteriori* analysis, which I consider the most important aspect of this thesis, to higher dimensions and to discuss the difficulties in extending the full results. It also includes a section discussing further developments required to make the analysis applicable to engineering problems. Further discussions about extensions to higher dimensions can be found in the respective chapters of the thesis.

1.1 Energy Minimization: Global or Local?

For the following discussion it is useful to have a concrete example in mind. Let us therefore assume that an elastic body can be described by a state variable y from a space of admissible configurations $\mathscr{A}$. To each deformation $y \in \mathscr{A}$, we assign an energy $E(y) \in (-\infty, +\infty]$. For concrete examples the reader may refer to [8, 30, 79]. It has become customary in a large mathematical and engineering solid mechanics community to assume that the elastic body attains a global energy minimum, i.e., a configuration $y \in \mathscr{A}$ such that $E(y) \leq E(\tilde{y})$ for all $\tilde{y} \in \mathscr{A}$. Therefore, over the past decades, the direct method of the calculus of variations [35], together with its relatives such as the theory of Γ-convergence [39, 37], have become important mathematical tools for understanding mechanical systems.

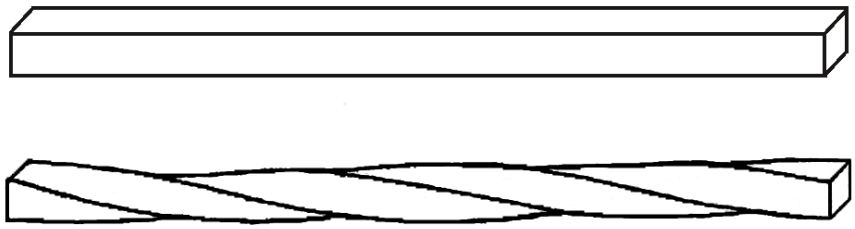

Figure 1.1: Two stable equilibria of an elastic body with the same boundary conditions. Modified from [30, p. 247].

The aim of the present section is to discuss this somewhat controversial postulate. For example, in [90] Tartar states: 'I like to call ultra-science-fiction elastic materials those... inexistent materials which instantaneously discover a minimum of their potential energy...'. Tartar's statement is intended to criticise, first, the fact that the evolution is entirely neglected (cf. [90, p.7], 'I do not think there is much Physics in a model where there is no time.'), and second, that the material is assumed to attain a *global* energy minimum.

In fact, for many mechanical systems simple counterexamples to the principle of global energy minimiation may be constructed. A well known counterexample from the theory of elasticity is the elastic rod shown in Figure 1.1. If the deformation is held fixed at both ends then the energy minimum is the undeformed state shown at the top of Figure 1.1. If the left-hand end is still held fixed, but the right-hand end is twisted by 360 degrees and then fixed in the same position as before, we obtain a new *stable equilibrium* under the same boundary conditions. Since, by allowing the right-hand face to deform freely, the elastic rod would return to its reference state, the twisted deformation must have a higher energy than the undeformed state and is therefore not *globally stable*. By iterating the process, it can be seen that the elastic energy has infinitely many *stable equilibria*.

While, based on these arguments, we may decide not to accept the principle of global energy minimization without a great degree of suspicion, we may still be interested in static equilibria only. It would seem a waste of resources to use dynamics to find a stable equilibrium of the system. Nevertheless, the stability of an equilibrium is inherently related to the evolution equation that is satisfied by the system as well as the perturbations from equilibrium which it is subjected to.

What is meant by a stable equilibrium (or meta-stable state) has to be decided in each particular case. Any general definition such as 'local energy minima' (with respect to a specific metric) is unlikely to be sufficiently strong for applications. Throughout

this thesis, only equilibria which lie in a uniformly convex basin of the energy functional are analyzed.

We proceed by considering two examples to further illustrate the discussion.

1.1.1 Microstructure

Since the work of Ball and James [10], the formation of microstructure in materials is often modelled as the (global) minimization of an energy functional of the type

$$E(y) = \int_\Omega W(x, y, \nabla y)\,\mathrm{d}x,$$

where Ω is a domain in $\mathbb{R}^d$ and W a stored energy function with two or more wells. (By contrast, an elastic stored energy function has a single well, namely $SO(d)$.) The gradient ∇y is not always a deformation gradient, but may sometimes be interpreted as a generalized measure of the deformation of atomistic lattices which describe a change of the lattice type.

For the purpose of this discussion it is sufficient though to consider the much simplified toy model

$$E(y) = \int_0^1 \left[\frac{1}{4}(y_x^2 - 1)^2 + y^2\right]\mathrm{d}x, \quad \mathscr{A} = \mathrm{W}_0^{1,4}(0,1). \tag{1.2}$$

It is easily seen that any sequence $y^{(j)} \in \mathscr{A}$ with gradients $y_x^{(j)} \in \{-1, 1\}$ and $y^{(j)} \to 0$ in L^2 satisfies $E(y^{(j)}) \to 0$ and hence, $\inf_{\mathscr{A}} E = 0$. However, if $E(y) = 0$ then $\|y\|_{\mathrm{L}^2}^2 = 0$ and hence $E(y) = E(0) = 1/4$ which implies that that E has no minimizer in the space $\mathscr{A}$.

Minimizing sequences for functionals such as (1.2) develop finer and finer oscillations. For this reason, the functional is often used as a cartoon for the formation of micro-structure in materials. An elaborate theory, based on a generalized notion of solution (Young measures), was developed to account for the non-existence of classical minimizers. See [67] for an introduction to variational models for microstructures and further references on the subject.

The theory of microstructures has greatly benefited from the study of the minimization of multi-well energies. It should therefore not be said that global minimization is wrong *per se*. One possible conclusion would be that the model (of minimizing the energy globally) has no direct physical interpretation but describes a mathematical process which bears similarities to the formation of microstructure in real materials. For example if, instead of minimizing a multi-well energy, we simply require that the

deformation gradients lie in the wells, then, depending on the wells' geometry, microstructure will also be predicted in many situations. While this problem is usually still studied in the context of energy minimization (cf. [79, Chapter 4]), it is in fact independent of the variational problem.

In §3.2, we look at (1.2) from the perspective of local energy minimization. It is demonstrated that typical local minimizers with respect to the $\mathrm{W}^{1,\infty}$-topology are stable and have a *finite microstructure*, i.e., a structure with a finite (as opposed to infinitesimal) length scale. It shoud be noted, however, that the equilibria computed in §3.2 are not local minimizers with respect to the $\mathrm{W}^{1,p}$-norm for any $p < \infty$. This demonstrates how crucial the notion of locality is which we choose to adopt.

1.1.2 Atomistic material models

As a second example, we consider a simplified model for an atomic chain. For $y = (y_j)_{j=0}^N \in \mathbb{R}^{N+1}$ define

$$E(y) = \sum_{j=1}^{N} J(y_j - y_{j-1}), \text{ where } J(z) = \begin{cases} z^2 - 2z, & \text{for } -\infty < z \leq 2 \\ 0, & \text{for } 2 \leq z < +\infty. \end{cases} \tag{1.3}$$

This energy represents a cartoon of the atomistic material models described in §1.2, taking into account nearest-neighbour interactions only and with a much simplified interaction potential.

One version of the Cauchy–Born hypothesis states that an atomistic body subjected to a small affine boundary displacement will follow this displacement in the bulk. A mathematical derivation of this important foundation of continuum mechanics is given in [32, 49] by considering global minima of a higher dimensional version of (1.3) but with a convex interaction potential. When the potential has sublinear growth, global minimization will typically not reproduce this behaviour.

To demonstrate this, let us consider the minimization problem

$$\min_{\substack{y_0=0,\\ y_N=N(1+\delta)}} E\big((y_j)_{j=0}^N\big). \tag{1.4}$$

Concerning the formulation of the boundary displacement, note that the minimum of $J(z)$ is attained at $z = 1$. The choice of boundary displacement we have made here scales linearly with the number of atoms, which is the macroscopically interesting relation. A different choice was made in [20] which we discuss briefly in Section 2.1.

Proposition 1.1 *If $\delta < N^{-1/2}$ then the affine state $y^a_j = (1+\delta)j$ is the unique solution of (1.4). If $\delta > N^{-1/2}$ then the set of solutions is given by*

$$\{y \in \mathbb{R}^{N+1} : y_0 = 0, y_N = N(1+\delta), y_i - y_{i-1} = 1 \text{ for all except one } i\}.$$

Proof. It is easy to see that, if a fractured state, i.e., a state $(y_i)_{i=0}^N$ where $J'(y_i - y_{i-1}) = 0$ for some i, is a minimizer then all but one interaction must satisfy $y_i - y_{i-1} = 1$. Furthermore, all such deformations have the same energy. On the other hand, if all interactions lie in the region over which J is convex then y^a is the only possible equilibrium. Hence, we only need to compare the energy of two atomic states.

Consider the 'fractured' deformation $y^f_j = j$ for $j = 0, 1, \ldots, N-1$ and $y^f_N = N(1+\delta)$. Then,

$$E_f(\delta) = E(y^f) = (N-1)J(1) + J(1+N\delta) = -(N-1).$$

The affine state $y^a_j = (1+\delta)j$, on the other hand, has the energy

$$E_a(\delta) = E(y^a) = N(\delta^2 - 1).$$

Thus, $E_f(\delta) < E_a(\delta)$ if, and only if $\delta > N^{-1/2}$. $\square$

Proposition 1.1, which is merely a review of well-known facts, shows that not only is the Cauchy–Born hypothesis violated, but in fact any material with a sufficient number of atoms breaks for arbitrarily small (in a macroscopic sense) boundary displacements or surface forces, if it were to attain its global energy minimum. This behaviour is in clear contradiction to observations and therefore, global minimization should be rejected for models of the type (1.3).

The result can easily be extended (though with a slightly weaker statement) to general atomistic interactions and to arbitrary dimensions. The only crucial condition is the sublinear growth of the potential.

Thus, atomistic energies are a class of models where the postulate of global energy minimization is clearly the wrong approach to finding equilibrium solutions. The analysis of Chapter 2 presents gradient flow dynamics, modelling local energy minimization, as an alternative. It should be emphasized from the outset that no claim is made regarding the physical relevance of the evolution. It is merely presented as a model for local energy minimization.

1.2 Introduction to Atomistic Material Models

The introduction to (semi-)empirical atomistic material models in this section follows to some extend the review papers by Liu *et al.* [54, 61] and Miller *et al.* [65], and the overview given in the thesis of Wilson [97].

Over the past three decades, engineering sciences have begun to analyze the properties of advanced materials directly from a study of their nanoscopic behaviour. It is now even becoming possible to synthesize tailor-design systems from the nanoscale upwards. With their exceptional mechanical and chemical properties such as low density, high stiffness, high strength, etc., such nanoscale materials find use in a wide range of applications, including material reinforcement, nanoelectronics, chemical sensing and many others. Since in many cases distinct non-continuum behaviour is observed, atomistic material models need to be employed. While physical and engineering research of nanoscale systems is advancing at an increasing rate, it is safe to say that the mathematical analysis of such systems is still in its infancy. It is the aim of this thesis to help fill this gap by opening up possible avenues for the mathematical study of atomistic models of materials and particularly their numerical treatment.

The state of an atomistic body is described by a state variable, which we shall always denote by y. We assume that to each atomistic state y, we can associate an energy $E(y) \in (-\infty, +\infty]$. The nature of the state variable and the energy functional depend on the particular atomistic model, the choice of which is typically governed by a balance of computational complexity and physical accuracy. One can perform calculations on systems with millions of atoms if a simple pair-potential energy is employed (cf. §1.2.1), but with only a few atoms if one uses a high level *ab initio* theory. In this section the most important atomistic models for engineering applications are described. It is not intended as a detailed review but only to provide an idea of the kind of atomistic models that are typically used for simulations in solid mechanics and materials science. No account of *ab initio* techniques which, due to their computational complexity, are not as widely used for applications in mechanics, is given. The reader may wish to refer to [54] for an introduction to the subject and to [47] for an attempt to apply the QC method to such a model.

We assume throughout that the atomistic body contains only one type of atom. Generalizations to complex lattices, at least at the theoretical level, are obvious.

1.2.1 Pair- and multi-body potentials

If we assume that the state of an atomistic system containing N atoms is described by the position of the atoms alone, we write the state variable as $y = (y_i)_{i=1}^N$ where $y_i \in \mathbb{R}^d$. The potential energy $E(y)$ can then be rewritten in terms of interactions involving up to N particles,

$$\begin{aligned} E(y) &= V^{(0)} + \sum_{i=1}^{N} V_i^{(1)} + \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} V_{i,j}^{(2)} \\ &\quad + \sum_{i=1}^{N-2} \sum_{j=i+1}^{N-1} \sum_{k=j+1}^{N} V_{i,j,k}^{(3)} + \cdots + V_{123\ldots N}^{(N)}. \end{aligned} \tag{1.5}$$

$V^{(0)}$ is mathematically irrelevant and can be taken to be zero. $V_i^{(1)}$ is the energy of the isolated atom i and is taken to be zero if it is in its ground state, which is always assumed in multi-body interaction systems. External forces could be accounted for in this term but we shall treat them separately. $V_{ij}^{(2)}$ is the two-body interaction between atoms i and j and is usually assumed to be of the form $V_{ij}^{(2)} = J(|y_i - y_j|)$. Truncating the series later increases the accuracy, but also the computational effort, of the calculation. A quick calculation shows that a summation which takes one second for 100 atoms at the two-body level would take approximately 33 seconds at the three-body level and nearly 14 minutes at the four-body level, assuming that the computational effort for each individual term is the same [97].

Pair potentials. Pair-potential energies are obtained by only calculating the interactions between pairs of atoms, i.e., by setting $V_{ijk}^{(3)} = \cdots = V_{12\ldots N}^{(N)} = 0$. Thus, E takes the form

$$E(y) = \sum_{i=1}^{N} \sum_{j=i+1}^{N} J(|y_i - y_j|), \tag{1.6}$$

where $y_i \in \mathbb{R}^d$ is the position of the ith atom and J denotes the potential of the force acting between the ith and the jth atom. Typical examples of atomistic interaction potentials are the Lennard–Jones 6-12 potential [57],

$$J(z) = Az^{-12} - Bz^{-6}, \tag{1.7}$$

or the Morse potential [66],

$$J(z) = \mathrm{e}^{-2\alpha(z-1)} - 2\mathrm{e}^{-\alpha(z-1)}. \tag{1.8}$$

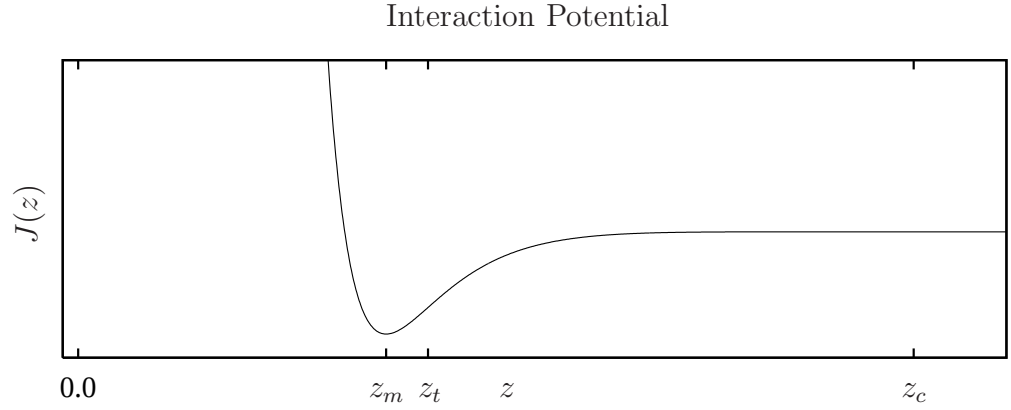

Figure 1.2: The shape of an atomistic interaction potential J with cut-off radius z_c. Here, z_m denotes the potential minimum and z_t the (unique) turning point.

The parameters A, B in the case of the Lennard–Jones potential and α in the case of the Morse potential depend on the type of interacting atoms. Since the interaction across a distance of more than a few atomic spacings is negligible, it is furthermore customary in practical computations to multiply the potential J with a cut-off function, for example,

$$\psi(z) = \begin{cases} z^4/(1+z^4), & \text{if } z \leq 0 \\ 0, & \text{if } z > 0. \end{cases} \tag{1.9}$$

The new potential $\tilde{J}(z) = J(z)\psi(z - z_c)$, where z_c is the cut-off radius, is then taken as the exact model and its parameters are fitted to the material under consideration. The shape of typical pair potentials such as (1.7) or (1.8) is shown in Figure 1.2.

A pair potential describes the interactions between particles with no relation to the other particles in the system. However, in metallic systems, the bonding electrons are delocalized over a large number of atoms and therefore the bonding between two atoms is highly dependent on the local environment. As a consequence, pair potentials fail to predict many important aspects of metallic systems. For example, the ratio of the vacancy energy (the energy required to remove an atom from the bulk and place it on the surface) to the cohesive energy (the energy of an atom in the bulk) is always overestimated as they do not account for the effect of the local environment at the vacancy on the bonding. Pair-potentials are therefore rarely used for the simulation of metallic systems, particularly, when quantitative results are required [97].

Many-body potential energy functions. The failings of pair potentials have led to the development of potentials for metals where the local environment of an atom is

incorporated into the potential through many-body effects. For example, the Axilrod–Teller potential energy [6] is given by

$$E(y) = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} J(|y_i - y_j|) + \sum_{i=1}^{N-2} \sum_{j=i+1}^{N-1} \sum_{k=j+1}^{N} V^{(3)}(y_i, y_j, y_k), \tag{1.10}$$

where J is typically a Lennard–Jones type potential and $V^{(3)}$ is given by

$$V^{(3)}(y_i, y_j, y_k) \sim \frac{1 + 3\cos(\gamma_i)\cos(\gamma_j)\cos(\gamma_k)}{(r_{ij} r_{jk} r_{ik})^3},$$

where r_{ij} is the distance between atoms i and j, and γ_i is the angle between the vectors $y_i - y_j$ and $y_i - y_k$.

More advanced examples include the Tersoff potential which is used to simulate Silicon and Germanium [91] or the class of Murrell–Mottram potentials [69, 68] which has been used successfully for the simulation of a number of metallic solids, for example noble metal clusters [97]. The angles γ_i play a crucial role in the Tersoff and Murrell–Mottram potentials as well.

1.2.2 The embedded atom model

The embedded atom method (EAM), developed by Daw and Baskes [38], is a popular atomistic model for solids which draws parallels with density functional theory (an *ab initio* method) and describes the bonding of atoms in terms of their local electronic densities. In its most basic form, the energy is given by

$$E(y) = \sum_{i=1}^{N} F(\overline{\rho}_i).$$

F is called the embedding energy and $\overline{\rho}_i$ the electron density at the site of the ith atom, given by

$$\overline{\rho}_i = \sum_{\substack{j=1 \\ j \neq i}}^{N} \rho(r_{ij}) \tag{1.11}$$

where $\rho(r_{ij})$ is the density of electrons of the jth atom at distance r_{ij}. Due to its fast decay, ρ is also multiplied by a cut-off potential such as (1.9). The EAM energy is usually complemented by a corrective pair potential energy and in some cases even three-body interaction terms. For our purposes it is sufficient to assume that an EAM energy is of the form

$$E(y) = \sum_{i=1}^{N} F(\overline{\rho}_i) + \sum_{i=1}^{N-1} \sum_{j=2}^{N} J(|y_i - y_j|). \tag{1.12}$$

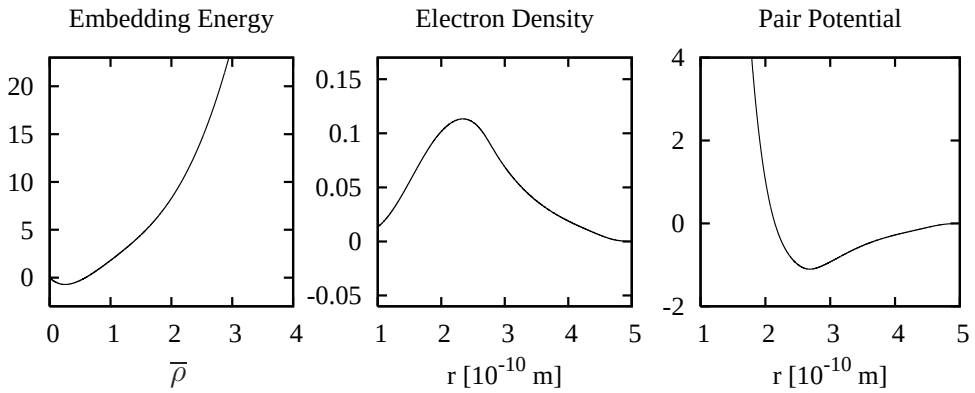

Figure 1.3: Embedding energy, electron density distribution and pair-interaction potential for an EAM model of α-titanium.

The functions F, ρ and J can in principle be derived from density functional theory [38] but are in practise obtained by fitting ansatz potentials to experimental data or *ab initio* simulations. For example, the EAM energy for a pure titanium system at zero-temperature [100] is given by

$$\begin{aligned}
\rho(r) &= \psi\left(\frac{r-r_c}{\varepsilon}\right)\left[A\mathrm{e}^{-\alpha_1(r-r_0)^2} + \mathrm{e}^{-\alpha_2(r-r_0')^2}\right], \qquad (1.13)\\
J(r) &= \psi\left(\frac{r-r_c}{\varepsilon}\right)\left[V_0\mathrm{e}^{-\beta_1(r-r_1)} + V_0'\left(\mathrm{e}^{-2\beta_2(r-r_1')} - 2\mathrm{e}^{-\beta_2(r-r_1')}\right) + \delta\right], \quad \text{and}\\
F(\overline{\rho}) &= F_0 + \frac{1}{2}F_2(\overline{\rho}-1)^2 + q_0(\overline{\rho}-1)^3 + \sum_{i=1}^{3} B_i(\overline{\rho}-1)^{3+i}.
\end{aligned}$$

The embedding energy F is in fact a truncated Taylor expansion near $\overline{\rho} = 1$. While some of the parameters in (1.13) were derived analytically from the properties of a titanium lattice, most were fitted from *ab initio* simulations. The optimized fitting parameters can be found in Table I in [100] and give rise to functions plotted in Figure 1.3.

For practical simulations, one would use splines of tabulated values which can be obtained, for a great variety of atomistic interactions, from online databases.

1.3 Overview of Multiscale Techniques

While microscale and nanoscale systems and processes are becoming more viable for engineering applications, our ability to model their performance numerically remains

limited. A modern desktop computer can simulate up to several millions of atoms while the simulation of realistic atomic systems requires at least tens of billions of atoms. Such systems cannot be modelled by continuum methods since they are too small but they cannot be modelled by atomistic methods either which are computationally too expensive. Therefore, *multiscale methods* are urgently needed for this class of problems.

In this section, an overview of several multiscale techniques which are commonly used in connection with material models is given. The purpose of this section is to put the work presented in this thesis into context. For a more thorough overview of multiscale methods for atomistic models see, for example, [34, 61] and references therein.

Top-down or bottom-up. It is customary to distinguish two fundamental classes of multiscale approaches: the top-down approach and the bottom-up approach. In the top-down approach a macroscopic model, for example continuum elasticity, is supplemented by information from the microscopic scale. For instance, a stored energy function which represents crystal symmetries could be regarded as a top-down multiscale model. Well-known examples of top-down models are the *heterogeneous multiscale method* [41, 44] and the *equation-free computing* by Kevrekidis *et al.* [50, 51]. In these methods, the *coarse* model is typically governed by a microscopic law which is simulated locally around an interpolation or quadrature point. Top-down models generally have the advantage that classical techniques can be used for their analysis. It is not always clear, however, what the relationship between different scales is and it is therefore not always straightforward to define a *modelling error*.

In contrast, bottom-up approaches assume a model on a fine scale as exact and *coarse-grain* it using appropriate techniques. From the point of view of modelling, they are clearly preferable. Particularly in atomistic simulations the bottom-up approach seems to be dominant. The distinction is not always clear-cut, however. The *local quasicontinuum method* (cf. §1.4) for example can be considered both a top-down or bottom-up method.

Concurrent methods. Another design decision for multiscale methods is the question whether computations on different scales should be performed independently or simultaneously. For example, it is possible to first compute a coarse solution and then correct it by localized computations on a finer scale wherever deemed necessary. An example of this principle is the *homogenized Dirichlet projection method* used to model heterogeneous structures [72, 99].

As experience has shown, the connection between microscale physics and macroscopic deformation cannot usually be neglected. For this reason, concurrent multiscale methods, which compute the coarse and microscopic variables simultaneously, are preferable and have become dominant. Two bottom-up, concurrent multiscale methods for coarse-graining atomistic material models, which have gained much attention in the literature, are the quasicontinuum (QC) method [65, 73] and the bridging-scales method [62, 89]. The QC method is one of the central topics of this dissertation. A thorough overview is given in §1.4 and a complete analysis is performed in one dimension in Chapters 4 and 5. The bridging- scales method will be briefly presented in the conclusion, as a possible avenue for further work.

1.4 The Quasicontinuum Method

Let us assume an atomistic model where the positions of the atoms are the only degrees of freedom. We use $\mathscr{V} = (\mathbb{R}^d)^N$ (N being the number of atoms) to denote the space of all possible atomistic deformations. The basic idea of the QC method is simple. Let $\mathscr{A}$ be a closed convex subset of $\mathscr{V}$ denoting the set of admissible deformations. All atomistic energies E introduced in §1.2 are continuously differentiable in their domains of definition (where they are finite). We are therefore looking for critical points of E in $\mathscr{A}$, i.e., deformations $y \in \mathscr{A}$ such that

$$E'(y; \tilde{y} - y) \geq 0 \qquad \forall \tilde{y} \in \mathscr{A}. \tag{1.14}$$

Since $\mathscr{A}$ is a subset of a finite-dimensional space it is *in principle* possible to compute solutions to (1.14) but due to the high number of degrees of freedom it cannot be realized in practise. However, by choosing an appropriate *coarse-grained* admissible set $\tilde{\mathscr{A}}$, the number of degrees of freedom can be reduced to a manageable amount. The Galerkin approximation of (1.14) in $\tilde{\mathscr{A}}$ is to find $Y \in \tilde{\mathscr{A}}$ such that

$$E'(Y; \tilde{Y} - Y) \geq 0 \qquad \forall \tilde{Y} \in \tilde{\mathscr{A}}. \tag{1.15}$$

Even for the simplest atomistic model, the evaluation of the variational inequality (1.15) has an equally high computational complexity. The energy E is therefore approximated by a simpler functional $\tilde{E}$, which approximates E on the space $\tilde{\mathscr{A}}$, and one tries to find critical points in $\tilde{\mathscr{A}}$, i.e., $Y \in \tilde{\mathscr{A}}$ such that

$$\tilde{E}'(Y; \tilde{Y} - Y) \geq 0 \qquad \forall \tilde{Y} \in \tilde{\mathscr{A}}. \tag{1.16}$$

It is important to note that $\tilde{E}$ is defined only on $\tilde{\mathscr{A}}$. In fact, the formulation (1.16) has the added advantage that it is not necessary to assume $\tilde{\mathscr{A}} \subset \mathscr{A}$.

The discussion up to this point is valid for any concurrent multiscale method for static equilibrium problems. What distinguishes the different methods is the construction of the coarse space $\tilde{\mathscr{A}}$ and the coarse energy $\tilde{E}$. The QC method achieves the coarse-graining by a finite element method. A set of representative atoms (*repatoms*) is chosen as the set of vertices of the finite element mesh. The atomistic domain is then triangulated so that each atom lies in the convex hull of $(d+1)$ repatoms and its movement is constraint by that of the repatoms. By taking every atom near a defect as a repatom but only few atoms away from it, the method achieves a continuum description of the bulk of the material while retaining a fully atomistic model of the defect.

1.4.1 Atomistic ground states

The interpolation procedure described above is only possible if the ground state of all atoms in an element can be described by a simple rule. Fortunately, such a rule is available for many solid materials. Particularly for metallic solids, the ground states usually observed have a regular arrangement of the atoms in a lattice, with small regions of impurities, called dislocations. Disregarding dislocations for the moment, the ground state of a metallic solid is usually a small perturbation of a subset of a rigid deformation of $\mathscr{L} = A\mathbb{Z}^d$ where the matrix $A \in \mathbb{R}^{d\times d}$ (which is not unique) defines the lattice orientation. The simplest lattices observed in nature are the body- centered cubic (bcc) lattice which is found, for example, in Na, K or Fe, and the face-centered cubic (fcc) also called cubic close-packed (ccp) lattice and can be found for example in Al, Cu, Au, Pb or Ag structures. The orientation matrices for bcc and fcc metals are respectively given by

$$A_{\text{bcc}} = \varepsilon \begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1 & 1/2 \\ 0 & 0 & 1/2 \end{bmatrix} \quad \text{and} \quad A_{\text{fcc}} = \varepsilon \begin{bmatrix} 1 & 1/2 & 1/2 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/2 \end{bmatrix},$$

where ε denotes the atomic spacing in the reference state.

Although regular structures such as bcc and fcc are observed in practise, little is known on a mathematically rigorous level. Only in two dimensions it was shown recently by Theil [92, Theorem 1.2], for a general class of pair potential energies, that any periodic ground state is (a rigid motion of a subset of) the triangular lattice defined

by

$$A_2 = \varepsilon \begin{bmatrix} 1 & 1/2 \\ 0 & \sqrt{3}/2 \end{bmatrix}.$$

1.4.2 Construction of the coarse space

For simplicity, we only consider the case where the reference state of the body is a perfect crystal, i.e. a subset of a regular lattice. Generalizations to defects such as dislocations or grain boundaries are possible but difficult to formulate in a mathematically rigorous framework and even more difficult to implement. Some further remarks in this direction are made in Chapter 6. Let us thus assume that the reference state of an atomistic body is a subset Ω of a regular lattice $\mathscr{L}$. This makes it possible to define an atomistic deformation as a map and we re-interpret the set of all atomistic deformations as

$$\mathscr{V} = \mathscr{V}(\Omega) = \{y : \Omega \to \mathbb{R}^d\}.$$

Let $\mathscr{N} = \{z_1, \dots, z_K\} \subset \Omega$ be the set of repatoms and let $\mathscr{T}$ be a regular partition of a subset of $\mathbb{R}^d$ into simplices with vertices in $\mathscr{N}$. The set of QC deformations is the set of 'continuous' splines of order one,

$$S^1(\mathscr{T}) = \big\{Y \in \mathscr{V}(\Omega) : \forall\, \kappa \in \mathscr{T}\ \exists\, b \in \mathbb{R}^d, F \in \mathbb{R}^{d\times d}\ \forall \xi \in \kappa \cap \Omega : Y(\xi) = b + F\xi\big\}.$$

In addition, for $Y \in S^1(\mathscr{T})$ we use $\bar{Y}$ to denote the piecewise affine interpolant of the nodal values of Y, i.e., $Y = \bar{Y}|_\Omega$, and $\nabla \bar{Y}_\kappa$ to denote its gradient in the element κ.

In practise, it is important that there are no 'holes' in an element, i.e., that $\kappa \cap \Omega = \kappa \cap \mathscr{L}$. This can be achieved by defining the domain Ω through the mesh $\mathscr{T}$, upon setting $\Omega = \mathscr{L} \cap \bigcup \mathscr{T}$.

The set $\mathscr{A}$ of admissible deformations is usually defined by 'Dirichlet boundary conditions'. To this end, let $\Omega_D \subset \Omega$ and $g\colon \Omega_D \to \mathbb{R}^d$ and define

$$\mathscr{A} = \{y \in \mathscr{V}(\Omega) : y|_{\Omega_D} = g\}.$$

The coarse set of admissible deformations is typically constructed by interpolation, i.e.,

$$\tilde{\mathscr{A}} = \big\{Y \in S^1(\mathscr{T}) : Y|_{\Omega_D \cap \mathscr{N}} = g|_{\Omega_D \cap \mathscr{N}}\big\}.$$

In practical applications it is usually entirely sufficient to assume that g is the restriction of an element of $S^1(\mathscr{T})$ to the boundary, in which case $\tilde{\mathscr{A}}$ is given by $\tilde{\mathscr{A}} = \mathscr{A} \cap S^1(\mathscr{T})$. Of course, other types of conditions are equally possible; for example, one could impose bounds on $y(\xi)$ which could describe a non-penetrable surface on which the body rests. They are easily integrated into the framework.

1.4.3 A weak Cauchy–Born rule

Mathematically, the construction of the coarse space $S^1(\mathscr{T})$ can be motivated using an approximation argument: if y is a 'smooth' deformation of Ω then there exists $Y \in S^1(\mathscr{T})$ approximating y in an appropriate sense. In the engineering literature, the construction is often justified by calling on the Cauchy–Born hypothesis which implies that a small affine deformation of a regular atomistic lattice is a stable equilibrium. The following weaker result is much easier to prove and sufficient for our purposes.

Proposition 1.2 *Let E be an EAM energy where J and ρ have a cut-off radius z_c. Let $Y \in S^1(\mathscr{T})$ be orientation-preserving and let $\xi \in \kappa \cap \Omega$ for some $\kappa \in \mathscr{T}$. If $B(Y(\xi), z_c) \subset Y(\kappa \cap \Omega)$ then*

$$\frac{\partial E}{\partial y_\xi}(Y) = 0.$$

Proof. The result follows immediately from symmetry considerations. □

It is important to be aware of this result when the presented QC model should be generalized to more complicated reference states. For example if, in its reference state, a defect is not in equilibrium, then it has to be fully triangulated. The reference state within a single element always needs to be in equilibrium. In the residual analysis in Chapter 6 the mathematical importance of this assumption will become apparent as well.

1.4.4 Summation rule approximations

Having constructed the approximation space $S^1(\mathscr{T})$, we are now confronted with the task of finding a good approximation to the energy functional E which can also be computed efficiently. To this end we modify the idea of quadrature rules used in continuum finite element analysis.

Suppose we want to approximate the sum

$$s = \sum_{\xi \in \Omega} f(\xi)$$

where $f \colon \Omega \to \mathbb{R}$. We can pick a set of summation points $\mathscr{N}_s \subset \Omega$ and summation weights $(w_z; z \in \mathscr{N}_s)$ and define

$$S = \sum_{z \in \mathscr{N}_z} w_z f(z). \tag{1.17}$$

The error $|s-S|$ can be bounded by estimates on finite differences of f. The sums that need to be approximated in the QC method are of a quite special nature, though, and an error analysis following the classical analysis of quadrature formulas is insufficient.

Several summation rules have been proposed for the QC method. Before we investigate some of them, we rewrite the atomistic energy in a more suitable form. We assume that the energy functional $E(y)$ can be written as

$$E(y) = \sum_{\xi \in \Omega} E_\xi(y), \tag{1.18}$$

where $E_\xi(y)$ is the contribution from the atom at the lattice site ξ. For example, a pair-potential energy can be written as

$$\begin{aligned} &\tfrac{1}{2} \sum_{\xi \in \Omega} \sum_{\xi' \in \Omega \setminus \{\xi\}} J(|y(\xi) - y(\xi')|) = \sum_{\xi \in \Omega} E_\xi(y), \quad \text{where} \\ &E_\xi(y) = \tfrac{1}{2} \sum_{\xi' \in \Omega \setminus \{\xi\}} J(|y(\xi) - y(\xi')|). \end{aligned}$$

Similarly, for the EAM model (1.12), (1.18) holds with

$$E_\xi(y) = F(\bar{\rho}_\xi) + \tfrac{1}{2} \sum_{\xi' \in \Omega \setminus \{\xi\}} J(|y(\xi) - y(\xi')|). \tag{1.19}$$

The local QC method. The local QC method can be used when there are no defects present in the material and only elastic deformation of the lattice occurs. Recall that all atomistic interactions, using the cut-off approximation described in §1.2, have only finite range. Hence, if a deformation $Y \in S^1(\mathscr{T})$ satisfies $Y(\xi) = b + F\xi$ in an element κ then, except near the 'boundary' of κ, the energy $E_\xi(Y)$ depends only on F but not on ξ. If the QC mesh $\mathscr{T}$ can be taken so coarse in relation to the atomic spacing that the majority of atoms lie in the bulk of the elements then the difference in energy of the atoms lying at the interfaces between elements can be neglected. The contribution then only depends on the macroscopic deformation gradient $F = \nabla \bar{Y}_\kappa$. Upon translating the atom into the origin, we can define

$$W_{\mathrm{CB}}(F) = E_0(y_F),$$

where $y_F(\xi) = F\xi$ for all $\xi \in \mathscr{L}$ and E_0 is given by (1.19) (replacing Ω by $\mathscr{L}$) if E is an EAM energy and by an appropriate formula otherwise. W_{CB} is usually called the Cauchy–Born stored energy function. The local QC method is then defined by

$$\tilde{E}(Y) = \sum_{\kappa \in \mathscr{T}} w_\kappa W_{\mathrm{CB}}(\nabla \bar{Y}_\kappa), \tag{1.20}$$

where the weights w_κ can either be defined by the number of atoms in the element κ, $w_\kappa = \sharp\kappa \cap \mathscr{L}$, or by the volume of the simplex, $w_\kappa = |\kappa|/N_{\rm ref}$ where $N_{\rm ref} = \sharp(0,1]^d \cap \mathscr{L}$.

If $\mathrm{dist}(\nabla \bar{Y}_\kappa, SO(d))$ is not too large, the element-wise summation rule (1.20) can clearly be recast in the more general framework for summation rules presented above, by taking for each $\kappa \in \mathscr{T}$ an arbitrary atom ξ near the centre of the element κ as the summation point.

Non-local QC. The local QC method loses is justification when the number of atoms in the bulk ceases to dominate the number of atoms at the surface of an element. This happens when the diameter of an element is less than, say, 50 atomic spacings.

In order to be able to reduce to a full atomistic description in a defect region, the trapezium rule is better suited. It is obtained by taking $\mathscr{N}_s = \mathscr{N}$ and choosing the summation weights such that, for functions $f \in S^1(\mathscr{T})$, the summation rule (1.17) is exact.

Cluster summation rules. Since the simple trapezium rule only accounts for atoms at the element interfaces, it tends to overestimate the actual energy. As an alternative, cluster summation rules have been introduced. For each $z \in \mathscr{N}$ we define a cluster C_z surrounding z but not overlapping with any other cluster. The set of summation points is then defined as $\mathscr{N}_s = \bigcup_{z\in\mathscr{N}} C_z$ and the summation weights are again defined by the requirement that for *affine* functions, the resulting summation rule is exact.

For smooth deformation, this rule should not differ much from the standard non-local QC method, or any other summation rule for that matter. Near the atomistic region, however, the summation rule becomes exact as, for sufficiently large clusters or sufficiently small elements, every atom becomes a summation point.

1.4.5 Analysis of the quasicontinuum method

The QC method was originally developed in two dimensions by Tadmor, Ortiz and Phillips [73] and has quickly become an important tool in nanoscale engineering applications. It was extended with adaptivity in [64]. A three-dimensional adaptive QC method is described in [55]. For a recent overview article, which also features many engineering applications, the reader may refer to [65].

Despite its growing popularity in the engineering community, the mathematical and numerical analysis of the QC method is still in its infancy. The first noteworthy analytical effort was by Lin [59] who considers the QC approximation of the reference state

(without boundary displacements or applied forces) of a one-dimensional Lennard–Jones model. He proves that the global energy minimum of the full atomistic as well as the reduced QC model lie in a region where the interaction potential is uniformly convex and uses these facts to derive an *a priori* error estimate.

E and Ming [42, 43] analyze the local QC method in the context of the heterogeneous multiscale method [41], which requires the assumption that a nearby smooth, elastic continuum solution is available. The error is estimated in terms of the atomic spacing in relation to the domain size as well as the mesh size.

In [60], Lin gives *a priori* error estimates for a modified version of the local QC method for purely elastic deformation in two dimensions without using such an assumption, but making instead a strong hypothesis (Assumptions 1. and 2. in [60]) on the exact solution of the atomistic model as well as on its QC approximation. Essentially, it is assumed what he was able to prove in one dimension, namely that both the exact and the QC solution lie in a region where the atomistic energy is convex. For lattice domains resembling smooth or convex sets this assumption seems intuitively reasonable but would be difficult to verify rigorously. For lattice domains with 'sharp', 're-entrant' boundary sections or defects we should not expect it to hold without further justification.

Finally, one should mention the work of Legoll *et al.* [15] where a multiscale method similar to the QC method is analyzed, however only nearest-neighbour interactions in one dimension are considered which makes it possible to compute the exact solutions analytically, similarly as in the present work in the proof of Lemma 2.14. So far, this has been the only work to consider defects in the analysis.

It is not too surprising that so little mathematical analysis is available for atomistic material models. For example, most techniques in continuum finite element analysis apply only if the governing equations are monotone, i.e., when the associated energy functional is convex which is grossly violated for atomistic problems. Furthermore, energy techniques such as Γ-convergence cannot be applied for two reasons: first, atomistic solutions are not global energy minima, and second, proving convergence alone is meaningless since the function space is finite-dimensional.

The goal is therefore to find new techniques that allow us to extend and unify the one-dimensional results and to make a higher-dimensional analysis possible without the overly strong assumptions made in previous work, and, in particular, allowing for the presence of defects. Only a small part of this program could be completed. Chapters 4 and 5 bring the analysis of the quasicontinuum method in one dimension to a satisfactory conclusion. In particular, it was possible to include defects in the

analysis. Furthermore, in contrast to the work of Legoll *et al.* [15], the technique employed shows a clear route to extend the results to higher dimensions. As opposed to the analysis of Lin [60], it does not require any assumptions that are difficult to justify. However, the program which is plotted out, will be very difficult to realize. Some of the results required do not seem easy to obtain. A partial extension of the *a posteriori* analysis and discussion of further possibilities and limitations is given in Chapter 6.

Chapter 2

Gradient Flows as a Selection Procedure

Motivated by the example given in §1.1.2, this chapter presents a possible concept for analyzing elastic energy functionals which do not satisfy the classical coercivity and weak lower semicontinuity conditions of the calculus of variations. The subject of study is the one-dimensional atomistic energy

$$E_{\text{atom}}\big((y_j)_{j=1}^N\big) = \sum_{j=1}^{N} \big[J(y_j - y_{j-1}) + f_j u_j\big], \tag{2.1}$$

where $N = 1, 2, \ldots$, and y_j are the positions of the atoms with $y_0 = 0$. The family (f_j) represents a linear applied force. Moreover, we assume that the Lennard–Jones type potential $J = J(z)$ satisfies

$$\begin{aligned} &J \in \mathrm{C}^2(0,\infty),\\ &J(z) = +\infty \text{ if } z \le 0 \text{ and } J(z) \to \infty \text{ as } z \to 0,\\ &J'(1) = 0,\, J''(z) > 0 \text{ in } (0, z_t), \text{ and}\\ &J \text{ is concave, increasing and bounded above in } (z_t, \infty), \end{aligned} \tag{2.2}$$

with $1 < z_t < +\infty$. The typical shape (there with a cut-off radius) is shown in Figure 1.2. The non-convexity of J lies much deeper than the geometric non-convexity of classical elasticity.

As we discussed at great length in §1.1, due to the sublinear growth of J, the energy in (2.1) should not be analyzed in terms of global minimization, as this would give unrealistic material behaviour. The most popular example given is that a material described by (2.1) would break for (almost) arbitrarily small loads if it were to attain a global minimum (cf. Proposition 1.1).

In general, for applications in mechanics, it is advantageous to consider metastable states. The difficulty here is that the number of critical points of E_{atom} tends to infinity

as $N \to \infty$. Thus, we require a selection criterion to single out the 'correct' equilibrium points. Theoretically, we should consider the natural dynamics of the material and let time tend to infinity to find its equilibrium state. Here, we take a considerably easier route and use $|\cdot|_{\mathrm{H}^1}$-gradient flow dynamics. Our justification for the gradient flow is merely to accept it as a simple model for local minimization. Concerning the choice of the metric, there are also strong mathematical reasons for choosing an $|\cdot|_{\mathrm{H}^1}$-gradient flow evolution which are outlined in Sections 2.2 and 2.3. The aim is to simply demonstrate a concept which gives physically more realistic results than the method of global minimization. The ideas in this chapter have also important applications for the numerical analysis of coarse-graining techniques such as the QC method [65], as they give an indication how numerical optimization methods can be stabilized (cf. Chapter 5).

The main goal of this chapter is to show that the $|\cdot|_{\mathrm{H}^1}$-gradient flow provides a selection criterion for critical points which results in good qualitative properties of the resulting equilibrium model. The simplicity of the one-dimensional model problem makes it possible to give complete results. While some techniques applied here are quite general, a generalization of the entire presentation to higher dimensions seems non-trivial; the related challenges are discussed in §2.5.

As an application of the idea to use gradient flows to analyze equilibrium points of non-convex energies, we consider the continuum limit of a rescaled version of the atomistic functional E_{atom} as the number of atoms N tends to infinity. The novelty is that we primarily consider the convergence of the gradient flow evolutions (Theorem 2.9), and obtain the convergence of the equilibria almost as an afterthought (cf. Theorem 2.13). This procedure gives a different and, one might argue, more realistic continuum limit than previous work; see §2.1 for a more extensive discussion. In addition, it shows that there is a strong relationship between the atomistic and continuum equilibria.

The local minimizers selected by the gradient flow are weak local minimizers, i.e., local minimizers with respect to the $\mathrm{W}^{1,\infty}$-norm. It is clear from the shape of the interaction potential (cf. Figure 1.2) and the comments at the end of §2.4 that this is in fact the only possibility. In any weaker topology, even the elastic critical points are not local minimizers of the energy. The same is true for fractured states but the interpretation of $\mathrm{W}^{1,\infty}$ would be more subtle in this case.

If we replace the Lennard–Jones potential by a potential which is smooth at the origin and therefore J' is Lipschitz-continuous, then the convergence analysis of the gradient flow requires only minor modifications of the classical convergence analysis of Galerkin discretizations of parabolic equations. For the approach in this chapter,

however, convergence of the energy is sufficient (cf. Theorem 2.5), which makes a result as general as Theorem 2.9 possible. To achieve this we use some ideas from [5, Chapter 4].

For the analysis of equilibria, we use a liminf condition for the slope of a family of functionals, the proof of which is based on the notion of λ-convexity. This condition was also used in Sandier and Serfaty [87] to analyse the convergence of gradient flows. Using the techniques in their paper, which has a different aim than the present work, the convergence would have to be obtained by compactness principles (which are not available in our case) rather than λ-convexity.

The analytical methods used in this chapter (particularly the notion of λ-convexity (2.6) and the evolutionary variational inequality (2.11)) are a direct adaption of the ideas from Chapter 4 and partly from Chapters 1–3 of the work of Ambrosio, Gigli and Savaré [5]. While those authors are primarily interested in the convergence of time discretizations of gradient flows and questions of existence and uniqueness, the goal here is to analyze the convergence of gradient flow evolutions along a family of functionals (see Theorem 2.5).

2.1 Continuum Limits of Atomistic Energies

Continuum limits of atomistic models have been studied by many authors in the past. Because it is customary, we consider the case of Dirichlet boundary conditions in this section only. To be able to compute a continuum limit we need to first rescale the energy (2.1) to a fixed, finite domain. The seemingly naive approach is to use a linear scaling of the energy as well as the boundary condition, which gives

$$E_N^{(1)}\big((y_j)_{j=0}^N\big) = \sum_{j=1}^{N} \frac{1}{N} J\big(N(y_j - y_{j-1})\big), \quad y_0 = 0, y_N = 1 + \delta. \tag{2.3}$$

If we assume that the body attains its global energy minimum then, for arbitrarily small boundary displacement δ, the deformation will not be a continuum state (compare §1.1 and Proposition 1.1). This fact is reflected by the Γ-limit of $E_N^{(1)}$ as $N \to \infty$ (see for example [18, 19] and references therein) which gives the energy

$$E^{(1)}(y) = \int_0^1 J^{**}(y')\,\mathrm{d}x, \qquad y(0) = 0, y(1) = 1 + \delta,$$

where J^{**} is the convex envelope of J.

Motivated by an analysis quite similar to Proposition 1.1, it was noticed by Braides *et al.* [20] that, if a different scaling is used, the Γ-limit becomes more interesting. If we define

$$E_N^{(2)}\big((u_j)_{j=0}^N\big) = \sum_{j=1}^N \Big[J\big(1+\sqrt{N}(u_j-u_{j-1})\big) - J(1)\Big], \quad u_0 = 0, u_N = \delta$$

then the Γ-limit turns out to be the Griffith functional [48],

$$G(u) = \alpha \int_0^1 |u'|^2 \,\mathrm{d}x + \beta \sharp S_u, \quad u(0) = 0, u(1) = \delta,$$

where S_u is the set of jump-discontinuities of the displacement u, $\alpha = \frac{1}{2}J''(1)$ and $\beta = \lim_{z\to\infty} J(z) - J(1)$. The boundary values of the possibly discontinuous functions u can be interpreted in a meaningful way. While it is interesting that the Griffith functional can be obtained in this way, it should be noted that this model is typically used for crack propagation only, not crack initiation. In one dimension, however, only crack initiation can be analyzed.

The philosophy adopted in this work is that the scaling of the functional $E_N^{(1)}$ is actually the natural one; only the process of passing to the continuum limit is flawed. It will be shown that, if the continuum limit is analyzed in terms of an appropriate evolution, then the resulting model is in fact a very realistic candidate.

One of the problems addressed in this chapter (cf. §2.4), is to find the stable equilibrium that the material would 'naturally' assume if we started in the reference configuration $y_i^N = x_i^N$, or a perturbation of it, and then applied forces. In Theorem 2.13 we show that the resulting equilibria represent the correct elastic behaviour. For this reason we prefer to work with surface forces rather than a prescribed displacement. This is, however, not a restriction. The entire convergence theory can also be repeated for Dirichlet conditions applied at both ends of the interval.

Connections to other models. Closest in spirit to the approach advocated here is the work by Blanc *et al.* [16]. Except for the fact that they consider far more complicated atomistic interactions in three dimensions, their continuum limit is the same. In fact, the work in this chapter may be seen as a small step towards a rigorous justification of the approach taken in [16]. From the point of view of numerical analysis, strong connections can be drawn to the local version of the quasicontinuum method [65]. In this respect, the results of E and Ming [43] have some similarities to our own.

In both of these works, the concept of global minimization of the energy is rejected and alternative means are sought to analyze equilibria of elastic energy functionals.

A similar approach is also taken by Rieger and Zimmer [83], who use a time-discrete gradient flow evolution of Young-measures to analyze material damage. In the slightly different setting of viscoelasticity [9, 80], it is shown that dynamics can prevent the formation of finer and finer microstructure and therefore the attainment of a global energy minimum.

The model presented here is not to be confused, however, with quasistatic or rate independent evolutions (see for example [36, 48] for fracture, or [27] for plasticity). In their time-discrete form, at every timestep, an equilibrium (typically a minimum) of a functional of the form

$$D(u_{j-1}, u) + E(u) \tag{2.4}$$

is sought, where D is a so-called dissipation metric. Rather, the gradient flow model we present here should be understood as a mechanism to find the equilibrium in the quasistatic evolution (2.4).

For results on the continuum manifestation of some further interesting atomistic effects such as finite-range interactions, the reader is referred to [29, 93].

Outline of the chapter. We begin in Section 2.2 by outlining the theoretical tools for the convergence analysis, a theory of gradient flows based on the notion of λ-convexity, and a corresponding approximation theory. We also review the notion of slope which is used to define the concept of critical points.

In Section 2.3, we prove the convergence of an atomistic gradient flow evolution to the $|\cdot|_{\mathrm{H}^1}$-gradient flow of a non-convex functional defined on H^1, giving a new type of continuum limit for atomistic functionals.

Finally, in Section 2.4, we analyze the resulting equilibrium solutions which are obtained when $t \to \infty$ in the gradient flow. We consider the case of small loads and show that the equilibria obtained are the physically observed elastic deformations and not the 'fractured' global energy minima.

Some numerical experiments in the fracture case are shown in [74].

2.2 Approximation of Gradient Flows of Non-Convex Energies

Let $\mathscr{H}$ be a Hilbert space with inner product $(\cdot,\cdot)$ and norm $\|\cdot\|$, let $\mathscr{A}$ be a closed convex subset of $\mathscr{H}$, and let $\phi\colon \mathscr{H} \to (-\infty, \infty]$. If ϕ is Fréchet differentiable at a point u, we denote the representation of its derivative (its gradient) by $\phi'(u)$. Second order derivatives are denoted by $\phi''(u; v_1, v_2)$. We denote the domain of definition of ϕ

by $D(\phi) = \{u \in \mathscr{H} : \phi(u) < \infty\}$. In fact, by using the convention $+\infty \leq +\infty$, we do not make much explicit use of the domain of definition. For example, the functional ϕ is convex if, and only if, $D(\phi)$ is convex and ϕ is convex in $D(\phi)$.

2.2.1 Evolutionary variational inequalities

Naively, we may call a curve $u \in \mathrm{C}^1(a, b; \mathscr{H})$ a gradient flow of ϕ, if

$$\dot{u}(t) = -\phi'(u(t)) \quad \forall t \in (a, b). \tag{2.5}$$

Equation (2.5) in infinite-dimensional spaces is usually restated only for convex functionals ϕ. The natural condition on ϕ, under which a considerable part of the theory of gradient flows for convex functionals can be recovered, is the condition of λ-convexity [5]. Let $\mathscr{A}$ be a closed, convex subset of $\mathscr{H}$. We say that ϕ is λ-convex in $\mathscr{A}$, for some $\lambda \in \mathbb{R}$, if

$$\phi\big((1-t)v_0 + tv_1\big) \leq (1-t)\phi(v_0) + t\phi(v_1) - \frac{\lambda}{2}t(1-t)\|v_0 - v_1\|^2 \\ \forall v_0, v_1 \in \mathscr{A}, \forall t \in (0,1). \tag{2.6}$$

To obtain a better feel for the meaning of λ-convexity, consider the following simple Proposition.

Proposition 2.1

(a) The functional ϕ is λ-convex in $\mathscr{A}$ if, and only if, $u \mapsto \phi(u) - \frac{\lambda}{2}\|u\|^2$ is convex in $\mathscr{A}$.

(b) One-sided Lipschitz continuity of the gradient:
If ϕ is differentiable at every point of $\mathscr{A}$ and satisfies

$$\big(\phi'(v_1) - \phi'(v_0), v_1 - v_0\big) \geq \lambda\|v_1 - v_0\|^2 \quad \forall\, v_1, v_0 \in \mathscr{A}, \tag{2.7}$$

then ϕ is λ-convex in $\mathscr{A}$.

(c) Boundedness below of the Hessian:
If ϕ is twice differentiable at every non-extremal point of $\mathscr{A}$ and

$$\phi''(u; v-u, v-u) \geq \lambda\|v-u\|^2 \quad \forall u, v \in \mathscr{A}, \tag{2.8}$$

then ϕ is λ-convex in $\mathscr{A}$. Moreover, if $D(\phi)$ is open and convex then ϕ is λ-convex in $D(\phi)$ if, and only if, $\phi''(u; v, v) \geq \lambda\|v\|^2$ for all $u \in D(\phi)$ and $v \in \mathscr{H}$.

(d) If $\phi = \phi_1 + \phi_2$, where $\phi_i \colon \mathscr{A} \to (-\infty, +\infty]$, ϕ_1 is λ_1-convex and ϕ_2 is λ_2-convex, then ϕ is $(\lambda_1 + \lambda_2)$-convex.

Proof. Throughout the proof, let $v_0, v_1 \in \mathscr{A}$, $v_t = (1-t)v_0 + tv_1$, and $F(v) = \phi(v) - \frac{\lambda}{2}\|v\|^2$.

The crucial observation is that $u \mapsto \frac{1}{2}\|u\|^2$ is 1-convex, in fact, we even have

$$\frac{1}{2}\|v_t\|^2 = (1-t)\frac{1}{2}\|v_0\|^2 + t\frac{1}{2}\|v_1\|^2 - \frac{1}{2}t(1-t)\|v_0 - v_1\|^2. \tag{2.9}$$

Suppose now that ϕ is λ-convex. Using (2.9), we have

$$\begin{aligned} F(v_t) &\leq (1-t)\phi(v_0) + t\phi(v_1) - \frac{\lambda}{2}t(1-t)\|v_0 - v_1\|^2 \\ &\quad -(1-t)\frac{\lambda}{2}\|v_0\|^2 - t\frac{\lambda}{2}\|v_1\|^2 + \frac{\lambda}{2}t(1-t)\|v_0 - v_1\|^2 \\ &= (1-t)F(v_0) + tF(v_1). \end{aligned}$$

On the other hand, if F is convex we may traverse the above inequality in the opposite direction, to obtain that ϕ is λ-convex.

The derivative of the mapping $v \mapsto \|v\|^2/2$ is $(v, \cdot)$, its second derivative is the inner product $(\cdot, \cdot)$. If ϕ satisfies (2.7) then

$$\big(F'(v_1) - F'(v_0), v_1 - v_0\big) \;=\; \big(\phi'(v_1) - \phi'(v_0), v_1 - v_0\big) - \lambda\|v_1 - v_0\|^2 \;\geq\; 0$$

which is equivalent to F being convex. If ϕ is twice differentiable and satisfies (2.8) then

$$F''(u; v-u, v-u) = \phi''(u; v-u, v-u) - \lambda\|v-u\|^2 \geq 0 \qquad \forall v \in \mathscr{A},$$

for all points $u \in \mathscr{A}$ which are not extremal and hence F is convex.

Conversely, if ϕ is λ-convex and twice differentiable in $D(\phi)$, and if $D(\phi)$ is open and convex then, for all $u \in D(\phi)$, for all $v \in \mathscr{H}$ and for sufficiently small s,

$$\phi(u) \leq \frac{1}{2}\phi(u+sv) + \frac{1}{2}\phi(u-sv) - s^2\frac{\lambda}{2}\|v\|^2,$$

where $t = 1/2$, $u = v_t$, $v_0 = u + sv$ and $v_1 = u - sv$ in (2.6). This inequality can be rearranged to

$$s^{-2}\Big(\phi(u+sv) - 2\phi(u) + \phi(u-sv)\Big) \geq \lambda\|v\|^2,$$

which, as $s \to 0$ gives the second result of item (c).

The last statement of the proposition is trivial. □

Note in particular that Proposition 2.1 gives us a characterization of λ-convexity in terms of the eigenvalues of ϕ''. If $\mathscr{A} = \mathscr{H}$ and if ϕ is twice differentiable then the largest value $\underline{\lambda}$ for which ϕ may be $\underline{\lambda}$-convex is given by

$$\underline{\lambda} = \inf_{u\in\mathscr{A}} \inf_{\substack{v\in\mathscr{H}\\ \|v\|=1}} \phi''(u;v,v),$$

which is the smallest eigenvalue of ϕ'' with respect to the norm $\|\cdot\|$.

Another important tool in the analysis of gradient flows in metric spaces is the local slope, $|\partial\phi|$. It measures the maximal descent of the functional ϕ at a given point. For example, if ϕ is differentiable at u then $|\partial\phi|(u) = \|\phi'(u)\|$. For a general functional ϕ, it is defined as

$$|\partial\phi|(u) = \limsup_{\substack{v\in\mathscr{A}\\ v\to u}} \frac{\big[\phi(u)-\phi(v)\big]^+}{\|v-u\|}. \tag{2.10}$$

With the help of the slope, we can compute a strong bound on the decay of λ-convex functionals.

Lemma 2.2 *Let ϕ be λ-convex in $\mathscr{H}$ and $u \in D(\phi)$, then*

$$\phi(v) \geq \phi(u) - |\partial\phi|(u)\|u-v\| + \frac{\lambda}{2}\|u-v\|^2 \qquad \forall v\in\mathscr{A}.$$

Proof. The definition of λ-convexity (2.6), taking $v_1 = v$ and $V_2 = u$, can be rewritten as

$$\begin{aligned}\phi(v) &\geq \phi(u) - \frac{\phi(u)-\phi((1-t)u+tv)}{t} + \frac{\lambda}{2}(1-t)\|u-v\|^2\\ &\geq \phi(u) - \frac{\big[\phi(u)-\phi((u+t(v-u))\big]^+}{t\|v-u\|}\,\|v-u\| + \frac{\lambda}{2}(1-t)\|u-v\|^2.\end{aligned}$$

Letting $t\to 0$ gives the desired bound. □

If a functional is λ-convex, then its gradient flows have an alternative characterization. Suppose that a curve $u\in \mathrm{C}^1(a,b;\mathscr{H})$ satisfies (2.5), where ϕ' satisfies (2.7). First, we write

$$\begin{aligned}\phi(u(t))-\phi(v) &= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}\phi\big(v+s(u(t)-v)\big)\,\mathrm{d}s\\ &= \int_0^1 \phi'\big(v+s(u(t)-v);u(t)-v\big)\,\mathrm{d}s \qquad \forall v\in\mathscr{H}.\end{aligned}$$

Combing this with (2.5), tested with $u(t)-v$, gives

$$\begin{aligned}&\big(\dot u(t),u(t)-v\big)+\phi(u(t))-\phi(v)\\ &\qquad= \int_0^1 \Big[\phi'\big(u(t)-(1-s)(u(t)-v)\big)-\phi'(u(t))\Big]\big(u(t)-v\big)\,\mathrm{d}s\\ &\qquad\leq -\lambda\|u(t)-v\|^2\int_0^1(1-s)\,\mathrm{d}s.\end{aligned}$$

Hence, we have shown that u also satisfies the evolutionary variational inequality

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u(t)-v\|^2+\frac{\lambda}{2}\|u(t)-v\|^2+\phi(u(t))\le\phi(v)\quad\forall\, v\in\mathscr{H},\forall\, t\in(a,b).$$

This inequality is the basis for a general theory of gradient flows in metric spaces, then called curves of maximal slope, developed in [5, Chapter 4]. Note, for example, that it makes sense to consider $u, v \in \mathscr{A}$ only, instead of all of $\mathscr{H}$. Theorem 2.3 is a translation of [5, Theorem 4.0.4] to the Hilbert space setting which is sufficient for our purposes.

Theorem 2.3 (Existence and uniqueness) *Let $\mathscr{A}$ be a closed, convex subset of a Hilbert space $\mathscr{H}$ and let $\phi\colon \mathscr{A} \to (-\infty,\infty]$ be (strongly) lower semi-continuous and λ-convex. For each $u_0 \in D(\phi)$, there exists a locally Lipschitz-continuous curve $u\colon [0,\infty) \to \mathscr{A}$ which is the unique solution of*

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u(t)-v\|^2+\frac{\lambda}{2}\|u(t)-v\|^2+\phi(u(t))\le\phi(v)\quad\forall\, v\in\mathscr{A},\ \text{for a.e. } t>0, \tag{2.11}$$

among all curves $v \in \mathrm{AC}_{\mathrm{loc}}(0,\infty;\mathscr{A})$, satisfying $v(0+) = u_0$.

Furthermore, $u(t)$ is a curve of maximal slope, i.e., $\phi(u(t)) \in \mathrm{AC}_{\mathrm{loc}}([0,+\infty))$ and

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi(u(t))\le-\frac{1}{2}\|\dot u\|^2-\frac{1}{2}|\partial\phi|^2(u)\qquad\text{for a.e. } t\in(0,\infty). \tag{2.12}$$

For the remainder of the chapter, we shall use the following definition of gradient flow.

Definition 2.4 *Let $\mathscr{A}$ be a convex, closed subset of a Hilbert space $\mathscr{H}$ and $\phi\colon \mathscr{A} \to (-\infty,\infty]$ a lower semi-continuous and λ-convex functional. We say that a locally Lipschitz-continuous curve $u\colon [0,\infty) \to \mathscr{A}$ is a gradient flow of ϕ in $\mathscr{A}$, if it satisfies* (2.11).

Remarks. 1. If ϕ is λ-convex then $|\partial\phi|$ is a metric subgradient for ϕ (cf. [5]). Hence, if (2.12) is satisfied then

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi(u(t))\le-\frac{1}{2}\|\dot u(t)\|^2-\frac{1}{2}|\partial\phi|(u(t))\le-\|\dot u(t)\||\partial\phi|(u(t))\le\frac{\mathrm{d}}{\mathrm{d}t}\phi(u(t)). \tag{2.13}$$

All inequalities must be equalities and therefore $\|\dot u(t)\| = |\partial\phi|(u(t))$ for a.e. t. This shows in particular that under appropriate smoothness condition (2.5), (2.11) and (2.12) are equivalent. ◀

2. In Definition 2.4 the term 'lower semi-continuous' refers to lower semi-continuity with respect to the strong topology of $\mathscr{H}$. Numerous examples of gradient flows of

convex functionals can be given, where the functional is not continuous with respect to the metric with respect to which the gradient flow is computed. The most common example is of course the Heat equation which is the gradient flow of the Dirichlet functional $\phi(u) = \frac{1}{2}\int_\Omega |\nabla u|^2 \,\mathrm{d}x$ with respect to the L^2-norm.

Also, in our situation, the energies (2.31) and (2.33) are not continuous with respect to H^1-convergence but they are lower semi-continuous (see Lemma 2.10). To see, for example, that E is not strongly continuous, take a deformation y with finite energy and modify it in a set A_δ such that $|A_\delta| = \delta$ to obtain a deformation y_δ such that $y_\delta' = y'$ in $(0,1) \setminus A_\delta$ and $y_\delta' = 0$ in A_δ. Clearly, $y_\delta \to y$ in H^1 but $E(y_\delta) = +\infty$ for all δ. Immediately, one can also construct sequences $y_j \to y$ for which $E(y_j)$ is finite but $E(y_j) \to \infty$ as $j \to \infty$. Corresponding constructions can also be made for the functionals E_N.

Since a strong control on the variable is given by the evolution (which is a gradient flow with respect to the Hilbert space norm) only strong lower semi-continuity is required in all proofs. This is in contrast to the direct method of the calculus of variations, where the usage of compactness arguments requires weak lower semi-continuity. In fact, since J is non-convex, the functional (2.31) is not weakly lower semi-continuous. ◀

2.2.2 Approximation of gradient flows

Based on the evolutionary variational inequality (2.11), an abstract convergence theory for gradient flows in a general metric setting for λ-convex functionals was developed in [77]. Theorem 2.5 below is one result therein which is relevant for the Hilbert space setting in the present work.

Theorem 2.5 *Let $\mathscr{A}$ be a closed, convex subset of a Hilbert space $\mathscr{H}$ and, for $N \in \mathbb{N}$, let $\phi, \phi_N \colon \mathscr{A} \to (-\infty, \infty]$ be functionals defined on $\mathscr{A}$. Let $u^0 \in D(\phi)$ and $u_N^0 \in D(\phi_N)$ be given initial values, and assume that the following conditions are satisfied:*

(i) Lower Semi-Continuity: *The functionals ϕ and ϕ_N ($N \in \mathbb{N}$) are (strongly) lower semi-continuous.*

(ii) Uniform λ-Convexity: *There exists $\lambda \in \mathbb{R}$, such that ϕ and ϕ_N, $N \in \mathbb{N}$, are λ-convex.*

(iii) Equi-Coercivity: *There exists $\gamma \geq 0$ such that*

$$\inf_{N\in\mathbb{N}} \inf_{v\in\mathscr{A}} \left[\phi_N(v) + \gamma\|v - u_N^0\|^2\right] = m^* > -\infty.$$

(iv) Convergence of the initial data: $\sup_{N\in\mathbb{N}} \phi_N(u_N^0) = M^* < \infty$ *and* $\|u_N^0 - u^0\| \to 0$ *as* $N \to \infty$.

(v) Consistency: *There exists a constant* $c_1 > 0$ *such that, for a.e.* $t \in (0,\infty)$,

$$\limsup_{N\to\infty} \big(\phi(u_N) - \phi_N(u_N)\big) \le 0, \text{ and } \phi(u_N) \le c_1(1 + [\phi_N(u_N)]^+ + \|u_N\|^2).$$

(vi) Best approximation error: *For every* $N \in \mathbb{N}$, *there exists a Borel-measurable curve* $v_N\colon (0,\infty) \to \mathscr{A}$, *so that* $v_N \to u$ *in* $\mathrm{L}^2_{\mathrm{loc}}([0,\infty);\mathscr{H})$, *and*

$$\phi_N(v_N(t)) \to \phi(u(t)) \text{ and } \phi_N(v_N(t)) \le c_2(1 + [\phi(u(t))]^+ + \|u(t)\|^2) \qquad \text{for a.e. } t,$$

where u *is the gradient flow of* ϕ *with initial data* u^0.

Then the gradient flows (in the sense of Definition 2.4) u_N *of* ϕ_N *with initial values* u_N^0 *converge in* $\mathrm{L}^\infty_{\mathrm{loc}}([0,\infty);\mathscr{H})$ *to the gradient flow* u *of* ϕ *with initial value* u^0.

Proof. Let u and u_N respectively satisfy

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u(t)-v\|^2 + \frac{\lambda}{2}\|u(t)-v\|^2 + \phi(u(t)) \le \phi(v) \quad \forall\, v \in \mathscr{A}, \text{ and} \tag{2.14}$$

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u_N(t)-v_N\|^2 + \frac{\lambda}{2}\|u_N(t)-v_N\|^2 + \phi_N(u_N(t)) \le \phi_N(v_N) \quad \forall\, v_N \in \mathscr{A}. \tag{2.15}$$

The existence of these curves is guaranteed by conditions (i) and (ii). We test (2.14) with $v = u_N$ and, since typically $D(\phi_N) \subsetneq D(\phi)$, choose a recovery sequence (v_N), i.e., a sequence satisfying (vi) to test (2.15). Adding (2.14) and (2.15) gives

$$\begin{aligned}\frac{1}{2}\Big[\frac{\mathrm{d}}{\mathrm{d}t}\Big(\|u(t)-u_N(s)\|^2 + \|u_N(t)-v_N(s)\|^2\Big)\Big]_{s=t} \\ + \frac{\lambda}{2}\Big(\|u(t)-u_N(t)\|^2 + \|u_N(t)-v_N(t)\|^2\Big) \\ + \phi(u(t)) + \phi_N(u_N(t)) \le \phi(u_N(t)) + \phi_N(v_N(t)).\end{aligned} \tag{2.16}$$

We now add and subtract terms in order to bring (2.16) into a form amenable to an application of Gronwall's inequality. Since u and u_N are locally Lipschitz continuous, we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\|u(t)-u_N(t)\|^2 = \Big[\frac{\mathrm{d}}{\mathrm{d}t}\Big(\|u(t)-u_N(s)\|^2 + \|u_N(t)-u(s)\|^2\Big)\Big]_{s=t} \qquad \text{for a.e. } t \in (0,\infty). \tag{2.17}$$

We can use (2.17) to rearrange (2.16) as

$$\begin{aligned}
&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u(t)-u_N(t)\|^2+\lambda\|u(t)-u_N(t)\|^2\\
&\qquad\leq\ \Big(\phi_N(v_N(t))-\phi(u(t))\Big)+\Big(\phi(u_N(t))-\phi_N(u_N(t))\Big)\\
&\qquad\quad+\frac{\lambda}{2}\Big(\|u(t)-u_N(t)\|^2-\|v_N(t)-u_N(t)\|^2\Big)\\
&\qquad\quad+\frac{1}{2}\Big[\frac{\mathrm{d}}{\mathrm{d}t}\Big(\|u_N(t)-u(s)\|^2-\|u_N(t)-v_N(s)\|^2\Big)\Big]_{s=t}\\
&=\ E_1+E_2+E_3+E_4.
\end{aligned}\tag{2.18}$$

The term E_3 can be estimated using the inverse triangle inequality and Cauchy's inequality, which give

$$\begin{aligned}
E_3 &= \frac{\lambda}{2}\Big(\|u_N-u\|+\|u_N-v_N\|\Big)\Big(\|u_N-u\|-\|u_N-v_N\|\Big)\\
&\leq \frac{|\lambda|}{2}\Big(2\|u_N-u\|+\|u-v_N\|\Big)\|u-v_N\|\\
&\leq \frac{|\lambda|}{2}\|u_N-u\|^2+|\lambda|\|u-v_N\|^2.
\end{aligned}\tag{2.19}$$

We invoke the identity

$$\|u_N(t)-u(s)\|^2-\|u_N(t)-v_N(s)\|^2=2(u_N(t),u(s)-v_N(s))+\|u(s)\|^2-\|v_N(s)\|^2$$

to deduce that

$$E_4\leq\|\dot{u}_N\|\|u-v_N\|.\tag{2.20}$$

Combining (2.18) with (2.19) and (2.20), we arrive at

$$\begin{aligned}
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u-u_N\|^2+\frac{\tilde{\lambda}}{2}\|u-u_N\|^2 \ &\leq\ \big(\phi_N(v_N)-\phi(u)\big)+\big(\phi(u_N)-\phi_N(u_N)\big)\\
&\quad+\frac{|\lambda|}{2}\|v_N-u\|^2+\frac{1}{2}\|\dot{u}_N\|\|v_N-u\|,
\end{aligned}$$

where $\tilde{\lambda}=\lambda-|\lambda|/2$. Using Gronwall's inequality, we obtain

$$\begin{aligned}
\mathrm{e}^{2\tilde{\lambda}T}\|u(T)-u_N(T)\|^2&\leq\|u(0)-u_N(0)\|^2+2\int_0^T\mathrm{e}^{2\tilde{\lambda}t}\big(\phi(u_N)-\phi_N(u_N)\big)\,\mathrm{d}t\\
&\quad+2\int_0^T\mathrm{e}^{2\tilde{\lambda}t}\Big[\big(\phi_N(v_N)-\phi(u)\big)+|\lambda|\|v_N-u\|^2+\|\dot{u}_N\|\|v_N-u\|\Big]\,\mathrm{d}t.
\end{aligned}\tag{2.21}$$

Using (2.24) and condition (v) we obtain an integrable bound on $\max(\phi(u_N)-\phi_N(u_N),0)$. We can therefore apply Fatou's lemma to deduce that

$$\limsup_{N\to\infty}\int_0^T\big(\phi(u_N)-\phi_N(u_N)\big)\,\mathrm{d}t\leq 0.$$

Similarly, it follows from (vi) that

$$\begin{aligned}\limsup_{N\to\infty}\int_0^T \big(\phi_N(v_N)-\phi(u)\big)\,\mathrm{d}t &\le 0, \quad \text{and}\\ \lim_{N\to\infty}\int_0^T \|v_N-u\|^2\,\mathrm{d}t &= 0.\end{aligned}$$

For the last term in (2.21) we apply the Cauchy–Schwarz inequality and the stability estimate (2.23) to obtain the bound

$$\int_0^T \|\dot u_N\|\|u-v_N\|\,\mathrm{d}t \le \left(C\int_0^T \|u-v_N\|^2\,\mathrm{d}t\right)^{1/2}$$

which also tends to zero as $N\to\infty$.

Finally, using (iv) to show that the initial condition converges, we can deduce that, for each $T>0$, $\sup_{t\le T}\|u_N(t)-u(t)\|\to 0$, as $N\to\infty$. □

The following stability estimates were already given in [5]. The proof is repeated here, in order to highlight the independence of the constants of N.

Lemma 2.6 *Suppose that the hypotheses (ii) and (iii) of Theorem 2.5 are satisfied; then there exist constants C_1, C_2 depending only on γ, m^* and M^* such that*

$$\|u_N(T)-u_N(0)\|^2 \le C_1\mathrm{e}^{C_2T}, \tag{2.22}$$

$$\int_0^T \|\dot u_N(t)\|^2\,\mathrm{d}t \le M^*-m^*+\gamma C_1\mathrm{e}^{C_2T}, \text{ and} \tag{2.23}$$

$$\phi_N(u_N(T)) \ge m^*-\gamma C_1\mathrm{e}^{C_2T}. \tag{2.24}$$

Proof. Condition (iii) of Theorem 2.5 gives the estimate

$$\phi_N(u_N(T)) \ge m^*-\gamma\|u_N(T)-u_N(0)\|^2. \tag{2.25}$$

Using (2.12) and (2.13), we can bound the L^2-norm of the velocity $\dot u_N$ by

$$\int_0^T \|\dot u_N\|^2\,\mathrm{d}t \le \phi_N(u_N(0))-\phi(u_N(T)) \le M^*-m^*+\gamma\|u_N(T)-u_N(0)\|^2. \tag{2.26}$$

We use (2.26) to compute a bound on $\|u_N(T)-u_N(0)\|^2$,

$$\begin{aligned}\frac{1}{2}\|u_N(T)-u_N(0)\|^2 &= \frac{1}{2}\int_0^T \frac{\mathrm{d}}{\mathrm{d}t}\|u_N(t)-u_N(0)\|^2\,\mathrm{d}t\\ &\le \int_0^T \|\dot u_N\|\|u_N(t)-u_N(0)\|\,\mathrm{d}t\\ &\le \frac{\varepsilon}{2}\int_0^T \|\dot u_N(t)\|^2\,\mathrm{d}t+\frac{1}{2\varepsilon}\int_0^T \|u_N(t)-u_N(0)\|^2\,\mathrm{d}t\\ &\le \frac{\varepsilon}{2}\Big[M^*-m^*+\gamma\|u_N(T)-u_N(0)\|^2\Big]+\frac{1}{2\varepsilon}\int_0^T \|u_N(t)-u_N(0)\|^2\,\mathrm{d}t.\end{aligned}$$

We can clearly choose ε, depending only on γ, to obtain constants C_1 and C_2 such that

$$\|u_N(T) - u_N(0)\|^2 \leq C_1 + C_2 \int_0^T \|u_N(t) - u_N(0)\|^2 \,\mathrm{d}t \qquad \forall T > 0.$$

An application of Gronwall's inequality gives (2.22).

Upon inserting (2.22) into respectively (2.26) and (2.25), we obtain (2.23) and (2.24). □

To conclude this section, we state a result from [5, Theorem 4.0.4 (v)], on the implicit Euler approximation of a gradient flow, which we will use frequently in Section 2.4.

Lemma 2.7 *Let ϕ be λ-convex in $\mathscr{A}$ and let $t_i = i\tau$, for $i = 0, 1, \ldots$, define a partition of $[0, \infty)$, with $0 < \tau < 1/\min(0, -\lambda)$. Let $u_0 \in \mathscr{A}$, and let the family $(u_i)_{i=1,2,\ldots}$ be defined by*

$$u_i = \mathrm{argmin}_{\mathscr{A}} \left[v \mapsto \frac{\|v - u_{i-1}\|^2}{2\tau} + \phi(v) \right].$$

Let $u(t)$ be the gradient flow of ϕ with $u(0) = u_0$ and let $\bar{u}_\tau(t)$ be the piecewise constant interpolant of (u_i), i.e.,

$$\bar{u}_\tau(0) = 0 \quad \text{and} \quad \bar{u}_\tau(t) = u_i \quad \text{if } t_{i-1} < t \leq t_i.$$

Then, $\bar{u}_\tau(t) \to u(t)$ in $\mathrm{L}^\infty_{\mathrm{loc}}([0, \infty), \mathscr{H})$, as $\tau \to 0$.

2.2.3 The slope

So far, we have defined and analysed gradient flow evolutions. However, we are mostly interested in analyzing the resulting equilibria, which can often be obtained by letting time tend to infinity. A natural concept of equilibrium, or critical point, is given by the local slope which we have defined in (2.10). We say that $u^* \in \mathscr{H}$ is a critical point of the functional ϕ, if $|\partial\phi|(u^*) = 0$. This is a necessary and sufficient condition for u^* to be a stationary point of the gradient flow. Note also that the functional (2.33) which we will analyze is not differentiable in H^1 and that the notion of slope is a genuine extension. The following lemma can be used in some situations to show that an accumulation point of the sequence of critical points of approximate functionals ϕ_N must again be a critical point. This result is still true in metric spaces [77]. Note that conditions (2.27) and (2.28) describe Γ-convergence (cf. [37, 39]) of the family ϕ_N in the strong topology of $\mathscr{H}$ with limit ϕ.

Lemma 2.8 *Let $\mathscr{A}$ be a closed convex subset of a Hilbert space, let $\phi, \phi_N \colon \mathscr{A} \to (-\infty, \infty]$ be λ-convex, with a uniform λ, and suppose that the conditions*

$$v_N \to v \Rightarrow \phi(v) \le \liminf_{N\to\infty} \phi_N(v_N) \tag{2.27}$$

$$\forall v \in \mathscr{A}\ \exists (v_N)_{N\in\mathbb{N}} \subset \mathscr{A} \text{ s.t. } v_N \to v \text{ and } \phi(v) = \lim_{N\to\infty} \phi(v_N) \tag{2.28}$$

are satisfied. Then, the slopes satisfy the liminf *condition*

$$u_N \to u \Rightarrow |\partial\phi|(u) \le \liminf_{N\to\infty} |\partial\phi_N|(u_N). \tag{2.29}$$

Proof. The crucial observation [5, Theorem 2.4.9] is that for λ-convex functionals, the slope can be rewritten as

$$|\partial\phi|(u) = \sup_{v\neq u} \left[\frac{\phi(u)-\phi(v)}{\|u-v\|} + \frac{\lambda}{2}\|u-v\|^2\right]^+.$$

Let $u_N \to u$, and for some fixed $v \neq u$ let $(v_N)_{N\in\mathbb{N}}$ be a recovery sequence for v, satisfying (2.28). Then, for sufficiently large N, we have

$$\begin{aligned}
&\left[\frac{\phi(u)-\phi(v)}{\|u-v\|} + \frac{\lambda}{2}\|u-v\|^2\right]^+ \\
\le\ &\left[\frac{\liminf_{N\to\infty}\phi_N(u_N) - \lim_{N\to\infty}\phi_N(v_N)}{\lim_{N\to\infty}\|u_N-v_N\|} + \frac{\lambda}{2}\lim_{N\to\infty}\|u_N-v_N\|^2\right]^+ \\
\le\ &\liminf_{N\to\infty}\left[\frac{\phi_N(u_N)-\phi_N(v_N)}{\|u_N-v_N\|} + \frac{\lambda}{2}\|u_N-v_N\|^2\right]^+ \\
\le\ &\liminf_{N\to\infty} |\partial\phi_N|(u_N)
\end{aligned}$$

Taking the supremum over $v \neq u$, we obtain (2.29). □

2.3 Convergence of an Atomistic Evolution

In Section 2.1, it was explained how different scalings of the atomistic energy E_{atom} give rise to different continuum limits. We have adopted the point of view that a linear scaling of all terms considered is the most natural choice. For the forces we assume that $f_N = O(1)$ and $f_j = O(1/N)$ for $1 \le j \le N-1$, i.e., f_N represents a surface traction. It is then natural to consider the rescaled energy

$$E_N\big((y_j^N)_{j=1}^N\big) = \sum_{j=1}^N \varepsilon_N \left[J\left(\frac{y_j^N - y_{j-1}^N}{\varepsilon_N}\right) - f_j^N (y_j^N + y_{j-1}^N)/2\right] - g y_N^N, \tag{2.30}$$

where $\varepsilon_N = 1/N$. The family $(f_i^N)_{i=1}^N$ defines a linear body force, which we assume is obtained by averaging an L^1 function, i.e.,

$$f_i^N = \frac{1}{\varepsilon_N} \int_{x_{i-1}^N}^{x_i^N} f(x)\,\mathrm{d}x,$$

where $x_i^N = i/N$, for each $i \in \mathbb{Z}$. The scalar g describes a linear surface force. For technical reasons, we may wish to impose an L^∞ bound on the deformations, i.e., we shall assume that $y_i^N \le M$, where $M \in (z_t, \infty]$.

To rewrite E_N as an integral functional it is customary to identify the atomistic deformation with a piecewise affine function. To this end, we define the set of admissible atomistic deformations to be

$$\mathscr{A}_N := \big\{ v \in \mathrm{H}^1(0,1) : v(0) = 0, v \le M, \text{ and } v \text{ is piecewise affine w.r.t. } (x_i^N) \big\}.$$

Letting

$$\begin{aligned} y_N'(x) &= \frac{y_i^N - y_{i-1}^N}{\varepsilon_N} \quad \text{if } x \in (x_{i-1}^N, x_i^N), \text{ and} \\ y_N(x) &= \int_0^x y_N'(x)\,\mathrm{d}x, \end{aligned}$$

y_N is the piecewise-affine interpolant of (y_i^N) and y_N' is its weak derivative, and we have in particular that $y_N \in \mathscr{A}_N$. Thus, we can rewrite E_N as

$$E_N(y_N) = \int_0^1 \big[J(y_N') - f_N y_N \big]\,\mathrm{d}x - g y_N(1) \qquad \text{for } y_N \in \mathscr{A}_N, \tag{2.31}$$

where f_N is the piecewise constant interpolant of f with

$$f_N(x) = f_N^i \quad \text{for } x \in (x_{i-1}, x_i). \tag{2.32}$$

In the formulation (2.31) it becomes obvious, that the non-convexity is with respect to the deformation gradient. In order to balance it out with the evolution, we need to consider the gradient flow with respect to the $|\cdot|_{\mathrm{H}^1}$-seminorm, which is in fact a norm in the spaces $\mathscr{A}_N$. We shall show below, though it is already quite obvious at this point, that the functionals E_N are uniformly λ-convex in the $|\cdot|_{\mathrm{H}^1}$-seminorm. Therefore, from Theorem 2.5, we expect the correct limit energy with respect to the $|\cdot|_{\mathrm{H}^1}$-gradient flow evolution to be

$$E(y) = \int_0^1 \big[J(y') - f y \big]\,\mathrm{d}x - g y(1), \tag{2.33}$$

defined for $y \in \mathscr{A} := \{ v \in \mathrm{H}^1(0,1) : v(0) = 0, v \le M \}$.

While it is possible to consider gradient flows with respect to the full H^1-norm as well, the analysis of equilibria becomes significantly more technical. All results can, however, be translated to the H^1-norm case [76].

Theorem 2.9 states that the (atomistic) $|\cdot|_{\mathrm{H}^1}$-gradient flow of E_N in $\mathscr{A}_N$ converges to the (continuum) $|\cdot|_{\mathrm{H}^1}$-gradient flow of E in $\mathscr{A}$. We embed $\mathscr{A}_N$ in $\mathscr{A}$ by setting $E_N(y) = +\infty$ if $y \in \mathscr{A} \setminus \mathscr{A}_N$.

Theorem 2.9 *Let $y^0 \in D(E)$, and let $y_N^0 \in \mathscr{A}_N$ be the piecewise affine interpolant of y^0 with respect to the mesh (x_i^N). Then, the $|\cdot|_{\mathrm{H}^1}$-gradient flow y_N of E_N with initial data y_N^0 converges in $\mathrm{L}^\infty_{\mathrm{loc}}([0,\infty);\mathscr{A})$ to the $|\cdot|_{\mathrm{H}^1}$-gradient flow y of E with initial data y^0.*

The convergence proof consists of three steps: first, establishing the λ-convexity of the functionals; second, estimating the perturbations caused by the discrete forcing term; and third, constructing a recovery sequence for the solution which satisfies condition (vi) of Theorem 2.5.

Lemma 2.10 *With respect to the norm $|\cdot|_{\mathrm{H}^1}$, the functionals E and E_N ($N = 1, 2, \dots$) are λ-convex in $\mathscr{A}$, with $\lambda = \min_{z>0} J''(z)$, and lower semi-continuous.*

Proof. For the λ-convexity as well as the lower semi-continuity, note that the linear, continuous terms need not be considered and we assume without loss of generality that $f, g \equiv 0$. In the spirit of Proposition 2.1, we define $F(z) = J(z) - (\lambda/2)z^2$. By the definition of λ, $F''(y) \geq 0$ whenever $y > 0$, hence F is convex in $(0,\infty)$. Since $F(z) = +\infty$ for $z \leq 0$, F is convex on $\mathbb{R}$. Therefore, the functional

$$G(y) = \int_0^1 \left(J(y') - \frac{\lambda}{2}|y'|^2 \right) \mathrm{d}x \;=\; \int_0^1 F(y')\,\mathrm{d}x$$

is convex as well which implies, by Proposition 2.1, that E is λ-convex. Since $E(y) = G(y) - \frac{\lambda}{2}|y|^2_{\mathrm{H}^1}$, a sum of a convex and a continuous functional, E is lower semicontinuous. To see that E_N is lower semi-continuous, simply note that under the assumption that $f, g \equiv 0$, $E_N = E|_{\mathscr{A}_N}$ where $\mathscr{A}_N$ is convex and closed and hence the proof carries over to E_N as well. $\square$

Lemma 2.11 *If $f \in \mathrm{L}^1(0,1)$, then, for every $v \in \mathscr{A}$, we have*

$$\left| \int_0^1 (f_N - f)v\,\mathrm{d}x \right| \leq |v|_{\mathrm{H}^1} \|f - f_N\|_{\mathrm{L}^1(0,1)}, \text{ and}$$
$$\|f - f_N\|_{\mathrm{L}^1} \to 0 \quad \text{as } N \to \infty,$$

where f_N is defined as in (2.32).

Proof. Hölder's inequality gives

$$\Big| \int_0^1 (f_N - f)v \,\mathrm{d}x \Big| \le \|v\|_{\mathrm{L}^\infty} \|f - f_N\|_{\mathrm{L}^1(0,1)}.$$

Using $v(0) = 0$, we also have $\|v\|_{\mathrm{L}^\infty} \le \|v'\|_{\mathrm{L}^1} \le |v|_{\mathrm{H}^1}$, which gives the first result. The convergence $\|f_N - f\|_{\mathrm{L}^1} \to 0$ follows from the fact that f_N is the L^2-projection of f onto the piecewise constant functions with respect to the mesh (x_i^N), using also the density of $\mathrm{L}^2(0,1)$ in $\mathrm{L}^1(0,1)$. $\square$

Lemma 2.12 *Let E and E_N be respectively given by (2.33) and (2.31), where $f \in \mathrm{L}^1(0,1)$ and f_N satisfies (2.32). For every $y \in \mathscr{A}$ with $E(y) < +\infty$, the piecewise affine, continuous interpolants v_N of y with respect to the mesh (x_i^N) satisfy*

$$|v_N - y|_{\mathrm{H}^1} \to 0, E_N(v_N) \to E(y) \quad \text{as } N \to \infty,$$
$$|v_N|_{\mathrm{H}^1} \le |y|_{\mathrm{H}^1}, \text{ and } \quad E_N(v_N) \le \big[2\|f\|_{\mathrm{L}^1}^2 + \sup_{z \ge 1} J(z)\big] + E(y) + 2|y|_{\mathrm{H}^1}^2.$$

Proof. Let $y \in \mathscr{A}$ and let v_N be the piecewise affine interpolant with respect to the mesh (x_i^N). Applying Jensen's inequality to

$$\int_{x_{i-1}^N}^{x_i^N} v_N' \,\mathrm{d}x = \int_{x_{i-1}^N}^{x_i^N} y' \,\mathrm{d}x,$$

and summing over i, we get $\|v_N'\|_{\mathrm{L}^2(0,1)} \le \|y'\|_{\mathrm{L}^2(0,1)}$. It follows from standard interpolation error estimates and a simple density argument that $|y - v_N|_{\mathrm{H}^1} \to 0$ as $N \to \infty$.

To compute the bounds on the energy as well and to show its convergence, we start with the lower-order terms. Jensen's inequality gives $\|f_N\|_{\mathrm{L}^1} \le \|f\|_{\mathrm{L}^1}$ and, as in the proof of Lemma 2.11, $\|v_N\|_{\mathrm{L}^\infty} \le |y|_{\mathrm{L}^\infty} \le |y|_{\mathrm{H}^1}$. Thus, we have

$$\begin{aligned}
-\int_0^1 f_N v_N \,\mathrm{d}x &= -\int_0^1 fy \,\mathrm{d}x + \int_0^1 \big[f(y - v_N) + (f - f_N)v_N\big] \,\mathrm{d}x \\
&\le -\int_0^1 fy \,\mathrm{d}x + \|f\|_{\mathrm{L}^1}\|y - v_N\|_{\mathrm{L}^\infty} + \|f - f_N\|_{\mathrm{L}^1}\|v_N\|_{\mathrm{L}^\infty} \\
&\le -\int_0^1 fy \,\mathrm{d}x + \|f\|_{\mathrm{L}^1}|y - v_N|_{\mathrm{H}^1} + \|f - f_N\|_{\mathrm{L}^1}|v|_{\mathrm{H}^1} \qquad (2.34) \\
&\le -\int_0^1 fy \,\mathrm{d}x + 2\|f\|_{\mathrm{L}^1}^2 + 2|y|_{\mathrm{H}^1}^2. \qquad (2.35)
\end{aligned}$$

Using Lemma 2.11 and the fact that $v_N(1) = y(1)$ for all $N \in \mathbb{N}$, we obtain from (2.34) and (2.35),

$$-\int_0^1 f_N v_N \,\mathrm{d}x - g v_N(1) \to -\int_0^1 fy \,\mathrm{d}x - gy(1) \quad \text{as } N \to \infty, \text{ and} \qquad (2.36)$$
$$-\int_0^1 f_N v_N \,\mathrm{d}x - g v_N(1) \le -\int_0^1 fy \,\mathrm{d}x - gy(1) + 2\|f\|_{\mathrm{L}^2(0,1)}^2 + 2|y|_{\mathrm{H}^1}^2.$$

To deal with the higher-order terms, let $J(z) = J_0(z) + J_1(z)$ where $J_0(z) = J^{**}(z)$. In the interval (x_{i-1}^N, x_i^N), we have $v_N' = N \int_{x_{i-1}^N}^{x_i^N} y' \, \mathrm{d}x$ and, using Jensen's inequality $J_0(v_N') \leq N \int_{x_{i-1}^N}^{x_i^N} J_0(y') \, \mathrm{d}x$ (note that $1/N$ is the length of the interval). If we define

$$a_N(x) = N \int_{x_{i-1}^N}^{x_i^N} J_0(y') \, \mathrm{d}x + \sup_{z \geq 1} J(z), \qquad \text{for } x \in (x_{i-1}^N, x_i^N),$$

then $J(v_N') \leq a_N(x)$ a.e. in $(0,1)$ and

$$\int_0^1 a_N(x) \, \mathrm{d}x = \int_0^1 J_0(y') \, \mathrm{d}x + \sup_{z \geq 1} J(z) =: A.$$

In particular, we also have

$$\int_0^1 J(v_N') \, \mathrm{d}x \leq \int_0^1 J(y') \, \mathrm{d}x + \sup_{z \geq 1} J(z),$$

which, together with (2.36) gives

$$E_N(v_N) \leq \left[2\|f\|_{\mathrm{L}^1}^2 + \sup_{z \geq 1} J(z) \right] + E(y) + 2|y|_{\mathrm{H}^1}^2. \tag{2.37}$$

Since $x \mapsto J_0(y'(x)) \in \mathrm{L}^1(0,1)$, we have, by a version of Lebesgue's differentiation theorem (Section 1.7, Corollary 2, [46])

$$\lim_{N \to \infty} a_N(x) = J_0(x) + \sup_{z \geq 1} J(z) \qquad \text{for a.e. } x \in (0,1),$$

and similarly, $v_N' \to y'$ a.e. in $(0,1)$. Using Fatou's Lemma, and the fact that J is continuous in $(0, \infty)$, we have

$$\begin{aligned}
2A - \limsup_{N \to \infty} \int_0^1 \left| J(v_N') - J(y') \right| \mathrm{d}x &= \liminf_{N \to \infty} \int_0^1 \left[2a_N - |J(v_N') - J(y')| \right] \mathrm{d}x \\
&\geq \int_0^1 \liminf_{N \to \infty} \left[2a_N - |J(v_N') - J(y')| \right] \mathrm{d}x \\
&= 2 \int_0^1 \left[J_0(y') + \sup_{z \geq 1} J(z) \right] \mathrm{d}x \\
&= 2A,
\end{aligned}$$

and hence, using also (2.36), we have $E(v_N) \to E(y)$ as $N \to \infty$ $\square$

We have now assembled all results to prove Theorem 2.9.

Proof of Theorem 2.9. The result is a straightforward application of Theorem 2.5, using the preparations of this Section.

Conditions (i) and (ii) were shown in Lemma 2.10. Condition (iii), the equi-coercivity, follows from the fact that J is bounded below and the forcing term is Lipschitz continuous. Condition (iv), the convergence of the initial data is guaranteed by standard interpolation error results as well as Lemma 2.12. Condition (v) is controlled by Lemma 2.11, since E_N and $E|_{\mathscr{A}_N}$ differ only in the forcing term.

Let $v_N(t)$ be the piecewise affine interpolant of $y(t)$. Using Lemma 2.12, to obtain (vi), we only need to show, that $t \mapsto v_N(t)$ is Borel measurable. In fact, it is fairly easy to see that it is even continuous. Since in one dimension, $\mathrm{H}^1(0,1)$ is embedded in $\mathrm{C}[0,1]$, the mapping $t \mapsto y(t)$ lies in $\mathrm{C}(0,\infty;\mathrm{C}[0,1])$ and hence $t \mapsto y(t,x)$ is continuous as well. Since

$$v_N(t,x) = \sum_{j=1}^{N} y(t, x_j^N)\varphi_j^N(x),$$

where the φ_j^N are Lipschitz functions, this shows that $v \in \mathrm{C}(0,\infty;\mathrm{H}^1)$. $\square$

2.4 Convergence of Equilibria

In this section we show that the gradient flows are a selection criterion which can be used to recover correct elastic behaviour even when the energy has sublinear growth.

The convergence result of Theorem 2.9 suggests the following procedure: for sufficiently small forces, there should be a critical point y_N^*, in fact a strict local minimum, of the atomistic functional E_N, such that ${y_N^*}' < z_t$, i.e., the deformation gradient lies in the region where J is convex. Hence, the gradient flow for sufficiently close starting points should converge to y_N^* as $t \to \infty$ and the deformation gradient should remain within the region where J is convex. Since the atomistic gradient flow converges to the continuum gradient flow, the continuum deformation gradient should remain in this region as well and therefore converge to a critical point in that set which should be the limit of the y_N^*. By y^* being a critical point of ϕ, we mean that $|\partial\phi|(y^*) = 0$, where $|\partial\phi|(y)$ is the $|\cdot|_{\mathrm{H}^1}$-slope of ϕ at y (compare Section 2.2.3).

The main difficulty is to show that the critical points y_N^* are 'uniform local minimizers' in the sense that we do not require perturbations to tend to zero as $N \to \infty$.

Before we start with the suggested program, let us note that it would be quite easy to show all results for the continuum problem directly. However, we wish to show here that the elastic critical point of the continuum functional (2.33) arises as the limit of the elastic critical points of the atomistic functionals (2.31). Furthermore, it is an interesting feature of the analysis that all information about the continuum functional can be obtained from the knowledge about the atomistic evolution.

Theorem 2.13 *Let E_N, $N = 1, 2, \ldots$, and E be defined respectively by* (2.31) *and* (2.33)*, and assume that $|g| + \|f\|_{\mathrm{L}^1(0,1)} < J'(z_t)$ (cf.* (2.2)*).*

(a) There exist critical points y_N^* of E_N in $\mathscr{A}_N$, such that ${y_N^*}' < z_t$. These equilibria are stable in the sense that any $|\cdot|_{\mathrm{H}^1}$-gradient flow y_N of E_N with $y_N'(0,x) < z_t$ satisfies $\lim_{t\to\infty} y_N(t) = y_N^*$ in $\mathrm{H}^1(0,1)$.

(b) There exists a critical point $y^* \in \mathscr{A}$ of E such that $\lim_{N\to\infty} y_N^* = y^*$ and $\lim_{t\to\infty} y(t) = y^*$ in H^1, for every $|\cdot|_{\mathrm{H}^1}$- gradient flow y of E with $y'(0,x) \le z_t - \epsilon$ for some $\epsilon > 0$.

(c) If, in addition, $f \equiv 0$, then $y_N^* = y^*$ are affine.

On the one hand, Theorem 2.13 shows that the derived continuum model has the correct qualitative and quantitative behaviour for small loads. On the other hand, it shows that in this situation, the atomistic model behaves essentially like a continuum. In particular, note that point (c) is the Cauchy–Born hypothesis for the model presented.

Note also, that not all proofs in this section are 'optimal'. Especially the final proof of Theorem 2.13 is more technical than it needs to be. The purpose of this discussion is to show that most of the techniques used here can be applied to more general problems.

The proof of Theorem 2.13 requires some preparation in the form of several Lemmas which assemble information about the atomistic gradient flow. Let $\mathscr{B}$ be the set of all deformations whose gradient remains in the region where J is convex, i.e., we define

$$\mathscr{B}_\epsilon = \{v \in \mathscr{A} : v'(x) \le z_t - \epsilon \text{ for a.e. } x \in (0,1)\}, \tag{2.38}$$

and $\mathscr{B} = \mathscr{B}_0$.

Lemma 2.14 *Suppose that $|g| + \|f\|_{\mathrm{L}^1(0,1)} \le J'(z_t - \epsilon)$ for some $\epsilon > 0$; then there exists a unique critical point y_N^* of E_N in the set $\mathscr{B}_\epsilon$. The point y_N^* satisfies*

$${y_N^*}'(x) = (J')^{-1}(F_j^N) \le z_t - \epsilon \quad \text{for } x_{j-1}^N < x < x_j^N, \tag{2.39}$$

where F_j^N is defined by (2.40)*.*

Proof. We compute the critical point by a change of variables. For $y_N \in \mathscr{A}_N$, let $r_j^N = (y_j^N - y_{j-1}^N)/\varepsilon_N$. Then, setting

$$\tilde{f}_i^N = \begin{cases} \frac{1}{2} f_1^N, & \text{if } i = 0, \\ \frac{1}{2}(f_i^N + f_{i+1}^N), & \text{if } 1 \le i \le N-1 \\ \frac{1}{2} f_N^N, & \text{if } i = N, \end{cases}$$

we have, using $y_0^N = 0$,

$$\begin{aligned} E_N(y_N) &= \sum_{j=1}^{N} \varepsilon_N J(r_j^N)) - \sum_{j=0}^{N} \varepsilon_N \tilde{f}_j^N y_j^N - g y_N^N \\ &= \sum_{j=1}^{N} \varepsilon_N J(r_j^N)) - \sum_{j=1}^{N} \varepsilon_N \tilde{f}_j^N \sum_{i=1}^{j} \varepsilon_N r_i^N - g \sum_{i=1}^{N} \varepsilon_N r_i^N \\ &= \sum_{j=1}^{N} \varepsilon_N J(r_j^N)) - \sum_{i=1}^{N} \varepsilon_N r_i^N \Big[g + \sum_{j=i}^{N} \varepsilon_N \tilde{f}_j^N \Big] \\ &= \sum_{j=1}^{N} \varepsilon_N \big[J(r_j^N) - F_j^N r_j^N \big], \end{aligned}$$

where

$$F_i^N \;=\; g + \sum_{j=i}^{N} \varepsilon_N \tilde{f}_j^N \;=\; g + \frac{\varepsilon_N}{2}(f_i^N + f_N^N) + \sum_{j=i+1}^{N-1} \varepsilon_N f_j^N. \tag{2.40}$$

To compute r_j^N, we differentiate E_N with respect to r_j^N, which gives the equation

$$\frac{\partial E_N(y_N)}{\partial r_j^N} = \varepsilon_N \big[J'(r_j^N) - F_j^N \big] = 0 \quad \text{for } j = 1, \dots, N,$$

or, equivalently, $J'(r_j^N) = F_j^N$. We estimate F_j^N, using the assumption that $\|f\|_{\mathrm{L}^1} + |g| \le J'(z_t - \epsilon)$, by

$$\begin{aligned} |F_j^N| &= \left| g + \frac{1}{2} \int_{x_{j-1}^N}^{x_j^N} f(x)\,\mathrm{d}x + \int_{x_j^N}^{x_{N-1}^N} f(x)\,\mathrm{d}x + \frac{1}{2} \int_{x_{N-1}^N}^{1} f(x)\,\mathrm{d}x \right| \\ &\le |g| + \int_{x_{j-1}^N}^{1} |f(x)|\,\mathrm{d}x \\ &\le |g| + \|f\|_{\mathrm{L}^1(0,1)} \\ &\le J'(z_t - \epsilon). \end{aligned} \tag{2.41}$$

In the region $\{z < z_t\}$, $J'(z)$ is strictly increasing and hence invertible. Therefore,

$$r_j^N = (J')^{-1}(F_j^N) \le z_t - \epsilon$$

describes the unique critical point of E_N in $\mathscr{B}_\epsilon$. $\square$

Lemma 2.15 *Under the conditions of Lemma 2.14, if $y_N \colon [0,\infty) \to \mathscr{A}_N$ is an $|\cdot|_{\mathrm{H}^1}$-gradient flow of E_N with $y_N(0) \in \mathscr{B}_\epsilon$ then $y_N(t) \in \mathscr{B}_\epsilon$ for all $t > 0$.*

Proof. Consider the time-discrete approximation $(U_N(t_j))_{j=0,1,\dots}$, as described in Lemma 2.7, for some fixed, sufficiently small time-step τ. Let $R_N^i(t_j)$ be as in the proof of Lemma 2.14. Then, $R_N(t_j)$ minimizes

$$\frac{1}{2\tau}\|R_N(t_j) - R_N(t_{j-1})\|_{L^2}^2 + E_N(R_N(t_j)). \tag{2.42}$$

As in the proof of Lemma 2.14, we compute the Euler–Lagrange equation in terms of $R_N^i(t_j)$. At the minimum, the equation

$$\frac{1}{\tau}\Big(R_N^i(t_j) - R_N^i(t_{j-1})\Big) = F_j^N - J'(R_N^i(t_j))$$

has to be satisfied. For sufficiently small τ, there is a unique solution. Now assume inductively that $R_N^i(t_{j-1}) \le z_t - \epsilon$. To show that $R_N^i(t_j) \le z_t - \epsilon$, assume this is not true. Then $F_j^N - J'(R_N^i(t_j)) < 0$, which gives a contradiction. Hence, we have that for all $i = 1, \dots, N$ and $j \in \mathbb{N}$, $R_N^i(t_j) \le z_t - \epsilon$. As $\tau \to 0$, the discrete solution converges to the gradient flow y_N and hence $y_N' \le z_t - \epsilon$ a.e. in $(0,1)$. □

Corollary 2.16 *Under the conditions of Lemma 2.14, every $|\cdot|_{\mathrm{H}^1}$-gradient flow y_N with $y_N(0) \in \mathscr{B}_\epsilon$ satisfies the evolutionary variational inequality*

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|y_N - v|_{\mathrm{H}^1}^2 + \frac{\alpha}{2}|y_N - v|_{\mathrm{H}^1}^2 + E_N(y_N) \le E_N(v) \quad \forall v \in \mathscr{B}_\epsilon, \tag{2.43}$$

where $\alpha = \min_{z \le z_t - \epsilon} J''(z) > 0$. In particular, we have

$$|y_N(t) - y_N^*|_{\mathrm{H}^1} \le \mathrm{e}^{-\alpha t}|y_N(0) - y_N^*|_{\mathrm{H}^1}.$$

Proof. We set $\tilde{E}_N = E_N|_{\mathscr{B}_\epsilon}$ and show that y_N is also a gradient flow for $\tilde{E}_N$ by considering the minimization problem (2.42) again. (Note that this procedure is equivalent to replacing E_N outside of $\mathscr{B}_\epsilon$ by a uniformly convex functional.) Since the minimizer remains in $\mathscr{B}_\epsilon$, it is also the minimizer of

$$\frac{1}{2\tau}\|R_N(t_j) - R_N(t_{j-1})\|_{\mathrm{L}^2}^2 + \tilde{E}_N(R_N(t_j)),$$

and hence the limit of the time-discretizations must also be the gradient flow of $\tilde{E}_N$. By arguing as in the proof of Lemma 2.10, we find that $\tilde{E}_N$ is α-convex (i.e. λ-convex with $\lambda = \alpha$), and hence y_N satisfies (2.43) if we replace E_N with $\tilde{E}_N$. For $v \in \mathscr{B}_\epsilon$, however, the functionals are the same.

On testing (2.43) with $v = y_N^*$, and multiplying the resulting inequality by $\mathrm{e}^{2\alpha t}$, we obtain

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathrm{e}^{\alpha t}|y_N(t) - y_N^*|_{\mathrm{H}^1}\Big)^2 \le \mathrm{e}^{\alpha t}\big(E_N(y_N^*) - E_N(y_N(t))\big) \le 0.$$

Integrating from 0 to T gives the result. $\square$

Proof of Theorem 2.13. Lemmas 2.14, 2.15, and 2.16 immediately imply item (a) and we only need to establish the facts about the continuum limit. Note that almost all of the following analysis is independent of the specific structure of the problem. The only crucial condition which we require, is that $y_N(t) \to y(t)$ as $N \to \infty$, for every $t \geq 0$, and $y_N(t) \to y_N^*$ as $t \to \infty$, uniformly in N.

For item (b), we first need to show that, given an initial condition $y(0)$ for the 'continuum' $|\cdot|_{\mathrm{H}^1}$-gradient flow satisfying the assumptions of the theorem, there exist 'atomistic' initial conditions $y_N(0)$ which satisfy the assumptions of Lemma 2.15. Let $y'(0,x) \leq z_t - \epsilon$ for a.e. $x \in (0,1)$. Letting $y_N(0,x)$ be the piecewise affine interpolant of $y(0,x)$, we have

$$y_N'(0,x) = \frac{1}{\varepsilon_N} \int_{x_{i-1}^N}^{x_i^N} y'(0,x)\,\mathrm{d}x \leq z_t - \epsilon, \quad x \in (x_{i-1}^N, x_i^N).$$

Therefore, the atomistic $|\cdot|_{\mathrm{H}^1}$-gradient flows with starting point $y_N'(0,\cdot)$ converge uniformly in N (compare Corollary 2.16) to the equilibria y_N^*, computed in item (a) or Lemma 2.14. We use this fact to estimate

$$\begin{aligned} |y_N^* - y_{N'}^*|_{\mathrm{H}^1} &\leq |y_N^* - y_N(t)|_{\mathrm{H}^1} + |y_N(t) - y_{N'}(t)|_{\mathrm{H}^1} + |y_{N'}(t) - y_{N'}^*|_{\mathrm{H}^1} \\ &\leq 2\mathrm{const.e}^{-\alpha t} + |y_N(t) - y_{N'}(t)|_{\mathrm{H}^1}, \end{aligned}$$

thus showing that $(y_N^*)_{N=1,2,\ldots}$ is a Cauchy-sequence. We denote its limit in H^1 by y^*. To see that $y(t) \to y^*$ as $t \to \infty$, consider

$$|y(t) - y^*|_{\mathrm{H}^1} \leq \inf_{N=1,2,\ldots} (|y(t) - y_N(t)|_{\mathrm{H}^1} + |y_N(t) - y_N^*|_{\mathrm{H}^1} + |y_N^* - y^*|_{\mathrm{H}^1}) \leq \mathrm{const.e}^{-\alpha t}.$$

We have shown that the 'discrete' equilibria y_N^* converge to a 'continuum' deformation y^* and that $y(t) \to y^*$.

The fact that y^* is a critical point of E is easily verified by hand, but in fact this follows from the general theory as well, using the concepts introduced in Section 2.2.3. It is straightforward to show that the functionals E_N $\Gamma(\mathrm{H}^1)$-converge to E in the strong H^1 topology. We merely note the limsup condition (2.28) is given by Lemma 2.12 while for the liminf condition (2.27) E and E_N can be decomposed into a convex, lower semicontinous part and a continuous, uniformly convergent part (compare also the prove of λ-convexity in Lemma 2.10).

Since the functionals E and $(E_N)_{N=1,2,\ldots}$ are also uniformly λ-convex, Lemma 2.8, shows that

$$|\partial E|(y^*) \leq \liminf_{N\to\infty} |\partial E_N|(y_N^*) = 0,$$

where $|\partial E_{(N)}|$ denotes the $|\cdot|_{\mathrm{H}^1}$-local slope of the functionals $E_{(N)}$. $\square$

Remark. It may not come as a surprise that the continuum 'elastic' critical point computed in Theorem 2.13 is actually not a local minimizer with respect to the H^1-topology. Indeed, let us assume that $f \equiv 0$ and $0 < g < J'(z_t)$ and define the curve $s \mapsto v(s)$ by

$$v'(s) = y^{*\prime} + \frac{1}{s}\chi_{(1/2,1/2+s^k)}.$$

It is straightforward to establish that for $k \geq 2p$, $v \in \mathrm{C}^{0,1/p}(0, s_0; \mathrm{W}^{1,p})$ and $E(v(s)) < E(y^*)$, where $s_0 > 0$ and $\mathrm{C}^{0,1/p}$ denotes the usual space of Hölder continuous functions. Thus, the critical point y^* is not an H^1-local minimum of the energy $E(y)$. This is also reflected by the fact that we only allow $\mathrm{W}^{1,\infty}$ perturbations in Theorem 2.13.

Why, we should ask ourselves, is this not in contradiction with Theorem 2.13? If there exists a curve along which the energy decreases, should the gradient flow not find this curve? The explanation is that the curve $v(s)$ which we have constructed is not absolutely continuous in $\mathrm{H}^1(0,1)$ and hence is not a candidate for the gradient flow evolution. An interesting question is whether there actually can exist an absolutely continuous curve starting in y^* along which the energy decreases strictly. A negative answer would lead to an interesting selection criterion for equilibria. It would in particular imply that the choice of evolution is not so crucial after all, as such equilibria would be stable under any 'sufficiently smooth' evolution. ◀

2.5 Remarks on Extensions to 2D and 3D

The simple problem which we have investigated in this chapter has a fair amount of one-dimensional structure. Although many of the techniques developed here can be readily generalized, the extension to two and three dimensions, which is of great importance to the modeling of material behaviour, is not entirely trivial.

The first difficulty to notice is that the passage to higher dimensions in a simple nearest-neighbour system based on the Lennard–Jones potential suffers from a loss of λ-convexity, since the atomistic deformations do not necessarily have to remain orientation preserving. By cutting off the Lennard–Jones potential at the origin, a process which is intuitively reasonable but difficult to justify rigorously, the convergence of the gradient flow can be recovered completely. A more interesting, and mathematically much more challenging alternative would be to consider a gradient flow with respect to a different metric, which may allow the blow-up behaviour of the Lennard–Jones potential, but such a metric seems to be presently unavailable.

To analyse elastic equilibria, it is necessary to obtain L^∞ bounds on the deformation gradient. This step poses the biggest challenge in higher dimensions as these bounds cannot be computed explicitly anymore. One possible avenue to obtain them would be to use the implicit function theorem, for which uniform bounds can be constructed with a slightly refined analysis. It would be necessary, however, that the solution of the linearized system lies in $W^{1,\infty}(\Omega)$, which can only be obtained in some very restrictive cases, e.g., with smooth domains and Dirichlet boundary conditions. At re-entrant corners (such as a crack tip) or interfaces between Dirichlet and Neumann boundaries, the nearest neighbour model is too simple to describe the material behaviour accurately.

If finite-range interactions are added to the energy functional, both the convergence theory for gradient flows and the analysis of elastic equilibria remains essentially unchanged, provided that uniform L^∞ bounds on the gradients can be provided. The case of infinite-range interactions is completely unclear.

Finally, it should be noted that different evolutions can be analyzed as well. For example, if the potential J is cut off at the origin, it is straightforward to extend the convergence result from the gradient flow evolution to linear viscoelasticity following, for example, the theory developed in [26]. It is more difficult in this setting, however, to analyze the resulting stationary points in similar detail.

Chapter 3

A Posteriori Existence in Numerical Computations

In Chapter 2, gradient flows were introduced as a possible concept for the analysis of local minima of non-convex functionals. In the present chapter, an alternative which is far better suited for applications in numerical analysis is presented. The technique extends and strengthens existing *a posteriori* error analyses and makes it possible to derive the existence of exact solutions from the computation, even when it is not known *a priori* whether a solution exists. Since the methodology to obtain such results is quite general and is widely applicable, an abstract analysis is given in this chapter, together with two interesting applications from continuum numerical analysis.

The basic idea is simple. Suppose we are solving the nonlinear equation $\mathscr{F}(u) = 0$, where $\mathscr{F}\colon \mathscr{X} \to \mathscr{Y}^*$, where $\mathscr{X}, \mathscr{Y}$ are Banach spaces and $\mathscr{Y}^*$ is the topological dual of $\mathscr{Y}$. A point u of the domain of definition of $\mathscr{F}$ is called regular if $\mathscr{F}'(u)^{-1}$ exists and is bounded (cf. [95, Proposition 2.1]; though here we use a slightly more general definition). *A posteriori* error estimates for nonlinear problems are typically formulated in the following way (cf. [94, Proposition 2.1], [95, Proposition 2.1] or [3, Lemma 9.5]): *If u is a regular solution to $\mathscr{F}(u) = 0$ and U is a numerical solution which is sufficiently close to u then* $\|u - U\|_{\mathscr{X}} \lesssim \|\mathscr{F}(U)\|_{\mathscr{Y}^*} \|\mathscr{F}'(u)^{-1}\|_{L(\mathscr{Y}^*, \mathscr{X})}$.

In this chapter, we make use of the following simple observation which seems to have gone unnoticed so far: the approximation U, which we may have computed numerically, solves the equation $\mathscr{G}(U) = 0$, where $\mathscr{G}(v) = \mathscr{F}(v) - \mathscr{F}(U)$. Thus, by reversing the role of approximate and exact solution, we see that a solution u, satisfying $\mathscr{F}(u) = 0$ could be considered the approximate solution to the new problem $\mathscr{G}(U) = 0$ and its residual is again $\|\mathscr{F}(U)\|_{\mathscr{Y}^*}$. In this new situation we only need to assume that U is regular rather than a *nearby exact solution* u which we usually do not know to exist *a priori*. More precisely, we shall prove that, *if a regular point $U \in \mathscr{X}$ has a sufficiently small*

residual $\|\mathscr{F}(U)\|_{\mathscr{Y}^*}$ *then there exists a nearby solution* u *to the equation* $\mathscr{F}(u) = 0$ *such that* $\|u - U\|_{\mathscr{X}} \lesssim \|\mathscr{F}(U)\|_{\mathscr{Y}^*}\|\mathscr{F}'(U)^{-1}\|_{L(\mathscr{Y}^*,\mathscr{X})}$.

In an abstract Banach space setting, we give two general results. First, Theorem 3.3 provides an asymptotically optimal strategy based on Lipschitz continuity of $\mathscr{F}'$. Second, Theorem 3.5 is a result based on a continuation argument, which is intended to outline a form for more specific strategies that may be preferable if sufficient analytical information about the problem is available. Both results are obtained by correctly interpreting the Inverse Function Theorem in Banach spaces and carefully tracking the constants in a suitably chosen versions of its proof.

In §3.3 we use the semilinear Laplace equation

$$-\Delta u + f(x,u) = 0, \qquad u|_{\partial\Omega} = 0,$$

to demonstrate further details of the abstract idea, based on Theorem 3.3, which are required for its practical implementation. In particular, it is shown how the required stability constants can be computed in a Hilbert space setting.

In §3.2 we look at a problem where the existence of solutions is not as clear. We compute local minimizers of a double-well energy arising in the mathematical theory of microstructures. It is shown how the *a posteriori* existence idea can be used to demonstrate the existence of a rich class of stable solutions to the Euler–Lagrange equations.

Of course, the main application is given in Chapter 5. The large number of metastable states of atomistic functionals makes it a compelling example for the *a posteriori* existence idea.

For evolution equations where the uniqueness of the solution is usually guaranteed, a similar procedure can give improved *a posteriori* error bounds based on local stability properties rather than global ones. We refer to [13] where this idea is used for Ginzburg–Landau type equations.

The reverse question of existence of a numerical solution near an exact solution has been extensively studied; see for example [23, 40].

3.1 Abstract Results

Let $\mathscr{X}$ be a Banach space, and let $\mathscr{Y}$ be a Banach space with topological dual $\mathscr{Y}^*$. We denote the duality pairing between $\mathscr{Y}$ and $\mathscr{Y}^*$ by $\langle\cdot,\cdot\rangle$. For $v \in \mathscr{X}$ and $R > 0$, we use $B(v,R)$ to denote the closed ball with centre v and radius R in $\mathscr{X}$. Let $\mathscr{A}$ be an

open subset of $\mathscr{X}$ and let $\mathscr{F}\colon \mathscr{A} \to \mathscr{Y}^*$. For example, the operator $y \mapsto -\mathrm{div}\sigma(\nabla y)$ with Dirichlet boundary conditions could be understood in the sense

$$\langle \mathscr{F}(y), \varphi \rangle = \int_\Omega \sigma(\nabla y) : \nabla \varphi \, \mathrm{d}x \qquad \forall \varphi \in \mathscr{Y},$$

where $\mathscr{X}$ and $\mathscr{Y}$ are appropriately chosen function spaces (cf. §3.2 and §3.3 for further detail).

We say that $\mathscr{F}$ is differentiable at a point $u \in \mathscr{A}$ if it is Gateaux differentiable at that point, i.e., if there exists a bounded linear operator $\mathscr{F}'(u) \in L(\mathscr{X}, \mathscr{Y}^*)$ such that, for $v \in \mathscr{X}$, we have

$$\lim_{h \to 0} \left\| h^{-1}\big(\mathscr{F}(u+hv) - \mathscr{F}(u)\big) - \mathscr{F}'(u)v \right\|_{\mathscr{Y}^*} = 0.$$

To avoid a cluttered notation, we shall always use $\|T\|$ to denote the operator norm of a bounded linear operator T between Banach spaces. It will always be clear from the context which spaces are meant.

We begin by stating a simple Lemma from functional analysis which will motivate our definition of regular points.

Lemma 3.1 *Let $T\colon \mathscr{X} \to \mathscr{Y}^*$ be a bounded linear operator; then (i) and (ii) are equivalent:*

(i) The range of T, denoted $\mathrm{range}(T)$*, is closed and $T\colon \mathscr{X} \to \mathrm{range}(T)$ is one-to-one.*

(ii) T is bounded and

$$\alpha := \inf_{\substack{u \in \mathscr{X} \\ \|u\|_{\mathscr{X}} = 1}} \sup_{\substack{\varphi \in \mathscr{Y} \\ \|\varphi\|_{\mathscr{Y}} = 1}} \langle Tu, \varphi \rangle > 0. \tag{3.1}$$

If (i) or (ii) are satisfied then $T^{-1} : \mathrm{range}(T) \to \mathscr{X}$ is bounded and $\|T^{-1}\| = 1/\alpha$.

Proof. To show that (ii) implies that $\mathrm{range}(T)$ is closed, let $Tu_j \to v$ in $\mathscr{Y}^*$. From (3.1) it follows that

$$\alpha \|u_j - u_k\|_{\mathscr{X}} \le \sup_{\substack{\varphi \in \mathscr{Y} \\ \|\varphi\| = 1}} \langle T(u_j - u_k), \varphi \rangle = \|T(u_j - u_k)\|. \tag{3.2}$$

Since (Tu_j) is Cauchy, so is (u_j); thus, there exists $u \in \mathscr{X}$ such that $u_j \to u$. Since T is bounded, it follows that $Tu_j \to Tu = v$ and hence $\mathrm{range}(T)$ is closed. The fact that T is one-to-one follows directly from (3.2).

If (i) is satisfied, the Open Mapping Theorem implies that T as well as T^{-1} : range$(T) \to \mathscr{X}$ are bounded. To show that (i) also implies that $\alpha > 0$, note that (3.1) can be written equivalently as

$$\begin{aligned}
\alpha &= \inf_{u\in\mathscr{X}\setminus\{0\}} \sup_{\varphi\in\mathscr{Y}\setminus\{0\}} \frac{\langle Tu,\varphi\rangle}{\|u\|_{\mathscr{X}}\|\varphi\|_{\mathscr{Y}}} \\
&= \inf_{v\in\mathrm{range}(T)\setminus\{0\}} \sup_{\varphi\in\mathscr{Y}\setminus\{0\}} \frac{\langle v,\varphi\rangle}{\|T^{-1}v\|_{\mathscr{X}}\|\varphi\|_{\mathscr{Y}}} \\
&= \inf_{v\in\mathrm{range}(T)\setminus\{0\}} \frac{\|v\|_{\mathscr{Y}^*}}{\|T^{-1}v\|_{\mathscr{X}}} \\
&= \|T^{-1}\|^{-1},
\end{aligned}$$

where $\|T^{-1}\|$ denotes the operator norm in $L(\mathrm{range}(T),\mathscr{X})$. This concludes the proof of equivalence of (i) and (ii) and also shows the last statement of the lemma. □

Motivated by Lemma 3.1, if $u \in \mathscr{A}$ and $\mathscr{F}$ is differentiable at u, we define

$$\alpha(u) = \inf_{\substack{v\in\mathscr{X}\\ \|v\|_{\mathscr{X}}=1}} \sup_{\substack{\varphi\in\mathscr{Y}\\ \|\varphi\|_{\mathscr{Y}}=1}} \langle \mathscr{F}'(u)v,\varphi\rangle, \qquad u\in\mathscr{A}.$$

If (i) or (ii) holds in Lemma 3.1 with $T = \mathscr{F}'(u)$ then $\alpha(u) = \|\mathscr{F}'(u)^{-1}\|^{-1}$, where $\|\mathscr{F}'(u)^{-1}\|$ denotes the operator norm in $L(\mathrm{range}(\mathscr{F}'(u)),\mathscr{X})$. Otherwise, $\alpha(u) = 0$. For the sake of a more attractive notation, we shall mostly use $\alpha(u)$ rather than $\mathscr{F}'(u)^{-1}$.

Definition 3.2 *We say that a point $u \in \mathscr{A}$ is regular if $\mathscr{F}$ is differentiable at u and $\alpha(u) > 0$.*

Both abstract *a posteriori* existence theorems are essentially reformulations of the Inverse Function (or Implicit Function) Theorem. The crucial step is to interpret them correctly and to track all constants in the proofs. The two results, Theorem 3.3 and Theorem 3.5 are fully equivalent. While Theorem 3.3 follows immediately from Theorem 3.5, the proof of the latter makes heavy use of the Inverse Function Theorem 3.4 which is a corollary of the former. Extensive comments on both results can be found below.

Theorem 3.3 *Suppose that $\mathscr{F}$ is differentiable in $\mathscr{A}$ and that $U \in \mathscr{A}$ is regular. For $\rho \in [1/2, 1]$ set $R = R(\rho) = \rho^{-1}\alpha(U)^{-1}\|\mathscr{F}(U)\|_{\mathscr{Y}^*}$ and let $L = L(\rho)$ be the Lipschitz constant of $\mathscr{F}'$ in $B(U,R)\cap\mathscr{A}$. If $B(U,R)\subset\mathscr{A}$, if*

$$\mathscr{F}(U) \in \mathrm{range}(\mathscr{F}'(\theta)) \qquad \forall\theta\in B(U,R), \quad \text{and if} \tag{3.3}$$

$$\alpha(U)^{-2}\|\mathscr{F}(U)\|_{\mathscr{Y}^*} \;\le\; \rho(1-\rho)/L \tag{3.4}$$

then there exists a unique $u \in B(U,R)$ such that $\mathscr{F}(u) = 0$. Furthermore, we have the error estimate

$$\|u - U\|_{\mathscr{X}} \leq \rho^{-1}\alpha(U)^{-1}\|\mathscr{F}(U)\|_{\mathscr{Y}^*}. \tag{3.5}$$

Proof. First, we determine a radius R and a stability constant for the set $B(U,R)$. To this end, we construct an optimal fraction $\rho \in (0,1]$, letting R depend on ρ, such that

$$\inf_{v \in B(U,R)} \alpha(v) \geq \rho\alpha(U).$$

In this case, the best possible error estimate that we can achieve would be $\|u-U\|_{\mathscr{X}} \leq \rho^{-1}\alpha(U)^{-1}\|\mathscr{F}(U)\|_{\mathscr{Y}^*}$. Hence, the smallest possible radius such that u is contained in $B(U,R)$ is given by $R = R(\rho) = \rho^{-1}\alpha(U)^{-1}\|\mathscr{F}(U)\|_{\mathscr{Y}^*}$. For any larger radius, we obtain an unnecessarily large Lipschitz constant, while for a smaller radius we could not possibly conclude that $u \in B(U,R)$.

Let $L = L(\rho)$ be the Lipschitz constant of $\mathscr{F}'$ in $B(U,R)$. For $\theta \in B(U,R)$, we can estimate

$$\begin{aligned}
\alpha(\theta) &= \inf_{\substack{v \in \mathscr{X} \\ \|v\|_{\mathscr{X}}=1}} \sup_{\substack{\varphi \in \mathscr{Y} \\ \|\varphi\|_{\mathscr{Y}}=1}} \langle \mathscr{F}'(\theta)v, \varphi \rangle \\
&\geq \inf_{\substack{v \in \mathscr{X} \\ \|v\|_{\mathscr{X}}=1}} \sup_{\substack{\varphi \in \mathscr{Y} \\ \|\varphi\|_{\mathscr{Y}}=1}} \langle \mathscr{F}'(U)v, \varphi \rangle - LR \\
&= \alpha(U) - LR.
\end{aligned}$$

Thus, in order to have $\alpha(\theta) \geq \rho\alpha(U)$, we require $LR \leq (1-\rho)\alpha(U)$, which is equivalent to (3.4). We see, in particular, that if (3.4) is satisfied then every point $v \in B(U,R)$ is regular and satisfies $\alpha(v) \geq \rho\alpha(U)$.

While L is, to some extent, dependent on the choice of ρ it is reasonable to assume that it remains roughly constant, particularly when the residual (and hence R) is small. Hence, the value $\rho = 1/2$ is roughly optimal in that (3.4) is most easily satisfied in this case. For smaller ρ the resulting error estimate deteriorates and (3.4) becomes more difficult to satisfy as well — hence the assumption that $\rho \geq 1/2$.

The remainder of the proof involves a careful tracking of the constants in the proof of the Inverse Function Theorem. The fixed point iteration used here is an adaption of the argument used in [63] and [78].

We wish to prove the existence of a solution to $\mathscr{F}(u) = 0$ in $B(U,R)$. Since $\mathscr{F}'$ is continuous in $B(U,R)$, we can use Taylor's Theorem to obtain, for $v \in B(U,R)$,

$$\mathscr{F}(v) - \mathscr{F}(U) = \int_0^1 \mathscr{F}'(U + \tau(v-U))\,\mathrm{d}\tau \cdot (v-U) =: \mathscr{F}'_{U,v}(v-U). \tag{3.6}$$

Assume, for the moment, that $u \in B(U,R)$ satisfies $\mathscr{F}(u) = 0$ so that $\mathscr{F}(u) - \mathscr{F}(U) = -\mathscr{F}(U)$. Applying (3.6) to the right-hand side, we find that $\mathscr{F}(u) = 0$ is equivalent to

$$\mathscr{F}'_{U,u}(u - U) = -\mathscr{F}(U).$$

We now define the map $\mathcal{L}\colon B(U,R) \to \mathscr{X}$ by

$$\mathscr{F}'_{U,v}(\mathcal{L}(v) - U) = -\mathscr{F}(U). \tag{3.7}$$

To show that the map is well-defined, we first use the Integral Mean Value Theorem to infer the existence of $\theta_v \in \operatorname{conv}\{U, v\}$ such that $\mathscr{F}'_{U,v} = \mathscr{F}'(\theta_v)$. Since $\theta_v \in B(U,R)$, it follows that $\mathscr{F}'_{U,v}$ is an isomorphism from $\mathscr{X}$ to $\operatorname{range}(\mathscr{F}'(\theta_v))$ which, by (3.3) contains $\mathscr{F}(U)$, and in particular that (3.7) has a unique solution. In summary, an element $u \in B(U,R)$ satisfies $\mathscr{F}(u) = 0$ if, and only if, u is a fixed point of $\mathcal{L}$.

To show that $\mathcal{L}$ maps $B(U,R)$ into itself, we multiply (3.7) by $(\mathscr{F}'_{U,v})^{-1}$ to infer

$$\|\mathcal{L}(v) - U\|_{\mathscr{X}} \le \|(\mathscr{F}'_{U,v})^{-1}\| \|\mathscr{F}(U)\|_{\mathscr{Y}^*} \le \rho^{-1}\alpha(U)^{-1}\|\mathscr{F}(U)\|_{\mathscr{Y}^*} = R.$$

To show that $\mathcal{L}$ is a contraction of $B(U,R)$ let $v_1, v_2 \in B(U,R)$; then

$$\mathscr{F}'_{U,v_i}(\mathcal{L}(v_i) - U) = -\mathscr{F}(U), \qquad i = 1, 2.$$

Subtracting these two equations, we obtain

$$\begin{aligned} \mathscr{F}'_{U,v_1}(\mathcal{L}(v_1) - \mathcal{L}(v_2)) &= -\mathscr{F}'_{U,v_1}(\mathcal{L}(v_2) - U) + \mathscr{F}'_{U,v_2}(\mathcal{L}(v_2) - U) \\ &= -(\mathscr{F}'_{U,v_1} - \mathscr{F}'_{U,v_2})(\mathcal{L}(v_2) - U). \end{aligned}$$

Multiplying by $(\mathscr{F}'_{U,v_1})^{-1}$ and using

$$\begin{aligned} &\Big\| \int_0^1 \big[\mathscr{F}'(U + \tau(v_1 - U)) - \mathscr{F}'(U + \tau(v_2 - U))\big]\, \mathrm{d}\tau \Big\| \\ &\qquad \le \int_0^1 \big\|\mathscr{F}'(U + \tau(v_1 - U)) - \mathscr{F}'(U + \tau(v_2 - U))\big\|\, \mathrm{d}\tau \\ &\qquad \le \int_0^1 \tau L \|v_1 - v_2\|_{\mathscr{X}}\, \mathrm{d}\tau \;=\; \tfrac{1}{2} L \|v_1 - v_2\|_{\mathscr{X}} \end{aligned}$$

and (3.4) we obtain

$$\begin{aligned} \|\mathcal{L}(v_1) - \mathcal{L}(v_2)\|_{\mathscr{X}} &\le \tfrac{1}{2}\, L\, \|v_1 - v_2\|_{\mathscr{X}} \|(\mathscr{F}'_{U,v_1})^{-1}\| R \\ &\le \tfrac{1}{2}\, L\, \|v_1 - v_2\|_{\mathscr{X}} \left(\rho^{-1}\alpha(U)^{-1}\right) \left(\rho^{-1}\alpha(U)^{-1}\|\mathscr{F}(U)\|_{\mathscr{X}}\right) \\ &\le \tfrac{1}{2}\rho^{-1}(1 - \rho)\, \|v_1 - v_2\|_{\mathscr{X}}. \end{aligned}$$

It follows that, if $\frac{1}{2}(1-\rho)/\rho < 1$, which is true whenever $\rho > 1/3$, then $\mathcal{L}$ is a contraction of $B(U,R)$ with contractivity $(1-\rho)/(2\rho)$ and must therefore have a unique fixed point u in $B(U,R)$ which is the unique solution of $\mathscr{F}(u) = 0$ in $B(U,R)$. Given our definition of R, the error estimate follows immediately from the fact that $u \in B(U,R)$. $\square$

Corollary 3.4 *Suppose that $\mathscr{F}$ is continuously differentiable in $\mathscr{A}$, that $U \in \mathscr{A}$ is regular and that there exists $R > 0$ such that $B(U,R) \subset \mathscr{A}$ and $\mathscr{F}'$ is Lipschitz continuous in $B(U,R)$. Then there exists $\delta > 0$ such that for all $f \in B(\mathscr{F}(U),\delta)$ satisfying $\mathscr{F}(U) - f \in \bigcap_{\theta\in B(U,R)} \operatorname{range}(\mathscr{F}'(\theta))$ there exists $u \in B(U,R)$ such that $\mathscr{F}(u) = f$.*

Proof. The result follows immediately upon setting $\rho = 1/2$ and replacing $\mathscr{F}$ by $\mathscr{F} - f$ in Theorem 3.3. $\square$

With the help of the Inverse Function Theorem, the following continuation result can be easily shown.

Theorem 3.5 *Let $\mathscr{F}$ be continuously differentiable in $\mathscr{A}$ and let $U \in \mathscr{A}$. Suppose that there exists $R > 0$ such that $B(U,R) \subset \mathscr{A}$, that $\mathscr{F}'$ is locally Lipschitz continuous in $B(U,R)$,*

$$\|\mathscr{F}(U)\|_{\mathscr{Y}^*} \leq R \inf_{v\in B(U,R)} \alpha(v), \qquad \text{and that} \tag{3.8}$$

$$\mathscr{F}(U) \in \operatorname{range}(\mathscr{F}'(\theta)) \qquad \forall \theta \in B(U,R). \tag{3.9}$$

Then there exists $u \in B(U,R)$ (unique if $\inf_{v\in B(U,R)} \alpha(v) > 0$) such that $\mathscr{F}(u) = 0$. Furthermore, we have the estimate

$$\|u - U\|_{\mathscr{X}} \leq \big[\inf_{v\in B(U,R)} \alpha(v)\big]^{-1} \|\mathscr{F}(U)\|_{\mathscr{Y}^*}. \tag{3.10}$$

Proof. Set $\alpha = \inf_{v\in B(U,R)} \alpha(v)$ and note that, unless $\mathscr{F}(U) = 0$, which we exclude without loss of generality, we have implicitly assumed that $\alpha > 0$. For $t \in [0,1]$, let $f_t = (1-t)\mathscr{F}(U)$.

Suppose that $u_t \in B(U,R)$ is a solution to $\mathscr{F}(u_t) = f_t$. By Taylor's Theorem there exists $\theta_t \in \operatorname{conv}\{U, u_t\}$ such that $\mathscr{F}(u_t) - \mathscr{F}(U) = \mathscr{F}'(\theta_t)(u_t - U)$. Upon testing with $\varphi \in \mathscr{Y}$, $\|\varphi\|_{\mathscr{Y}} = 1$, we obtain

$$\alpha\|u_t - U\|_{\mathscr{X}} \leq \alpha(\theta_t)\|u_t - U\|_{\mathscr{X}} \leq \|f_t - \mathscr{F}(U)\|_{\mathscr{Y}^*} \leq t\|\mathscr{F}(U)\|_{\mathscr{Y}^*} \leq t\alpha R \tag{3.11}$$

which implies that $u_t \in B(U,tR)$.

Given a solution u_t, we use Corollary 3.4 to compute solutions u_s for $s \in [t, t+\delta]$, for some $\delta > 0$.

For $t = 0$, we obviously have $u_0 = U$. Hence there exists $T > 0$ such that, for $t \in [0, T)$ there exists $u_t \in B(U, tR)$ satisfying $\mathscr{F}(u_t) = f_t$. Let $T \in (0, 1]$ be maximal and let $t_j \uparrow T$. From

$$\alpha \|u_{t_j} - u_{t_k}\|_{\mathscr{X}} \leq \|\mathscr{F}(u_{t_j}) - \mathscr{F}(u_{t_k})\|_{\mathscr{Y}^*} \leq |t_j - t_k| \|\mathscr{F}(U)\|$$

it follows that (u_{t_j}) is a Cauchy sequence and hence there exists $u_T \in B(U, R)$ such that $u_{t_j} \to u_T$ as $j \to \infty$. Since $\mathscr{F}$ is continuous in $B(U, R)$, $\mathscr{F}(u_T) = f_T$. Since we can apply Corollary 3.4 again, this contradicts the maximality of T unless $T = 1$.

The uniqueness and error estimate in the case $\alpha > 0$ follow immediately. □

Remarks. 1. While the estimation of the residual $\|\mathscr{F}(U)\|_{\mathscr{Y}^*}$ is more or less classical (cf. [94, 95]) the numerical computation of the stability constant $\alpha(U)$ does not seem to be considered common practise. There are essentially two options available.

- If the geometry of the equation $\mathscr{F}(u) = 0$ is not too complicated it may be possible to compute the stability sets $B(U, R)$ and (a bound for) the corresponding stability constant $\alpha = \inf_{v \in B(U,R)} \alpha(v)$ directly. An example of how this can be done is given in §3.2.

- In some situations it may, however, be easier to compute a local Lipschitz constant for $\mathscr{F}'$, particularly when $\mathscr{F}'$ is globally Lipschitz continuous. In that case, Theorem 3.3 may be preferable. Only the stability constant $\alpha(U)$ has to be computed which gives some additional flexibility. ◄

2. Theorem 3.5 still holds if $\mathscr{F}'$ is not locally Lipschitz continuous. By using the fixed point iteration

$$\mathscr{F}'(U)(\mathcal{L}(u) - u) = f - \mathscr{F}(u)$$

in place of (3.7), Corollary 3.4 remains true even if $\mathscr{F}'$ is only assumed to be continuous at U and satisfies $\mathscr{F}(v) \in \text{range}(\mathscr{F}'(U))$ for all $v \in B(U, R)$. Consequently, in Theorem 3.5, if $\mathscr{F}'$ is not Lipschitz continuous, one would have to assume that $\mathscr{F}'$ is continuous in $B(U, R)$ and that $\mathscr{F}(v) \in \text{range}(\mathscr{F}'(w))$ for all $v, w \in B(U, R)$.

If $\mathscr{X} = \mathscr{Y}$ and if this space is reflexive then no condition of this type is required. ◄

3. It can be seen from

$$\Big[\inf_{v\in B(U,R)}\alpha(v)\Big]\,\|u-U\|_{\mathscr{X}}\le\|\mathscr{F}(U)-\mathscr{F}(u)\|_{\mathscr{Y}^*}\le\Big[\sup_{v\in B(U,R)}\|\mathscr{F}'(v)\|\Big]\,\|U-u\|_{\mathscr{X}},$$

that the error estimates in Theorems 3.3 and 3.5 are quasioptimal. ◀

4. One advantage of Theorem 3.3 over Theorem 3.5 is that it provides a straightforward strategy for the verification of the *a posteriori* existence condition and adaptive mesh refinement:

(1) For $\rho=1/2$, compute an upper bound $\eta(U)$ for the residual $\|\mathscr{F}(U)\|_{\mathscr{Y}^*}$, a lower bound $\tilde\alpha(U)$ for the inf-sup constant $\alpha(U)$, and (a bound for) the Lipschitz constant L of $\mathscr{F}'$ in $B(U,R)$.

(2) If $\eta(U)/\tilde\alpha(U)^2>1/4L$, use $\eta(U)\le q\,\tilde\alpha(U)^2/4L$, where $q\in(0,1)$, as a refinement criterion and continue at (1).

(3) Otherwise, let

$$\rho=\frac12+\sqrt{\frac14-\frac{L\eta(U)}{\tilde\alpha(U)^2}},$$

which maximizes ρ subject to (3.4), assuming that L does not change. This gives the error estimate

$$\|u-U\|_{\mathscr{X}}\le\frac{\eta(U)}{\rho\tilde\alpha(U)}.$$ ◀

5. If Theorem 3.5 is used, then R cannot be identified quite so easily. One possibility would be the following:

(1) Compute an upper bound $\eta(U)$ for $\|\mathscr{F}(U)\|_{\mathscr{Y}^*}$ and a lower bound $\tilde\alpha(U)$ for $\alpha(U)$.

(2) Set $R_0=\eta(U)/\tilde\alpha(U)$, fix $0<q<1$ and, for $R_j=R_0q^j$, $j=0,1,\dots,J$, compute a stability estimate

$$\tilde\alpha_j\le\inf_{\theta\in B(U,R_j)}\alpha(\theta).$$

(3) If $\eta(U)>\tilde\alpha_jR_j$ for all j, use $\eta(U)\le q'\max_{j=0,\dots,J}\tilde\alpha_jR_j$, where $q'\in(0,1)$, as a refinement criterion and continue at (1).

(4) Otherwise, find j minimizing R_j (or equivalently maximizing $\tilde\alpha_j$) subject to the constraint $\eta(U)\le\tilde\alpha_jR_j$ to obtain the error estimate

$$\|u-U\|_{\mathscr{X}}\le\frac{\eta(U)}{\tilde\alpha_j}.$$

The above procedures do not cover the case $\tilde{\alpha}(U) \leq 0$. If this situation occurs, it is in general not clear how to proceed. It could either mean that the estimate is not good enough, that the solution is in fact unstable, or even that no exact solution corresponding to the numerically computed one exists. ◀

3.2 Local Minimizers of a Non-Convex Functional

As a first application of the *a posteriori* existence idea, we give an example from materials science. Consider the energy functional

$$\mathcal{I}(u) = \int_0^1 \left[W(u_x) + \tfrac{1}{2}u^2 \right] \mathrm{d}x, \tag{3.12}$$

where W is non-convex. The prototypical example is the double-well energy $W(z) = \frac{1}{4}(z^2-1)^2$. While we restrict the analysis to this particular choice, most of the ideas used are applicable to more general stored energy densities with multiple wells.

If we try to minimize $\mathcal{I}$ in $\mathrm{W}_0^{1,4}(0,1)$, a space in which $\mathcal{I}$ is coercive, we see that any sequence $u^{(j)}$ such that $\|u^{(j)}\|_{\mathrm{L}^2} \to 0$ and $u_x^{(j)} \in \{1,-1\}$ satisfies $\mathcal{I}(u^{(j)}) \to 0$. However, if $\mathcal{I}(u) = 0$, it follows that $u = 0$ and hence $\mathcal{I}(u) \geq W(0) > 0$ and thus the minimizer is not attained.

Minimizing sequences for functionals such as (3.12) develop finer and finer oscillations. For this reason, the functional is often used as a cartoon for the formation of microstructure in materials. A theory of a generalized notion of solution (Young measures) was developed to account for the non-existence of classical minimizers. Using the *a posteriori* existence idea, however, we can quite easily find another class of solutions, previously identified in [9] using entirely different techniques, namely local minimizers which have a *finite structure*. See [67] for an introduction to variational models for microstructures and further references on the subject. For the numerical treatment of multiwell energies using Young measure related ideas, see [12, 28].

We use a Galerkin finite element method to discretize $\mathcal{I}$. Let K be the number of elements and let $\mathcal{T}$ be the finite element mesh with nodes $x_i = i/K$, $i = 0, \ldots, K$. The finite element space $S_0^1(\mathcal{T})$ is defined as

$$S_0^1(\mathcal{T}) = \{u \in \mathrm{H}_0^1(0,1) : u|_{(x_{i-1},x_i)} \text{ is affine for } i = 1, \ldots, K\}.$$

The numerical problem is to (locally) minimize $\mathcal{I}$ in $S_0^1(\mathcal{T})$. To connect the (local) minimization problem to the abstract analysis of §3.1 we define $\mathscr{F}$ and $\mathscr{F}'$, formally

for the moment, by

$$\langle \mathscr{F}(u), \varphi \rangle = \mathcal{I}'(u;\varphi) = \int_0^1 \left[W'(u_x)\varphi_x + u\varphi \right] \mathrm{d}x, \quad \text{and} \tag{3.13}$$

$$\langle \mathscr{F}'(u)v, \varphi \rangle = \mathcal{I}''(u;v,\varphi) = \int_0^1 \left[W''(u_x)v_x\varphi_x + v\varphi \right] \mathrm{d}x. \tag{3.14}$$

A first numerical experiment without the *a posteriori* existence analysis reveals two important facts:

(i) The gradients U_x of the numerically computed local minima U are distributed between the two wells, i.e, $W''(U_x) > 0$ in all experiments performed.

(ii) All computed equilibria are clearly $\mathrm{W}^{1,\infty}$-functions, in the sense that $\|U_x\|_{\mathrm{L}^\infty}$ remains bounded as the mesh size tends to zero.

This suggests that we should use the trial space $\mathscr{X} = \mathrm{W}_0^{1,\infty}(0,1)$ rather than $\mathrm{W}_0^{1,4}(0,1)$ as seems to be suggested by the growth conditions on W. In fact, we can easily see in the next proposition that the $\mathrm{W}^{1,\infty}$-topology is the weakest topology with respect to which we can expect to find local minimizers of $\mathcal{I}$ at all.

Proposition 3.6 *Fix $p \in [4,\infty)$ and $u \in \mathrm{W}_0^{1,p}(0,1)$. Then, for each $\varepsilon > 0$ there exists $u_\varepsilon \in \mathrm{W}_0^{1,p}(0,1)$ such that $\|u_\varepsilon - u\|_{\mathrm{W}^{1,p}} \le \varepsilon$ and $\mathcal{I}(u_\varepsilon) < \mathcal{I}(u)$.*

Proof. Suppose first that $u = 0$. Then the second derivative with respect to u_x of the stored energy density is negative and therefore u cannot be a $\mathrm{W}^{1,p}$-local minimizer.

Now let $u \in \mathrm{W}_0^{1,p}(0,1) \setminus \{0\}$ and denote $m := \max u$. With out loss of generality we assume that $m > 0$. For each $t \in (0,m]$ let A_t be an interval of length at most t such that $u \ge m - t$ in A_t. Since $u \in \mathrm{C}[0,1]$ such intervals exist.

Upon replacing u by an oscillatory function with gradient in $\{-1,1\}$ in the interval A_t we obtain a function $u_t \in \mathrm{W}_0^{1,p}(0,1)$ such that

$$u_t = u \ \text{ in } [0,1] \setminus A_t, \quad \text{ and } \ (u_t)_x \in \{-1,1\} \text{ and } 0 \le u_t \le m - t \text{ in } A_t.$$

It follows immediately that $\mathcal{I}(u_t) < \mathcal{I}(u)$ for all t. Furthermore, we see from

$$\|u_t - u\|_{\mathrm{W}^{1,p}(0,1)} \le \|u\|_{\mathrm{W}^{1,p}(A_t)} + \|u_t\|_{\mathrm{W}^{1,p}(A_t)} \le \|u\|_{\mathrm{W}^{1,p}(A_t)} + (t + m^p t)^{1/p},$$

that $\|u_t - u\|_{\mathrm{W}^{1,p}} \to 0$ as $t \to 0$. $\square$

Remark. While Proposition 3.6 makes a strong argument for the use of $\mathrm{W}^{1,\infty}$ in our analysis, it must be noted that this requires that we disregard the structure of the

minimization problem and reformulate the Euler–Lagrange equations in a new functional framework where they are only formally equivalent to the "natural" framework of $\mathrm{W}_0^{1,4}(0,1)$. It is unclear whether a similar analysis can be performed if the operator $\mathscr{F}$ is taken as the gradient of $\mathcal{I}$ in the natural variational space $\mathrm{W}_0^{1,4}$, i.e., as a map from $\mathrm{W}_0^{1,4}(0,1)$ to its dual $\mathrm{W}^{-1,4/3}(0,1)$. This question is left for investigation at a later time. ◀

After this introductory discussion, we begin with an investigation of the inf-sup constant.

Proposition 3.7 *Let $u \in \mathrm{W}^{1,\infty}(0,1)$ such that $W_0 \le W''(u_x)$ for a.e. $x \in (0,1)$, then*

$$\alpha(u) = \inf_{\substack{v \in \mathrm{W}_0^{1,\infty} \\ |v|_{\mathrm{W}^{1,\infty}}=1}} \sup_{\substack{\varphi \in \mathrm{W}_0^{1,1} \\ |\varphi|_{\mathrm{W}^{1,1}}=1}} \mathcal{I}''(u;v,\varphi) \ge \frac{W_0 \min(W_0,1)}{2\min(W_0,1)+1} =: \tilde{\alpha}(u).$$

Proof. Let $v \in \mathrm{W}_0^{1,\infty}(0,1)$ and $|v|_{\mathrm{W}^{1,\infty}} = 1$. By the definition of $|\cdot|_{\mathrm{W}^{1,\infty}}$, for each $\varepsilon > 0$, there exists a measurable set A_ε such that

$$\frac{1}{|A_\varepsilon|}\left|\int_{A_\varepsilon} v_x \,\mathrm{d}x\right| \ge 1-\varepsilon.$$

Without loss of generality, we assume that $\int_{A_\varepsilon} v_x \,\mathrm{d}x \ge |A_\varepsilon|(1-\varepsilon)$. Since $\int_0^1 v_x \,\mathrm{d}x = 0$ it follows that there exists another measurable set B_ε, with positive measure, such that $\int_{B_\varepsilon} v_x \,\mathrm{d}x \le 0$. We define

$$\tilde{\varphi}_x(x) = \begin{cases} \frac{1}{2}|A_\varepsilon|^{-1}, & \text{if } x \in A_\varepsilon \\ -\frac{1}{2}|B_\varepsilon|^{-1}, & \text{if } x \in B_\varepsilon \\ 0, & \text{otherwise.} \end{cases}$$

If follows that $\tilde{\varphi} \in \mathrm{W}_0^{1,1}(0,1)$ with $|\tilde{\varphi}|_{\mathrm{W}^{1,1}} = 1$ and furthermore,

$$\int_0^1 W''(u_x) v_x \tilde{\varphi}_x \,\mathrm{d}x \ge \tfrac{1}{2} W_0.$$

For $t \in (0,1)$, we define

$$\varphi = c\big(t\tilde{\varphi} + (1-t)v/|v|_{\mathrm{W}^{1,1}}\big),$$

where $c \ge 1$ is such that $|\varphi|_{\mathrm{W}^{1,1}} = 1$. Testing $\mathcal{I}''(u;v,\cdot)$ with φ, we obtain

$$\begin{aligned} \mathcal{I}''(u;v,\varphi) &\ge (1-t)\mathcal{I}''(u;v,v)/|v|_{\mathrm{W}^{1,1}} + t\mathcal{I}''(u;v,\tilde{\varphi}) \\ &\ge (1-t)\min(W_0,1)\|v\|_{\mathrm{H}^1}^2/|v|_{\mathrm{W}^{1,1}} + \tfrac{1}{2}tW_0 - t\|v\|_{\mathrm{L}^2}\|\tilde{\varphi}\|_{\mathrm{L}^2} \end{aligned}$$

We estimate $|v|_{\mathrm{W}^{1,1}} \le \|v\|_{\mathrm{H}^1}$ and $\|\tilde\varphi\|_{\mathrm{L}^2} \le \|\tilde\varphi\|_{\mathrm{L}^\infty} \le \frac12 |\tilde\varphi|_{\mathrm{W}^{1,1}} = \frac12$ to arrive at

$$\mathcal{I}''(u;v,\varphi) \ge \Big[(1-t)\min(W_0,1) - \tfrac12 t\Big]\|v\|_{\mathrm{H}^1} + \tfrac12 t W_0.$$

The first term an the right-hand side vanishes if we set

$$t = \min(W_0,1)/\big(\min(W_0,1) + 1/2\big). \quad \square$$

Proposition 3.7 suggests that we should use $\mathscr{Y} = \mathrm{W}_0^{1,1}(0,1)$ as the test space. Based on this definition, we can now prove that $\mathcal{I}$ is twice continuously differentiable and that $\mathcal{I}''$ is locally Lipschitz continuous.

Proposition 3.8 *Let $\mathcal{I} : \mathrm{W}_0^{1,\infty} \to \mathbb{R}$ be given by* (3.12)*. Its formal derivative $\mathscr{F} = \mathcal{I}'$, defined by* (3.13)*, maps $\mathrm{W}_0^{1,\infty}$ into $(\mathrm{W}_0^{1,1})^*$ and is differentiable in $\mathrm{W}_0^{1,\infty}$ with Lipschitz continuous derivative $\mathscr{F}' = \mathcal{I}''$ given by* (3.14)*.*

Proof. Since $W \in \mathrm{C}^3(\mathbb{R})$, if $u \in \mathrm{W}_0^{1,\infty}$ then $W'(u_x) \in \mathrm{L}^\infty$ and hence $\mathscr{F}(u) \in \mathscr{Y}^*$.

To see that $\mathscr{F}$ is differentiable, we note that the formal derivative coincides with the directional derivative and that $\mathscr{F}'(u)$ defines a bounded linear operator from $\mathrm{W}_0^{1,\infty}$ to $(\mathrm{W}_0^{1,1})^*$.

If $u_1, u_2, \in \mathrm{W}_0^{1,\infty}$ then

$$\begin{aligned}\big|\langle(\mathscr{F}'(u_1) - \mathscr{F}'(u_2))v,\varphi\rangle\big| &\le \int_0^1 \big|W''(u_{1,x}) - W''(u_{2,x})\big||v_x||\varphi_x|\,\mathrm{d}x \\ &\le \|W''(u_{1,x}) - W''(u_{2,x})\|_{\mathrm{L}^\infty}\end{aligned}$$

for all $v \in \mathrm{W}_0^{1,\infty}$ with $\|v_x\|_{\mathrm{L}^\infty} = 1$ and $\varphi \in \mathrm{W}_0^{1,1}$ with $\|\varphi_x\|_{\mathrm{L}^1} = 1$. Since $W \in \mathrm{C}^3(\mathbb{R})$, W'' is Lipschitz continuous in any bounded set which immediately implies Lipschitz continuity of $\mathscr{F}'$ in bounded subsets of $\mathrm{W}_0^{1,\infty}(0,1)$. $\square$

Next, we estimate the residual in the appropriate dual topology.

Proposition 3.9 *Let $U \in S_0^1(\mathcal{T})$ be a critical point of $\mathcal{I}$ in $S_0^1(\mathcal{T})$. Then,*

$$\begin{aligned}\|\mathcal{I}'(U)\|_{\mathscr{Y}^*} &\le \max_{k=1,\dots,K}\eta_k =: \eta(U), \quad \text{where} \\ \eta_k &= h_k\|U\|_{\mathrm{L}^\infty(x_{k-1},x_k)}. \end{aligned} \tag{3.15}$$

Proof. Let $\varphi \in \mathrm{W}_0^{1,1}(0,1)$ and let Φ be its piecewise affine interpolant, i.e. $\Phi \in S_0^1(\mathcal{T})$ and $\Phi(x_i) = \varphi(x_i)$, $i = 0,\dots,K$. By Galerkin orthogonality, we have $\mathcal{I}'(U;\varphi) = \mathcal{I}'(U;\varphi - \Phi)$. In each element (x_{k-1},x_k), since $W'(U_x)$ is piecewise constant, we have

$$\int_{x_k-1}^{x_k} W'(U_x)(\varphi - \Phi)_x\,\mathrm{d}x = 0$$

and, using the Poincaré inequality for $\mathrm{W}_0^{1,1}(0,1)$,

$$\begin{aligned}\int_{x_{k-1}}^{x_k} U(\varphi-\Phi)\,\mathrm{d}x &\leq \|U\|_{\mathrm{L}^\infty(x_{k-1},x_k)}\|\varphi-\Phi\|_{\mathrm{L}^1(x_{k-1},x_k)}\\ &\leq \tfrac{1}{2}h_k\|U\|_{\mathrm{L}^\infty(x_{k-1},x_k)}\|\varphi_x-\Phi_x\|_{\mathrm{L}^1(x_{k-1},x_k)}.\end{aligned}$$

A straightforward computation gives $\|\varphi_x-\Phi_x\|_{\mathrm{L}^1(x_{k-1},x_k)} \leq 2|\varphi|_{\mathrm{W}^{1,1}(x_{k-1},x_k)}$. In summary, we obtain

$$\begin{aligned}|\mathcal{I}'(U;\varphi)| &= |\mathcal{I}'(U;\varphi-\Phi)| \leq \sum_{k=1}^{K} h_k\|U\|_{\mathrm{L}^\infty(x_{k-1},x_k)}|\varphi|_{\mathrm{W}^{1,1}(x_{k-1},x_k)}\\ &\leq \max_{k=1,\dots,K} h_k\|U\|_{\mathrm{L}^\infty(x_{k-1},x_k)}\,|\varphi|_{\mathrm{W}^{1,1}(0,1)}\end{aligned}$$

which implies the stated result. □

The last ingredient required to be able to apply Theorem 3.5 is to show that the residual $\mathcal{I}'(U)$ of a numerical solution U lies in the range of $\mathcal{I}''(v)$ for all v in a neighbourhood of U.

Proposition 3.10 *Let $u, v \in \mathrm{W}_0^{1,\infty}(0,1)$ and suppose that $0 < W_0 \leq W''(v_x)$ for a.e. $x \in (0,1)$. Then, $\mathcal{I}'(u) \in \mathrm{range}(\mathcal{I}''(v))$.*

Proof. Let the symmetric bilinear form $a = \mathcal{I}''(v) : \mathrm{W}_0^{1,\infty}(0,1) \times \mathrm{W}_0^{1,1}(0,1) \to \mathbb{R}$ be given by

$$a(w,\varphi) = \int_0^1 \left[W''(v_x)w_x\varphi_x + w\varphi\right]\mathrm{d}x \qquad \forall w \in \mathrm{W}_0^{1,\infty}(0,1),\ \forall \varphi \in \mathrm{W}_0^{1,1}(0,1).$$

The residual, $\ell = \mathscr{F}(u)$ is given by

$$\ell(\varphi) = \langle \mathscr{F}(u), \varphi\rangle = \int_0^1 \left[W'(u_x)\varphi_x + u\varphi\right]\mathrm{d}x \qquad \forall \varphi \in \mathrm{W}_0^{1,1}(0,1).$$

It is straightforward to see that a can be defined on $\mathrm{H}_0^1(0,1)$ and that on this Hilbert-space, a and ℓ satisfy the conditions of the Lax–Milgram theorem. Hence, there exists $w \in \mathrm{H}_0^1(0,1)$ such that

$$a(w,\varphi) = \ell(\varphi) \qquad \forall \varphi \in \mathrm{H}_0^1(0,1). \tag{3.16}$$

To show that $w \in \mathrm{W}_0^{1,\infty}$, we use an appropriate test function in (3.16). For $j \in \mathbb{N}$, let A_j be a measurable set such that $|A_j| \downarrow 0$ and $|A_j|^{-1}\int_{A_j} w_x\,\mathrm{d}x \uparrow \operatorname{ess.sup} w_x$. We may assume without loss of generality that $w_x \geq 0$ in A_j and, since $\int_0^1 w_x\,\mathrm{d}x = 0$,

there exist measurable sets B_j such that $|B_j| > 0$ and $w_x \le 0$ in B_j. As in the proof of Proposition 3.7, we define

$$\varphi_x^{(j)}(x) = \begin{cases} \frac{1}{2}|A_j|^{-1}, & \text{if } x \in A_j \\ -\frac{1}{2}|B_j|^{-1}, & \text{if } x \in B_j \\ 0, & \text{otherwise.} \end{cases} \tag{3.17}$$

Testing (3.16) with φ_j we obtain

$$\begin{aligned} \tfrac{1}{2}W_0|A_j|^{-1}\textstyle\int_{A_j} w_x \,\mathrm{d}x &\le \displaystyle\int_0^1 W''(v_x)w_x\varphi_x^{(j)}\,\mathrm{d}x \\ &= \displaystyle\int_0^1 \Big[W'(u_x)\varphi_x + u\varphi - w\varphi\Big]\,\mathrm{d}x \\ &\le \|W'(u_x)\|_{\mathrm{L}^\infty} + \|u-w\|_{\mathrm{L}^1}\|\varphi\|_{\mathrm{L}^\infty}. \end{aligned}$$

Since $\mathrm{L}^1(0,1)$ is embedded in $\mathrm{H}^1(0,1)$ and $\mathrm{L}^\infty(0,1)$ in $\mathrm{W}^{1,1}(0,1)$ it follows that ess.sup w_x is finite. □

3.2.1 Numerical results

In our numerical experiments, we use an initial guess for the optimization algorithm of the form

$$U_0(x) = \sum_{m=1}^{M} a_m \sin(m\pi x),$$

where the coefficients a_m are generated randomly. The function U_0 is used as the initial guess for a proximal point optimization algorithm. In this optimization method, to compute the ℓth step U_ℓ, we (locally) minimize the functional

$$\Phi_\ell(v) = \frac{\gamma_\ell}{2}|v - U_{\ell-1}|_{\mathrm{H}^1}^2 + \mathcal{I}(v)$$

over the finite element space $S_0^1(\mathcal{T})$. The parameter γ_ℓ is chosen adaptively. The effect of the proximal point algorithm is essentially that of a gradient flow, discretized in time with maximal step lengths. Typically, the parameter γ_ℓ is chosen zero for some step, in which case the algorithm reduces to Newton's method. For more detail on the implementation, see §5.3.

Once we have obtained an equilibrium U of $\mathcal{I}$ in $S_0^1(\mathcal{T})$, we compute its residual estimate $\eta(U)$ (cf. Proposition 3.9). The radius R is computed using the procedure suggested in the fifth remark in §3.1. In all experiments only the case $\eta \le \tilde{\alpha}R$ occurred and no refinement was necessary. Some examples of the computed equilibria are shown in Figure 3.1.

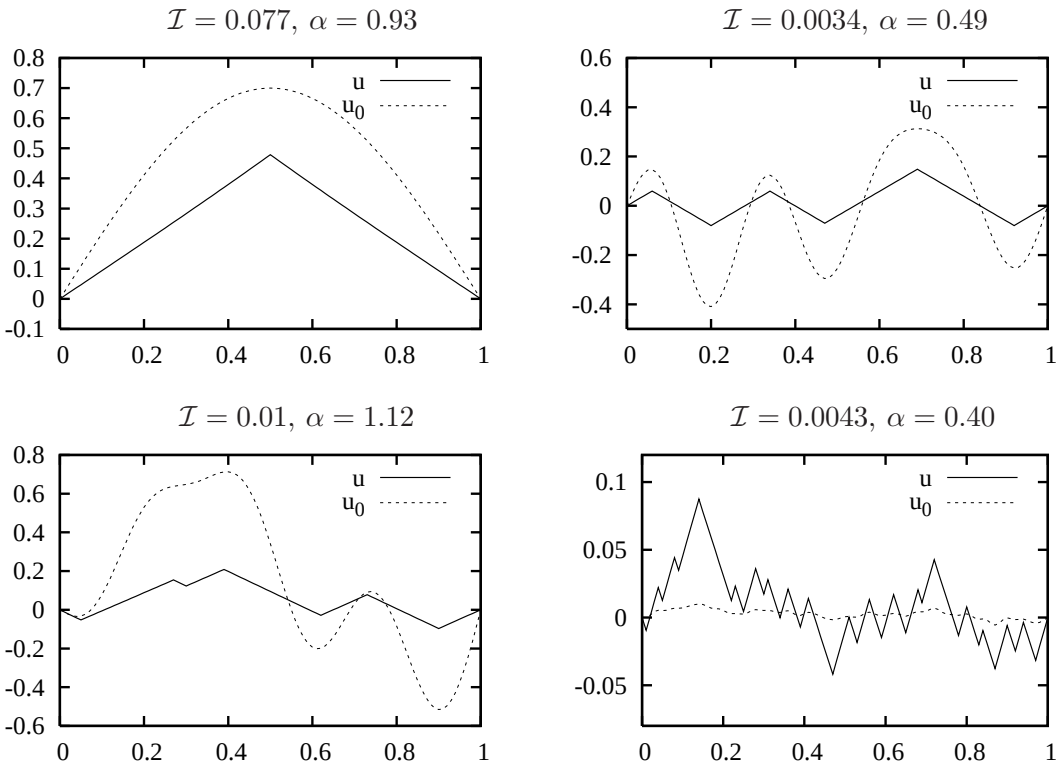

Figure 3.1: Examples of metastable states of the non-convex energy (3.12) as well as the initial conditions used by the optimization method (a proximal point algorithm) to compute them. For all numerical solutions, the existence of a nearby exact solution and an error in the $|\cdot|_{W^{1,\infty}}$-semi-norm at most 10^{-2} is guaranteed.

Since the one-dimensional situation considered here is quite far from three-dimensional reality where the mechanism creating microstructure is not a term such as $\frac{1}{2}u^2$ but a complicated mixing process based on a compatibility condition between gradients, the results presented here can only give very limited information on the physics of microstructure. They do, however, serve as a demonstration that the *a posteriori* existence idea can in principle be applied to highly nonlinear problems with only negligable additional computational effort.

It should also be noted that meta-stable states were analyzed previously in [9]. The theory of viscous dynamics developed therein predicts that the arbitrarily fine microstructure observed when minimizing the energy "almost never" occurs and thus agrees with the results presented in this section.

3.3 A Hilbert Space Example

In this section, we apply the idea of *a posteriori* existence to the semilinear Laplace equation,

$$-\Delta u + f(x,u) = 0, \quad u|_{\partial\Omega} = 0. \tag{3.18}$$

Since the nonlinearity is of a lower order, one can make use of the Hilbert space structure of the problem which makes it possible to compute the stability constant in the numerical solution via eigenvalue computations. While the numerical example remains one-dimensional, the analysis is performed in higher dimensions as well, which is intended to show that the *a posteriori* existence idea is not only restricted to one-dimensional toy problems. For simplicity, we assume that f is differentiable with respect to u and that f_u is globally Lipschitz-continuous in the sense that

$$\big|f_u(x,u_1) - f_u(x,u_2)\big| \le L_{f_u}|u_1 - u_2| \qquad \text{for a.e. } x \in \Omega \quad \forall u_1, u_2 \in \mathbb{R}.$$

In this case the following weak form of the nonlinear operator is well-defined.

Let $d \le 6$ and let Ω be a domain in $\mathbb{R}^d$, i.e., a bounded open and connected set. We set $\mathscr{X} = \mathscr{Y} = \mathrm{H}^1_0(\Omega)$ equipped with the norm $|u|_{\mathrm{H}^1} = \|\nabla u\|_{\mathrm{L}^2}$. The operator $\mathscr{F}\colon \mathrm{H}^1_0(\Omega) \to \mathrm{H}^{-1}(\Omega)$ is defined by

$$\langle \mathscr{F}(u), \varphi\rangle = \int_\Omega \Big[\nabla u \cdot \nabla\varphi + f(x,u)\varphi\Big]\,\mathrm{d}x \qquad \forall \varphi \in \mathrm{H}^1_0(\Omega). \tag{3.19}$$

Proposition 3.11 *The operator $\mathscr{F}$ is differentiable in $\mathrm{H}^1_0(\Omega)$ with derivative*

$$\big\langle \mathscr{F}'(u)v, \varphi\big\rangle = \int_\Omega \Big[\nabla v \cdot \nabla\varphi + f_u(x,u)v\varphi\Big]\,\mathrm{d}x. \tag{3.20}$$

$\mathscr{F}'$ is globally Lipschitz continuous with Lipschitz constant $L_{\mathscr{F}'} = L_{f_u} C_S^3$, where C_S is the Sobolev embedding constant of $\mathrm{H}_0^1(\Omega)$ in $\mathrm{L}^3(\Omega)$.

Proof. The proof is a straightforward application of Taylor's Theorem, the generalised Hölder inequality and Sobolev's embedding of $\mathrm{H}_0^1(\Omega)$ in $\mathrm{L}^3(\Omega)$. □

To discretize the equation $\mathscr{F}(u) = 0$ by a Galerkin finite element method, let $\mathcal{T}$ be a regular partition (cf. [22]) of Ω into d-simplices κ with diameter h_κ, and define

$$S_0^p(\mathcal{T}) = \big\{ V \in \mathrm{H}_0^1(\Omega) : V|_\kappa \text{ is a polynomial of degree } p, \forall \kappa \in \mathcal{T} \big\}.$$

Furthermore, let $\mathcal{E}$ be the set of interior $(d-1)$-dimensional faces in $\mathcal{T}$ and, for $e \in \mathcal{E}$, denote $h_e = \operatorname{diam}(e)$ and ν_e any unit normal vector to e. The finite element method (for simplicity we do not consider quadrature approximations) is then defined by

$$\int_\Omega \Big[\nabla U \cdot \nabla \Phi + f(x,U)\Phi \Big] \,\mathrm{d}x = 0 \qquad \forall \Phi \in S_0^p(\mathcal{T}). \tag{3.21}$$

Following Verfürth [95, Chapter 1], we obtain the following residual estimate. For a detailed discussion of the constant $C(\Omega, \mathcal{T})$, which depends only on the quality of the mesh (cf. [96]). A sketch of the proof is included, mainly to provide a 'continuum model' for the discussion in Chapter 6.

Proposition 3.12 *Let $U \in S_0^p(\mathcal{T})$ satisfy* (3.21)*; then*

$$\begin{aligned} \|\mathscr{F}(U)\|_{\mathrm{H}^{-1}} &\leq C(\Omega,\mathcal{T}) \Big\{ \sum_{e\in\mathcal{E}} \eta_e^2 + \sum_{\kappa\in\mathcal{T}} \eta_\kappa^2 \Big\}^{1/2} =: \eta(U), \quad \text{where} \\ \eta_e^2 &= h_e \int_e \Big| \Big[\frac{\partial U}{\partial \nu_e} \Big] \Big|^2 \,\mathrm{d}s, \quad \text{and} \\ \eta_\kappa^2 &= h_\kappa^2 \int_\kappa \big| -\Delta U + f(x,U) \big|^2 \,\mathrm{d}x. \end{aligned} \tag{3.22}$$

Proof. Let $\varphi \in \mathrm{H}_0^1(\Omega)$ and let $\Phi \in S_0^1(\mathcal{T})$ be defined by setting

$$\Phi(z) = \frac{1}{|B_z|} \int_{B_z} \varphi \,\mathrm{d}x,$$

where B_z is a suitably chosen neighbourhood of the mesh vertex z. This can be done so that

$$\begin{aligned} \|\varphi - \Phi\|_{\mathrm{L}^2(\kappa)} &\leq C_\kappa h_\kappa \|\nabla\varphi\|_{\mathrm{L}^2(T_\kappa)} \qquad \forall \kappa \in \mathcal{T}, \text{ and} \\ \|\varphi - \Phi\|_{\mathrm{L}^2(e)} &\leq C_e h_e^{1/2} \|\nabla\varphi\|_{\mathrm{L}^2(T_e)} \qquad \forall e \in \mathscr{E}, \end{aligned} \tag{3.23}$$

where C_e and C_κ depend only on the local mesh quality and T_κ and T_e are the unions of all elements which respectively intersect with κ and e.

Since $\langle \mathscr{F}(U), \Phi \rangle = 0$, we have

$$\begin{aligned} \langle \mathscr{F}(U), \varphi \rangle &= \langle \mathscr{F}(U), \varphi - \Phi \rangle \\ &= \int_\Omega \big[\nabla U \cdot \nabla(\varphi - \Phi) + f(x,U)(\varphi - \Phi) \big] \,\mathrm{d}x. \end{aligned}$$

Integrating the first term by parts elementwise, we obtain

$$\begin{aligned} \langle \mathscr{F}(U), \varphi \rangle &= \sum_{\kappa \in \mathcal{T}} \Big[\int_{\partial\kappa} \frac{\partial U}{\partial \nu_\kappa} \big(\varphi - \Phi\big) \,\mathrm{d}s + \int_\kappa \big(-\Delta U + f(x,U) \big)(\varphi - \Phi) \,\mathrm{d}x \Big] \\ &= \sum_{e \in \mathscr{E}} \int_e \Big[\frac{\partial U}{\partial \nu_e} \Big] \big(\varphi - \Phi\big) \,\mathrm{d}s + \sum_{\kappa \in \mathcal{T}} \int_\kappa \big(-\Delta U + f(x,U) \big)(\varphi - \Phi) \,\mathrm{d}x. \end{aligned}$$

The residual estimate follows immediately from an application of the Cauchy–Schwarz inequality and the interpolation error estimates (3.23).

For the full details of the proof see [95, Chapter 1] and [96]. □

To avoid an unnecessarily technical discussion of the stability constant, we shall view (3.19) as a minimisation problem again, i.e., to *find* $u \in \mathrm{H}^1_0(\Omega)$ *(locally) minimising the functional*

$$\mathcal{I}(u) = \int_\Omega \Big[\frac{1}{2} |\nabla u|^2 + g(x,u) \Big] \,\mathrm{d}x, \tag{3.24}$$

where $f = g_u$. Thus, we are looking for solutions U to the Galerkin method (3.21) for which $\mathscr{F}'(U)$ is positive. In this case, $\mathscr{F}'(U)$ is an isomorphism and it can be easily seen that

$$\|\mathscr{F}'(U)^{-1}\|^{-1} = \inf_{\substack{\varphi \in \mathrm{H}^1_0(\Omega) \\ \|\nabla\varphi\|_{\mathrm{L}^2} = 1}} \Big\langle \mathscr{F}'(U)\varphi, \varphi \Big\rangle = \inf_{\substack{\varphi \in \mathrm{H}^1_0(\Omega) \\ \|\nabla\varphi\|_{\mathrm{L}^2} = 1}} \int_\Omega \Big[|\nabla\varphi|^2 + f_u(x,U)\varphi^2 \Big] \,\mathrm{d}x.$$

Thus, computing $\|\mathscr{F}'(U)^{-1}\|$ amounts to finding the smallest H^1_0-eigenvalue of $\mathscr{F}'(U)$, i.e., the smallest $\lambda \in \mathbb{R}$ for which there exists $v \in \mathrm{H}^1_0(\Omega)$ such that the pair (λ, v) is a solution of the eigenvalue-problem

$$\int_\Omega \Big[\nabla v \cdot \nabla\varphi + f_u(x,U) v\varphi \Big] \,\mathrm{d}x = \lambda \int_\Omega \nabla v \cdot \nabla\varphi \,\mathrm{d}x \qquad \forall \varphi \in \mathrm{H}^1_0(\mathcal{T}). \tag{3.25}$$

The strong form of (3.25) is

$$-\Delta v + f_u(x,U) v = -\lambda \Delta v.$$

The Galerkin finite element approximation of (3.25) is to find $(\Lambda, V) \in \mathbb{R} \times S_0^p(\mathcal{T})$ such that

$$\int_\Omega \Big[\nabla V \cdot \nabla \Phi + f_u(x, U) V \Phi\Big] \,\mathrm{d}x = \Lambda \int_\Omega \nabla V \cdot \nabla \Phi \,\mathrm{d}x \qquad \forall \Phi \in S_0^p(\mathcal{T}), \tag{3.26}$$

and such that Λ is minimal. A warning should be issued at this point. It is not clear that the smallest eigenvalue of (3.26) approximates the smallest eigenvalue of (3.25). While L^2-eigenvalues of the operator $-\Delta + c\mathrm{Id}$ are in general well seperated and are 'clustered' only at infinity, its H^1-eigenvalues are clustered around 1. This follows from rewriting the eigenvalue problem as

$$(-\Delta)^{-1} c v = (1 - \lambda) v,$$

and noting that $v \mapsto cv$ is bounded in L^2 while the operator $(-\Delta)^{-1}$ is compact. Since their composition must also be compact, it follows that the eigenvalues of $v \mapsto (-\Delta)^{-1} cv$ are clustered at zero, and hence those of the original problems are clustered at one. In particular, this implies that the H^1-eigenvalue problem may be unstable.

The following discussion is based on ideas in [56], to which the reader is also referred to for further references on the adaptive solution of (L^2-) eigenvalue problems.

Set $c(x) = f_u(x, U(x))$ and let (Λ, V) be any solution of (3.26). Obviously, we have $\lambda \leq \Lambda$. Let v be the elliptic projection of V onto the eigenspace of λ; then

$$\int_\Omega \Big[\nabla v \cdot \nabla V + cvV - \lambda \nabla v \cdot \nabla V\Big] \,\mathrm{d}x = 0.$$

We add and subtract $\Lambda(\nabla v, \nabla V)$ to obtain

$$\int_\Omega \Big[\nabla v \cdot \nabla V + cvV - \Lambda \nabla v \cdot \nabla V\Big] \,\mathrm{d}x + (\Lambda - \lambda) \int_\Omega \nabla v \cdot \nabla V \,\mathrm{d}x = 0.$$

Using the fact that v is the elliptic projection of V, we can rearrange this equality to yield

$$(\Lambda - \lambda) \|\nabla v\|_{\mathrm{L}^2}^2 = - \int_\Omega \Big[\nabla v \cdot \nabla V + cvV - \Lambda \nabla v \cdot \nabla V\Big] \,\mathrm{d}x.$$

At this point, we need to assume that $\|\nabla v\|_{\mathrm{L}^2}^2 \geq 1 - \delta$ for some $\delta \in [0, 1)$. Since v and $v - V$ are orthogonal, this is equivalent to $\|v - V\|_{\mathrm{L}^2}^2 \leq \delta < 1$. In this case, we have

$$\Lambda - \lambda \leq \frac{-1}{1 - \delta} \int_\Omega \Big[\nabla v \cdot \nabla V + cvV - \Lambda \nabla v \cdot \nabla V\Big] \,\mathrm{d}x.$$

The estimation of the residual on the right-hand side, using the usual procedure of Galerkin orthogonality, integration by parts and a Clément-type interpolation error estimate (cf. [95, Chapter 1] and the proof of Proposition 3.12), gives the following result.

Proposition 3.13 *Let πV be the elliptic projection of V onto the eigenspace of λ. If $\delta = \|\nabla(\pi V - V)\|^2_{\mathrm{L}^2} < 1$ then*

$$\begin{aligned} \Lambda - \lambda &\leq \frac{C(\Omega,\mathcal{T})}{1-\delta}\Big\{|1-\Lambda|\sum_{e\in\mathcal{E}}\theta_e^2 + \sum_{\kappa\in\mathcal{T}}\theta_\kappa^2\Big\} =: \theta(\Lambda,V), \quad \text{where} \qquad (3.27)\\ \theta_e^2 &= h_e\int_e \Big|\Big[\frac{\partial V}{\partial \nu_e}\Big]\Big|^2 \mathrm{d}s, \quad \text{and} \\ \theta_\kappa^2 &= h_\kappa^2\int_\kappa \Big| -(1-\Lambda)\Delta V + f_u(x,U)V\Big|^2 \mathrm{d}x. \end{aligned}$$

Of course, $\delta \leq 1$ is always satisfied and $\delta < 1$ unless πV and V are orthogonal. Hence, except if $f_u(x,U)$ is highly irregular, or eigenvalues are excessively clustered near zero, the condition $\|\nabla(\pi V - V)\|^2_{\mathrm{L}^2} < 1$ should not pose a problem in practise. Care has to be taken, however, and the factor $1/(1-\delta)$ which multiplies the error estimate and which has to be postulated should be justified in each situation.

Having obtained a solution U to the Galerkin method (3.21), we compute the smallest eigenvalue of $\mathscr{F}'(U)$ in $S_0^p(\mathcal{T})$ using (3.26), and define the local ellipticity constant

$$c_0(U) = \begin{cases} \Lambda - \theta(\Lambda,V), & \text{if } \delta = |\pi V - V|^2_{\mathrm{H}^1} < 1 \\ 0, & \text{otherwise.} \end{cases}$$

If $c_0(U) > 0$ then $\mathscr{F}'(U)$ is positive and satisfies $\|\mathscr{F}'(U)^{-1}\| \leq c_0(U)^{-1}$. This formulation is purely hypothetical of course, since we cannot know whether V approximates the eigenspace of λ. It is demonstrated in §3.3.1 how this situation may be dealt with in practise.

Combined with the findings of this section, Theorem 3.3 with $\rho = 1/2$ gives the following *a posteriori* existence result.

Theorem 3.14 *Let $U \in S_0^p(\mathcal{T})$ be a solution of (3.21). If $c_0(U) > 0$ and*

$$\frac{\eta(U)}{c_0(U)^2} \leq \big(4L_{f_u}C_S^3\big)^{-1}$$

then there exists a solution $u \in \mathrm{H}_0^1(\Omega)$ of (3.19) (a strict local minimum of the energy $\mathcal{I}$ defined in (3.24)) which satisfies

$$\|\nabla(u-U)\|_{\mathrm{L}^2} \leq 2\,\frac{\eta(U)}{c_0(U)}.$$

3.3.1 Numerical results

To avoid the tedious computation of the constants $C(\Omega, \mathcal{T})$ featuring in Propositions 3.12 and 3.13, we restrict our experiments to the one-dimensional setting. The polynomial degree of the finite element method is $p = 1$. We choose the nonlinear reaction term

$$f(x, u) = u^2 - a$$

which corresponds to the energy density $g(x, u) = \frac{1}{3}u^3 - au$, where a is a real constant. It is easy to see that $\inf_{\mathrm{H}^1_0} \mathcal{I} = -\infty$ for every $a \in \mathbb{R}$. We are therefore interested in determining values of a for which 'metastable' states of $\mathcal{I}$, i.e., solutions of (3.18) at which $\mathcal{I}''$ is positive, exist. It is quite straightforward to see analytically that for $a \geq 0$ there exists a stable solution for which $u \geq 0$. This can immediately be extended to a less than, but sufficiently close to zero. With the *a posteriori* existence it is possible to find a relatively sharp lower bound on a for which solutions exist.

A slightly improved estimation of the Lipschitz constant of $\mathscr{F}'$ gives

$$L_{\mathscr{F}'} \leq \pi^{-2}$$

Upon using the nodal interpolant rather than the Clément interpolant for the residual estimates (cf. Proposition 3.9), we also obtain

$$\|\mathscr{F}(U)\|^2_{\mathscr{Y}^*} \leq \sum_{k=1}^{K} \eta_k^2 =: \eta(U), \quad \text{where}$$
$$\eta_k = \frac{h_k^2}{\pi^2} \|f(U)\|^2_{\mathrm{L}^2(x_{k-1}, x_k)},$$

in place of (3.22), and

$$\Lambda - \lambda \leq \frac{1}{1-\delta} \sum_{k=1}^{K} \theta_k^2 =: \theta(\Lambda, V), \quad \text{where}$$
$$\theta_k = \frac{h_k^2}{\pi^2} \|f_u(U)V\|^2_{\mathrm{L}^2(x_{k-1}, x_k)},$$

in place of (3.27).

In all experiments, it was sufficient to compute the three smallest discrete eigenvalues in order to obtain a trustworthy stability constant. The smallest discrete eigenvalue was an isolated eigenvalue near zero. The second was always larger than 0.7 while the third was always larger than 0.9. This indicates clearly that all further eigenvalues will cluster closely around 1. The shape of the eigenfunction corresponding to the smallest

it.	$\sharp\mathcal{T}$	λ	c_0	$4L\eta/c_0^2$	err. est.
1	20	0.070	0.051	30.6	—
2	39	0.039	0.032	27.8	—
3	77	0.027	0.024	18.0	—
4	153	0.022	0.021	8.12	—
5	305	0.021	0.021	3.06	—
6	609	0.021	0.021	1.09	—
7	803	0.021	0.021	0.95	0.049

Table 3.1: Details of the refinement iterations, based on Remark 4 in §3.1, for the model problem (3.18) with $a = -22.6$. $\sharp T$ denotes the number of mesh nodes, λ the smallest discrete eigenvalue and c_0 the resulting stability constant. The convergence of the stability constant indicates the existence of a nearby exact solution.

discrete eigenvalue was smooth and therefore it was save to assume that δ was small. To be save, however, $\delta = 9/10$ was used.

We describe two experiments, taking $a = -22.6$ and $a = -22.61$. In both cases, for different mesh sizes, a local minimum of the energy functional was computed. It is well-known of course that a local minimum of the finite element method may not correspond to any critical point of the exact problem. The *a posteriori* existence method is able to capture such a situation as demonstrated in the following.

Our first experiment, the refinement iterations of which are shown in Table 3.1, is with $a = -22.6$. We observe that the stability constant converges, while the value $4L\eta/c_0^2$ decreases quite rapidly. As it decays below 1, the resulting error estimate is $|u - U|_{\mathrm{H}^1} \leq 0.049$.

In contrast, in Table 3.2, which shows the experiment for $a = -22.61$, the stability constant clearly tends to zero while the number $4L\eta/c_0^2$ tends to $+\infty$. After the fifth iteration, at the very latest, one should become highly suspicious of the computed numerical solution. And indeed, at the next refinement iteration, the local optimization method did not terminate, and the energy 'converged' to $-\infty$.

Conclusion

In this chapter, an extension and refinement of the methodology for *a posteriori* (error) analysis of nonlinear equations was suggested. Based on the numerical computation or analytical estimation of a local stability constant, the assumption usually employed that an exact nearby solution exists may be omitted. Two general strategies in an abstract Banach space setting were presented. Possible applications are:

it.	$\sharp\mathcal{T}$	λ	c_0	$4L\eta/c_0^2$	err. est.
1	20	0.067	0.048	35.1	—
2	39	0.033	0.026	42.6	—
3	77	0.016	0.014	55.8	—
4	153	0.0075	0.0066	87.2	—
5	305	0.0019	0.0016	521	—

Table 3.2: Details of the refinement iterations, based on Remark 4 in §3.1, for the model problem (3.18) with $a = -22.61$. $\sharp T$ denotes the number of mesh nodes, λ the smallest discrete eigenvalue and c_0 the resulting stability constant. The divergence of the stability constant indicates that the numerical solution is spurious, i.e., it does not correspond to an exact solution of (3.18).

- Rigorous error bounds in adaptive numerical computations;
- Experimental investigation of solutions to equations where no general existence theory is available but linearized stability estimates can be obtained; for example, a survey of parameter regimes.

Two examples were given to demonstrate some practical aspects of the *a posteriori* existence idea and to show that it can indeed be employed in practically relevant situations. Both examples were of such a nature that with sufficiently high effort, similar results could have been obtained analytically. This is a typical effect in the application of *a posteriori* existence. If the analytical knowledge about an equation is very poor then it would indeed be very difficult to find estimates for the stability constant. Improved methodologies that use as little analytical knowledge as possible are under investigation.

The nonlinearity of the second example is only of lower order which makes it possible to use Hilbert space techniques only. By contrast, the techniques used in the example of §3.2 are truly one-dimensional, as is the analysis in Chapter 5. The use of the $\mathrm{W}^{1,\infty}$-topology makes it by no means straightforward to extend it to higher dimensions. This is currently the biggest restriction of the work presented. Were it lifted, a wealth of new applications would present themselves.

Chapter 4

A Priori Analysis of the Quasicontinuum Method in One Dimension

The lack of a convincing analysis of the QC method, even in one dimension, was outlined in detail in §1.4.5. The purpose of this chapter is to close this gap by developing a technique to derive optimal *a priori* error estimates for stable solutions. We shall see that even in the one-dimensional case the analytical effort is considerable. In order to keep the presentation simple, we consider only pair-interaction energies with interaction potentials of Lennard–Jones type, but this should not be a true restriction. A detailed description of our model problem is given in §4.2.1 and of the version of the QC method analyzed in §4.2.2.

While, to some extent, the computations to obtain the coercivity are contained in [59] the novelty of the approach is to look at coercivity and stability with respect to the $\mathrm{w}_\varepsilon^{1,\infty}$-norm, a discrete version of the $\mathrm{W}^{1,\infty}$ Sobolev norm which is defined in §4.1. This makes it possible to give precise conditions under which local monotonicity assumptions, such as Assumptions 1. and 2. in [60], are justified. Furthermore, to the best of the author's knowledge, the case of defects has so far not been analyzed for a realistic QC model with long-range interactions.

Probably the most remarkable feature of atomistic models is the large number of critical points. Already in one dimension, it is fairly straightforward to see for many problems that the number of solutions is at least as large as the number of atoms in the body. Therefore, error estimates must necessarily be restricted to local results. Due to the possibility of fracture, stability of solutions can only be obtained with respect to the $\mathrm{w}_\varepsilon^{1,\infty}$-norm or a similar topology. The basic idea is to show that if the mesh is

able to resolve the exact solution (this can be measured in terms of the interpolation error) then there exists a nearby QC solution for which an error estimate holds.

Ultimately, however, the results are only realistic in the elastic case, in the sense that we can actually expect to find the QC solution which satisfies the error estimate that we derive. For example, when an exact solution we wish to approximate is a fractured state then we can prove under certain conditions that there exists a nearby QC solution, however, we should not expect to find it numerically. If only one atom lies on the wrong side of the crack then the error in the discrete $\mathrm{w}^{1,\infty}_{\varepsilon}$-norm cannot 'converge' to zero.

Therefore, in Chapter 5, for the *a posteriori* error analysis of the QC method, we reverse the role of the exact and the QC solution. We derive bounds on the residual of the QC solution and show that, if the QC solution is stable and its residual sufficiently small, there exists an exact solution of the atomistic model for which we give an *a posteriori* error bound.

4.1 Discrete Function Spaces

It will be notationally convenient to define discrete versions of the usual Sobolov norms. First, for $u = (u_i)_{i=0}^N \in \mathbb{R}^{N+1}$, we introduce the discrete derivatives

$$u_i' = \frac{u_i - u_{i-1}}{\varepsilon}, \quad i = 1, \ldots, N \quad \text{and} \quad u_i'' = \frac{u_{i+1} - 2u_i + u_{i-1}}{\varepsilon^2}, \quad i = 1, \ldots, N-1,$$

where ε is a lattice parameter that can be adjusted to the problem at hand and should roughly be the distance between two neighbouring atoms in an undeformed state. For $1 \le p < \infty$, $u \in \mathbb{R}^{N+1}$, $0 \le i_1 \le i_2 \le N$, we define the (semi-)norms

$$\begin{aligned}
\|u\|_{\ell^p_\varepsilon(\!(i_1,i_2)\!)} &= \left(\sum_{i=i_1}^{i_2} \varepsilon |u_i|^p \right)^{1/p}, \\
|u|_{\mathrm{w}^{1,p}_\varepsilon(\!(i_1,i_2)\!)} &= \left(\sum_{i=i_1+1}^{i_2} \varepsilon |u_i'|^p \right)^{1/p}, \quad \text{and} \\
|u|_{\mathrm{w}^{2,p}_\varepsilon(\!(i_1,i_2)\!)} &= \left(\sum_{i=i_1+1}^{i_2-1} \varepsilon |u_i''|^p \right)^{1/p}.
\end{aligned}$$

For $p = \infty$, we define the corresponding versions,

$$\begin{aligned}
\|u\|_{\ell^\infty_\varepsilon(\!(i_1,i_2)\!)} &= \max_{i=i_1,\ldots,i_2} |u_i|, \\
|u|_{\mathrm{w}^{1,\infty}_\varepsilon(\!(i_1,i_2)\!)} &= \max_{i=i_1+1,\ldots,i_2} |u_i'|, \quad \text{and} \\
|u|_{\mathrm{w}^{2,\infty}_\varepsilon(\!(i_1,i_2)\!)} &= \max_{i=i_1+1,\ldots,i_2-1} |u_i''|.
\end{aligned}$$

Sums or maxima taken over empty sets are understood to be zero. If the label $((i_1, i_2))$ is omitted we mean $i_1 = 0, i_2 = N$. For reasons that will become apparent below, we will only require the $p = 1$ and $p = \infty$ versions of these (semi-)norms in our analysis. $B(y, R)$ is understood to be the closed ball, centre y, radius R, with respect to the $\mathrm{w}^{1,\infty}_\varepsilon$-semi-norm.

For $u, v \in \mathbb{R}^{N+1}$, we define the bilinear form

$$\langle u, v \rangle_\varepsilon = \sum_{i=0}^{N} \varepsilon u_i v_i.$$

4.1.1 Functionals

We fix the notation for derivatives of functionals. Let $\phi\colon \mathbb{R}^{N+1} \to \mathbb{R}$ be differentiable at a point $u \in \mathbb{R}^{N+1}$. We understand the derivative of ϕ in u as a linear functional $\phi'(u) = \phi'(u;\cdot)\colon \mathbb{R}^{N+1} \to \mathbb{R}$ defined by

$$\phi(u + v) = \phi(u) + \phi'(u; v) + o(|v|), \quad \text{as } v \to 0,$$

where $|v|$ denotes the Euclidean norm of v. Similarly, if ϕ is twice differentiable at $u \in \mathbb{R}^{N+1}$, the second derivative of ϕ at u is a symmetric bilinear form $\phi''(u) = \phi''(u;\cdot,\cdot)\colon \mathbb{R}^{N+1} \times \mathbb{R}^{N+1} \to \mathbb{R}$ defined by

$$\phi(u + v) = \phi(u) + \phi'(u; v) + \phi''(u; v, v) + o(|v|^2), \quad \text{as } v \to 0.$$

When ϕ' is interpreted as a linear functional we may also write $\phi'(u; v) = \phi'(u)v$. Similarly, we shall write $\phi''(u)v$ for the linear functional defined by $\phi''(u; v, \cdot)$.

4.1.2 Auxiliary results

In this section, we collect some results that are used throughout this chapter, and which are closely related to, and whose formulation and proof do not differ much from, their corresponding continuum versions. The important fact to note is that all bounds are uniform in ε.

Lemma 4.1 *Let $(g_i)_{i=1}^L \in \mathbb{R}^L$ and $\sum_{i=1}^L g_i = 0$; then*

$$|g_i| \le L^{-1} \sum_{k=2}^{L} |g_k - g_{k-1}| \phi_{i,k}, \qquad i = 1, \dots, L, \tag{4.1}$$

where $\phi_{i,k} = k - 1$ for $k = 2, \dots, i$ and $\phi_{i,k} = L - k + 1$ for $k = i + 1, \dots, L$.

Proof. Set $\varepsilon = 1$ and let $i \in \{1, \ldots, L\}$; then

$$\begin{aligned}
|g_i| &= \left| g_i - L^{-1} \sum_{j=1}^{L} g_j \right| \\
&= L^{-1} \left| \sum_{j=1}^{L} (g_i - g_j) \right| \\
&\leq L^{-1} \sum_{j=1}^{i-1} |g_i - g_j| + L^{-1} \sum_{j=i+1}^{L} |g_i - g_j| \\
&\leq L^{-1} \sum_{j=1}^{i-1} \sum_{k=j+1}^{i} |g'_k| + L^{-1} \sum_{j=i+1}^{L} \sum_{k=i+1}^{j} |g'_k| \\
&= L^{-1} \sum_{k=2}^{i} |g'_k| \sum_{j=1}^{k-1} 1 + L^{-1} \sum_{k=i+1}^{L} |g'_k| \sum_{j=k}^{L} 1 \\
&= L^{-1} \sum_{k=2}^{i} |g'_k|(k-1) + L^{-1} \sum_{k=i+1}^{L} |g'_k|(L-k+1). \quad \square
\end{aligned}$$

Lemma 4.2 (Discrete Friedrichs and Poincaré Inequalities) *Suppose that $L \geq 1$, and that $(f_i)_{i=0}^{L} \in \mathbb{R}^{L+1}$ and $(g_i)_{i=1}^{L} \in \mathbb{R}^{L}$ such that $f_0 = f_L = 0$ and $\sum_{i=1}^{L} g_i = 0$. For $p \in \{1, \infty\}$ we have*

$$\|f\|_{\ell^p_\varepsilon(\!(0,L)\!)} \leq \frac{1}{2}(\epsilon L)|f|_{\mathrm{w}^{1,p}_\varepsilon(\!(0,L)\!)}, \quad \text{and} \tag{4.2}$$

$$\|g\|_{\ell^p_\varepsilon(\!(1,L)\!)} \leq \frac{1}{2}(\epsilon L)|g|_{\mathrm{w}^{1,p}_\varepsilon(\!(1,L)\!)}. \tag{4.3}$$

Proof. First, we note that all occurrences of ε can be removed from the results by simple cancellation. Furthermore, the inequalities are trivial if $L = 1$. Thus, we assume without loss of generality that $\varepsilon = 1$ and $L \geq 2$.

We begin with the case $p = 1$. To obtain (4.2), consider

$$\begin{aligned}
\sum_{i=0}^{L} |f_i| &= \sum_{i=1}^{L-1} |f_i| \\
&= \frac{1}{2} \sum_{i=1}^{L-1} \left[\Big| \sum_{j=1}^{i} (f_j - f_{j-1}) \Big| + \Big| \sum_{j=i+1}^{L} (f_j - f_{j-1}) \Big| \right] \\
&\leq \frac{1}{2} \sum_{i=1}^{L-1} \sum_{j=1}^{L} |f_j - f_{j-1}| \\
&= L \frac{1}{2} \Big(1 - \frac{1}{L} \Big) \sum_{j=1}^{L} |f_j - f_{j-1}|.
\end{aligned}$$

To obtain (4.3), we sum inequality (4.1) over $i = 1, \ldots, L$ to obtain

$$\begin{aligned}
\sum_{i=1}^{L} |g_i| &\leq L^{-1} \sum_{i=1}^{L} \sum_{k=2}^{i} |g'_k|(k-1) + L^{-1} \sum_{i=1}^{L} \sum_{k=i+1}^{L} |g'_k|(L-k+1) \\
&= L^{-1} \sum_{k=2}^{L} |g'_k| \sum_{i=k}^{L} (k-1) + L^{-1} \sum_{k=2}^{L} |g'_k| \sum_{i=1}^{k-1} (L-k+1) \\
&= \frac{2}{L} \sum_{k=2}^{L} |g'_k|(k-1)(L-k+1) \\
&\leq 2L \max_{k=2,\ldots,L} \Big(\frac{k-1}{L}\Big)\Big(1 - \frac{k-1}{L}\Big) \sum_{k=2}^{L} |g'_k| \;\leq\; \frac{L}{2} \sum_{k=2}^{L} |g'_k|.
\end{aligned}$$

For $p = \infty$, suppose that $|f_i| = \max_{j=0,\ldots,L} |f_j|$; then

$$\max_{j=0,\ldots,L} |f_j| = |f_i| \leq \sum_{j=1}^{i} |f_j - f_{j-1}| \leq i \max_{j=1,\ldots,L} |f_j - f_{j-1}|.$$

Similarly, we also have

$$\max_{j=0,\ldots,L} |f_j| = |f_i| \leq \sum_{j=i+1}^{L} |f_j - f_{j-1}| \leq (L-i) \max_{j=1,\ldots,L} |f_j - f_{j-1}|,$$

and therefore,

$$\max_{j=0,\ldots,L} |f_j| \leq \min(i, L-i) \max_{j=1,\ldots,L} |f_j - f_{j-1}|,$$

which gives (4.2) with $p = \infty$.

Using Lemma 4.1, we have, for each $i = 1, \ldots, L$,

$$\begin{aligned}
|g_i| &\leq L^{-1} \sum_{j=2}^{i} |g'_j|(j-1) + L^{-1} \sum_{j=i+1} |g'_j|(L-j+1) \\
&\leq \frac{1}{L} \max_{j=2,\ldots,L} |g'_j| \frac{1}{2}\big[i(i-1) + (L-i)(L-i+1)\big] \\
&= \frac{1}{2L} \max_{j=2,\ldots,L} |g'_j| \big[L^2 + L - 2Li + 2i^2 - 2i\big] \\
&= \frac{1}{2L} \max_{j=2,\ldots,L} |g'_j| \big[L(L-1) - 2(L-i)(i-1)\big] \\
&\leq L\Big(\frac{1}{2} - \frac{1}{2L}\Big) \max_{j=2,\ldots,L} |g'_j|. \quad \square
\end{aligned}$$

Note that (4.2) and (4.3) are of course valid for any p with constants independent of ε. Furthermore the optimal Friedrichs constants $C_{p,L}$ and Poincaré constants $\bar{C}_{p,L}$ in

the cases $p \in \{1, 2, \infty\}$ satisfy

$$\begin{aligned} C_{1,L} &= \frac{1}{2} - \frac{1}{2L}, \quad \bar{C}_{1,L} = \begin{cases} 1/2, & \text{if } L \text{ is even,} \\ (1/2) - (1/2L), & \text{if } L \text{ is odd.} \end{cases}, \\ C_{\infty,L} &= \begin{cases} 1/2, & \text{if } L \text{ is even,} \\ (1/2) - (1/2L), & \text{if } L \text{ is odd.} \end{cases}, \text{ and } \bar{C}_{\infty,L} = \frac{1}{2} - \frac{1}{2L}, \text{ and} \\ \frac{1}{\pi} &= \lim_{L\to\infty} C_{2,L} \leq C_{2,L} = \bar{C}_{2,L} = \frac{1}{2L\sin(\pi/(2L))} \leq C_{2,2} = 8^{-1/2}, C_{2,1} = \bar{C}_{2,1} = 0. \end{aligned}$$

In one dimension we also have the following embedding inequality.

Lemma 4.3 *Let $(f_i)_{i=0}^L \in \mathbb{R}^{L+1}$ with $f_0 = f_L = 0$. Then,*

$$\|f\|_{\ell^\infty_\varepsilon(\!(0,L)\!)} \leq \frac{1}{2}|f|_{\mathrm{w}^{1,\infty}_\varepsilon(\!(0,L)\!)}.$$

Proof. For each $i \in \{1, \ldots, N-1\}$, we have

$$\begin{aligned} |f_i| &\leq \sum_{j=1}^{i} \varepsilon |f'_j| \text{ as well as} \\ |f_i| &\leq \sum_{j=i+1}^{N} \varepsilon |f'_j|. \end{aligned}$$

Adding the two inequalities gives the desired result. □

Finally, we combine the estimates of Lemma 4.2 to obtain the following interpolation error estimates.

Theorem 4.4 (Bounds on the Interpolation Error) *Suppose that $(f_i)_{i=0}^L \in \mathbb{R}^{L+1}$ and let*

$$F_i = f_0 + \frac{i}{L}(f_L - f_0)$$

be the affine interpolant of f. Then, for $p \in \{1, \infty\}$,

$$|f - F|_{\mathrm{w}^{1,p}_\varepsilon(\!(0,L)\!)} \leq \frac{1}{2}(\varepsilon L)|f|_{\mathrm{w}^{2,p}_\varepsilon(\!(0,L)\!)}, \quad \text{and} \tag{4.4}$$

$$\|f - F\|_{\ell^p_\varepsilon(\!(0,L)\!)} \leq \frac{1}{4}(\varepsilon L)^2 |f|_{\mathrm{w}^{2,p}_\varepsilon(\!(0,L)\!)}. \tag{4.5}$$

Proof. First note that the grid function $\tilde{f} = f - F$ satisfies $\tilde{f}_0 = \tilde{f}_L = 0$ and therefore $\sum_{i=1}^L \tilde{f}'_i = 0$. Inequality (4.4) therefore follows directly from (4.3).

The estimate (4.5) can be obtained by applying first (4.2) and then (4.3),

$$\begin{aligned} \|f - F\|_{\ell^p_\varepsilon(\!(0,L)\!)} &\leq \frac{1}{2}(\epsilon L)\|(f-F)'\|_{\ell^p_\varepsilon(\!(1,L)\!)} \\ &\leq \frac{1}{4}(\epsilon L)^2 \|f''\|_{\ell^p_\varepsilon(\!(1,L-1)\!)} \\ &= \frac{1}{4}(\epsilon L)^2 |f|_{\mathrm{w}^{2,p}_\varepsilon(\!(0,L)\!)}. \quad \square \end{aligned}$$

4.2 Model Problem and QC Approximation

4.2.1 The atomistic model problem

Fix $N \in \mathbb{N}$. Each vector $y = (y_i)_{i=0}^N \in \mathbb{R}^{N+1}$ represents a state of an atomistic body, consisting of $N+1$ atoms. To each such *deformation* we associate a pair-potential energy

$$E(y) = \sum_{i=1}^{N} \sum_{j=0}^{i-1} J(y_i - y_j).$$

Upon defining the *lattice parameter* $\varepsilon = 1/N$, and writing y_i instead of εy_i we can rescale the energy to

$$E(y) = \sum_{i=1}^{N} \sum_{j=0}^{i-1} \varepsilon J\big(\varepsilon^{-1}(y_i - y_j)\big), \tag{4.6}$$

without changing the problem. Such a scaling highlights the practically relevant case where ε is small in comparison to the length-scale of the problem. For examples of interaction potentials see §1.2.1.

For the scope of this chapter, we shall not assume that the potential has a cut-off radius, but instead analyze the error that is committed with this simplification. We assume throughout this chapter that there exist $z_0 \in [-\infty, +\infty)$, $z_m, z_t \in \mathbb{R}$ such that $z_0 < z_t/2 < z_m < z_t$, and

$$\begin{aligned} &J \in \mathrm{C}^3(z_0, \infty), \quad J'(z_m) = 0, \quad J''(z_t) = 0, \\ &J(z) \to +\infty \text{ as } z \to z_0+, \quad J(z) = +\infty \quad \forall z \le z_0, \\ &J''(z) \ge 0 \quad \forall z \in (0, z_t] \quad \text{and} \quad J''(z) \le 0 \quad \forall z \in [z_t, \infty). \end{aligned} \tag{4.7}$$

The only condition which is not natural is the assumption that $z_t/2 < z_m$ which considerably simplifies the analysis and is not a true restriction — any realistic interaction potential should satisfy this. For example, assume that J is the Morse potential with $\alpha = 5.0$. In that case, $J'(z_t/2) \approx -659$ while $J'(z_t) \approx 0.024$, i.e., it requires an amount of force, several orders of magnitude larger to compress the specimen beyond $z_t/2$ than it takes to break it.

Before we define what we mean by an atomistic solution, we recall from our discussion in the introduction that atomistic deformations are typically only meta-stable states rather than global minimizers. This has also been amply demonstrated in Chapter 2.

We consider a 'Dirichlet' problem where the atomistic deformation is prescribed at the endpoints. It would also be possible, and in fact easier, to consider a problem

with a Dirichlet condition at one end and a Neumann condition at the other end of the interval. We define the set of admissible deformations as

$$\mathscr{A} = \{y \in \mathbb{R}^{N+1} : y_0 = 0, y_N = y_N^D\} \quad \text{and} \quad \mathscr{A}_0 = \{y \in \mathbb{R}^{N+1} : y_0 = y_N = 0\}. \tag{4.8}$$

Each $f \in \mathbb{R}^{N+1}$ represents a linear body force. The atomistic problem is to *find a critical point of the functional* $E(y) - \langle f, y\rangle_\varepsilon$ *in* $\mathscr{A}$. From the assumptions we have made on the interaction potential it follows that E is differentiable at every point which has finite energy. Thus, a critical point y of $E(y) - \langle f, y\rangle_\varepsilon$ in $\mathscr{A}$ with finite energy must satisfy

$$E'(y; v) = \langle f, v\rangle_\varepsilon \qquad \forall v \in \mathscr{A}_0. \tag{4.9}$$

If y satisfies (4.9), we say that $E'(y) = f$ in $\mathscr{A}$.

Elastic deformations are those whose gradient is sufficiently close to z_m, in a region where the potential J is convex. Such solutions exist whenever f is *sufficiently small*. This is measured with respect to the *dual norm*

$$\|f\|_* = \max_{\substack{v \in \mathscr{A}_0 \\ |v|_{\mathrm{w}_\varepsilon^{1,1}} = 1}} \langle f, v\rangle_\varepsilon.$$

Since we can interpret f as a linear functional, we can extend the definition of the dual norm to linear maps $\ell \colon \mathscr{A}_0 \to \mathbb{R}$ by

$$\|\ell\|_* = \max_{\substack{v \in \mathscr{A}_0 \\ |v|_{\mathrm{w}_\varepsilon^{1,1}} = 1}} |\ell(v)|.$$

For future reference, we define the quantities

$$\rho_1(z) = \sum_{r=2}^{\infty} r|J'(rz)|, \quad \text{and} \tag{4.10}$$

$$\rho_2(z_1, z_2) = \sum_{r=1}^{\infty} r^2 \min_{z_1 \le z \le z_2} J''(rz_m), \tag{4.11}$$

which are important in the analysis of existence and stability of elastic deformations. The quantity $\rho_1(z)$ is an estimate for the residual of the affine deformation $y_i = zi/N$ which we use to derive the existence of a reference state. We shall assume throughout that ρ_1 is continuous in a neighbourhood of z_m which, for the Lennard–Jones and the Morse potentials, follows from elementary calculus. The number $\rho_2(z_1, z_2)$ is used to estimate the inf-sup constant of E'' in the set $\{z_1 \le y_i' \le z_2\}$. For the analysis of the QC approximation, we will also use

$$\rho_3(z_1, z_2) = \sum_{r=1}^{\infty} r^2 \max_{z_1 \le z \le z_2} |J''(rz)|, \tag{4.12}$$

which is a Lipschitz constant of E' in the set $\{z_1 \le y_i' \le z_2\}$.

4.2.2 Quasicontinuum approximation

A QC mesh $\mathscr{T}$ is defined by choosing indices $0 = t_0 < t_1 < \cdots < t_K = N$ and setting $\mathscr{T} = \{t_0, \ldots, t_K\}$. For each $k = 1, \ldots, K$, we set $h_k = \varepsilon(t_k - t_{k-1})$, the physical length of the kth element. The set of piecewise affine deformations is given by

$$S^1(\mathscr{T}) = \Big\{ V \in \mathbb{R}^{N+1} : V_i = \frac{t_k - i}{t_k - t_{k-1}} V_{t_{k-1}} + \frac{i - t_{k-1}}{t_k - t_{k-1}} V_{t_k}, \text{ if } t_{k-1} \le i \le t_k \Big\}.$$

We define the set of admissible QC deformations and QC test functions respectively as

$$\mathscr{A}(\mathscr{T}) = \mathscr{A} \cap S^1(\mathscr{T}) \quad \text{and} \quad \mathscr{A}_0(\mathscr{T}) = \mathscr{A}_0 \cap S^1(\mathscr{T}).$$

For convenience, we sometimes use the notation $\overline{V}_k = V_{t_k}$ and $\overline{V}'_k = V'_{t_k}$ for the nodal values of an $S^1(\mathscr{T})$ function. For our analysis it is also necessary to define the interpolant $\Pi\colon \mathbb{R}^{N+1} \to S^1(\mathscr{T})$ by $\Pi u = (\Pi u_i)_{i=0}^N$ and

$$\Pi u_{t_k} = u_{t_k}, \qquad k = 0, \ldots, K.$$

Note that if $y \in \mathscr{A}$ then $\Pi y \in \mathscr{A}(\mathscr{T})$.

The Galerkin approximation of (4.9) in $\mathscr{A}(\mathscr{T})$ is to *find critical points of* $E(Y) - \langle Y, f\rangle_\varepsilon$ *in* $\mathscr{A}(\mathscr{T})$. Any such critical point $Y \in \mathscr{A}$ must satisfy

$$E'(Y; V) = \langle f, V\rangle_\varepsilon \qquad \forall\, V \in \mathscr{A}_0(\mathscr{T}). \tag{4.13}$$

However, in view of the long-range atomistic interaction, which, for the purpose of evaluating the energy and its derivatives still necessitates the computation of very large sums, it is helpful to make some further approximations to the energy functional. First, it is common to replace J by a cut-off potential $\tilde{J}$, which vanishes outside a certain cut-off radius ρ_c. If the deformation gradient is bounded away from zero, then the number of atoms over which one needs to sum is bounded by a small integer. This purely one-dimensional effect means that it is unnecessary to make any further (summation-rule type) approximations to the atomistic energy; thus we define

$$\tilde{E}(Y) = \sum_{i=1}^{N} \sum_{j=0}^{N-1} \varepsilon \tilde{J}\big(\varepsilon^{-1}(Y_i - Y_j)\big).$$

For the analysis of the coercivity of the QC approximation we will need the quantity

$$\tilde{\rho}_2(z_1, z_2) = \sum_{r=1}^{\infty} r^2 \min_{z_1 \le z \le z_2} \tilde{J}''(rz).$$

To approximate the body force potential, we can use a so-called summation rule, i.e., a discrete version of a quadrature rule. In order to recover the full atomistic problem in the limit, it is reasonable to employ a trapezium rule. Thus, we define the discrete bilinear form

$$\langle f, v \rangle_{\mathscr{T}} = \sum_{i=0}^{N} \varepsilon \Pi(fv)_i.$$

The full QC approximation to (4.9) is then to *find* $Y \in \mathscr{A}(\mathscr{T})$ *satisfying*

$$\tilde{E}'(Y; V) = \langle f, V \rangle_{\mathscr{T}} \qquad \forall\, V \in \mathscr{A}_0(\mathscr{T}). \tag{4.14}$$

4.3 Elastic Deformation

In the first part of this chapter, we consider elastic deformation only.

Theorem 4.5 *Let J satisfy the assumptions of §4.2.1 and, in addition, assume that there exists an $R \in (0, \min(z_m - z_t/2, z_t - z_m))$ such that $\rho_1(z_m) < R\, \rho_2(z_m - R, z_m + R)$. Then the following hold:*

(a) *Coercivity: There exist $z_1, z_2 \in \mathbb{R}$, independent of ε, such that $z_1 < z_m < z_2 < z_t$ and*

$$\inf_{y \in \mathscr{Z}_e} \inf_{\substack{u \in \mathscr{A}_0 \\ |u|_{\mathrm{w}^{1,\infty}_\varepsilon} = 1}} \sup_{\substack{v \in \mathscr{A}_0 \\ |v|_{\mathrm{w}^{1,1}_\varepsilon} = 1}} E''(y; u, v) \geq \tfrac{1}{2} \rho_2(z_1, z_2) =: c_0 > 0, \tag{4.15}$$

where $\mathscr{Z}_e = \{y \in \mathbb{R}^{N+1} : z_1 \leq y_i' \leq z_2, \text{ for } i = 1, \ldots, N\}$.

(b) *Existence: There exist $\delta_1, \delta_2 > 0$, independent of ε, such that for every $y_N^D \in \mathbb{R}$ with $|y_N^D - z_m| < \delta_1$ and for every $f \in \mathbb{R}^{N+1}$ with $\|f\|_* \leq \delta_2$, there exists a solution y_f of (4.9) in $\mathscr{Z}_e$.*

(c) *Stability: Let y_f, y_g be solutions to (4.9) in $\mathscr{Z}_e \cap \mathscr{A}$, corresponding respectively to the right-hand sides $f, g \in \mathbb{R}^{N+1}$; then*

$$|y_f - y_g|_{\mathrm{w}^{1,\infty}_\varepsilon} \leq c_0^{-1} \|f - g\|_*.$$

Theorem 4.5 is of theoretical relevance in that it gives a relatively sharp condition under which elastic solutions to (4.9) exist and are stable. It furthermore directly relates the shape of the interaction potential to the coercivity of the energy. In practice, we would numerically determine a region where E'' is coercive and then prove that it contains a reference state, using the condition $\rho_1(z_m) < \min(z_m - z_1, z_2 - z_m)\rho_2(z_1, z_2)$. We demonstrate this in §4.5.

For the formulation and proof of the *a priori* error bound, there are several options. One could simply formulate a QC version of the existence theorem and prove that the elastic QC solution satisfies an error estimate. However, it seems more illuminating to make fewer assumptions on the structure of the problem, and impose stronger assumptions on a particular solution instead.

For any given $f \in \mathbb{R}^{N+1}$ and a solution $y \in \mathscr{A}$ of (4.9), we identify three error sources: the interpolation error,

$$\mathcal{E}_1 = |y - \Pi y|_{\mathrm{w}_\varepsilon^{1,\infty}}, \tag{4.16}$$

the perturbation of the linear form,

$$\mathcal{E}_2 = \max_{\substack{V \in \mathscr{A}_0(\mathscr{T}) \\ |V|_{\mathrm{w}_\varepsilon^{1,1}} = 1}} \big| \langle f, V \rangle_{\mathscr{T}} - \langle f, V \rangle_\varepsilon \big|, \tag{4.17}$$

and the perturbation of the energy,

$$\mathcal{E}_3 = \max_{Y \in \mathscr{A}(\mathscr{T}) \cap \mathscr{Z}_e} \max_{\substack{V \in \mathscr{A}_0(\mathscr{T}) \\ |V|_{\mathrm{w}_\varepsilon^{1,1}} = 1}} \big| E'(Y;V) - \tilde{E}'(Y;V) \big|. \tag{4.18}$$

Theorem 4.6

(a) *Let $\mathscr{Z}_e$ be defined as in Theorem 4.5; then,*

$$\min_{Y \in S^1(\mathscr{T}) \cap \mathscr{Z}_e} \min_{\substack{U \in \mathscr{A}_0(\mathscr{T}) \\ |U|_{\mathrm{w}_\varepsilon^{1,\infty}} = 1}} \max_{\substack{V \in \mathscr{A}_0(\mathscr{T}) \\ |V|_{\mathrm{w}_\varepsilon^{1,1}} = 1}} E''(Y;U,V) \;\geq\; \tfrac{1}{2}\rho_2(z_1,z_2) = c_0, \text{ and} \tag{4.19}$$

$$\max_{Y \in S^1(\mathscr{T}) \cap \mathscr{Z}_e} \max_{\substack{U \in \mathscr{A}_0(\mathscr{T}) \\ |U|_{\mathrm{w}_\varepsilon^{1,\infty}} = 1}} \max_{\substack{V \in \mathscr{A}_0(\mathscr{T}) \\ |V|_{\mathrm{w}_\varepsilon^{1,1}} = 1}} E''(Y;U,V) \;\leq\; \rho_3(z_1,z_2) =: c_1. \tag{4.20}$$

(b) *Let $y \in \mathscr{Z}_e \cap \mathscr{A}$ be a solution of (4.9) and define $R = \min_{i=1,\dots,N} \min(z_2 - y_i', y_i' - z_1)$. Assume, furthermore, that the QC mesh $\mathscr{T}$ is sufficiently fine so that*

$$c_1\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3 \leq c_0 R. \tag{4.21}$$

Then, there exists a solution $Y \in \mathscr{A}(\mathscr{T}) \cap \mathscr{Z}_e$ of (4.14) which satisfies

$$|y - Y|_{\mathrm{w}_\varepsilon^{1,\infty}} \leq c_0^{-1}\Big((c_0 + c_1)\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3\Big).$$

If $\tilde{\rho}_2(z_1,z_2) > 0$, then the QC solution is unique in $\mathscr{A}(\mathscr{T}) \cap \mathscr{Z}_e$.

(c) *The error quantities $\mathcal{E}_1, \mathcal{E}_2$, and $\mathcal{E}_3$ can be bounded as follows:*

$$\begin{aligned}
\mathcal{E}_1 &\leq \tfrac{1}{2} \max_{k=1,\dots,K} h_k |y|_{\mathrm{w}_\varepsilon^{2,\infty}(\!(t_{k-1},t_k)\!)}, && (4.22)\\
\mathcal{E}_2 &\leq \max_{k=1,\dots,K} h_k^2 \max\big(|f|_{\mathrm{w}_\varepsilon^{2,\infty}(\!(t_{k-1},t_k)\!)}, && (4.23)\\
&\qquad\qquad 2|f|_{\mathrm{w}_\varepsilon^{1,\infty}(\!(t_{k-1}+1,t_k)\!)} + 2|f|_{\mathrm{w}_\varepsilon^{1,\infty}(\!(t_{k-1},t_k-1)\!)}\big), \text{ and}\\
\mathcal{E}_3 &\leq \sum_{r=1}^{\infty} r \max_{z_1\leq z\leq z_2} \big|\tilde{J}'(rz) - J'(rz)\big|. && (4.24)
\end{aligned}$$

4.3.1 Coercivity of the atomistic problem

For this fairly straightforward but tedious analysis it is convenient to rewrite the energy and its derivatives in the following form. First, we rewrite E as

$$E(y) = \sum_{i=1}^{N}\sum_{j=1}^{i} \varepsilon J\left(\sum_{k=j}^{i} y_k'\right). \tag{4.25}$$

For the moment we will only need E'', however, for future reference we first compute E' which can be written in the form

$$\begin{aligned}
E'(y;w) &= \sum_{i=1}^{N}\sum_{j=1}^{i} \varepsilon J'\left(\sum_{k=j}^{i} y_k'\right)\left(\sum_{n=j}^{i} w_n'\right)\\
&= \sum_{i=1}^{N}\sum_{j=1}^{i}\sum_{n=j}^{i} \varepsilon w_n' J'\big(\varepsilon^{-1}(y_i - y_{j-1})\big)\\
&= \sum_{i=1}^{N}\sum_{n=1}^{i} \varepsilon w_n' \sum_{j=1}^{n} J'\big(\varepsilon^{-1}(y_i - y_{j-1})\big)\\
&= \sum_{n=1}^{N} \varepsilon w_n' \left(\sum_{i=n}^{N}\sum_{j=1}^{n} J'\big(\varepsilon^{-1}(y_i - y_{j-1})\big)\right)\\
&= \sum_{n=1}^{N} \varepsilon F_n'(y) w_n', && (4.26)
\end{aligned}$$

where

$$F_n'(y) = \sum_{i=n}^{N}\sum_{j=1}^{n} J'\big(\varepsilon^{-1}(y_i - y_{j-1})\big).$$

If E is twice differentiable at a point y, then $E''(Y)$ is most conveniently written in the form

$$\begin{aligned} E''(y;v,w) &= \sum_{i=1}^{N}\sum_{j=1}^{i} \varepsilon J''\big(\varepsilon^{-1}(y_i - y_{j-1})\big)\left(\sum_{m=j}^{i} v'_m\right)\left(\sum_{n=j}^{i} w'_n\right) \\ &= \sum_{n=1}^{N} \varepsilon w'_n \sum_{i=n}^{N}\sum_{j=1}^{n}\sum_{m=j}^{i} v'_m J''\big(\varepsilon^{-1}(y_i - y_{j-1})\big) \\ &= \sum_{n=1}^{N} \varepsilon w'_n \sum_{i=n}^{N}\sum_{m=1}^{i}\sum_{j=1}^{n\wedge m} v'_m J''\big(\varepsilon^{-1}(y_i - y_{j-1})\big) \\ &= \sum_{n=1}^{N}\sum_{m=1}^{N} \varepsilon w'_n v'_m \left(\sum_{i=m\vee n}^{N}\sum_{j=1}^{n\wedge m} J''\big(\varepsilon^{-1}(y_i - y_{j-1})\big)\right) \\ &= \sum_{n=1}^{N}\sum_{m=1}^{N} \varepsilon F''_{nm} v'_m w'_n, \end{aligned} \tag{4.27}$$

where

$$F''_{nm}(y) = \sum_{i=m\vee n}^{N}\sum_{j=1}^{n\wedge m} J''\big(\varepsilon^{-1}(y_i - y_{j-1})\big).$$

Our aim in this section is to identify a set of deformations,

$$\mathscr{Z}_e = \{y \in \mathscr{A} : z_1 \le y'_i \le z_2\},$$

with $z_1 < z_m < z_2 < z_t$ for which $E''(y)$ satisfies the inf-sup condition

$$\min_{y\in\mathscr{Z}_e} \min_{\substack{u\in\mathscr{A}_0 \\ |u|_{\mathrm{w}^{1,\infty}_\varepsilon}=1}} \max_{\substack{v\in\mathscr{A}_0 \\ |v|_{\mathrm{w}^{1,1}_\varepsilon}=1}} E''(y;u,v) \ge c_0 > 0.$$

For convenience, we have assumed in §4.2.1 that $z_m > z_t/2$, and hence we may assume here that $z_1 \ge z_t/2$ as well. This implies that

$$\begin{cases} J''(z) > 0, & \text{for } z_1 \le z \le z_2, \text{ and} \\ J''(z) \le 0, & \text{for } z \ge 2z_1, \end{cases} \tag{4.28}$$

and consequently $F''_{nm} \le 0$ whenever $n \ne m$.

The proof of the inf-sup condition is based on an argument related to row diagonally dominant matrices. Fix $u \in \mathscr{A}_0$ and choose $p, q \in \{1, \dots, N\}$ such that u'_p is maximal and u'_q is minimal. Since $u \in \mathscr{A}_0$ we have $\sum_{i=1}^{N} u'_i = 0$ and hence $u'_p \ge 0$ and $u'_q \le 0$. We define the test function v by

$$v'_i = \begin{cases} \frac{1}{2}\varepsilon^{-1}, & \text{if } i = p, \\ -\frac{1}{2}\varepsilon^{-1}, & \text{if } i = q, \text{ and} \\ 0, & \text{otherwise.} \end{cases}$$

It is clear from the definition of v that $v \in \mathscr{A}_0$ and $|v|_{\mathrm{w}^{1,1}_\varepsilon} = 1$. Let $P = \{i : u_i' > 0\}$ and $Q = \{i : u_i' < 0\}$. Using (4.27), we have

$$\begin{aligned} E''(y;u,v) &= \sum_{n=1}^{N}\sum_{m=1}^{N} \varepsilon F''_{nm}(y) u_n' v_m' \\ &= \frac{1}{2\varepsilon}\sum_{n=1}^{N} \varepsilon F''_{np}(y) u_n' - \frac{1}{2\varepsilon}\sum_{n=1}^{N} \varepsilon F''_{nq}(y) u_n' \\ &= \frac{1}{2}F''_{pp}(y)u_p' + \frac{1}{2}\sum_{n\neq p} F''_{np}u_n' - \frac{1}{2}F''_{qq}(y)u_q' - \sum_{n\neq q} F''_{nq}(y)u_n'. \end{aligned}$$

Using (4.28), we see that for $n \neq m$ we have $F''_{nm}(y) \leq 0$. Hence, we obtain

$$\begin{aligned} 2E''(y;u,v) &\geq F''_{pp}(y)u_p' + \sum_{m\in P\setminus\{p\}} F''_{mp}(y)u_m' - F''_{qq}(y)u_q' - \sum_{m\in Q\setminus\{q\}} F''_{mq}(y)u_m' \\ &\geq u_p'\Big[F''_{pp}(y) + \sum_{m\in P\setminus\{p\}} F''_{mp}(y)\Big] + (-u_q')\Big[F''_{qq}(y) + \sum_{m\in Q\setminus\{q\}} F''_{mq}(y)\Big] \\ &\geq |u|_{\mathrm{w}^{1,\infty}_\varepsilon} \sum_{m=1}^{N} F''_{mn}(y), \end{aligned} \tag{4.29}$$

where $n \in \{p, q\}$. Thus, to prove the coercivity estimate (4.15), we need to show that the matrix $(F''_{nm})_{n,m=1,\dots,N}$ is strictly row diagonally dominant; more precisely, we need to obtain a lower bound on the sum in the last expression. To do so, we split the sum as follows:

$$\begin{aligned} \sum_{m=1}^{N} F''_{nm}(y) &= \sum_{m=1}^{n-1}\sum_{i=n}^{N}\sum_{j=1}^{m} J''\big(\varepsilon^{-1}(y_i - y_{j-1})\big) + \sum_{m=n+1}^{N}\sum_{j=1}^{n}\sum_{i=m}^{N} J''\big(\varepsilon^{-1}(y_i - y_{j-1})\big) \\ &\quad + \sum_{j=1}^{n}\sum_{i=n}^{N} J''\big(\varepsilon^{-1}(y_i - y_{j-1})\big). \end{aligned}$$

For all pairs (i, j) with $i \geq j$ we bound

$$J''(\varepsilon^{-1}(y_i - y_{j-1})) \geq \min_{z_1\leq z\leq z_2} J''\big((i-j+1)z\big) =: \underline{J}''(i-j+1),$$

which we use to estimate

$$\begin{aligned} \sum_{m=1}^{N} F''_{nm}(y) &\geq \sum_{m=1}^{n-1}\sum_{i=n}^{N}\sum_{j=1}^{m} \underline{J}''(i-j+1) + \sum_{m=n+1}^{N}\sum_{j=1}^{n}\sum_{i=m}^{N} \underline{J}''(i-j+1) \\ &\quad + \sum_{j=1}^{n}\sum_{i=n}^{N} \underline{J}''(i-j+1). \end{aligned} \tag{4.30}$$

In the first triple-sum, we exchange the order of summation three times to obtain

$$\begin{aligned}\sum_{m=1}^{n-1}\sum_{i=n}^{N}\sum_{j=1}^{m}\underline{J}''(i-j+1) &= \sum_{i=n}^{N}\sum_{j=1}^{n-1}\sum_{m=j}^{n-1}\underline{J}''(i-j+1)\\ &= \sum_{j=1}^{n-1}\sum_{i=n}^{N}(n-j)\underline{J}''(i-j+1)\\ &\geq \sum_{j=1}^{n-1}(n-j)\sum_{r=n-j+1}^{\infty}\underline{J}''(r),\end{aligned}$$

where we used the fact that $\underline{J}''(r) \leq 0$ for $r \geq 2$. We change the order of summation again,

$$\begin{aligned}\sum_{j=1}^{n-1}(n-j)\sum_{r=n-j+1}^{\infty}\underline{J}''(r) &= \sum_{r=2}^{\infty}\underline{J}''(r)\sum_{j=n-r+1}^{n-1}(n-j)\\ &= \frac{1}{2}\sum_{r=2}^{\infty}r(r-1)\underline{J}''(r),\end{aligned}$$

where we used $\sum_{j=n-r+1}^{n-1}(n-j) = r(r-1)/2$. Similarly, for the second triple-sum in (4.30), we obtain

$$\sum_{m=n+1}^{N}\sum_{j=1}^{n}\sum_{i=m}^{N}\underline{J}''(i-j+1) \geq \frac{1}{2}\sum_{r=2}^{\infty}r(r-1)\underline{J}''(r).$$

For the third term in (4.30), we have

$$\begin{aligned}\sum_{j=1}^{n}\sum_{i=n}^{N}\underline{J}''(i-j+1) &\geq \sum_{j=1}^{n}\sum_{r=n-j+1}^{\infty}\underline{J}''(r)\\ &= \sum_{r=1}^{\infty}\sum_{j=n-r+1}^{n}\underline{J}''(r)\\ &= \sum_{r=1}^{\infty}r\underline{J}''(r).\end{aligned}$$

On combining this with the previously obtained bounds, and recalling the definition (4.11), we finally arrive at

$$\sum_{m=1}^{N}F''_{nm}(y) \geq \sum_{r=1}^{\infty}r^2\underline{J}''(r) = \rho_2(z_1,z_2). \tag{4.31}$$

Therefore, returning to (4.29), we obtain

$$\max_{\substack{v\in\mathscr{A}_0\\ |v|_{\mathrm{w}_\varepsilon^{1,1}}=1}} E''(y;u,v) \geq c_0|u|_{\mathrm{w}_\varepsilon^{1,\infty}}, \tag{4.32}$$

where $c_0 = \frac{1}{2}\rho_2(z_1, z_2)$. We refer to Appendix 4.5 for specific values of z_1, z_2 and c_0 for the Lennard–Jones and the Morse potential.

4.3.2 Proof of Theorem 4.5

The proof of Theorem 4.5 as well as the extension to fracture solutions in §4.4 rely on a local existence result which is a direct corollary of Theorem 3.5.

Lemma 4.7 *Let $\|\cdot\|$ be a norm in $\mathscr{A}_0$, $R > 0$ and $\tilde{y} \in \mathscr{A}$, and define $\mathscr{Z} = \{y \in \mathscr{A} : \|y - \tilde{y}\| \le R\}$. Suppose, further, that:*

(i) $\Phi\colon \mathbb{R}^{N+1} \to (-\infty, +\infty]$ *is three times continuously differentiable in $\mathscr{Z}$;*

(ii) $\Phi'(\tilde{y}) = \tilde{f}$ *in $\mathscr{A}$; and*

(iii) *there exists $c_0 > 0$ such that for all $y \in \mathscr{Z}$,*

$$c_0 \le \min_{\substack{u \in \mathscr{A}_0 \\ \|u\|=1}} \max_{\substack{v \in \mathscr{A}_0 \\ |v|_{\mathrm{w}_\varepsilon^{1,1}}=1}} \Phi''(y; u, v). \tag{4.33}$$

Then, for each $f \in \mathbb{R}^{N+1}$ satisfying $\|f - \tilde{f}\|_* \le c_0 R$, there exists a unique $y \in \mathscr{Z}$ such that $\Phi'(y) = f$ in $\mathscr{A}$. Furthermore, the solution y satisfies

$$\|y - \tilde{y}\| \le c_0^{-1} \|f - \tilde{f}\|_*. \tag{4.34}$$

Proof. The result follows from Theorem 3.5 by setting $\mathscr{F} = \Phi'$, $\|\cdot\|_{\mathscr{X}} = \|\cdot\|$ and $\|\cdot\|_{\mathscr{Y}} = |\cdot|_{\mathrm{w}_\varepsilon^{1,1}}$. Note furthermore that in finite dimensions, $c_0 > 0$ implies (3.9). □

Lemma 4.7 gives a clear path to the proof of Theorem 4.5. We have already established the necessary conditions for coercivity in the previous section.

To show the existence of a reference state, we define the deformation $y_i^D = \varepsilon i y_N^D$, where we assume that $z_1 < y_N^D < z_2$, and estimate the residual $E'(y^D)$. It is more convenient to do this in the following alternative representation of E':

$$E'(y; v) = \sum_{n=1}^{N-1} E_n'(y) v_i \qquad \forall y \in \mathscr{A}, \forall v \in \mathscr{A}_0, \tag{4.35}$$

where

$$E_n'(y) = \sum_{i=0}^{n-1} J'\big(\varepsilon^{-1}(y_n - y_i)\big) - \sum_{i=n+1}^{N} J'\big((\varepsilon^{-1}(y_i - y_n)\big), \qquad n = 1, \dots, N-1.$$

Using the embedding inequality $\|v\|_{\ell^\infty_\varepsilon} \le \frac{1}{2}|v|_{\mathrm{w}^{1,1}_\varepsilon}$ (cf. Lemma 4.3) we can estimate

$$|E'(y;v)| \le \sum_{n=1}^{N-1} |E'_n(y)|\|v\|_{\ell^\infty_\varepsilon} \le \frac{1}{2}\sum_{n=1}^{N-1} |E'_n(y)||v|_{\mathrm{w}^{1,1}_\varepsilon},$$

which implies that

$$\|E'(y)\|_* \le \frac{1}{2}\sum_{n=1}^{N-1} |E'_n(y)|. \tag{4.36}$$

For $y = y^D$, we have

$$E'_n(y^D) = \sum_{i=0}^{2n-N-1} J'((n-i)y^D_N) - \sum_{i=2n+1}^{N} J'((i-n)y^D_n),$$

and, taking absolute values,

$$|E'_n(y)| \le \sum_{r=n\wedge(N-n)+1}^{\infty} |J'(ry^D_n)|.$$

Thus, we can estimate

$$\begin{aligned}\|E'(y^D)\|_* &\le \frac{1}{2}\sum_{n=1}^{N-1}|E'_n(y^D)| \le \frac{1}{2}\sum_{n=1}^{\infty}\sum_{r=n+1}^{\infty}|J'(ry^D_N)| \\ &\le \frac{1}{2}\sum_{r=2}^{\infty}\sum_{n=1}^{r-1}|J'(ry^D_N)| \le \frac{1}{2}\sum_{r=2}^{\infty}(r-1)|J'(ry^D_N)| = \frac{1}{2}\rho_1(y^D_N).\end{aligned}$$

We now apply Lemma 4.7 with $\Phi = E$, $\|\cdot\| = |\cdot|_{\mathrm{w}^{1,\infty}_\varepsilon}$, $\tilde{y} = y^D$ and $f = 0$. If

$$\tfrac{1}{2}\rho_1(y^D_N) \le \tfrac{1}{2}\rho_2(z_1,z_2) \times \min(z_2 - y^D_N, y^D_N - z_1), \tag{4.37}$$

there exists a reference state $y^* \in \mathscr{A}$ satisfying (4.9) with $f = 0$. From the stability estimate (4.34), we infer that

$$|y^* - y^D_N|_{\mathrm{w}^{1,\infty}_\varepsilon} \le c_0^{-1}\|E'(y^D)\|_* \le \frac{\rho_1(y^D_N)}{\rho_2(z_1,z_2)}.$$

If the inequality in (4.37) is strict, there exists an $R > 0$ such that $\{y \in \mathscr{A} : |y - y^*|_{\mathrm{w}^{1,\infty}_\varepsilon} \le R\} \subset \mathscr{Z}_e$. Thus, for $\|f\|_* \le c_0 R =: \delta$, there exists a unique solution to (4.9) in $\mathscr{Z}_e$.

To complete the proof of Theorem 4.5 we only need to show that the numbers z_1, z_2 satisfying our assumptions exist. This, however, follows immediately from the assumption that $\rho_1(z_m) < R\rho_2(z_m - R, z_m + R)$ and that ρ_1 is continuous.

4.3.3 Coercivity of the QC approximation

In order to apply a similar technique as in §4.3.2 to prove the existence of a QC solution near an exact solution, we need to show that E'' is also coercive in $\mathscr{A}_0(\mathscr{T})$, i.e., that there exists a constant $\tilde{c}_0 > 0$ such that, for all $Y \in \mathscr{Z}_e \cap \mathscr{A}(\mathscr{T})$, we have

$$\inf_{\substack{U\in\mathscr{A}_0(\mathscr{T})\\ |U|_{\mathrm{w}^{1,\infty}_\varepsilon}=1}} \sup_{\substack{V\in\mathscr{A}_0(\mathscr{T})\\ |V|_{\mathrm{w}^{1,1}_\varepsilon}=1}} E''(Y;U,V) \geq \tilde{c}_0.$$

To this end, fix $U \in \mathscr{A}_0(\mathscr{T})$ and pick $p, q \in \{1, \dots, K\}$ such that $\overline{U}'_p$ is maximal and $\overline{U}'_q$ is minimal. Similarly as before, we also let $P = \{i : \overline{U}'_i > 0\}$ and $Q = \{i : \overline{U}'_i < 0\}$, and we define

$$\overline{V}'_k = \begin{cases} \frac{1}{2}h_p^{-1}, & \text{if } k = p, \\ -\frac{1}{2}h_q^{-1}, & \text{if } k = q, \text{ and} \\ 0, & \text{otherwise.} \end{cases}$$

This gives

$$\begin{aligned} E''(Y;U,V) &= \sum_{n=1}^{N}\sum_{m=1}^{N} \varepsilon F''_{nm}(Y) U'_n V'_m \\ &= \frac{1}{2h_p}\sum_{n=1}^{N}\sum_{m=t_{p-1}+1}^{t_p} \varepsilon F''_{nm}(Y) U'_n - \frac{1}{2h_q}\sum_{n=1}^{N}\sum_{m=t_{q-1}+1}^{t_q} \varepsilon F''_{nm}(Y) U'_n \\ &\geq \frac{\overline{U}'_p}{2h_p}\sum_{m=t_{p-1}+1}^{t_p} \varepsilon \sum_{n\in P} F''_{nm}(Y) - \frac{\overline{U}'_q}{2h_q}\sum_{m=t_{q-1}+1}^{t_q} \varepsilon \sum_{n\in Q} F''_{nm}(Y). \end{aligned}$$

Using the estimate (4.31), we obtain

$$\begin{aligned} E''(Y;U,V) &\geq \frac{\overline{U}'_p}{2h_p}\sum_{m=t_{p-1}+1}^{t_p} \varepsilon\rho_2(z_1,z_2) - \frac{\overline{U}'_q}{2h_q}\sum_{m=t_{q-1}+1}^{t_q} \varepsilon\rho_2(z_1,z_2) \\ &\geq c_0 |U|_{\mathrm{w}^{1,\infty}_\varepsilon}, \end{aligned}$$

where $c_0 = \frac{1}{2}\rho_2(z_1, z_2)$, i.e., we have the same inf-sup constant as in the case of the full test-space $\mathscr{A}_0$.

If we now replace E by $\tilde{E}$ in all the above computations, we obtain instead

$$\min_{Y\in\mathscr{A}(\mathscr{T})\cap\mathscr{Z}_e} \min_{\substack{U\in\mathscr{A}_0(\mathscr{T})\\ |U|_{\mathrm{w}^{1,\infty}_\varepsilon}=1}} \max_{\substack{V\in\mathscr{A}_0(\mathscr{T})\\ |V|_{\mathrm{w}^{1,1}_\varepsilon}=1}} \tilde{E}''(Y;U,V) \geq \tfrac{1}{2}\tilde{\rho}_2(z_1,z_2). \tag{4.38}$$

4.3.4 Proof of Theorem 4.6

Stimulated by the *a priori* error analysis in [78] and the proof of Theorem 3.3, we begin by rewriting the QC approximation as a fixed-point problem. To this end assume that $Y \in \mathscr{A}(\mathscr{T}) \cap \mathscr{Z}_e$ satisfies (4.14). Let $y \in \mathscr{A} \cap \mathscr{Z}_e$ be an exact solution and let Πy be its interpolant. We then have, for all $V \in \mathscr{A}_0(\mathscr{T})$,

$$\begin{aligned}\int_0^1 E''\big(\Pi y + \tau(Y - \Pi y); Y - \Pi y, V\big)\,\mathrm{d}\tau &= E'(Y;V) - E'(\Pi y;V) \qquad (4.39)\\ &= E'(Y;V) - \tilde{E}'(Y;V) + \langle f, V\rangle_{\mathscr{T}} - \langle f, V\rangle_{\varepsilon} + E'(y;V) - E'(\Pi y;V) =: \ell_Y(V).\end{aligned}$$

In fact, we see that Y is a solution of (4.14) if, and only if, it solves (4.39) which we rewrite as a fixed point problem. Let $\varphi \in \mathscr{A}(\mathscr{T}) \cap \mathscr{Z}_e$. We define the fixed point map $\mathcal{L}\colon \mathscr{A}(\mathscr{T}) \cap \mathscr{Z}_e \to \mathscr{A}(\mathscr{T})$, $Y_\varphi = \mathcal{L}(\varphi)$ by

$$\int_0^1 E''\big(\Pi y + \tau(\varphi - \Pi y); Y_\varphi - \Pi y, V\big)\,\mathrm{d}\tau = \ell_\varphi(V) \qquad \forall V \in \mathscr{A}_0(\mathscr{T}). \qquad (4.40)$$

By the Integral Mean Value Theorem, there exists $\theta \in \operatorname{conv}\{\varphi, \Pi y\} \subset \mathscr{Z}_e$ such that $\int_0^1 E''(\Pi y + \tau(\varphi - \Pi y))\,\mathrm{d}\tau = E''(\theta)$. Hence, if $c_0 > 0$, the map $\mathcal{L}$ is well defined and we can rewrite (4.40) as

$$E''(\theta; Y_\varphi - \Pi y, V) = \ell_\varphi(V) \qquad \forall V \in \mathscr{A}_0(\mathscr{T}).$$

Upon taking the supremum over all $V \in \mathscr{A}_0(\mathscr{T})$ with $|V|_{\mathrm{w}^{1,1}_\varepsilon} = 1$ we obtain

$$c_0 |Y_\varphi - \Pi y|_{\mathrm{w}^{1,\infty}_\varepsilon} \le \max_{\substack{V \in \mathscr{A}_0(\mathscr{T})\\ |V|_{\mathrm{w}^{1,1}_\varepsilon} = 1}} |\ell_\varphi(V)| \le c_1\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3,$$

where c_1 is a Lipschitz constant for E' in $\mathscr{Z}_e$ and $\mathcal{E}_i$, $i = 1, 2, 3$, are defined at the beginning of §4.3. Thus, in order for $\mathcal{L}$ to map $\mathscr{A}(\mathscr{T}) \cap \mathscr{Z}_e$ into itself, it is sufficient that

$$c_1\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3 \le c_0 \min_{i=1,\dots,N} \min(\Pi y_i' - z_1, z_2 - \Pi y_i').$$

Since $\Pi y_{t_k} = y_{t_k}$ for $k = 0, \dots, K$, it follows that

$$\sum_{i=t_{k-1}+1}^{t_k} \varepsilon y_i' - h_k\big(\overline{\Pi y}\big)_k' = 0,$$

and hence $\min(\Pi y_i' - z_1, z_2 - \Pi y_i') \le R$. We conclude that if (4.21) is satisfied then $\mathcal{L}$ maps $\mathscr{A}(\mathscr{T}) \cap \mathscr{Z}_e$ into itself. The Implicit Function Theorem implies that $\mathcal{L}$ is continuous. Therefore, by Brouwer's fixed point theorem, $\mathcal{L}$ has a fixed point Y in

$\mathscr{A}(\mathscr{T})\cap\mathscr{Z}_e$. From our discussion above it follows that Y is a solution to (4.14). From (4.38) we see that if $\tilde{\rho}_2(z_1,z_2)>0$ then the QC solution is unique in $\mathscr{A}(\mathscr{T})\cap\mathscr{Z}_e$. This concludes the proof of part (b) of Theorem 4.6. We are only left to prove the stated bounds on c_1, and $\mathcal{E}_i$, $i=1,2,3$.

To bound E'' in $\mathscr{Z}_e$, we compute

$$\begin{aligned}
|E''(\theta;U,V)| &= \sum_{n=1}^{N}\sum_{m=1}^{N}\varepsilon|F''_{nm}(\theta)||U'_n||V'_m| \\
&\le |U|_{\mathrm{w}_\varepsilon^{1,\infty}}\sum_{m=1}^{N}\varepsilon|V'_m|\sum_{n=1}^{N}|F''_{nm}(\theta)| \\
&\le |U|_{\mathrm{w}_\varepsilon^{1,\infty}}|V|_{\mathrm{w}_\varepsilon^{1,1}}\max_{m=1,\dots,N}\sum_{n=1}^{N}|F''_{nm}(\theta)|.
\end{aligned}$$

We can bound the sum in the last term by a computation identical to that in (4.31) except that the signs are reversed, and thus we obtain (4.20).

To bound $\mathcal{E}_1$ we simply use Theorem 4.4 with $p=\infty$. For $\mathcal{E}_2$, we use Theorem 4.4 with $p=1$ to estimate

$$\begin{aligned}
\big|\langle f,V\rangle_{\mathscr{T}}-\langle f,V\rangle_\varepsilon\big| &\le \sum_{i=1}^{N}\varepsilon\big|\Pi(fV)_i-f_iV_i\big| \\
&\le \sum_{k=1}^{K}h_k^2|\Pi(fV)|_{\mathrm{w}_\varepsilon^{2,1}(\!(t_{k-1},t_k)\!)}.
\end{aligned}$$

For $i=t_{k-1}+1,\dots,t_k-1$, using the fact that $V_i''=0$, we have

$$\begin{aligned}
(fV)_i'' &= \varepsilon^{-2}(f_{i+1}V_{i+1}-2f_iV_i+f_{i-1}V_{i-1}) \\
&= \frac{f_{i+1}-2f_i+f_{i-1}}{\varepsilon^2}V_i+\frac{f_{i+1}-f_i}{\varepsilon}\frac{V_{i+1}-V_i}{\varepsilon}+\frac{f_i-f_{i-1}}{\varepsilon}\frac{V_i-V_{i-1}}{\varepsilon}.
\end{aligned}$$

Thus, using the discrete Friedrichs inequality (4.2), we obtain

$$\begin{aligned}
\big|\langle f,V\rangle_{\mathscr{T}}-\langle f,V\rangle_\varepsilon\big| &\le \sum_{k=1}^{K}h_k^2\Big[|f|_{\mathrm{w}_\varepsilon^{2,\infty}(\!(t_{k-1},t_k)\!)}\|V\|_{\ell^1_\varepsilon(\!(t_{k-1}+1,t_k-1)\!)} \\
&\qquad +(|f|_{\mathrm{w}_\varepsilon^{1,\infty}(\!(t_{k-1}+1,t_k)\!)}+|f|_{\mathrm{w}_\varepsilon^{1,\infty}(\!(t_{k-1},t_k-1)\!)})|V|_{\mathrm{w}_\varepsilon^{1,1}(\!(t_{k-1},t_k)\!)}\Big] \\
&\le \max_{k=1,\dots,K}h_k^2\max\big(|f|_{\mathrm{w}_\varepsilon^{2,\infty}(\!(t_{k-1},t_k)\!)},2|f|_{\mathrm{w}_\varepsilon^{1,\infty}(\!(t_{k-1}+1,t_k)\!)} \\
&\qquad +2|f|_{\mathrm{w}_\varepsilon^{1,\infty}(\!(t_{k-1},t_k-1)\!)}\big)\big(\|V\|_{\ell^1_\varepsilon}+\tfrac{1}{2}|V|_{\mathrm{w}_\varepsilon^{1,1}}\big).
\end{aligned}$$

We apply (4.2) to estimate $\|V\|_{\ell^1_\varepsilon}\le\frac{1}{2}|V|_{\mathrm{w}_\varepsilon^{1,1}}$ and thus prove the bound (4.23).

Finally, using (4.26), the bound (4.24) on $\mathcal{E}_3$ follows from

$$\begin{aligned} |E'(\theta;V) - \tilde{E}'(\theta;V)| &\leq \sum_{n=1}^{N} \varepsilon |F_n'(\theta) - \tilde{F}_n'(\theta)||V_n'| \\ &\leq \max_{n=1,\dots N} |F_n'(\theta) - \tilde{F}_n'(\theta)||V|_{\mathrm{w}^{1,1}_\varepsilon}, \end{aligned}$$

and a computation that is identical to the one leading to (4.37).

4.4 Fracture

We now look at a class of solutions of the atomistic model (4.9) with a single defect — a fracture. We fix an index $\xi \in \{1,\dots,N\}$ and consider deformations $y \in \mathscr{A}$ such that $y'_\xi \gg z_t$ while $z_1 \leq y'_i \leq z_2 < z_t$ for $i \neq \xi$. The *fracture* is the broken interaction between the two atoms at y_ξ and $y_{\xi-1}$. Elastic states and fractured states with a single crack are the only stable steady states in one dimension. If at least two gradients y'_i, y'_j are greater than or equal to z_t, it can be easily seen that $E''(y)$ has at least one negative eigenvalue (cf. §5.2.2).

However, even with a single fracture, it should be apparent from the analysis of §4.3.1 that we cannot expect (4.32) to hold when $|u|_{\mathrm{w}^{1,\infty}_\varepsilon} = |u'_\xi|$ since $J''(u'_\xi) \approx 0$. We therefore change the norm in which we analyze the error into the norm $|\cdot|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}}$ defined by

$$|u|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}} = \max_{\substack{i=1,\dots,N \\ i\neq\xi}} |u'_i|.$$

Since we have imposed a Dirichlet condition at both endpoints, $|\cdot|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}}$ is indeed a norm on $\mathscr{A}_0$. We use $B_f(y,R)$ to denote the balls, centre y and radius R, with respect to the $|\cdot|_{\mathrm{w}^{1,\infty}_\varepsilon}$-semi-norm. As was hinted above, we define

$$\mathscr{Z}_f = \{y \in \mathbb{R}^{N+1} : y'_\xi \geq z_f \text{ and } z_1 \leq y'_i \leq z_2 \text{ for } i = 1,\dots,N, i \neq \xi\},$$

where the constants z_i satisfy $z_1 < z_m < z_2 < z_t$, and z_f is sufficiently large (which we will make precise).

In order to simplify the proofs of coercivity we assume that

$$J'''(z) \geq 0 \qquad \text{for } z \geq z_f. \tag{4.41}$$

This typically imposes a negligible lower bound on z_f. We shall also need a further measure of stability,

$$\rho_{2,f}(z_f, z_1) = \sum_{r=0}^{\infty} (r+1) J''(z_f + r z_1).$$

The definition of $\rho_{2,f}$ does not involve z_2 because we have assumed (4.41). The function $\tilde{\rho}_{2,f}$ corresponding to the cut-off potential $\tilde{J}$ is defined analogously. In order to be able to neglect the effect of long-range interactions across the crack, we assume that

$$\forall\, a > 0\ \forall\, z_1 \geq z_t/2\ \exists\, z_D = z_D(a, z_1):\ N\rho_{2,f}\big(N(z_D - z_t), z_1\big) \geq -a. \tag{4.42}$$

This would typically involve a growth condition for J'', for example, $|J''(z)| \lesssim z^{-k}$, for some $k > 3$ and z sufficiently large.

Theorem 4.8 *Let J satisfy the assumptions of §4.2.1 as well as conditions (4.41) and (4.42). Assume also that there exists $R \in (0, \min(z_m - z_t/2, z_t - z_m))$ such that $2\rho_1(z_m) < R\rho_2(z_m - R, z_m + R)$. Then the following hold:*

(a) *Coercivity: There exist $z_1 < z_m < z_2 < z_t$ independent of ε, and $z_f = O(\varepsilon^{-1})$ such that*

$$\inf_{y\in\mathscr{Z}_f} \inf_{\substack{u\in\mathscr{A}_0\\ |u|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}}=1}} \sup_{\substack{v\in\mathscr{A}_0\\ |v|_{\mathrm{w}^{1,1}_{\varepsilon}}=1}} E''(y;u,v) \geq \tfrac{1}{2}\big(\rho_2(z_1,z_2) + 2N\rho_{2,f}(z_f,z_1)\big) =: c_0 > 0, \tag{4.43}$$

where $\mathscr{Z}_f$ is defined as above.

(b) *Existence: There exist $\delta_1, \delta_2 > 0$, independent of ε, such that for every $y_N^D \in \mathbb{R}$ with $y_N^D \geq z_m + \delta_1$ and for every $f \in \mathbb{R}^{N+1}$ with $\|f\|_* \leq \delta_2$, there exists a solution y_f of (4.9) in $\mathscr{Z}_f$.*

(c) *Stability: Let y_f, y_g be solutions to (4.9) in $\mathscr{Z}_f \cap \mathscr{A}$ with respective right-hand sides f and g; then*

$$|y_f - y_g|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}} \leq c_0^{-1}\|f - g\|_*.$$

For the QC error bounds, let $\mathcal{E}_1 = |y - \Pi y|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}}$ and let $\mathcal{E}_2$ and $\mathcal{E}_3$ be defined as in §4.3.

Theorem 4.9 *Let J satisfy the conditions of §4.2.1 as well as (4.41) and (4.42), and let $\mathscr{Z}_f$ be defined as above. Furthermore, assume that $\{\xi - 1, \xi\} \subset \mathscr{T}$.*

(a) *We have the coercivity and continuity estimates*

$$\inf_{\theta\in\mathscr{Z}_f} \min_{\substack{U\in\mathscr{A}_0(\mathscr{T})\\ |U|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}}=1}} \max_{\substack{V\in\mathscr{A}_0(\mathscr{T})\\ |V|_{\mathrm{w}^{1,1}_{\varepsilon}}=1}} E''(\theta;U,V) \geq \tfrac{1}{2}\big(\rho_2(z_1,z_2) + 2N\rho_{2,f}(z_f,z_1)\big) =: c_0, \quad \text{and} \tag{4.44}$$

$$\max_{Y\in S^1(\mathscr{T})\cap\mathscr{Z}_f} \max_{\substack{U\in\mathscr{A}_0(\mathscr{T})\\ |U|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}}=1}} \max_{\substack{V\in\mathscr{A}_0(\mathscr{T})\\ |V|_{\mathrm{w}^{1,1}_{\varepsilon}}=1}} E''(Y;U,V) \leq \rho_3(z_1,z_2) =: c_1. \tag{4.45}$$

(b) *Suppose that $z_f > z_t$ is sufficiently large so that $c_0 > 0$ (cf. (4.42)). Let $y \in \mathscr{Z}_f \cap \mathscr{A}$ be a solution of (4.9) and define $R = \min_{i\neq\xi} \min(z_2 - y'_i, y'_i - z_1)$. Assume furthermore that the QC mesh $\mathscr{T}$ is sufficiently fine so that*

$$c_1\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3 \leq c_0 \min\big(R, \varepsilon(y'_\xi - z_f)\big). \tag{4.46}$$

Then, there exists a solution $Y \in \mathscr{A}(\mathscr{T}) \cap \mathscr{Z}_f$ of the QC method (4.14) which satisfies

$$|y - Y|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}} \leq c_0^{-1}\Big((c_0 + c_1)\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3\Big).$$

If $\tilde{\rho}_2(z_1, z_2) + 2N\tilde{\rho}_{2,f}(z_f, z_1) > 0$ then the QC solution is unique in $\mathscr{A}(\mathscr{T}) \cap \mathscr{Z}_f$.

(c) *The error quantities $\mathcal{E}_1$ and $\mathcal{E}_2$ satisfy the same bounds as in Theorem 4.6, while $\mathcal{E}_3$ is now bounded by*

$$\begin{aligned} \mathcal{E}_3 \;\leq\; \sum_{r=1}^{\infty} r \max\Big[&\max_{z_1\leq z\leq z_2} \big|\tilde{J}'(rz) - J'(rz)\big|, \\ &\max_{z_1\leq z\leq z_2} \big|\tilde{J}'(z_f + (r-1)z) - J'(z_f + (r-1)z)\big|\Big]. \end{aligned} \tag{4.47}$$

As we remark in §4.4.4, the condition (4.46) is not overly restrictive. We may think, for example that $z_f = O(\varepsilon^{-1})$ and $y'_\xi \geq 2z_f$. In that case, the upper bound required on the error terms is independent of ε.

4.4.1 Coercivity of the atomistic problem

For the proof of coercivity in the case of fracture we make use of the fact that the fracture problem can, to some extent, be seen as a combination of two Neumann problems. Fix $y \in \mathscr{Z}_f$ and $u \in \mathscr{A}_0$. Upon multiplying u by (-1), we may assume without loss of generality that $u'_p = |u|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}}$. Let $P = \{i : u'_i > 0\}$ and $Q = \{j : u'_j < 0\}$ and define

$$v'_n = \begin{cases} \frac{1}{2}\varepsilon^{-1}, & \text{if } n = p, \\ -\frac{1}{2}\varepsilon^{-1}, & \text{if } n = \xi, \\ 0, & \text{otherwise.} \end{cases}$$

In that case,

$$\begin{aligned} E''(y; u, v) &= \sum_{n=1}^{N}\sum_{m=1}^{N} \varepsilon F''_{nm}(y) u'_n v'_m \\ &= \sum_{n=1}^{N} \varepsilon u'_n \Big[F''_{np}(y)\frac{1}{2\varepsilon} - F''_{n\xi}(y)\frac{1}{2\varepsilon}\Big] \\ &\geq \frac{1}{2}\sum_{n\in P} u'_n F''_{np}(y) - \frac{1}{2}\sum_{n\in Q} u'_n F''_{n\xi}(y). \end{aligned}$$

If we divide the sum over $n \in P$ into those indices which lie on the same side of the fracture as p and the rest, we can estimate $F''_{np} \geq F''_{n\xi}$ for those n which lie on the opposite side of the fracture from p (cf. condition (4.41)). If we assume, without loss of generality, that $p < \xi$, we obtain

$$E''(y;u,v) \geq \frac{1}{2}\sum_{n<\xi} |u'_n| F''_{np}(y) + \sum_{n\neq\xi} |u'_n| F''_{n\xi}(y) + |u'_\xi| F''_{\xi\xi}(y).$$

Since $u \in \mathscr{A}_0$, we have $|u'_\xi| \leq (N-1)|u|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}}$ and hence, we obtain

$$E''(y;u,v) \geq \frac{1}{2}|u|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}} \Big[\sum_{n<\xi} F''_{np}(y) + \sum_{n\neq\xi} F''_{n\xi}(y) + (N-1)F''_{\xi\xi}(y) \Big], \tag{4.48}$$

For the first sum in (4.48) we can use the same procedure as in the elastic case, i.e,

$$\sum_{n<\xi} F''_{np}(y) \geq \rho_2(z_1, z_2),$$

while the second sum as well as $F''_{\xi\xi}$ should be practically zero. In this regime the forces should be so weak that we can make fairly crude estimates. Using assumption (4.41) we have $F''_{n\xi} \geq F''_{\xi\xi}$ for all n and hence only need to estimate $F''_{\xi\xi}$,

$$\begin{aligned} F''_{\xi\xi}(y) &= \sum_{i=\xi}^{N}\sum_{j=1}^{\xi} J''\big(\varepsilon^{-1}(y_i - y_{j-1})\big) \geq \sum_{i=\xi}^{N}\sum_{j=1}^{\xi} J''\big(z_f + (i-j)z_1\big) \\ &\geq \sum_{j=1}^{\xi}\sum_{r=\xi-j}^{\infty} J''(z_f + rz_1) = \sum_{r=0}^{\infty}\sum_{j=\xi-r}^{\xi} J''(z_f + rz_1) \\ &= \sum_{r=0}^{\infty}(r+1)J''(z_f + rz_1) = \rho_{2,f}(z_f, z_1). \end{aligned}$$

Putting everything together, we obtain

$$E''(y;u,v) \geq \tfrac{1}{2}\big(\rho_2(z_1,z_2) + 2(N-1)\rho_{2,f}(z_f,z_1)\big)|u|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}}. \tag{4.49}$$

4.4.2 Proof of Theorem 4.8

First, let us finalize the question of coercivity. To this end, let $c'_0 = \frac{1}{2}\rho_2(z_1,z_2)$, which we assume to be positive, and choose

$$z_f = N\big(z_D(\alpha c'_0, z_1) - z_t\big),$$

where $\alpha \in (0,1)$ is a ratio that we shall determine in a moment. In that case, (4.43) holds with $c_0 = (1-\alpha)c'_0$.

In order to use Lemma 4.7 as in §4.3.2, we need to characterize the balls with respect to the $|\cdot|_{\mathrm{w}^{1,\infty}_{\varepsilon}}$-semi-norm. This is achieved in the following lemma.

Lemma 4.10 *Suppose that* $y_N^D \geq z_D(\alpha c_0', z_1)$. *Then*

$$B(\tilde{y}, R) \subset \mathscr{Z}_f \qquad \forall \tilde{y} \in \mathscr{Z}_f, \quad \forall R \leq \min_{n \neq \xi} \min(z_2 - \tilde{y}_n', \tilde{y}_n' - z_1). \tag{4.50}$$

Proof. If $\tilde{y} \in \mathscr{Z}_f$ and $y \in B_f(\tilde{y}, R)$ then

$$\varepsilon y_\xi' = y_N^D - \sum_{i \neq \xi} \varepsilon y_\xi' \geq z_D - z_2 \geq z_D - z_t = \varepsilon z_f. \quad \square$$

Thanks to Lemma 4.10, the coercivity estimate (4.43) holds for all $y \in B(\tilde{y}, R)$ which makes it possible to use Lemma 4.7.

As in §4.3.2 we use Lemma 4.7 to construct a reference state. Let y^D be a preliminary reference state defined as follows,

$$(y_i^D)' = \begin{cases} i\varepsilon z_m, & \text{if } i < \xi \\ y_N^D - z_m(1 - i\varepsilon), & \text{if } i \geq \xi. \end{cases}$$

As in the elastic case, we estimate the residual of y^D. Fix $n \in \mathbb{N}$ and assume, without loss of generality, that $n < \xi$. Since $z_f \geq z_t$ and $J''(z) < 0$ for $z > z_t$ it follows that J' is decreasing in that domain. In particular, we have $|J'(z_f + z)| \leq |J'(z_1 + z)|$ whenever $z \geq z_1$. Using this fact, and otherwise closely following the computations in §4.3.2, we have

$$|E_n'(y^D)| \leq \sum_{r=n\wedge(\xi-n)+1}^{\infty} \big|J'(rz_m)\big|.$$

Summing over $n < \xi$, we obtain

$$\sum_{n<\xi} |E_n'(y^D)| \leq \frac{1}{2} \sum_{r=2}^{\infty} (r-1)|J'(rz_m)|.$$

We now add the terms with $n \geq \xi$ which gives

$$\|E'(y^D)\|_* \leq \rho_1(z_m). \tag{4.51}$$

Setting $\Phi = E$, $\|\cdot\| = |\cdot|_{\mathrm{w}_\varepsilon^{1,\infty}}$, $\tilde{y} = y^D$, $\tilde{f} = E'(\tilde{y})$ and $f = 0$ in Lemma 4.7 we can deduce the existence of $y^* \in \mathscr{Z}_f$, satisfying $E'(y^*) = 0$. We note that

$$|y_i^{*\prime} - z_m| \leq c_0^{-1} \|E'(y^D)\|_* \leq \frac{2\rho_1(z_m)}{(1-\alpha)\rho_2(z_1, z_2)}, \qquad i \neq \xi. \tag{4.52}$$

If the conditions of Theorem 4.8 are satisfied, then there exists $\alpha > 0$, independent of ε, such that

$$2\frac{\rho_1(z_m)}{(1-\alpha)\rho_2(z_1, z_2)} < R,$$

which implies that $y^* \in \mathrm{int}(\mathscr{Z}_f \cap \mathscr{A})$. All results of Theorem 4.8 now follow from another application of Lemma (4.7) setting $\Phi = E$, $\|\cdot\| = |\cdot|_{\mathrm{w}_\varepsilon^{1,\infty}}$, $\tilde{y} = y^*$ and $\tilde{f} = 0$. In particular, it is sufficient to assume that $y_N^D \geq z_D(\alpha c_0', z_1)$.

4.4.3 Coercivity of the QC approximation

First of all, we note that the assumption of Theorem 4.9 allows us to assume that $\{\xi-1,\xi\}\subset\mathscr{T}$. This is in fact a necessary condition to make an approximation of a fracture in $\mathrm{w}^{1,\infty}_{\varepsilon,f}$ possible.

Let $Y\in\mathscr{Z}_f$ and $U\in\mathscr{A}_0(\mathscr{T})$. Following §4.4.1 and §4.3.3 we assume that $\overline{U}'_p=|U|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}}$ and define the test function V by

$$\overline{V}'_k=\begin{cases}\frac{1}{2}h_p^{-1}, & \text{if } k=p\\ -\frac{1}{2}\varepsilon^{-1}, & \text{if } k=\xi\\ 0, & \text{otherwise.}\end{cases}$$

Then, assuming again without loss of generality that $t_p<\xi$, and using (4.41), we have

$$\begin{aligned}
E''(Y;U,V) &= \sum_{n=1}^{N}\sum_{m=1}^{N}\varepsilon F''_{nm}(Y)U'_nV'_m\\
&= \frac{\varepsilon}{2h_p}\sum_{n=1}^{N}\sum_{m=t_{p-1}+1}^{t_p}F''_{nm}(Y)U'_n-\frac{1}{2}\sum_{n=1}^{N}F''_{n\xi}(Y)U'_n\\
&\geq \frac{\varepsilon}{2h_p}\sum_{m=t_{p-1}+1}^{t_p}\Big[\sum_{n<\xi}F''_{nm}(Y)U'_n+\sum_{n\geq\xi,n\in P}F''_{n\xi}(Y)U'_n\Big]-\sum_{n\in Q}F''_{n\xi}(Y)U'_n\\
&\geq \frac{\varepsilon}{2h_p}\sum_{m=t_{p-1}+1}^{t_p}\sum_{n<\xi}F''_{nm}(Y)U'_n-\frac{1}{2}\sum_{n\neq\xi}F''_{n\xi}(Y)|U'_n|-\frac{1}{2}|U'_\xi|F''_{\xi\xi}(Y).
\end{aligned}$$

We estimate the first term as in §4.3.3 and the second and third term as in §4.4.1 which gives

$$E''(Y;U,V)\geq\frac{1}{2}|U|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}}\big(\rho_2(z_1,z_2)+2N\rho_{2,f}(z_f,z_1)\big),$$

and thus (4.44). If E is replaced by $\tilde{E}$, we have instead

$$\tilde{E}''(Y;U,V)\geq\frac{1}{2}|U|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}}\big(\tilde{\rho}_2(z_1,z_2)+2N\tilde{\rho}_{2,f}(z_f,z_1)\big).\tag{4.53}$$

4.4.4 Proof of Theorem 4.9

To prove the QC error estimate we can repeat the fixed point argument of §4.3.4 almost verbatim. Only two modifications need to be made. First, as in the existence proof of §4.4.2 we need to show that a solution of the linearized problem appearing in the fixed point argument lies in $\mathscr{Z}_f$. This can be done by the same argument as in the proof of Lemma 4.10, if we choose y_N^D sufficiently large. This method was suitable for the existence theorem where we needed to construct a reference solution. Now, however,

the reference solution is given by the exact solution y which allows us to follow a more general approach.

As in §4.3.4 let $Y_\varphi = \mathcal{L}(\varphi)$, then,

$$\begin{aligned} \varepsilon(Y_\varphi)'_\xi &= y_N^D - \sum_{i\neq\xi} \varepsilon(Y_\varphi)'_i \\ &= \sum_{i=1}^{N} \varepsilon \Pi y'_i - \sum_{i\neq\xi} \varepsilon(Y_\varphi)'_i \\ &\geq \varepsilon y'_\xi - |\Pi y - Y_\varphi|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}}. \end{aligned}$$

Hence, in order to guarantee $Y_\varphi \in \mathscr{Z}_f$, we require

$$y'_\xi \geq z_f + N|\Pi y - Y_\varphi|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}}.$$

This may seem an insurmountable requirement at first but remember that y'_ξ is typically of order N. For $|\Pi y - Y_\varphi|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}}$ we have the estimate

$$|Y_\varphi - \Pi y|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}} \leq c_0^{-1}\big(c_1\mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3\big).$$

Hence, if (4.46) holds, then we can deduce the existence of a QC solution in the set $\mathscr{Z}_f$.

Our second modification of the proof of §4.3.4 is to compute a new bound for $\mathcal{E}_3$. We use (4.26) again to estimate

$$\begin{aligned} |E'(\theta;V) - \tilde{E}'(\theta;V)| &= \sum_{n=1}^{N} \varepsilon |F'_n(\theta) - \tilde{F}'_n(\theta)\Big||V'_n| \\ &\leq |V|_{\mathrm{w}^{1,1}_{\varepsilon}} \max_{n=1,\ldots,N} |F'_n(\theta) - \tilde{F}'_n(\theta)\Big|. \end{aligned}$$

For each n, we have

$$|F'_n(\theta) - \tilde{F}'_n(\theta)| \leq \sum_{i=1}^{n}\sum_{j=n}^{N} \big|J'\big(\varepsilon^{-1}(\theta_i - \theta_{j-1})\big) - \tilde{J}'\big(\varepsilon^{-1}(\theta_i - \theta_{j-1})\big)\big|.$$

As in the elastic case, we can estimate and rearrange this sum to obtain (4.47).

4.5 Computation of Coercivity Regions

In this short appendix, we confirm that the hypotheses made on the interaction potential can indeed be satisfied. With the use of simple MATLAB scripts it is straightforward to compute possible values for z_1, z_2 and, in the fracture case, for z_f. Only the elastic

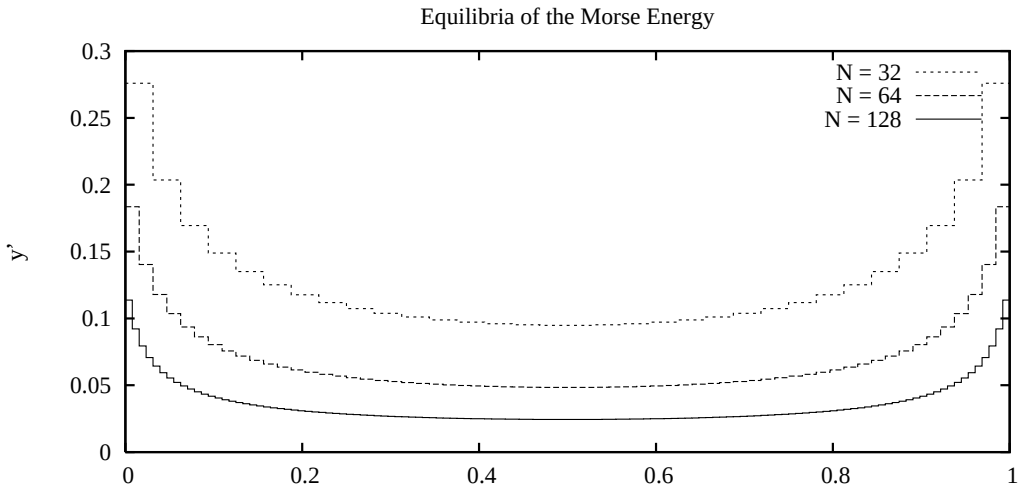

Figure 4.1: Equilibria of the atomistic energy (4.6), where J is the Morse potential (1.8) with $\alpha = 1.0$.

case is included here since the additional requirements of the fracture case are very easily met given the fast decay of most interaction potentials.

We choose the constants in the Lennard–Jones potential so that its minimum lies at $z = 1$,

$$J(z) = z^{-12} - 2z^{-6}.$$

Hence, we have $z_m = 1$ and $z_t = (13/7)^{1/6} \approx 1.11$. If we choose $z_1 = 0.88$ and $z_2 = 1.06$, we obtain $\rho_2(z_1, z_2) \approx 12.5$. Furthermore, we have $\rho_1(z_m) \approx 0.2$ which guarantees the existence of a reference state for sufficiently small boundary displacements.

The Morse potential is slightly less forthcoming in this respect. First, we note that $z_m = 1$ and $z_t = 1+\alpha^{-1}\log(2)$. If we choose $\alpha = 1$ in (1.8) we obtain $\rho_2(z_m, z_m) \approx -3.8$ and we have therefore no hope of constructing an equilibrium with the technique we have used. This does not mean that E has no equilibrium in this case. In fact, the mere existence of a global energy minimum can be easily deduced by a compactness argument. However, numerical experiments shown in Figure 4.1 indicate that those equilibria are extremely unstable and bear no resemblance to the observed equilibria of metallic materials. Furthermore, there seems to be no convergence of those equilibria to a continuum as $N \to \infty$.

If we make the well steeper, however, we can achieve coercivity. Already for $\alpha = 4$, we can choose $z_1 = 0.9$ and $z_2 = 1.08$ to obtain $\rho_2(z_1, z_2) \approx 5.38$. Since $\rho_1(z_m) \approx 0.3$ is follows that $0.8 \times \rho_2(z_1, z_2) > \rho_1(z_m)$ and hence there exists a reference state for sufficiently small boundary displacements.

Finally, we should note that the steeper the basin of convexity around z_m (the larger α in the Morse potential case) the better the bounds become.

Concluding Remarks

Clearly, the relative completeness of our results was primarily due to the one-dimensional setting that we have chosen. However, the fundamental approach to error estimation for the QC method, namely a fixed point argument based on the $\mathrm{w}^{1,\infty}_{\varepsilon}$-norm (or its modifications), can be employed in higher dimensions as well. In fact, it can be seen that, if we assume coercivity directly, rather than proving it as we have done here, then Theorem 4.6 and 4.9 carry over immediately. However, the main difference between ellipticity (cf. Lin [60]) and the inf-sup condition which we have employed is that the inf-sup condition does not automatically translate to subspaces. Even worse, it may be expected that the inf-sup constant in two and three dimensions is not independent of ε. A possible alternative that will be investigated is to use the techniques in [82] where $\mathrm{W}^{1,\infty}$-convergence of a finite element method is proved without using an inf-sup condition. A further option would be to prove an inf-sup condition for a $\mathrm{W}^{1,p}$-like topology, where $p > d$ and to use inverse estimates to obtain a convergence rate of order $h^{1-d/p}$ in the $\mathrm{W}^{1,\infty}$-norm. Thus, for sufficiently small h (but independent of ε) we could again conclude that a QC solution exists in a neighbourhood of an exact solution.

Another fact worth noting is that the $\mathrm{w}^{1,\infty}_{\varepsilon}$-norm employed in the elastic analysis is actually equivalent (in the sense that the constants are independent of ε) to the *energy norm* $u \mapsto \|E''(\theta)u\|_*$, for any $\theta \in \mathscr{Z}_e$. Similarly, the $\mathrm{w}^{1,\infty}_{\varepsilon,f}$-norm from the analysis in §4.4 is equivalent to $u \mapsto \|E''(\theta)u\|_*$, for any $\theta \in \mathscr{Z}_f$. This unifies, to some extend, the analysis of stable critical points.

Chapter 5

A Posteriori Analysis and Adaptive Algorithms for the Quasicontinuum Method in One Dimension

One of the most remarkable features of atomistic models is the large number of metastable states. Already in one dimension, it is fairly straightforward to see for many problems that the number of stable equilibria of the energy is at least as large as the number of atoms in the body. Therefore, error estimates must necessarily be restricted to local results. Due to the possibility of fracture, stability of solutions can only be obtained with respect to the $\mathrm{w}_\varepsilon^{1,\infty}$-norm which was introduced in §4.1, or a similar topology. As a consequence of this lack of global monotonicity and stability, the *a priori* error estimates in Chapter 4 are, except possibly in the case of elastic deformation, of purely theoretical value. For example, when an exact solution we wish to approximate is a fractured state then we can prove under come natural conditions that there exists a nearby QC solution; however, we should not expect to find it numerically. If only one atom lies on the wrong side of the crack then the error in $\mathrm{w}_\varepsilon^{1,\infty}$-norm cannot 'converge' to zero, even if every atom in the body is a repatom.

To the best of the author's knowledge, no work has been carried out so far on the *a posteriori* error analysis of the QC method. Only some experimental results, using a gradient-averaging technique for the computation of error indicators, were published in [55]. The goal of the present chapter is to present a rigorous approach to *a posteriori* error estimation for the QC method, using the strategy presented in Chapter 3. We derive estimates on the residual of the QC solution (cf. Theorem 5.2 and §5.2.1) and on an inf-sup constant which measures its stability (cf. Theorem 5.3 and §5.2.3) and show in Theorem 5.1 that, if certain natural conditions are satisfied, there exists an exact solution of the atomistic model for which an *a posteriori* error estimate holds.

We then apply this idea to the development of an adaptive optimization algorithm based on minimizing movements and proximal point methods. We prove for this algorithm that it is possible to choose the parameters in such a way that at each step of the optimization there exists an exact solution of a related problem whose distance to the numerical solution is less than a given tolerance.

A crucial ingredient in almost all *a posteriori* error analysis is a sound knowledge of the (local) stability of the equations. In Chapter 4 we have derived a number of such results for one-dimensional atomistic models, which we will make heavy use of.

Atomistic model problem The model problem will be the same as the one presented in §4.2.1, except that we shall now assume again that the exact potential J has a cut-off radius z_c. In this case, the atomistic energy E can be computed exactly and does not have to be approximated by the cut-off energy $\tilde{E}$.

We assume that there exist $z_0 \in [-\infty, +\infty)$, $z_m, z_t, z_c \in \mathbb{R}$ such that $z_0 < z_m < z_t < z_c$, and

$$\begin{aligned} &J \in \mathrm{C}^3(z_0, \infty), \quad J'(z_m) = 0, \quad J''(z_t) = 0, \quad J(z) = 0 \quad \forall z \geq z_c, \\ &J(z) \to +\infty \text{ as } z \to z_0+, \qquad J(z) = +\infty \quad \forall z \leq z_0, \\ &J''(z) \geq 0 \quad \forall z \in (z_0, z_t] \quad \text{and} \quad J''(z) \leq 0 \quad \forall z \in [z_t, \infty). \end{aligned} \tag{5.1}$$

Using these slightly modified assumptions on the atomistic interactions, the atomistic model problem is again to *find $y \in \mathscr{A}$ such that*

$$E'(y; v) = \langle f, v \rangle_\varepsilon \qquad \forall v \in \mathscr{A}_0. \tag{5.2}$$

The residual of any deformation y is the linear functional $v \mapsto E'(y; v) - \langle f, v \rangle_\varepsilon$. Since we shall analyse the error in the $\mathrm{w}^{1,\infty}_\varepsilon$-norm, we shall measure the residual in the corresponding dual norm defined in §4.2.1 for each linear functional $\ell \colon \mathscr{A}_0 \to \mathbb{R}$ by

$$\|\ell\|_* = \max_{\substack{v \in \mathscr{A}_0 \\ |v|_{\mathrm{w}^{1,1}_\varepsilon} = 1}} |\ell(v)|.$$

For the definition of the QC mesh $\mathscr{T}$ and the QC spaces $S^1(\mathscr{T})$, $\mathscr{A}(\mathscr{T})$ and $\mathscr{A}_0(\mathscr{T})$ we use the same notation as in §4.2.2. The QC approximation to (5.2) is then to *find $Y \in \mathscr{A}$ satisfying*

$$E'(Y; V) = \langle f, V \rangle_{\mathscr{T}} \qquad \forall V \in \mathscr{A}_0. \tag{5.3}$$

5.1 Adaptive Optimization Algorithm

5.1.1 Proximal point methods

As long as the deformation is purely elastic, it is easy to find the critical points of the QC functional, either directly by a Newton method or, if necessary, using a continuation principle.

However, in order to find critical points with defects, we need to supplement the QC method with an optimization algorithm. The choice of optimization methods for unconstrained optimization available is quite overwhelming. However, most algorithms are based on one of two principles and are accordingly divided into linesearch methods and trust region methods. As opposed to naive steepest descent or Newton methods, linesearch and trust region methods can be formulated in such a way that they are globally convergent. Let ϕ be the functional to be minimized and let x_ℓ denote the iterates generated by the respective optimization method.

In a linesearch method, after the ℓth iterate has been obtained, a search direction p_ℓ is computed, usually involving information about the gradient $\phi'(x_\ell)$ and the Hessian $\phi''(x_\ell)$. This must be done in such a way that p_k is a descent direction for ϕ, i.e., $\phi(x_\ell + \alpha p_\ell)$ is decreasing for $0 \leq \alpha$ sufficiently small. For the $(\ell+1)$th step, a step length α_ℓ which gives sufficient energy decrease is computed to obtain $x_{\ell+1} = x_\ell + \alpha_\ell p_\ell$.

In a typical trust region method, at the ℓth iterate, a quadratic model such as $m_k(x) = \phi(x_k) + \phi'(x_k; x - x_k) + \phi''(x_k; x - x_k, x - x_k)$ of the objective function ϕ is built. In addition, a trust region radius Δ_k is defined which measures the quality of the model m_k. The $(k+1)$th step then requires to (approximately) solve the problem

$$x_{\ell+1} \in \operatorname*{argmin}_{\|x - x_\ell\| \leq \Delta_k} m_k(x),$$

where $\|\cdot\|$ is a suitably chosen norm. The trust region radius is dynamically adjusted during the optimization process.

Our choice fell on a class, related to trust region methods, called *proximal point algorithms* (PPA) which was originally formulated as a method to compute roots of monotone operators [84]. The ℓth step of the proximal point method for finding a root of a maximal monotone operator $T : \mathscr{H} \to 2^{\mathscr{H}}$ defined on the Hilbert space $\mathscr{H}$, starting with $x_0 \in \mathscr{H}$, is to solve the inclusion problem

$$0 \in T(x) + \gamma_\ell (x - x_{\ell-1}), \tag{5.4}$$

where (γ_ℓ) is a bounded sequence of positive real numbers. It was shown in [84] that the sequence (x_ℓ) is well defined (in the sense that each x_ℓ exists and is unique), converges

weakly to a root of T if one exists, and is unbounded otherwise. While not often used in practise, the PPA is an important theoretical tool for the analysis of optimization methods [86].

If T is the (sub-)gradient of a convex functional ϕ then (5.4) is equivalent to finding

$$x_\ell \in \operatorname{argmin}\Big[\phi(x) + \frac{\gamma_\ell}{2}\|x - x_{\ell-1}\|^2\Big], \tag{5.5}$$

where $\|\cdot\|$ is the norm in $\mathscr{H}$, and x_ℓ converges to a minimizer of ϕ. In this new form, the relationship to minimizing movements [5] and the gradient flow approach of Chapter 2 becomes immediately apparent. This is in fact where the intuition to use a PPA for the QC method originates.

Namely, if ϕ is not convex then the theory for PPAs does not apply. However, motivated by the analysis in Chapter 2, if $\phi = E$, we can choose $\|\cdot\| = |\cdot|_{\mathrm{w}^{1,2}_\varepsilon}$ which, at least in the case of nearest-neighbour interactions, convexifies each step of the PPA as we have seen in Lemma 2.10. In §5.2.2, this analysis is extended to long-range interactions and it is shown in §5.3.2 that, for sufficiently large γ_ℓ, the ℓth step of the PPA is well-defined in this case as well. We shall demonstrate, furthermore, that a subsequence of the family generated by the PPA converges to a critical point. The PPA is formulated in such a way that it terminates in a finite number of iterations (choosing $\gamma_\ell = 0$ for sufficiently large ℓ) if the Hessian at this point is positive definite.

Admittedly, the choice of optimization method is guided to a large extent by mathematical convenience. For a trust region method, each iteration requires the solution of a variational inequality which would make the analysis far more involved. It is conceivable, however, that some of the results in this chapter can also be shown for trust region methods, taking the $\mathrm{w}^{1,\infty}_\varepsilon$-semi-norm as the trust region norm. For most practical linesearch methods on the other hand, it is not easy to find a relationship to measure the error of the current iterate. Nevertheless, the numerical experiments of §5.3 indicate that our PPA is quite competitive for the QC method.

An analysis for a similar method, which is essentially a linearly implicit Euler method for the system $\dot{u} = -\phi'(u)$, and a discussion of the relationship to trust region and line search methods is given in [53].

5.1.2 An adaptive PPA for the QC method

Let $Y^{(0)}$ be an *initial guess* for the PPA. This is usually provided from the previous step of a quasistatic process. The ℓth step of the PPA is to *find a critical point* $Y^{(\ell)}$ *of the functional*

$$Y \mapsto \tilde{\Phi}_\ell(Y) = \frac{\gamma_\ell}{2}\|Y - Y^{(\ell-1)}\|^2 + E(Y) - \langle Y, f\rangle_{\mathscr{T}}.$$

The norm $\|\cdot\|$ and the penalty parameters γ_ℓ should be chosen appropriately, to suit the structure of the functional. For example, if E is λ-convex with respect to the norm $\|\cdot\|$, then, for sufficiently large γ_ℓ, the functional $\tilde{\Phi}_\ell$ is strictly convex (cf. Proposition 2.1). The atomistic functional E, defined in (4.6) is λ-convex with respect to the $|\cdot|_{\mathrm{w}^{1,2}_\varepsilon}$-semi-norm (cf. Lemma 5.5 (i)). However, Lemma 5.5 (ii) shows that the lower bound on the Hessian tends to $-\infty$ as $N \to \infty$. This worst case cannot occur in practise though, as shown in Lemma 5.5 (iii), but only under 'infinite compression', as $y_i' \to 0$ for some i. Indeed, if the deformations are restricted to $y_i' \geq z'$ then $|\cdot|^2_{\mathrm{w}^{1,2}_\varepsilon}$ can in fact be used to convexify E in this region. The choice $\|\cdot\| = |\cdot|_{\mathrm{w}^{1,\infty}_\varepsilon}$ seems therefore reasonable.

We supplement the proximal point algorithm described in the previous section with an adaptive procedure. At each step we will, if necessary, adapt the mesh. For $\ell = 0, 1, \ldots$, let $\mathscr{T}_\ell$ be QC meshes and let $Y^{(0)} \in \mathscr{A}(\mathscr{T}_0)$ be a starting guess. Recalling that E is differentiable at all deformations which have finite energy, for the ℓth step of the PPA we wish to *find* $Y^{(\ell)} \in \mathscr{A}(\mathscr{T}_\ell)$ *satisfying*

$$\tilde{\Phi}_\ell'(Y^{(\ell)}; V) = 0 \qquad \forall V \in \mathscr{A}_0(\mathscr{T}_\ell), \tag{5.6}$$

where

$$\tilde{\Phi}_\ell(Y) = \frac{\gamma_\ell}{2}|Y - Y^{(\ell-1)}|^2_{\mathrm{w}^{1,2}_\varepsilon} + E(Y) - \langle f, Y\rangle_{\mathscr{T}_\ell},$$

and $\tilde{\Phi}_\ell'$ is given by

$$\tilde{\Phi}_\ell'(Y; V) = \sum_{i=1}^N \varepsilon\big(Y_i - Y_i^{(\ell-1)}\big)' V_i' + E'(Y; V) - \langle f, V\rangle_{\mathscr{T}_\ell}.$$

The PPA (5.6) can be interpreted as the implicit Euler discretization of the gradient flow of $E - \langle f, \cdot\rangle_{\mathscr{T}_\ell}$ with respect to the $\mathrm{w}^{1,2}_\varepsilon$-semi-norm. As such, it is in principle possible to analyze the error of this *discrete evolution*. We could imagine that the PPA is applied first to the full atomistic problem, then to the QC approximation and analyze the error between those two discrete evolutions. However, due to the lack of convexity, the resulting error estimates typically overestimate the actual error by several orders of magnitude. Instead, motivated by standard practice in the field of ODE solvers, we shall analyze the local error instead, i.e., the error committed by replacing the full space $\mathscr{A}$ by the QC space $\mathscr{A}(\mathscr{T}_\ell)$ in the ℓth step of the PPA. Unlike for the numerical solution of ODEs, since we are only interested in the efficient computation of a static equilibrium, this is entirely justified. We therefore also define the functional

$$\Phi_\ell(y) = \frac{\gamma_\ell}{2}|y - Y^{(\ell-1)}|^2_{\mathrm{w}^{1,2}_\varepsilon} + E(y) - \langle f, y\rangle_\varepsilon,$$

and the corresponding problem to *find* $y^{(\ell)} \in \mathscr{A}$ *such that*

$$\Phi'_\ell(y^{(\ell)}; v) = 0 \qquad \forall v \in \mathscr{A}_0. \tag{5.7}$$

The embedding of the error analysis and adaptivity into the optimization method achieves a considerable increase in performance. Otherwise the mesh would have to be adapted after the termination of the optimization method and the entire optimization step repeated, which would be particularly cumbersome during the formation of defects (unless the location of defect is known *a priori*) when optimization can be a computationally expensive process.

5.2 *A Posteriori Existence* and Error Estimates

In this section, we develop the theory required for an adaptive implementation of the PPA described in Section 5.3. We analyze a single step of the PPA only; thus we omit the sub- and super-scripts ℓ throughout and replace $Y^{(\ell-1)}$ by $Y^{(0)}$. Furthermore, we make the simplifying assumption that, if a mesh is coarsened, then $Y^{(0)}$ is assumed to be a member of the coarse space. We shall approximately enforce this condition by a requirement that an element may be coarsened only if the resulting interpolation error is sufficiently small. This simplification allows us to assume that $Y^{(0)} \in \mathscr{A}(\mathscr{T})$.

In the following theorem, $\|\cdot\|$ should be taken either as $|\cdot|_{\mathrm{w}^{1,\infty}_\varepsilon}$ or as $|\cdot|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}}$. Note that if we set $\gamma = 0$ then Theorem 5.1 gives an *a posteriori* existence result for an atomistic solution. In the residual estimate in Theorem 5.2, we understand sums over empty sets to be zero.

Theorem 5.1 (*A Posteriori* Existence) *Let $\|\cdot\|$ be a norm in $\mathscr{A}_0$. Let $Y \in \mathscr{A}$ and let $R(Y)$, $\mu(Y)$ and $\eta(Y)$ be non-negative numbers satisfying*

$$0 < \mu(Y) \;\leq\; \min_{\substack{y\in\mathscr{A}\\ \|y-Y\|\leq R(Y)}} \min_{\substack{u\in\mathscr{A}_0\\ \|u\|=1}} \max_{\substack{v\in\mathscr{A}_0\\ |v|_{\mathrm{w}^{1,1}_\varepsilon}=1}} \Phi''(Y;u,v), \quad \text{and} \tag{5.8}$$

$$\|\Phi'(Y)\|_* \;\leq\; \eta(Y). \tag{5.9}$$

If $\eta(Y) \leq \mu(Y)R(Y)$, then there exists $y \in \mathscr{A}$ satisfying $\Phi'(y) = 0$ in $\mathscr{A}$ such that

$$\|y - Y\| \leq \frac{\eta(Y)}{\mu(Y)}. \tag{5.10}$$

Proof. This theorem is an application of Theorem 3.5 with $\|\cdot\|_{\mathscr{X}} = \|\cdot\|$, $\|\cdot\|_{\mathscr{Y}} = |\cdot|_{\mathrm{w}^{1,1}_\varepsilon}$ and $\mathscr{F} = \Phi'$. $\square$

Theorem 5.2 (Residual Bound) *Let $Y \in \mathscr{A}(\mathscr{T})$ satisfy $\tilde{\Phi}'(Y) = 0$ in $\mathscr{A}(\mathscr{T})$, then*

$$\|\Phi'(Y)\|_* \leq \max_{k=1,\dots,K} \eta_{r,k} + \max_{k=1,\dots,K} \eta_{s,k} =: \eta(Y),$$

where

$$\eta_{r,k} = \max_{i=t_{k-1}+1,\dots,t_k} \left| \sum_{j=t_{k-1}+1}^{i-1} \Phi'_j(Y) - \sum_{j=i}^{t_k-1} \Phi'_j(Y) \right|, \tag{5.11}$$

$$\eta_{s,k} = h_k^2 \max\big(|f|_{\mathrm{w}_\varepsilon^{2,\infty}(\!(t_{k-1},t_k)\!)}, 2|f|_{\mathrm{w}_\varepsilon^{1,\infty}(\!(t_{k-1}+1,t_k)\!)} + 2|f|_{\mathrm{w}_\varepsilon^{1,\infty}(\!(t_{k-1},t_k-1)\!)}\big). \tag{5.12}$$

In particular, if $t_k - t_{k-1} = 1$ then $\eta_{r,k} + \eta_{s,k} = 0$.

A proof of Theorem 5.2 as well as a detailed discussion about the concrete evaluation of the residual terms and their interpretation and comparison with residual estimates in continuum mechanics is given in §5.2.1.

Theorem 5.3 (Stability Estimate)

(a) *For each $y \in \mathbb{R}^{N+1}$ with $z' = \min_{i=1,\dots,N} y'_i$, we have*

$$\min_{\substack{u\in\mathscr{A}_0 \\ |u|_{\mathrm{w}_\varepsilon^{1,\infty},1}=1}} \max_{\substack{v\in\mathscr{A}_0 \\ |v|_{\mathrm{w}_\varepsilon^{1,1},1}=1}} \Phi''(y;u,v) \geq \frac{1}{2}\big(\gamma + \min_{i=1,\dots,N} J''(y'_i) - \rho_\infty(z')\big), \tag{5.13}$$

$$\text{where} \quad \rho_\infty(z') = \sum_{r=2}^{\infty} r^2 \max_{z\geq rz'} |J''(z)|. \tag{5.14}$$

(b) *If, in addition, $y'_\xi \geq z_c$, then*

$$\min_{\substack{u\in\mathscr{A}_0 \\ |u|_{\mathrm{w}_{\varepsilon,f}^{1,\infty},1}=1}} \max_{\substack{v\in\mathscr{A}_0 \\ |v|_{\mathrm{w}_\varepsilon^{1,1},1}=1}} \Phi''(y;u,v) \geq \frac{1}{2}\big(\gamma + \min_{\substack{i=1,\dots,N \\ i\neq\xi}} J''(y'_i) - \rho_\infty(z')\big). \tag{5.15}$$

A proof of Theorem 5.3 is contained in §5.2.3. While Theorems 5.1 and 5.2 are generic, it must be emphasized that Theorem 5.3 provides a good estimate only if the deformation y has a generic but nevertheless very specific structure, namely elastic (small) deformation with possibly one single fracture. If $\min_{i=1,\dots,N} y'_i < z_t/2$, then the bound is not sharp. On the other hand, if $\max_{i=1,\dots,N} y'_i \geq z_t$ then $\mu(y)$ is zero or negative; cf. §5.2.3.

5.2.1 Residual bounds

Let $Y \in \mathscr{A}(\mathscr{T})$ be a QC solution, i.e., suppose that $\Phi'(Y) = 0$ in $\mathscr{A}(\mathscr{T})$. To bound its residual $\|\Phi'(Y)\|_*$, we use the usual Galerkin orthogonality argument to obtain

$$\begin{aligned}\Phi'(Y;u) &= \Phi'(Y;u-\Pi u) + \Phi'(Y;\Pi u) \\ &= \Phi'(Y;u-\Pi u) + \Big(\Phi'(Y;\Pi u) - \tilde{\Phi}'(Y;\Pi u)\Big) \qquad \forall u \in \mathscr{A}_0, \quad (5.16)\end{aligned}$$

where Πu is the nodal interpolant defined in §4.2.2. The second term in (5.16) was already estimated in §4.3.4. Using (4.2) and $|\Pi u|_{\mathrm{w}_\varepsilon^{1,1}} \leq |u|_{\mathrm{w}_\varepsilon^{1,1}}$, which can be verified by a straightforward computation, we obtain

$$|\Phi'(Y;\Pi u) - \tilde{\Phi}'(Y;\Pi u)| \leq \max_{k=1,\ldots,K} \eta_{s,k}\, |u|_{\mathrm{w}_\varepsilon^{1,1}}, \tag{5.17}$$

where $\eta_{s,k}$ is defined by (5.12).

For the first term in (5.16), we note that

$$\Phi'(Y;v) = \sum_{i=1}^{N-1} \Phi_i'(Y) v_i, \tag{5.18}$$

where

$$\Phi_i'(Y) = E_i'(Y) - \varepsilon f_i \qquad \forall\, i \in \{0,\ldots,N\} \setminus \mathscr{T}, \tag{5.19}$$

and we take $v = u - \Pi u$. For each $i \in \{t_{k-1}+1, \ldots, t_k - 1\}$, using the fact that v_i vanishes for $i = t_{k-1}$ and for $i = t_k$, we can write v_i as

$$v_i = \frac{1}{2}\left(\sum_{j=t_{k-1}+1}^{i} \varepsilon v_j' - \sum_{j=i+1}^{t_k} \varepsilon v_j' \right).$$

Inserting this into (5.18) and rearranging the summation gives

$$\begin{aligned}\Phi'(Y;v) &= \frac{1}{2}\sum_{k=1}^{K}\left[\sum_{i=t_{k-1}+1}^{t_k-1} \Phi_i'(Y) \sum_{j=t_{k-1}+1}^{i} \varepsilon v_j' - \sum_{i=t_{k-1}+1}^{t_k-1} \Phi_i'(Y) \sum_{j=i+1}^{t_k} \varepsilon v_j' \right] \\ &= \frac{1}{2}\sum_{k=1}^{K}\left[\sum_{j=t_{k-1}+1}^{t_k-1} \varepsilon v_j' \sum_{i=j}^{t_k-1} \Phi_i'(Y) - \sum_{j=t_{k-1}+2}^{t_k} \varepsilon v_j' \sum_{i=t_{k-1}+1}^{j-1} \Phi_i'(Y) \right] \\ &= \frac{1}{2}\sum_{k=1}^{K}\left[\sum_{j=t_{k-1}+1}^{t_k} \varepsilon v_j' \sum_{i=j}^{t_k-1} \Phi_i'(Y) - \sum_{j=t_{k-1}+1}^{t_k} \varepsilon v_j' \sum_{i=t_{k-1}+1}^{j-1} \Phi_i'(Y) \right].\end{aligned}$$

Note that, in the last line, sums over empty sets (whenever the lower summation index is larger than the upper summation index) may occur; each such empty sum is

considered to be zero. We use this convention in order to avoid complicated formulae. Upon setting

$$R_j = \frac{1}{2}\left[\sum_{i=j}^{t_k-1} \Phi_i'(Y) - \sum_{i=t_{k-1}+1}^{j-1} \Phi_i'(Y)\right] \qquad \text{for } t_{k-1} < j < t_k,$$

using the same summation convention as above, we obtain

$$\Phi'(Y;v) = \sum_{j=1}^{N} \varepsilon v_j' R_j. \tag{5.20}$$

An application of Hölder's inequality together with

$$|v|_{\mathrm{w}_\varepsilon^{1,1}(\!(t_{k-1},t_k)\!)} \le |u|_{\mathrm{w}_\varepsilon^{1,1}(\!(t_{k-1},t_k)\!)} + |\Pi u|_{\mathrm{w}_\varepsilon^{1,1}(\!(t_{k-1},t_k)\!)} \le 2|u|_{\mathrm{w}_\varepsilon^{1,1}(\!(t_{k-1},t_k)\!)}$$

gives the bound

$$\begin{aligned} |\Phi'(Y;v)| &\le 2\sum_{k=1}^{K} \Big[\max_{j=t_{k-1}+1,\dots,t_k} |R_j|\Big] |u|_{\mathrm{w}_\varepsilon^{1,1}(\!(t_{k-1},t_k)\!)} \\ &\le \max_{k=1,\dots,K} \eta_{r,k}\, |u|_{\mathrm{w}_\varepsilon^{1,1}}, \end{aligned} \tag{5.21}$$

where $\eta_{r,k}$ is defined by (5.11). Combining (5.17) with (5.21) we obtain

$$|\Phi'(Y;u)| \le \Big(\max_{k=1,\dots,K} \eta_{r,k} + \max_{k=1,\dots,K} \eta_{s,k}\Big) |u|_{\mathrm{w}_\varepsilon^{1,1}}$$

which concludes the proof of Theorem 5.2.

Formula (5.11) is not necessarily straightforward to implement. We therefore briefly discuss some interesting aspects of the residual estimate and an upper bound which reveals its structure and gives a form amenable to implementation. To this end, let us first assume that only nearest and next-nearest neighbour interactions occur, i.e., $\min_i Y_i' \ge z_c/3$. In that case, we can rewrite (5.19) as

$$\Phi_i'(Y) = J'(Y_{i-1}' + Y_i') + J'(Y_i') - J'(Y_{i+1}') - J'(Y_{i+1}' + Y_{i+2}') - \varepsilon f_i.$$

If $i \in \{t_{k-1}+1, \dots, t_k - 1\}$, then we always have $J'(Y_i') - J'(Y_{i+1}') = 0$. For $i \in \{t_{k-1}+2, \dots, t_k - 2\}$ we also have

$$J'(Y_{i-1}' + Y_i') - J'(Y_{i+1}' + Y_{i+2}') = J'(2\overline{Y}_k') - J'(2\overline{Y}_k') = 0.$$

Therefore, if $t_k - t_{k-1} \ge 3$, the auxiliary variables R_j can be estimated by

$$\begin{aligned} |R_j| &\le \frac{1}{2}(h_k - \varepsilon)\|f\|_{\ell_\varepsilon^\infty(\!(t_{k-1}+1,t_k-1)\!)} \\ &\quad + \frac{1}{2}\Big(\big|J'(2\overline{Y}_k') - J'(\overline{Y}_k' + \overline{Y}_{k-1}')\big| + \big|J'(2\overline{Y}_k') - J'(\overline{Y}_{k+1}' + \overline{Y}_k')\big|\Big). \end{aligned} \tag{5.22}$$

Similarly, if $t_k - t_{k-1} = 2$, then

$$|R_j| \le \frac{1}{2}(h_k - \varepsilon)\|f\|_{\ell^\infty_\varepsilon(\!(t_{k-1}+1,t_k-1)\!)} + \frac{1}{2}\big|J'(\overline{Y}'_{k+1} + \overline{Y}'_k) - J'(\overline{Y}'_k + \overline{Y}'_{k-1})\big|. \tag{5.23}$$

If $t_k - t_{k-1} = 1$, then obviously $\eta_{r,k} = 0$.

The first term in (5.22) and (5.23) is the same as the one we would have obtained in the continuum theory, except that the factor $h_k - \varepsilon$ would have been simply h_k. The second term in (5.22) and (5.23) is a purely atomistic effect and highlights the non-local interaction of the atoms. It represents a force at the interface between two elements which has not been fully resolved by the QC approximation.

For the practical computation of the indicators $\eta_{r,k}$, the following proposition which is a generalization of the above discussion is useful.

Proposition 5.4 *Suppose that $J(z) = 0$ for $z \ge z_c$. If $\min(i - t_{k-1}, t_k - i) \ge z_c/\overline{Y}'_k$, then $E'_i(Y) = 0$. In particular, we have*

$$\eta_{r,k} \le (h_k - \varepsilon)\|f\|_{\ell^\infty_\varepsilon(\!(t_{k-1}+1,t_k-1)\!)} + \sum_{\substack{i\in\{t_{k-1}+1,\dots,t_k-1\}\\ \min(i-t_{k-1},t_k-i)<z_c/\overline{Y}'_k}} |E'_i(Y)|.$$

Proof. Fix $k \in \{1, \dots, K\}$. If $i \in \{t_{k-1} + 1, \dots, t_k - 1\}$ then the derivative with respect to the penalty term vanishes and therefore,

$$\Phi'_i(Y) = \sum_{j=0}^{i-1} J'\big(\varepsilon^{-1}(Y_i - Y_j)\big) - \sum_{j=i+1}^{N} J'\big(\varepsilon^{-1}(Y_j - Y_i)\big) - \varepsilon f_i.$$

Since Y is affine in the set $\{t_{k-1}, \dots, t_k\}$, we have $Y_{i+j} - Y_i = Y_i - Y_{i-j}$ for $j = 1, \dots, r$ where $r = \min(i - t_{k-1}, t_k - i)$ and therefore

$$\Phi'_i(Y) = \sum_{j=0}^{i-r-1} J'\big(\varepsilon^{-1}(Y_i - Y_j)\big) - \sum_{j=i+r+1}^{N} J'\big(\varepsilon^{-1}(Y_j - Y_i)\big) - \varepsilon f_i.$$

For the remaining differences, we have $\varepsilon^{-1}|Y_j - Y_i| \ge r\overline{Y}'_k$ and hence $J'(\varepsilon^{-1}|Y_j - Y_i|) = 0$ if $r \ge z_c/\overline{Y}'_k$. $\square$

5.2.2 Spectral analysis of E''

Having provided a computable estimate on the residual, the crucial missing ingredient for the implementation of an adaptive algorithm is a technique that allows us to determine the inf-sup constant $\mu(Y)$ of Φ''. Instead, in a first step, we analyze the eigenvalues of E''. This analysis will be equally important in the practical implementation

of the optimization algorithm (cf. §5.3.2) and will also show which situations we need to focus on when discussing $\mu(Y)$. Furthermore, in order to justify our formulation of the PPA, we still need to investigate whether E is λ-convex.

For each $y \in \mathbb{R}^{N+1}$ let $\lambda(y)$ be the smallest eigenvalue of $E''(y)$ with respect to the $|\cdot|_{\mathrm{w}_\varepsilon^{1,2}}$-norm,

$$\lambda(y) = \inf_{\substack{u \in \mathscr{A}_0 \\ |u|_{\mathrm{w}_\varepsilon^{1,2}}=1}} E''(y;u,u).$$

Let $D(E)$ be the convex open subset of $\mathbb{R}^{N+1}$ where E is finite, and hence twice differentiable. Recall from Proposition 2.1 that E is λ-convex in $D(E)$ with respect to the $|\cdot|_{\mathrm{w}_\varepsilon^{1,2}}$-norm if, and only if, there exists a constant $\lambda \in \mathbb{R}$ such that

$$\lambda(y) \geq \lambda \quad \forall y \in D(E). \tag{5.24}$$

In the following lemma, we analyze the λ-convexity of E in three steps. We prove (i) that E is indeed λ-convex for some $\lambda = \lambda_N$, but (ii) that $\lambda_N \to -\infty$ as $N \to \infty$. This seemingly bad result is however remedied in step (iii) which demonstrates that the worst case only occurs in unphysical situations.

Lemma 5.5

(i) For each N there exists a constant C_N such that

$$\inf_{\substack{y \in \mathbb{R}^{N+1} \\ E(y)<\infty}} \lambda(y) \geq C_N. \tag{5.25}$$

(ii) For each $N \in \mathbb{N}$ there exists $y^{(N)} \in \mathscr{A}$ such that $E(y^{(N)}) < \infty$ and $\lambda(y^{(N)}) \to -\infty$.

(iii) For each $z' > 0$, we have

$$\sup_{\substack{y \in \mathbb{R}^{N+1} \\ y' \geq z'}} \sup_{\substack{u \in \mathbb{R}^{N+1} \\ |u|_{\mathrm{w}_\varepsilon^{1,2}}=1}} \left|E''(y;u,u)\right| \leq \sum_{r=1}^{\infty} r^2 \max_{z \geq rz'} |J''(z)|.$$

Proof. For statement (i) it is convenient to use the norm equivalence in finite dimensions and consider $\|\cdot\|_{\ell^2}$-eigenvalues instead of $|\cdot|_{\mathrm{w}_\varepsilon^{1,2}}$-eigenvalues. Fix a deformation $y \in \mathbb{R}^{N+1}$ at which $E''(y)$ exists. For each $u \in \mathscr{A}_0$, we have

$$\begin{aligned} E''(y;u,u) &= \sum_{n=1}^{N-1}\sum_{m=1}^{N-1} E''_{nm}(y)u_n u_m, \quad \text{where} \\ E''_{nm}(y) &= \begin{cases} \varepsilon^{-1}\sum_{i\neq n} J''\big(\varepsilon^{-1}|y_i - y_n|\big), & \text{if } n = m, \\ -\varepsilon^{-1} J''\big(\varepsilon^{-1}|y_n - y_m|\big), & \text{if } n \neq m. \end{cases} \end{aligned}$$

By Gershgorin's theorem, for each eigenvalue λ of the matrix $(E''_{nm}(y))_{n,m=1}^{N-1}$ there exists an n such that

$$|E''_{nn}(y) - \lambda| \leq \sum_{m \neq n} |E''_{nm}(y)|,$$

and hence

$$\begin{aligned}
\lambda &\geq E''_{nn}(y) - \sum_{m \neq n} |E''_{nm}(y)| \\
&\geq \varepsilon^{-1} \sum_{m \neq n} \Big[J''\big(\varepsilon^{-1}|y_n - y_m|\big) - \big|J''\big(\varepsilon^{-1}|y_n - y_m|\big)\big| \Big] \\
&\geq \sum_{m \neq n} 2\varepsilon^{-1} \min\big(0, J''(\varepsilon^{-1}|y_n - y_m|)\big) \\
&\geq 2\varepsilon^{-2} \min_{z \geq z_0} J''(z).
\end{aligned}$$

This concludes the proof of item (i).

Since part (ii) is a statement about the limit as $N \to \infty$, we can assume without loss of generality that N is arbitrarily large. Let $z'' > z_t$ be the point where J'' is minimal. Since we assumed that $J''(z_t) = 0$ and $J''(z_m) > 0$ such a point exists and $J''(z'') < 0$. Fix $\delta > 0$ such that $J''(z) \leq J''(z'')/2$ for all $z \in [z'' - \delta, z'' + \delta]$. We construct a deformation for which 'many' interactions are in the interval $[z'' - \delta, z'' + \delta]$. To this end, let $y'_i = z_m$ for $i = 1, \ldots, i_1$ where i_1 is maximal such that

$$\sum_{i=1}^{i_1} \varepsilon y'_i \leq z'' - \delta.$$

Furthermore, let y'_{i_1+1} be such that

$$\sum_{i=1}^{i_1+1} \varepsilon y'_i = z'' - \delta.$$

Note that the value i_1 is independent of ε. For $i = i_1 + 2, \ldots, N-1$, let $y'_i = 2\delta/N$ so that $\varepsilon^{-1}(y_i - y_0) \in [z'' - \delta, z'' + \delta]$ for $i = i_1 + 1, \ldots, N-1$. Finally let y'_N be an arbitrary value so that a prescribed boundary displacement is satisfied. Upon choosing N sufficiently large, y'_N may be assumed to be greater than or equal to the cut-off radius z_c.

Next, let the test function u be such that $u'_1 = \frac{1}{\sqrt{2}}\varepsilon^{-1/2}$ and $u'_N = -\frac{1}{\sqrt{2}}\varepsilon^{-1/2}$, then $|u|_{\mathrm{w}_\varepsilon^{1,2}} = 1$ and

$$\begin{aligned}
E''(y; u, u) &= \varepsilon F''_{11}(y)(u'_1)^2 + \varepsilon F''_{NN}(y)(u'_N)^2 + 2\varepsilon F''_{1N}(y) u'_1 u'_N \\
&= \tfrac{1}{2}\big(F''_{11}(y) + F''_{NN}(y) - 2F''_{1N}(y)\big).
\end{aligned}$$

Each of these terms can be easily bounded as follows:

$$\begin{aligned} F''_{11}(y) &= \sum_{i=1}^{N} J''\big(\varepsilon^{-1}(y_i - y_0)\big) \leq \text{Const.} + (N - i_1 - 2)J''(z'')/2, \\ F''_{NN}(y) &= \sum_{i=0}^{N-1} J''\big(\varepsilon^{-1}(y_N - y_i)\big) = 0, \quad \text{and} \\ F''_{1N}(y) &= J''\big(\varepsilon^{-1}(y_N - y_0)\big) = 0. \end{aligned}$$

Consequently, $E''(y; u, u) \leq c(1 - N)$, when N is sufficiently large, which tends to $-\infty$ as $N \to \infty$.

Finally, to prove item (iii) we estimate

$$\begin{aligned} |E''(y; u, u)| &\leq \sum_{i=1}^{N}\sum_{j=1}^{N} \varepsilon |F''_{ij}(y)||u'_i||u'_j| \\ &\leq \|F''\|\, |u|^2_{\mathrm{w}^{1,2}_\varepsilon}, \end{aligned}$$

where $\|F''\|$ denotes the ℓ^2-operator norm of the matrix $F'' = (F''_{ij}(y))_{i,j=1}^N$. $\|F''\|$ is the largest eigenvalue (in magnitude) of F'' for which, by Gershgorin's Theorem,

$$\max_{j=1,\ldots,N} \sum_{i=1}^{N} |F''_{ij}|$$

is an upper bound

Using similar computations as for the determination of ρ_2 and ρ_3 in §4.3.1, this value can be bounded by

$$\sum_{i=1}^{N} |F''_{ij}| \leq \sum_{r=1}^{\infty} r^2 \max_{z \geq rz'} |J''(z)|. \quad \square$$

The knowledge of eigenvalues is only important for the numerical optimization algorithm. Since our analysis is set in the $\mathrm{w}^{1,\infty}_\varepsilon$ topology, they do not enter the error estimate. However, we wish to note one important situation where the discrepancy of the eigenvalues of E'' restricted to $\mathscr{A}_0(\mathscr{T})$ and those of E'' in $\mathscr{A}_0$ is considerable, namely, when a deformation gradient in an element $\{t_{k-1}, \ldots, t_k\}$ of length strictly greater than one enters the non-convex region. In this case it is possible that E'' has only positive eigenvalues in $\mathscr{A}_0(\mathscr{T})$ but several negative or zero eigenvalues in $\mathscr{A}_0$. Thus, we may seriously underestimate the possible decrease of energy by local minimization.

Consider for example the case of fracture of a large element. An optimization algorithm would happily accept this state as a local minimum, unless it is somehow

able to recognize along the way that a direction of energy-decay was missed by the QC space. After the fracture has been created, refinement of the element does not help since the gradient may already have entered the cut-off region. These situations need to be detected, either by the adaptive procedure, or through a user-defined mesh.

More generally, we prove that any atomistic state with more than one fracture (for example across one large element, or in two different places) cannot be stable.

Proposition 5.6 *If $y \in \mathbb{R}^{N+1}$ with $y'_p \geq z_t$ and $y'_q \geq z_t$, where $p < q \in \{1, \ldots, N\}$, then $\lambda(y) \leq 0$. If y'_p or y'_q is strictly greater than but sufficiently close to z_t then $\lambda(y) < 0$.*

Proof. We perturb y with the displacement u given by

$$u'_i = \begin{cases} -\varepsilon^{-1/2}, & \text{if } i = p, \\ \varepsilon^{-1/2}, & \text{if } i = q, \text{ and} \\ 0, & \text{otherwise.} \end{cases}$$

Then, $|u|^2_{\mathrm{w}^{1,2}_\varepsilon} = 2$ and

$$\begin{aligned} E''(y;u,u) &= F''_{pp} + F''_{qq} - 2F''_{pq} \\ &= \sum_{i=p}^{N}\sum_{j=1}^{p} J''\big(\varepsilon^{-1}(y_i - y_{j-1})\big) + \sum_{i=q}^{N}\sum_{j=1}^{q} J''\big(\varepsilon^{-1}(y_i - y_{j-1})\big) \\ &\qquad -2\sum_{i=q}^{N}\sum_{j=1}^{p} J''\big(\varepsilon^{-1}(y_i - y_{j-1})\big) \\ &= \sum_{i=q}^{p-1}\sum_{j=1}^{q} J''\big(\varepsilon^{-1}(y_i - y_{j-1})\big) + \sum_{i=q}^{N}\sum_{j=q+1}^{p} J''\big(\varepsilon^{-1}(y_i - y_{j-1})\big). \end{aligned}$$

Since $y'_p, y'_q \geq z_t$ it follows that $J''(\varepsilon^{-1}(y_i - y_{j-1})) \leq 0$ for all i and j appearing in the last two sums. If either y'_p or y'_q is not too large then this expression is negative. Hence the result follows. $\square$

The proof of Proposition 5.6 reveals, in fact, that negative eigenvalues undetected by the QC method can occur even when $\overline{Y'_k}$ is less than (but sufficiently close to) z_t. Furthermore, it shows that for full-range interactions (without cut-off potential) we always have a negative eigenvalue if two or more fractures are present. Thus, we should rule out such cases from the class of stable configurations.

5.2.3 Estimation of the inf-sup constant

Our analysis in Chapter 4 showed that the inf-sup constants with respect to the $|\cdot|_{\mathrm{w}_\varepsilon^{1,p}}$-norms ($p = 1, \infty$) can be computed using the diagonal dominance of the Hessian matrices (F''_{nm}). Clearly, we do not wish to compute the entire matrix Φ''. In the following discussion we will provide an efficient way to compute the inf-sup constant which is heavily based on our analysis in Chapter 4. While our analysis is not sufficiently general to cover all possible types of solutions, we believe that it is sufficient for most purposes. We shall demonstrate this in §5.3.

Recall from §4.3.1 that the inf-sup constant for $E''(Y)$ can be bounded below by

$$\min_{\substack{u\in\mathscr{A}_0\\ |u|_{\mathrm{w}_\varepsilon^{1,\infty}}=1}} \max_{\substack{v\in\mathscr{A}_0\\ |v|_{\mathrm{w}_\varepsilon^{1,1}}=1}} E''(Y;u,v) \geq \min_{n=1,\dots,N} \frac{1}{2}\Big(F''_{nn}(Y) - \sum_{n\neq m} |F''_{nm}(Y)|\Big).$$

This was obtained by testing with v given by

$$v'_i = \begin{cases} \frac{1}{2}\varepsilon^{-1}, & \text{if } i = p \\ -\frac{1}{2}\varepsilon^{-1}, & \text{if } i = q \\ 0, & \text{otherwise,} \end{cases}$$

where $p, q \in \{1, \dots, N\}$ such that $u'_p = \max u'_i$ and $u'_q = \min u'_i$. Recall also that $u \in \mathscr{A}_0$ and hence u'_j must change sign. This makes it clear that adding $\gamma\mathrm{Id}$ to the matrix $(F''_{nm}(Y))_{n,m=1}^N$ corresponds to simply adding $\frac{1}{2}\gamma$ to the estimate.

We recall from §4.3.1, where the same computation was performed, that we can rewrite the stability factor as

$$\begin{aligned}\gamma + F''_{nn} - \sum_{m\neq n} |F''_{nm}| \;\;\geq\;\; & \gamma + J''(y'_n) - \sum_{i=n}^{N}\sum_{j=1}^{n-1} \big|J''\big(\varepsilon^{-1}(y_i - y_{j-1})\big)\big| \\ & - \sum_{m=1}^{n-1}\sum_{i=n}^{N}\sum_{j=1}^{m} \big|J''\big(\varepsilon^{-1}(y_i - y_{j-1})\big)\big| - \sum_{m=n+1}^{N}\sum_{i=m}^{N}\sum_{j=1}^{n} \big|J''\big(\varepsilon^{-1}(y_i - y_{j-1})\big)\big|.\end{aligned}$$

Upon setting $z' = \min_{i=1,\dots,N}$ and

$$\overline{J''}(r) = \max_{z\geq rz'} |J''(z)|$$

we obtain the bound

$$\gamma + F''_{nn} - \sum_{m\neq n} |F''_{nm}| \geq \gamma + J''(y'_n) - \sum_{r=2}^{\infty} r^2 \overline{J''}(r),$$

which concludes the proof of Theorem 5.3 (a). Since $J''(z) \leq 0$, this seemingly gross simplification is more-or-less sharp if $z' \geq z_t/2$.

Remarks. 1. It is important to note that ρ_∞ is in fact very simple to compute efficiently. If z' is not very close to zero then the calculation of ρ_∞ only involves the computation of a relatively small finite sum. ◄

2. Although in the computations in §5.3.6 it was not necessary, it may in general be advantageous not to estimate $\min_i y'_i$ globally but only locally. This could be crucial when $\min_i y'_i$ is significantly smaller than $\max_i y'_i$, for example, if $\max_i y'_i$ is close to z_t but $\min_i y'_i \le z_m$. ◄

Corollary 5.7 *For every $z' > 0$ there exists $\gamma_0 \ge 0$ such that, for all $\gamma \ge \gamma_0$, and for all $y \in \mathbb{R}^{N+1}$ with $y' \ge z'$, $\Phi''(y)$ has a positive inf-sup constant. Furthermore, as γ tends to infinity, so does the inf-sup constant.*

Consider the situation where the solution y has a fracture. In this case, we have no hope of ever obtaining a reasonable bound for the inf-sup constant of $E''(y)$ without regularization. In this case, we may replace the $\mathrm{w}^{1,\infty}_{\varepsilon}$-norm by the $\mathrm{w}^{1,\infty}_{\varepsilon,f}$-norm which we used in the analysis of fracture in §4.4. If $\gamma = 0$ then we can obtain (5.15) by the same computations as in §4.4.1. We only note that for $y'_\xi \ge z_c$ the term $\rho_{2,f}(z_f, z_1)$, defined in §4.4.1, vanishes and thus (5.15) follows immediately from Theorem 4.8. In general, however, we have to be more careful.

Let $y \in \mathscr{A}$ with $y'_\xi \ge z_c$. Let $u \in \mathscr{A}_0$ and define $P = \{i : u'_i > 0\}$ and $Q = \{i : u'_i < 0\}$. Without loss of generality, we assume that $u'_p = |u|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}}$ for some $p \in P$. We distinguish two cases. If $u'_\xi \le 0$, we proceed as in §4.4.1, setting $v'_p = \frac{1}{2}\varepsilon^{-1}$ and $v'_\xi = -\frac{1}{2}\varepsilon^{-1}$ to obtain

$$\Phi''(y; u, v) \ge \frac{1}{2}\Big(\gamma + F''_{pp}(y) - \sum_{n\in Q} \big|F''_{n\xi}(y)\big|\Big)|u|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}} + \frac{1}{2}\gamma|u'_\xi|,$$

which, using the fact that all interactions across the *crack* vanish, leads to (5.15). The extra term $\frac{1}{2}\gamma|u'_\xi|$ is simply neglected.

On the other hand, if $u'_\xi > 0$ then there exists $q \in Q$ such that y'_q is minimal. In this case, we proceed as in the case without fracture and note that the term u'_ξ appears nowhere in the calculation since $v'_\xi = 0$ and $J''(\varepsilon^{-1}(y_i - y_{j-1})) = 0$ whenever $i \ge \xi$ and $j \le \xi$. This concludes the proof of Theorem 5.3.

Finally, we formulate a result that will help us to determine the balls $B_f(Y, R) = \{y \in \mathscr{A} : |y - Y|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}} \le R\}$. In particular, since we can only compute the inf-sup constant with respect to $|\cdot|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}}$ if $y'_\xi \ge z_c$, we need to determine under what conditions all elements $y \in B_f(Y, R)$ satisfy this property.

Proposition 5.8 *Let $Y \in \mathscr{A}$ with $Y'_\xi > z_c$; then,*

$$y'_\xi \geq z_c \qquad \forall y \in B_f\big(Y, \varepsilon(Y'_\xi - z_c)\big).$$

Proof. For $y \in B_f(Y, R)$ we have

$$\begin{aligned} y'_\xi &= \varepsilon^{-1}\big(y_N^D - \sum_{i\neq\xi} \varepsilon y'_i\big) \\ &= \varepsilon^{-1}\big(y_N^D - \sum_{i\neq\xi} \varepsilon Y'_i + \sum_{i\neq\xi}(Y'_i - y'_i)\big) \\ &\geq Y'_\xi - NR. \end{aligned}$$

Hence, for $R \leq \varepsilon(Y'_\xi - z_c)$ the required property holds. □

5.3 Implementation and Numerical Examples

5.3.1 Implementation of the basic PPA

We begin by describing the implementation of a general PPA without an adaptive procedure. Let $\phi\colon \mathbb{R}^N \to (-\infty, +\infty]$ be twice continuously differentiable in its domain of definition $D(\phi)$ and let ϕ'' be locally Lipschitz continuous. Furthermore, let $\|\cdot\|$ be a euclidean semi-norm on $\mathbb{R}^N$ with respect to which ϕ is λ-convex. Let $(\cdot,\cdot)$ be the symmetric bilinear form associated with $\|\cdot\|$. Abusing notation, we define the dual norm of a linear functional f by $\|f\|_* = \sup_{\|u\|=1} f(u)$.

Suppose we are given an initial guess $y^{(0)}$ and penalty parameters $\gamma_\ell \geq 0$. The ℓth step of the PPA is to find a solution to the problem

$$\gamma_\ell(y - y^{(\ell-1)}, u) + \phi'(y; u) = 0 \qquad \forall u \in \mathbb{R}^N. \tag{5.26}$$

The efficiency of the method will depend on how well the norm $\|\cdot\|$ is chosen for the particular functional ϕ.

In order to guarantee that a local solution of the ℓth step (5.26) has lower energy than the previous iterate, we should try to guarantee that (5.26) is a locally convex problem. A necessary condition for this is that the parameter γ_ℓ is at least as large as $-\lambda(y^{(\ell-1)})$, where $\lambda(y)$ is the algebraically smallest (generalized) eigenvalue of $\phi''(y)$, i.e.,

$$\lambda(y) = \min_{\substack{u\in\mathbb{R}^N \\ \|u\|=1}} \phi''(y; u, u).$$

More generally, we define a local curvature estimate λ_ℓ which is initially set to $\lambda(y^{(\ell-1)})$ but may be adjusted during the computation of the ℓth step. In addition, we define a

non-negative parameter POSFAC which is updated during the computation and can be interpreted as a measure for the change of curvature of the problem. The initial guess and the updating procedure will be discussed in the following paragraphs. We set

$$\gamma_\ell = \max\big(\texttt{POSFAC} - \lambda_\ell, 0\big).$$

This allows the algorithm to reduce to a direct solution (Newton's method) whenever $\lambda_\ell > \texttt{POSFAC}$. If $\lambda(y^{(0)}) > 0$ then we initialize POSFAC to $\lambda(y^{(0)})/2$ so that the PPA reduces to Newton's method whenever possible. Otherwise, POSFAC is initialized to an arbitrary positive number.

The nonlinear equation (5.26) is solved using Newton's method with the residual termination criterion $\|\Phi'_\ell\|_* \leq \texttt{NEWTON\_TOL}$ where NEWTON_TOL is a positive predefined parameter. If γ_ℓ tends to infinity then the problem becomes essentially quadratic in the limit and hence Newton's method should terminate in a single step. This is made precise in Proposition 5.11. We therefore define three further parameters MAXIT, HIGHIT and LOWIT. If the number of Newton iterations required to solve (5.26) is larger than MAXIT, we repeat the step with an increased value of γ_ℓ. This is achieved by multiplying POSFAC by a constant predefined factor POSFAC_INC. If Newton's method terminates in at most MAXIT iterations then the step is accepted. If the number of iterations is at least HIGHIT, we increase POSFAC for the next step. If the number of iterations is less than LOWIT then the number of iterations is decreased by multiplying POSFAC by a constant, the predefined factor POSFAC_DEC.

Finally, we monitor the local curvature during the Newton iteration. Suppose that $Y^{\langle s\rangle}$ is the sth iteration of Newton's method for solving (5.26). If $\gamma_\ell + \lambda(Y^{\langle s\rangle}) \leq 0$ we interrupt the Newton iteration and repeat the PPA step with an updated curvature value $\lambda_\ell = \lambda(Y^{\langle s\rangle})$.

The algorithm described so far was sufficient in practise to achieve that each step of the PPA decreases the energy. However, in principle, this may fail. We therefore implement another test to see whether $\Phi_\ell(y^{(\ell)}) \leq \Phi_\ell(y^{(\ell-1)})$ at the ℓth step, and reject it and increase POSFAC if this is not the case.

The PPA terminates if either $\|y^{(\ell)} - y^{(\ell-1)}\| \leq \texttt{STEPTOL}$ or if $\|\phi'(y^{(\ell)})\|_* \leq \texttt{RESTOL}$.

The parameter values that were used in all non-adaptive computations are

$$\begin{aligned}
&\texttt{MAXIT} = 20, \quad \texttt{LOWIT} = 5, \quad \texttt{HIGHIT} = 12,\\
&\texttt{POSFAC_INC} = 4, \quad \texttt{POSFAC_DEC} = 1/4,\\
&\texttt{STEPTOL} = 0.0, \quad \texttt{RESTOL} = 10^{-8}, \quad \texttt{NEWTON_TOL} = 10^{-8}.
\end{aligned}$$

5.3.2 Convergence analysis

Despite the complexity of the heuristics used in the formulation of the PPA it is possible to give a fairly complete convergence theory. We first show that each step of the PPA is well defined. Based on the properties required for the termination of each PPA step, we prove that any accumulation point of the sequence of iterates satisfies a certain residual bound. Finally, we show that if the Hessian at an accumulation point is positive definite then the algorithm terminates after finitely many steps.

Proposition 5.9 *Let $\|\cdot\|$ be a Euclidean norm on $\mathbb{R}^N$ with respect to which ϕ is λ-convex and three times continuously differentiable. Then, for each $y^{(\ell-1)} \in \mathbb{R}^N$ there exists $\overline{\gamma} \geq 0$, depending only on $\|\phi'\|_*$, $\|\phi''\|$ and $\mathrm{Lip}(\phi'')$ in a neighbourhood of $y^{(\ell-1)}$, such that, for $\gamma_\ell \geq \overline{\gamma}$,*

(i) Newton's method for (5.26) *terminates in a single iteration, and*

(ii) $\Phi(Y^{\langle 1\rangle}) \leq \Phi(y^{(\ell-1)})$.

Proof. Let $M \in \mathbb{R}^{N\times N}$ be the positive definite matrix defining the inner product $(\cdot,\cdot)$ associated with the norm $\|\cdot\|$. Upon dividing (5.26) by γ_ℓ, we obtain the equivalent equation

$$\mathscr{F}(y) = M(y - y^{(\ell-1)}) + \gamma_\ell^{-1}\phi'(y^{(\ell)}) = 0. \tag{5.27}$$

Fix $R > 0$ such that $B(y^{(\ell-1)}, R) \subset D(\phi)$ and let

$$M_1 = \sup_{y\in B(y^{(\ell-1)},R)} \|\phi'(y)\|_*, \quad M_2 = \sup_{y\in B(y^{(\ell-1)},R)} \sup_{\substack{u\in\mathbb{R}^N \\ \|u\|=1}} \phi''(y;u,u),$$

and let L be the Lipschitz constant of ϕ'' in $B(y^{(\ell-1)}, R)$.

It follows that $\mathscr{F}'$ is Lipschitz continuous in $B(y^{(\ell-1)}, R)$ with Lipschitz constant L/γ_ℓ. Furthermore, as $\gamma_\ell \to \infty$, the inf-sup constant of $\mathscr{F}'(y^{(\ell-1)})$, i.e., the smallest eigenvalue with respect to the norm $\|\cdot\|$ tends to one. Hence, by Theorem 3.3, for $\gamma_\ell \geq \bar{\gamma}_1 = \bar{\gamma}_1(L, M_1)$ there exists a solution $y \in B(y^{(\ell-1)}, R)$ satisfying $\mathscr{F}(y) = 0$ or equivalently (5.26).

The first iteration of Newton's method, denoted by Y, is given by

$$\big[\gamma_\ell M + \phi''(y^{(\ell-1)})\big](Y - y^{(\ell-1)}) = -\phi'(y^{(\ell-1)}). \tag{5.28}$$

Multiplying by $(Y - y^{(\ell-1)})$, we obtain

$$(\gamma_\ell + \lambda)\|Y - y^{(\ell-1)}\|^2 \leq \|\phi'(y^{(\ell-1)})\|_*\|Y - y^{(\ell-1)}\| \tag{5.29}$$

and hence, for $\gamma_\ell \geq \bar{\gamma}_2 = \bar{\gamma}_2(M_1)$, $Y \in B(U,R)$. If we denote $H = \int_0^1 \phi''(y^{(\ell-1)} + \tau(y - y^{(\ell-1)}))\,\mathrm{d}\tau$, then by Taylor's Theorem (cf. also the proof of Theorem 3.3) y satisfies

$$(\gamma_\ell M + H)(y - y^{(\ell-1)}) = -\phi'(y^{(\ell-1)}). \tag{5.30}$$

Subtracting (5.30) from (5.28), we obtain

$$(\gamma_\ell M + \phi''(y^{(\ell-1)}))(Y - y) = (\phi''(y^{(\ell-1)}) - H)(y - y^{(\ell-1)}), \tag{5.31}$$

which implies

$$(\gamma_\ell + \lambda)\|Y - y\| \leq \tfrac{1}{2}L\|y - y^{(\ell-1)}\|^2.$$

Using (5.30) it follows that $\|y - y^{(\ell-1)}\| \leq \|\phi'(y^{(\ell-1)})\|_*/(\gamma_\ell + \lambda)$ and hence

$$\|Y - y\| \leq \tfrac{1}{2}LM_1^2/(\gamma_\ell + \lambda)^3.$$

For $\Phi'_\ell(Y)$, this gives the estimate

$$\|\Phi'_\ell(Y)\|_* = \|\Phi'_\ell(Y) - \Phi'_\ell(y)\|_* \leq (\gamma_\ell + M_2)\,\|Y - y\| \leq C(L, M_1)\frac{\gamma_\ell + M_2}{(\gamma_\ell + \lambda)^3}.$$

Hence, for sufficiently large $\gamma_\ell \geq \bar{\gamma}_3 = \bar{\gamma}_3(M_1, M_2, L)$, the termination criterion $\|\Phi'_\ell(Y)\|_* \leq$ `NEWTON_TOL` is satisfied.

Finally, to show that $\Phi_\ell(Y) \leq \Phi_\ell(y^{(\ell-1)})$ we test (5.28) with $Y - y^{(\ell-1)}$, giving

$$(\gamma_\ell + \lambda)\|Y - y^{(\ell-1)}\|^2 \leq -\phi'(y^{(\ell-1)}; Y - y^{(\ell-1)}),$$

which we use to obtain

$$\begin{aligned}
\Phi(Y) &= \Phi(y^{(\ell-1)}) + \Phi'(y^{(\ell-1)}; Y - y^{(\ell-1)}) + \tfrac{1}{2}\Phi''(\theta; Y - y^{(\ell-1)}, Y - y^{(\ell-1)}) \\
&= \Phi(y^{(\ell-1)}) - \tfrac{1}{2}\Phi''(y^{(\ell-1)}; Y - y^{(\ell-1)}, Y - y^{(\ell-1)}) \\
&\quad + \tfrac{1}{2}\Big[\Phi''(\theta; Y - y^{(\ell-1)}, Y - y^{(\ell-1)}) - \Phi''(y^{(\ell-1)}; Y - y^{(\ell-1)}, Y - y^{(\ell-1)})\Big] \\
&\leq \Phi(y^{(\ell-1)}) - \tfrac{1}{2}(\gamma_\ell + \lambda)\|Y - y^{(\ell-1)}\|^2 + L\,\|Y - y^{(\ell-1)}\|^3,
\end{aligned}$$

Thus, given the bound (5.29) we have for $\|Y - y^{\ell-1}\|$, for $\gamma_\ell \geq \bar{\gamma}_4 = \bar{\gamma}_4(M_1, L)$ it follows that $\Phi(Y) \leq \Phi(y^{(\ell-1)})$.

Setting $\bar{\gamma} = \max\{\bar{\gamma}_i : i = 1, \ldots, 4\}$ proves the desired result. $\square$

Proposition 5.9 shows that, if `POSFAC` is initially set to a positive value and if `POSFAC_INC` $>$ 1, then the formulation of the PPA presented in §5.3.1 defines a sequence $(y^{(\ell)})_{\ell=0}^{\infty}$ for which $\phi(y^{(\ell)}) \leq \phi(y^{(\ell-1)}) - \frac{1}{2}\gamma_\ell\|y^{(\ell)} - y^{(\ell-1)}\|^2$ and $\|\Phi'_\ell(y^{(\ell)})\|_* \leq$ `NEWTON_TOL`.

The following corollary establishes some asymptotic convergence of the PPA.

Corollary 5.10 *Let $\|\cdot\|$ be a Euclidean norm in $\mathbb{R}^N$, and $\phi : \mathbb{R}^N \to (-\infty, +\infty]$ be λ-convex with respect to $\|\cdot\|$. Suppose that ϕ is three times continuously differentiable and that ϕ'' is Lipschitz continuous in level sets of ϕ.*

Let $(y^{(\ell)})_{\ell=0}^{\infty}$ be the sequence of the PPA described in §5.3.1. If $\|y^{(\ell)}\|$ is bounded then

$$\|\phi'(y^{(\ell)})\|_* \leq \texttt{NEWTON\_TOL} + r_\ell,$$

where $r_\ell \to 0$ as $\ell \to \infty$.

Proof. Since the sequence $(y^{(\ell)})_{\ell\in\mathbb{N}}$ is bounded, the constant $\bar{\gamma}$ in Theorem 5.9 can be chosen uniformly for all ℓ and hence $\gamma_\ell \leq \bar{\gamma}$ for all ℓ.

Suppose, for contradiction, that $\gamma_\ell \|y^{(\ell)} - y^{(\ell-1)}\|$ does not converge to zero. Since γ_ℓ is bounded above, there exists a subsequence $\ell_j \uparrow \infty$ such that

$$\|y^{(\ell_j)} - y^{(\ell_j-1)}\| \geq \delta > 0 \qquad \forall j \in \mathbb{N}.$$

Furthermore, since $y^{(\ell)}$ is bounded, γ_ℓ must be bounded below by a constant $\underline{\gamma} > 0$.

From the definition of the PPA and Proposition 5.9 (ii) it follows that

$$\phi(y^{(\ell_j)}) \leq \phi(y^{(\ell_j-1)}) - \frac{1}{2}\underline{\gamma}\delta^2$$

which implies that $\phi(y^{(\ell)}) \downarrow -\infty$ and contradicts the fact (cf. Lemma 2.2) that ϕ is bounded below on bounded sets. Thus, we conclude that

$$\gamma_\ell \|y^{(\ell)} - y^{(\ell-1)}\| \to 0 \qquad \text{as } \ell \to \infty, \tag{5.32}$$

which immediately implies the required statement. □

An important feature of the PPA is that, if $\texttt{POSFAC} \leq \lambda_\ell$ then $\gamma_\ell = 0$, i.e., it is possible that the PPA reduces to Newton's method. This situation should occur whenever the iteration enters the basin of attraction of a critical point at which the Hessian is positive definite. With this result, we conclude our analysis of the PPA.

Proposition 5.11 *Let ϕ satisfy the conditions of Theorem 5.10 and let $\texttt{LOWIT} \geq 1$ and $\texttt{POSFAC\_DEC} < 1$. Let $y^* \in \mathbb{R}^N$ be a critical point of ϕ such that $\phi''(y^*)$ is positive definite. Then, there exists an $R > 0$ such that, if for some ℓ_0, $y^{(\ell_0)} \in B(y^*, R)$, then $y^{(\ell)} \in B(y^*, R)$ for all $\ell \geq \ell_0$. Furthermore there exists $\ell_1 \geq \ell_0$ such that for $\ell \geq \ell_1$, $\gamma_\ell = 0$. For $\ell \geq \ell_1$, $y^{(\ell)}$ converges to y^* q-quadratically.*

Proof. Since $\phi''(y^*)$ is positive definite, there exists $R_0 > 0$ such that $\phi''(y) \geq \lambda_0 > 0$ for all $y \in B(y^*, R_0)$.

We prove that for $y^{(\ell-1)}$ sufficiently close to y^*, Newton's method for (5.26) terminates in a single step, regardless of the value of γ_ℓ. Let $R_1 \leq R_0$ be sufficiently small so that $\|\phi'(y)\|_* \leq \lambda_0/L$, for all $y \in B(y^*, R_1)$, where L is the Lipschitz constant of ϕ'' in $B(y^*, R_0)$. Let $Y = Y^{\langle 1 \rangle}$ be the first iteration of Newton's method for (5.26), i.e.,

$$\Phi''_\ell(y^{(\ell-1)})(Y - y^{(\ell-1)}) = -\phi'(y^{(\ell-1)}).$$

Using $\phi'(y^*) = 0$ we obtain, for some $\theta \in \operatorname{conv}\{y^*, y^{(\ell-1)}\}$,

$$\begin{aligned} \Phi''_\ell(y^{(\ell-1)})(Y - y^*) &= -\Phi''_\ell(y^{(\ell-1)})(y^* - y^{(\ell-1)}) + \phi'(y^*) - \phi'(y^{(\ell-1)}) \\ &= -\gamma_\ell M(y^* - y^{(\ell-1)}) + (\phi''(\theta) - \phi''(y^{(\ell-1)}))(y^* - y^{(\ell-1)}). \end{aligned}$$

It thus follows that

$$(\gamma_\ell + \lambda_0)\|Y - y^*\| \leq \gamma_\ell \|y^* - y^{(\ell-1)}\| + L\|y^* - y^{(\ell-1)}\|^2.$$

For $\|y^{(\ell-1)} - y^*\| \leq \lambda_0/L$ we therefore have

$$\|Y - y^*\| \leq \|y^{(\ell-1)} - y^*\|. \tag{5.33}$$

To estimate $\Phi'_\ell(Y)$, consider

$$\begin{aligned} \Phi'_\ell(Y) &= \phi'_\ell(y^{(\ell-1)}) + \Phi''_\ell(\theta)(Y - y^{(\ell-1)}) \\ &= \phi'_\ell(y^{(\ell-1)}) + \Phi''_\ell(y^{(\ell-1)})(Y - y^{(\ell-1)}) + \big[\Phi''_\ell(\theta) - \Phi''_\ell(y^{(\ell-1)})\big](Y - y^{(\ell-1)}) \\ &= \big[\Phi''_\ell(\theta) - \Phi''_\ell(y^{(\ell-1)})\big](Y - y^{(\ell-1)}), \end{aligned}$$

which implies

$$\|\Phi'_\ell(Y)\|_* \leq L\|Y - y^{(\ell-1)}\|^2. \tag{5.34}$$

Thus, if R_1 is chosen sufficiently small (depending only on the Lipschitz constant L) then $\|\Phi'_\ell(Y)\|_* \leq$ `NEWTON_TOL`.

Since `POSFAC_DEC` < 1, `POSFAC` is decreased by a constant factor. Hence, after finitely many steps, `POSFAC` $\leq \lambda_0$. Since Newton's iteration terminates after the first step which always remains in $B(y^*, R_1)$ it follows that $\lambda_\ell \geq \lambda_0$ for all $\ell \geq \ell_0$ and hence $\gamma_\ell = 0$ for all sufficiently large ℓ.

For $\ell \geq \ell_1$, each step of the PPA is precisely one iteration of Newton's method. Hence, by (5.34), $y^{(\ell)}$ converges to y^* q-quadratically; cf. also [88, Theorem 5.3.2]. $\square$

5.3.3 PPA versus Optimization Toolbox

We compare the implementation of the PPA to the large-scale trust region method (the command `fminunc` with appropriate set of parameters) of MATLAB's optimization toolbox. Our benchmark is a QC model problem, $\phi(\overline{Y}) = E(Y) - \langle Y, f\rangle_{\mathscr{T}}$ and $\|\cdot\| = |\cdot|_{\mathrm{w}^{1,2}_{\varepsilon}}$, defined as follows.

First we determine a stress-free reference state by (approximately) solving $E'(\hat{y}) = 0$ with a Dirichlet condition on only the left-hand end of the domain. The atomistic potential is the Morse potential with $\alpha = 5.0$ and cut-off radius $z_c = 2.7$. We define the applied body-force by

$$f_i = \begin{cases} 0.03, & \text{if } i \geq \xi \\ -0.03, & \text{if } i < \xi. \end{cases}$$

This non-smooth body-force creates a stress intensifier between the two atoms at sites $\xi - 1$ and ξ which is where we should physically expect fracture to occur. The constant 0.03 is rather arbitrary. It is sufficiently small so that the body-force does not dominate the equation but sufficiently large so that the QC method should be able to find the correct fracture. The QC mesh, with roughly 50 nodes, is constructed such that nodes are clustered with full atomistic resolution at both ends of the atomistic domain as well as around ξ, and scaled smoothly in between.

We then successively solve for $Y(t)$ satisfying $E'(Y(t)) = f$ subject to the boundary conditions $Y_0(t) = 0$ and $Y_N(t) = \hat{y}_N + t$, for

$$t = 0.0, 0.025, 0.05, 0.075, 0.1, 0.115, 0.1215, 0.1245, 0.1257, 0.15. \tag{5.35}$$

The initial condition for each step is obtained by adding an affine function to the previously obtained equilibrium and so that it satisfies the new boundary condition.

Since the two methods, the PPA and the trust region method, are very different both in terms of their design and implementation it is hard to compare them directly. For example, our PPA uses a direct method to solve the linear systems while `fminunc` uses a conjugate gradient method. Furthermore, there are no provisions in `fminunc` to take the specific structure of the energy functional into account, which we have done with our choice of penalization norm $|\cdot|_{\mathrm{w}^{1,2}_{\varepsilon}}$. All we can offer therefore are rough qualitative remarks that are intended to demonstrate the efficiency of our PPA, specifically for the QC method.

Table 5.1 summarizes some results which highlight the performance of the two algorithms. For the first 8 quasistatic steps, the two methods perform very similarly. Essentially, both reduce to a Newton method. Note that while each iteration of the

	PPA		Opt. Toolbox	
q.s. step	iterations	linear systems	iterations	CG iterations
1	1	1	0	0
2	1	3	2	17
3	1	3	2	19
4	1	3	2	19
5	1	3	3	29
6	1	3	3	31
7	1	4	3	31
8	1	4	3	33
9	42	203	–	–
10	1	5	19	475

Table 5.1: Iteration count and linear system count for the proximal point algorithm and MATLAB's trust region method `fminunc`.

PPA is an application of Newton's method, each iteration of the trust region method is only one iteration of Newton's method.

The two methods only start to differ significantly in the presence of defects. In step 9, when the fracture forms, the trust region method `fminunc` failed to converge in 10^4 steps while the PPA converged in well under 100 iterations. While we had expected that our PPA would perform much better in the formation of the defect — it is after all designed specifically for this purpose — we are somewhat puzzled by the bad performance of `fminunc`. It is perhaps even more surprising that `fminunc` fails to recognise that, for the final step, a simple Newton iteration is again sufficient.

The performance of the PPA does not deteriorate as N becomes larger. For $N = 10^4$, the step 9 required 37 iterations and 175 linear systems, for $N = 10^5$ it required 43 iterations and the solution of 161 linear systems. However, while for $N = 10^3$ the fracture index was close to ξ (depending on the specific choice of parameters between ξ and $\xi + 3$), for $N = 10^5$ the algorithm computed an unphysical fracture at an element larger than the atomic spacing. Problems such as this can be solved by adding an adaptive procedure to the optimization algorithm.

As a final verdict on the performance of our PPA, we would have to test it on a greater variety of benchmark problems and compare it to more advanced optimization packages, such as TRON [58], GALAHAD [52] or KNITRO [25]. We have obtained some evidence though that our optimization method can achieve good performance for the highly non-convex optimization problems occurring in the QC method.

5.3.4 Adaptivity in the PPA

We now add an adaptive procedure into the innermost loop of the PPA, the solution of the equation (5.26), applied to the QC method. A flowchart of the resulting adaptive PPA which is detailed in the following paragraphs is shown in Figure 5.1.

Suppose that we are given an initial condition $Y^{(0)} \in \mathscr{A}(\mathscr{T}^{(0)})$ and suppose furthermore, that we have already computed the $(\ell-1)$th step $Y^{(\ell-1)} \in \mathscr{A}(\mathscr{T}^{(\ell-1)})$. To compute $Y^{(\ell)}$ we choose a mesh $\mathscr{T}^{(\ell)}$, initially set to $\mathscr{T}^{(\ell)} = \mathscr{T}^{(\ell-1)}$, and solve

$$\tilde{\Phi}'_\ell(Y;U) = 0 \qquad \forall U \in \mathscr{A}_0(\mathscr{T}). \tag{5.36}$$

We solve this equation by the procedure described in §5.3.1. We also compute the eigenvalues of E'' in $\mathscr{A}_0(\mathscr{T})$ in order to obtain the curvature parameters λ_ℓ. This is motivated by the fact that the solution of (5.36) only depends on the QC eigenvalues but not on the eigenvalues of the full atomistic problem. Our only modification is to define a new non-negative parameter `POSFAC_A`,

$$\texttt{POSFAC\_A} = -\min\big(0, \min_{i=1,\dots,N} J''(y'_i) - \rho_\infty(\min_i y'_i)\big), \tag{5.37}$$

and redefine

$$\gamma_\ell = \max\big(0, \texttt{POSFAC} + \texttt{POSFAC\_A} - \lambda_\ell\big),$$

which allows additional control on the penalization. In particular, it guarantees that the penalization does not tend to zero and we thus avoid an overrefinement of the mesh before the current iterate has entered a region of coercivity. The definition (5.37) is used when the error is measured in the $|\cdot|_{\mathrm{w}^{1,\infty}_\varepsilon}$-semi-norm while we use

$$\texttt{POSFAC\_A} = -\min\big(0, \min_{i\neq\xi} J''(y'_i) - \rho_\infty(\min_i y'_i)\big)$$

instead if the error is measured in the $|\cdot|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}}$-semi-norm. As it turned out, the adaptive version of the PPA was less well able to cope with a large number of Newton iterations which create small penalty terms γ_ℓ which in turn require a much more refined mesh. By changing the parameters to

$$\texttt{LOWIT} = 3, \quad \texttt{HIGHIT} = 5, \quad \texttt{MAXIT} = 10,$$

the performance did not deteriorate significantly while the mesh size was stabilized.

Suppose now that (5.36) has a QC solution $\tilde{Y}$ that was accepted by the PPA. We estimate the error committed and, if necessary, refine the mesh and repeat the step.

To this end, we compute the residual bound $\eta = \eta(\tilde{Y})$, using Theorem 5.2 and Proposition 5.4. These values are passed to a search algorithm which tries to find

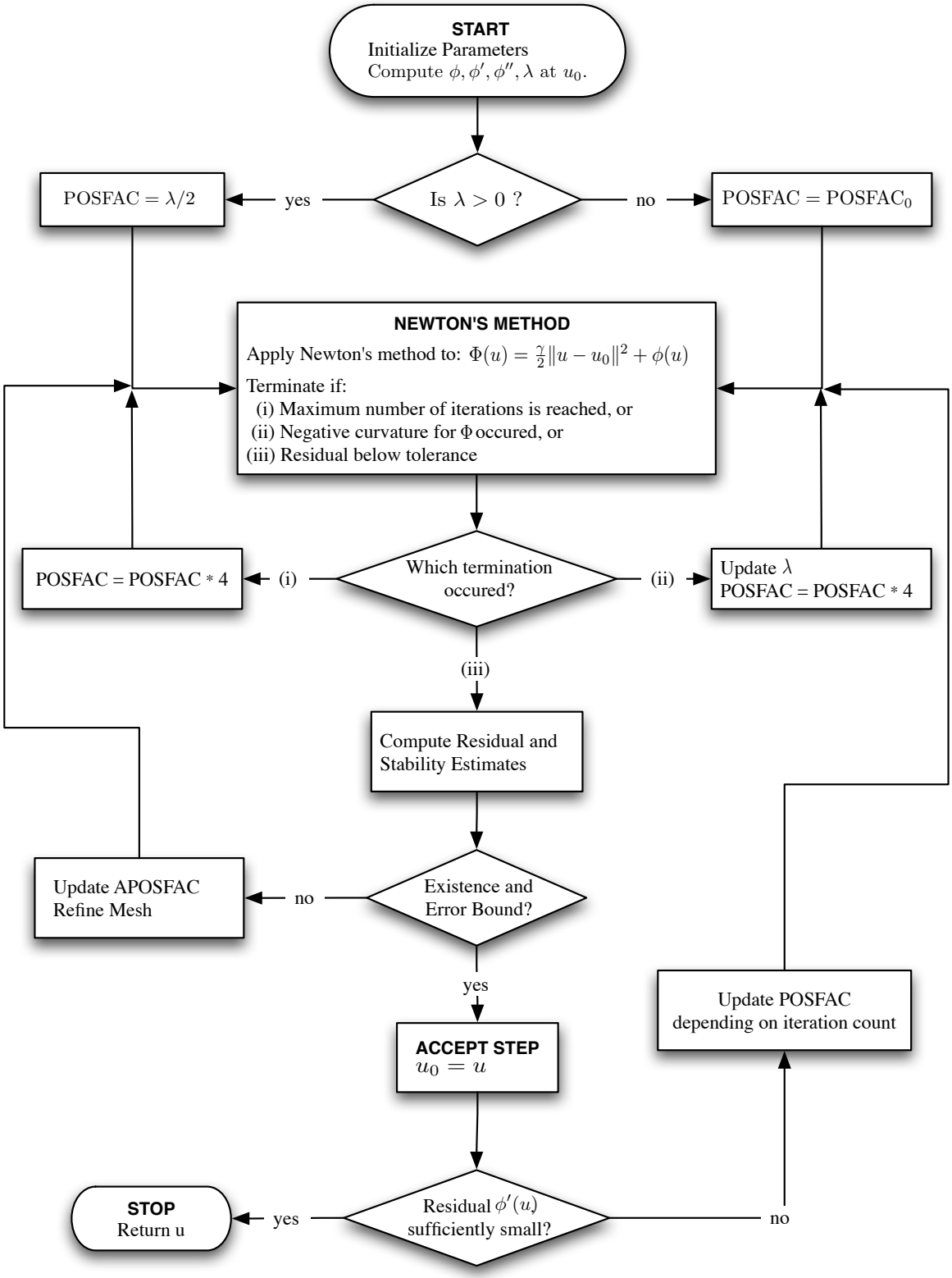

Figure 5.1: Flow-chart of the adaptive proximal point algorithm which is described in Sections 5.3.1, 5.3.2, 5.3.3 and 5.3.4.

optimal radii (if they exist) R and R_f such that $\eta/\mu \leq R$ and $\eta/\mu_f \leq R_f$ where μ and μ_f are the respective inf-sup constants in $B(\tilde{Y}, R)$ and $B_f(\tilde{Y}, R_f)$ with respect to the norms $|\cdot|_{\mathrm{w}^{1,\infty}_{\varepsilon}}$ and $|\cdot|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}}$. These constants can be computed using Theorem 5.3. In order to determine admissible radii R_f, we use Proposition 5.8. The following situations can now occur.

1. If no radius R_f exists such that $\eta \leq \mu_f R_f$, then we use the $|\cdot|_{\mathrm{w}^{1,\infty}_{\varepsilon}}$-norm in the analysis:

   1.1 There exists R such that $\eta/\mu \leq R$: Find R for which this holds and for which μ is maximal. Use $\eta/\mu \leq \texttt{TOL}$ as a refinement criterion. If $\eta/\mu \leq \texttt{TOL}$ set $Y^{(\ell)} = \tilde{Y}$ and increase ℓ by one to continue the PPA. Otherwise, use the refinement criterion to obtain a new mesh $\mathscr{T}_\ell$ and repeat the step.

   1.2 There exists no R for which $\eta/\mu \leq R$: Find R such that μR is maximal and use $\eta \leq \mu R$ as a refinement criterion to obtain a new mesh $\mathscr{T}_\ell$ with which to repeat the ℓth PPA step.

   1.3 There exists no $R > 0$ such that $\mu > 0$: Recompute `POSFAC_A` with the new stability estimate and repeat the ℓth PPA step.

2. If there exists a radius $R_f \leq \tilde{Y}'_\xi - z_c$ such that $\eta \leq \mu_f R_f$, then we use the $|\cdot|_{\mathrm{w}^{1,\infty}_{\varepsilon,f}}$-norm in the analysis: Take $\eta/\mu_f \leq \mathrm{TOL}$ as a refinement criterion. If it is satisfied, set $Y^{(\ell)} = \tilde{Y}$ and increase ℓ by one to continue the PPA. Otherwise, compute a new mesh $\mathscr{T}_\ell$ and repeat the ℓth PPA step.

5.3.5 Mesh coarsening

Ideally, an adaptive finite element method should have the capability to refine as well as coarsen a mesh. Translated to our context, a typical criterion to mark an element for coarsening would be

$$\eta_k/\mu \leq q \times \min(\mathrm{TOL}, R),$$

where $q \in (0, 1)$. This is based on the assumption that, say, doubling the size of an element should increase the residual roughly by a factor of two. However, in our atomistic problems we have no such property. This can best be seen by considering a fractured element which, as we discussed, must have length one and hence the residual in this element vanishes. Such an element would always be marked for coarsening. Because of this (and similar) difficulties a rigorous analysis of the coarsening procedure is difficult and only a heuristic idea is presented which seems to work well in practice.

We define a *pseudo-residual* $\tilde{\eta}_k$ which measures the residual in the kth element as if it were a large element. To this end, we recall from the discussion in §5.2.1 that most of the time only nearest and next-nearest neighbour interactions contribute to the energy. More generally, it is reasonable to assume that the residual contribution from long-range interactions can simply be neglected. Thus, we define

$$\tilde{\eta}_k = h_k \|f\|_{\ell^\infty_\varepsilon(\!(t_{k-1},t_k)\!)} + \left|J'(2\overline{Y}'_k) - J'(\overline{Y}'_k + \overline{Y}'_{k-1})\right| + \left|J'(2\overline{Y}'_k) - J'(\overline{Y}'_{k+1} + \overline{Y}'_k)\right|$$

with a suitable modification for boundary elements, and choose the coarsening criterion

$$\tilde{\eta}_k/\mu \leq q \times \min(\mathrm{TOL}, R)$$

in Case 1. of §5.3.4, and

$$\tilde{\eta}_k/\mu_f \leq q \times \min(\mathrm{TOL}, R_f),$$

in the Case 2. of §5.3.4. In our computations, we chose $q = 1/4$.

As a second criterion, we require that the interpolation error committed during coarsening is less than a specified tolerance, which should be a fraction of the tolerance TOL.

The coarsening is performed at the same time as the error estimation and mesh refinement.

5.3.6 Numerical example

We use the benchmark example from §5.3.3 to test our adaptive implementation. In the very first step of the quasistatic evolution the user has to supply an initial condition in the form of a QC mesh and the nodal values. This initial mesh has to be chosen so that the summation error term η^s_k in the residual can be neglected and does not have to be computed in the adaptive procedure. Consequently, we have also implemented a further safeguard in the coarsening procedure, preventing it to remove any nodes which are present in the original mesh.

For our particular example, it is sufficient to choose $\mathscr{T} = \{0, \xi - 1, \xi, N\}$ as the initial mesh. The benchmark example is run with $N \in \{10^3, 10^4, 10^5, 10^6\}$ and $\mathrm{TOL} \in \{10^{-2}, 10^{-3}\}$. The performance of the adaptive PPA is described in Table 5.2 and Figures 5.2 – 5.4 and in the following discussion.

In Table 5.2 we notice immediately that that number of iterations of the PPA seems to be roughly independent of both the tolerance level and the number of atoms. The higher number of PPA iterations at the 7th and 8th step were caused by the

	$N = 10^3$		$N = 10^4$		$N = 10^5$		$N = 10^6$	
q.s. step	10^{-2}	10^{-3}	10^{-2}	10^{-3}	10^{-2}	10^{-3}	10^{-2}	10^{-3}
1	1	2	1	2	1	2	1	2
2	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1
5	1	1	1	1	1	1	1	1
6	1	1	1	1	1	1	1	1
7	3	1	4	1	4	1	5	1
8	4	4	5	4	4	4	4	4
9	66	65	72	69	79	87	85	103
10	1	1	1	1	1	1	1	1

Table 5.2: Number of iterations of the adaptive PPA for the benchmark problem described in §5.3.3, for $N \in \{10^3, 10^4, 10^5, 10^6\}$ and TOL $\in \{10^{-2}, 10^{-3}\}$.

introduction of the parameter `APOS_FAC`. Near the turning point, the coercivity constant becomes quite small and the adaptive PPA is more 'careful' in this case.

Similarly, we see in Figure 5.2 that the number of degrees of freedom (DOFs) required to meet the tolerance depends only on TOL but not on N. These results indicate the robustness of the adaptive algorithm.

Next, in Figure 5.3 we analyze the efficiency of our error estimates. This test was only performed for $N = 10^3$ as it requires the computation of the full atomistic solution. For all tests, the efficiency index, the ratio between the estimated and the true error, lies between 2 and 8. In particular, the efficiency index does not explode as we approach the bifurcation point in step 8 of the quasistatic evolution.

Finally, in Figure 5.4, we plot the entire history of the adaptive PPA for the 9th quasistatic step which is the most interesting. This is done for $N = 10^3$ and TOL $\in \{10^{-2}, 10^{-3}\}$. We notice that the two *discrete evolutions* behave very similarly. As the local curvature estimate λ_ℓ becomes more and more negative, the penalization parameter increases. As the PPA iterations converge to the equilibrium, the stability (described by the constant μ_f in Theorem 5.3) increases and hence the number of DOFs required to meet the tolerance decreases as well. Note that between the 20th and 40th PPA iteration the number of DOFs are the same for TOL $= 10^{-2}$ and TOL $= 10^{-3}$. This indicates that the error tolerance was overwritten by the *a posteriori* existence condition.

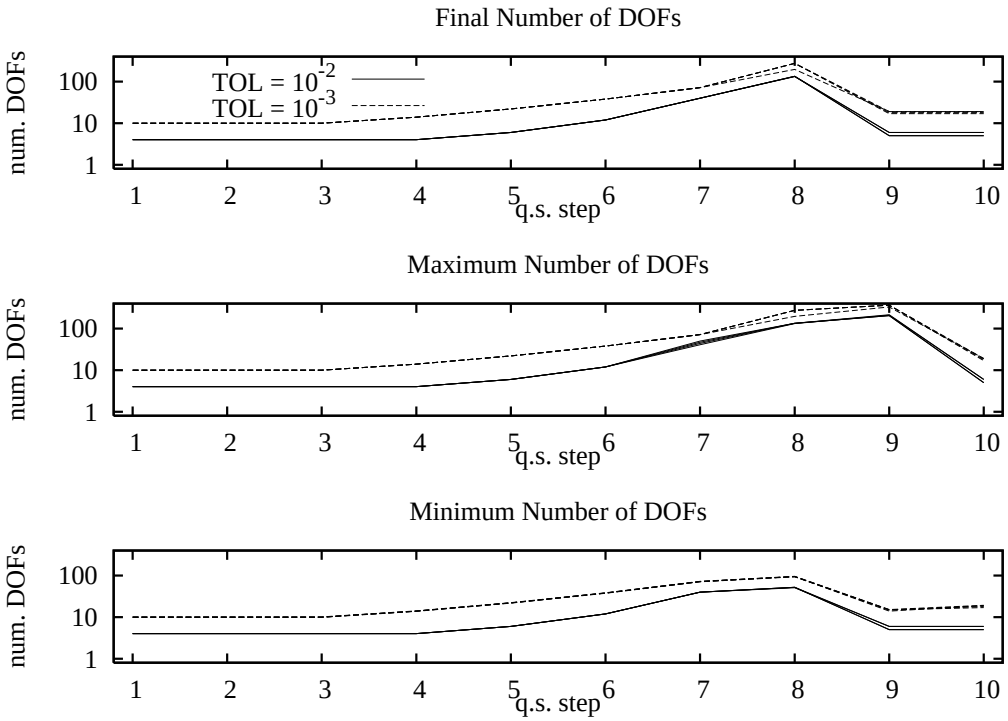

Figure 5.2: Number of DOFs in the adaptive PPA for the benchmark problem described in §5.3.3. We plot the maximum number of DOFs during each quasistatic iteration, the minimum number of DOFs and also the final number after termination of the PPA.

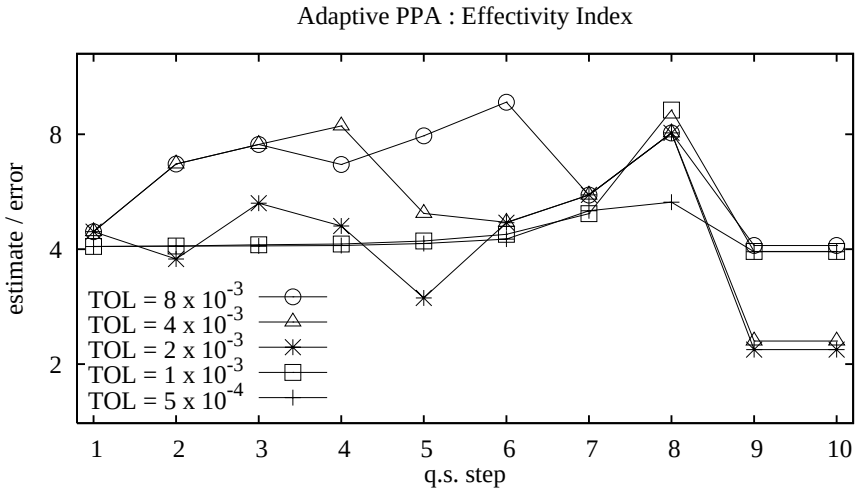

Figure 5.3: Efficiency index (ratio between estimated and true error) for the final error estimate (after termination of the PPA) for each quasistatic step of the benchmark problem described in §5.3.3. All tests are performed with $N = 10^3$.

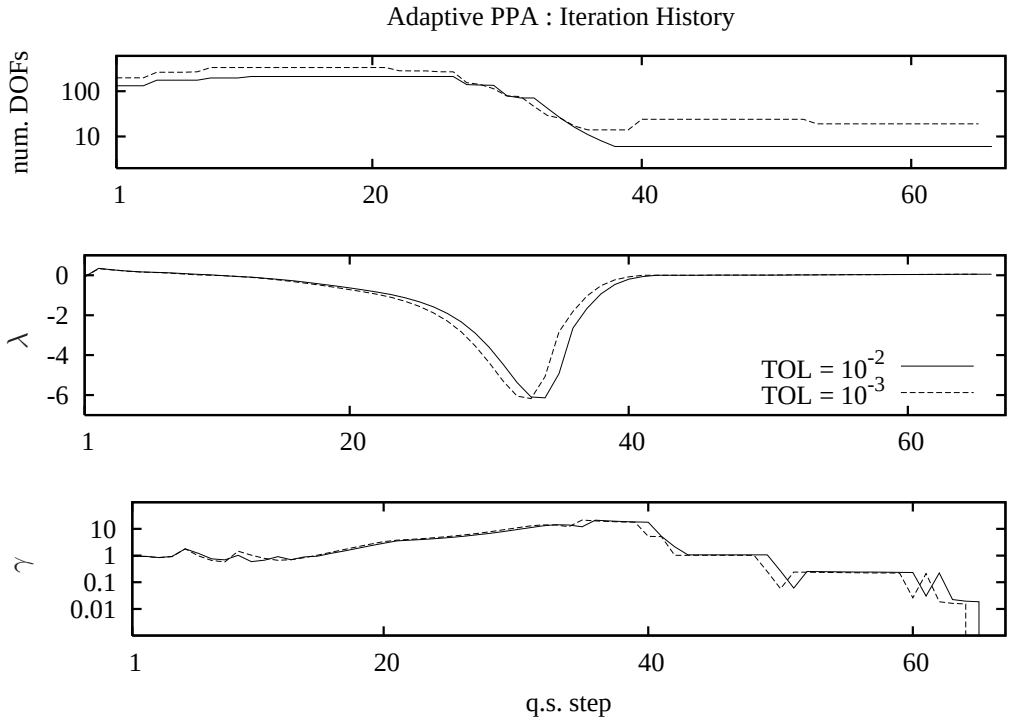

Figure 5.4: Iteration history of the adaptive proximal point algorithm for the 9th quasistatic step of the benchmark problem described in §5.3.3 where $N = 10^3$ and $\mathrm{TOL} \in \{10^{-2}, 10^{-3}\}$.

Conclusion

In this chapter, an *a posteriori* existence and error analysis for the quasicontinuum method in one dimension was presented. In particular, we consciously avoided to assume the existence of a nearby exact solution to the atomistic model, but instead the *a posteriori existence* technique developed Chapter 3 was used which made it possible to deduce the existence of atomistic solutions from the computation.

The *a posteriori* error analysis and existence results were integrated into an adaptive optimization method based on proximal point algorithms, and the numerical experiments presented in §5.3 demonstrate the effectiveness of the approach.

While we have seen that the stability-analysis of atomistic equations and the structure of the residual in the quasi-continuum method bear many similarities to the analysis of continuum problems in one dimension, this is no longer true in two or three space dimensions. The abstract *a posteriori* existence result remains valid of course; however, generalizing the stability estimates to higher dimensions seems a non-trivial challenge. In particular, it should be expected that the inf-sup constants with respect to similar norms depend on ε. See §6.3 for further comments on this issue.

Chapter 6

Outlook and Open Problems

Almost all results of the first five chapters were either abstract or restricted to one dimension. This is a clear limitation of the work in this thesis, particularly for its practical application. The goal of this chapter is to investigate, at least partially, where extensions are possible and which results are truly restricted to one dimension. Concentrating on the *a posteriori* analysis, which is the practically most relevant part of the thesis, we shall investigate the possibility of computing residual bounds for higher dimensional atomistic models, using the argument sketched out in the proof of Proposition 3.12.

To simplify the analysis, we shall make several assumptions on the atomistic model and its QC approximation which somewhat restrict the generality of the presentation in §1.4. First, we assume that the space of admissible deformations and the space of test functions coincide, i.e., $\mathscr{V} = \mathscr{A}$. Dirichlet boundary conditions can still be modelled in this context by adding a penalty term such as

$$\text{const.} \sum_{\xi \in \mathcal{N}_D} |y(\xi) - y_D(\xi)|^2$$

to the atomistic energy. Similarly, bound constraints can be modelled by barrier functions. It should, however, be possible to generalize the analysis to *strong* constraints by modifying the Clément operators Π and Π_c defined in §6.1.3.

Furthermore, to avoid having to distinguish many different cases, we assume that the domain Ω is 'convex'. To make this precise, let $\bar{\Omega}$ denote the union of all QC elements and assume that this set is convex. The only point where we will use this assumption directly is the Poincaré inequality (6.7) in the proof of Lemma 6.4. When this restriction is lifted some changes in the analysis, laid out in Remark 1 at the end of §6.2, are required.

In contrast with the introductory presentation in Chapter 1 we shall, however, not require that the reference domain Ω is a subset of a regular lattice. The only condition that we need is that a weak Cauchy–Born rule such as Proposition 1.2 holds. This is made precise in §6.2 where the exact requirement will become apparent.

6.1 Connecting Discrete and Continuum

If the approach in the proof of Proposition 3.12 should be followed more or less closely then the crucial condition would be to define 'discrete star-shaped sets' and extend the interpolation and trace inequalities, or to somehow link the discrete deformations to continuum deformations which would make it possible to use continuum techniques for the definition and approximation error analysis of the Clément operator.

The choice made here is to follow the latter route. This decision has not been made out of convenience but because of the fact that the discrete geometries of QC finite elements (or unions thereof) have far inferior properties than classical finite elements. Consider, for example, the element shown in Figure 6.1. Although in continuum analysis a high quality element, the marked vertex is not connected to the rest of the element through the nearest neighbour relation. It is effects such as this one that make a direct discrete analysis very inefficient. Note that such effects do not vanish when working with continuum methods instead, however, by switching between continuum and discrete variables only at the beginning and at the end of the analysis, the connection has to be made only on a global level where it is more easily controlled and understood.

6.1.1 Voronoi tesselations

There are many possibilities how an atomistic deformation can be interpreted as a continuum deformation. Possibly the most commonly used is spline interpolation. For example, the two-dimensional triangular lattice has a natural triangulation which could be used to interpolate an atomistic deformation using piecewise affine splines. However, such an interpolation would be difficult to use at corners where the resulting domain may become disconnected. Instead, we shall use a partition of the continuum domain $\bar{\Omega}$ into Voronoi cells.

Definition 6.1 *Let Z be a finite subset of $\mathbb{R}^d$. The Voronoi tesselation of $\mathbb{R}^d$ with respect to Z is the collection $V(Z) = \{C_z; z \in Z\}$ of the closed, convex sets*

$$C_z = \{x \in \mathbb{R}^d : |x - z| \leq |x - z'| \quad \forall z' \in Z\}.$$

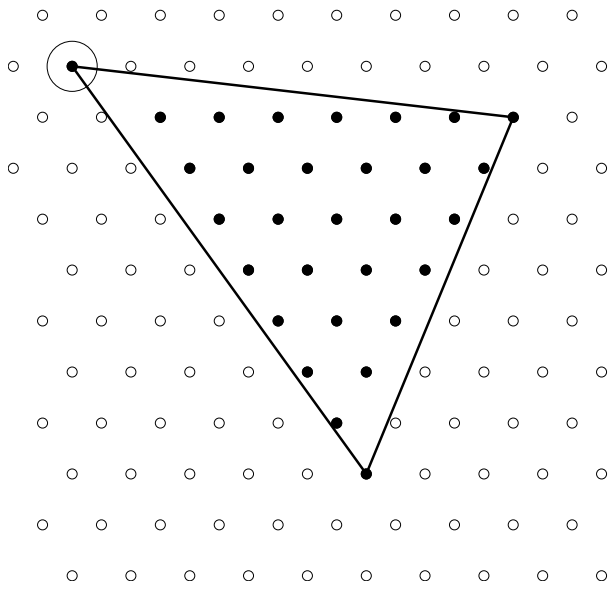

Figure 6.1: A disconnected QC element: the marked element vertex is not connected via the nearest neighbour relation to the rest of the element.

Let $V = V(\Omega)$ be the Voronoi tesselation associated with the set Ω (cf. Figure 6.2 for an example). We define the *lifting operator* $\mathscr{R} : \mathscr{A} \to \mathrm{PC}$, where PC denotes the space of piecewise constant functions of $\mathbb{R}^d$, via

$$\mathscr{R}y(x) = y(\xi) \qquad \forall\, x \in \mathrm{int}(C_\xi) \quad \forall \xi \in \Omega. \tag{6.1}$$

Two atoms ξ, ξ' are called nearest neighbours in Ω if $\mathcal{H}^{d-1}(C_\xi \cap C_{\xi'}) > 0$. In that case we write $x \overset{\Omega}{\sim} x'$.

Since we shall frequently integrate over the Voronoi cells restricted to $\bar{\Omega}$, we define $\bar{C}_\xi = C_\xi \cap \bar{\Omega}$. For each $\xi \in \Omega$ let $r_\xi = \sup_{x \in \bar{C}_\xi} |x - \xi|$. For an atomistic domain without 'holes', where the distance between any two nearest neighbours is roughly ε, it is reasonable to assume that

$$r_\xi \leq \varepsilon \qquad \forall \xi \in \Omega. \tag{6.2}$$

As possible dual norms for the residual, we use the BV norm in $\bar{\Omega}$ (cf. §A.1), given by

$$\begin{aligned} \|y\|_{\mathscr{V}^{1,1}} &= \|\mathscr{R}y\|_{\mathrm{BV}(\bar{\Omega})} = \|\mathscr{R}y\|_{\mathrm{L}^1(\bar{\Omega})} + |D(\mathscr{R}y)|(\bar{\Omega}) \\ &= \sum_{\xi \in \Omega} |\bar{C}_\xi||y(\xi)| + \sum_{\xi \overset{\Omega}{\sim} \xi'} \mathcal{H}^{d-1}(\bar{C}_\xi \cap \bar{C}_{\xi'})\, |y(\xi) - y(\xi')|, \end{aligned} \tag{6.3}$$

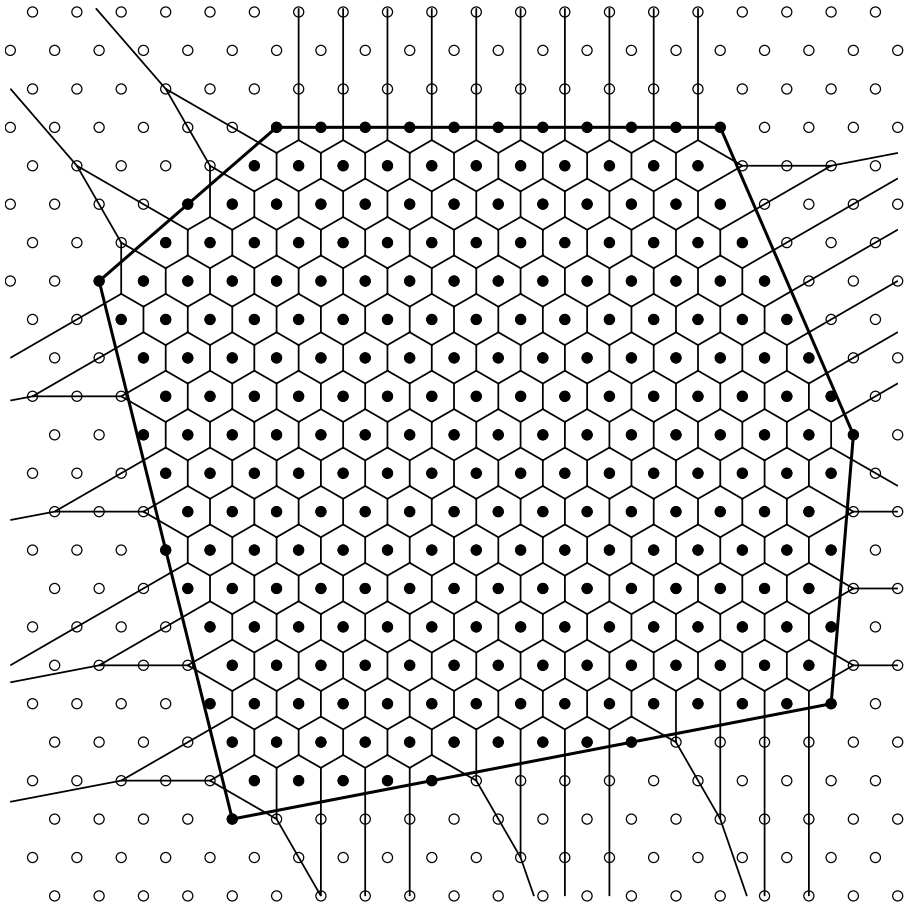

Figure 6.2: Voronoi tesselation of a convex atomistic domain.

and the corresponding semi-norm

$$|y|_{\mathscr{V}^{1,1}} \;=\; |D(\mathscr{R}y)|(\bar{\Omega}) \;=\; \sum_{\xi \overset{\Omega}{\sim} \xi'} \mathcal{H}^{d-1}(\bar{C}_\xi \cap \bar{C}_{\xi'})\, |y(\xi) - y(\xi')|. \tag{6.4}$$

These (semi-)norms are convenient extension of the $\mathrm{w}^{1,1}_\varepsilon$-(semi-)norms which we used in the one-dimensional analysis. Consider particularly the following proposition which establishes the equivalence between the $\mathscr{V}^{1,1}$-(semi-)norms and other discrete norms.

Proposition 6.2 *For each $\varphi \in \mathscr{V}$, we have*

$$C_V^{-1}\|\mathscr{R}\varphi\|_{\mathrm{L}^1(\bar{\Omega})} \le \sum_{\xi\in\Omega} \varepsilon^d |\varphi(\xi)| \le c_V^{-1}\|\mathscr{R}\varphi\|_{\mathrm{L}^1(\bar{\Omega})}, \quad \text{where}$$
$$c_V = \min_{\xi\in\Omega} \varepsilon^{-d}|\bar{C}_\xi| \;\text{ and }\; C_V = \max_{\xi\in\Omega} \varepsilon^{-d}|\bar{C}_\xi|.$$

and

$$C_V'^{-1}|\varphi|_{\mathscr{V}^{1,1}} \le \sum_{\xi \overset{\Omega}{\sim} \xi'} \varepsilon^{d-1}|\varphi(\xi)-\varphi(\xi')| \le c_V'^{-1}|\varphi|_{\mathscr{V}^{1,1}}, \quad \text{where}$$

$$c_V' = \min_{\xi \overset{\Omega}{\sim} \xi'} \varepsilon^{1-d}\mathcal{H}^{d-1}(\bar{C}_\xi \cap \bar{C}_{\xi'}) \quad \text{and} \quad C_V' = \max_{\xi \overset{\Omega}{\sim} \xi'} \varepsilon^{1-d}\mathcal{H}^{d-1}(\bar{C}_\xi \cap \bar{C}_{\xi'}).$$

Proof. Both results follow immediately from (6.3) and (6.4). □

While the constants C_V, c_V', and C_V' are only required to demonstrate the equivalence of the $\mathscr{V}^{1,1}$-(semi-)norms, the constant c_V plays a prominent role in the analysis of the residual. In particular, the computed residual bound may tend to infinity as $c_V \to 0$. Figure 6.2 demonstrates clearly what kind of behaviour might be expected, at least for subsets of the two-dimensional triangular lattice. In generic situations, C_V and C_V' cannot be too large (cf. condition (6.2)), however, at corners with a small angle, the volume of the cell tends rapidly to zero which causes a deterioration of the constants c_V and c_V'.

It should be remarked, however, that so far we have not assumed any particular structure on the atomistic ground state. Neither should this be necessary as the following proposition demonstrates. It shows that, for practical purposes, the constant c_V depends only on the geometry of the domain $\bar{\Omega}$ but not on the atomistic ground state.

Proposition 6.3 *If for any two points $\xi, \xi' \in \Omega$, $|\xi - \xi'| \ge \varepsilon$ is satisfied, then*

$$\min_{\xi \in \Omega} \varepsilon^{-d}|C_\xi| \ge 2^{-d} v_d$$

where v_d denotes the volume of the d-dimensional unit ball.

Suppose furthermore that $\bar{\Omega}$ is a convex polygonal set which satisfies the following interior cone condition: there exists $r > 0$ and an angle $\alpha > 0$ such that, for each $\xi \in \Omega$, there exists a cone $C(\xi)$ centered in ξ with opening angle α, satisfying $C(\xi) \cap B(\xi, r) \subset \bar{\Omega}$. In that case, the constant c_V depends only on d, α and possibly on r.

Proof. For each $\xi, \xi' \in \Omega$ and for each $x \in B(\xi, \varepsilon/2)$, we have

$$|x - \xi'| \ge |\xi - \xi'| - |x - \xi| \ge \varepsilon - \varepsilon/2 = \varepsilon/2,$$

which shows that $B(\xi, \varepsilon/2)$ is contained in C_ξ. Hence the first result follows.

The second result can be obtained in the same way by replacing the ball $B(\xi, \varepsilon/2)$ by the set $C(\xi) \cap B(\xi, r \wedge \varepsilon/2)$. □

For none of the other three parameters can a similar result be obtained without further specific information about the topology of the reference state. It should not be too difficult, however, to obtain similar results for regular lattices.

6.1.2 Notation

Using the nearest-neighbour relation in Ω defined in the previous section, some additional notation can be established. Let

$$\mathscr{E} = \{e = \kappa_e \cap \kappa_e' : \mathcal{H}^{d-1}(\kappa_e \cap \kappa_e') > 0\} \cup \{e = \partial\bar{\Omega} \cap \kappa : \kappa \in \mathscr{T}, \mathcal{H}^{d-1}(\kappa \cap \partial\bar{\Omega}) > 0\}$$

be the set of $(d-1)$-dimensional faces associated with the mesh $\mathscr{T}$. Throughout, we shall use κ_e and κ_e' to denote the two neighbouring elements of the face e if e is an interior face and κ_e the single element associate to e if it is a boundary face. To each face we also associate a unit normal ν_e, which is identical to the outward unit normal to κ_e on this face. By $\mathscr{E}_\mu$ we denote all those faces which are *fully refined*, i.e.,

$$\mathscr{E}_\mu = \{e \in \mathscr{E} : \xi \overset{\Omega}{\sim} \xi' \quad \forall \xi, \xi' \in \mathscr{N} \cap e\},$$

and we denote $\mathscr{E}_M = \mathscr{E} \setminus \mathscr{E}_\mu$. Correspondingly, we also define the microscopic and macroscopic domains by

$$\Omega_\mu = \{\xi \in \mathscr{N} : \exists \xi' \in \mathscr{N} \text{ s.t. } \xi \sim_\Omega \xi'\} \quad \text{and} \quad \Omega_M = \Omega \setminus \Omega_\mu.$$

6.1.3 Definition and analysis of the Clément operator

For each $z \in \mathscr{N}$, define

$$T_z = \bigcup_{\substack{\kappa \in \mathscr{T} \\ z \in \kappa}} \kappa,$$

and let $B_z = B(z, \rho_z) \cap \bar{\Omega}$ be a (convex) section of a closed ball with centre z and radius ρ_z such that $B_z \subset T_z$.

For each $z \in \mathscr{N}$ we define

$$h_z = \sup_{x \in \partial T_z} |x - z| \quad \text{and} \quad \gamma_z = h_z/\rho_z. \tag{6.5}$$

The scalar γ_z is called the chunkiness factor of T_z (cf. [22, Definition 4.2.16]). For macroscopic elements, the nodal values $\Pi\varphi(z)$ of the Clément operator will be defined by averaging $\mathscr{R}y$ over B_z. The error thus committed can be controlled by the following result. We denote the average of an integrable function over a measurable set A by $(u)_A = |A|^{-1} \int_A u \, \mathrm{d}x$.

Lemma 6.4 *Let $u \in \mathrm{BV}(T_z)$, then*

$$\|u - (u)_{B_z}\|_{\mathrm{L}^1(T_z)} \leq \rho_z\big(\gamma_z^d + \gamma_z^d/d - 1/d\big)|Du|(T_z).$$

Proof. This proof is a modification of the proof of [96, Lemma 4.1], which covers the H^1 situation. To simplify notation we drop the subscripts and write $B = B_z$, $\rho = \rho_z$, etc.. We furthermore assume, without loss of generality, that $z = 0$ and that $(u)_{B_z} = 0$.

Since T is the finite union of closed, convex sets, it is star-shaped with respect to the origin. Using the local approximation of BV functions by smooth functions (cf. [46, Sec. 5.2.2]), there exists a sequence $u_j \in \mathrm{BV}(T) \cap \mathrm{C}^\infty(\mathrm{int}(T))$ such that $u_j \to u$ strictly in BV, i.e., $u_j \to u$ strongly in L^1 and $|Du_j|(K) \to |Du|(K)$ as $j \to \infty$. Hence, we can assume without loss of generality that $u \in \mathrm{C}^\infty(T)$. We assume furthermore that $(u)_{B_z} = 0$ and that $z = 0$.

We write

$$\|u\|_{\mathrm{L}^1(T)} = \|u\|_{\mathrm{L}^1(B)} + \|u\|_{\mathrm{L}^1(T\setminus B)}. \tag{6.6}$$

Let Σ be the subset of the unit sphere in $\mathbb{R}^n$ such that, for each $\sigma \in \Sigma$, the ray $t\sigma$, $t \geq 0$, points into T. For each $\sigma \in \Sigma$, let $r(\sigma)\sigma \in \partial T$. For the second term in (6.6), we compute

$$\begin{aligned}
\|u\|_{\mathrm{L}^1(T\setminus B)} &= \int_\Sigma \int_\rho^{r(\sigma)} t^{d-1}|u(t\sigma)|\,\mathrm{d}t\,\mathrm{d}s(\sigma) \\
&\leq \int_\Sigma \int_\rho^{r(\sigma)} t^{d-1}\big|u(t\sigma) - u(\rho\sigma)\big|\,\mathrm{d}t\,\mathrm{d}s(\sigma) + \int_\Sigma \int_\rho^{r(\sigma)} t^{d-1}\big|u(\rho\sigma)\big|\,\mathrm{d}t\,\mathrm{d}s(\sigma) \\
&=: S_1 + S_2.
\end{aligned}$$

To obtain a bound on S_1, consider

$$\begin{aligned}
S_1 &= \int_\Sigma \int_\rho^{r(\sigma)} t^{d-1}\Big|\int_\rho^t \partial_r u(r\sigma)\,\mathrm{d}r\Big|\,\mathrm{d}t\,\mathrm{d}s(\sigma) \\
&\leq \rho^{1-d}\int_\Sigma \int_\rho^{r(\sigma)} t^{d-1}\int_\rho^t r^{d-1}|\partial_r u(r\sigma)|\,\mathrm{d}r\,\mathrm{d}t\,\mathrm{d}s(\sigma) \\
&\leq \frac{1}{d}\rho^{1-d}(h^d - \rho^d)\int_\Sigma \int_\rho^{r(\sigma)} r^{d-1}|\partial_r u(r\sigma)|\,\mathrm{d}r\,\mathrm{d}s(\sigma) \\
&\leq \frac{\rho}{d}(\gamma^d - 1)\|\nabla u\|_{\mathrm{L}^1(T\setminus B)}.
\end{aligned}$$

For S_2, we estimate

$$\begin{aligned}
S_2 &= \frac{1}{d}\int_\Sigma \big(r(\sigma)^d - \rho^d\big)|u(\rho\sigma)|\,\mathrm{d}s(\sigma) \\
&\leq \frac{\rho}{d}\int_\Sigma \Big[\frac{h^d}{\rho^d} - 1\Big]\rho^{d-1}|u(\rho\sigma)|\,\mathrm{d}s(\sigma) \\
&= \frac{\rho}{d}(\gamma^d - 1)\int_\Sigma \rho^{d-1}|u(\rho\sigma)|\,\mathrm{d}s(\sigma).
\end{aligned}$$

The last term can be bounded as follows,

$$\begin{aligned}\int_\Sigma \rho^{d-1}|u(\rho\sigma)|\,\mathrm{d}s(\sigma) &= \int_\Sigma \rho^{d-1}\Big|\int_0^\rho \partial_r\Big[\Big(\frac{r}{\rho}\Big)^d u(r\sigma)\Big]\,\mathrm{d}r\Big|\,\mathrm{d}s(\sigma)\\ &= \int_\Sigma \rho^{d-1}\Big|\int_0^\rho \Big[\Big(\frac{r}{\rho}\Big)^d \partial_r u(r\sigma) + \frac{dr^{d-1}}{\rho^d}u(r\sigma)\Big]\,\mathrm{d}r\Big|\,\mathrm{d}s(\sigma)\\ &\le \int_\Sigma \rho^{-1}\int_0^\rho r^d\big|\partial_r u(r\sigma)\big|\,\mathrm{d}r\,\mathrm{d}s(\sigma) + d\int_\Sigma \rho^{-1}\int_0^\rho r^{d-1}|u(r\rho)|\,\mathrm{d}r\,\mathrm{d}s(\sigma)\\ &\le \|\nabla u\|_{\mathrm{L}^1(B)} + \frac{d}{\rho}\|u\|_{\mathrm{L}^1(B)}.\end{aligned}$$

Combining all our estimates, we obtain

$$\|u\|_{\mathrm{L}^1(T)} \le \gamma^d\|u\|_{\mathrm{L}^1(B)} + \frac{\rho}{d}(\gamma^d-1)\|\nabla u\|_{\mathrm{L}^1(T)}.$$

Using the fact the $(u)_B = 0$, can employ the Poincaré inequality [1]

$$\|u\|_{\mathrm{L}^1(B)} \le \tfrac{1}{2}\mathrm{diam}(B)\|\nabla u\|_{\mathrm{L}^1(B)} \le \rho\|\nabla u\|_{\mathrm{L}^1(B)}, \tag{6.7}$$

which holds uniformly for all convex sets, to deduce the desired result. □

For microscopic elements it is advantageous to let the Clément operator coincide with the nodal interpolant. It will become apparent below that this should be done at least for any repatom $z \in \mathscr{N}$ which belongs to a microscopic face. It can be enforced by the condition

$$B_\xi \subset C_\xi \qquad \forall \xi \in \Omega_\mu.$$

We are now in a position to define and analyze the Clément operator. For each $z \in \mathscr{N}$ let ϕ_z be the associated hat-function, i.e., $\phi_z \in S^1_0(\mathscr{T})$ and $\phi_z(z') = \delta_{z,z'}$ for all $z, z' \in \mathscr{N}$. Note that $\{\phi_z : z \in \mathscr{N}\}$ is a partition of unity for $\bar\Omega$. We define the Clément operator by

$$\Pi\varphi(\xi) = \sum_{z\in\mathscr{N}} (\mathscr{R}\varphi)_{B_z}\phi_z(\xi) \qquad \forall \varphi \in \mathscr{V}. \tag{6.8}$$

To avoid a cluttered notation, we shall not distinguish between $\Pi\varphi$ and $\mathscr{R}\Pi\varphi$, i.e, $\Pi\varphi$ is a piecewise constant function defined on all of $\mathbb{R}^d$. Since it is awkward to estimate the error between φ and $\Pi\varphi$ directly, we also define the *continuous* Clément operator

$$\Pi_{\mathrm{c}}\varphi = \sum_{z\in\mathscr{N}} (\mathscr{R}\varphi)_{B_z}\phi_z, \tag{6.9}$$

which is a continuous, piecewise affine function of $\bar\Omega$. Rather than estimating $\|\mathscr{R}\varphi - \Pi\varphi\|_{\mathrm{L}^1}$ directly, we shall estimate $\|\mathscr{R}\varphi - \Pi_{\mathrm{c}}\varphi\|_{\mathrm{L}^1}$ and $\|\Pi_{\mathrm{c}}\varphi - \Pi\varphi\|_{\mathrm{L}^1}$.

Lemma 6.5 *(i) For each $z \in \mathscr{N}$ let $M_{P,z} = (\gamma_z^d + \gamma_z^d/d - 1/d)$, then,*

$$\|\mathscr{R}\varphi - \Pi_c\varphi(z)\|_{L^1(T_z)} \leq \rho_z M_{P,z} |D(\mathscr{R}\varphi)|(T_z). \tag{6.10}$$

(ii) Setting $M_{C,z} = d(1 + M_{P,z})$, we have

$$\sum_{\substack{e\in\mathscr{E}\\ z\in e}} \|\phi_z(\mathscr{R}\varphi - \Pi_c\varphi(z))\|_{L^1(e)} \leq M_{C,z} |D(\mathscr{R}\varphi)|(T_z). \tag{6.11}$$

(iii) The gradient $\nabla\Pi_c\varphi$ can be bounded by

$$\|\nabla\Pi_c\varphi\|_{L^1(\bar{\Omega})} \leq (d+1)\big[\max_{z\in\mathscr{N}} M_{P,z}\big]\, |D(\mathscr{R}\varphi)|(\bar{\Omega}). \tag{6.12}$$

Proof. The estimate (6.10) follows immediately from Lemma 6.4.

To prove (ii), for each $e \in \mathscr{E}$ such that $z \in e$, Lemma 6.6 below implies

$$\|\phi_z\,(\mathscr{R}\varphi - \Pi_c\varphi(z))\|_{L^1(e)} \leq \Big[\rho_z^{-1}\|\mathscr{R}\varphi - \Pi_c\varphi(z)\|_{L^1(\kappa_e)} + |D(\mathscr{R}\varphi)|(\kappa_e)\Big]. \tag{6.13}$$

For each $\kappa \subset T_z$ there are at most d faces $e \subset T_z$ for which $\kappa = \kappa_e$ appears in the above estimate and hence

$$\sum_{\substack{e\in\mathscr{E}\\ z\in e}} \|\phi_z(\mathscr{R}\varphi - \Pi_c\varphi(z))\|_{L^1(e)} \leq d\rho_z^{-1}\|\mathscr{R}\varphi - \Pi_c\varphi(z)\|_{L^1(T_z)} + d|D(\mathscr{R}\varphi)|(T_z).$$

Using (6.10), we obtain

$$\sum_{\substack{e\in\mathscr{E}\\ z\in e}} \|\phi_z(\mathscr{R}\varphi - \Pi_c\varphi(z))\|_{L^1(e)} \leq d(1 + M_{P,z})|D(\mathscr{R}\varphi)|(T_z)$$

which concludes the proof of (ii).

For the third part of the Lemma, note first that since $\{\phi_z : z \in \mathscr{N}\}$ is a partition of unity, it follows that

$$\sum_{z\in\mathscr{N}} \nabla\phi_z = \nabla \sum_{z\in\mathscr{N}} \phi_z = \nabla 1 = 0,$$

and hence

$$\begin{aligned} \nabla\Pi_c\varphi &= \nabla\Pi_c\varphi - \sum_{z\in\mathscr{N}} \nabla\phi_z\mathscr{R}\varphi \\ &= \sum_{z\in\mathscr{N}} \nabla\phi_z\big[\Pi_c\varphi(z) - \mathscr{R}\varphi\big]. \end{aligned} \tag{6.14}$$

We estimate $|\nabla\phi_z|_\infty$ by

$$|\nabla\phi_z|_\infty \le \max_{\kappa\subset T_z} \max_{z',z''\in\mathcal{N}\cap\kappa} |\phi_z(z') - \phi_z(z'')|/|z' - z''|.$$

Since ϕ_z vanishes at all nodes except z, it follows that

$$|\nabla\phi_z|_\infty \le \max_{z'\in\mathcal{N}\cap T_z} 1/|z' - z| \le \rho_z^{-1}.$$

Thus, taking the modulus and integrating over $\bar\Omega$ gives

$$\begin{aligned}
\|\nabla\Pi_{\rm c}\varphi\|_{{\rm L}^1(\bar\Omega)} &\le \sum_{z\in\mathcal{N}} \|\nabla\phi_z\|_{{\rm L}^\infty} \|\mathscr{R}\varphi - \Pi_{\rm c}\varphi(z)\|_{{\rm L}^1(T_z)} \\
&\le \sum_{z\in\mathcal{N}} \rho_z^{-1} M_{P,z}\rho_z |D(\mathscr{R}\varphi)|(T_z) \\
&\le (d+1)\big[\max_{z\in\mathcal{N}} M_{P,z}\big]\, |D(\mathscr{R}\varphi)|(\bar\Omega). \quad \square
\end{aligned}$$

Lemma 6.6 *Let $z\in\mathcal{N}$ and let e be a surface of κ which contains z. Then, for any $\psi\in{\rm BV}(\bar\Omega)$ there holds*

$$\|\phi_z\psi\|_{{\rm L}^1(e)} \le \rho_z^{-1}\|\psi\|_{{\rm L}^1(\kappa)} + |D\psi|(\kappa). \tag{6.15}$$

Proof. The proof of this Lemma is a modification of Lemmas 3.1 and 3.2 in [96], which cover the ${\rm H}^1$ case. Without loss of generality we assume again that ψ is smooth.

Let $\alpha\in\mathbb{R}^d$ be the unit vector pointing along the edge of the simplex κ which connects z and the face where ϕ_z vanishes. For each $x'\in e$, let $T(x')\ge 0$ be such that $x' + T(x')\alpha$ lies on that face. Then

$$\begin{aligned}
\big|\phi_z(x')\psi(x')\big| &= \Big|\int_{t=0}^{T(x')} \frac{\rm d}{{\rm d}t}\big[\phi_z(x'+t\alpha)\psi(x'+t\alpha)\big]\,{\rm d}t \\
&\le \int_{t=0}^{T(x')} \Big[|\partial_\alpha\phi_z(x'+t\alpha)||\psi(x'+t\alpha)| + |\phi_z(x'+t\alpha)||\partial_\alpha\psi(x'+t\alpha)|\Big]\,{\rm d}t \\
&\le |\partial_\alpha\phi_z| \int_{t=0}^{T(x')} |\psi(x'+t\alpha)|\,{\rm d}t + \int_{t=0}^{T(x')} |\nabla\psi(x'+t\alpha)|_1\,{\rm d}t.
\end{aligned}$$

Integrating with respect to x' over e gives

$$\|\phi_z\phi(x')\|_{{\rm L}^1(e)} \le |\partial_\alpha\phi_z|\|\psi\|_{{\rm L}^1(\kappa)} + \|\nabla\psi\|_{{\rm L}^1(\kappa)}.$$

By looking along the line $z + t\alpha$, $0\le t\le T(z)$, we see that $\partial_\alpha\phi_z$ is given by

$$|\partial_\alpha\phi_z| = T(\alpha)^{-1} \le \rho_z^{-1}. \quad \square$$

6.2 $\mathscr{V}^{1,1}$-Residual Estimate

We assume that no body forces are applied and that the exact Galerkin problem is solved, i.e., no summation rule is employed. We therefore analyze the residual E' which is given by the formula

$$E'(y;\varphi) = \sum_{\xi\in\Omega} \varepsilon^{d-1} E'_\xi(y)\cdot\varphi(\xi), \tag{6.16}$$

where $E'_\xi(y) = \varepsilon^{1-d}(\partial E/\partial y_\xi)(y)$, directly. To motivate the chosen scaling, consider the case of a pair potential energy (cf. §1.2.1) which can be written as

$$E(y) = \frac{1}{2}\sum_{\xi\in\Omega}\sum_{\xi'\in\Omega\setminus\{\xi\}} J(|y(\xi)-y(\xi')|).$$

Since the distance between nearest neighbours ξ and ξ' should be roughly ε and the number of atoms in the body, $\sharp\Omega \approx |\Omega|\varepsilon^{-d}$, in order to obtain a non-dimensional scaling of the energy, we rescale J by setting $J_\varepsilon(r) = \varepsilon^{-d}J(\varepsilon^{-1}r)$. This gives

$$E(y) = \frac{1}{2}\sum_{\xi\in\Omega}\sum_{\xi'\in\Omega} \varepsilon^d J_\varepsilon\big(\varepsilon^{-1}|y(\xi)-y(\xi')|\big).$$

By dropping the subscript in J_ε, the residual is given by

$$E'_\xi(y) = \frac{1}{2}\sum_{\xi'\in\Omega\setminus\{\xi\}} J'\big(\varepsilon^{-1}|y(\xi)-y(\xi')|\big)\frac{\xi-\xi'}{|\xi-\xi'|},$$

which is a non-dimensional term of order one. A similar argument can be given for EAM models.

Y is a critical point of the Galerkin approximation of E in $\tilde{\mathscr{A}}$ if

$$E'(Y;\Phi) = 0 \qquad \forall\Phi\in\tilde{\mathscr{A}}. \tag{6.17}$$

Therefore,

$$E'(Y;\varphi) = E'(Y;\varphi-\Pi\varphi) = \sum_{\xi\in\Omega}\varepsilon^{d-1}E'_\xi\cdot(\varphi(\xi)-\Pi\varphi(\xi)) \qquad \forall\varphi\in\mathscr{A},$$

where Π is defined by (6.8) and $E'_\xi = E'_\xi(Y)$, a notation which we adopt for the sake of brevity. Here it becomes apparent why we should insist on keeping the scaling parameter ε in the analysis. Otherwise, we might quite happily use a simple Hölder inequality to estimate the residual which would become of order ε^{-1}.

If the ground state Ω is a regular lattice in κ then we have shown in Proposition 1.2 that E'_ξ vanishes at atomistic sites which are in the bulk of κ. Similarly as in

Proposition 3.12, only interactions across element interfaces contribute to the residual. More generally, let us assume directly that, if $\bar{Y}$ is orientation preserving then

$$B(Y(\xi), z_c) \subset \bar{Y}(\kappa) \Rightarrow E'_\xi(Y) = 0 \tag{6.18}$$

holds. This assumption is still slightly strong since an atomistic deformation will not in general be orientation-preserving, particularly at defects. For concrete applications the statement should be localized.

Using (6.18), we may restrict the sum to those atomic sites where the residual is non-zero, and replace the sum over atoms by an integral over the respective Voronoi cells. Furthermore, we recall that $\Pi\varphi$ was constructed to coincide with φ in each cell $\bar{C}_\xi$ where $\xi \in \Omega_\mu$. Thus, we have

$$\begin{aligned} E'(Y;\varphi) &= \sum_{\xi\in\Omega_M} \varepsilon^{d-1} E'_\xi \cdot \big(\varphi(\xi) - \Pi\varphi(\xi)\big) \\ &= \sum_{\substack{\xi\in\Omega_M \\ E'_\xi\neq 0}} \varepsilon^{d-1} \frac{1}{|\bar{C}_\xi|} \int_{C_\xi} E'_\xi \cdot \big(\mathscr{R}\varphi(x) - \Pi\varphi(x)\big)\,\mathrm{d}x \\ &\leq \sum_{\substack{\xi\in\Omega_M \\ E'_\xi\neq 0}} \varepsilon^{-1} \int_{C_\xi} \frac{\varepsilon^d |E'_\xi|_\infty}{|\bar{C}_\xi|} \big|\mathscr{R}\phi - \Pi\phi\big|_1 \,\mathrm{d}x. \end{aligned} \tag{6.19}$$

As mentioned above, atoms in the bulk of an element have a zero residual. We shall therefore group all atoms such that each atom ξ with $E'_\xi \neq 0$ is associated to one or more faces.

While (except in some trivial cases) the faces $e \in \mathscr{E}$ have no immediate interpretation on the atomistic level, depending on the rate of compression of its neighbouring elements κ_e and possibly κ'_e, we can define a radius r_e and an 'atomistic face' $S_e \subset \bar{\Omega}$ (both depending on the deformation Y) such that the following conditions hold:

$$e \subset S_e \subset \{x' + t\nu_e : x' \in e, -r_e \leq t \leq r_e\} \qquad \forall e \in \mathscr{E}_M, \tag{6.20}$$

$$|S_e \cap S_{e'}| = 0 \qquad \forall e' \neq e \in \mathscr{E}_M, \quad \text{and} \tag{6.21}$$

$$\bigcup_{\substack{\xi\in\Omega_M \\ E'_\xi=0}} \bar{C}_\xi \subset \bigcup_{e\in\mathscr{E}_M} S_e. \tag{6.22}$$

The heights r_e should be chosen minimal ($\max_e r_e$ is minimal) subject to these conditions. It is clear that such radii r_e exist but they may be of order one which would be useless for our analysis.

With the help of the assumption (6.2), however, it is geometrically evident that we may choose

$$r_e = \varepsilon(z_c+1)\,\max\Big(\big|\nabla \bar{Y}_{\kappa_e}\nu_e\big|^{-1}, \big|\nabla \bar{Y}_{\kappa'_e}\nu_e\big|^{-1}\Big),$$

where the maximum is taken only over the first entry if e is a boundary surface.

Subdividing the sum (6.19) into sums over the atomistic faces S_e we obtain

$$E'(Y;\varphi) \;\leq\; \sum_{e\in\mathscr{E}_M}\Big[\max_{\xi\in\Omega\cap S_e}\frac{\varepsilon^d|E'_\xi|_\infty}{|\bar{C}_\xi|}\Big]\,\varepsilon^{-1}\int_{S_e}|\mathscr{R}\varphi-\Pi\varphi|\,\mathrm{d}x. \tag{6.23}$$

Note that while atoms may appear in the contribution from several faces, the sets over which the integrals are taken are disjoint. The estimate (6.23) almost seems like a sum of surface integrals and it is indeed possible to reduce the integrals in such a way. This is particularly straightforward if S_e is a simple cylinder and only slightly more technical in our situation.

Lemma 6.7 *Let e be a $(d-1)$-dimensional planar face and let S be a convex subset of its the cylinder $\{x\in\mathbb{R}^d : x = x' + t\nu_e, 0\le t\le r_e, x'\in e\}$, then*

$$\|\psi\|_{\mathrm{L}^1(S)} \le r_e\Big(\|\psi\|_{\mathrm{L}^1(e)} + |D\psi|(S)\Big) \qquad \forall\psi\in\mathrm{BV}(S).$$

Proof. Using the strict approximation of BV functions by smooth functions, we assume without loss of generality that $\psi\in\mathrm{C}^1(S)$.

Let $x = x' + r\nu_e \in S$, where $r\in[-r_e, r_e]$, then

$$\begin{aligned}
|\psi(x)| &= \Big|\psi(x') + \int_{t=0}^{r}\frac{\mathrm{d}}{\mathrm{d}t}\psi(x'+t\nu_e)\,\mathrm{d}t\Big| \\
&\le |\psi(x')| + \int_{t=0}^{r}|\partial_{\nu_e}\psi(x'+t\nu_e|\,\mathrm{d}t \\
&\le |\psi(x')| + \int_{t=0}^{r}|\nabla\psi(x'+\nu_e)|_1\,\mathrm{d}t.
\end{aligned} \tag{6.24}$$

For each $r\in\mathbb{R}$, let $e(r)$ be the convex subset of e such that $e + r\nu_e\in S$. Integrating (6.24) over $x'\in e(r)$, we obtain

$$\|\psi\|_{\mathrm{L}^1(e(r)+r\nu_e)} \le \|\psi\|_{\mathrm{L}^1(e)} + \int_{t=0}^{r}\int_{x'\in e(r)}|\nabla\psi(x'+t\nu_e)|_1\,\mathrm{d}s(x')\,\mathrm{d}t.$$

Integration over $r\in(0,r_e)$ gives

$$\begin{aligned}
\|\psi\|_{\mathrm{L}^1(S)} &\le r_e\|\psi\|_{\mathrm{L}^1(e)} + \int_{r=0}^{r_e}\int_{t=0}^{r}\int_{x'\in e(r)}|\nabla\psi(x'+t\nu_e)|_1\,\mathrm{d}s(x')\,\mathrm{d}t\,\mathrm{d}r \\
&= r_e\|\psi\|_{\mathrm{L}^1(e)} + \int_{t=0}^{r_e}\int_{r=t}^{r_e}\int_{x'\in e(r)}|\nabla\psi(x'+t\nu_e)|_1\,\mathrm{d}s(x')\,\mathrm{d}r\,\mathrm{d}t \\
&= r_e\|\psi\|_{\mathrm{L}^1(e)} + r_e\|\nabla\psi\|_{\mathrm{L}^1(S)}. \quad \square
\end{aligned}$$

To simplify notation, denote $m_e = \max_{\xi\in\Omega\cap S_e} \varepsilon^d |E'_\xi|_\infty/|\bar{C}_\xi|$. If we set $\rho_e = r_e/\varepsilon$ and apply Lemma 6.7 to (6.23) we would obtain

$$E'(Y;\varphi) \le \sum_{e\in\mathscr{E}_M} m_e \Big[2\rho_e \|\mathscr{R}\varphi - \Pi\varphi\|_{\mathrm{L}^1(e)} + \rho_e |D(\mathscr{R}\varphi - \Pi\varphi)|(S_e) \Big].$$

Thus, we would be required to obtain bounds on $|D\psi|(S_e)$ which seems a difficult task. It can be avoided, however, by first estimating (6.23) by

$$E'(Y;\varphi) \;\le\; \sum_{e\in\mathscr{E}_M} m_e\, \varepsilon^{-1} \Big[\|\mathscr{R}\varphi - \Pi_{\mathrm{c}}\varphi\|_{\mathrm{L}^1(S_e)} + \|\Pi_{\mathrm{c}}\varphi - \Pi\varphi\|_{\mathrm{L}^1(S_e)} \Big]$$

and applying Lemma 6.7 only to the first term, which gives

$$\begin{aligned} E'(Y;\varphi) \;\le\; \sum_{e\in\mathscr{E}_M} m_e \Big[& 2\rho_e \|\mathscr{R}\varphi - \Pi_{\mathrm{c}}\varphi\|_{\mathrm{L}^1(e)} \\ & + \rho_e |D(\mathscr{R}\varphi - \Pi_{\mathrm{c}}\varphi)|(S_e) + \varepsilon^{-1} \|\Pi_{\mathrm{c}}\varphi - \Pi\varphi\|_{\mathrm{L}^1(S_e)} \Big]. \end{aligned} \tag{6.25}$$

For the first term inside the sum of (6.25) we can use the continuum technique,

$$\begin{aligned} \|\mathscr{R}\varphi - \Pi_{\mathrm{c}}\varphi\|_{\mathrm{L}^1(e)} &= \int_e \Big| \mathscr{R}\varphi - \sum_{z\in\mathscr{N}\cap e} \phi_z (\mathscr{R}\varphi)_{T_z} \Big|_1 \,\mathrm{d}x \\ &\le \sum_{z\in\mathscr{N}\cap e} \int_e \big| (\mathscr{R}\varphi - (\mathscr{R}\varphi)_{T_z})\phi_z \big|_1 \,\mathrm{d}x \\ &\le \sum_{z\in\mathscr{N}\cap e} \|(\mathscr{R}\varphi - \Pi_{\mathrm{c}}\varphi(z))\,\phi_z\|_{\mathrm{L}^1(e)}. \end{aligned} \tag{6.26}$$

This construction prepares the estimate for an application of (6.11). The second and third terms are best estimated collectively over all of $\bar{\Omega}$. Since the sets S_e are not overlapping, we can write

$$\begin{aligned} E'(Y;\varphi) \;\le\; & \sum_{e\in\mathscr{E}_M} \Big[m_e\rho_e |D(\mathscr{R}\varphi - \Pi_{\mathrm{c}}\varphi)|(S_e) + m_e\varepsilon^{-1} \|\Pi_{\mathrm{c}}\varphi - \Pi\varphi\|_{\mathrm{L}^1(S_e)} \\ & \qquad + 2\rho_e m_e \sum_{z\in\mathscr{N}\cap e} \|(\mathscr{R}\varphi - \Pi_{\mathrm{c}}\varphi(z))\phi_z\|_{\mathrm{L}^1(e)} \Big] \end{aligned} \tag{6.27}$$

$$\begin{aligned} \le\; & \max_{e\in\mathscr{E}_M} \Big[2\rho_e m_e \Big] \cdot \Big[\tfrac{1}{2} |D(\mathscr{R}\varphi - \Pi_{\mathrm{c}}\varphi)|(\bar{\Omega}) + \tfrac{1}{2} \|\Pi_{\mathrm{c}}\varphi - \Pi\varphi\|_{\mathrm{L}^1(\bar{\Omega})} \\ & \qquad + \sum_{z\in\mathscr{N}} \sum_{\substack{e\in\mathscr{E} \\ z\in e}} \|(\mathscr{R}\varphi - \Pi_{\mathrm{c}}\varphi(z))\phi_z\|_{\mathrm{L}^1(e)} \Big]. \end{aligned} \tag{6.28}$$

We continue by estimating $\|\Pi_{\mathrm{c}}\varphi - \Pi\varphi\|_{\mathrm{L}^1(\bar{\Omega})}$ which we can rewrite as a sum over Voronoi cells. In each cell we use the following lemma.

Lemma 6.8 *Let T be star-shaped with respect to a point ξ and let $\psi \in \mathrm{C}^1(T)$, $\psi(\xi) = 0$. Then, setting $h_T = \sup_{x\in T} |x - \xi|$ we have*

$$\|\psi\|_{\mathrm{L}^1(T)} \leq h_T \|\nabla\psi\|_{\mathrm{L}^1(T)}.$$

Proof. Let Σ be the unit sphere in $\mathbb{R}^d$ and, for each $\sigma \in \Sigma$, let $r(\sigma) = \sup_{\xi+t\sigma\in T} t$. Then,

$$\begin{aligned}
\int_T |\psi(x)|\,\mathrm{d}x &= \int_{\sigma\in\Sigma}\int_{r=0}^{r(\sigma)} |\psi(\xi + r\sigma)|\,\mathrm{d}r\,\mathrm{d}s(\sigma) \\
&= \int_{\sigma\in\Sigma}\int_{r=0}^{r(\sigma)} \Big|\int_{t=0}^{r} \frac{\mathrm{d}}{\mathrm{d}t}\psi(\xi + t\sigma)\,\mathrm{d}t\Big|\,\mathrm{d}r\,\mathrm{d}s(\sigma) \\
&\leq \int_{\sigma\in\Sigma}\int_{r=0}^{r(\sigma)}\int_{t=0}^{r(\sigma)} |\nabla\psi(\xi + t\sigma)|_1\,\mathrm{d}t\,\mathrm{d}r\,\mathrm{d}s(\sigma) \\
&\leq h_T\|\nabla\psi\|_{\mathrm{L}^1(T)}. \quad \square
\end{aligned}$$

As a consequence of Lemma 6.8 and assumption (6.2) we obtain the estimate

$$\varepsilon^{-1}\|\Pi_{\mathrm{c}}\varphi - \Pi\varphi\|_{\mathrm{L}^1(\bar\Omega)} \leq \|\nabla\Pi_{\mathrm{c}}\varphi\|_{\mathrm{L}^1(\Omega)} \qquad \forall\varphi \in \mathscr{V}, \tag{6.29}$$

which allows us to tackle this term together with the first term in (6.28), for which we have

$$|D(\mathscr{R}\varphi - \Pi_{\mathrm{c}}\varphi)|(\bar\Omega) \leq |D(\mathscr{R}\varphi)|(\bar\Omega) + \|\nabla\Pi_{\mathrm{c}}\varphi\|_{\mathrm{L}^1(\bar\Omega)}.$$

To bound $\|\nabla\Pi_{\mathrm{c}}\varphi\|_{\mathrm{L}^1(\bar\Omega)}$ we can use Lemma 6.5 (iii) and we obtain

$$\tfrac{1}{2}|D(\mathscr{R}\varphi - \Pi_{\mathrm{c}}\varphi)|(\bar\Omega) + \tfrac{1}{2}\varepsilon^{-1}\|\Pi_{\mathrm{c}}\varphi - \Pi\varphi\|_{\mathrm{L}^1(\bar\Omega)} \leq \Big[\tfrac{1}{2} + (d+1)\max_{z\in\mathscr{N}} M_{P,z}\Big]\,|D(\mathscr{R}\varphi)|(\bar\Omega). \tag{6.30}$$

For the third term in (6.28), we use Lemma 6.5 (ii) to estimate

$$\begin{aligned}
\sum_{z\in\mathscr{N}}\sum_{\substack{e\in\mathscr{E}_M\\ z\in e}} \|(\mathscr{R}\varphi - \Pi_{\mathrm{c}}\varphi(z))\phi_z\|_{\mathrm{L}^1(e)} &\leq \big[\max_{z\in\mathscr{N}} M_{C,z}\big] \sum_{z\in\mathscr{N}} |D(\mathscr{R}\varphi)|(T_z) \\
&\leq (d+1)\big[\max_{z\in\mathscr{N}} M_{C,z}\big]\,|D(\mathscr{R}\varphi)|(\bar\Omega). \quad (6.31)
\end{aligned}$$

Combining (6.30) and (6.31) with (6.28) we finally obtain the following residual estimate:

$$\max_{\substack{\varphi\in\mathscr{V}\\ |\varphi|_{\mathscr{V}1,1}=1}} \big|E'(Y;\varphi)\big| \leq C(\mathscr{T}) \max_{e\in\mathscr{E}_M}\Big[\rho_e \max_{\xi\in\Omega\cap S_e} |E'_\xi(Y)|_\infty\Big]$$

where

$$C(\mathscr{T}) = \begin{cases} 5 + 18\max_{z\in\mathscr{N}}\gamma_z^2, & \text{if } d = 2, \\ 16 + 28\max_{z\in\mathscr{N}}\gamma_z^3, & \text{if } d = 3. \end{cases}$$

Remarks. 1. A large part of the analysis presented in this section seems quite sharp. Some particular improvements seem possible, however. First, the Poincaré inequality used to estimate $\|\psi-(\psi)_B\|_{\mathrm{L}^1(B)}$ in the proof of Lemma 6.4 is optimal only for sets which are essentially one-dimensional. Using a Poincaré inequality specifically designed for balls and sections of balls should improve the interpolation error estimates in Lemma 6.5 significantly. Second, it would be of a great advantage to move the constants $M_{P,z}$ and $M_{C,z}$ inside the maximum taken over edges rather than estimating them globally. For example, if the mesh is strongly graded near a fully refined region the constants $M_{P,z}$ and $M_{C,z}$ become fairly large. Since in such a region, the actual residual $E'_\xi(Y)$ should be small, it would be able to balance this effect and hide the large constants from the residual estimate. Many further points where minor improvements are possible can be found throughout the analysis of this and the previous section. ◀

2. To evaluate the error indicators $\max_{\xi\in S_e} \varepsilon^d|E'_\xi(Y)|_\infty/|\bar{C}_\xi|$, it should be sufficient to take a small representative cluster at each face and one at each node of the mesh and take the maximum over those sets only. It is not clear, however, how to make this approximation precise. ◀

3. To include body forces in the analysis, the residual should be split into two parts, as in the continuum analysis. While the first part, which corresponds to (6.16), can be treated as before, the second part requires interpolation error estimates on the elements which can be easily obtained using (6.10). ◀

4. Nowhere have we used the fact that the cells $\bar{C}_\xi$ related to the atomic sites ξ are Voronoi cells. We have not even used the fact that they are convex. It should therefore be possible to modify a Voronoi tesselation in order to obtain an improved lower bound on the constant c_V. ◀

5. Depending on the goal of the computation, it may be preferable to use different norms for the analysis of the residual. For example, if we wish to use a norm related to the $\mathrm{W}^{1,p}$-Sobolev-norms, $p \in (1,\infty]$, we could proceed as follows. Suppose that any two points $\xi, \xi' \in \Omega$ satisfy $|\xi - \xi'| \geq \varepsilon$ so that $B(\xi, \varepsilon/2)$ is contained in C_ξ. We can then use the smooth partition of unity

$$\psi_\xi(x) = \int_{C_\xi} (\varepsilon/2)^{-d}\eta\big((\varepsilon/2)^{-1}(x-y)\big)\,\mathrm{d}y,$$

where η is a standard mollifier with support in $B(0,1)$, to define

$$\mathscr{R}_S\varphi = \sum_{\xi\in\Omega} \varphi(\xi)\psi_\xi,$$

to construct a C^∞ lifting of the space $\mathscr{V}$. This should make it possible to analyze the residuals with respect to $W^{1,p}$-like norms for Sobolev indices $p > 1$ with the same techniques as above. ◀

6. Finally, it should be remarked that no summation rule approximation was analyzed. For a practical residual estimate this would have to be considered as well. ◀

6.3 Conclusion, Open Problems and Future Directions

This thesis has only scratched the surface of a potentially vast area of mathematics that has only begun to develop. Naturally, many questions have been left open. To conclude we list some of those questions and some further related problems which are interesting candidates for future research.

Inf-sup Constants. In order to fully generalize the *a priori* analysis of Chapter 4 and the *a posteriori* analysis of Chapter 5, it will be necessary to quantify the relevant inf-sup constants in higher dimensions. The continuous case can give us little guidance here. For example, there are several results showing, under suitable conditions, that operators of the form $u \mapsto -\mathrm{div}(A(x)\nabla u)$ are topological isomorphisms from $W_0^{1,p}(\Omega)$ to $(W_0^{1,p'}(\Omega))^*$. These are, however, usually restricted to specific values of p which certainly do not include $p = \infty$. Even for $p \in (1, \infty)$ the results are usually of an abstract form.

One direction for further research is therefore to quantify these results. It is possible, in several ways to rewrite the inf-sup problem as a linear program (a quadratic optimization problem with linear equality and inequality constraints) and try to solve it numerically. This should give us some indication as to what kind of results we might expect and to guide us for further research in this direction.

A-Priori Analysis. There are at least three promising possibilities for the extension of the *a priori* error analysis of Chapter 4. The first is based on inf-sup constants and, if these can be computed, is obvious.

A second possibility is based on the weighted norm technique of Rannacher and Scott [82]. By estimating the error of a Galerkin projections of a regularized Green's function they are able to prove optimal $W^{1,\infty}$ error estimates for piecewise linear finite

element approximations. However, their technique seems, at present, restricted to quasi-uniform meshes.

Also using the assumption of quasi-uniformity of the mesh it would be possible to use the ideas in [78]. Suppose that, for the linearized problem, we have an error estimate of the type $|y - Y|_{\mathscr{W}^{1,p}} \lesssim h^k$ where $|\cdot|_{\mathscr{W}^{1,p}}$ denotes a discrete $\mathrm{W}^{1,p}$-like semi-norm and h is the mesh size. Using norm equivalence in finite dimensional spaces, one can then show that, for quasi-uniform meshes, we have $|y - Y|_{\mathscr{W}^{1,\infty}} \lesssim h^{k-d/p}$. Thus, by raising the polynomial degree to $k = 2$ and letting $p = 2$, or by letting $k = 1$ and $p > d$, we can deduce a $\mathscr{W}^{1,\infty}$-bound on the error, which in turn, allows the application of the fixed point argument again.

The two latter ideas, while certainly more realistic to pursue, suffer from the assumption of quasi-uniformity of the mesh which is grossly violated for the quasicontinuum method.

Complete and Improved Residual Estimates. The analysis of §6.1 and §6.2 shows a clear path how residual estimates can be computed for the QC method in a very general setting. Combined with inf-sup estimates it is conceivable that the *a posteriori* analysis of Chapter 5 can be at least partially extended to higher dimensions. Of course, it is always possible, instead of using the *a posteriori* existence technique, to revert to assuming that an exact solution exists nearby. In this case it may be advantageous to use residuals with respect to dual norms other than the $\mathscr{W}^{1,1}$-norm. Remark 5 in §6.2 gives a clue as to how these might be obtained.

Goal Oriented Adaptivity. In engineering applications, QC simulations are usually performed with a specific goal in mind. It is often not important to obtain a good approximation to an exact solution in a particular norm but *only* an approximation to a certain quantity of interest. This may include an average displacement field, a critical force, or the energy of a dislocation. Reviews of different goal oriented adaptive techniques can be found in [7, 11, 14].

While an entirely rigorous analysis of goal oriented adaptive techniques is usually difficult, particularly for nonlinear problems, they have been demonstrated to provide highly efficient mesh refinement criteria for static problems.

Other Multiscale Methods. The QC method is only one example from the large class of multiscale methods for atomistic models for solids. Particularly for static

problems it provides maximal flexibility; however, other methods may be advantageous in specific situations.

For example, the *bridging-scales method* [62] constructs a coarse variable by defining a region $\Omega_\mu \subset \Omega$ (here, we understand Ω and Ω_μ as continuum domains) and decompose the deformation into a macroscopic deformation y_M and an additional microscopic displacement u_μ which is only computed in Ω_μ, i.e.,

$$y(\xi) = y_M(\xi) + u_\mu(\xi).$$

The relationship between y_M and y is controlled by a projection operator. This procedure has the advantage that no mesh grading is necessary in order to obtain a full atomistic region. An extension of the *a posteriori* analysis to this and other alternative methods would be an interesting project.

Finally, we should mention another direction which is not directly related to the work in this thesis but which cannot be neglected in any list of open problems in the field of multiscale methods for atomistic material models. As opposed to the QC method, the bridging-scales method is mainly used for dynamic simulations. The biggest difficulty is to obtain a correct continuum representation for the high frequency waves in an atomistic region and to thus be able to transfer the energy between Ω_μ and $\Omega_M = \Omega$. Mathematical problems originating from dynamical situations such as this, pose one of the biggest challenges for multiscale modelling.

Appendix A

Supplementary Material

Some basic background of linear and nonlinear analysis, function spaces as well as finite element methods is assumed in this thesis. While linear (functional) analysis cannot be covered here and is assumed as a prerequisite (see [85] or [98] for an elementary introduction), a brief introduction to functions spaces, the calculus of variations, and to finite element methods is given in the following three sections. While not strictly required to be able to follow the analysis in the main body of the thesis, it may provide a helpful background.

A.1 Function Spaces

Good introductions to Sobolev spaces can be found in introductory books on partial differential equations such as [45] or books on finite element methods [22]. A more detailed treatment can be found in [2] or [46]. The latter can also be used as an introduction to functions of bounded variation which are also used in the last chapter of this thesis. Another excellent reference on functions of bounded variation is [4]. The present section can only serve as a review and to fix the notation.

A.1.1 Sobolev spaces

Let Ω be a domain (an open, connected set) in $\mathbb{R}^d$. For any measurable set $A \subset \Omega$ we say $A \subset\subset \Omega$ if $\bar{A}$ is bounded and $\bar{A} \subset \Omega$. We define the set of test functions by

$$\mathscr{D}(\Omega) = \{\varphi \in \mathrm{C}^\infty(\Omega) : \mathrm{supp}(\varphi) \subset\subset \Omega\}.$$

Throughout, we shall identify any two measurable functions $u_1, u_2 : \Omega \to \mathbb{R}^m$ if $u_1(x) = u_2(x)$ for a.e. $x \in \Omega$. For each measurable function $u : \Omega \to \mathbb{R}^m$ we define

$$\begin{aligned} \|u\|_{\mathrm{L}^p(\Omega)} &= \left(\int_\Omega |u|_p^p \, \mathrm{d}x \right)^{1/p}, \qquad \text{for } p \in [1, \infty), \text{ and} \\ \|u\|_{\mathrm{L}^\infty(\Omega)} &= \operatorname*{ess.sup}_{x \in \Omega} |u(x)|_\infty, \end{aligned}$$

where $|\cdot|_p$ denotes the ℓ^p-norm on $\mathbb{R}^m$. We use $\mathrm{L}^p(\Omega)^m$ to denote the space of vector-valued measurable functions with finite L^p-norm. For scalar functions, the superscript is dropped. We also define $\mathrm{L}^1_{\mathrm{loc}}(\Omega)$ to be the set of scalar measurable functions u of Ω such that $\|u\|_{\mathrm{L}^1(C)}$ is finite for each set $C \subset\subset \Omega$. The fundamental theorem of the calculus of variations (this follows for example from [46, Section 4.2]) states that, for each $u \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$,

$$u = 0 \text{ for a.e. } x \in \Omega \text{ iff. } \int_\Omega u\varphi \, \mathrm{d}x = 0 \qquad \forall \varphi \in \mathscr{D}(\Omega). \tag{A.1}$$

A function $u \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ is said to be weakly differentiable if there exists a function $g \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)^d$ such that

$$\int_\Omega g \cdot \varphi \, \mathrm{d}x = -\int_\Omega u \, \mathrm{div}\varphi \, \mathrm{d}x \qquad \forall \varphi \in \mathscr{D}(\Omega)^d.$$

In this case, we use $\nabla u = g$ to denote its gradient and $\partial u / \partial x_j$ to denote the jth component of ∇u, $j = 1, \ldots, d$. If $u \in \mathrm{L}^1_{\mathrm{loc}}(\Omega)$ and is weakly differentiable, for $p \in [1, \infty]$, we define the Sobolev semi-norms

$$|u|_{\mathrm{W}^{1,p}(\Omega)} = \|\nabla u\|_{\mathrm{L}^p(\Omega)},$$

and the Sobolev norms

$$\begin{aligned} \|u\|_{\mathrm{W}^{1,p}(\Omega)} &= \left(\|u\|^p_{\mathrm{L}^p(\Omega)} + |u|^p_{\mathrm{W}^{1,p}(\Omega)} \right)^{1/p}, \qquad p \in [1, \infty), \text{ and} \\ \|u\|_{\mathrm{W}^{1,\infty}(\Omega)} &= \max\left(\|u\|_{\mathrm{L}^\infty(\Omega)}, |u|_{\mathrm{W}^{1,\infty}(\Omega)} \right). \end{aligned}$$

The spaces of locally integrable functions with finite Sobolev norms are denoted by $\mathrm{W}^{1,p}(\Omega)$. As is customary, we furthermore identify the symbols $\mathrm{H}^1 \equiv \mathrm{W}^{1,2}$. Higher order weak derivatives and the corresponding higher order Sobolev spaces can be defined analogously. Sobolev spaces are Banach spaces and, for $p \in (1, \infty)$ are reflexive [2, Theorem 3.5] (a Banach space $\mathscr{X}$ is called reflexive if its canonical embedding in $\mathscr{X}^{**}$ is a topological isomorphism).

Weakly differentiable functions share many properties with classically differentiable functions. The reason for this is that for many domains, classically differentiable functions are dense in Sobolev spaces; cf. for example [46, Section 4.2] or [45, Section 5.3].

A.1.2 The Dirichlet problem

As an application, consider the Dirichlet problem

$$-\Delta u + u = f \quad \text{in } \Omega, \quad u = 0 \text{ on } \partial\Omega, \tag{A.2}$$

where $f \in \mathrm{L}^2(\Omega)$. Upon multiplying by a test function, formally integrating by parts, and invoking (A.1), (A.2) can be shown to be formally equivalent to

$$\int_\Omega \Big[\nabla u \cdot \nabla \varphi + u\varphi\Big] \,\mathrm{d}x = \int_\Omega f\varphi \,\mathrm{d}x \qquad \forall \varphi \in \mathscr{D}(\Omega).$$

Let $\mathrm{H}^1_0(\Omega)$ be the closure of $\mathscr{D}(\Omega)$ with respect to the H^1-norm. Since both sides in the variational form are continuous in the topology induced by the H^1-norm, we may reformulate (A.2) as: *Find* $u \in \mathrm{H}^1_0(\Omega)$ *such that*

$$\int_\Omega \Big[\nabla u \cdot \nabla v + uv\Big] \,\mathrm{d}x = \int_\Omega fv \,\mathrm{d}x \qquad \forall v \in \mathrm{H}^1_0(\Omega), \tag{A.3}$$

which is called the weak form of the Laplace equation (as opposed to the strong form (A.2)). A closer look reveals that the bilinear form on the left-hand side of (A.3) is in fact an inner product on $\mathrm{H}^1_0(\Omega)$ which induces the H^1-norm. Let us denote this inner product by $(\cdot,\cdot)_{\mathrm{H}^1}$ so that we can rewrite (A.3) as $(u,v)_{\mathrm{H}^1} = \ell(v)$, where $\ell(v)$ is the bounded linear functional given by $\ell(v) = \int_\Omega fv \,\mathrm{d}x$. Hence, it follows from the Riesz representation theorem for $\mathrm{H}^1_0(\Omega)$ that (A.3) has a unique solution.

A.1.3 Functions of bounded variation

We also review functions of bounded variation which are used in Chapter 6. For $u \in \mathrm{L}^1(\Omega)$, and any open subset A of Ω the *total variation of* u *in* A is defined by

$$|Du|(A) = \sup_{\substack{\varphi \in \mathrm{C}^1(A), \operatorname{supp}(\varphi) \subset\subset A \\ \|\varphi\|_{\mathrm{L}^\infty(A)} \le 1}} \int_\Omega u \operatorname{div} \varphi \,\mathrm{d}x,$$

and the BV-norm in Ω by $\|u\|_{\mathrm{BV}(\Omega)} = \|u\|_{\mathrm{L}^1(\Omega)} + |Du|(\Omega)$. We say that a function $u \in \mathrm{L}^1(\Omega)$ has bounded variation if $\|u\|_{\mathrm{BV}(\Omega)} < +\infty$, and collect these functions into the space $\mathrm{BV}(\Omega)$. As is the case with Sobolev spaces, it turns out that $\mathrm{BV}(\Omega)$ is a Banach space. If $u \in \mathrm{BV}(\Omega)$ then there exists a vector-valued, signed Radon measure Du such that

$$-\int_\Omega u \operatorname{div} \varphi \,\mathrm{d}x = \int_\Omega \varphi \,\mathrm{d}Du \qquad \forall \varphi \in \mathrm{C}^1(\Omega),$$

i.e, the measure Du, called the variation of u, can be considered the weak (or better distributional) derivative of u (cf. [46, Section 5.1]). It can be shown to have the decomposition

$$Du = \nabla u \mathcal{L}^d + [u]\nu_u^\top \mathcal{H}^{d-1}|_{S_u} + D_c u,$$

where $\nabla u \in \mathrm{L}^1(\Omega)$, $\mathcal{L}^d$ is the d-dimensional Lebesgue measure, S_u is a $(d-1)$-dimensional set of 'weak jump discontinuities' with unit normal ν_u, $[u]$ is the jump across S_u, $\mathcal{H}^{d-1}$ is the $(d-1)$-dimensional Hausdorff surface measure, and $D_c u$ is a measure which is singular with respect to $\mathcal{L}^d$ as well as $\mathcal{H}^{d-1}|_{S_u}$ and is called the Cantor part of Du. Piecewise $\mathrm{W}^{1,1}$ functions have bounded variation without Cantor part.

Unlike Sobolev functions, BV functions cannot be approximated in the BV-norm by smooth functions. However, they can be approximated by smooth functions in the strict topology of BV, which is often sufficient. Let the metric ρ be defined by

$$\rho(u,v) = \|u-v\|_{\mathrm{L}^1(\Omega)} + \big||Du|(\Omega) - |Dv|(\Omega)\big| \qquad \forall u, v \in \mathrm{BV}(\Omega).$$

Equipped with ρ, BV is a metric space and the topology induced by ρ is called the strict topology. For any domain Ω, for any $u \in \mathrm{BV}(\Omega)$ there exists a sequence $u_j \in \mathrm{C}^\infty(\Omega) \cap \mathrm{BV}(\Omega)$ such that $\rho(u, u_j) \to 0$ as $j \to \infty$ [46, Section 5.2.2].

A.2 Calculus of Variations

To fix the notation, we assume that $\mathscr{X}$ and $\mathscr{Y}$ are separable Banach spaces (i.e. Banach spaces which contain a countable set which is dense) with topological duals $\mathscr{X}^*$ and $\mathscr{Y}^*$ (the spaces of bounded linear functionals from respectively $\mathscr{X}$ or $\mathscr{Y}$ to $\mathbb{R}$). The norm associated with a Banach space $\mathscr{X}$ is denoted by $\|\cdot\|_{\mathscr{X}}$, and so forth. The space of bounded linear mappings between $\mathscr{X}$ and $\mathscr{Y}$ is denoted by $L(\mathscr{X}, \mathscr{Y})$. In the following sections we give a brief introduction to the most basic methods of the calculus of variations.

A nice elementary introduction to the calculus of variations can be found in [21]. For more advanced material, particularly for applications in solid mechanics, see [35] or [79].

A.2.1 The direct method

Let $\mathscr{X}^*$ be a dual Banach space, i.e., let $\mathscr{X}^*$ be the dual of a Banach space $\mathscr{X}$, and let $\phi : \mathscr{X} \to (-\infty, +\infty]$. We wish to find

$$u \in \operatorname*{argmin}_{\mathscr{X}} \phi. \tag{A.4}$$

A powerful technique to prove the existence of solutions to (A.4) is the *direct method of the calculus of variations*.

Suppose that ϕ satisfies the coercivity condition

$$\|u\|_{\mathscr{X}^*} \to \infty \Rightarrow \phi(u) \to \infty. \tag{A.5}$$

Let $(u_j)_{j\in\mathbb{N}} \subset \mathscr{X}$ be a minimizing sequence for ϕ, i.e,

$$\phi(u_j) \to \inf_{\mathscr{X}} \phi \quad \text{as } j \to \infty.$$

If the domain of ϕ, $D(\phi) = \{u \in \mathscr{A} : \phi(u) < +\infty\}$ is non-empty, then each minimizing sequence must be bounded. By the Banach–Alaoglu theorem [85, Theorem 3.15], it follows that (u_j) is precompact in the weak-$*$ topology of $\mathscr{X}^*$, i.e., there exists $u \in \mathscr{X}^*$ and a subsequence (w.l.o.g. not relabelled) such that $u_j \overset{*}{\rightharpoonup} u$ weakly-$*$ in $\mathscr{X}^*$.

At this point the crucial assumption for the direct method comes into play. If ϕ is *sequentially weakly-$*$ lower semicontinuous*, i.e.,

$$\phi(v) \leq \liminf_{j\to\infty} \phi(v_j),$$

whenever $v_j \overset{*}{\rightharpoonup} v$, then it follows that u, constructed above, is a solution to (A.4).

A.2.2 Euler–Lagrange equations

Let $\mathscr{X}, \mathscr{Y}$ be normed, linear spaces, let $\mathscr{A}$ be an open subset of $\mathscr{X}$ and let $\mathscr{F} : \mathscr{A} \to \mathscr{Y}$. $\mathscr{F}$ is said to be Gateaux-differentiable at $u \in \mathscr{A}$ if there exists a map $T \in L(\mathscr{X}, \mathscr{Y})$ such that, for every $v \in \mathscr{X}$,

$$\lim_{\substack{h\to 0 \\ h\in\mathbb{R}}} |h|^{-1} \|\mathscr{F}(u+hv) - \mathscr{F}(u) - hTv\|_{\mathscr{Y}} = 0.$$

In this case, we write $T = \mathscr{F}'(u)$. $\mathscr{F}$ is said to be Fréchet differentiable if it is Gateaux differentiable and

$$\lim_{\substack{v\to 0 \\ v\in\mathscr{X}}} \|v\|_{\mathscr{X}}^{-1} \|\mathscr{F}(u+v) - \mathscr{F}(u) - \mathscr{F}'(u)v\|_{\mathscr{Y}} = 0.$$

A classical technique in the calculus of variations is to make use of Banach space differentiation to derive necessary conditions for solutions to (A.4). Let $\phi : \mathscr{X} \to \mathbb{R}$, and let $\mathscr{A}$ be a closed, convex subset of $\mathscr{X}$. Suppose that $u \in \mathscr{A}$ is a local minimizer of ϕ in $\mathscr{A}$ (i.e., locality is understood in the topology of $\mathscr{X}$ with respect to which ϕ is differentiable) and that ϕ is Gateaux differentiable at u. Then, for all $v \in \mathscr{A}$, we have

$$\phi(u) \leq \phi(v) = \phi(u) + h\phi'(u; v-u) + o(h) \quad \text{as } h \to 0.$$

Letting $h \to 0$, it follows that

$$\phi'(u; v - u) \geq 0 \qquad \forall v \in \mathscr{A}. \tag{A.6}$$

(A.6) is the prototype of a variational inequality. If $\mathscr{A}$ is an affine subspace then we have equality in (A.6). In this case, the resulting equation $\phi'(u) = 0$ is called the Euler–Lagrange equation of ϕ.

More generally, a point u satisfying (A.6) is called a critical point. If ϕ is convex then the the set of critical points and the set of global minimizers coincide. In general, the structure of critical points can be quite complicated. However, as in the finite-dimensional case, we can at least say the following: if $\phi'(u) = 0$, if ϕ' is Fréchet differentiable at u and $\phi''(u)$ is positive definite, then u is a strict local minimum of ϕ.

In practise, the derivation of the Euler–Lagrange equation can be subtle. Consider, for example, the integral functional

$$\phi(u) = \int_\Omega f(u(x), \nabla u(x)) \, \mathrm{d}x,$$

where Ω is an open subset of $\mathbb{R}^d$ and $f = f(u, p)$ is continuously differentiable in $\mathbb{R} \times \mathbb{R}^d$. Clearly, ϕ is well-defined and differentiable on the function space $\mathrm{W}^{1,\infty}(\Omega)$. The resulting Euler–Lagrange equations are

$$\int_\Omega \Big[f_u(u, \nabla u) v + f_p(u, \nabla u) \cdot \nabla v \Big] \, \mathrm{d}x = 0 \qquad \forall v \in \mathrm{W}^{1,\infty}(\Omega). \tag{A.7}$$

In many situations, (A.7) does not have a solution in $\mathrm{W}^{1,\infty}(\Omega)$. In which other topologies can ϕ then be differentiated? For example, if f satisfies the growth conditions

$$|f(u, g)| \leq C_1(1 + |u|^p + |g|^p) \quad \text{and} \quad |f_u(u, g)| + |f_g(u, g)| \leq C_1(1 + |u|^{p-1} + |g|^{p-1})$$

then it can be shown that ϕ is Gateaux differentiable in $\mathrm{W}^{1,p}(\Omega)$.

A.3 Finite Element Methods

Finite element methods are flexible techniques for the numerical approximation of partial differential equations and many other types of mathematical models. For a thorough introduction see [22].

Let Ω be a polygonal domain in $\mathbb{R}^d$ and let $\mathcal{T}$ be a *finite element mesh*, a collection of open simplices $\kappa \subset \mathbb{R}^d$ such that $\cup_{\kappa \in \mathcal{T}} \bar{\kappa} = \bar{\Omega}$ and $\kappa \cap \kappa' = \emptyset$ if $\kappa \neq \kappa' \in \mathcal{T}$. The construction and the specific properties of the mesh $\mathcal{T}$ are crucial to the success of the finite element method, both analytically and algorithmically. For example, in

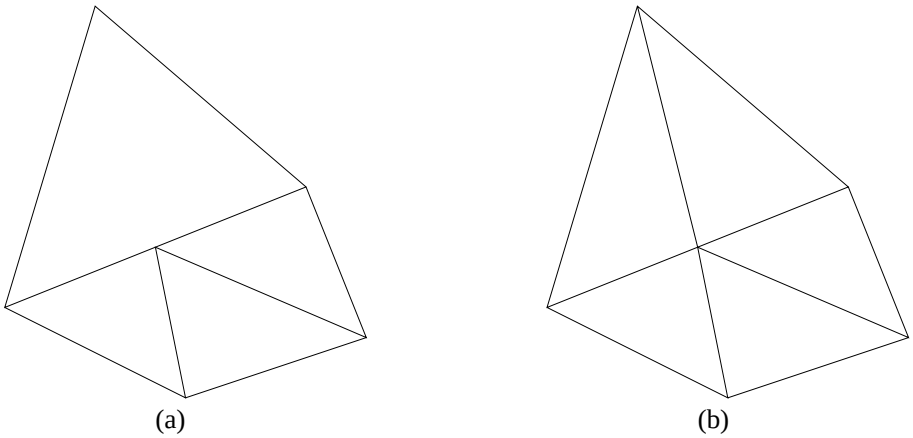

Figure A.1: A mesh with a hanging node (a) and a regular mesh (b).

two dimensions, it is typically assumed that the mesh has no hanging nodes, i.e., if $\bar{\kappa} \cap \bar{\kappa}' \neq \emptyset$ then either, this intersection consists of exactly one common edge of κ and κ', or of one common vertex (cf. Figure A.1). In three dimensions, if $\bar{\kappa} \cap \bar{\kappa}' \neq \emptyset$, then the intersection must consist of exactly one common face, of one common edge or of one common vertex. Such meshes are called regular. Irregular meshes may also be used, but their discussion goes beyond the scope of this appendix.

For $k \in \mathbb{N} \setminus \{0\}$ let P_k be the set of polynomials of degree up to k in $\mathbb{R}^d$ and let $S^k(\mathcal{T})$ be defined by

$$S^k(\mathcal{T}) = \big\{ v \in \mathrm{C}(\bar{\Omega}) : v|_\kappa \in P_k \qquad \forall\, \kappa \in \mathcal{T} \big\}.$$

It should be intuitively clear that this definition is not very useful for arbitrary partitions. Only if a mesh is regular (Figure A.1 (b)) or has a very specific structure, is $S^k(\mathcal{T})$ rich enough to approximate Sobolev spaces.

As a first example of a finite element method we discretize (A.3). To this end, let $S_0^k(\mathcal{T}) = S^k(\mathcal{T}) \cap \mathrm{H}_0^1(\Omega)$. The Galerkin finite element approximation to (A.3) is to *find* $U \in S_0^k(\mathcal{T})$ *such that*

$$\int_\Omega \big[\nabla U \cdot \nabla V + UV \big] \,\mathrm{d}x = \int_\Omega fV \,\mathrm{d}x \qquad \forall\, V \in S_0^k(\mathcal{T}), \tag{A.8}$$

or, in short, $(U, V)_{\mathrm{H}^1} = \ell(V)$. By the Riesz representation theorem for $S_0^k(\mathcal{T})$, (A.8) has a unique solution. Let u be the solution to (A.3) and let U be the solution to (A.8). Since $S_0^k(\mathcal{T}) \subset \mathrm{H}_0^1(\Omega)$, we have

$$(u - U, V)_{\mathrm{H}^1} = (u, V)_{\mathrm{H}^1} - (U, V)_{\mathrm{H}^1} = \ell(V) - \ell(V) = 0 \qquad \forall V \in S_0^k(\mathcal{T})$$

and hence

$$\|u-U\|_{\mathrm{H}^1}^2 = (u-U, u-U)_{\mathrm{H}^1} = (u-U, u-V)_{\mathrm{H}^1} \leq \|u-U\|_{\mathrm{H}^1}\|u-V\|_{\mathrm{H}^1}$$

for all $V \in S_0^k(\mathcal{T})$. Dividing by $\|u-U\|_{\mathrm{H}^1}$, we obtain the best approximation error estimate

$$\|u-U\|_{\mathrm{H}^1} = \inf_{V \in S_0^k} \|u-V\|_{\mathrm{H}^1}. \tag{A.9}$$

In general, a Galerkin finite element method is not a projection as in the case of the Dirichlet problem, but estimates similar to (A.9) still hold in many situations.

With (A.9), the error estimation is reduced to an approximation problem. We shall not go into any detail in this area but only mention that this is the point in the finite element analysis where the specific construction of the mesh, for example its regularity, is of crucial importance. For further material see, for example, [22, Chapter 4]. For error estimation with respect to different norms, see [22, Chapter 8].

In the following two sections, we briefly outline two techniques which can be used to analyze the finite element method for nonlinear problems.

A.3.1 Error estimates for nonlinear equations

Consider the partial differential equation

$$-\operatorname{div} S(\nabla u) = f \ \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega. \tag{A.10}$$

Suppose furthermore that S is uniformly elliptic, i.e.,

$$(S(F)-S(G)) : (F-G) \geq M_1|F-G|^2,$$

and Lipschitz continuous, i.e., $|S(F)-S(G)| \leq L|F-G|$, where $M_1 > 0$ and $L \in \mathbb{R}$. A possible weak form for (A.10) is to *find* $u \in \mathrm{H}_0^1(\Omega)$ *such that*

$$B(u;v) = \ell(v) \qquad \forall v \in \mathrm{H}_0^1, \text{ where} \tag{A.11}$$
$$B(u;v) = \int_\Omega S(\nabla u) : \nabla v \,\mathrm{d}x \quad \text{and} \quad \ell(v) = \int_\Omega f v \,\mathrm{d}x.$$

The finite element Galerkin discretization to (A.11) is to *find* $U \in S_0^k(\mathcal{T})$ *such that*

$$B(U;V) = \ell(V) \qquad \forall V \in S_0^k(\mathcal{T}).$$

We leave the solubility of (A.10) and (A.11) aside but note that both equations can be shown to posses unique solutions. By a generalization of the argument in the linear

case, we have

$$\begin{aligned} M_1|u-U|^2_{\mathrm{H}^1} &\leq \big|B(u;u-U)-B(U;u-U)\big| \\ &= \big|B(u;u-V)-B(U;u-V)\big| \\ &\leq L|u-U|_{\mathrm{H}^1}|u-V|_{\mathrm{H}^1}, \end{aligned}$$

and thus obtain the quasi-optimal error estimate

$$|u-U|_{\mathrm{H}^1} \leq \frac{L}{M_1} \inf_{V\in S^k_0(\mathcal{T})} |u-V|_{\mathrm{H}^1}.$$

If S is not *globally* elliptic and/or Lipschitz continuous, but only *locally* in a convex set $\mathcal{M}\subset\mathbb{R}^d$, then, by assuming that $\nabla u, \nabla U \in \mathcal{M}$ the same results can be recovered. In fact, in some situations, this assumption can be made rigorous by use of a fixed point argument such as the one in Chapter 4. This would go beyond the scope of this section, however, and the reader is referred to [63] and [78] for further examples.

A.3.2 Variational convergence analysis

The direct method outlined in §A.2.1 has the great advantage that only the energy is required for its analysis. This makes it usually clear which function space (namely one in which ϕ is coercive) one should work in and considerably simplifies the analysis.

To motivate the following discussion note that (A.3) is the Euler–Lagrange equation of the Dirichlet functional in $\mathrm{H}^1_0(\Omega)$ while (A.8) is the Euler Lagrange equation of the Dirichlet functional in $S^k_0(\mathcal{T})$. This is generally true: the finite element discretization of an Euler–Lagrange equation is 'usually' the Euler–Lagrange equation of the same functional restricted to the finite element space.

Thus, let $\phi : \mathrm{W}^{1,p}(\Omega) \to (-\infty,+\infty]$, where $p\in(1,\infty)$, be a strongly continuous, sequentially weakly (and since $\mathrm{W}^{1,p}(\Omega)$ is reflexive, also weakly-$*$) lower semi-continuous and coercive functional, i.e.,

$$\begin{aligned} u_j \to u &\;\Rightarrow\; \phi(u_j)\to\phi(u), \\ u_j \rightharpoonup u &\;\Rightarrow\; \phi(u)\leq \liminf_{j\to\infty}\phi(u_j), \quad \text{and} \\ \|u_j\|_{\mathrm{W}^{1,p}}\to\infty &\;\Rightarrow\; \phi(u_j)\to\infty. \end{aligned}$$

Furthermore, for each $N\in\mathbb{N}$, let $\mathcal{T}_N$ be a finite element mesh such that

$$\forall\, u\in \mathrm{W}^{1,p}\ \exists V_N\in S^k(\mathcal{T}_N), N\in\mathbb{N} : V_N\to u \ \text{ in } \mathrm{W}^{1,p}.$$

The *energy Galerkin finite element method* is to find $U_N\in S^k(\mathcal{T}_N)$ such that $\phi(U_N) = \inf_{S^k(\mathcal{T}_N)}\phi$. By the direct method, such U_N exist and by the coercivity of ϕ we may

extract a subsequence U_{N_j}, $N_j \uparrow \infty$, such that $U_{N_j} \rightharpoonup u$ for some $u \in \mathrm{W}^{1,p}(\Omega)$. We shall now prove that u is a minimizer of ϕ in $\mathrm{W}^{1,p}(\Omega)$. To this end, let $v \in \mathrm{W}^{1,p}(\Omega)$ and let $V_N \in S^k(\mathcal{T}_N)$, $V_N \to v$ as $N \to \infty$, and consider

$$\phi(u) \leq \liminf_{j\to\infty} \phi(u_{N_j}) \leq \limsup_{j\to\infty} \phi(U_{N_j}) \leq \lim_{j\to\infty} \phi(V_{N_j}) = \phi(v).$$

Since v was arbitrary, it follows that $u \in \operatorname{argmin} \phi$. Furthermore, by setting $v = u$, it follows that all inequalities in the above chain are equalities and hence $\phi(U_{N_j}) \to \phi(u)$. The result can be strengthened in several ways. For example, if the minimizer of ϕ is unique then the entire sequence U_N converges (weakly). In some cases (depending on the specific structure of ϕ) it is even possible to deduce strong convergence of the sequence U_{N_j} (or U_N). To demonstrate this, we may consider the p-Laplacian energy functional

$$\phi(u) = \int_\Omega \Big[|\nabla u|^p - fu \Big] \,\mathrm{d}x,$$

restricted to $\mathrm{W}^{1,p}_0(\Omega)$, where $f \in \mathrm{L}^{p'}(\Omega)$ and $p \in (1,\infty)$. If $U_{N_j} \rightharpoonup u$ weakly in $\mathrm{W}^{1,p}(\Omega)$ and $\phi(u_{N_j}) \to \phi(u)$ then, in particular, $\|\nabla U_{N_j}\|_{\mathrm{L}^p} \to \|\nabla u\|_{\mathrm{L}^p}$ which, together with the stated weak convergence implies strong the convergence $\|\nabla(u - U_{N_j})\|_{\mathrm{L}^p} \to 0$ (using Clarkson's Theorem that L^p is uniformly convex [31] and a straightforward computation). The reader is referred to [75] for more detail and further concrete examples.

The argument outlined in this section is a special case of Γ-convergence [37, 39]. While undeniably elegant, it has a crucial flaw that is not usually mentioned in numerical works based on Γ-convergence. It requires the global minimization of the energy ϕ in the discrete spaces $S^k(\mathcal{T})$ which, unless ϕ is convex, is generally not tractable.

References

[1] G. Acosta and R. G. Durán. An optimal Poincaré inequality in L^1 for convex domains. *Proc. Amer. Math. Soc.*, 132(1):195–202, 2004.

[2] R. A. Adams. *Sobolev spaces*. Academic Press, New York-London, 1975. Pure and Applied Mathematics, Vol. 65.

[3] M. Ainsworth and J. T. Oden. *A Posteriori Error Estimation in Finite Element Analysis*. Pure and Applied Mathematics (New York). Wiley-Interscience [John Wiley & Sons], New York, 2000.

[4] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of Bounded Variation and Free Discontinuity Problems*. Oxford Mathematical Monographs. The Clarendon Press Oxford University Press, New York, 2000.

[5] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient Flows in Metric Space and in the Space of Probability Measures*. Birkhäuser Verlag, 2005.

[6] B. M. Axilrod and E. Teller. Interaction of van der Waals type between three atoms. *J. Chem. Phys.*, 11:299–300, 1943.

[7] I. Babuska and T. Strouboulis. *The Finite Element Method and its Reliability*. Oxford University Press, 2001.

[8] J. M. Ball. Convexity conditions and existence theorems in nonlinear elasticity. *Arch. Rat. Mech. Anal.*, 63, 1977.

[9] J. M. Ball, P. J. Holmes, R. D. James, R. L. Pego, and P. J. Swart. On the dynamics of fine structure. *J. Nonlinear Sci.*, 1(1):17–70, 1991.

[10] J. M. Ball and R. D. James. Fine phase mixtures as minimizers of energy. *Arch. Rational Mech. Anal.*, 100(1):13–52, 1987.

[11] W. Bangerth and R. Rannacher. *Adaptive Finite Element Methods for Differential Equations*. Birkhäuser Verlag, 2003.

[12] S. Bartels. *Numerical Analysis of Some Non-Convex Variational Problems*. PhD thesis, Christian–Albrechts-Universität zu Kiel, 2001.

[13] S. Bartels. A posteriori error analysis for time-dependent Ginzburg–Landau type equations. *Numer. Math.*, 99(4):557–583, 2005.

[14] R. Becker and R. Rannacher. An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numer.*, 10:1–102, 2001.

[15] X. Blanc, C. Le Bris, and F. Legoll. Analysis of a prototypical multiscale method coupling atomistic and continuum mechanics. *M2AN Math. Model. Numer. Anal.*, 39(4):797–826, 2005.

[16] X. Blanc, C. Le Bris, and P.-L. Lions. From molecular models to continuum mechanics. *Arch. Ration. Mech. Anal.*, 164(4):341–381, 2002.

[17] B. Bourdin, G. A. Francfort, and J.-J. Marigo. Numerical experiments in revisited brittle fracture. *J. Mech. Phys. Solids*, 48(4):797–826, 2000.

[18] A. Braides, G. Dal Maso, and A. Garroni. Variational formulation of softening phenomena in fracture mechanics: the one-dimensional case. *Arch. Ration. Mech. Anal.*, 146(1):23–58, 1999.

[19] A. Braides and M. S. Gelli. Continuum limits of discrete systems without convexity hypotheses. *Math. Mech. Solids*, 7(1):41–66, 2002.

[20] A. Braides, A. Lew, and M. Ortiz. Effective cohesive behavior of layers of interatomic planes. *Arch. Ration. Mech. Anal.*, 180(2):151–182, 2006.

[21] U. Brechtken-Manderscheid. *Introduction to the calculus of variations*. Chapman and Hall Mathematics Series. Chapman & Hall, London, 1991. Translated from the German by P. G. Engstrom.

[22] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*, volume 15 of *Texts in Applied Mathematics*. Springer-Verlag, New York, second edition, 2002.

[23] F. Brezzi, J. Rappaz, and P.-A. Raviart. Finite-dimensional approximation of nonlinear problems. I. Branches of nonsingular solutions. *Numer. Math.*, 36(1):1–25, 1980/81.

[24] A. Buffa and C. Ortner. Variational convergence of DGFEM. work in progess.

[25] R. H. Byrd, J. Nocedal, and R. A. Waltz. KNITRO: An integrated package for nonlinear optimization. In *Large-scale nonlinear optimization*, volume 83 of *Nonconvex Optim. Appl.*, pages 35–59. Springer, New York, 2006.

[26] C. Carstensen and G. Dolzmann. Time-space discretization of the nonlinear hyperbolic system $u_{tt} = \operatorname{div}(\sigma(Du) + Du_t)$. *SIAM J. Numer. Anal.*, 42(1):75–89, 2004.

[27] C. Carstensen, K. Hackl, and A. Mielke. Non-convex potentials and microstructures in finite-strain plasticity. *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.*, 458(2018):299–317, 2002.

[28] C. Carstensen and P. Plecháč. Numerical solution of the scalar double-well problem allowing microstructure. *Math. Comp.*, 66(219):997–1026, 1997.

[29] M. Charlotte and L. Truskinovsky. Linear elastic chain with a hyper-pre-stress. *J. Mech. Phys. Solids*, 50(2):217–251, 2002.

[30] P. G. Ciarlet. *Mathematical Elasticity, Volume I: Three-Dimensional Elasticity.* North Holland, 1988.

[31] James A. Clarkson. Uniformly convex spaces. *Trans. Amer. Math. Soc.*, 40(3):396–414, 1936.

[32] S. Conti, G. Dolzmann, B. Kirchheim, and S. Müller. Sufficient conditions for the validity of the Cauchy–Born rule close to $SO(n)$. Technical Report 85/2005, Max Planck Institute für Mathematik in den Naturwissenschaften, 2005.

[33] B. A. Coomes, H. Koçak, and K. J. Palmer. Rigorous computational shadowing of orbits of ordinary differential equations. *Numer. Math.*, 69(4):401–421, 1995.

[34] W. A. Curtin and R. E. Miller. Atomistic/continuum coupling in computational materials science. *Modelling Simul. Mater. Sci. Eng.*, 11:33–68, 2003.

[35] B. Dacorogna. *Direct Methods in the Calculus of Variations*, volume 78 of *Applied Mathematical Sciences*. Springer-Verlag, Berlin, 1989.

[36] G. Dal Maso and R. Toader. A model for the quasi-static growth of brittle fractures based on local minimization. *Math. Models Methods Appl. Sci.*, 12(12):1773–1799, 2002.

[37] G. DalMaso. *An Introduction to* Γ*-Convergence*. Birkhäuser, Boston, 1993.

[38] M. S. Daw and M. I. Baskes. Embedded-atom method: Derivation and application to impurities, surfaces, and other defects in metals. *Physical Review B*, 20, 1984.

[39] E. De Giorgi and T. Franzoni. Su un tipo di convergenza variazionale. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur. (8)*, 58(6):842–850, 1975.

[40] M. Dobrowolski and R. Rannacher. Finite element methods for nonlinear elliptic systems of second order. *Math. Nachr.*, 94:155–172, 1980.

[41] W. E and B. Engquist. The heterogeneous multiscale methods. *Commun. Math. Sci.*, 1(1):87–132, 2003.

[42] W. E and P. Ming. Analysis of the local quasicontinuum method. Preprint.

[43] W. E and P. Ming. Analysis of multiscale methods. *J. Comput. Math.*, 22(2):210–219, 2004. Special issue dedicated to the 70th birthday of Professor Zhong-Ci Shi.

[44] W. E, P. Ming, and P. Zhang. Analysis of the heterogeneous multiscale method for elliptic homogenization problems. *J. Amer. Math. Soc.*, 18(1):121–156, 2005.

[45] L. C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998.

[46] L. C. Evans and R. F. Gariepy. *Measure Theory and Fine Properties of Functions*. Studies in Advanced Mathematics. CRC Press, Boca Raton, FL, 1992.

[47] M. Fago, R. L. Hayes, E. A. Carter, and M. Ortiz. Density-functional-theory-based local quasicontinuum method: Prediction of dislocation nucleation. *Physical Review B (Condensed Matter and Materials Physics)*, 70(10):100102, 2004.

[48] G. A. Francfort and J.-J. Marigo. Revisiting brittle fracture as an energy minimization problem. *J. Mech. Phys. Solids*, 46(8):1319–1342, 1998.

[49] G. Friesecke and F. Theil. Validity and failure of the Cauchy–Born hypothesis in a two-dimensional mass-spring lattice. *J. Nonlinear Sci.*, 12(5):445–478, 2002.

[50] C.W. Gear, J.M. Hyman, I.G. Kevrekidis, P.G. Kevrekidis O. Runborg, and C. Theodoropoulos. Equation-free, coarse-grained multiscale computation: enabling microscopic simulators perform system-level tasks. Preprint, 2006.

[51] W. Gear, I. G. Kevrekidis, and C. Theodoropoulos. "Coarse" integration/ bifurcation analysis via microscopic simulator: micro-Galerkin methods. *Comp. Chem. Engng.*, 26:941–963, 2002.

[52] N. I. M. Gould, D. Orban, and P. L. Toint. GALAHAD, a library of thread-safe Fortran 90 packages for large-scale nonlinear optimization. *ACM Trans. Math. Software*, 29(4):353–372, 2003.

[53] D. J. Higham. Trust region algorithms and timestep selection. *SIAM J. Numer. Anal.*, 37(1):194–210, 1999.

[54] E. G. Karpov, W. K. Liu, H. S. Park, and S. Zhang. An introduction to computational nanomechanics and materials. *Comput. Methods. Appl. Mech. Engrg.*, 193:1529–1578, 2004.

[55] J. Knap and M. Ortiz. An analysis of the quasicontinuum method. *J. Mech. Phys. Solids*, 49:1899–1923, 2001.

[56] M. G. Larson. A posteriori and a priori error analysis for finite element approximations of self-adjoint elliptic eigenvalue problems. *SIAM J. Numer. Anal.*, 38(2):608–625, 2000.

[57] J.E. Lennard–Jones. Cohesion. *Proc. Phys. Soc.*, 43:461–482, 1931.

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

[59] P. Lin. Theoretical and numerical analysis for the quasi-continuum approximation of a material particle model. *Math. Comp.*, 72(242):657–675, 2003.

[60] P. Lin. Convergence analysis of a quasi-continuum approximation for a two-dimensional material without defects. to appear in SIAM J. Num. Ana., 2006.

[61] W. K. Liu and H. S. Park. An introduction and tutorial on multiple-scale analysis in solids. *Comput. Methods Appl. Mech. Engrg.*, 193:1733–1772, 2004.

[62] W. K. Liu and G. J. Wagner. Coupling of atomistic and continuum simulations using a bridging scale decomposition. *J. Comp. Phys.*, 190:249–274, 2003.

[63] C. G. Makridakis. Finite element approximations of nonlinear elastic waves. *Math. Comp.*, 61(204):569–594, 1993.

[64] R. Miller, R. Phillips, M. Ortiz, D. Rodney, V. B. Shenoy, and E. B. Tadmor. An adaptive finite element approach to atomic-scale mechanics—the quasicontinuum method. *J. Mech. Phys. Solids*, 47(3):611–642, 1999.

[65] R. E. Miller and E. B. Tadmor. The quasicontinuum method: Overview, applications and current directions. *Journal of Computer-Aided Materials Design*, 9:203–239, 2003.

[66] P. M. Morse. Diatomic molecules according to the wave mechanics. II. Vibrational levels. *Phys.Rev.*, 34:57–64, 1929.

[67] S. Müller. Variational Models for Microstructure and Phase Transitions. Lecture Notes no. 2, Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig, 1998.

[68] J. N. Murrell. The many-body expansion of the potential-energy function for elemental clusters. *Int. J. Quant. Chem.*, 37:95–102, 1990.

[69] J. N. Murrell and R. E. Mottram. Potential-energy functions for atomic solids. *Mol. Phys.*, 69:571–585, 1990.

[70] M. Negri. A discontinuous finite element approximation of free discontinuity problems. *Adv. Math. Sci. Appl.*, 15(1):283–306, 2005.

[71] M. Negri and C. Ortner. Numerical analysis of Griffith's model for fracture. Work in progress.

[72] J. Tinsley Oden, Kumar Vemaganti, and Nicolas Moës. Hierarchical modeling of heterogeneous solids. *Comput. Methods Appl. Mech. Engrg.*, 172(1-4):3–25, 1999.

[73] M. Ortiz, R. Phillips, and E. B. Tadmor. Quasicontinuum analysis of defects in solids. *Philosophical Magazine A*, 73(6):1529–1563, 1996.

[74] C. Ortner. Gradient flows as a selection procedure for equlibria of non–convex energies. To appear in SIAM J. Math. Anal.

[75] C. Ortner. Γ-limits of Galerkin discretizations with quadrature. Technical Report 04/26, Oxford University Computing Laboratory, 2004.

[76] C. Ortner. Continuum limits of an atomistic energy based on local energy minimization. Technical Report NA05/11, Oxford University Computing Laboratory, 2005.

[77] C. Ortner. Two variational techniques for the approximation of curves of maximal slope. Technical Report NA05/10, Oxford University Computing Laboratory, 2005.

[78] C. Ortner and E. Süli. Discontinuous Galerkin finite element approximation of nonlinear second-order elliptic and hyperbolic systems. Technical Report NA06/05, Oxford University Computing Laboratory, 2006.

[79] P. Pedregal. *Variational Methods in Nonlinear Elasticity*. SIAM, 2000.

[80] R. L. Pego. Phase transitions in one-dimensional nonlinear viscoelasticity: admissibility and stability. *Arch. Rational Mech. Anal.*, 97(4):353–394, 1987.

[81] M. Plum. Computer-assisted enclosure methods for elliptic differential equations. *Linear Algebra Appl.*, 324(1-3):147–187, 2001. Special issue on linear algebra in self-validating methods.

[82] R. Rannacher and R. Scott. Some optimal error estimates for piecewise linear finite element approximations. *Math. Comp.*, 38(158):437–445, 1982.

[83] M. O. Rieger and J. Zimmer. Young measure flow as a model for damage. Preprint, University of Bath, 2005.

[84] R. T. Rockafellar. Monotone operators and the proximal point algorithm. *SIAM J. Control Optimization*, 14(5):877–898, 1976.

[85] W. Rudin. *Functional Analysis*. International Series in Pure and Applied Mathematics. McGraw-Hill Inc., New York, second edition, 1991.

[86] A. Ruszczyński. *Nonlinear Optimization*. Princeton University Press, Princeton, NJ, 2006.

[87] E. Sandier and S. Serfaty. Gamma-convergence of gradient flows with applications to Ginzburg-Landau. *Comm. Pure Appl. Math.*, 57(12):1627–1672, 2004.

[88] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis*, volume 12 of *Texts in Applied Mathematics*. Springer-Verlag, New York, third edition, 2002. Translated from the German by R. Bartels, W. Gautschi and C. Witzgall.

[89] S. Tang, T. Y. Hou, and W. K. Liu. A mathematical framework of the bridging scale method. *Internat. J. Numer. Methods Engrg.*, 65(10):1688–1713, 2006.

[90] L. Tartar. Does Nature minimize or conserve energy? Center for Nonlinear Analysis Summer School, Multiscale Problems in Nonlinear Analysis, 2001.

[91] J. Tersoff. A new empirical model for the structural properties of silicon. *Phys. Rev. Lett.*, 56(6):632–635, 1986.

[92] F. Theil. A proof of crystallization in two dimensions. *Comm. Math. Phys.*, 262(1):209–236, 2006.

[93] L. Truskinovsky and A. Vainchtein. The origin of nucleation peak in transformational plasticity. *J. Mech. Phys. Solids*, 52(6):1421–1446, 2004.

[94] R. Verfürth. A posteriori error estimates for nonlinear problems. Finite element discretizations of elliptic equations. *Math. Comp.*, 62(206):445–475, 1994.

[95] R. Verfürth. *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. John Wiley & Sons Ltd., 1996.

[96] R. Verfürth. Error estimates for some quasi-interpolation operators. *M2AN Math. Model. Numer. Anal.*, 33(4):695–713, 1999.

[97] N.T. Wilson. *The Structure and Dynamics of Noble Metal Clusters*. PhD thesis, University of Birmingham, 2000.

[98] K. Yosida. *Functional analysis*, volume 123 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, sixth edition, 1980.

[99] T. I. Zohdi, J. T. Oden, and G. J. Rodin. Hierarchical modeling of heterogeneous bodies. *Comput. Methods Appl. Mech. Engrg.*, 138(1-4):273–298, 1996.

[100] R. R. Zope and Y. Mishin. Interatomic potentials for atomistic simulations of the Ti-Al system. *Phys. Rev. B*, 68:024102, July 2003.