On the asymptotically stochastic computational modeling of microstructures

Dennis D. Cox, Petr Klouček, Daniel R. Reynolds

Abstract

We consider a class of alloys and ceramics with equilibria described by non-attainable infima of non-quasiconvex variational integrals. Such situations frequently arise when atomic lattice structure plays an important role at the mesoscopic continuum level.

We prove that standard variational approaches associated with gradient based relaxation of non-quasiconvex integrals in Banach or Hilbert spaces are not capable of generating relaxing sequences for problems with non-attainable structure.

We introduce a variational principle suitable for the computational purposes of approaching non-attainable infima of variational integrals. We demonstrate that this principle is suitable for direct calculations of the Young Measures on a computational example in one dimension.

The new variational principle provides the possibility to approximate crystalline microstructures using a Fokker–Planck equation at the meso-scale. We provide an example of such a construction.

The successful relaxation of strain energies of constrained crystalline materials allows us to determine which particular atomic structures produce mesoscopic strain. Such information can be used, e.g., to decide whether or not the stress-induced martensitic transformation can occur at the tip of a propagating crack. If the structure of the material does allow for the martensitic transformation then the crack propagation may be stopped by a toughening accompanying the transformation from a less stable and more ductile state, say a tetragonal lattice structure, to a more stable and less ductile, say, monoclinic lattice arrangement.

The model we consider here does not account for a surface energy, hence it lacks any spatial scale. Phenomenologically such models can be used only to investigate which variants, from the set of all the

Keywords: Microstructures; Crystalline materials; Weak white noise; Steepest descent algorithm; Weak convergence; Calculus of variations

1. Introduction

We address in this paper the necessity of using a stochastic variational principle (SVP) leading to the Fokker–Planck equation describing the meso-scale properties of various materials undergoing a martensitic transformation. The motivation for this approach stems from the theory presented here which indicates that descent relaxation methods based on a pseudo-gradient will not relax non-attainable non-quasiconvex problems.

The successful relaxation of strain energies of constrained crystalline materials allows us to determine which particular atomic structures produce mesoscopic strain. Such information can be used, e.g., to decide whether or not the stress-induced martensitic transformation can occur at the tip of a propagating crack. If the structure of the material does allow for the martensitic transformation then the crack propagation may be stopped by a toughening accompanying the transformation from a less stable and more ductile state, say a tetragonal lattice structure, to a more stable and less ductile, say, monoclinic lattice arrangement.

The model we consider here does not account for a surface energy, hence it lacks any spatial scale. Phenomenologically such models can be used only to investigate which variants, from the set of all the
The paper is organized as follows. We describe a relation between microstructures and transformation toughening in Section 2. We recall how non-attainable states can be represented using Radon measures in Section 3. We describe a particular construction of a pseudo-gradient in Section 4. We prove that gradient based relaxation fails to achieve stochastic states in Section 5. We introduce the constrained finite element method in Section 6. The next two sections introduce a notion of weak white noise and a corresponding variational principle. Section 9 demonstrates our variational approach by direct calculations of a Young Measure and its underlying probabilities.

2. Transformation toughening in zirconia

The purpose of this section is to discuss how certain material properties can be drastically altered by changing just the symmetry of their lattice structure, i.e., by promoting certain specific arrangements of their atomic lattices. We rely in this section on the overview article [13].

Transformation toughening is the increase in fracture toughness of a material that is the direct result of a phase transformation occurring at the tip of an advancing crack. The discovery of transformation toughening in zirconia ceramics [12] indicates that traditionally brittle ceramics can reach fracture toughness four or more times higher when it undergoes the martensitic transformation. This is a lattice-distortive, virtually diffusionless structural change having a dominant deviatoric component and associated shape change such that strain energy dominates the kinetics and morphology during the transformation [9]. Hence, zirconia ceramics are a good candidate for applications where toughness is required and where an advantage of wear resistance, low density and of high melting point characterizing ceramics can be taken advantage of.

The martensitic transformation is triggered by the nucleation strain at the interface between a progressing crack and the bulk zirconia. This ceramic undergoes the tetragonal to monoclinic transformation. The mathematical challenge associated with modeling this lattice transformation is to find the microscopic distribution of martensitic variants participating in the mesoscopic strain. The distribution of these variants determines the structure of the atomic lattice. The lattice structure has a direct relation to the transformation toughening: the more complex is the microstructure, the more difficult it is for the crack to propagate. The drawing and caption are reproduced from [13].
monoclinic variants (also low-temperature variants) are given by [2]:

\[ F_i = (I + \eta(-1)^i) \varepsilon_2 \otimes \varepsilon_1 \sum_{i=1}^{N} d_i \varepsilon_2 \otimes \varepsilon_1, \]

where \( \eta > 0, \ i = 1, 2. \) (2)

We refer to [13] for discussion of observations and measurements identifying these matrices during the transformation to monoclinic zirconia.

Simple conditions for the existence of rank-one connection between martensitic wells are given in [11]. These conditions are used to determine the structure of the set \( F \) in (2) constraining the deformation gradients. The rank-one connection is required by the Hadamard jump condition to allow for a continuous deformation with given gradients.

Using the classical variational approach, we can formulate the problem (1) as follows. Let us assume that there exists a strain density \( W \), having its local minimum distributed along the orbits:

\[ \{ QF | i = 1, 2 \}, \quad \text{where} \quad Q \in \text{SO}(N). \] (3)

The density must inherit the symmetry of the tetragonal phase, i.e.:

\[ W(QF) = W(F) \quad \text{for any} \quad Q \in \mathcal{G}, \] (4)

where \( \mathcal{G} \) is the symmetry group of the tetragonal phase.

The density \( W \) must satisfy the principle of frame indifference, i.e.:

\[ W(RF) = W(F) \quad \text{for any rotation} \quad R \in \text{SO}(N). \] (5)

We also assume the following growth and coerciveness conditions for the density \( W \):

\[ C_1 |F|^p + 1 \geq W(F) \geq C_2 |F|^p - C_3 \] (6)

for some suitable \( p > N \) and \( C_1 > 0, i = 1, 2, 3. \) It is also reasonable to assume that, in the extreme case:

\[ W(F) \to +\infty \quad \text{if} \quad \det F \to 0+. \] (7)

We then define

\[ E(u) = \int_{\Omega} W(\nabla u(x)) \, dx \] (8)

and we investigate

\[ \inf\{E(u) | u \in W^{1,p}(\Omega), \ u = g \text{ on} \ \partial \Omega \} \quad \text{for some suitable} \quad p > N. \] (9)

The difficulty with finding the infimum or minimum in (9) is that we have to deal with a profoundly non-quasiconvex problem which often has unattainable structure. The properties of the minimizer depend strongly on the choice of the Dirichlet boundary condition \( g \).

The complexity of microstructures corresponding to (9) can be considerable, and it may be accompanied by a massive lack of uniqueness. Demonstration of these difficulties can be found in [15, Section 4.7].

3. Representation of non-attainable states

A typical relaxing sequence of (9) will converge weakly, likely weakly-*., to its expected value, or average, in an appropriate Sobolev space. Since we do not expect the variational integral in (8) to be weakly lower semicontinuous, we assume that

\[ \liminf_{n \to \infty} \int_{\Omega} W(\nabla u_n(x)) \, dx < \int_{\Omega} W(\nabla u(x)) \, dx. \] (10)

Here, \( \{u_n\}_{n \in N} \subset W^{1,p}(\Omega) \) is a relaxing sequence and \( u \in W^{1,p}(\Omega) \) is its weak limit. Translation of (10) into the framework of material science would imply that the function \( u \), representing the averaged deformation of a zirconia ceramic, does not carry any pointwise information. Yet, a precise description of the spatial composition of deformation gradients is needed to describe the transformation toughening as explained in Section 2. The lattice structure is also important and necessary as an input for thermodynamic models [16] (Fig. 2).

In order to overcome this difficulty we may suppose that for any density \( W \in C^0(\Omega) \) which has at least a quadratic growth at infinity there exists a function \( W \in L^1(\Omega) \) such that

\[ \liminf_{n \to \infty} \int_{\omega} W(\nabla u_n(x)) \, dx = \int_{\omega} W(x) \, dx \] (11)

for any \( \omega \) which is a compact subset of \( \Omega \). The function \( W \) does not remember anything about the equilibrium structure of the material. This is encoded into
the density $W$. Though more importantly, the function $\tilde{W}$ does tell us about the mesoscopic distribution of the elastic energy density $W$.

The function $\tilde{W}$ can be obtained via integral representation using a Radon probability measure $\mu_\lambda$. Namely:

$$\tilde{W}(x) = \int_{\mathbb{R}^m} W(y) d\mu_\lambda(y).$$

(12)

The Radon measure $\mu_\lambda$ has a very natural structure in the case of alloys undergoing simple symmetry phase change. In the case of bi-stable alloys, it is possible to show that [15]:

$$\mu_\lambda = \lambda(x)\delta_{F_1} + (1 - \lambda(x))\delta_{F_2},$$

(13)

where $F_1, F_2$ are the deformation gradients participating in the equilibrium deformation. The function $\lambda(x)$ represents the probability that at a point $x$ the material is deformed with the deformation gradient $F_1$. Usually, this ratio is called the volume fraction.

It follows directly from its definition that

$$\lambda(x) = \lim_{r \to 0^+} \lim_{n \to \infty} \lim_{\varepsilon \to 0^+} \frac{\text{meas}\{y \in B_R(x) | \|D\varepsilon_n(x) - F_1\| \leq \varepsilon\}}{\text{meas}(B_R(x))}.$$

(14)

where $D\varepsilon_n$ are the gradients of the relaxing sequence of the strain energy. For (14) to be valid, the minimizing sequence must become stochastic in its derivative. This represents a major challenge for the optimization, numerical and computational approaches. We believe that the desirable methods for direct calculations of Young Measures $\mu_\lambda$ in (13) are those which generate a stochastic state from an initial deterministic one. This problem has not been solved. It seems that it is not well understood what methods can achieve this goal.

The crux of this problem can be represented by an almost 70-year-old question posed by Heisenberg: what physical processes would serve as the source of an origin of stochastic behavior?

It is possible to construct such sequences (processes) using a self-similar periodic construction, e.g. [5]. It is however impossible to reconstruct such sequences computationally using "off-the-shelf" tools. We address this issue in Section 5. We strive to compute the volume fraction using a Stochastic Variational Principle (SVP) that will guarantee that any minimizing sequence becomes asymptotically...
A pseudo-gradient prohibits the convergence otherwise. Any kind of finite-dimensional approximation, which is in the presented form relatively inexpensive to compute. The pseudo-gradient is often used in critical point theory\[8\].

The relaxation of (9) in Banach spaces is complicated since it is very expensive to obtain the steepest descent direction. We instead use a pseudo-gradient which is in the presented form relatively inexpensive to compute. The pseudo-gradient is often used in critical point theory\[8\].

Let us assume that \(E\) is a \(C^1\)-functional on a Banach space \(V\). The substitution for a gradient in spaces lacking a scalar product is the following cone of locally Lipschitz continuous vector fields \(g(u) : V \setminus \{0\} \mapsto V\) such that \[18\]:

\[\|
g(u)\|_V < 2 \min\{\|dE(u, \cdot)\|_{V^*}, 1\}, \tag{15}\]

\[dE(u, g(u)) > \min\{\|dE(u, \cdot)\|_{V^*}, 1\}\|dE(u, \cdot)\|_{V^*}, \tag{16}\]

where \(dE(u, \cdot)\) denotes the Gateaux derivative of \(E\) which is a linear functional on \(V\). A remarkable result\[19, Lemma 3.2\] guarantees that any \(E \in C^1(V)\) admits a pseudo-gradient vector field. This result guarantees that a gradient flow is well defined and asymptotically leads to descent to a critical point. In particular, this means that any vector field satisfying the conditions (15) and (16) can be used to define a gradient descent method for minimization of variational integrals such as (8). Assuming that the energy \(E, (8)\), is defined on \(V = W^{1,p}\), \(p > 2\), we define a particular pseudo-gradient below.

First, let us denote the distributional gradient of the energy \(E\) by \(g\), i.e., for any \(u \in W^{1,p}(\Omega)\), the functional \(G \in W^{-1,q}(\Omega, \mathbb{R}^N)\), where \(W^{-1,q}(\Omega, \mathbb{R}^N)\) denotes the dual space of \(W^{1,q}(\Omega, \mathbb{R}^N)\), \(1/p + 1/q = 1\), is given by the variational relation:

\[\langle G(u), \psi \rangle_{W^{-1,q}(\Omega, \mathbb{R}^N), W^{1,q}(\Omega, \mathbb{R}^N)} = \int_{\Omega} \sum_{i=1}^N f_i(x) \frac{\partial \psi(x)}{\partial x_i} + f_0(x) \psi(x) \, dx,\]

\[\|F\|_{W^{-1,q}(\Omega, \mathbb{R}^N)} = \left( \int_{\Omega} \sum_{i=1}^N |f_i|^q \right)^{1/q}, \tag{21}\]

for all \(\psi \in W^{1,q}(\Omega, \mathbb{R}^N)\), where \(A, B \in \mathbb{R}^{N \times N}\). We have the following Lemma.

**Lemma 4.1** (pseudo-gradient). Let us assume that \(\Omega\) is bounded with smooth boundary. Let \(\tilde{g}(u) \in W^{1,q}_0(\Omega, \mathbb{R}^N)\) be defined by

\[\Delta \tilde{g}(u)(x) = \text{div}(\nabla \tilde{g}(u)(x))(\nabla \tilde{g}(u)(x))^{-1}g(u), \quad x \in \Omega, \quad \tilde{g}(u)(x) = 0 \text{ on } \partial \Omega, \tag{18}\]

where \(|\cdot|\) is a matrix norm and \(\tilde{g}(u) \in W^{1,q}_0(\Omega, \mathbb{R}^N)\) is given by

\[\tilde{g}(u) = -\Delta^{-1}G(u). \tag{19}\]

Here \(-\Delta^{-1} : W^{-1,q}(\Omega, \mathbb{R}^N) \mapsto W^{1,q}_0(\Omega, \mathbb{R}^N)\). Finally we define

\[g(u) = \tilde{g}(u)(\tilde{g}(u))^{p-2} \in W^{1,q}_0(\Omega, \mathbb{R}^N). \tag{20}\]

Then \(g(u)\) is a pseudo-gradient of the strain energy \(E\) in the sense of the conditions (15) and (16) with \(V = W^{1,q}(\Omega, \mathbb{R}^N)\).

**Remark 4.2.** We have \(g(u) \in W^{1,q}_0(\Omega, \mathbb{R}^N)\) in view of the growth conditions (6).

**Remark 4.3.** We note that the \(-\Delta^{-1}\) mapping is isometric from \(W^{-1,q}(\Omega, \mathbb{R}^N)\) to \(W^{1,q}_0(\Omega, \mathbb{R}^N)\) \[1\]. Hence, \(\|\tilde{g}(u)\|_{W^{1,q}_0(\Omega, \mathbb{R}^N)} = \|G(u)\|_{W^{-1,q}(\Omega, \mathbb{R}^N)}\).

**Remark 4.4.** The pre-gradient \(\tilde{g}(u)\) is unique for a given \(G\). Namely, assuming that \(\partial \Omega\) is piecewise \(C^\infty\), and taking \(F \in W^{-1,q}(\Omega)\) represented by

\[\langle F, \psi \rangle_{W^{1,q}(\Omega, \mathbb{R}^N), W^{-1,q}(\Omega, \mathbb{R}^N)} = \int_{\Omega} \sum_{i=1}^N f_i(x) \frac{\partial \psi(x)}{\partial x_i} + f_0(x) \psi(x) \, dx,\]

we have \[14, Theorem 3.8\]. Let \(p \in [2, \infty)\), then there exists a unique solution \(\tilde{g}(u) \in W^{1,q}(\Omega, \mathbb{R}^N), u(x) = \)
\[ u_0(x) \text{ on } \partial \Omega, \quad u_0 \in W^{1,q}(\Omega), \]
\[ \int_{\Omega} \nabla u(x) \nabla \varphi(x) \, dx = \langle F, \varphi \rangle, \]
\[ \text{for all } \varphi \in W_0^{1,q}(\Omega). \quad (22) \]
Moreover, there exists a constant \( C(q, \Omega) \) such that
\[ \| \hat{u} \|_{W^{1,q}(\Omega)} \leq C(q, \Omega) \| u_0 \|_{W^{1,q}(\Omega)} + \| F \|_{W^{-1,q}(\Omega)}. \quad (23) \]

**Proof of Lemma 4.1.** We observe that (18) yields
\[ \int_{\Omega} \nabla \hat{u}(x) \nabla \varphi(x) \, dx = \int_{\Omega} \nabla \varphi(x) \, dx. \quad (24) \]
The Hölder inequality and (24) yields
\[ \int_{\Omega} |\hat{u}(x)|^p \, dx \leq \int_{\Omega} |\varphi(x)|^p \, dx \quad (25) \]
It follows from Remark 4.4 and (18) that there exists a constant \( C \), depending on \( p \) and \( \Omega \), such that
\[ \int_{\Omega} |\nabla \hat{u}(x)|^q \, dx \leq C(p, \Omega)^q \int_{\Omega} |\nabla \varphi(x)|^q \, dx. \quad (26) \]
Hence (25) and (26) yield
\[ \int_{\Omega} |\nabla \hat{u}(x)|^p \, dx = C(p, \Omega)^p \int_{\Omega} |\nabla \varphi(x)|^q \, dx \quad (27) \]
Since
\[ \int_{\Omega} |\nabla \hat{u}(x)|^p \, dx \]
\[ = C(p, \Omega)^p \int_{\Omega} |\nabla \varphi(x)|^q \, dx \]
\[ = C(p, \Omega)^p \int_{\Omega} |\nabla \varphi(x)| \cdot |\nabla \hat{u}(x)| \, dx \]
\[ = C(p, \Omega)^p \int_{\Omega} D(\nabla u(x)) : \nabla \hat{u}(x) \, dx, \quad (28) \]
we have either with respect to the semi-norm \( W_{0,1}^{1,p}(\Omega) \)
or from the Poincaré inequality (assuming that the constant appearing in it is bounded by 1):
\[ \| \hat{u} \|_{W_{0,1}^{1,p}(\Omega)} \]
\[ = C(p, \Omega)^p \| D(\nabla u(x)) \| \| \hat{u}(x) \|_{W_{0,1}^{1,p}(\Omega)} \]
\[ \leq C(p, \Omega)^p \| D(\nabla u(x)) \| \| \hat{u}(x) \|_{W_{0,1}^{1,p}(\Omega)}. \quad (29) \]
Thus (20) and (29) yield
\[ \| \hat{u}(x) \|_{W_{0,1}^{1,p}(\Omega)} \]
\[ = \| \hat{u}(x) \|_{W_{0,1}^{1,p}(\Omega)} \]
\[ \leq C(p, \Omega)^p \| D(\nabla u(x)) \| \| \hat{u}(x) \|_{W_{0,1}^{1,p}(\Omega)}. \quad (30) \]
Thus we assume \( C(p, \Omega)^p < 1 \)
\[ \| \hat{u} \|_{W_{0,1}^{1,p}(\Omega)} \]
\[ \leq \| \hat{u} \|_{W_{0,1}^{1,p}(\Omega)} \]
\[ \leq C(p, \Omega)^p \| D(\nabla u(x)) \| \| \hat{u}(x) \|_{W_{0,1}^{1,p}(\Omega)}. \quad (32) \]
Thus
\[ \sup_{\| \hat{u} \|_{W_{0,1}^{1,p}(\Omega)} \leq 1} \| dE(\hat{u}, \varphi) \| \]
\[ \leq \sup_{\| \hat{u} \|_{W_{0,1}^{1,p}(\Omega)} \leq 1} \| \nabla \hat{u}(x) \cdot \nabla \varphi(x) \| \]
\[ \leq C(p, \Omega)^p \| D(\nabla u(x)) \| \| \hat{u}(x) \|_{W_{0,1}^{1,p}(\Omega)}. \quad (33) \]
Hence for \( q \in (1, 2) \), i.e., for \( p > 2 \):
\[ \| dE(\hat{u}, \cdot) \|_{L^{1/q}(W_{0,1}^{1,p}(\Omega))} \leq \| dE(\hat{u}, \cdot) \|_{L^{1/q}(W_{0,1}^{1,p}(\Omega))}, \quad (34) \]
Assuming \( \| dE(\hat{u}, \cdot) \|_{W_{0,1}^{1,q}(\Omega)} < 1 \), taking some \( \varepsilon \in (0, 1) \) so that \( q + \varepsilon < 2 \), and using (33), we obtain
\[ \| dE(\hat{u}, \cdot) \|_{W_{0,1}^{1,q}(\Omega)} \]
\[ \leq \| dE(\hat{u}, \cdot) \|_{L^{1/q}(W_{0,1}^{1,p}(\Omega))} \]
\[ \leq \| dE(\hat{u}, \cdot) \|_{W_{0,1}^{1,q}(\Omega)} \| dE(\hat{u}, \cdot) \|_{W_{0,1}^{1,q}(\Omega)} \]
\[ = \| dE(\hat{u}, \cdot) \|_{W_{0,1}^{1,q}(\Omega)} \| dE(\hat{u}, \cdot) \|_{W_{0,1}^{1,q}(\Omega)} \]
\[ \times \| dE(\hat{u}, \cdot) \|_{W_{0,1}^{1,q}(\Omega)}. \quad (35) \]
It remains to show that
\[ \| \hat{u}(x) \|_{W_{0,1}^{1,p}(\Omega)} \]
\[ \leq \| \hat{u}(x) \|_{W_{0,1}^{1,p}(\Omega)} \]
\[ \leq C(p, \Omega)^p \| D(\nabla u(x)) \| \| \hat{u}(x) \|_{W_{0,1}^{1,p}(\Omega)}. \quad (36) \]
It follows from (29) and (31) that
\[ \| \hat{u}(x) \|_{W_{0,1}^{1,p}(\Omega)} \]
\[ \leq \| \hat{u}(x) \|_{W_{0,1}^{1,p}(\Omega)} \]
\[ \leq C(p, \Omega)^p \| D(\nabla u(x)) \| \| \hat{u}(x) \|_{W_{0,1}^{1,p}(\Omega)}. \quad (37) \]
continuity of the embedding of $W^{1,p}$ functions, $C(\Omega)$

5. Failure of pseudo-gradient based relaxation

The purpose of this section is to show that for any $u_0 \in W^{1,p}(\Omega, \mathbb{R}^N)$, $p > N$, the sequence $\{u_n\}_{n \in \mathbb{N}} \subset W^{1,p}(\Omega, \mathbb{R}^N)$, $p > N$, generated during a gradient navigated relaxation of (8) with the Dirichlet boundary data such that (9) does not admit minimizer converges strongly to a function in $W^{1,p}(\Omega, \mathbb{R}^N)$. The minimizer represents a local minimum of the stored energy $E$. The derivatives of the minimizing sequence thus do not converge to a measure-valued distribution even though the infimum of the stored energy is not attainable in any Sobolev space. Namely:

$u_n \rightharpoonup u$ strongly in $W^{1,1}(\Omega, \mathbb{R}^N)$, and consequently, $\lim_{n \to \infty} E(u_n) = E(u) > 0$.\hspace{0.5cm} (41)

This result indicates that we need a different structure of the strain energy in order to obtain the measure-valued distributions. We address this modification in Section 8.

We assume that the density $W$ has the following structural properties:

$E(u) \geq \lambda p^2 \int_\Omega \|
abla u(x) - (\Pi \nabla u(x))\|^p \, dx, \quad \lambda > 0,$\hspace{0.5cm} (42a)

$\left| D^2 W(Q) \right|_{N^2 \times N^2} \leq A \left| Q \right|_{N \times N}^{p-2}, \quad Q \in M^{N \times N},$\hspace{0.5cm} (42b)

$1 \leq \|u_n\|_{W^{1,p}(\Omega, \mathbb{R}^N)} \leq A_0,$\hspace{0.5cm} (42c)

$\sqrt{\lambda} - \frac{A_{p-2}}{\lambda^{p/2(p-1)}} > 0,$\hspace{0.5cm} (42d)

$\frac{A_{p-2}}{\sqrt{\lambda}} \leq 1,$\hspace{0.5cm} (42e)

where $\Pi$ is a projection on the set of equilibria of $E$ given by (3).

We prove the following two lemmas that will be used in the proof of Theorem 5.3.

**Lemma 5.1.** Let the sequence $\{u_n\}_{n=0}^\infty \subset W^{1,p}(\Omega, \mathbb{R}^N)$, $p > N$, be generated by the descent algorithm, Definition 4.5. Let us assume that (42) holds true. Let

$\lim_{n \to \infty} E(u_n) = J.$\hspace{0.5cm} (43)

Then for any $t \in [0, 1]$ we have

$C(p, \Omega, \lambda, A) \|E(u_n + tg(u_n)) - J\|^{1-1/p} \leq \|g(u_n + tg(u_n))\|_{W^{1,p}(\Omega, \mathbb{R}^N)}^{1-1/p},$\hspace{0.5cm} (44)

where

$C(p, \Omega, \lambda, A) \equiv C(p, \Omega) \lambda^{-1} \left( \frac{A_{p-2}}{\lambda^{p/2(p-1)}} \right).$

The constant $C(p, \Omega)$ originates in the estimate (23).

**Proof.** We abbreviate $u_n(t) \equiv u_{n+1} + tg(u_n)$
and we assume without loss of generality that \( J = 0 \).

It follows from the Hadamard jump condition [2] that there exists a \( \Pi \in C^0(\Omega) \) such that

\[
\nabla u_N \in SO(N)F_1 \cup SO(N)F_2,
\]

\[
\|u_N - u_{\Pi}\|_{L^1(\Omega)} \leq C_1 \|\nabla u_N - \nabla u_{\Pi}\|_{L^1(\Omega)R^{n+1}},
\]

where \( \Pi \) is a projection on the set of equilibria of \( E \). Consequently, \( E(u_N) = 0 \).

Hence

\[
E(u_N) = E(u_N(t)) - E(u_N(t)) = \int_0^T \int_\Omega \left[ \frac{1}{2} \nabla (\nabla u_N(t)) : (\nabla u_N(t)) - \nu u_N(t) \right] d\tau dx - \frac{1}{2} \int_\Omega (\nabla u_N(t) : (\nabla u_N(t)) - \nu u_N(t)) d\tau.
\]

Using the assumption (42b), the last integral in (49) can be estimated by

\[
\int_0^T \int_\Omega \left[ \frac{1}{2} \nabla (\nabla u_N(t)) : (\nabla u_N(t)) - \nu u_N(t) \right] d\tau dx - \frac{1}{2} \int_\Omega (\nabla u_N(t) : (\nabla u_N(t)) - \nu u_N(t)) d\tau.
\]
The last inequality is obtained as follows. We have
\[ C_1(\|\nabla u_n\|_{L^p(\Omega;\mathbb{R}^N)} + 1) \geq E(u_n) \]
\[ \geq \lambda^{p/2} \|\nabla u_n(t) - u_n(t)\|_{L^2(\Omega;\mathbb{R}^N)} \]  
Thus, in view of (42b) and (42c)
\[ 1 \geq \frac{2C_1}{\sqrt{\lambda}} \geq \frac{\|\nabla u_n(t) - u_n(t)\|_{L^2(\Omega;\mathbb{R}^N)}}{\|\nabla u_n\|_{L^2(\Omega;\mathbb{R}^N)}} \]  
Consequently, it follows from (54) that
\[ \left( \int_{\Omega} (\int_{\Omega} |\nabla g(u_n)|^{p-2} \nabla g(u_n) \cdot \nabla d\mu)^{p-2} \right)^{1/p} \leq A^{p-2}. \]  
Thus (49), (51) and (57) yield
\[ E(u_n(t)) \leq C(p, \Omega)^{-p/q} \|g(u_n(t))\|_{H^1_0(\Omega)} \]
\[ \times \frac{1}{\sqrt{\lambda}} \left( \int_{\Omega} |\nabla u_n(t)|^{1/p} + \lambda^{p/2} \|\nabla u_n\|_{L^2(\Omega;\mathbb{R}^N)} \right)^{1/p}. \]  
Hence
\[ 1 - \frac{A^{p-2}}{\lambda^{p/2}} E(u_n(t)) \]
\[ \leq C(p, \Omega)^{-p/q} \frac{1}{\sqrt{\lambda}} \|g(u_n(t))\|_{H^1_0(\Omega)} \]
\[ \times \left( \int_{\Omega} |\nabla u_n(t)|^{1/p} + \lambda^{p/2} \|\nabla u_n\|_{L^2(\Omega;\mathbb{R}^N)} \right)^{1/p}. \]  
which yields (44). \( \square \)

**Lemma 5.2.** Let \( \{g(u_n)\}_{n \in \mathbb{N}} \subset W^{1,p}(\Omega; \mathbb{R}^N) \), \( p > N \), be the sequence of pseudo-gradients, defined by (18), (19) and (20), corresponding to the sequence \( \{u_n\}_{n \in \mathbb{N}} \subset W^{1,p}(\Omega; \mathbb{R}^N) \), \( p > N \), generated by the descent algorithm, Definition 4.5. Then there exists a positive finite constant \( \beta \), independent of \( n \), such that
\[ \int_{\Omega} \nabla g(u_{n+1}) : \nabla g(u_n) \, dx \]
\[ \geq \beta \left( \int_{\Omega} |g(u_{n+1}) + t g(u_n)| \right) \]  
\[ \times \left( \int_{\Omega} |g(u_n)| \right)^{1/p} \]  
for \( t \in [0, a_n] \). \( \beta \) is defined as
\[ \beta = \max_{n \in \mathbb{N}} (1 - (a_n)^2)^{-1} \]
from (63) and the inequality \( a^2 + b^2 \geq 2ab \). It remains to verify (62).

Since for any \( \phi \in W^{1,2}_0(\Omega; \mathbb{R}^N) \):
\[ \int_{\Omega} \nabla g(u_n) : \nabla \phi \, dx = \int_{\Omega} D W(u_n) : \nabla \phi \, dx, \]
\[ \int_{\Omega} \nabla g(u_{n+1} + t g(u_n)) : \nabla \phi \, dx \]
\[ = \int_{\Omega} D W(u_{n+1} + t g(u_n)) : \nabla \phi \, dx \]
\[ = \int_{\Omega} D W(u_n + (1 - a_n) g(u_n)) : \nabla \phi \, dx, \]  
we have in view of the assumptions (42b) and (42c) for \( t \in [0, a_n] \)
Theorem 5.3. Let the sequence \( \{u_n\}_{n=1}^\infty \subset W_0^{1,p}(\Omega, \mathbb{R}^N) \), \( p > N \), be generated by the descent algorithm, Definition 4.5. Let us assume that \( E(u_0) < 1 \) and let us assume that (42) holds true. Then there exists a finite constant \( C \) such that

\[
\sum_{n=0}^\infty a_n \| g(u_n) \|_{W_0^{1,p}(\Omega, \mathbb{R}^N)} \leq C < +\infty.
\]  

Proof. We denote again

\[
g_n(t(x)) := g(t(u_n)) \psi(x).
\]

Lemmas 5.1 and 5.2 and the definition of the descent algorithm (39) yield for any \( \theta \in [0, 1/2] \):

\[
\frac{d}{dt} E(u_{n+1} + t g(u_n))^{\theta} \cdot \nabla g(u_n) \frac{dx}{dx} = \theta \cdot D W(\nabla u_{n+1} + t \nabla g(u_n)) \cdot \nabla g(u_n) \\
\int_{\Omega} D W(\nabla u_{n+1} + t \nabla g(u_n)) \cdot \nabla g(u_n) \, dx = \frac{d}{dt} E(u_{n+1} + t g(u_n))^{\theta-1}.
\]

Lemma 5.2

\[
\sum_{n=0}^\infty a_n \| g(u_n) \|_{W_0^{1,p}(\Omega, \mathbb{R}^N)} \leq C < +\infty.
\]  

Integrating \( (d/dt) E(u_{n+1} + t g(u_n))^{\theta} \) over \( (0, a_n) \) we obtain

\[
E(u_{n+1} + t \alpha g(u_n))^{\theta} - E(u_{n+1})^{\theta} \geq C\rho_0 \| g(u_n) \|_{W_0^{1,p}(\Omega, \mathbb{R}^N)}^{\theta-1}.
\]  

Since \( E(u_{n+1} + t \alpha g(u_n)) = E(u_n) \), summing up (69) over \( n \in \mathbb{N} \) we obtain

\[
E(u_n)^{\theta} - \lim_{n \to \infty} E(u_n)^{\theta} \geq C\rho_0 \sum_{n=0}^\infty a_n \| g(u_n) \|_{W_0^{1,p}(\Omega, \mathbb{R}^N)}^{\theta-1},
\]

which proves (67).

Theorem 5.4 (strong convergence). Let the sequence \( \{u_n\}_{n=1}^\infty \subset W_0^{1,p}(\Omega, \mathbb{R}^N) \), \( p > N \), be generated by the descent algorithm, Definition 4.5. Let us assume that \( E(u_0) < 1 \) and let us assume that (42) holds true. Then there exists a function \( u \in W_0^{1,p}(\Omega, \mathbb{R}^N) \) such that

\[
u_n \to u, \quad \text{strongly in } W_0^{1,p}(\Omega, \mathbb{R}^N).
\]  

Consequently

\[
\lim_{n \to \infty} E(u_n) = E(u) > \inf_{v \in W_0^{1,p}(\Omega, \mathbb{R}^N)} \{ E(v) \}.
\]  

Proof. Let us assume that the sequence \( \{u_n\}_{n=1}^\infty \subset W_0^{1,p}(\Omega, \mathbb{R}^N) \) generated by the descent algorithm, Definition 4.5, minimizes the energy \( E \). Then there exists its weak limit \( u \in W_0^{1,p}(\Omega, \mathbb{R}^N) \). Thus

\[
\int_{\Omega} (D u_n(x) - D u(x)) \psi(x) \, dx = \sum_{n=0}^\infty \int_{\Omega} (D u_n(x) - D u_{n+1}(x)) \psi(x) \, dx
\]

for all \( \psi \in L^2(\Omega, \mathbb{R}^N) \).

Since

\[
\int_{\Omega} (D u_n(x) - D u_{n+1}(x)) \psi(x) \, dx = \int_{\Omega} \left( \frac{\nabla u_n}{\nabla g(u_n)} \right) \psi(x) \, dx,
\]

we have, using the Poincaré–Friedrichs inequality:
\[ \int_{\Omega} \left( \partial_u(x) - D_u(x) \partial(x) \right) \varphi(x) \, dx \]
\[ \leq C \| \varphi \|_{L^q(\Omega, \mathbb{R}^N)} \sum_{m=n}^{\infty} \alpha^m \| g(u_m) \|_{W^{1,p}(\Omega, \mathbb{R}^N)}. \] (75)

It follows from Theorem 5.3 that
\[ \sum_{m=n}^{\infty} \alpha^m \| g(u_m) \|_{W^{1,p}(\Omega, \mathbb{R}^N)} \to 0 \text{ as } n \to \infty. \] (76)

Hence
\[ \partial_u \to u \text{ strongly in } W^{1,p}(\Omega, \mathbb{R}^N). \] (77)

6. Langevin equation and the Fokker–Planck dynamics

Based on the result of Section 5, we assume that the elastic density \( W = W(Du) \) describing equilibria of crystalline materials with non-attainable structure should be written in the form (Fig. 3):
\[ W(s) = W_{\text{meso}}(s) + W_{\text{micro}}(s), \quad s \in \mathbb{R}^{N \times N}. \] (78)

The contribution \( W_{\text{meso}} \) encodes the information about the equilibrium state of a given material and the \( W_{\text{micro}} \) contribution guarantees that any minimizing sequence becomes weak white noise in the sense of Definition 7.1. In particular, we assume that
\[ \partial Du \cdot W_{\text{micro}}(Du)(x,t) = \sigma(Du(x,t)) \eta(x,t). \] (79)

where \( \eta \) becomes asymptotically (as \( t \to \infty \)) weak white noise. A possible construction yielding such a microscaling term using subgrid projection method is studied both analytically and computationally in [6].

We interpret the gradient flow in \( W^{-1}(\Omega) \)-topology as a Langevin equation having the form
\[ \frac{d}{dt} p(t, u) = \text{div}(\partial_u W_{\text{micro}}(Du(t,x)) + \sigma(Du(t,x)) \eta(x,t)). \] (80)

Writing (80) in its weak form and using a piecewise affine finite element approximation of \( u \), we obtain a system of ODEs for the coefficients \( \alpha^i = (\alpha_1, \ldots, \alpha_N) \) which now become functions of time. Namely:
\[ \frac{d}{dt} \alpha(t) = A^{-1}(b(\alpha) + \sum_i \hat{\eta}(t)). \] (81)

The Fokker–Planck equation corresponding to (81) for the probability density \( p = p(t,u) > 0 \) reads [17]:
\[ \frac{d}{dt} p(t,u) = \text{div}(A^{-1}(b(u)p(t,u)) + \frac{1}{2} D_2^2 A^{-1}(\sum_i (\partial u_i)p(t,u))). \] (82)

Let \( \mathcal{E} \) denote the expectation operator. Thus
\[ \mathcal{E}(\partial u(T))(t) = \int_{\Omega} \psi_T(y) \, dy, \quad \text{where} \]
\[ p_T(t,y) = \sum_{i=1}^N \alpha_i(t) \psi_i(y), \quad \eta_T(t,y) = \sum_{i=1}^N \alpha_i(t) \psi_i(y). \] (83)

The function \( p_T(t,y) \) represents the probability density for \( \alpha_i \). We note that for sharp values of \( \alpha_i \) we have \( p_T(t,y) \approx \delta_{\alpha_i}(y) \). Consequently, we define the

---

**Fig. 3.** The periodic or quasi-periodic structures often found in micrographs of metallic alloys or zirconia ceramics are invisible on the Langevin scale due to the stochastic nature of the description at the atomic scale. These structures are recovered by solutions of the Fokker–Planck equations. The picture on the right is a photo of a complex martensitic microstructure by C. Chu and R.D. James [4].
7. Multi-dimensional weak white noise

We have shown in Section 5 that the descent algorithm, Definition 4.5, applied to the double-well problem does not generate relaxing sequences. In other words, there exists a deterministic limiting state describing the infimum of any minimizing sequence to adhere to any state where possible spatial correlation can occur. This mechanism prohibits the minimizing sequences to converge to any of the many local minima of the variational integral, and leads to the representation (89) , the averaged state (the weak limit) with numeric noise.

Typically this leads almost immediately to a convergence of the minimizing sequence to the nearest local minimum. The evaluation of the volume fraction cannot provide a reliable approximation. In computational practice, the initial iteration for descent algorithms is obtained by a superposition of the averaged state (the weak limit) with numeric noise.

This approach is adopted in [7].

8. An SVP in one dimension

For the one-dimensional analog of the SVP, we consider the case where the set of allowable gradients \( \mathcal{A} = \{ \pm 1 \} \), and the body \( \Omega = (0, 1) \). In addition to the coercitivity of the variational integral in (8), we also assume, similar to (42a), that

\[
W_{\text{mean}}(u') \geq \lambda \, \text{dist}(u', [\pm 1])^2, \quad \lambda > 0.
\]

This is enough to guarantee that any minimizing sequence will converge macroscopically to a function \( g \), i.e.:

\[
u_n \rightarrow g \text{ weakly in, e.g., } W^{1,2}(0, 1).
\]

In order to obtain a computationally feasible problem, we insist that among all (uncountably many) possible minimizing sequences having the property (89), the ones which become asymptotically weak white noise...
are computationally desirable. Our approach consists in the introduction of a $W_{\text{micro}}(x, t)$ term which can be relaxed only by such sequences. We demonstrate a possible construction on a one-dimensional problem in the framework of finite element approximation. Let us assume that $u_0$ is an element of some finite element space defined on a regular mesh with size $h$. Then we define

$$W_{\text{micro}}(u_k(x)) \triangleq \frac{1}{N^2} \sum_{i=0}^{N/2} (c_i^k(x) - 1)^2,$$

where

$$c_i^k = N_0 h^2 \int_0^1 u_i'(x) - g_i(x) \frac{e^{-2\pi ik x}}{\sqrt{1-g_i^2(x)}} \, dx.$$  

Here $g_i$ represents the projection of the weak limit $g$ in the given finite element space. We evaluate the coefficients $c_i^k$ using the fast Fourier transform. Let us set $h = 1/n$.

**Theorem 8.1.** There exist asymptotically weak white noises sequences $\{u_{k,n}(x)\}_{k,n}$.

**Proof.** We give a sketch of the proof. We construct independent random variables $Y_{i_1}, Y_{i_2}, \ldots, Y_{i_n}$ taking value $\pm 1$ with $P[Y_{i_1} = 1] = (1 + g'(x)/2)$ and $P[Y_{i_1} = -1] = (1 - g'(x)/2)$. Let

$$u_{k,n}(x) = \sum_{i=1}^{N_0} Y_i \phi_i[x/(1/n)](x).$$  

where $\phi_A(x)$ denotes the characteristic function of the set $A$, i.e., $\phi_A(x) = 1$ if $x \in A$ and $\phi_A(x) = 0$ if $x \notin A$. Standard probabilistic arguments are used to show that this is an asymptotically weak white noise with probability one; hence there exist sequences satisfying the definition. □

Suppose now that $u_k$ is a statistical white noise, meaning that

$$E \left[ \frac{u_k(x) - g_k(x)}{\sqrt{1 - g_k^2(x)}} \right] = \begin{cases} 1 & \text{if } x \text{ and } y \text{ are in the same } A_j, \\ 0 & \text{otherwise}. \end{cases}$$

Here, $A_j$ are partition intervals. Then

$$E \left[ (u_k(x) - g_k(x))^2 \right] = N_0 \int_0^1 \int_0^1 E \left[ \frac{u_k(x) - g_k(x)}{\sqrt{1 - g_k^2(x)}} \frac{u_k(y) - g_k(y)}{\sqrt{1 - g_k^2(y)}} \right] e^{-2\pi i k (x-y)} \, dx \, dy.$$  

By the white noise property, the expectation inside the integral is 0 unless $x$ and $y$ belong to the same interval.

---

Fig. 4. The solutions (left) and their derivatives (right) using both the standard descent method (top) and the proposed SVP method (bottom). Note that the standard descent method produces considerably wilder behavior in both the solution and derivative.
If we suppose that the partition intervals are uniform, i.e., $A_j = [a_j, b_j) = [(j - 1)/N, j/N)$, then the last expression is

$$
\frac{N^2}{2\pi^2 k^2} \left[ 1 - \cos \left( \frac{2\pi k}{N} \right) \right] = \frac{N^2}{4\pi^2 k^2} \left[ \frac{2\pi^2 k^2}{N^2} + \mathcal{O} \left( \frac{1}{N^3} \right) \right] = \frac{1}{2} + \mathcal{O} \left( \frac{1}{N^2} \right)
$$

in the partition, in which case the expectation is 1.

Thus

$$
E\left( \chi_h^2 \right) = N_h \sum_{j=1}^{N_h} \int_{A_j} \int_{A_j} \exp\left(-2\pi i k(x - y)\right) dx \, dy
$$

$$
= \frac{1}{2\pi^2 k^2} \sum_{j=1}^{N_h} \left( 1 - \cos \left( 2\pi k(b_j - a_j) \right) \right).
$$

(94)
as $N \to \infty$. Hence, we see that the statistical white noise property implies that “on average” the $W_{\text{micro}}$ term will be minimized. In fact, one can show that if $a_{k}^\ast$ is a statistical white noise, then the random variables $\varepsilon_{k}^\ast$ will not converge to 1 with probability 1—the variance does not even go to 0. This is related to the fact that the periodogram is not a consistent estimator of the spectral density function. We may be able to achieve better results by using an analogue of the spectral density estimates which “smooth” the periodogram. More discussion on these topics may be found in [3].

9. One-dimensional computational example

Using the SVP introduced in Section 8, we attempt to solve the simple one-dimensional problem with constant target function $g(x)=0$, $x \in [0,1]$. We use gradient-based relaxation. Using a finite element discretization containing a regular mesh of 2048 elements, the proposed variational principle indeed performed much better than descent algorithm, Definition 4.5, applied to non-perturbated functional in the generation of the proper microscopic structure. Figs. 4–6 show the differences.

Acknowledgements

Cox was supported in part by the grant NSF DMS–9971797. The other two authors were supported in part by the grant NSF DMS–0107539, by the Los Alamos National Laboratory Computer Science Institute (LACSI) through LANL contract number 03891-99-23, as part of the prime contract W-7405-ENG-36 between the Department of Energy and the Regents of the University of California, by the grant NASA SECTP NAG5–8136, and by the grant from TRW Foundation.

References

Dennis D. Cox received his PhD in Mathematics from the University of Washington in 1980. He has held positions in mathematics and statistics at the Universities of Wisconsin and Illinois before joining the Department of Statistics at Rice University in 1993. His primary research interests are in statistical function estimation, spatial statistics, design and analysis of computer experiments, Monte Carlo methods, and stochastic dynamical systems.

Petr Klouček received his PhD in Mathematics from the Charles University, Prague, Czech Republic, in 1991. He has held positions in ETH Zurich, EPF Lausanne and University of Minnesota before he joined Department of Computational and Applied Mathematics at Rice University in 1996. His primary research interests are nonlinear partial differential equations applied to problems in materials science, finite element method applied to approximation of measure-valued solutions of non-quasiconvex problems and mathematical modeling of electrolytic processes.

Daniel R. Reynolds received his PhD in May 2003 from the Rice University for his work on a thermodynamic model governing the physics of shape memory alloy wires.