The Computational Modeling of Crystalline Materials Using a Stochastic Variational Principle [⋆]

Dennis Cox[1], Petr Klouček[2], and Daniel R. Reynolds[2]

[1] Department of Statistics, Rice University, 6100 Main Street,
Houston, TX 77005, USA
[2] Department of Computational and Applied Mathematics, Rice University,
6100 Main Street, Houston, TX 77005, USA

Abstract. We introduce a variational principle suitable for the computational modeling of crystalline materials. We consider a class of materials that are described by non-quasiconvex variational integrals. We are further focused on equilibria of such materials that have non-attainment structure, i.e., Dirichlet boundary conditions prohibit these variational integrals from attaining their infima. Consequently, the equilibrium is described by probablity distributions. The new variational principle provides the possibility to use standard optimization tools to achieve stochastic equilibrium states starting from given initial deterministic states.

1 INTRODUCTION

Recently, there has been a great deal of research into the modeling and use of so-called Smart Materials. These materials have the ability to undergo internal physical transformations, which may be used to do work in ways and places that traditional engineering materials cannot. Such new materials include composites, ceramics, liquid crystals, biomaterials, ferromagnetics and shape memory alloys. Advances in micro-machines, damping mechanisms, high-resolution displays and superconductors are a few of the applications for which they are being designed.

These materials are singled out by their unique ability to undergo a temperature dependent crystalline phase transformation, known as the Martensitic Transformation. At higher temperatures Shape Memory Alloys exhibit a stiff, cubic crystalline lattice structure known as the Austenitic phase. In lower temperatures, these alloys change to exhibit a more easily deformable tetragonal lattice structure, having many crystallographically equivalent states. The reason for the name comes from the process where, if one begins by establishing a reference configuration in the Austenitic phase, then cools

it to the Martensitic phase and deforms the body, the material will resume to the original
"remembered" reference shape when the temperature is raised.

At the root of this transformation is the structure Helmholtz free energy for these
materials. At high temperatures, this energy has a quadratic structure with respect to
the lattice deformation gradient. Thus at the thermodynamically stable minimum, the
material will exhibit one stable, undeformed state. At low temperatures, however, the
Helmholtz free energy has many minima corresponding to each of the stable Martensitic
states (24 in the case of Nickel-Titanium).

Furthermore, various engineering constants such as the specific heat and elastic
modulus change depending on which state the material is in. Thus if some of a Shape
Memory Alloy is in the Martensitic phase, while the rest is in the Austenitic phase,
these constants can vary throughout the body. Hence models using Shape Memory Al-
loys must encompass the macroscopic scale of the overall system, while also retaining
information about the microscopic scale of the crystal lattice.

2   EQUILIBRIUM CONFIGURATIONS OF CRYSTALLINE MATERIALS

Shape memory materials are a notorious example of Crystalline Materials. These alloys
develop a striated microcrystalline structure: a mosaic of lamella of coherent compound
twins, formed by alternation of phases with symmetry related atomic lattices. Math-
ematically, the deformation gradients are constrained to attain values from a certain
finite set of matrices while subjected to Dirichlet boundary conditions. Their equilib-
rium properties can thus be described by differential inclusions. Given a set of matrices
$\mathcal{A} \subset \mathbb{R}^{m \times n}$, and given a boundary function $g \in W^{1,\infty}(\Omega, \mathbb{R}^m)$, we seek a Lipschitz
map u such that

$$
\begin{aligned}
Du \in \mathcal{A}, &\quad \text{a.e. in } \Omega, \\
u = g, &\quad \text{on } \partial\Omega,
\end{aligned}
\tag{1}
$$

where $\Omega \subset \mathbb{R}^n$, $n = 1, 2, 3$, is an open set, $u : \Omega \mapsto \mathbb{R}^m$, and Du is a matrix of the
first derivatives in $\mathbb{R}^{m \times n}$.

The existence of solution(s) to (1) is far from being understood, especially for
$m > 2$. We refer to [1] for extensive treatment of this problem. It seems that if $Dg \in$
Interior co-$\mathcal{A}$ then the problem has a dense set of solutions (in the sense of Baire cate-
gory) while if $Dg \in \partial$ co-$\mathcal{A}$ then the solution can be characterized only in the stochastic
sense. If $\text{dist}(Dg, \text{co-}\mathcal{A}) > 0$ then there is no solution. The symbol "co-" stands for
various convex hulls of $\mathcal{A}$: quasi-convex, polyconvex, and rank-1 convex hull.

The differential inclusion (1) can be posed as a constrained minimization problem

$$
\inf \left\{ \int_\Omega W(Dv(x)) \, dx \,|\, v \in W^{1,\infty}(\Omega, \mathbb{R}^m), v = g, \text{ on } \partial\Omega \right\},
\tag{2}
$$

where W is a smooth positive density that vanishes on $SO(n)\mathcal{A}$, $SO(n)$ denotes the
set of Simple Orthogonal rotations in $\mathbb{R}^n$.

We address the following problem. Let us assume that $Dg \in \partial$ co-$\mathcal{A}$, and let as
assume that this condition guarantees existence of a minimizing sequence $\{v_n\}_{n \in \mathbb{N}}$

of the problem (2) which converges weakly-$*$ to a macroscopic state. The assumption $Dg \in \partial$ co-$\mathcal{A}$ is compatible with lack of lower semicontinuity, i.e. the infimum in (2) is not expected to be attainable. We wish to construct a *stochastic variational principle* which guarantees that standard optmization tools can be used to relax the constrained minimization problem (2).

2.1   LATTICE SYMMETRY PHASE CHANGE

Alloys such as Nickel-Titanium or Indium-Thallium are typical examples of Shape Memory Materials. These alloys exhibit multiscale domain patterns described by gradients of weakly differentiable maps $u : \mathbb{R}^n \to \mathbb{R}^n$ that can come close to the solution of the following *differential inclusion*:

$$
\begin{aligned}
& Du \in \{F_1, F_2\}, \qquad \text{a.e. in } \Omega \subset \mathbb{R}^n, \quad n = 2, 3, \\
& F_i \in M^{n \times n}, \quad i = 1, 2, \\
& u(x) = (\lambda_1 F_1 + \lambda_2 F_2)\, x, \quad x \in \partial\Omega, \\
& \lambda_i > 0, \quad i = 1, 2, \quad \lambda_1 + \lambda_2 = 1, \\
& F_2 = F_1 + a \otimes b, \qquad a, b \in \mathbb{R}^n, \quad (a \otimes b)_{ij} = a_i b_j.
\end{aligned} \tag{3}
$$

The matrices F_i are assumed to be positive definite and linearly independent. Typical examples of these matrices are associated with crystallographic theories, [3]. We note that there does not exist any functional representation for the solution of (3). This is because it is necessary to create oscillations in the gradients with unlimited frequency to meet the boundary condition between the Austenite and Martensite states.

3   REPRESENTATION OF NON-ATTAINABLE STATES

Typical minimizing sequence of (2) will converge weakly, likely weakly-$*$, to its expected value, i.e., an average, in an appropriate Sobolev space. Since we do not expect the variational integral in (2) to be weakly lower semicontinous, we assume that

$$
\liminf_{n \to \infty} \int_{\Omega} W(Du_n(x))\, dx < \int_{\Omega} W(Du(x))\, dx. \tag{4}
$$

Here, $\{u_n\}_{n \in \mathbb{N}} \subset W^{1,\infty}(\Omega)$ is a minimizing sequence and $u \in W^{1,\infty}(\Omega)$ is its weak-$*$ limit. Translation of (4) into the framework of material science would imply that the function u representing, e.g., the averaged deformation of a composite or Shape Memory Alloy, does not carry any pointwise information. Yet, a precise description of the spatial composition of deformation gradients is needed to access the elastic properties of these materials, and is also necessary as an input for thermodynamical models [5].

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 $\overline{W} \in L^1(\Omega)$ such that

$$
\lim_{n \to \infty} \int_{\omega} W(Du_n(x))\, dx = \int_{\omega} \overline{W}(x)\, dx, \tag{5}
$$

for any ω which is a compact subset of Ω. The function $\overline{W}$ does not remember anything about the equilibrium structure of the material. This is encoded into the density W. Though more importantly, the function $\overline{W}$ does tell us about the pointwise distribution of the elastic energy density W. We can obtain the spatial microstructure of the gradients Du_n for large n by comparing at a given point x the value $\overline{W}(x)$ with $W(s)$ for some $s \in \mathbb{R}^{m \times n}$ belonging to an equlibrium set of W.

The function $\overline{W}$ can be obtained by integral representation using a Radon probablity measure μ_x. Namely,

$$\overline{W}(x) = \int_{\mathbb{R}^{m \times n}} W(y) \, d\mu_x(y). \tag{6}$$

The representation (6) can be derived by a standard Riezs argument [6]. We note that if the minimizing sequence $\{u_n\}_{n \in \mathbb{N}}$ would converge strongly to its weak limit then

$$\mu_x = \delta_{Du(x)}, \qquad \text{for almost all } x \in \Omega, \tag{7}$$

where $\delta_{Du(x)}$ is the Dirac measure assigning 1 to an open subset $\mathcal{D}$ of $\mathbb{R}^{n \times m}$ if $Du(x) \in \mathcal{D}$ and 0 otherwise. This can be seen immediately from the approximation

$$W(Du)(x) \sim \sum_i W(y_i) \delta_{Du(x)}(\Delta y_i) \sim \int_{\mathbb{R}^{m \times n}} W(y) \, d\delta_{Du(x)}(y). \tag{8}$$

The key observation is that the Radon measure μ_x can be approximated. Indeed, we can construct a measure applicable to any Borel subset of the domain Ω by, for almost all x,

$$\mu_{x,R,Du_n}(\mathcal{D}) \overset{\text{def}}{=} (\text{meas}(B_R(x)))^{-1} \, \text{meas}\left(\{y \in B_R(x) \,|\, Du_n(x) \in \mathcal{D}\}\right). \tag{9}$$

It is quite direct to show that [3]

$$\lim_{R \to 0_+} \lim_{n \to \infty} \int_{\mathbb{R}^{n \times m}} f(y) \mu_{x,R,Du_n}(y) = \int_{\mathbb{R}^{n \times m}} f(y) \, d\mu_x(y),$$
$$\text{for any } f \in L^1\left(\mathbb{R}^{n \times m}\right). \tag{10}$$

First since

$$\int_{\mathbb{R}^{n \times m}} f(y) \mu_{x,R,Du_n}(y) \sim \sum_i f(y_i) \mu_{x,R,Du_n}(\Delta y_i), \tag{11}$$

and realizing that

$$\sum_i f(y_i) \mu_{x,R,Du_n}(\Delta y_i) \sim (\text{meas}(B_R(x)))^{-1} \int_{B_R(x)} f(Du_n(y)) \, dy, \tag{12}$$

and defining the averaged measure

$$\mu_{x,R} \overset{\text{def}}{=} (\text{meas}(B_R(x)))^{-1} \int_{B_R(x)} \mu_x \, dy, \tag{13}$$

we obtain from (11)-(13), and in view of the weak-$*$ convergence

$$f(Du_n) \rightharpoonup \overline{f} = \int_{\mathbb{R}^{n \times m}} f(y)\mu_x(y), \tag{14}$$

that

$$
\begin{aligned}
&\lim_{n \to \infty} \int_{\mathbb{R}^{n \times m}} f(y)\mu_{x,R,Du_n}(y) \\
&= \lim_{n \to \infty} (\text{meas}(B_R(x)))^{-1} \int_{B_R(x)} f(Du_n(y))\, dy \\
&= \lim_{n \to \infty} (\text{meas}(B_R(x)))^{-1} \int_{B_R(x)} \int_{\mathbb{R}^{n \times m}} f(y)\, d\mu_x(y)\, dy \\
&= \lim_{n \to \infty} \int_{\mathbb{R}^{n \times m}} f(y)\mu_{x,R}(y)\, dy.
\end{aligned} \tag{15}
$$

It follows from the Lebesque differentiation theorem that

$$\lim_{R \to 0_+} \lim_{n \to \infty} \int_{\mathbb{R}^{n \times m}} f(y)\mu_{x,R}(y)\, dy = \int_{\mathbb{R}^{n \times m}} f(y)\, d\mu_x(y) \tag{16}$$

which verifies (10).

The Radon measure μ_x has a very natural structure in the case of various alloys undergoing symmetry phase change described in Section 2.1. The structure can be seen quite directly from its approximation (9). It is possible to show that

$$\mu_x = \lambda(x)\delta_{F_1} + (1 - \lambda(x))\delta_{F_2} \tag{17}$$

in the case (3), [4]. The corresponding smooth and positive energy density W must then vanish on the set

$$\{RF_i^T F_i R^T \mid R \in SO(n)\}. \tag{18}$$

The function $\lambda = \lambda(x)$ represents the probablity that at a point x the material is deformed with the deformation gradient F_1. Usually, this ratio is called *volume fraction*. It follows directly from the definition (9) that

$$\lambda(x) = \lim_{r \to 0_+} \lim_{R \to 0_+} \lim_{n \to \infty} \frac{\text{meas}\left(\{y \in B_R(x) \mid \|Du_n(x) - F_1\| \le r\}\right)}{\text{meas}(B_R(x))}. \tag{19}$$

For (19) to be valid, the minimizing sequence must become stochastic in its derivative. This represents a major challenge for the optimization, numerical and computational approaches. Desirable methods are those which can generate a stochastic state from the initial deterministic one. It is easy to construct such sequences using self-similar *periodic* construction. It is however impossible to reconstruct such sequences computationally using "off-the-shelf" tools. We address this issue briefly in the next Section 4. We strive to compute the volume fraction using a *stochastic variational principle* that will guarantee that any minimizing sequence becomes asymptotically a *weak white noise*, c.f., Section 5. This seems to be the key selection principle. Such sequences can be obtained computationally despite all the limitations implied by any kind of finite dimensional approximation, which prohibit the convergence otherwise.

4 WEAK WHITE NOISE

We have shown in [2] that the Steepest Descent applied to (2) with $\mathcal{A}$ corresponding to the double-well problem, i. e. $\mathcal{A} = F_1 \cup F_2$, $\operatorname{rank}(F_1 - F_2) = 1$, $n = 2, 3$, does not generate relaxing sequences. In other words, there exists a deterministic limiting state corresponding to a deterministic initial guess and not a limiting stochastic state describing the infimum of the energy. This means in terms of any minimizing sequence $\{u_n\}_{n\in\mathbb{N}}$ that it converges strongly to a function in the norm of some Sobolev space. In this case, the reconstruction procedure (19) for computing the volume fraction cannot provide a reliable approximation.

The structure of the argument used in [2] indicates that the Steepest Descent itself is not at fault. Rather, it is the construction of the minimization problem. The reason is that an energy density that is absolutely stable on $SO(n)\mathcal{A}$ ought to include coupling between the micro- and meso-scales in order to provide a pathway on the microscale local lattice distortion level to connect the absolutely stable states.

In the computational practice, the initial iteration for the Steepest Descent is obtained by a superposition of the averaged state (the weak limit) with superimposed numeric noise. Typically this leads almost immediately to a convergence of the minimizing sequence to the nearest local minimum. The evalution of the volume fraction based on such sequences using (19) yields results with almost 50% error. We demonstrate this phenomenon in Section 6.

The variational principle we introduce imposes a forcing mechanism prohibiting the minimizing sequences to adhere to any state where possible spatial correlation can occur. This mechanism prohibits the minimizing sequence to converge to any of the many (possibly countably many) local minima of the variational integral (2).

The following Definition establishes how we interpret the notion of weak white noise in the discussed framework.

Definition 41 *We say that a sequence $\{u_n\}_{n\in\mathbb{N}} \subset W^{1,\infty}(\Omega, \mathbb{R}^m)$, $\Omega \subset \mathbb{R}^n$, $n = 1, 2, 3$, $m = 1, 2, 3$, that converges weakly-* in $W^{1,\infty}(\Omega, \mathbb{R}^m)$ to a Lipschitz continuous function g, becomes asymptotically weak white noise if the following three conditions hold true. Let*

$$z_n(x) \overset{def}{=} Du_n(x) - Dg(x) \in \mathbb{R}^{m\times n}. \tag{20}$$

The first condition is

$$E_\omega[z_n(x)] = \frac{1}{\operatorname{meas}(\omega)} \int_\omega z_n(x)\, dx \to 0, \qquad \textit{for all open } \omega \subset \Omega, \textit{ as } n \to \infty.$$

The second and the third conditions are that

for almost all $\tau \in \mathbb{R}$ the covariance operator E has the following two properties

$$E[z_n(x) \otimes z_n(x + \tau), \omega] = \tfrac{1}{\operatorname{meas}(\omega)} \int_\omega z_n(x) \otimes z_n(x + \tau)\, dx\, \sigma^{-1}(\omega)$$

$$\to \begin{cases} I \in \mathbb{R}^{m^2 \times n^2}, & \textit{if } \tau = 0, \\ 0 \in \mathbb{R}^{m^2 \times n^2}, & \textit{if } \tau \neq 0, \end{cases} \qquad \textit{for all open } \omega \subset \Omega, \textit{ as } n \to \infty,$$

$$\tag{21}$$

where σ represents "standard deviation" given by

$$\sigma(\omega) \overset{\text{def}}{=} E[z_n \otimes z_n, \omega]^{1/2}. \tag{22}$$

We assume that the functions Du_n are extended periodically onto $\mathbb{R}^{m \times n}$ for $u_n(x + \tau)$ to be defined where $x + \tau \notin \Omega$. We note that E is non-negative.

5 ONE DIMENSIONAL STOCHASTIC VARIATIONAL PRINCIPLE

We define the Principle in one spatial dimension in order to remove unnecessary technical issues. We assume that the density $W = W(u')$ in (2) should be written as

$$W(x, s) \overset{\text{def}}{=} W_{\text{meso}}(s) + W_{\text{micro}}(s), \qquad s \in \mathbb{R}^1. \tag{23}$$

The contribution W_{meso} encodes the information about the equlibrium state of a given material and the W_{micro} contribution guarantees that any minimizing sequence becomes *weak white noise* in the sense of Definition 41.

In addition to the coercitivity of the variational integral in (2), we also assume that

$$W_{\text{meso}}(u') \geq \Lambda \operatorname{dist} \{u', \{\pm 1\}\}, \qquad \Lambda > 0. \tag{24}$$

This is enough to guarantee that any minimizing sequence will convenverge macroscopically to a function g, i.e.,

$$u_n \rightharpoonup g \qquad \text{weakly in e.g. } W^{1,2}\left((0,1)\right). \tag{25}$$

In order to obtain a computationally feasible problem, we insist that among all (uncountably many) possible minimizing sequences having the propoerty (25), the ones which become asymptotically weak white noise are computationally desirable. Our approach consists in the introduction of a $W_{\text{micro}}(x, s)$ 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_h is an element of some finite element space defined on a regular mesh with size h. Then we define

$$\begin{aligned} W_{\text{micro}}(u_h')(x) &\overset{\text{def}}{=} \frac{1}{N_h} \sum_{k=0}^{N_h/2} \left(z_h^k(x)\overline{z}_h^k(x) - 1\right)^2, \qquad \text{where} \\ z_h^k(x) &\overset{\text{def}}{=} \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{u_h'(x) - g_h'(x)}{\sqrt{1 - g_h'^2(x)}} \exp(-ikx) \, dx. \end{aligned} \tag{26}$$

Here g_h represents the projection of the weak limit g in the given finite element space. We evaluate the coefficients z_h^k using the fast Fourier Tranform. Let as set $h = 1/n$. We have the following Theorem.

Theorem 51 *There exists asymptotically weak white noises sequences $u_{1/n}'$.*

We give a sketch of the proof. We construct independent random variables Y_{1n}, Y_{2n}, ... Y_{nn} taking value ± 1 with $P[Y_{kn} = 1] = (1 + g'(x))/2$ and $P[Y_{kn} = -1] = (1 - g'(x))/2$. Let

$$u'_{1/n}(x) = \sum_{i=1}^{n} Y_i \phi_{[(i-1)/n, i/n)]}(x), \tag{27}$$

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 that there exist sequences satisfying the definition. From this discussion we see immediately the following

Theorem 52 *Any asymptotically weak white noise sequence $u'_{1/n}$ satisfies*

$$W_{micro}(u'_{1/n}) \to 0 \text{ as } n \to \infty. \tag{28}$$

We conjecture that the converse is also true – i.e. that if $W_{micro}(u'_{1/n}) \to 0$ then $u'_{1/n}$ is an asymptotically weak white noise sequence in the sense of Definition 41.

6   ONE DIMENSIONAL COMPUTATIONAL EXAMPLE

We present in this section a computational example in which we approximate a solution of the following differential inclusion

$$\begin{aligned} u'(x) &\in \{-1, 1\}, \qquad \text{a.e. in } (0, 1), \\ u(x) &= 0, \qquad \text{for all } x \in (0, 1). \end{aligned} \tag{29}$$

We construct the approximate solutions as minimizing sequences $\{u_n\}_{n\in\mathbb{N}} \subset W^{1,4}(0, 1)$ generated by the Steepest Descent minimization of the functional

$$\begin{aligned} &\mathcal{J}(u_h, h^{3/2}) \\ &\stackrel{\text{def}}{=} \frac{1}{2} \int_0^1 |u_h(x)|^2 \, dx + \frac{1}{4} \int_0^1 \left| u'_h(x)^2 - 1 \right|^2 \, dx + h^{3/2} \int_0^1 W_{\text{micro}}(u'_h) \, dx, \end{aligned} \tag{30}$$

where the density W_{micro} is defined by (26). We note that the differential inclusion (29) is the one dimensional analog of (1) with $\mathcal{A} = \{-1, 1\}$.

The calculations described in the caption to Figure 1 show clear advantage in using the stochastic term in getting the appropriate relaxing sequences.
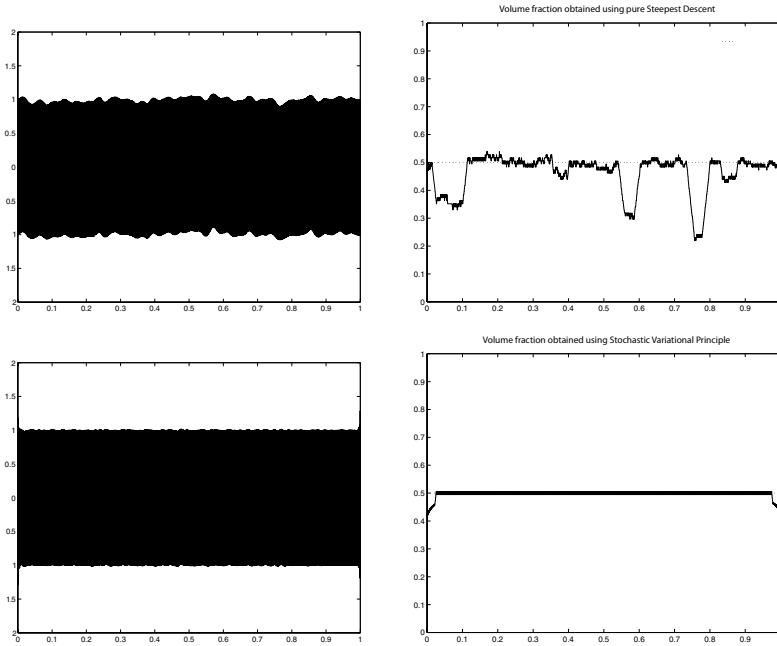
Fig. 1. These calculations are done on the regular mesh with $h = 1/2048$. The upper two pictures show the derivatives (left) and the volume fraction $\lambda = \lambda(x)$ (right) for the sequence obtained by the Steepest descent algorithm applied to (30) without the contribution of the term $h^{3/2} \int_0^1 W_{\mathrm{micro}}(u'_h) \, dx$. We note that $\lambda(x) = 1/2$ for any $x \in (0, 1)$ for the Differential inclusion (29). The lower two pictures show the same quantities obtained by the Steepest descent method applied to relaxation of the functional (30).

References

1. B. Dacorogna and P. Marcellini, *Implicit partial differential equations*, Birkhäuser, 2000.
2. P. Klouček, *The steepest descent minimization of double-well stored energies does not yield vectorial microstructures*, Rice University (2001), Technical report 01-04, Department of Computational and Applied Mathematics.
3. Luskin M., *On the computation of crystalline microstructure*, Acta Numerica (1996), 191–257.
4. P. Pedregal, *Variational methods in nonlinear elasticity*, SIAM, Philadelphia, 2000.
5. D. Reynolds, *Vibration damping using martenstic phase transformation*, (2002), Ph.D. Thesis, Rice University.
6. W. P. Ziemer, *Weakly differentiable functions*, Graduate Texts in Math., Springer-Verlag, New York, 1989.