NAMT
95-012

An Extended Variational Principle

Richard Jordan
Carnegie Mellon University

David Kinderlehrer
Carnegie Mellon University

Research Report No. 95-NA-012

June 1995

Sponsors

U.S. Army Research Office
Research Triangle Park
NC 27709

National Science Foundation
1800 G Street, N.W.
Washington, DC 20550

To appear in
Topics in partial differential equations and applications: Collected papers in honor of Carlo Pucci,
Marcellini, P., Talenti, G., and Vesentini, E., editors

An Extended Variational Principle

Dedicated to Carlo Pucci

RICHARD JORDAN, Center for Nonlinear Analysis, Carnegie Mellon University, Pittsburgh, PA[1]

DAVID KINDERLEHRER, Center for Nonlinear Analysis, Carnegie Mellon University, Pittsburgh, PA[1]

1 INTRODUCTION

The paradigm of the calculus of variations is the functional

$$I(v) = \int_\Omega W(\nabla v)\, dx. \tag{1.1}$$

In the absence of quasiconvexity of W, I may fail to assume its minimum in a given admissible class, even when this class is very reasonable. In this situation, we may seek to represent the solution in terms of the oscillatory statistics developed by a minimizing sequence, or its Young Measure, [37]. For definiteness, suppose that

$$c(|\lambda|^p - 1)^+ \le W(\lambda) \le C(1 + |\lambda|^p), \quad \lambda \in \mathbb{M}, 0 < c < C, \ 1 < p < \infty. \tag{1.2}$$

where $\mathbb{M}$ denotes $m \times n$ matrices and that $\Omega \subset \mathbb{R}^n$ is a domain with smooth boundary. Given $u_o \in H^{1,p}(\Omega,\mathbb{R}^m)$, consider the variational principle

$$\inf_V I(v) \qquad \text{where } V = u_o + H_0^{1,p}(\Omega,\mathbb{R}^m). \tag{1.3}$$

In recent years we have learned much about minimizing sequences (u^k) of (1.1) and their Young Measures. For example, (u^k) may be chosen so that in addition to

$$u^k \rightharpoonup u \text{ in } V,$$

[1] This research supported by the Air Force Office of Scientific Research, the Army Research Office, and the National Science Foundation

we have that there is a $g \in L^1(\Omega)$ for which

$$|\nabla u^k|^p \rightharpoonup g \text{ in } L^1(\Omega) \tag{1.4}$$

and the Young Measure representation is valid for $\psi \in C(M)$ satisfying

$$|\psi(\lambda)| \leq C(1 + |\lambda|^p) \tag{1.5}$$

This means that there is a family $\nu = (\nu_x)_{x \in \Omega}$ of probability measures on M such that when ψ satisfies (1.5)

$$\psi(\nabla u^k) \rightharpoonup \bar{\psi} \text{ in } L^1(\Omega) \quad \text{where} \quad \bar{\psi}(x) = \int_M \psi(\lambda)\, d\nu_x(\lambda) \text{ in } \Omega \text{ a.e.} \tag{1.6}$$

In particular, the limit gradient ∇u is recaptured as

$$\nabla u(x) = \int_M \lambda\, d\nu_x(\lambda) \text{ in } \Omega \text{ a.e. and} \tag{1.7}$$

$$\int_{\Omega \times M} W(\lambda)\, d\nu_x(\lambda) dx = \lim_{k \to \infty} \int_\Omega W(\nabla u^k)\, dx = \inf_V I(v). \tag{1.8}$$

This suggests introducing a class A of measures μ on $\Omega \times M$, for example, the Young Measures μ generated by a sequences of gradients (∇v^k), $v^k \in V$, and considering, in place of (1.4), the variational principle, with $d\mu(x,\lambda) = d\mu_x(\lambda)dx$,

$$\inf_A I_W(\mu) \quad \text{with} \quad I_W(\mu) = \langle T,\mu \rangle = \int_{\Omega \times M} W(\lambda)\, d\mu(x,\lambda). \tag{1.9}$$

The functional $\langle T,\mu \rangle$ is a linear function on A. Thus every variational principle is linear.

To what extent can we formulate principles in the form (1.9) so they become meaningful? In this note we shall discuss, somewhat informally, this isssue. We shall also address its natural successor, namely, the analysis of questions which depend in some more complicated fashion on the measure and are not linear functionals. To offer a short and readable account, we shall be guided by a precept of Carlo Pucci's old friend Hans Lewy: emphasize the obvious and eschew the difficult.

2 FRAMEWORK AND BASIC PROPERTIES

We introduce a few notations and review briefly the general features of Young Measures. Let

$$E^p = \{ \psi \in C(\mathbf{R}^N): \lim_{|\lambda| \to \infty} \frac{\psi(\lambda)}{1 + |\lambda|^p} \text{ exists} \}, \; 1 \leq p < \infty. \tag{2.1}$$

E^p is a separable Banach Space whose dual we denote by $E^{p\prime}$. For technical reasons, it has an advantage over the inseparable space of functions X^p defined by the inequality (1.5). Given a domain $\Omega \subset \mathbf{R}^n$, note that, [10],

$$L^1(\Omega; E^p)' = L^\infty(\Omega; E^{p'}).$$

A measure $\mu \in L^\infty(\Omega; E^{p'})$ is a parametrized measure or Young Measure if and only if there is a sequence $f^k \in L^p(\Omega;\mathbb{R}^N)$ and $f \in L^p(\Omega;\mathbb{R}^N)$ and $g \in L^1(\Omega)$ such that

$$f^k \rightharpoonup f \text{ in } L^p, \quad |f^k|^p \rightharpoonup g \text{ in } L^1, \text{ and} \tag{2.2}$$

$$\psi(f^k) \rightharpoonup \bar{\psi} \text{ in } L^1(\Omega) \quad \text{where} \quad \bar{\psi}(x) = \int_M \psi(\lambda)\, d\nu_x(\lambda) \text{ in } \Omega \text{ a.e. for } \psi \in E^p.$$

The formula in (2.2) is called the "Young Measure representation" and it permits us to represent the statistics of the oscillations developed by the sequence (f^k) by means of the formula

$$\bar{\psi}(a) = \lim_{\rho \to 0} \lim_{k \to \infty} \frac{1}{|B_\rho|} \int_{B_\rho(a)} \psi(f^k)\, dx, \quad a \in \Omega \text{ a.e.} \tag{2.3}$$

A situation of special interest is when $f^k = \nabla u^k$, that is, when (f^k) is a sequence of gradients. In this case we call μ an $H^{1,p}$ - Young measure, or simply a gradient Young Measure, when the value of p is not of concern.

An important feature of Young Measures as they appear in a variational context is their duality with lower semicontinuous functionals exhibited by means of Jensen's Inequality [24,25,26]. The characterizing property of parametrized measures in $L^\infty(\Omega; E^{p'})$ is Jensen's Inequality for convex functions, cf. (2.11). Gradient Young Measures are dual to quasiconvex functions: a parametrized measure $\nu \in L^\infty(\Omega; E^{p'})$ is a gradient Young Measure if and only if

(i) $$\varphi(F(a)) \le \int_M \varphi(\lambda)\, d\nu_a(\lambda) \quad \text{where} \quad F(a) = \int_M \lambda\, d\nu_a(\lambda) \text{ in } \Omega \text{ a.e. for every quasiconvex } \varphi \in E^p, \tag{2.4}$$

(ii) there is a $u \in H^{1,p}(\Omega,\mathbb{R}^m)$ with $F(x) = \nabla u(x)$, and (2.5)

(iii) $$\Psi(x) = \int_M |\lambda|^p\, d\nu_x(\lambda) \in L^1(\Omega). \tag{2.6}$$

Condition (i) is akin to a local condition, [25,26],

(i)' ν_a is a homogeneous gradient Young Measure for $a \in \Omega$ a.e. (2.4)'

There is an extensive literature on this subject beginning with L. C. Young's own interpretation in control theory [37]. Unfortunately, we lack space to adequately cite all the important recent work. These methods were enhanced and generalized by Tartar [34,35], who studied conservation laws and compensated compactness. The past ten years have witnessed further extensions. The tool of the Young Measure has become fundamental to the study of microstructure in solids, where the weak continuity properties of the minors of (∇u^k) have made it possible to establish kinematical restrictions on minimum energy configurations. This has led to new understanding of structural phase transformations, [4,5,6,7,8,9,17,18,19]. The interpretation of coherent structures in turbulence by means of statistical equilibrium theory has led to maximum entropy principles and a Young Measure description of most probable states via the theory of large deviations, [20,21,31,32,36]. The Young Measure is an extremely useful device for understanding nonlinear processes across widely disparate length scales.

Although the presentation here may be viewed in some ways as a synthesis of the two areas discussed above, our motivation was an attempt to investigate the nature of metastability in some physical systems. We treat this in a separate paper. Details and extensions of our discussion here are in [22,23].

The first issue at hand is what happens, in general, to a bounded sequence of measures in $L^\infty(\Omega; E^{p'})$. A sequence $\mu^k \in L^\infty(\Omega; E^{p'})$ of parametrized measures with

$$\| \mu^k \| = \int_{\Omega\times\mathbf{R}^N} (1 + |\lambda|^p)\, d\mu^k(x,\lambda) \leq M < \infty \tag{2.7}$$

admits a subsequence, not relabeled, such that

$$\mu^k \overset{*}{\rightharpoonup} \tau \in L^\infty(\Omega; E^{p'}), \tag{2.8}$$

where τ need not be a probability measure on $\mathbf{R}^N$, as we know well. Nonetheless, we may isolate from τ a probability measure μ, and it is this we wish to clarify. For simplicity, assume that the μ^k are homogeneous, i.e., independent of $x \in \Omega$, so we may regard $\mu^k \in E^{p'}$. It follows from elementary methods that there is a measure $\mu \in C_o(\mathbb{R}^N)'$ such that

$$\mu^k \overset{*}{\rightharpoonup} \mu \text{ in } C_o(\mathbb{R}^N)'. \tag{2.9}$$

We claim

(a) $\mu \in E^{p'}$ and

(b) $\displaystyle\int_{\mathbf{R}^N} \psi\, d\mu = \lim_{k\to\infty} \int_{\mathbf{R}^N} \psi\, d\mu^k$ whenever $\psi \in E^p$ with

$$\lim_{|\lambda|\to\infty} \frac{\psi(\lambda)}{1 + |\lambda|^p} = 0.$$

Part (a) follows from the Monotone Convergence Theorem. Part (b) follows from (a) and the bound (2.7). In particular

$$\int_{\mathbf{R}^N} d\mu = \lim_{k\to\infty} \int_{\mathbf{R}^N} d\mu^k = 1,$$

so μ is a probability measure.

Returning now to the characterization of $\tau \in E^{p'}$ with

$$\mu^k \overset{*}{\rightharpoonup} \tau \text{ in } E^{p'},$$

it is easy to check that $\tau \geq 0$ and may be expressed (as the measure on $\mathbf{R}^N \cup \{\infty\}$)

$$\tau = c\,\delta_\infty + \mu, \quad c \geq 0,$$

where $\mu \in E^{p'}$ satisfies (a) and (b),

$$\langle \delta_\infty, \psi \rangle = \lim_{|\lambda| \to \infty} \frac{\psi(\lambda)}{1 + |\lambda|^p}, \quad \text{and} \quad c = \lim_{r \to \infty} \langle T, \chi_{\{|\lambda| \geq r\}} |\lambda|^p \rangle.$$

Because of this, we are justified in referring to μ as the *probability measure determined by the* (μ^k). Similarly, for $\tau \in L^\infty(\Omega; E^{p'})$ satisfying (2.8), there is a parametrized family of probability measures $\mu = (\mu_x)_{x \in \Omega} \in L^\infty(\Omega; E^{p'})$ and $\gamma \in L^1(\Omega)$ such that

$$\tau = \gamma \delta_\infty + \mu, \quad \gamma \geq 0, \quad \text{and}$$

$$\int_{\Omega \times \mathbf{R}^N} \psi \, d\mu = \lim_{k \to \infty} \int_{\Omega \times \mathbf{R}^N} \psi \, d\mu^k \quad \text{whenever} \quad \psi \in L^\infty(\Omega; E^p) \text{ with} \tag{2.10}$$

$$\lim_{|\lambda| \to \infty} \frac{\psi(x, \lambda)}{1 + |\lambda|^p} = 0 \text{ a.e. in } \Omega$$

The characterization theorem for Young Measures ensures that the conditions (2.2) are satisfied: there is a sequence $f^k \in L^p(\Omega; \mathbb{R}^N)$ which generates μ.

It is easy to verify Jensen's Inequality directly in this situation from (a) and (b). Let $\varphi \in E^p$, $p > 1$, be convex and let

$$g(\lambda) = \varphi(\lambda_o) + L \cdot (\lambda - \lambda_o) \leq \varphi(\lambda), \quad \lambda_o = \int_{\mathbf{R}^N} \lambda \, d\mu_{x_o}(\lambda), \; x_o \in \Omega.$$

Then from (b),

$$\varphi(\lambda_o) = \lim_{k \to \infty} \int_{\mathbf{R}^N} g(\lambda) \, d\mu^k_{x_o}(\lambda) = \int_{\mathbf{R}^N} g(\lambda) \, d\mu_{x_o}(\lambda) \leq \int_{\mathbf{R}^N} \varphi(\lambda) \, d\mu_{x_o}(\lambda), \text{ in } \Omega \text{ a.e.} \tag{2.11}$$

This property does not seem so obvious for other constraints, in particular, if μ^k are gradient Young Measures, is μ a gradient Young Measure?

Proposition 2.1 *Let $\mu^k \in L^\infty(\Omega; E^{p'})$ be a weak* convergent sequence of gradient Young Measures and let μ be the parametrized measure determined by the (μ^k). Then $\mu \in L^\infty(\Omega; E^{p'})$ is a gradient Young Measure.*

The localization property (2.4)' makes it possible to assume that the (μ^k) are homogeneous. Each μ^k is generated by a sequence $v^{k,j} \in H^{1,p}(\Omega, \mathbb{R}^m)$ with

$$\| v^{k,j} \|_{H^{1,p}(\Omega)} \leq C.$$

Choosing a diagonal sequence w^k of the $v^{k,j}$ we obtain a sequence which generates μ as an $H^{1,q}$ Young Measure for $q < p$ and as a biting Young Measure, cf. [26]. From Theorem1.1 of [26], we are assured that μ is an $H^{1,p}$ Young Measure. Note that the sequence which generates μ as as an $H^{1,p}$ Young Measure is not in general the (w^k).

The localization feature makes it possible to construct variations of the Proposition. For example, we may specify a subdomain $\Omega' \subset \Omega$ such that $\mu^k|_{\Omega'}$ are gradient Young Measures. Then $\mu|_{\Omega'}$ is also a gradient Young Measure. In the next section we outline a different method.

The variational principle (1.9) may now be placed into a rigorous context. We choose A to be the gradient Young Measures generated by sequences in V, or what is the same, the gradient Young Measures for which $u \in V$ in (2.5).

Proposition 2.2 *Suppose that* $\nu^k \in A \subset L^\infty(\Omega;E^{p'})$ *with*

$$I_W(\nu^k) \to \inf_A I_W(\mu). \tag{2.12}$$

Then there is a subsequence, not relabeled, of the ν^k *and a* $\nu \in A$ *(a gradient Young Measure) such that*

$$\nu^k \overset{*}{\rightharpoonup} \nu \quad \textit{and} \quad I_W(\nu) = \inf_A I_W(\mu). \tag{2.13}$$

From (1.2) we have that (ν^k) are bounded in $L^\infty(\Omega;E^{p'})$ and hence, after extraction of a subsequence, have a weak limit of the form

$$\tau = \gamma\,\delta_\infty + \nu \quad \text{with } \gamma \geq 0 \text{ and } \nu \in A. \tag{2.14}$$

It is easy to check from this that $\gamma = 0$ so that $(2.13)_1$ holds.

Clearly, in this first try at the "every variational principle is linear" idea, the lower bound in (1.2) is crucial, cf. [12,13,28].

3 LOCAL CONSTRAINTS

A local constraint on a parametrized measure might be one determined by some quasiconvex functions rather than all of them. For example, write $\mathbf{R}^N = \mathbb{M}\times\mathbf{R}^d$, where $\mathbb{M}$ denotes $m \times n$ matrices, and write $\lambda = (A,\alpha)$, $A \in \mathbb{M}$ and $\alpha \in \mathbf{R}^d$. Suppose that ν satisfies Jensen's Inequality for functions φ satisfying, $|\Omega| = 1$,

$$\varphi(F,p) \leq \int_\Omega \varphi(F + \nabla\zeta, p + q)\,dx \quad \text{for } \zeta \in H_0^{1,\infty}(\Omega,\mathbf{R}^m) \text{ and}$$
$$q \in L^\infty(\Omega;\mathbf{R}^d) \text{ with } \int_\Omega q\,dx = 0. \tag{3.1}$$

We would expect ν to be generated by a sequence of the form $(\nabla u^k, p^k)$, and this is in fact the case, [14]. Here we wish to give a different result in a similar direction.

Given $0 \leq \varphi \in E^p$ satisfying (3.1), determine the set $K_\varphi \subset L^\infty(\Omega;E^{p'})$ of parametrized measures ν for which

$$\psi(F,p) \leq \int_{\mathbf{M}\times\mathbf{R}^d} \psi(A,\alpha)\,d\nu_x(A,\alpha) \quad \text{where } (F,p) = \int_{\mathbf{M}\times\mathbf{R}^d} (A,\alpha)\,d\nu_x(A,\alpha), \; x \in \Omega \text{ a.e.},$$

whenever $\psi \in E^p$ satisifes (3.1) and $\psi \leq \varphi$. (3.2)

Theorem 3.1 *Let $\mu^k \in K_\varphi$ be a weak* convergent sequence and let μ be the parametrized measure determined by (μ^k). Then $\mu \in K_\varphi$.*

To prove this, we really only have to show that if $\varphi \in E^p$ is nonnegative and satisfies (3.1), then there is a sequence, which we shall call $(\psi^s)^\#$ such that $(\psi^s)^\# \in E^{sp}$, $0 < s < 1$, and $(\psi^s)^\# \to \varphi$, pointwise as $s \to 1^-$, $(\psi^s)^\# \leq \varphi$. For this, introduce

$$\psi^s(A,\alpha) = \begin{cases} \varphi(A,\alpha) & \text{in } \{\varphi(A,\alpha) \leq 1\} \\ \varphi(A,\alpha)^s & \text{in } \{\varphi(A,\alpha) > 1\} \end{cases} \tag{3.3}$$

and its "relaxation"

$$(\psi^s)^\#(A,\alpha) = \inf \int_\Omega \psi^s(A + \nabla\zeta, \alpha + q)\, dx \in E^{sp} \tag{3.4}$$

for $\zeta \in H_0^{1,\infty}(\Omega,\mathbb{R}^m)$ and $q \in L^\infty(\Omega;\mathbb{R}^d)$ with $\int_\Omega q\, dx = 0$. Note that we may choose Ω = Q, a unit cube, and replace the boundary condition on ζ by one of periodicity. To show that $(\psi^s)^\#$ converges to φ, we avail ourselves of the argument of Marcellini and Sbordone [29] to establish that a minimizing sequence for (3.4) may be chosen equi-integrable in L^p. This argument combines the Ekeland Distance Lemma [11] with the Meyers - Elcrat [30] form of the reverse Hölder Inequality.

Obviously, the theorem above implies the other statements we have been discussing.

4 EXTENDED VARIATIONAL PRINCIPLES

We now seek to enlarge the scope of questions amenable to Young Measure methods with a new paradigm. Let

$$\begin{aligned} &\psi \in E^p \text{ satisfy } \psi \geq 0 \text{ and} \\ &\varphi \in C(\mathbb{R}^+) \text{ satisfy } \rho\varphi(\rho) \text{ convex and increasing for large } \rho \text{ and} \\ &\int_\Omega \varphi\left(\frac{d\nu}{d\mu}\right) d\nu \geq 0 \text{ for all probability measures } \nu \in E^{p'}. \end{aligned} \tag{4.1}$$

For a fixed probablility measure $\mu^o \in E^{p'}$, with $d\mu = d\mu^o dx$ and $\sigma > 0$, consider the functional

$$I_\sigma(\nu) = \int_{\Omega\times\mathbf{R}^N} \psi\, d\nu + \sigma \int_{\Omega\times\mathbf{R}^N} \varphi\left(\frac{d\nu}{d\mu}\right) d\nu, \quad \nu \in L^\infty(\Omega;E^{p'}), \tag{4.2}$$

with the convention that $I_\sigma(\nu) = +\infty$ if ν fails to be absolutely continuous with respect to μ. It will be clear in what follows that μ^o need not be a probability measure. The most common choice of φ is $\varphi(\rho) = \log \rho$ and in this case the Kullback Entropy is minus the second integral. (4.1) is satisfied in this case. Also in this case, the minimizer of (4.2) is given by

$$d\nu^\sigma = \frac{1}{Z(\sigma)} e^{-\frac{\psi}{\sigma}} d\mu, \text{ where } Z(\sigma) = \int_{\mathbf{R}^N} e^{-\frac{\psi}{\sigma}} d\mu^o \tag{4.3}$$

is the well known "partition function," when the competition is among all parametrized measures and $Z(\sigma)$ is finite. We wish to make prominent some extremely elementary properties of (4.2).

Theorem 4.1 *Suppose that* $I_\sigma(\nu^*) < +\infty$ *for some* $\nu^* \in K$ *and* $\sigma = \sigma_o$. *Then*

$$\inf_K \int_{\Omega\times\mathbf{R}^N} \psi \, d\nu = \lim_{\sigma\to 0} \inf_K I_\sigma(\nu) \tag{4.4}$$

If $\sigma < \sigma'$, then $I_\sigma(\nu) \le I_{\sigma'}(\nu)$, whence $f(\sigma) = \inf_K I_\sigma(\nu)$ is decreasing as $\sigma \to 0$. Let

$$a = \inf_K \int_{\Omega\times\mathbf{R}^N} \psi \, d\nu$$

and given $\varepsilon > 0$, choose ν such that

$$\int_{\Omega\times\mathbf{R}^N} \psi \, d\nu \le a + \varepsilon.$$

Then

$$\int_{\Omega\times\mathbf{R}^N} \psi \, d\nu + \sigma \int_{\Omega\times\mathbf{R}^N} \varphi\left(\frac{d\nu}{d\mu}\right) d\nu \le a + \varepsilon + \sigma \int_{\Omega\times\mathbf{R}^N} \varphi\left(\frac{d\nu}{d\mu}\right) d\nu$$

and

$$\inf_K I_\sigma(\nu) \le a + \varepsilon + \sigma \int_{\Omega\times\mathbf{R}^N} \varphi\left(\frac{d\nu}{d\mu}\right) d\nu ,$$

whence

$$\lim_{\sigma\to 0} \inf_K I_\sigma(\nu) \le a + \varepsilon.$$

Suppose that the set where ψ assumes its minimum is compact and

$$\inf_K \int_{\Omega\times\mathbf{R}^N} \psi \, d\nu = |\Omega| \min \psi.$$

Assume that we have in hand ν^σ such that $I_\sigma(\nu^\sigma) = \inf_K I_\sigma(\nu)$. Then the (ν^σ) are bounded in $L^\infty(\Omega;E^{p'})$ and thus, according to our discussion, we may select a subsequence (not relabeled) which determines a parametrized measure $\nu^o \in L^\infty(\Omega;E^{p'})$. (Note that ν^o is not generally absolutely continuous with respect to μ.) From the Monotone Convergence Theorem (viz. the argument we used to prove (a) in §2),

$$\int_{\Omega\times\mathbf{R}^N} \psi \, d\nu^o \le \liminf_{\sigma\to 0} \int_{\Omega\times\mathbf{R}^N} \psi \, d\nu^\sigma \le \lim_{\sigma\to 0} I_\sigma(\nu^\sigma) = |\Omega| \min \psi.$$

It follows that $\operatorname{supp} \nu^o \subset \{\psi = \min \psi\}$ is compact and $\nu^\sigma \overset{*}{\rightharpoonup} \nu^o$ in $L^\infty(\Omega;E^{p'})$ and, of course, ν^o realizes the minimum for $\sigma = 0$. Facts like these are well known for the particular case of (4.3) but are proved by direct computation. Passing to the limit as $\sigma \to 0$ gives the stationary distribution of a random variable and is related to the simulated annealing algorithm. In practical situations, it is quite important to choose the sequence of σ carefully, [16].

We now address the existence question.

Theorem 4.2 *Let $K \subset L^\infty(\Omega;EP')$ be a set of parametrized measures which enjoys the closure property*

if $\mu^k \in K$ and $\mu^k \overset{*}{\rightharpoonup} \tau$ in $L^\infty(\Omega;EP')$, then the parametrized measure μ determined by (μ^k) satisfies $\mu \in K$. (4.5)

Assume that (4.1) *is satisfied and that $\varphi(\rho) \to \infty$ as $\rho \to \infty$. Then there exists a parametrized measure minimizer $\nu^\sigma \in K$ of I_σ. Of course $\nu^\sigma << \mu$.*

For a minimizing sequence ν^k, write $d\nu^k = \rho^k d\mu$, $\rho^k \in L^1(\Omega\times\mathbf{R}^N,\mu)$. Then

$$\sup \int_{\Omega\times\mathbf{R}^N} \varphi(\rho^k)\rho^k\, d\mu < \infty \quad \text{with} \quad \lim_{|\rho|\to\infty} \frac{\varphi(\rho)\rho}{\rho} = \infty.$$

This is a well known condition for weak relative compactness in L^1 and it follows from the de la Vallée Poussin Criterion that there is a subsequence of the (ρ^k) weakly convergent in L^1. A standard technique shows that the second term of I_σ is lower semi-continuous with respect to weak convergence in L^1 from which it is easy to show that the weakly convergent subsequence converges to a minimizer.

Theorem 4.3 *Assume that* (4.1) *is satisfied and set* $g(\rho) = \varphi(\rho)\rho$. *Suppose that*

$$\frac{d^2}{d\rho^2} g > 0 \text{ for } \rho > 0 \;, \quad \text{range}\,(g')^{-1} \subset (0,\infty),$$

and that there is a constant α such that

$$\int_{\Omega\times\mathbf{R}^N} \rho^*(\lambda)\, d\mu^o(\lambda) = 1 \quad \text{where} \quad \rho^*(\lambda) = (g')^{-1}\left(\frac{\alpha - \psi(\lambda)}{\sigma}\right).$$

Then $\nu^* = \rho^*\mu$ is an absolute minimizer of I_σ.

Minimization falls into this framework. Let $K \subset \mathbf{R}^N$ be compact and let μ^o denote normalized Lebesgue measure on K. Consider a function $\psi \in C^2$, for example, and the functional

$$I_\sigma(\nu) = \int_K \psi\, d\nu + \sigma \int_K \log\left(\frac{d\nu}{d\mu}\right) d\nu, \quad \nu \in EP',\ \sigma > 0.$$

As mentioned above, the minimizer is given by (4.3). Assuming that ψ has M isolated global minimizers at $\lambda_1, \ldots, \lambda_M$ in K, our theorem asserts that

$$\nu^\sigma \overset{*}{\rightharpoonup} \sum c_i\, \delta_{\lambda_i} \text{ as } \sigma \to 0 \text{ where } c_i \geq 0 \text{ and } \sum c_i = 1.$$

This can also be established by calculating the Taylor expansion of ψ, which reveals that the c_i are related to $\nabla^2\psi(\lambda_i)$. These considerations form the basis of the simulated annealing algorithm for global minimization. More generally, there is an intimate connection between the type of variational principle we are discussing and the stationary Fokker-Planck Equation, [1,15,16].

5 Some examples

5.1 The Langevin function and constrained theory

Consider an ensemble of N identical particles governed in equilibrium by thermal motion in, say, $-1 \leq \xi \leq 1$, each of which tends to orient under a field f with strength $\eta\xi$. The probability distribution $d\nu^\sigma = \rho^\sigma d\xi$ which describes the state of a given particle at "temperature" σ is the extremal of

$$I_\sigma(\nu) = -\int_K f\eta\xi\rho \, d\xi + \sigma\int_K \rho\log\rho \, d\xi, \quad K = [-1,1], \tag{5.1}$$

and is given by the formula (4.3). Note that since K is compact, we may arrange that (4.1) is satisfied. The expected value of the "strength", the state of a particle, is

$$\eta\langle \xi \rangle = \eta \int_K \xi\rho^\sigma d\xi = \eta\, L\left(\frac{f\eta}{\sigma}\right), \text{ where } L(x) = \coth x - \frac{1}{x} \tag{5.2}$$

is the Langevin Function. The expected strength of the ensemble is

$$S = N\eta\, L\left(\frac{f\eta}{\sigma}\right). \tag{5.3}$$

Langevin used this analysis successfully to explain paramagnetism [27]. For small values of x, $L(x) = \frac{1}{3}x + O(x^3)$, and leads to the notion of the susceptibility

$$\chi = \frac{S}{f} = \frac{1}{3}\frac{N\eta^2}{\sigma}. \tag{5.4}$$

Note the interesting scaling properties here. If the N particles are grouped into clusters of size M where each cluster responds as a unit, then the expected strength of a cluster is $M\eta\langle M\xi \rangle$. The expected strength of the ensemble and the susceptibility are

$$S_M = \frac{N}{M}M\eta L\left(\frac{fM\eta}{\sigma}\right) \quad \text{and} \quad \chi_M = \frac{S_M}{f} = \frac{1}{3}\frac{MN\eta^2}{\sigma}. \tag{5.5}$$

Thus the effective susceptibility is enhanced when the particles can act as clusters. This is among the mechanisms considered in magnetic nanocomposites, a subject of current research, [33]. The linear approximation to $L(x)$ breaks down if M is large, which leads to an optimization problem for M.

A second view is given by what we call the "constrained theory," [3]. Here a nonlinear elastic body, for example, is assumed to reside in a collection of potential wells Σ even when subjected to a modest constant field T. This gives rise to the functional

$$\inf_K - \int_{\Omega\times\Sigma} T\cdot\lambda \, d\nu \quad \text{where } K = \text{gradient Young Measures with support in } \Sigma. \tag{5.6}$$

The solution of (5.6) may be approximated by extemals of

$$I_\sigma(\nu) = -\int_{\Omega\times\Sigma} T\cdot\lambda \, d\nu + \sigma \int_{\Omega\times\Sigma} \varphi\left(\frac{d\nu}{d\mu}\right) d\nu,$$

where μ is a fixed reference measure, e.g., a Gaussian. A generalized Langevin Function is given by

$$L\left(\frac{T}{\sigma}\right) = \frac{1}{\sigma}\int_\Sigma T\cdot\lambda \, d\nu^\sigma(\lambda) . \tag{5.7}$$

Thus, in view of Theorem 4.1, the constrained theory may be realized as the zero temperature limit of a system governed by thermal motion confined to a given collection of potential wells.

5.2 Coherent structures identified by maximum entropy principles

We present an example of the analysis related to a problem that arises in modeling coherent structures in 2D microtearing turbulence and in 2D magnetohydrodynamic turbulence. Consider the functional

$$F_\sigma(\nu) = \frac{1}{2}\int_{\Omega\times\mathbb{R}^N} (|A|^2 + |\alpha|^2) d\nu + \sigma \int_{\Omega\times\mathbb{R}^N} \log\left(\frac{d\nu}{d\mu}\right) d\nu \quad (A \in \mathbb{M}^{n\times m}, \alpha \in \mathbb{R}^n) \tag{5.8}$$

$$d\mu = d\mu^o(A,\alpha)dx, \quad d\mu^o(A,\alpha) = \frac{1}{(2\pi)^{N/2}} e^{-\frac{1}{2}(|A|^2 + |\alpha|^2)} \, dA d\alpha \cdot \tag{5.9}$$

a standard Gaussian on $\mathbb{R}^N$, $N = nm + n$. We shall minimize F_σ over the set K defined as the set of parametrized measures $\nu \in L^\infty(\Omega;E^{2'})$ such that ν is an $H^1(\Omega) \times L^2(\Omega)$ Young Measure satisfying

$$J_1(v) = \frac{1}{2}\int_\Omega |v|^2 dx = J_1^o \quad \text{and} \quad J_2(v) = \int_\Omega v\cdot q \, dx = J_2^o, \quad \text{where} \tag{5.10}$$

$$(\nabla v(x), q(x)) = \int_{\mathbb{R}^N} (A,\alpha) \, d\nu_x(A,\alpha) \quad \text{with} \quad (v,q) \in H_o^1(\Omega;\mathbb{R}^m) \times L^2(\Omega;\mathbb{R}^n). \tag{5.11}$$

The statement that ν is an $H^1(\Omega) \times L^2(\Omega)$ Young Measure means that $\nu \in K_\varphi$ whenever φ satisfies the conditions of (3.1)-(3.2). It is easy to check that K has the closure property (4.5) by using Theorem 3.1 and the Rellich Compactness Theorem. Hence by Theorem 4.3, we are assured the existence of

$$\nu^\sigma \in K: F_\sigma(\nu^\sigma) = \min_K F_\sigma(\nu) \tag{5.12}$$

with first moment $(\nabla u^\sigma, p^\sigma)$ and from Theorem 4.1 we know that

$$\inf_{K} \int_{\Omega\times\mathbb{R}^N} \psi \, dv = \lim_{\sigma\to 0} \inf_{K} I_\sigma(v), \quad \text{where} \quad \psi(A,\alpha) = \frac{1}{2}(|A|^2 + |\alpha|^2).$$

Let

$$E(v) = \frac{1}{2}\int_{\Omega\times\mathbb{R}^N} (|A|^2 + |\alpha|^2)dv \quad \text{and} \quad E^\#(v,q) = \frac{1}{2}\int_\Omega (|\nabla v|^2 + |q|^2)dx. \qquad (5.13)$$

By Jensen's Inequality, $E^\#(v,q) \leq E(v)$ when the relation (5.11) holds with equality when $v_x = \delta_{(\nabla v(x), q(x))}$, so it follows that, with A the set of (v,q) which satisfy (5.10) and $(5.11)_2$,

$$\inf_A E^\#(v,q) \leq \inf_K E(v) = \mathit{inf}.$$

From this we see that *inf* is attained at $v^o \in K$ with $v^o_x = \delta_{(\nabla u(x), p(x))}$, where $(u,p) \in A$ minimizes $E^\#$. (Note that the constants J^o_i must be given consistently so that the class K is not empty.) There are multipliers a and b such that

$$u \in H^1_0(\Omega;\mathbb{R}^m): \quad -\Delta u + au + bp = 0 \quad \text{and} \quad bu + p = 0 \ \text{in } \Omega, \qquad (5.14)$$

and so it turns out that u is an eigenfunction corresponding to the smallest eigenvalue of $-\Delta$ with homogeneous boundary conditions.

Although we have invented this problem to illustrate the method, it has several interesting features. The solution v^σ $(\sigma > 0)$ is for each x a Gaussian with mean $(\nabla u^\sigma(x), p^\sigma(x))$, each component having variance $\sigma/(1+\sigma)$, and with independent components, i.e., zero covariance. Let μ^σ denote the Young Measure which is for each x Gaussian with mean $(\nabla u(x), p(x))$, with the same variance $\sigma/(1+\sigma)$, and with independent components. Then

$$F_\sigma(v^\sigma) = (\sigma+1)\, E^\#(u^\sigma,p^\sigma) + \sigma\,|\Omega|\log\left(\left(\frac{1+\sigma}{\sigma}\right)^{N/2}\right) \quad \text{and}$$
$$F_\sigma(\mu^\sigma) = (\sigma+1)\, E^\#(u,p) + \sigma\,|\Omega|\log\left(\left(\frac{1+\sigma}{\sigma}\right)^{N/2}\right). \qquad (5.15)$$

Thus $(u^\sigma,p^\sigma) = (u,p)$ and the Young Measure delivers the limit deformation at each "temperature".

Finally, v^σ minimizes F_σ over the larger set K' of $L^2(\mathbb{R}^N)$ Young measures defined analogously to K with the boundary condition applied to the curl-free part of the appropriate portion of the first moment, as determined by the Helmholtz decomposition.

ACKNOWLEDGEMENTS

We would like to thank Yury Grabovsky and Pablo Pedregal for helpful conversations.

REFERENCES

[1] Aluffi-Pentini, F., Parisi, V., and Zirilli, F. 1985 Global optimization and stochastic differential equations, J. Opt. Theory and Appl., 47, 1-16

[2] Ball, J. M. 1989 A version of the fundamental theorem for Young measures, *PDE's and continuum models of phase transitions,* Lecture Notes in Physics, 344,(Rascle, M., Serre, D., and Slemrod, M., eds.) Springer, 207-215

[3] Ball, J. M. and James, R. 1987 Fine phase mixtures as minimizers of energy, Arch. Rat. Mech. Anal.,100, 15-52

[4] Ball, J. M. and James, R. to appear

[5] Ball, J.M. and James, R. 1991 Proposed experimental tests of a theory of fine microstructure and the two well problem, Phil. Trans. Roy. Soc. Lond. A338, 389-450

[6] Battacharya, K. 1991 Wedge-like microstructure in martensite, Acta Metal., 39, 2431-2444

[7] Battacharya, K. 1992 Self accomodation in martensite, Arch. Rat. Mech. Anal., 120, 201-244

[8] Chipot, M. and Kinderlehrer, D. 1988 Equilibrium configurations of crystals, Arch. Rat. Mech. Anal., 103, 237-277

[9] DeSimone, A. 1993 Energy minimizers for large ferromagnetic bodies, Arch. Rat. Mech. Anal., 125, 99-144

[10] Edwards, R.E. 1965 *Functional Analysis*, Holt, Rinehart, and Winston

[11] Ekeland, I. 1979 Nonconvex minimiztion problems, Bull. AMS, 1.3, 443-474

[12] Fonseca, I. and Maly, J. Relaxation of multiple integrals below the growth exponent, Anal. Nonlineare (to appear)

[13] Fonseca, I. and Marcellini, P. Relaxation of multiple integrals in subcritical Sobelov spaces, J. Geom. Anal.. (to appear)

[14] Fonseca, I., Kinderlehrer, D., and Pedregal, P. 1994 Energy functionals depending on elastic strain and chemical composition, Calc. Var., 2, 283-313

[15] Gardiner, C.W. 1990 *Handbook of Stochastic Methods*, Second Edition, Springer-Verlag

[16] Geman, S. and Hwang, C.-R. 1986 Diffusions for global optimization, SIAM J. Control, 24.5, 1031-1045

[17] James, R. D. and Kinderlehrer, D. Laminate structures in martensite (to appear)

[18] James, R. D. and Kinderlehrer, D. 1989 Theory of diffusionless phase transitions, *PDE's and continuum models of phase transitions,* Lecture Notes in Physics, 344,(Rascle, M., Serre, D., and Slemrod, M., eds.) Springer, 51-84

[19] James, R.D. and Kinderlehrer, D. 1993 A theory of magnetostriction with application to TbDyFe2, Phil. Mag. B, 68.2, 237-274

[20] Jordan, R. 1995 A statistical equilibrium model of coherent structures in magnetohydrodynamics, Nonlinearity, 8, 585-613

[21] Jordan, R. Maximum entropy states and coherent structures in two-dimensional microtearing turbulence (to appear)

[22] Jordan, R. and Kinderlehrer, D. in preparation

[23] Jordan, R., Kinderlehrer, D., and Pedregal, P. in preparation

[24] Kinderlehrer, D. and Pedregal, P. 1991 Charactérisation des mesures de Young associées à un gradient, C.R.A.S. Paris, 313, 765-770

[25] Kinderlehrer, D. and Pedregal, P. 1991 Characterizations of Young measures generated by gradients, Arch. Rat. Mech. Anal., 115, 329-365

[26] Kinderlehrer, D. and Pedregal, P. 1994 Gradient Young measures generated by sequences in Sobolev spaces, J. Geom. Anal., 4, 59-90

[27] Langevin, P. 1905 J. Phys., 4, 678

[28] Marcellini, P. 1991 Regularity and existence of solutions of elliptic equations with p,q growth conditions, J. Diff. Eqns., 90, 1-30

[29] Marcellini, P. and Sbordone, C. 1983 On the existence of minima of multiple integrals in the calculus of variations, J. Math. Pures. Appl., 62, 1-9

[30] Meyers, N. and Elcrat, A. 1975 Some results on regularity for solutions of non linear elliptic systems and quasiregular functions, Duke Math. J., 42 (1), 121-136

[31] Robert, R. 1991 A maximum entropy principle for two-dimensional perfect fluid dynamics, J. Stat. Phys., 65, 531-553

[32] Robert, R. and Sommeria, J. 1991 Statistical equilibrium states for two-dimensional flows, J. Fluid Mech., 229, 291-310

[33] Shull, R. 1995 Magnetic applications of nanocomposite materials, lecture at Carnegie Mellon University

[34] Tartar, L. 1979 Compensated compactness and applications to partial differential equations, *Nonlinear analysis and mechanics: Heriot Watt Symposium, Vol I V*(Knops, R., ed.) Pitman Res. Notes in Math. 39, 136-212

[35] Tartar, L. 1983 The compensated compactness method applied to systems of conservation laws, *Systems of nonlinear partial differential equations* (Ball, J. M., ed) Riedel

[36] Turkington, B. and Jordan, R. Turbulent relaxation of a megnetofluid: a statistical equilibrium model, to appear

[37] Young, L. C. 1969 *Lectures on calculus of variations and optimal control theory*, W.B. Saunders