# Young measures and minimization problems of mechanics

**J. M. Ball**
Department of Mathematics, Heriot–Watt University

and

**G. Knowles**
Department of Electrical Engineering, Imperial College

## 1. INTRODUCTION

In this article we discuss the use of Young measures in various minimization problems drawn from continuum mechanics. Our main application is to Gibbsian thermostatics of mixtures, but we also touch on recent work on elastic crystals.

The Young measure $(\nu_x)_{x\in\Omega}$ corresponding to a sequence of functions $z^{(j)}:\Omega\to\mathbb{R}^k$, where $\Omega\subset\mathbb{R}^n$, gives the limiting probability distribution as $j\to\infty$ of values of $z^{(j)}(y)$ as points $y$ are chosen uniformly at random from a small neighbourhood of $x$. (This intuitive description is made precise in Ball (1989).) If $f:\mathbb{R}^k\to\mathbb{R}$ is continuous then whenever the weak $L^1$ limit of $f(z^{(j)})$ exists it is given by the expectation $\langle \nu_x, f\rangle = \int_{\mathbb{R}^k} f(\lambda)\,\mathrm{d}\nu_x(\lambda)$.

In recent years the Young measure has become a popular tool for the study of nonlinear partial differential equations, following the influential work of Tartar (1979) and DiPerna (1983) on hyperbolic systems. However, the use we make of it here is the one for which it was originally introduced by L. C. Young (see Young, 1937, 1969; MacShane, 1940, 1978), namely as a device for generalizing the idea of a minimizer for problems of the calculus of variations and for describing the behaviour of minimizing sequences.

The plan of the article is as follows:

In Section 2 we outline necessary technical results concerning Young measures.

In Section 3 we prove general results of a classical type for the minimization of integrals of the form

$$I(u) = \int_\omega f(x, u(x))\,\mathrm{d}x \tag{1.1}$$

in the set $\mathscr{A}(a)$ of integrable maps $u:\omega\to\mathbb{R}^k$ satisfying the linear constraint

$$\int_\omega u(x)\,\mathrm{d}x = a, \tag{1.2}$$

where $a \in \mathbb{R}^k$ is given and where $f: \omega \times \mathbb{R}^k \to (-\infty, \infty]$ satisfies appropriate continuity and growth hypotheses. We do *not* assume that $f(x, \cdot)$ is convex. It is proved that the minimum of $I$ in $\mathscr{A}(a)$ is attained. A generalized problem for Young measures is formulated, each Young measure minimizer being the limit of a sequence of minimizers of $I$. Our choice of proofs makes frequent use of ideas of Berliocchi & Lasry (1973), and we also apply Aumann's theorem on set-valued integration.

In Section 4 we apply the results of Section 3 to the problem of minimization of the free energy of an $N$ component fluid mixture. We treat both miscible and immiscible mixtures, studying also the effect of gravity on the behaviour of minimizing sequences.

Finally, in Section 5 we briefly remark on recent research concerning Young measure minimizers in nonlinear elasticity. This case differs from that for fluid mixtures in that the minimum free energy is in general no longer attained. Minimizing sequences generate microstructure as they approach the lower bound for the energy; such microstructure is a typical feature of crystal morphology.

This article is an updated and supplemented version of unpublished work carried out several years ago and partially announced in Ball (1984).

## 2. YOUNG MEASURES

The following version of the fundamental theorem concerning the existence of Young measures and their relation to weak convergence is given in Ball (1989); the essence of the result is well known (see, for example, Berliocchi & Lasry (1973), Tartar (1979), Balder (1984) and the other references in Ball (1989)).

*Theorem 2.1* Let $\Omega \subset \mathbb{R}^n$ be Lebesgue measurable, let $K \subset \mathbb{R}^k$ be closed, and let $z^{(j)}: \Omega \to \mathbb{R}^k$, $j = 1, 2, \ldots$, be a sequence of Lebesgue measurable functions satisfying $z^{(j)}(\cdot) \to K$ in measure as $j \to \infty$, i.e. given any open neighbourhood $U$ of $K$ in $\mathbb{R}^k$

$$\lim_{j \to \infty} \operatorname{meas} \{x \in \Omega : z^{(j)}(x) \notin U\} = 0.$$

Then there exists a subsequence $z^{(\mu)}$ of $z^{(j)}$ and a family $(\nu_x)$, $x \in \Omega$, of positive measures on $\mathbb{R}^k$, depending measurably on $x$, such that

(i) $|\nu_x| \stackrel{\text{def}}{=} \int_{\mathbb{R}^k} \mathrm{d}\nu_x \leqslant 1$ for a.e. $x \in \Omega$,

(ii) $\operatorname{supp} \nu_x \subset K$ for a.e. $x \in \Omega$, and

(iii) $f(z^{(\mu)}) \stackrel{*}{\rightharpoonup} \langle \nu_x, f \rangle = \int_{\mathbb{R}_k} f(\lambda)\, \mathrm{d}\nu_x(\lambda)$

in $L^\infty(\Omega)$ for each continuous function $f: \mathbb{R}^k \to \mathbb{R}$ satisfying

$$\lim_{|\lambda| \to \infty} f(\lambda) = 0.$$

Suppose further that given any $R > 0$ there exists a continuous nondecreasing function

$g_R:[0,\infty)\to\mathbb{R}$, with $\lim_{t\to\infty} g_R(t)=\infty$, such that

$$\sup_{\mu}\int_{\Omega\cap B_R} g_R(|z^{(\mu)}(x)|)\,dx<\infty, \tag{2.1}$$

where $B_R=B(0,R)$. Then $\|v_x\|=1$ for a.e. $x\in\Omega$ (i.e. $v_x\in P(\mathbb{R}^k)=\{$probability measures on $\mathbb{R}^k\}$ a.e.), and given any measurable subset $A$ of $\Omega$

$$f(z^{(\mu)})\to\langle v_x,f\rangle \qquad \text{in } L^1(A) \tag{2.2}$$

for any continuous function $f:\mathbb{R}^k\to\mathbb{R}$ such that $\{f(z^{(\mu)})\}$ is sequentially weakly relatively compact in $L^1(A)$.

*Brief sketch of proof*

Denote by $C_0(\mathbb{R}^k)$ the Banach space of continuous functions $f:\mathbb{R}^k\to\mathbb{R}$ satisfying $\lim_{|\lambda|\to\infty} f(\lambda)=0$, with the norm $\|f\|_{C^0}=\sup_{\lambda\in\mathbb{R}^k}|f(\lambda)|$. By the Riesz representation theorem the dual space $C_0(\mathbb{R}^k)^*$ of $C_0(\mathbb{R}^k)$ is isometrically isomorphic to the Banach space $M(\mathbb{R}^k)$ of bounded Radon measures on $\mathbb{R}^k$. We identify $z^{(j)}$ with the mapping $v^{(j)}:\Omega\to M(\mathbb{R}^k)$ defined by

$$v^{(j)}(x)=\delta_{z^{(j)}(x)}. \tag{2.3}$$

Then $v^{(j)}$ is a bounded sequence in the Banach space $L^\infty_w(\Omega;M(\mathbb{R}^k))$ of essentially bounded weak* measurable mappings $\mu:\Omega\to M(\mathbb{R}^k)$. Since $C_0(\mathbb{R}^k)$ is separable, $L^\infty_w(\Omega;M(\mathbb{R}^k))$ is isometrically isomorphic to the dual space of $L^1(\Omega;C_0(\mathbb{R}^k))$, and hence there exists a subsequence $v^{(\mu)}$ of $v^{(j)}$ and an element $v=(v_x)$ of $L^\infty_w(\Omega;M(\mathbb{R}^k))$ such that $v^{(\mu)}\overset{*}{\rightharpoonup} v$ in $L^\infty_w(\Omega;M(\mathbb{R}^k))$. This is easily seen to imply (iii), and (i) follows by weak* lower semicontinuity of the norm. It is not hard to prove that $v_x\geqslant 0$, $\operatorname{supp} v_x\subset K$, $v$ inheriting these properties from the sequence $v^{(j)}$. The remaining assertions of the theorem may be proved via approximations of 1 and $f\in C(\mathbb{R}^k)$ by functions from $C_0(\mathbb{R}^k)$. For the details see Ball (1989). □

## 3. AN ABSTRACT MINIMIZATION PROBLEM

Let $\omega\subset\mathbb{R}^n$ be a bounded domain. We study the problem of minimizing

$$I(u)=\int_\omega f(x,u(x))\,dx \tag{3.1}$$

in the set $\mathscr{A}(a)$ of integrable maps $u:\omega\to\mathbb{R}^k$ satisfying the linear constraint

$$\int_\omega u(x)\,dx=a, \tag{3.2}$$

where $a\in\mathbb{R}^k$ is given. We suppose that $f:\omega\times\mathbb{R}^k\to(-\infty,\infty]$ is a *normal integrand*, that is

(a) $f(x,\cdot)$ is lower semicontinuous (l.s.c.) for a.e. $x\in\omega$, and
(b) there exists a Borel measurable function $\tilde f:\omega\times\mathbb{R}^k\to[0,\infty]$ such that $\tilde f(x,\cdot)=f(x,\cdot)$ for a.e. $x\in\omega$.

We also assume that $f$ satisfies the growth condition

$$f(x,v) \geqslant \phi(|v|) \quad \text{for a.e. } x\in\omega \text{ and all } v\in\mathbb{R}^k, \tag{3.3}$$

for some function $\phi\colon[0,\infty)\to(-\infty,\infty]$ which is bounded below and such that $\lim_{t\to\infty}\phi(t)/t=\infty$. Without loss of generality we suppose that $\phi$ is convex, nondecreasing and l.s.c. Under these hypotheses $I(u)$ is well defined (either finite or $+\infty$) for every measurable $u\colon\omega\to\mathbb{R}^k$. Finally we suppose that

$$\inf_{\mathscr{A}(b)} I < \infty, \tag{3.4}$$

for every $b$ in some neighbourhood of $a$.

We denote by $f^{**}$ the lower convex envelope of $f$ with respect to $v$. Then $f^{**}\colon\omega\times\mathbb{R}^k\to(-\infty,\infty]$ is also a normal integrand satisfying the growth condition (3.3) with the same $\phi$ (see Berliocchi & Lasry (1973, p. 159), Ekeland & Temam (1974, p. 229)). The first approach to the study of minimizing sequences for the problem (3.1), (3.2) is encapsulated in the following theorem, which is a special case of a result of Ekeland & Temam (1974, p. 266):

*Theorem 3.1* The minimum of

$$I^{**}(u) = \int_\omega f^{**}(x,u(x))\,\mathrm{d}x \tag{3.5}$$

in $\mathscr{A}(a)$ is attained, and

$$\inf_{\mathscr{A}(a)} I^{**} = \inf_{\mathscr{A}(a)} I. \tag{3.6}$$

Given any minimizing sequence $u^{(j)}$ of $I$ in $\mathscr{A}(a)$, there exists a subsequence $u^{(\mu)}$ converging weakly in $L^1(\omega;\mathbb{R}^k)$ to a minimizer $u^{**}$ of $I^{**}$ in $\mathscr{A}(a)$. Conversely, given any minimizer $u^{**}$ of $I^{**}$ in $\mathscr{A}(a)$ there exists a minimizing sequence $u^{(m)}$ of $I$ in $\mathscr{A}(a)$ converging weakly in $L^1(\omega;\mathbb{R}^k)$ to $u^{**}$.

*Remark 3.2* In fact more than this is true. The minimum of $I$ in $\mathscr{A}(a)$ is itself attained (see Theorem 3.3; for other existence assertions 'without convexity' see Olech (1970), Aubert & Tahraoui (1979), Marcellini (1980), Cesari (1983)), and the minimizing sequence $u^{(m)}$ can be taken to be a sequence of minimizers (see the proof). However these facts are due to special features of our variational problem, and the corresponding statements are not in general true for other problems of the calculus of variations.

Theorem 3.1 can be criticized on the grounds that it does not give optimal information about the values in $\mathbb{R}^k$ taken by minimizing sequences of $I$. In fact the theorem does not distinguish between different $f$ having the same lower convex envelope $f^{**}$. More precise information can most conveniently be expressed using the Young measure. This leads to a second approach to the study of (3.1), (3.2), in which a new minimization problem among Young measures is introduced.

Let $X = L^\infty_w(\omega; M(\mathbb{R}^k))$, $M=\{v\in X\colon v_x\in P(\mathbb{R}^k) \text{ a.e. } x\in\omega\}$. Given $v\in M$ we define

$$\hat{I}(v) = \int_\omega \langle v_x, f(x,\cdot)\rangle\,\mathrm{d}x,$$

$$= \int_\omega \int_{\mathbb{R}^k} f(x, v)\, \mathrm{d}\nu_x(v)\, \mathrm{d}x. \tag{3.7}$$

The map $\nu \mapsto \hat{I}(\nu)$ is well defined, affine, and weak* l.s.c. from $M$ to $(-\infty, \infty]$ (Berliocchi & Lasry, 1973, p. 143). Note that if $u: \omega \to \mathbb{R}^k$ is measurable then

$$\hat{I}(\delta_{u(\cdot)}) = I(u).$$

Let $\hat{\mathscr{A}}(a)$ denote the set of those $\nu \in M$ such that $\langle \nu_x, v \rangle = \int_\omega v\, \mathrm{d}\nu_x(v)$ is integrable and

$$\int_\omega \langle \nu_x, v \rangle \mathrm{d}x = a. \tag{3.8}$$

The new variational problem is to minimize $\hat{I}$ in $\hat{\mathscr{A}}(a)$.

Given $m \in \mathbb{R}^k$, define $f_m: \omega \times \mathbb{R}^k \to (-\infty, \infty]$ by

$$f_m(x, v) = f(x, v) - m \cdot v. \tag{3.9}$$

We denote by $S_m(x)$ the set of minimizers of $f_m(x, \cdot)$. Note that

$$S_m(x) = \partial f(x, \cdot)^{-1}(m), \tag{3.10}$$

where we recall that if $F: \mathbb{R}^k \to (-\infty, \infty]$ then the subdifferential $\partial F(v)$ of $F$ at the point $v \in \mathbb{R}^k$ is defined to be the set

$$\partial F(v) = \{m \in \mathbb{R}^k : F(w) \geqslant F(v) + m \cdot (w - v) \text{ for all } w \in \mathbb{R}^k\}.$$

*Theorem 3.3*

(i) The minimum of $\hat{I}$ in $\hat{\mathscr{A}}(a)$ is attained.

(ii) $\inf\limits_{\hat{\mathscr{A}}(a)} \hat{I} = \inf\limits_{\mathscr{A}(a)} I.$ $\quad (3.11)$

(iii) The minimizers of $\hat{I}$ in $\hat{\mathscr{A}}(a)$ are precisely those $\bar{\nu} \in \hat{\mathscr{A}}(a)$ such that

$$\operatorname{supp} \bar{\nu}_x \subset S_m(x) \qquad \text{a.e. } x \in \omega, \tag{3.12}$$

for some $m \in \mathbb{R}^k$.

(iv) The minimum of $I$ in $\mathscr{A}(a)$ is attained. The minimizers are precisely those $\bar{u} \in \mathscr{A}(a)$ such that

$$\bar{u}(x) \in S_m(x) \qquad \text{a.e. } x \in \omega. \tag{3.13}$$

(v) Given any minimizing sequence $u^{(j)}$ of $I$ in $\mathscr{A}(a)$, there exists a subsequence $u^{(\mu)}$ such that $\delta_{u^{(\mu)}(\cdot)}$ converges weak* in $X$ to a minimizer $\bar{\nu}$ of $\hat{I}$ in $\hat{\mathscr{A}}(a)$.

(vi) Given any minimizer $\bar{\nu}$ of $\hat{I}$ in $\hat{\mathscr{A}}(a)$, there exists a sequence $u^{(m)}$ of minimizers of $I$ in $\mathscr{A}(a)$ such that $\delta_{u^{(m)}(\cdot)}$ converges weak* in $X$ to $\bar{\nu}$.

*Proof*

(i) Let $\nu^{(j)}$ be a minimizing sequence for $\hat{I}$ in $\hat{\mathscr{A}}(a)$. Since $\nu^{(j)}$ is bounded in $X$, there exists a subsequence $\nu^{(\mu)}$ converging weak* to some $\bar{\nu} \in X$. It follows from general results of Berliocchi & Lasry (1973, pp. 150, 153) that $\bar{\nu} \in \hat{\mathscr{A}}(a)$; however, for completeness we give a proof. First, from the definition of weak* convergence it is easily shown that $\bar{\nu}_x \geqslant 0$ a.e. Next, since $\| \bar{\nu} \|_X \leqslant \underline{\lim}_{\mu \to \infty} \| \nu^{(\mu)} \|_X = 1$, it follows that $|\nu_x| \leqslant 1$ a.e. Using (3.3), (3.4) we have that for all sufficiently large $\mu$ and all $k \geqslant 1$,

$$k^{-1}\phi(k)\int_{\omega}\int_{\{|v|\geqslant k\}}|v|\,\mathrm{d}\nu_x^{(\mu)}(v)\,\mathrm{d}x \leqslant \int_{\omega}\langle \nu_x^{(\mu)}, \phi(|\cdot|)\rangle\,\mathrm{d}x \leqslant C < \infty, \tag{3.14}$$

for some constant $C$. Define $\vartheta^k \in C_0(\mathbb{R}^m)$ by

$$\vartheta^k(v) = \begin{cases} 1 & \text{for } |v| \leqslant k, \\ 1+k-|v| & \text{for } k \leqslant |v| \leqslant k+1, \\ 0 & \text{for } |v| \geqslant k+1. \end{cases} \tag{3.15}$$

Then from (3.14)

$$\begin{aligned} \left|\int_{\omega}(\langle \bar{\nu}_x, \vartheta^k(\cdot)\rangle - 1)\,\mathrm{d}x\right| &= \left|\lim_{\mu\to\infty}\int_{\omega}\langle \nu_x^{(\mu)}, \vartheta^k(\cdot) - 1\rangle\,\mathrm{d}x\right| \\ &\leqslant \lim_{\mu\to\infty}\int_{\omega}\nu_x^{(\mu)}([k,\infty))\,\mathrm{d}x \leqslant C\phi(k)^{-1}. \end{aligned} \tag{3.16}$$

Letting $k \to \infty$ in (3.16), it follows from the monotone convergence theorem that

$$\int_{\omega}(1 - |\bar{\nu}_x|)\,\mathrm{d}x = 0. \tag{3.17}$$

Hence $|\bar{\nu}_x| = 1$ a.e., so that $\bar{\nu}_x \in P(\mathbb{R}^k)$ a.e. Let $g^k(v) = \vartheta^k(v)v$. Then

$$\begin{aligned} \left|\int_{\omega}\langle \bar{\nu}_x, g^k(\cdot)\rangle\,\mathrm{d}x - a\right| &= \left|\lim_{\mu\to\infty}\int_{\omega}\langle \nu_x^{(\mu)}, g^k(\cdot)\rangle\,\mathrm{d}x - a\right| \\ &\leqslant \left|\lim_{\mu\to\infty}\int_{\omega}\langle \nu_x^{(\mu)}, g^k(v) - v\rangle\,\mathrm{d}x\right| \\ &\leqslant Ck\phi(k)^{-1}, \end{aligned}$$

where we have used (3.14). Letting $k \to \infty$ we deduce from the monotone convergence theorem that $\int_{\omega}\langle \bar{\nu}_x, v\rangle\,\mathrm{d}x = a$, and thus $\bar{\nu} \in \hat{\mathscr{A}}(a)$ as claimed.

Since $\hat{I}(\cdot)$ is weak* l.s.c. it follows that $\hat{I}(\bar{\nu}) = \inf_{\hat{\mathscr{A}}(a)}\hat{I}$, so that $\bar{\nu}$ is a minimizer.

(iii) Suppose $\bar{\nu} \in \hat{\mathscr{A}}(a)$ satisfies (3.12) for some $m \in \mathbb{R}^k$. Then for a.e. $x \in \omega$,

$$f(x,w) \geqslant f(x,v) + m\cdot(w - v) \quad \text{for all } v \in \operatorname{supp}\bar{\nu}_x, \quad w \in \mathbb{R}^k. \tag{3.18}$$

Let $\nu \in \hat{\mathscr{A}}(a)$. Integrating (3.18) with respect to $\bar{\nu}_x$ and $\nu_x$ we obtain

$$\langle \nu_x, f(x,\cdot)\rangle \geqslant \langle \bar{\nu}_x, f(x,\cdot)\rangle + m\cdot(\langle \nu_x, v\rangle - \langle \bar{\nu}_x, v\rangle) \quad \text{a.e. } x \in \omega. \tag{3.19}$$

Integrating over $\omega$ we deduce that $\hat{I}(\nu) \geqslant \hat{I}(\bar{\nu})$, proving that $\bar{\nu}$ is a minimizer.

Conversely, let $\bar{\nu}$ minimize $\hat{I}$ in $\hat{\mathscr{A}}(a)$. Consider the function $G(b) \stackrel{\text{def}}{=} \inf_{\hat{\mathscr{A}}(b)}\hat{I}$. $G$ is obviously convex, and, by a minor adaptation of the proof of (i), $G$ is l.s.c. From (3.4), $a \in \operatorname{int}\{G < \infty\}$. Hence $\partial G(a)$ is nonempty (cf. Ekeland & Temam, 1974, pp. 12, 22). Let $m \in \partial G(a)$. Then

$$\int_{\omega}\langle \nu_x - \bar{\nu}_x, f(x,v) - m\cdot v\rangle\,\mathrm{d}x \geqslant 0 \quad \text{for all } \nu \in M. \tag{3.20}$$

Let $z: \omega \to \mathbb{R}^k$ be measurable and such that

$$f_m(x, z(x)) = \min_{v \in \mathbb{R}^k} f_m(x, v).$$

Such a $z$ exists (cf. Ekeland & Temam, 1974, p. 220). Let $v_x = \delta_{z(x)}$. By (3.20),

$$\int_\omega \left[ \min_{v \in \mathbb{R}^k} f_m(x, v) - \langle \bar{v}_x, f_m(x, \cdot) \rangle \right] \mathrm{d}x \geqslant 0. \tag{3.21}$$

But the integrand in (3.21) is nonpositive. Hence, for a.e. $x \in \omega$, $\operatorname{supp} \bar{v}_x$ is contained in the set of minimum points of $f_m(x, \cdot)$, which gives (3.12).

(ii), (iv) Let $\bar{v}$ be a minimized of $\hat{I}$ in $\hat{\mathscr{A}}(a)$, and let $m$ be such that (3.12) holds. Consider the set-valued function $S_m(\cdot)$. By (a) and (3.12), for a.e. $x \in \omega$, $S_m(x)$ is nonempty and closed. Also, on account of (3.4) there exists $v \in L^1(\omega; \mathbb{R}^k)$ such that $a(\cdot) \overset{\text{def}}{=} f_m(\cdot, v(\cdot)) \in L^1(\omega)$. Hence

$$\gamma(x) \overset{\text{def}}{=} \min_{z \in \mathbb{R}^k} f_m(x, z) \leqslant a(x) \quad \text{a.e. } x \in \omega.$$

By (3.12) there exist $d > 0$ and $c$ such that $f_m(x, z) \geqslant c + d|z|$ for a.e. $x \in \omega$ and all $z$. Hence

$$|z| \leqslant d^{-1}[a(x) - c] \qquad \text{for all } z \in S_m(x), \quad \text{a.e. } x \in \omega,$$

and thus $S_m(\cdot)$ is integrably bounded. In particular, $S_m(x)$ is bounded for a.e. $x \in \omega$ and hence its convex hull $\operatorname{co} S_m(x)$ is closed (Rudin, 1973, p. 72). Further, since $\gamma(\cdot)$ is measurable (Ekeland & Temam, 1974, p. 220), $f_m(\cdot, \cdot) - \gamma(\cdot)$ is a normal integrand, and hence without loss of generality may be assumed Borel measurable. So for any open set $E \subset \mathbb{R}^k$, the set

$$\{(x, v) \in \omega \times E : f_m(x, v) = \gamma(x)\} = \{(x, v) : v \in S_m(x) \cap E\}$$

is Borel measurable. Its projection onto $\omega$ is $\{x \in \omega : S_m(x) \cap E \text{ is nonempty}\}$, which is therefore Borel measurable. Hence $S_m(\cdot)$ is measurable. Since $S_m(\cdot)$ is measurable, closed, nonempty and integrably bounded, by Aumann's theorem (Aumann, 1965; Berliocchi & Lasry, 1973, p. 164; Clarke, 1983)

$$\int_\omega S_m(x)\,\mathrm{d}x = \int_\omega \operatorname{co} S_m(x)\,\mathrm{d}x. \tag{3.22}$$

(The integral of a measurable, nonempty and closed set-valued function $F$ from $\omega$ to subsets of $\mathbb{R}^k$ is the nonempty set

$$\int_\omega F(x)\,\mathrm{d}x = \left\{ \int_\omega z(x)\,\mathrm{d}x : z \in L^1(\omega; \mathbb{R}^k) \text{ such that } z(x) \in F(x) \text{ a.e. } x \in \omega \right\}.)$$

Now since by (iii) $\operatorname{supp} \bar{v}_x \subset S_m(x)$ a.e., and since $\operatorname{co} S_m(x)$ is closed a.e., it follows that $\langle \bar{v}_x, v \rangle \in \operatorname{co} S_m(x)$ a.e. Therefore $a \in \int_\omega \operatorname{co} S_m(x)\,\mathrm{d}x$, and thus by (3.22) there exists a $\bar{u} \in \mathscr{A}(a)$ such that $\bar{u}(x) \in S_m(x)$ a.e. By (iii) $\delta_{\bar{u}(\cdot)}$ is a minimizer of $\hat{I}$ in $\hat{\mathscr{A}}(a)$, and since $\inf_{\hat{\mathscr{A}}(a)} \hat{I} \leqslant \inf_{\mathscr{A}(a)} I$ it follows that $\bar{u}$ minimizes $I$ in $\mathscr{A}(a)$ and that (3.11) holds. Because of (3.11), $\delta_{\bar{u}(\cdot)}$ minimizes $\hat{I}$ in $\hat{\mathscr{A}}(a)$ for any minimizer $\bar{u}$ of $I$ in $\mathscr{A}(a)$, so that (3.13) follows from (iii).

(v) This follows immediately from (ii) and the proof of (i).

(vi) Let $\bar{\nu}$ be a minimizer of $\hat{I}$ in $\hat{\mathscr{A}}(a)$, and let $m$ be such that (3.12) holds. Let

$$\Delta = \left\{ \delta_{\bar{u}(\cdot)} : \bar{u} \in L^1(\omega; \mathbb{R}^k), \bar{u}(x) \in S_m(x) \text{ a.e.}, \int_\omega \bar{u}\, dx = a \right\},$$

$$N = \left\{ \bar{\nu} \in M : \operatorname{supp} \bar{\nu}_x \subset S_m(x) \text{ a.e.}, \int_\omega \langle \bar{\nu}_x, v \rangle\, dx = a \right\}.$$

We prove that $N$ is the weak* closure of $\Delta$ in $X$. The result then follows from (iii), (iv) and the metrizability of the weak* topology of $X$ when restricted to $M$. The proof of (i) shows that $N$ is weak* closed, so we just need to prove that $N \subset$ weak* closure $(\Delta)$. Suppose not. Then there exist $\bar{\nu} \in N$ and a weak* open neighbourhood $U$ of $\bar{\nu}$ disjoint from $\Delta$. We may assume that $U$ consists of those $\nu \in X$ such that

$$\left| \int_\omega \langle \bar{\nu}_x - \nu_x, \psi_i(x, v) \rangle\, dx \right| < \varepsilon_i, \qquad i = 1, \ldots, L, \tag{3.23}$$

where $\psi_i \in L^1(\omega; C_0(\mathbb{R}^k))$, $\varepsilon_i > 0$. Since each $\psi_i$ is a normal integrand we may assume without loss of generality that $\psi = (\psi_1, \ldots, \psi_L)$ is Borel measurable. Define

$$\Gamma(x) = \{(\psi(x, v), v) : v \in S_m(x)\} \subset \mathbb{R}^{k+L}.$$

Then $\Gamma(\cdot)$ is closed, nonempty, and integrably bounded (since $S_m(\cdot)$ is). To show that $\Gamma(\cdot)$ is measurable it suffices to show that the set

$$T \overset{\text{def}}{=} \{x \in \omega : (\psi(x, v), v) \in E_1 \times E_2 \text{ for some } v \in S_m(x)\}$$

is measurable for any given open subsets $E_1 \subset \mathbb{R}^L$, $E_2 \subset \mathbb{R}^k$. But $T$ is the projection onto $\omega$ of the set

$$\{(x, v) : v \in S_m(x) \cap E_2\} \cap \{(x, v) : \psi(x, v) \in E_1\}. \tag{3.24}$$

The first set in (3.24) is Borel measurable as shown in the proof of (ii), (iv) above, while the second is Borel measurable since $\psi$ is. Thus $T$ is measurable.

We apply again Aumann's theorem to $\Gamma(\cdot)$. Since $(\langle \bar{\nu}_x, \psi(x, v) \rangle, \langle \bar{\nu}_x, v \rangle) \in \operatorname{co} \Gamma(x)$ a.e., it follows that there exists $\bar{u} \in L^1(\omega; \mathbb{R}^k)$ with $\bar{u}(x) \in S_m(x)$ a.e. such that

$$\int_\omega \psi_i(x, \bar{u}(x))\, dx = \int_\omega \langle \bar{\nu}_x, \psi_i(x, v) \rangle\, dx, \qquad i = 1, \ldots, L,$$

$$\int_\omega \bar{u}(x)\, dx = \int_\omega \langle \bar{\nu}_x, v \rangle\, dx = a.$$

But then $\nu = \delta_{\bar{u}(\cdot)} \in U \cap \Delta$, a contradiction. □

*Proof of Theorem 3.1* Applying Theorem 3.3 to $f^{**}$, we see that the minimum of $I^{**}$ on $\hat{\mathscr{A}}(a)$ is attained (this could of course also be proved directly via standard lower semicontinuity results). Let $u^{**}$ be a minimizer. By Berliocchi & Lasry (1973, Lemme 3, p. 161), there exists $\bar{\nu} \in M$ such that

$$f^{**}(x, u^{**}(x)) = \int_\omega \langle \bar{\nu}_x, f(x, \cdot) \rangle\, dx, \quad \langle \bar{\nu}_x, v \rangle = u^{**}(x), \quad \text{a.e. } x \in \omega. \tag{3.25}$$

Hence $\bar{v}\in\hat{\mathscr{A}}(a)$ and $\hat{I}(\bar{v})=\inf_{\mathscr{A}(a)} I^{**}\leqslant\inf_{\mathscr{A}(a)} I$. By (ii) $\bar{v}$ is a minimizer of $\hat{I}$ in $\hat{\mathscr{A}}(a)$ and (3.6) holds. Furthermore, by (vi) there exists a sequence $u^{(m)}$ of minimizers of $I$ in $\mathscr{A}(a)$ such that $\delta_{u^{(m)}(\cdot)}$ converges weak* in $X$ to $\bar{v}$. Since by (3.3) and the de la Vallée Poussin criterion $u^{(m)}$ is sequentially weakly relatively compact in $L^1(\omega;\mathbb{R}^k)$, this implies in particular by Theorem 2.1 that $u^{(m)}$ converges weakly in $L^1(\omega;\mathbb{R}^k)$ to $\langle\bar{v}_x, v\rangle = u^{**}$. Similarly, given any minimizing sequence $u^{(j)}$ of $I$ in $\mathscr{A}(a)$, by (v) there exists a subsequence $u^{(\mu)}$ such that $\delta_{u^{(\mu)}(\cdot)}$ converges weak* in $X$ to a minimizer $\bar{v}$ of $\hat{I}$ in $\hat{\mathscr{A}}(a)$; thus $u^{(\mu)}$ converges weakly in $L^1(\omega;\mathbb{R}^k)$ to $\langle\bar{v}_x, v\rangle \overset{\text{def}}{=} u^{**}$, and by convexity (3.25) holds, showing that $u^{**}$ is a minimizer of $I^{**}$ in $A(a)$. □

## 4. EQUILIBRIUM OF FLUIDS

In this section we apply Theorems 3.1, 3.3 to the problem of the equilibrium of single and multicomponent (heterogeneous) fluids, in the spirit of Gibbs (1873).

### A. Minimization of the free energy for miscible mixtures

We consider first the case of an $N$ component homogeneous miscible fluid mixture filling a bounded open container $\omega\subset\mathbb{R}^3$ whose boundary is held at a constant temperature $\vartheta$. By *miscible* we mean that each of the $N$ fluid components can be present at every point $x=(x_1,x_2,x_3)\in\omega$. Physically, one can think of the mixing in a miscible mixture as taking place at a molecular length scale. Continuum theories of immiscible mixtures typically also have the property that each component can be present at every point, but the viewpoint we will adopt in subsection B for the modelling of immiscibility will be different.

We denote by $\rho_i(x), i=1,\ldots,N$, the *specific mass densities* of the fluid components at the point $x\in\omega$. The *state* of the fluid is the vector field $\rho:\omega\to\mathbb{R}^N$ given by

$$\rho(x)=(\rho_1,\ldots,\rho_N)(x). \tag{4.1}$$

The *mass density* of the mixture is given by

$$M(\rho)(x)=\sum_{i=1}^{N}\rho_i(x). \tag{4.2}$$

We suppose that at the temperature $\vartheta$ the *specific Helmholtz free energy* $\psi(x)$ of the fluid is given by the constitutive relation

$$\psi=\psi_\vartheta(\rho_1,\ldots,\rho_N)=\psi_\vartheta(\rho), \tag{4.3}$$

where $\psi_\vartheta:\mathbb{R}^N\to(-\infty,\infty]$ is bounded below, lower semicontinuous, and satisfies the following conditions:

(H1) $D\overset{\text{def}}{=}\{\rho\in\mathbb{R}^N:\psi_\vartheta(\rho)<\infty\}$ consists of an open subset $U$ of the positive orthant $\mathbb{R}^N_+=\{\rho\in\mathbb{R}^N:\rho_i\geqslant 0\}$ together with part of the boundary $\partial U$ of $U$, and $\psi_\vartheta$ is locally bounded above in $U$,

(H2) $\partial\psi_\vartheta(\rho)$ is empty for all $\rho\in D\cap\partial U$,

(H3) $\lim_{|\rho|\to\infty}|\rho|^{-1}\psi_\vartheta(\rho)=\infty.$

*Example 4.1* For a van der Waals fluid the free energy has the form (with $N = 1$)

$$\psi_\vartheta(\rho) = \begin{cases} -s\rho^2 + k\rho\vartheta\log\left(\dfrac{\rho}{b-\rho}\right) - c\rho\vartheta\log\vartheta - d\rho\vartheta + e\rho & \text{if } \rho\in[0, b), \\ +\infty \quad \text{otherwise,} \end{cases} \tag{4.4}$$

where $s > 0, b > 0, c > 0, k > 0, d$ and $e$ are constants (cf. Landau & Lifschitz, 1970). In this case $D = [0, b), U = (0, b)$ and $\psi_\vartheta(0) = 0, \psi_\vartheta(b) = \infty$. Clearly $\psi_\vartheta(\cdot)$ is l.s.c. and satisfies (H1), (H3), while (H2) holds since $\lim_{\rho\to 0+} \partial\psi_\vartheta(\rho)/\partial\rho = -\infty$.

*Example 4.2* For a mixture of $N$ noninteracting and infinitely compressible fluids we can take

$$\psi_\vartheta(\rho) = \sum_{i=1}^{N} \psi^i(\rho_i), \tag{4.5}$$

where, for each $i, \psi^i(t) = \infty$ if $t < 0, \psi^i$ is finite and continuous on $[0, \infty), \psi^i(0) = 0$, $\lim_{t\to 0+} \psi^i(t)/t = -\infty, \lim_{t\to\infty} \psi^i(t)/t = \infty$. In this case $D = \mathbb{R}^N_+, U = (0, \infty)^N$, and it is easily verified that $\psi_\vartheta$ is l.s.c. and satisfies (H1)–(H3).

We suppose the body force to be gravitational with potential energy density $gM(\rho)(x)x_3$, where $g \geqslant 0$ is a constant. The volumetric heat supply is assumed zero. A well known argument (see, for example, Duhem, 1911; Ericksen, 1966; Coleman & Dill, 1973; Ball, 1984; Ball & Knowles, 1986) based on the existence of a Lyapunov function, the *availability* or *ballistic free energy*, for the dynamical equations, then suggests that states of the fluid approaching equilibrium will be associated with minimizing sequences for the free energy functional

$$I(\rho) = \int_\omega [\psi_\vartheta(\rho(x)) + gM(\rho)(x)x_3]\, dx, \tag{4.6}$$

subject to the mass constraint

$$\int_\omega \rho(x)\, dx = a, \tag{4.7}$$

where $a = (a_1, \ldots, a_N)\in U$.

We are now in a position to apply Theorems 3.1, 3.3 to the minimization problem (4.6), (4.7) with $f(x, \rho) = \psi_\vartheta(\rho) + gM(\rho)x_3$. It is easily seen that $f$ is a normal integrand. The growth condition (3.3) follows from (H3) since $\omega$ is bounded, while (3.4) holds taking $\rho(x) \equiv b\in U$. Hence both theorems apply. In particular, by part (iii) of Theorem 3.3 $\bar{\nu}\in\hat{\mathscr{A}}(a)$ is a minimizing Young measure of $\hat{I}$ in $\hat{\mathscr{A}}(a)$ if and only if there exists an $m\in\mathbb{R}^N$ such that $\operatorname{supp}\bar{\nu}_x \subset S_m(x)$ a.e. Suppose for simplicity that $\psi_\vartheta\in C^1(U)$. Then by (H2)

$$\psi_\vartheta(\rho) + gM(\rho)x_3 - m\cdot\rho = \gamma(x_3), \tag{4.8}$$

$$D_\rho\psi_\vartheta(\rho) = m - gx_3\tau, \tag{4.9}$$

for all $\rho\in S_m(x)$, where $\gamma(x_3) = \min_\rho[\psi_\vartheta(\rho) + gM(\rho)x_3 - m\cdot\rho]$ and $\tau = (1, \ldots, 1)$. Define the *pressure* $p = p(\rho)$ and the *vector of chemical potentials* $\mu = \mu(\rho)$ by

$$p = \rho \cdot D_\rho \psi_\vartheta(\rho) - \psi_\vartheta(\rho), \qquad \mu = D_\rho \psi_\vartheta(\rho). \tag{4.10}$$

Then it follows from (4.8), (4.9) that

$$p(\rho) = -\gamma(x_3), \quad \mu(\rho) = m - g\tau x_3 \qquad \text{for all } \rho \in S_m(x). \tag{4.11}$$

In particular,

$$\langle \bar{v}_x, p(\cdot) \rangle = -\gamma(x_3), \quad \langle \bar{v}_x, \mu(\cdot) \rangle = m - g\tau x_3 \qquad \text{a.e. } x \in \omega, \tag{4.12}$$

for any minimizer $\bar{v}$.

We split the discussion into two cases:

*Case 1* ($g = 0$, i.e. zero gravity)

Here the set $S_m = S_m(x)$ is independent of $x$ and may contain more than one point, and in general $\bar{v}_x$ is not a Dirac mass. For example, for a van der Waals fluid with free energy given by (4.4) with $sb > (\frac{3}{2})^3 k\vartheta$ the graph of $\psi_\vartheta$ in $[0, b)$ is not convex and has a common tangent with end-points $\rho^{(1)}, \rho^{(2)}$ satisfying $0 < \rho^{(1)} < \frac{1}{3}b < \rho^{(2)}$ and slope $m^*$ (see Fig. 1). If $(\text{meas}\,\omega)^{-1} a \in (\rho^{(1)}, \rho^{(2)})$ then $a = (\text{meas}\,\omega)(\lambda\rho^{(1)} + (1-\lambda)\rho^{(2)})$ for some $\lambda \in (0, 1)$, and thus

$$\bar{v}_x = \lambda\delta_{\rho^{(1)}} + (1-\lambda)\delta_{\rho^{(2)}} \tag{4.13}$$

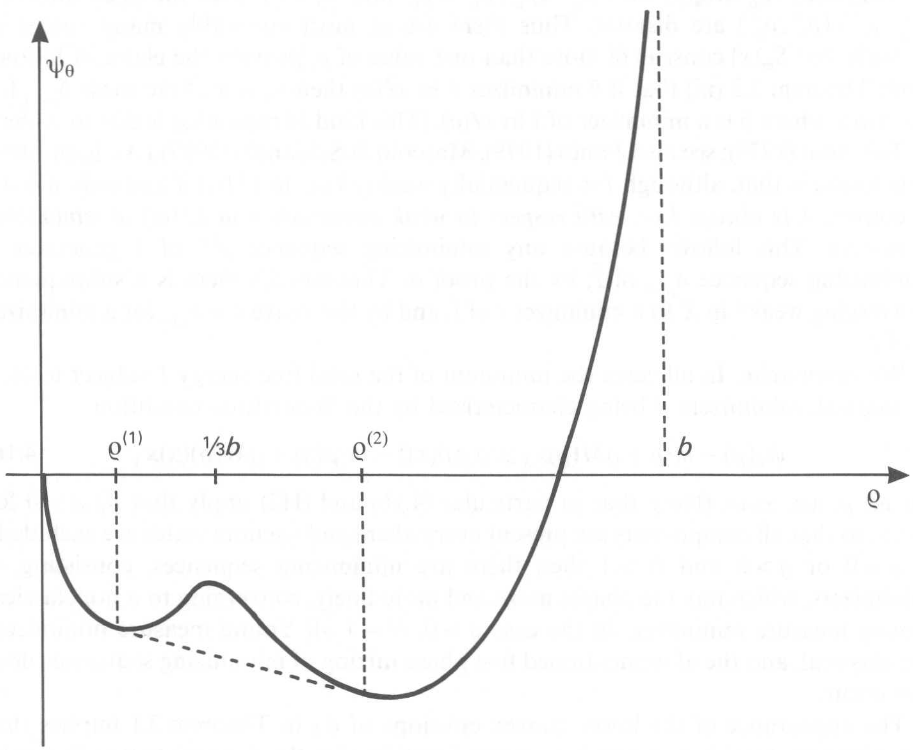


Fig. 1. The graph of $\psi_\vartheta$ for a van der Waals fluid.

satisfies $\operatorname{supp}\bar{\nu}_x \subset S_{m^*}$ and $\int_\omega \langle \bar{\nu}_x, \rho \rangle \, dx = a$. Hence by Theorem 3.3(iii) $\bar{\nu}$ is a minimizer.

*Case 2* ($g > 0$, i.e. positive gravity)
We suppose first that there is more than one component to the mixture (i.e. $N > 1$). This case is similar to Case 1 in that $S_m(x)$ may contain more than one point a.e., and $\bar{\nu}_x$ need not be a Dirac mass. To construct an example, suppose that $N = 2$ and that $f$ is such that

$$f(\rho_1, \rho_2) = f^{**}(\rho_1, \rho_2) = (\rho_1 + \rho_2 - 2)^2 \quad \text{if } (\rho_1 - 1)^2 + (\rho_2 - 1)^2 \leqslant \tfrac{1}{4}.$$

Let $\omega \subset \{x \in \mathbb{R}^3 : |x_3| < 2g^{-1}\}$. Then the line segment

$$K_x = \{(\rho_1, \rho_2) : \rho_1 + \rho_2 = 2 - \tfrac{1}{2} g x_3, (\rho_1 - 1)^2 + (\rho_2 - 1)^2 \leqslant \tfrac{1}{4}\},$$

is contained in $S_m(x)$ for $m = 0$, and so any $\bar{\nu}$ such that $\operatorname{supp}\bar{\nu}_x \subset K_x$ a.e. is a minimizer of $\hat{I}$ in $\hat{\mathscr{A}}(a)$ for some $a$.

Next we suppose that $N = 1$. We claim that, for any $m \in \mathbb{R}$, $S_m(x)$ is a singleton for a.e. $x \in \omega$. In fact, let $\rho_a \in S_m(x)$, $\rho_b \in S_m(y)$ with $x_3 < y_3$. Then

$$\psi_\vartheta(\rho_b) - m\rho_b + g\rho_b x_3 \geqslant \psi_\vartheta(\rho_a) - m\rho_a + g\rho_a x_3, \tag{4.14}$$

$$\psi_\vartheta(\rho_a) - m\rho_a + g\rho_a y_3 \geqslant \psi_\vartheta(\rho_b) - m\rho_b + g\rho_b y_3, \tag{4.15}$$

so that adding (4.14), (4.15) gives $(\rho_b - \rho_a)(x_3 - y_3) \geqslant 0$, whence $\rho_b \leqslant \rho_a$. Thus if $\rho_a^-$, $\rho_a^+ \in S_m(x)$, $\rho_b^-, \rho_b^+ \in S_m(y)$ with $\rho_a^- < \rho_a^+$, $\rho_b^- < \rho_b^+$ and $x_3 < y_3$, then the open intervals $(\rho_a^-, \rho_a^+)$, $(\rho_b^-, \rho_b^+)$ are disjoint. Thus there are at most countably many values of $x_3$ such that $S_m(x)$ consists of more than one value of $\rho$, proving the claim. It follows from Theorem 3.3 (iii) that if $\bar{\nu}$ minimizes $\hat{I}$ in $\hat{\mathscr{A}}(a)$ then $\bar{\nu}_x$ is a Dirac mass $\delta_{\bar{\rho}(x)}$ for a.e. $x \in \omega$, where $\bar{\rho}$ is a minimizer of $I$ in $\mathscr{A}(a)$. (This kind of reasoning is due to Aubert & Tahraoui (1979); see also James (1979), Mascolo & Schianchi (1987).) An interesting conclusion is that, although $I$ is sequentially weakly l.s.c. in $L^1(\omega)$ if and only if $\psi_\vartheta(\cdot)$ is convex, *$I$ is always l.s.c. with respect to weak convergence in $L^1(\omega)$ of minimizing sequences.* This follows because any minimizing sequence $\rho^{(j)}$ of $I$ generates a minimizing sequence $\delta_{\rho^{(j)}}$ of $\hat{I}$; by the proof of Theorem 3.3 there is a subsequence converging weak* in $X$ to a minimizer $\bar{\nu}$ of $\hat{I}$, and by the above $\bar{\nu} = \delta_{\bar{\rho}(\cdot)}$ for a minimizer $\bar{\rho}$ of $I$.

We summarize. In all cases the minimum of the total free energy $I$ subject to (4.7) is attained, minimizers $\bar{\rho}$ being characterized by the Weierstrass condition

$$\psi_\vartheta(\rho) - m\cdot\rho + gM(\rho)x_3 \geqslant \psi_\vartheta(\bar{\rho}(x)) - m\cdot\bar{\rho}(x) + gM(\bar{\rho})(x)x_3 \tag{4.16}$$

for all $\rho$, a.e. $x \in \omega$. (Note that in particular (4.16) and (H2) imply that $\bar{\rho}_i(x) > 0$ for a.e. $x$, so that all components are present everywhere and vacuum states are excluded.) If $g = 0$ or $g > 0$ and $N > 1$ then there are minimizing sequences, consisting of minimizers, which mix the phases more and more finely, converging to a nonclassical Young measure minimizer. In the case $g > 0$, $N = 1$ all Young measure minimizers are classical, and the aforementioned fine phase mixing of minimizing sequences does not occur.

The appearance of the lower convex envelope of $\psi_\vartheta$ in Theorem 3.1 implies that the minimum total free energy is a convex function of $a$; this is consistent with results of statistical physics for infinite volumes. For discussion and references see Thompson

(1972). Penrose & Lebowitz (1971), on the other hand, give an interesting interpretation within the context of equilibrium statistical physics of metastable portions of the graph of $\mathscr{A}_\vartheta$ for a single component fluid, i.e. parts where $\psi_\vartheta$ is convex but not equal to $\psi_\vartheta^{**}$. The Weierstrass 'stability' condition (4.16) shows that only values of $\rho$ where $\psi_\vartheta = \psi_\vartheta^{**}$ can occur in minimizers (and, in the sense of Young measures, in minimizing sequences).

If hypothesis (H2) is dropped, then we may still apply Theorems 3.1, 3.3, but minimizers may take values on $\partial U$. In particular, vacuum states may occur. For example, in the case $N = 1$, the problem of minimizing

$$I(\rho) = \int_\omega [k\rho^\gamma + g\rho x_3]\,dx, \tag{4.17}$$

subject to the constraints $\rho \geqslant 0$ a.e. and

$$\int_\omega \rho\,dx = a, \tag{4.18}$$

where $k > 0$, $\gamma > 1$ are constants, corresponds to that of equilibrium of a fixed mass of an adiabatically deforming polytropic gas under gravity. The minimizer is unique, and in the case $\omega = E \times (0, H)$, where $E \subset \mathbb{R}^2$ is bounded and open with area $\Delta$, is given for $H > h$ by

$$\bar{\rho}(x) = \begin{cases} [(k\gamma)^{-1} g(h - x_3)]^{1/(\gamma-1)} & \text{if } x_3 \leqslant h \\ 0 & \text{if } x_3 > h, \end{cases} \tag{4.19}$$

where

$$h = \left(\frac{a\gamma}{\Delta(\gamma-1)}\right)^{(\gamma-1)/\gamma} \left(\frac{k\gamma}{g}\right)^{1/\gamma}. \tag{4.20}$$

This can be verified by showing that $\bar{\rho}(x) \in S_{gh}(x)$ a.e. For the same problem treated in material coordinates see Ball (1988).

We could also contemplate applying Theorems 3.1, 3.3 to the Gibbs problem of maximizing the total entropy

$$I(\varepsilon, \rho) = \int_\omega \hat{\eta}(\varepsilon(x), \rho(x))\,dx, \tag{4.21}$$

of a miscible mixture subject to given energy

$$\int_\omega \varepsilon(x)\,dx = \alpha, \tag{4.22}$$

and given mass densities

$$\int_\omega \rho(x)\,dx = a. \tag{4.23}$$

Here $\varepsilon = \varepsilon(x)$ denotes the *specific internal energy* of the mixture and the *specific entropy* $\eta(x)$ is assumed to be given by the constitutive relation $\eta = \hat{\eta}(\varepsilon, \rho)$. This problem corresponds to the case of a thermally insulated boundary $\partial\omega$. However, for the existence proofs of Theorems 3.1, 3.3 to apply directly, it would be necessary to assume the growth condition

$$-\hat{\eta}(\varepsilon, \rho) \geqslant \phi(|\varepsilon| + |\rho|) \qquad \text{for all } \varepsilon \in \mathbb{R}, \quad \rho \in \mathbb{R}^N, \tag{4.24}$$

for some function $\phi:[0, \infty) \to (-\infty, \infty]$ which is bounded below and such that $\lim_{t\to\infty} \phi(t)/t = \infty$. This growth condition is unrealistic since we want the temperature $\vartheta = \partial\hat{\eta}/\partial\varepsilon$ to be positive. Analogous results could probably be proved by exploiting the positivity of $\vartheta$ (cf. Lin, 1989), but we do not explore this further here. The characterizations (3.12), (3.13) of maximizers do not require (4.24).

For (3.13), which corresponds to classical results of Gibbsian thermostatics, see, for example, Dunn & Fosdick (1980), Fosdick & Patiño (1986). Particular mention should be made of the work of Noll (1970), which roughly speaking concerns Young measure maximizers that are independent of $x$; this work also investigates the Gibbs phase rule (see also Man, 1985b). For later work in the same spirit we refer to Patiño (1987).

A further extension would be to include as a new variable the velocity, together with appropriate constraints corresponding to additional constants of motion (see Man, 1985a; Ball, 1984; Lin, 1989).

## B. Minimization of the free energy for immiscible mixtures

An *immiscible* mixture is one for which the components remain separate on a length scale that is sufficiently large with respect to molecular dimensions. An extensive review of theories of immiscible mixtures is given by Bedford & Drumheller (1983). Our purpose here is to illustrate how at the level of thermostatics Theorems 3.1, 3.3 can make a link between a theory in which at each point $x$ there is only one component present, and one in which all components may be present.

We model immiscibility by imposing the constraint that at most one of the densities $\rho_i(x)$ can be nonzero for a.e. $x \in \omega$. Equivalently, the function

$$H(\rho) = \begin{cases} 0 & \text{if, for some } j,\ \rho_i = 0 \text{ for } i \neq j,\ \rho_j \geqslant 0 \\ +\infty & \text{otherwise,} \end{cases} \tag{4.25}$$

which is the indicator function of the nonnegative coordinate axes, is zero for a.e. $x \in \omega$.

We suppose for simplicity that the fluid components are infinitely compressible, and that at the temperature $\vartheta$ the $i$th component of the mixture has the free energy function $\psi^i = \psi^i(\rho_i)$, where $\psi^i(t) = \infty$ if $t < 0$, $\psi^i$ is finite and continuous on $[0, \infty)$, $\psi^i(0) = 0$, and $\lim_{t\to\infty} \psi^i(t)/t = \infty$. We consider the problem of minimizing

$$J(\rho) = \int_\omega \left( \sum_{i=1}^N \psi^i(\rho_i(x)) + H(\rho(x)) + gM(\rho(x))x_3 \right) \mathrm{d}x, \tag{4.26}$$

subject to

$$\int_\omega \rho(x)\,\mathrm{d}x = a, \tag{4.27}$$

where $a_i > 0$ for all $i = 1, \ldots, N$.

We apply Theorems 3.1, 3.3 to the minimization problem (4.26), (4.27) with $f(x, \rho) = \sum_{i=1}^N \psi^i(\rho_i) + H(\rho) + gM(\rho)x_3$. Since $H$ is l.s.c., $f$ is a normal integrand. The growth condition (3.3) and hypothesis (3.4) are also easily verified. So both theorems apply. Since $H$ is not convex, the computation of $f^{**}(x, \rho)$ is not immediate even when the $\psi^i$ are themselves convex. What is clear, however, is that $f^{**}(x, \rho) < \infty$ for all $\rho \in \mathbb{R}^N_+$, so that the relaxed problem in Theorem 3.1 takes the form of that for a miscible mixture.

Define $F: \mathbb{R}^N \to (-\infty, \infty]$ by

$$F(\rho) \stackrel{\text{def}}{=} \sum_{i=1}^{N} \psi^i(\rho_i) + H(\rho). \tag{4.28}$$

*Theorem 4.3*

$$F^{**}(\rho) = \min\left\{ \sum_{i=1}^{N} \lambda_i (\psi^i)^{**}(m_i) : \lambda_i \geqslant 0, \sum_{i=1}^{N} \lambda_i = 1, \lambda_i m_i = \rho_i \right\}, \tag{4.29}$$

$$f^{**}(x, \rho) = F^{**}(\rho) + gM(\rho)x_3. \tag{4.30}$$

*Proof.* We denote by $Y_i$ the $i$th semi-axis $\{te_i : t \geqslant 0\}$ of $\mathbb{R}^N$, where $e_i$ is the unit vector in the $i$th direction. Since evidently $F^{**}(\rho) \geqslant (\psi^i)^{**}(\rho)$ for $\rho \in Y_i$, we may assume without loss of generality that each $\psi^i$ is convex.

We have that (cf. Ekeland & Temam, 1974, p. 260)

$$F^{**}(\rho) = \min\left\{ \sum_{k=1}^{N+1} \mu_k F(\rho^{(k)}) : \mu_k \geqslant 0, \sum_{k=1}^{N+1} \mu_k = 1, \sum_{k=1}^{N+1} \mu_k \rho^{(k)} = \rho \right\}. \tag{4.31}$$

Let $\rho \in \mathbb{R}^N_+$, and let $\mu_k, \rho^{(k)}$ realize the minimum in (4.31). If $\mu_k > 0$, then since $H(\rho^{(k)}) < \infty$ it follows that $\rho^{(k)}$ belongs to some $Y_i$. Let $P_1 = \{k : \mu_k > 0, \rho^{(k)} \in Y_1\}$ and $P_i = \{k : \mu_k > 0, \rho^{(k)} \in Y_i, \rho^{(k)} \neq 0\}, i = 2, \ldots, N$. If $P_i$ is nonempty let $\lambda_i = \sum_{P_i} \mu_k, r^{(i)} = \lambda_i^{-1} \sum_{P_i} \mu_k \rho^{(k)}$; otherwise set $\lambda_i = 0$, $r^{(i)} = 0$. Then $\sum_{i=1}^N \lambda_i = 1, \sum_{i=1}^N \lambda_i r^{(i)} = \rho$, and

$$\sum_{i=1}^{N} \lambda_i F(r^{(i)}) \leqslant \sum_{i=1}^{N} \lambda_i \sum_{P_i} \lambda_i^{-1} \mu_k F(\rho^{(k)}) = F^{**}(\rho), \tag{4.32}$$

where $\sum_{P_i}$ is zero if $P_i$ is empty. Let $m_i = r_i^{(i)}$. Then $\lambda_i m_i = \rho_i$. Since by (4.31) the left-hand side of (4.32) is not less than $F^{**}(\rho)$, we obtain (4.29). Since $M(\cdot)$ is linear, (4.30) follows. □

*Remark 4.4* If $\rho \in Y_i$ then $F^{**}(\rho) = (\psi^i)^{**}(\rho_i)$. In fact by the theorem $F^{**}(\rho) = \min\{\lambda(\psi^i)^{**}(m_i) : \lambda m_i = \rho_i, \lambda \in [0, 1]\}$, and since $(\psi^i)^{**}(0) = 0$ we have that $(\psi^i)^{**}(\rho_i) \leqslant \lambda(\psi^i)^{**}(m_i) + (1 - \lambda)(\psi^i)^{**}(0) \leqslant F^{**}(\rho)$.

The expression (4.29) may be easily computed in some special cases using Lagrange multipliers:

*Example 4.5* Let $\psi^i(\tau) = k_i \tau^\alpha, k_i > 0, \alpha > 1$. Then

$$F^{**}(\rho) = \left( \sum_{i=1}^{N} k_i^{1/\alpha} \rho_i \right)^\alpha. \tag{4.33}$$

*Example 4.6* Let $\psi^i(\tau) = k_i \tau \log \tau, k_i > 0, \alpha > 1$. Then

$$F^{**}(\rho) = \left( \sum_{i=1}^{N} k_i \rho_i \right) \log \left( \sum_{i=1}^{N} k_i \rho_i \right) - \sum_{i=1}^{N} \rho_i k_i \log k_i. \tag{4.34}$$

The numbers $\lambda_i$ in (4.29) may be interpreted as volume fractions, which are used as extra variables in theories of immiscible mixtures.

**Comments on theories incorporating density gradients and dynamics**

In all the above discussion we have ignored surface energy. This is often modelled by adding a term such as $\varepsilon|\nabla\rho(x)|^2$ to the integrand in (4.6), where $\varepsilon>0$ is constant, following the ideas of van der Waals (1893) and Cahn & Hilliard (1958). For the study of such problems and the taking of the limit $\varepsilon\to 0$ see, for example, Gurtin & Matano (1988), Modica (1987), Fonseca & Tartar (1989), and Baldo (1989). Adding such a term imposes strong conditions on the geometrical structure of minimizers, even in the limit $\varepsilon\to 0$, and prevents minimizing sequences from mixing the phases infinitely finely, thus converging to a Young measure minimizer that is not a Dirac mass.

Of course, there may nevertheless be dynamic mechanisms, consistent with thermodynamics, that induce finer and finer phase mixing down to a length scale at which surface energy takes over and prevents further mixing; such behaviour would lend interest to Young measure minimizers. Another possibility is that of fine mixing induced by negative surface energy. The rigorous study of such questions is in its infancy; for discussion and examples see Ball (1986), Pego (1987).

## 5. REMARKS ON YOUNG MEASURE MINIMIZERS IN NONLINEAR ELASTICITY

Consider a homogeneous elastic body occupying in some reference configuration the open set $\Omega\subset\mathbb{R}^3$. Configurations of the body are described by mappings $y:\Omega\to\mathbb{R}^3$. At some constant temperature we consider the problem of minimizing the total free energy

$$I(y)=\int_\Omega \psi(Dy(x))\,\mathrm{d}x \tag{5.1}$$

of the body subject to some boundary conditions, for example that

$$y|_{\partial\Omega_1} \text{ is given,} \tag{5.2}$$

for some subset $\partial\Omega_1$ of the boundary $\partial\Omega$. Here $Dy(x)$ denotes the deformation gradient and $\psi=\psi(A)$ the free energy function of the material, and we have ignored all other energy contributions.

In recent years it has emerged that there are materials for which (5.1), (5.2) provide a good mathematical model but for which the minimum is not attained. This has been established rigorously for certain models of elastic crystals by Ball & James (1987, 1989) (see also Chipot & Kinderlehrer, 1988; Fonseca, 1987; James & Kinderlehrer, 1989). The idea of investigating such possibilities has its roots in suggestions of Ericksen (see, for example, Ericksen, 1981).

In cases when the minimum is not attained, minimizing sequences form finer and finer *microstructures*, consisting, for example, of parallel layers in which the deformation gradient alternates between two essentially constant values. Such microstructure is frequently observed in both optical and electron micrographs of crystals, and can have a period of as little as a few atomic spacings (see, for example, the remarkable electron micrograph of microtwinning in NiMn of Baele *et al.*, 1987). Careful study of the minimization problem can lead to predictions concerning the geometry of such microstructures. As for the case of fluids, extra geometrical regularity is to be expected if surface energy is taken into account (see Parry, 1987; Fonseca, 1989); for example, this might be expected to give fine twinning a periodic structure (for a suggestive one-dimensional analysis see Müller, 1989).

If $y^{(j)}$ is a minimizing sequence for (5.1), (5.2), then under mild supplementary hypotheses there will be by Theorem 2.1 a subsequence $y^{(\mu)}$ such that $\delta_{Dy^{(\mu)}}$ converges weak* in $X = L^\infty(\Omega; M(\mathbb{R}^9))$ to a Young measure $\nu = (\nu_x)$. Note that $\nu_x$ is a probability measure on $3 \times 3$ matrices $M^{3\times 3} \cong \mathbb{R}^9$ for a.e. $x$. One could envisage analysing a generalized minimization problem among Young measures, in the spirit of (3.7), but this is made difficult by the fact that no characterization of Young measures that may be obtained as weak* limits of gradients is known. The properties of these Young measures that are used in Ball & James (1987, 1989) are those that result from the weak continuity of Jacobians (cf. Reshetnyak, 1968; Ball, Currie & Olver, 1981); more precisely, we have that if $Dy^{(j)}$ is bounded in $L^p(\Omega; \mathbb{R}^9)$ for some $p > 3$, then

$$\langle \nu_x, \operatorname{cof} A \rangle = \operatorname{cof} \langle \nu_x, A \rangle, \langle \nu_x, \det A \rangle = \det \langle \nu_x, A \rangle, \text{ a.e. } x \in \Omega, \tag{5.3}$$

where $\operatorname{cof} A$ denotes the matrix of cofactors of $A$. However, it is known that the relations (5.3) do not characterize Young measures coming from gradients.

Rather than describe in detail here the results that have been obtained for crystals, we confine ourselves to discussing a simple model example which shows how minimizing sequences are forced to have more and more microstructure in a way strongly resembling fine twinning in martensites. Let $Q = (0, 1)^2$, and consider the problem of minimizing

$$I(u) = \int_Q [(u_x^2 - 1)^2 + u_y^2]\, dx\, dy, \tag{5.4}$$

among scalar functions $u = u(x, y)$ satisfying the boundary condition

$$u|_{y=0} = 0. \tag{5.5}$$

In (5.4), $u_x, u_y$ denote the weak partial derivatives of $u$. We claim that the infimum of $I$ subject to (5.5) is zero, but that it is not attained. To prove the former statement, define $\bar{u}: \mathbb{R} \times (0, \infty) \to \mathbb{R}$ by

$$\bar{u}(x, y) = \begin{cases} x\phi(y) & \text{if } 0 \leqslant x \leqslant \frac{1}{2} \\ (1 - x)\phi(y) & \text{if } \frac{1}{2} \leqslant x \leqslant 1, \end{cases} \tag{5.6}$$

where $\phi(y) = y$ if $0 \leqslant y \leqslant 1$, $= 1$ if $y \geqslant 1$, extended as a 1-periodic function of $x$ to the whole of $\mathbb{R} \times (0, \infty)$. Then define $u^{(j)}(x, y) = j^{-1}\bar{u}(jx, jy)$. Now $Du^{(j)}(x, y) = (\bar{u}_x, \bar{u}_y)(jx, jy)$ is uniformly bounded and so

$$\lim_{j\to\infty} I(u^{(j)}) = \lim_{j\to\infty} \int_{Q \cap \{y \leqslant j^{-1}\}} [(u_x^{(j)2} - 1)^2 + u_y^{(j)2}]\, dx\, dy = 0. \tag{5.7}$$

Hence the infimum of $I$ subject to (5.5) is zero. It is not attained because any minimizer $u$ would satisfy $u_y \equiv 0$, which together with (5.5) implies that $u \equiv 0$ and hence that

$I(u)=1$, contradicting $\inf I=0$. The microstructure induced by $u^{(j)}$ consists of bands of width $(2j)^{-1}$ parallel to the $y$ axis in each of which $Du^{(j)}$ is constant, together with a boundary layer $0\leqslant y\leqslant j^{-1}$ to allow compatibility with the boundary condition (5.5).

We now determine the Young measure $\nu_{(x,y)}$ corresponding to $Du^{(j)}$ for any minimizing sequence $u^{(j)}$. We assume that a subsequence has already been extracted so that the Young measure is defined. In particular, $Du^{(j)}\rightharpoonup Du$ in $L^1(Q)$ for some $u$ satisfying (5.5). Since clearly $Du^{(j)}\to K\stackrel{\text{def}}{=}\{(-1,0),(1,0)\}$ in measure, by Theorem 2.1 we have that $\operatorname{supp}\nu_{(x,y)}\subset K$ a.e., i.e.

$$\nu_{(x,y)}=\lambda(x,y)\delta_{(-1,0)}+(1-\lambda(x,y))\delta_{(1,0)}, \tag{5.8}$$

where $0\leqslant\lambda(x,y)\leqslant 1$ a.e. Then $Du(x,y)=\langle\nu_{(x,y)},A\rangle=(1-2\lambda(x,y),0)$, which by (5.5) implies that $u=0$ and hence that $\lambda(x,y)=\frac{1}{2}$ a.e. We have thus proved that

$$\nu_{(x,y)}=\tfrac{1}{2}\delta_{(1,0)}+\tfrac{1}{2}\delta_{(-1,0)}. \tag{5.9}$$

In particular the Young measure is unique.

## REFERENCES


Aubert, G. and R. Tahraoui (1979) Théorèmes d'existence pour des problèmes du calcul des variations..., *J. Differential Equations*, **33**, 75–100.

Aumann, R. J. (1965) 'Integrals of set-valued functions', *J. Math. Anal. Appl.*, **12**, 1–12.

Baele, I., G. van Tendeloo and S. Amelinckx (1987) Microtwinning in Ni-Mn resulting from the $\beta$–$\vartheta$ martensitic transformation, *Acta Metallurgica*, **35**, 401–412.

Balder, E. J. (1984) A general approach to lower semicontinuity and lower closure in optimal control theory, *SIAM J. Control and Optimization*, **22**, 570–598.

Baldo, S. (1989) Minimal interface criterion for phase transitions in mixtures of Cahn-Hilliard fluids, *Annales de l'Inst. Henri Poincaré—Analyse Non Linéaire*, to appear.

Ball, J. M. (1984) Material instabilities and the calculus of variations, Proc. conference on *Phase transformations and material instabilities in solids*, Mathematics Research Center, University of Wisconsin, Academic Press, Publication No. 52, 1–20.

Ball, J. M. (1986) Minimizing sequences in thermomechanics, in Proc. Meeting on *Finite thermoelasticity*, Accademia Nazionale dei Lincei.

Ball, J. M. (1988) 'Loss of the constraint in convex variational problems', in 'Analyse mathématique et applications', Contributions en l'honneur de J.-L. Lions, Gauthier-Villars, 1988.

Ball, J. M. (1989) A version of the fundamental theorem for Young measures, Proceedings conference on *Partial Differential Equations and Continuum Models of Phase Transitions*, Nice, 1988, ed. D. Serre, to appear.

Ball, J. M., J. C. Currie and P. Olver (1981) Null Lagrangians, weak continuity, and variational problems of arbitrary order, *J. Functional Anal.*, **41**, 135–174.

Ball, J. M. and R. D. James (1987) Fine phase mixtures as minimizers of energy, *Arch. Rat. Mech. Anal.*, **100**, 13–52.

Ball, J. M. and R. D. James (1989) Proposed experimental tests for a theory of fine microstructure and the two well problem, to appear.

Ball, J. M. and G. Knowles (1986) Lyapunov functions for thermomechanics with spatially varying boundary temperatures, *Arch. Rat. Mech. Anal.*, **92**, 193–204.

Bedford, A. and D. S. Drumheller (1983) Theories of immiscible and structured mixtures, *Int. J. Engng. Sci.*, **21**, 863–960.

Berliocchi, H. and J. M. Lasry (1973) Intégrandes normales et mesures paramètres en calcul des variations, *Bull. Soc. Math. France*, **101**, 129–184.

Cahn, J. W. and J. E. Hilliard (1958) Free energy of a nonuniform system. I. Interfacial free energy, *J. Chem. Physics*, **28**, 258–267.

Cesari, L. (1983) *Optimization—Theory and Applications*, Springer.

Chipot, M. and D. Kinderlehrer (1988) Equilibrium configurations of crystals, *Arch. Rat. Mech. Anal.*, **103**, 237–278.
Clarke, F. W. (1983) *Optimization and Nonsmooth Analysis*, Wiley-Interscience.
Coleman, B. D. and E. H. Dill (1973) On thermodynamics and the stability of motion of materials with memory, *Arch. Rat. Mech. Anal.*, **51**, 1–53.
DiPerna, R. J. (1983) Convergence of approximate solutions to conservation laws, *Arch. Rat. Mech. Anal.*, **82**, 27–70.
Duhem, P. (1911) *Traité d'énergetique ou de thermodynamique générale*, Gauthier-Villars, Paris.
Dunn, J. E. and R. Fosdick (1980) The morphology and stability of material phases, *Arch. Rat. Mech. Anal.*, **74**, 1–99.
Ekeland, I. and R. Temam (1974) *Analyse convexe et problèmes variationnels*, Dunod.
Ericksen, J. L. (1966) Thermoelastic stability, *Proc. 5th National Cong. Appl. Mech.*, 187–193.
Ericksen, J. L. (1981) Continuous martensitic transitions in thermoelastic solids, *J. of Thermal Stresses*, **4**, 107–119.
Fonseca, I. (1987) Variational methods for elastic crystals, *Arch. Rat. Mech. Anal.*, **97**, 189–220.
Fonseca, I. (1989) Phase transitions of elastic solid materials, *Arch. Rat. Mech. Anal.*, **107**, 195–224.
Fonseca, I. (1989) Phase transitions of elastic solid materials, *Arch. Rat. Mech. Anal.*, to appear.
Fonseca, I. and L. Tartar (1989) The gradient theory of phase transitions for systems with two potential wells, *Proc. Royal Soc. Edinburgh A*, **111**, 89–102.
Fosdick, R. and J. Patiño (1986) On the Gibbsian thermostatics of mixtures, *Arch. Rat. Mech. Anal.*, **93**, 203–222.
Gibbs, J. W. (1873) Graphical methods in the thermodynamics of fluids, *Trans. Connecticut Acad.*, **2**, 309–342.
Gurtin, M. E. and H. Matano (1988) On the structure of equilibrium phase transitions within the gradient theory of fluids, *Quarterly of Appl. Math.*, **46**, 301–318.
James, R. D. (1979) Co-existent phases in the one-dimensional static theory of elastic bars, *Arch. Rat. Mech. Anal.*, **72**, 99–140.
James, R. D. and D. Kinderlehrer (1989) Theory of diffusionless phase transitions, Proceedings of conference on *Partial Differential Equations and Continuum Models of Phase Transitions*, Nice, 1988, ed. D. Serre, to appear.
Kinderlehrer, D. (1988) Remarks about equilibrium configurations of crystals, in *Material Instabilities in Continuum Mechanics*, ed. J. M. Ball, Oxford University Press, pp. 217–241.
Landau, L. D. and E. M. Lifshitz (1970) *Statistical Physics*, Pergamon, Oxford.
Lin, P. (1989) Maximization of the entropy for an elastic body free of surface traction, to appear.
MacShane, E. J. (1940) Generalized curves, *Duke Math. J.*, **6**, 513–536.
MacShane, E. J. (1978) The calculus of variations from the beginning through optimal control theory, in *Optimal Control and Differential Equations*, Academic Press.
Man, C.-S. (1985a) Dynamic admissible states, negative absolute temperature, and the entropy maximum principle, *Arch. Rat. Mech. Anal.*, **90**, 263–289.
Man, C.-S. (1985b) Material stability, the Gibbs conjecture and the first phase rule for substances, *Arch. Rat. Mech. Anal.*, **91**, 1–54.
Marcellini, P. (1980) Alcune osservazione sull' esistenza del minimo di integrali del calcolo delle variazioni senza ipotesi di convessità, *Rendiconti di Matematica*, **13**, 271–281.
Mascolo, E. and R. Schianchi (1987) Existence theorems in the calculus of variations, *J. Differential Equations*, **67**, 185–198.
Modica, L. (1987) Gradient theory of phase transitions and minimal interface criteria, *Arch. Rational Mech. Anal.*, **98**, 357–383.
Müller, S. (1989) Minimizing sequences for nonconvex functionals, phase transitions and singular perturbations, Proc. conference on '*Problems Involving Change of Type*', Springer, to appear.
Noll, W. (1970) On certain convex sets of measures and on phases of reacting mixtures, *Arch. Rat. Mech. Anal.*, **38**, 1–12.
Olech, C. (1970) Integrals of set-valued functions and linear optimal control problems, *Colloque sur la théorie mathématique du contrôle optimal*, C.B.R.M. Vander Louvain.
Parry, G. P. (1987) On shear bands in crystals, *J. Mech. Phys. Solids*, **35**, 367–382.
Patiño, J. G. S. (1987) On nonconvex optimization with integral constraints, *J. Optimization Theory and Applications*, **55**, 391–402.

Pego, R. L. (1987) Phase transitions in one-dimensional nonlinear viscoelasticity: admissibility and stability, *Arch. Rat. Mech. Anal.*, **97**, 353–394.

Penrose, O. and J. L. Lebowitz (1971) Rigorous treatment of metastable states in the van der Waals–Maxwell theory, *J. Statistical Physics*, **3**, 211–236.

Reshetnyak, Yu. G. (1968) Stability theorems for mappings with bounded excursion, *Siberian Math. J.*, **9**, 499–512.

Rudin, W. (1973) *Functional Analysis*, McGraw-Hill.

Tartar, L. (1979) Compensated compactness and applications to partial differential equations, in *Nonlinear Analysis and Mechanics, Heriot-Watt Symposium*, Vol. IV, Pitman Research Notes in Mathematics, pp. 136–192.

Thompson, C. J. (1972) *Mathematical Statistical Mechanics*, Macmillan, New York.

van der Waals, J. D. (1893) The thermodynamic theory of capillarity under the hypothesis of a continuous variation in density (in Dutch), *Verhandel. Konink. Akad. Weten. Amsterdam* (*Sec. 1*), **1**, No. 8.

Young, L. C. (1937) Generalized curves and the existence of an attained absolute minimum in the calculus of variations, *Comptes Rendus de la Société des Sciences et des Lettres de Varsovie, classe III*, **30**, 212–234.

Young, L. C. (1969) *Lectures on the Calculus of Variations and Optimal Control Theory*, Saunders, reprinted by Chelsea (1980).