MATH. SCAND. 54 (1984), 133–156

# SOME CONVEXITY QUESTIONS ARISING IN STATISTICAL MECHANICS

R. B. ISRAEL and R. R. PHELPS

**Abstract.**

An abstract version of a mathematical model for the classical theory of lattice gases leads quite naturally to the study of a class $\mathscr{P}$ of generalized "pressure" functions, which are convex and continuous on the Banach space $A(K)$ of affine real-valued continuous functions on a compact convex set $K$. The differentiability properties of the members of $\mathscr{P}$, as well as the extremal structure of the convex set $\mathscr{P}$ itself, are investigated by means of the Fenchel duality between the members of $\mathscr{P}$ and the members of a certain cone $\mathscr{H}$ of lower semicontinuous convex functions on $K$. New results on the differentiability of the physical pressure are obtained in this context.

## 1. Introduction.

An abstract version of a mathematical model used in statistical mechanics consists of a compact metric space $\Omega$, an infinite Abelian group $G$ of homeomorphisms of $\Omega$ (isomorphic to the group of translations of the $n$-dimensional lattice $\mathbf{Z}^n$) and the space $C(\Omega)$ of continuous real-valued functions on $\Omega$. We denote the group action by $\tau$ (so for each $x \in G$, $\tau^x$ is a homeomorphism of $\Omega$) and we assume $\tau$ is expansive, that is, there exists $\varepsilon > 0$ such that $\omega, \omega' \in \Omega$ and $d(\tau^x\omega, \tau^x\omega') \leqq \varepsilon$ for all $x \in G$ imply $\omega = \omega'$. If we let $\mathscr{L}$ be the closed linear span in $C(\Omega)$ of all functions of the form

$$f \circ \tau^x - f, \qquad x \in G,\ f \in C(\Omega)\,,$$

then $\mathscr{L}^\perp \subseteqq C(\Omega)^*$ consists of the $G$-invariant measures and we know from the amenability of $G$ that the weak* compact convex simplex $K$ of invariant probability measures is non-empty. The elements of $C(\Omega)$ are called "interactions" and those in $K$ are called "invariant states". There are two

Received August 4, 1982.

The second named author wishes to acknowledge helpful conversations with Professor Isaac Namioka. Research by the first named author was supported in part by Grant A-4015 from the Natural Sciences and Engineering Research Council of Canada while visiting the Department of Mathematics at Rutgers University.

related objects of study: one is the *pressure*, which (if finite) is a convex continuous function $P$ on $C(\Omega)$, and the other is the *mean entropy*, which is a bounded, nonnegative affine upper-semicontinuous function $h$ on $K$ (which is not, in general, continuous). The pressure has a number of interesting and special properties; letting $q: C(\Omega) \to C(\Omega)/\mathscr{L}$ denote the quotient map, one has

(P1) $$P(f+r) = P(f)+r, \quad f \in C(\Omega),\ r \in \mathbf{R}.$$

(P2) $$P(f) \leqq P(0), \quad f \leqq 0.$$

(P3) $$P(f) \geqq \|q(f)\|, \quad f \geqq 0.$$

(P4) $$P(f+g) = P(f) \quad \text{whenever } g \in \mathscr{L}.$$

Further properties turn out to be consequences of these four; for instance,

$$|P(f)-P(g)| \leqq \|f-g\|,\ P(f+g) \leqq P(f)+P(g) \text{ and } P(f) \leqq P(g) \text{ if } f \leqq g,$$

all of which are immediate from Proposition 2.2. (Property (P3) is an important one which seems not to have been noted explicitly.)

One of the central theorems in the subject is the variational principle, which relates the pressure $P$ and entropy $h$ by

$$P(f) = \sup\{\mu(f)+h(\mu) : \mu \in K\}, \quad f \in C(\Omega).$$

A dual version of this is

$$h(\mu) = \inf\{P(f)-\mu(f) : f \in C(\Omega)\}, \quad \mu \in K.$$

In the cases most often studied by mathematical physicists, $\Omega$ is the cartesian product $W^{\mathbf{Z}^n}$, where $W$ is a finite set, and $\tau$ is the natural action $(\tau^x\omega)_y = \omega_{y-x}$. These models are sometimes called "classical lattice gases". In "quantum lattice gases" the framework is somewhat different: $C(\Omega)$ is replaced by the self-adjoint elements of a certain non-abelian quasilocal C*-algebra $\mathscr{A}$, and $K$ consists of the $G$-invariant states on $\mathscr{A}$. All the properties of the pressure and the mean entropy which we have mentioned may be extended to the case of quantum lattice gases.

From a geometric point of view, perhaps the most remarkable theorem in this subject is the following one, which is valid for both the classical and quantum lattice gases.

1.1. THEOREM. *The metrizable simplex $K$ is a (the) Poulsen simplex, that is, its extreme points (the ergodic states) are weak* dense in $K$. In fact, for each $r \in \mathbf{R}$, those extreme points of $K$ which are contained in $\{\mu \in K: h(\mu) > r\}$ are weak* dense in this set.*

The fact that $K$ is the Poulsen simplex appears to have been proved first by Ruelle [11, p. 197] in the classical case. A proof of this for the quantum case is given by Bratteli and Robinson [3, vol. I p. 398], who attribute it to Ruelle. (See, also, the survey by Olsen [9].) The stronger assertion concerning the density of extreme points in the "slices" defined by $h$ was proved in [6, p. 94]. As will be seen later, it contains a great deal of information about differentiability of the pressure. We refer to [6] and the monographs by Ruelle [11, 12] for detailed expositions of the material touched on this section.

**2. Geometric reformulation.**

In what follows, we will put these matters in the setting of compact convex sets $K$ and the associated Banach spaces $A(K)$ of affine continuous real valued functions on $K$ (with supremum norm). (See [1], [2] or [10] for relevant background material.) The pressure will then be seen to be just one member of a *class* $\mathscr{P}$ of functions on $A(K)$ satisfying properties (P1)–(P3). That this is a reasonable approach arises from two facts. First, the Banach space of interactions studied in [6] and [12] admits a continuous linear mapping onto $C(\Omega)$ (see [12, p. 37]) in such a way that the physical pressure as originally defined is constant on the kernel of this map [12, p. 39]; this allows one to consider it as a function on $C(\Omega)$. Second, as we have noted, the pressure can actually be considered as a function on $C(\Omega)/\mathscr{L}$, and the latter is linearly isometric with $A(K)$ (below). Thus, any differentiability properties we obtain for the elements of $\mathscr{P}$ are automatically valid for the physical pressure.

2.1. PROPOSITION. *The space $C(\Omega)/\mathscr{L}$ is linearly isometric with $A(K)$, where $K$ is the simplex of $G$-invariant probability measures on $\Omega$.*

PROOF. Given an element of $C(\Omega)/\mathscr{L}$, that is, an element of the form $q(f)$ where $f \in C(\Omega)$, let $f^\wedge \in A(K)$ be defined by $f^\wedge(\mu) = \mu(f)$, $\mu \in K \subseteqq \mathscr{L}^\perp$. Clearly,

$$\|f^\wedge\| = \sup\{|\mu(f)| : \mu \in K\} \leqq \sup\{|\mu(f)| : \mu \in \mathscr{L}^\perp, \|\mu\| \leqq 1\}$$
$$= \|q(f)\| .$$

To establish the reverse inequality, given $q(f)$ choose, by the Hahn–Banach theorem, $L \in (C(\Omega)/\mathscr{L})^*$ such that $L(q(f)) = \|q(f)\|$ and $\|L\| = 1$. Since $(C(\Omega)/\mathscr{L})^*$ can be identified with $\mathscr{L}^\perp$, we can consider $L$ to be a signed invariant measure $\mu$ on $\Omega$ of norm 1. Since the total variation of an invariant measure is invariant also, we have $|\mu| \in K$. Now

$$\|q(f)\| = L(q(f)) = \mu(f)) = \mu(f) \leqq \||\mu|(f)| = |f^\wedge(|\mu|)| \leqq \|f^\wedge\| ,$$

so we conclude that the mapping $q(f) \to f^\wedge$ is an isometry. To see that it is a surjection, suppose that $g \in A(K)$. Any $\mu \in \mathscr{L}^\perp$ admits the unique Jordan decomposition $\mu = \mu^+ - \mu^- = a_1\mu_1 - a_2\mu_2$ where $a_1 = \mu^+(\Omega)$, $a_2 = \mu^-(\Omega)$ and $\mu_i \in K$, $i = 1, 2$. The linear functional

$$\mu \to a_1 g(\mu_1) - a_2 g(\mu_2)$$

is well-defined and, using weak* compactness and the fact that $g$ is affine and continuous, it is weak* continuous on bounded subsets of $\mathscr{L}^\perp$. By the Krein–Šmul'yan theorem, it is weak* continuous on $\mathscr{L}^\perp$ hence arises from an element $q(f)$ of $C(\Omega)/\mathscr{L}$, which completes the proof. [The above isometry is almost order-preserving: The positive cone $A(K)^+$ is the uniform closure of $q(C(\Omega)^+)$.]

The foregoing reformulation can also be carried out for quantum lattice gases.

In what follows, then, we will let $K$ be a compact convex set (in some locally convex space) and let $A(K)$ denote the real-valued affine continuous functions on $K$. We can embed $K$ in the obvious way in $A(K)^*$ and hence there is a natural *norm* topology on $K$, as well. Let $\mathscr{P}$ denote the family of all continuous convex functions $P$ on $A(K)$ which satisfy

(P1) $\quad P(f+r) = P(f)+r, \quad f \in A(K),\ r \in \mathbf{R}$.

(P2) $\quad P(f) \leqq P(0) \quad$ if $f \leqq 0$.

(P3) $\quad P(f) \geqq \|f\| \quad$ if $f \geqq 0$.

EXAMPLE. Consider $P_0(f) = \sup\{f(x) \colon x \in K\}$.

Recall the definition of the Fenchel dual $P^*$ of a convex function $P$: For any $\mu \in A(K)^*$,

(*) $$P^*(\mu) = \sup\{\mu(f) - P(f) : f \in A(K)\}.$$

This is always weak* lower semicontinuous and convex (possibly taking the value $+\infty$) and since $P$ is convex and lower semicontinuous, the dual of $P^*$ on $A(K)^{**}$ defined by

$$P^{**}(x^{**}) = \sup\{\langle x^{**}, \mu\rangle - P^*(\mu) : \mu \in A(K)^*\}$$

has the property that its restriction to $A(K)$ coincides with $P$, that is

(**) $$P(f) = \sup\{\mu(f) - P^*(\mu) : \mu \in A(K)\}.$$

The properties of $P$ turn out to be precisely what we need to show that $P^*$ is really a function on $K \subseteqq A(K)^*$; that is, $P^*(\mu) < +\infty$ if and only if $\mu \in K$. This is contained in the following proposition.

2.2. PROPOSITION. *For any $P \in \mathscr{P}$ the lower semicontinuous convex function $P^*$ satisfies*

(i) $P^*(\mu) \leqq 0$ *and* $P^*(\mu) \geqq -P(0)$ *for all* $\mu \in K$

(ii) $P^*(\mu) = +\infty$ *if* $\mu \in A(K)^* \setminus K$.

PROOF. i). The second assertion is immediate if one takes $f=0$ in (*). To prove the first assertion, suppose that $f \in A(K)$; then $f + \|f\| \geqq 0$, so by (P3), (P1) and the definition of the norm in $A(K)$ we have

$$\mu(f) + \|f\| \leqq \|(f + \|f\|)\| \leqq P(f + \|f\|) = P(f) + \|f\|$$

hence $\mu(f) - P(f) \leqq 0$ for all $f \in A(K)$. This implies (i).

(ii). If $\mu \notin K$, then either $\mu(1) \neq 1$ or $\mu(1) = 1$ (hence $\|\mu\| \geqq 1$) and $\|\mu\| > 1$. In the first case, let $f \equiv r$ in (*) to get (using (P1))

$$P^*(\mu) \geqq \mu(r) = r\mu(1) - P(0) - r = r[\mu(1) - 1] - P(0) .$$

By the correct choice of the sign of $r$ we can make the right side as a big as we wish, so $P^*(\mu) = +\infty$. If $\|\mu\| > 1$, then there exists $f \in A(K)$ such that $\mu(f) > \|f\|$. For any $r > 0$ we have $g \equiv r(f - \|f\|) \leqq 0$ so by (P2) we know that $P(g) \leqq P(0)$. Consequently,

$$P^*(\mu) \geqq \mu(g) - P(g) \geqq r[\mu(f) - \|f\|] - P(0)$$

for any $r > 0$, therefore $P^*(\mu) = +\infty$.

As a result of this proposition we can rewrite (**) in the form of a supremum over $K$:

$$(**)' \qquad P(f) = \sup\{\mu(f) - P^*(\mu) : \mu \in K\} .$$

Note that if we let $h = -P^*$, then this has the same form as the variational principle cited earlier; moreover, (*) now has the form

$$h(\mu) = -P^*(\mu) = \inf\{P(f) - \mu(f) : f \in A(K)\} ,$$

which is the dual version of the variational principle. In other words, *the variational principle* (*or its dual*) *is an instance of Fenchel duality.* What we have essentially shown is that once the properties (P1)–(P4) are established for the pressure, then the entropy can be uniquely defined by duality. One can reverse this, of course: If the entropy is given, then the pressure can be obtained from it by duality. The next result makes this correspondence more explicit. We will henceforth use $h$ for $P^*$; the special case of the mean entropy will therefore be the negative of one of our functions.

DEFINITION. Let $\mathscr{H}$ denote the convex cone of all bounded lower semicontinuous convex functions $h$ on $K$ for which $h \leqq 0$.

2.3. PROPOSITION. *If $h \in \mathscr{H}$, then the function $P_h$ defined on $A(K)$ by*

$$P_h(f) = \sup\{f(x) - h(x) : x \in K\}$$

*is a member of $\mathscr{P}$. The correspondence $h \to P_h$ is an order-reversing map of the convex cone $\mathscr{H}$ onto the convex set $\mathscr{P}$, and it is an isometry between the supremum metric on $\mathscr{H}$ and the supremum metric on $\mathscr{P}$. If K contains more than one point, then this map is not affine.*

PROOF. It is straightforward to verify that $P_h$ satisfies (P1)–(P3). If $h_1 \leqq h_2$, then clearly $P_{h_1}(f) \geqq P_{h_2}(f)$ for all $f$. We have already proved that the map $h \to P_h$ is onto. To show that it is an isometry, suppose that $h_1, h_2 \in \mathscr{H}$ and let $P_1, P_2$ denote the corresponding elements of $\mathscr{P}$. With the supremum norm on $\mathscr{H}$ we have $-h_1 \leqq -h_2 + \|h_1 - h_2\|$ and hence

$$P_1 \leqq P_2 + \|h_1 - h_2\| .$$

Reversing the roles of $h_1, h_2$ we get $|P_1 - P_2| \leqq \|h_1 - h_2\|$ and therefore

$$\|P_1 - P_2\|_\infty \leqq \|h_1 - h_2\| .$$

On the other hand, we can apply the same kind of argument by starting with $P_1 \leqq P_2 + \|P_1 - P_2\|_\infty$ and using $h_i(x) = \sup\{f(x) - P_i(f) : f \in A(K)\}$ to obtain the reverse inequality.

If $K$ contains more than one point we can choose $h_1, h_2 \in A(K)$ such that $h_1, h_2 \leqq 0$, $h_1 + h_2 = -1$ and $\inf h_i < -1/2$. Let $h = (1/2)(h_1 + h_2)$; then $h \equiv -1/2$ and therefore $P_h(0) = 1/2$. Also, $P_{h_i}(0) = \sup(-h_i) > 1/2$, so $P_h(0) \neq (1/2)[P_{h_1}(0) + P_{h_2}(0)]$, which proves that our correspondence is not affine.

## 3. Differentiablity properties of pressure functions.

One of the basic questions of physical interest is that of differentiability of the pressure. Recall that, since $A(K)$ is a separable Banach space whenever $K$ is metrizable (as it is in the physically significant case), Mazur's theorem is applicable and one can conclude that each member $P$ of $\mathscr{P}$ has a dense $G_\delta$ set of points of Gateaux differentiability. A stronger assertion would be that the set $G$ of all points $g$ in $A(K)$ such that $P$ is Gateaux differentiable at $\lambda g$, for all $\lambda \neq 0$, is also a dense $G_\delta$. This last assertion is untrue, as can readily be seen by examining simple examples when $K$ is the two dimensional simplex, say. It *is* true, however, for the physical pressure; this will be shown below. In fact, even more is true: Suppose that $S$ is any nonempty $\sigma$-compact subset of $A(K)$ [for instance, take $S = \{0\}$, or $S$ to be a countable dense set, or a dense countable union of finite dimensional subspaces of $A(K)$] and let $G_1$ be the set of all $g$ in $A(K)$ such that the pressure is Gateaux differentiable at $f+\lambda g$ whenever $f \in S$ and $\lambda \neq 0$; then $G_1$ is a dense $G_\delta$ subset of $A(K)$. This says that the set where the pressure is Gateaux differentiable contains many translates of many one-dimensional subspaces (the origin excluded). A more general assertion is the claim that this set contains many translates of many $n$-dimensional subspaces (the origin excluded); this is, in fact, contained in Theorem 3.2 below. We first recall the definition of the subdifferential of a convex function $P$ and its relationship to Gateaux and Fréchet differentiability.

DEFINITION. For $f \in A(K)$ the *subdifferential* $\partial P(f)$ of $P$ at $f$ is the set of all $\mu \in A(K)^*$ which satisfy

$$\mu(g)-\mu(f) \leqq P(g)-P(f), \qquad g \in A(K)\,.$$

This says that the continuous affine function on $A(K)$ defined by $g \to \mu(g) -\mu(f)+P(f)$ is dominated by $P$ and equals $P$ at $g=f$; geometrically, this means that the graph of this affine function (which is a hyperplane in $A(K) \times \mathbf{R}$) supports the convex epigraph of $P$ at $f$. The continuity of $P$ and the Hahn–Banach theorem guarantee that $\partial P(f)$ is always nonempty, and it is easily seen to be weak* compact and convex. Note that if $\mu \in \partial P(f)$, then for all $g \in A(K)$

$$\mu(g)-P(g) \leqq \mu(f)-P(f)$$

so that

$$P^*(\mu) \equiv \sup\{\mu(g)-P(g) : g \in A(K)\} = \mu(f)-P(f)\,.$$

Since this is finite, we conclude that $\mu \in K$. Thus, *for members of $\mathscr{P}$, the subdifferential is a subset of* $K$ and $x \in \partial P(f)$ if and only if $h(x)=f(x)-P(f)$. This can be rewritten as $x \in \partial P(f)$ if and only if $P(f)=f(x)-h(x)$; in view of Proposition 2.3, this is equivalent to saying that the supremum which characterizes $P$ is attained at $x$. An easily verified but important consequence of these observations is the fact (which we will frequently use) that *if $h \in \mathscr{H}$ is affine, then $\partial P_h(f)$ is a closed face of $K$*; in particular, the extreme points of $\partial P_h(f)$ are extreme points of $K$.

Differentiability of $P$ at a point $f$ can be characterized in terms of the subdifferential, as follows. (We may take these as the *definitions* of differentiability. See, for instance, [5, pp. 147–148].)

$P$ is *Gateaux differentiable* at $f$ if and only if $\partial P(f)$ consists of a single point.

$P$ is *Fréchet differentiable* at $f$ if and only if it is Gateaux differentiable there and the set-valued map $g \to \partial P(g)$ is norm-norm upper semicontinuous at $f$. Equivalently, $\partial P(f)=\{x\}$ and if $\|f_n-f\| \to 0$ and $x_n \in \partial P(f_n)$, then $\|x_n-x\| \to 0$.

There are two major ingredients in our results concerning differentiability. The first of these is the extreme point density property for mean entropy which was described in Theorem 1.1 and which we now single out.

DEFINITION. An element $h$ of $\mathscr{H}$ will be said to have property (D) provided

(D) $\quad \operatorname{ext}K \cap \{x \in K : h(x) < r\}$ is dense in $\{x \in K : h(x) < r\}$ ,

for all $r \in \mathbf{R}$ .

Theorem 1.1 says that the (negative of) the mean entropy has property (D). It is also clearly satisfied by any continuous $h$ on the Poulsen simplex.

The second major ingredient is Edwards' separation theorem for simplexes (see [1] or [2]).

3.1. THEOREM (D. A. Edwards). *Suppose that $K$ is a simplex and suppose that $h_1$ and $-h_2$ are upper semicontinuous convex functions on $K$ with $h_1 \leqq h_2$. Then there exists $g \in A(K)$ such that $h_1 \leqq g \leqq h_2$.*

We will also use the fact that the convex hull of finitely many extreme points of a simplex $K$ is a *face* of $K$.

3.2. THEOREM. *Suppose that $h$ is a bounded lower semicontinuous affine function on the metrizable simplex $K$ and that $h$ has property* (D). *Let $S$ be any $\sigma$-compact subset of $A(K)$ and for each $n \geqq 1$ let $G_n$ be the set of all $n$-tuples $\mathbf{g} = (g_1, g_2, \ldots, g_n)$ in $A(K)^n$ such that $P \equiv P_h$ is Gateaux differentiable at each point $f + \boldsymbol{\lambda} \cdot \mathbf{g} \equiv f + \sum_{i=1}^n \lambda_i g_i$ whenever $f \in S$ and $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_n) \in \mathbf{R}^n \setminus \{0\}$. Then $G_n$ is a dense $G_\delta$ subset of $A(K)^n$.*

PROOF. For each $a$ in $A(K)$ let

$$G_n^a = \{\mathbf{g} \in A(K)^n : a(x_1) = a(x_2) \text{ whenever } f \in S,\ \boldsymbol{\lambda} \in \mathbf{R}^n \setminus \{0\} \text{ and } x_1, x_2 \in \partial P(f + \boldsymbol{\lambda} \cdot \mathbf{g})\} .$$

By separability of $A(K)$, the set $G_n$ is a countable intersection of sets $G_n^a$, so it suffices to prove that each of these is a dense $G_\delta$. Fix $a$ in $A(K)$, then, and consider

$$C = \{f \in A(K) : \text{There exist } x_1, x_2 \in \partial P(f) \text{ with } a(x_1) \neq a(x_2)\} .$$

By using the compactness of $K$ it is not difficult to show that for each $i \geqq 1$ the set

$$C_i = \{f \in A(K) : \text{There exist } x_1, x_2 \in \partial P(f) \text{ with } a(x_1) \geqq a(x_2) + i^{-1}\}$$

is closed; it follows that $C = \bigcup C_i$ is an $F_\sigma$. By writing $S = \bigcup S_k$ and $\mathbb{R}^n \setminus \{0\} = \bigcup K_j$ where the $S_k$ and $K_j$ are compact, we can express $G_n^a$ as a countable intersection of the sets

$$H_{i,j,k} = \{\boldsymbol{g} \in A(K)^n : C_i \cap (S_k + K_j \cdot \boldsymbol{g}) = \varnothing\} .$$

If $\boldsymbol{g} \in H_{i,j,k}$, then dist $(S_k + K_j \cdot \boldsymbol{g}, C_i) > 0$; it follows that $H_{i,j,k}$ is open and hence that $G_n^a$ is a $G_\delta$. We must prove that $G_n^a$ is dense. For this, it suffices to prove that $G_1^a$ is dense, since the rest of the proof will be a consequence of the following induction argument: If $(g_1, \ldots, g_{n-1}) \in G_{n-1}^a$, let

$$S' = S + \text{span}\{g_1, \ldots, g_{n-1}\}$$

(this is again $\sigma$-compact) and consider the set

$$G^a(g_1, \ldots, g_{n-1}) = \{g_n \in A(K) : (g_1, \ldots, g_{n-1}, g_n) \in G_n^a\} .$$

One can verify that this contains the set $G_1^a$ (based now on $S'$) by using the fact that $(g_1, \ldots, g_{n-1}) \in G_{n-1}^a$. Assuming that sets of the form $G_1^a$ are dense in $A(K)$, then, we see that each $G^a(g_1, \ldots, g_{n-1})$ is itself dense in $A(K)$. If we now assume by induction that $G_{n-1}^a$ is dense in $A(K)^{n-1}$, it follows that $G_n^a$ (which contains the product $(g_1, \ldots, g_{n-1}) \times G^a(g_1, \ldots, g_{n-1})$ for each $(g_1, \ldots, g_{n-1})$ in $G_{n-1}^a$) is dense in $A(K)^n = A(K)^{n-1} \times A(K)$.

In order to show that $G_1^a$ is dense, write $G_1^a = G_+^a \cap G_-^a$, where $G_+^a$ (respectively $G_-^a$) is defined in the same way as $G_1^a$, but with the restriction $\lambda > 0$ (respectively $\lambda < 0$). Then the same argument as above shows that $G_+^a$ and $G_-^a$ are $G_\delta$ sets, and it suffices to prove that $G_+^a$ is dense. (By symmetry, $G_-^a = -G_+^a$ will also be dense.) Note that the set $G_+^a$ becomes even smaller if the set $S$ becomes bigger, so we may assume without loss of generality that $S$ is a dense linear subspace (since $A(K)$ is separable and the linear span of a $\sigma$-compact set is $\sigma$-compact). This implies that if $g \in G_+^a$, then so is $f + g$ for any $f \in S$. Thus, it suffices to show that $G_+^a$ is nonempty, since if $g \in G_+^a$, then $G_+^a$ will contain the dense set $g + S$.

We now construct an element $g$ in $G_+^a$. Let $S = \bigcup S_j$ with $S_j$ compact, and let $d$ be a metric on $K$. We may assume $a \in S_1$ and $\|a\| = 1$. Define inductively sequences $\delta_j > 0$, $x_i \in \text{ext}\, K$, $g_j \in A(K)$ as follows:

(1) Choose a decreasing sequence $\delta_j$ so that $|f(x_1) - f(x_2)| < 3^{-j}$ whenever $f \in \bigcup_{i=1}^j S_i \cup \{g_1, \ldots, g_{j-1}\}$ and $x_1, x_2 \in K$ with $d(x_1, x_2) < \delta_j$; this is possible since compact subsets of $A(K)$ are equicontinuous.

(2) Suppose that $x_1, \ldots, x_{m(j-1)} \in \text{ext}\, K$ have been chosen, for some integer $m(j-1) > 0$. Since $h$ is lower semicontinuous, for each $x' \in \text{ext}\, K$ the set

$$\{x \in K : d(x, x') < \delta_j \text{ and } h(x) > h(x') - 3^{-j}\}$$

is open. Moreover, since $h$ satisfies property (D), the union of these sets covers $K$. By compactness, we can choose $x_{m(j-1)+1},\ldots,x_{m(j)}$ in ext $K$ such that the corresponding sets cover $K$, hence for each $x \in K$ there is some $i$ with $i \leqq m(j)$, $d(x,x_i)<\delta_j$ and $h(x_i)<h(x)+3^{-j}$.

(3) Use Edwards' theorem to choose $g_j \in A(K)$ such that $g_j(x_i)=a(x_i)^2$, $i=1,\ldots,m(j)$ and $a^2 \leqq g \leqq 1$, by noting that the function which is defined to be equal to $a(x_i)^2$ at the finitely many extreme points $x_i$ and equal to 1 elsewhere is lower semicontinuous and concave, and clearly dominates the convex continuous function $a^2$.

Next, let $g=-\sum 2^{-j}g_j$. Suppose that $g$ were not in $G^a_+$. Then there would exist $f \in S$, $\lambda>0$ and $x',x'' \in \partial P(f+\lambda g)$ such that $\varepsilon=[a(x')-a(x'')]^2>0$. Since $h$ is affine we should have $x=(1/2)(x'+x'') \in \partial P(f+\lambda g)$. We claim, however, that for some $i$,

$$(f+\lambda g-h)(x_i) > (f+\lambda g-h)(x),$$

which implies that $x \notin \partial P(f+\lambda g)$. Indeed, given $j$ (to be specified below) choose $i \leqq m(j)$ so that $d(x,x_i)<\delta_j$ and $h(x_i)<h(x)+3^{-j}$. For $k \geqq j$ we have

$$\begin{aligned} g_k(x)-g_k(x_i) &\geqq (1/2)[g_k(x')+g_k(x'')]-a(x_i)^2 \\ &\geqq (1/2)[a(x')^2+a(x'')^2]-a(x_i)^2 \\ &= \varepsilon/4+a(x)^2-a(x_i)^2 . \end{aligned}$$

Since $a \in S_1$ and $d(x,x_i)<\delta_j$, we have $|a(x)-a(x_i)|<3^{-j}$ while $a(x)+a(x_j) \leqq 2$, so that $a(x)^2-a(x_i)^2>-2\cdot 3^{-j}$.

For $k<j$ we have $|g_k(x)-g_k(x_i)|<3^{-j}$ and $h(x)-h(x_i)>-3^{-j}$; also, $|f(x)-f(x_i)|<3^{-j}$ if $f \in \bigcup_{i=1}^{j} S_i$. Consequently,

$$\begin{aligned} (f+\lambda g-h)(x_i)-(f+\lambda g-h)(x) &> -2\cdot 3^{-j}+\lambda \sum 2^{-j}[g_k(x)-g_k(x_i)] \\ &> -2\cdot 3^{-j}-\lambda\left(\sum_{k=1}^{j-1} 2^{-k}\right)\cdot 3^{-j} \\ &\quad +\lambda \sum_{k=j}^{\infty} 2^{-k}(\varepsilon/4-2\cdot 3^{-j}) \\ &> 2^{-j+1}\lambda(\varepsilon/4-2\cdot 3^{-j})-(2+\lambda)3^{-j}, \end{aligned}$$

which is easily seen to be positive if $j$ is chosen to be sufficiently large.

For a similar result in a more physical setting, see [7].

The next result establishes a converse to the foregoing theorem, showing that the stronger differentiability conclusion *requires* property (D).

3.3 THEOREM. *Suppose that $h \in \mathscr{H}$ fails to satisfy property* (D); *then for some $n \geqq 1$ the subset of $A(K)^n$*

$$\{\boldsymbol{g} \in A(K)^n : \text{for some } \boldsymbol{\lambda} \in \mathbf{R}^n \setminus \{0\}, \partial P_h(\boldsymbol{\lambda} \cdot \boldsymbol{g}) \nsubseteq \operatorname{ext} K\}$$

*has nonempty interior. When h is affine, the condition $\partial P_h(\boldsymbol{\lambda} \cdot \boldsymbol{g}) \subseteq \operatorname{ext} K$ means that $P \equiv P_h$ is not Gateaux differentiable at $\boldsymbol{\lambda} \cdot \boldsymbol{g}$.*

PROOF. By hypothesis, there exists $r \in \mathbf{R}$ such that $\operatorname{ext} K \cap \{x \in K : h(x) < r\}$ is not dense in $\{x \in K: h(x) < r\}$ (and the latter is nonempty). Thus, there exist $x_0 \in K, \delta > 0, 0 < \varepsilon = r - h(x_0)$ and $f_1, f_2, \ldots, f_m \in A(K)$ with $\|f_i\| < 1$ such that if $|f_i(x) - f_i(x_0)| < \delta$ for $i = 1, 2, \ldots, m$, then either $x \notin \operatorname{ext} K$ or $h(x) > h(x_0) + \varepsilon$. By duality,

$$h(x_0) = \sup\{f(x_0) - P(f): f \in A(K)\} ,$$

so there exists $f_{m+1} \in A(K)$ such that $f_{m+1}(x_0) - P(f_{m+1}) > h(x_0) - \varepsilon/4$. Suppose first that $f_1, f_2, \ldots, f_{m+1}$ do not span $A(K)$.

Choose $f_{m+2} \in A(K)$ with $\|f_{m+2}\| < \varepsilon/8$ and $f_{m+2} \notin \operatorname{span}\{f_1, \ldots, f_m, f_{m+1}\}$. Now, let $U$ be the neighborhood of $(f_1, \ldots, f_{m+2})$ in $A(K)^{m+2}$ consisting of all $\boldsymbol{g} = (g_1, \ldots, g_{m+2})$ satisfying

$$\|f_i - g_i\| < \delta/3 \text{ and } \|g_i\| < 1 \quad \text{for } i = 1, 2, \ldots, m ,$$

$$g_{m+1}(x_0) - P(g_{m+1}) > h(x_0) - \varepsilon/4$$

$$\|g_{m+2}\| < \varepsilon/8 \text{ and } g_{m+2} \notin \operatorname{span}\{g_1, \ldots, g_{m+1}\} .$$

We claim that for each $\boldsymbol{g} \in U$ there exists $\boldsymbol{\lambda} \in \mathbf{R}^{m+2} \setminus \{0\}$ with $\partial P(\boldsymbol{\lambda} \cdot \boldsymbol{g}) \nsubseteq \operatorname{ext} K$. Given $\boldsymbol{g} \in U$, then, let

$$\gamma = \min\{\delta/3, \varepsilon(4\|g_{m+1}\|)^{-1}\} .$$

By the version of the Bishop–Phelps theorem given in [6, Corollary V.1.2], if $\mathscr{F} = \operatorname{span}\{g_1, \ldots, g_{m+1}\}$ (and recalling that the subdifferentials—that is, the support functionals—of $P$ are contained in $K$) we can find $x_1 \in K$ and $g \in \mathscr{F}$ such that $x_1 \in \partial P(g + g_{m+2})$,

$$\|g - g_{m+1}\| \leqq \gamma^{-1}[P(g_{m+2} + g_{m+1}) - (g_{m+2} + g_{m+1})(x_0) + h(x_0)]$$

and

$$|g'(x_0) - g'(x_1)| \leqq \gamma\|g'\| \quad \text{for all } g' \in \mathscr{F} .$$

Since $P$ is Lipschitz $-1$ and $\|g_{m+2}\| < \varepsilon/8$, we know that $P(g_{m+2} + g_{m+1}) < \varepsilon/8 + P(g_{m+1})$ and hence

$$P(g_{m+2}+g_{m+1})-(g_{m+2}+g_{m+1})(x_0)+h(x_0) <$$
$$< \varepsilon/8+P(g_{m+1})-g_{m+1}(x_0)+h(x_0)-g_{m+2}(x_0) < \varepsilon/2 .$$

Also, $|g_i(x_0)-g_i(x_1)|<\gamma\leqq\delta/3$ and $\|g_i-f_i\|<\delta/3$ so that $|f_i(x_0)-f_i(x_1)|<\delta$ for $i=1,2,\ldots,m$. By hypothesis, then, either $x_1 \notin \operatorname{ext} K$ or $h(x_1)>h(x_0)+\varepsilon$. The latter cannot hold, however, since $x_1 \in \partial P(g+g_{m+2})$ implies that $(g+g_{m+2}-h)(x_0)\leqq(g+g_{m+2}-h)(x_1)$ so that

$$\begin{aligned} h(x_1)-h(x_0) &\leqq g_{m+2}(x_1)-g_{m+2}(x_0)+g(x_1)-g(x_0) \\ &< \varepsilon/4+\gamma\|g\| \leqq \varepsilon/4+\gamma(\|g-g_{m+1}\|+\|g_{m+1}\|) \\ &< \varepsilon/4+\gamma(\gamma^{-1}\varepsilon/2+\|g_{m+1}\|) \leqq \varepsilon . \end{aligned}$$

Thus, $\partial P(g+g_{m+2})\nsubseteq \operatorname{ext} K$ and since $g$ and $g_{m+2}$ are linearly independent, we have $g+g_{m+2}=\boldsymbol{\lambda}\cdot\boldsymbol{g}$ for some $\boldsymbol{\lambda}\neq 0$. This completes the proof for the case where $f_1,\ldots,f_{m+1}$ do not span $A(K)$. If they do span $A(K)$, then an argument similar to that above (using 0 instead of $g_{m+2}$) shows that for $(g_1,\ldots,g_{m+1})$ in some neighborhood of $(f_1,\ldots,f_{m+1})$, there exists $g \in \operatorname{span}\{g_1,\ldots,g_{m+1}\}$ such that $\partial P(g)\nsubseteq \operatorname{ext} K$. To avoid the case $g=0$, choose a (possibly smaller) neighborhood of $(f_1,\ldots,f_{m+1})$ in which each $(g_1,\ldots,g_{m+1})$ span $A(K)$. Then for some constant $c$ we have $g+c=\boldsymbol{\lambda}\cdot\boldsymbol{g}$ with $\boldsymbol{\lambda}\neq 0$ and $\partial P(g+c) =\partial P(g)\nsubseteq \operatorname{ext} K$.

The situation concerning *Fréchet* differentiablity is much less positive. In fact, it is known [4] that the physical pressure is *nowhere* Fréchet differentiable. This result (Corollary 3.5) will drop out of the theorem which follows, which shows that — when $K$ is a simplex and $h$ is affine — Fréchet differentiability of $P_h$ at some point forces $P_h$ to be *affine* in a neighborhood of the point. This rather strong property has a number of equivalent formulations, one of which (assertion (v) below) negates property (D). It is somewhat remarkable that mere Gateaux differentiability of $P_h$ in an open set implies that $P_h$ is affine in that set; this is evident from the proof that (iv) implies (iii).

3.4. THEOREM. *Suppose that $h \in \mathscr{H}$ is affine and that $K$ is a simplex. The following assertions are equivalent.*

(i) *$P_h$ is Fréchet differentiable at some point of $A(K)$.*
(ii) *There exists $x \in \operatorname{ext} K$ such that $\partial P_h(f)=\{x\}$ for all $f$ in some nonempty open subset of $A(K)$.*
(iii) *$P_h$ is affine in some nonempty open subset of $A(K)$.*
(iv) *$P_h$ is Gateaux differentiable in some nonempty open subset of $A(K)$.*
(v) *There exists $x \in \operatorname{ext} K$ and a neighborhood $V$ of $x$ such that $h(x) <\inf\{h(y)\colon y \in V\cap \operatorname{ext} K,\ y\neq x\}$.*

PROOF. (i) implies (ii). Suppose that $P_h$ is Fréchet differentiable at $f \in A(K)$, with $\partial P_h(f)=\{x\}$, say, where $x \in \operatorname{ext} K$. Recall that for distinct extreme points $x, y$ of a simplex, one has $\|x-y\|=2$. (This is a corollary of Edwards' theorem.) It follows from this, from the fact that every subdifferential of $P_h$ contains an extreme point of $K$, and from the norm-norm upper semicontinuity of $\partial P_h$ at $f$ that one can find an open neighborhood $U$ of $f$ such that $\partial P_h(g)=\{x\}$ for each $g \in U$.

(ii) implies (iii). It is clear from (ii) that

$$P_h(g) = g(x)-h(x) \quad \text{for all } g \in U ,$$

so that $P_h$ is affine in $U$.

(iii) implies (i). If the continuous function $P_h$ is affine in some open set $U$, then $P_h$ will have the same subdifferential at each point of $U$; in particular, $\partial P_h$ will be norm-norm continuous at each such point and hence $P_h$ will be Fréchet differentiable in $U$.

It is obvious that (iii) implies (iv). To see that (iv) implies (iii), suppose that for some $\varepsilon>0$, the function $P_h$ is Gateaux differentiable in the $2\varepsilon$-neighborhood of the function $f \in A(K)$ but not affine in any neighborhood of $f$. Let $\partial P_h(f)=\{x\} \subseteqq \operatorname{ext} K$ and let $m=P_h(f)=f(x)-h(x)$. Since $P_h$ is not affine in any neighborhood of $f$ we can choose $f_1 \in A(K)$ and $y \in \operatorname{ext} \partial P_h(f_1)$ such that $\|f-f_1\|<\varepsilon/2$ and $y \neq x$. We have

$$f(y)-h(y) = (f-f_1)(y)+P_h(f_1) > P_h(f)-\varepsilon = m-\varepsilon .$$

To show that $P_h$ cannot be Gateaux differentiable in a $2\varepsilon$-neighborhood of $f$, it suffices to find $g \in A(K)$ such that $\|g\| \leqq \varepsilon$ and $\partial P_h(f+g)$ contains the two points $x$ and $y$, that is, $f+g-h$ attains its maximum on the face $F=[x,y]$. In order to use Edwards' separation theorem, we define $h_1, h_2$ on $K$ as follows: Let

$$h_1 = \begin{cases} m-(f-h) & \text{on } F \\ 0 & \text{elsewhere} \end{cases} \qquad h_2 = \min\,[\varepsilon, m-(f-h)] .$$

It is straightforward to verify that $h_1$ is convex and upper semicontinuous, that $h_2$ is concave and lower semicontinuous and that $h_1 \leqq h_2$. By Edwards' theorem, there exists $g \in A(K)$ such that $h_1 \leqq g \leqq h_2$. Since $h_1=h_2=m-(g-h)$ on $F$, the same is true of $g$, so $f+g-h=m$ on $F$ and (since $g \leqq h_2$), $f+g-h \leqq m$ on $K$. Finally, $0 \leqq g \leqq \varepsilon$, which completes this part of the proof.

(iv) implies (v). Suppose that (v) fails; we will show that (iv) cannot hold, by showing that for any $f \in A(K)$ and any $\varepsilon>0$ there exists $g \in A(K)$ such that $\|g\| \leqq \varepsilon$ and $\partial P_h(f+g)$ contains at least two extreme points. To this end, choose

$x \in \operatorname{ext} \partial P_h(f)$ and choose a neighborhood $V$ of $x$ in which $f$ satisfies $f > f(x) - \varepsilon/2$. Since (v) fails, there must exist $y \in V \cap \operatorname{ext} K$, $y \neq x$, such that $h(y) \leqq h(x) + \varepsilon/2$. Let $m = P_h(f) = f(x) - h(x)$; we have

$$h(y) \leqq f(x) - m + \varepsilon/2 < f(y) - m + \varepsilon$$

with (obviously) the same inequality at $x$. The construction used in proving that (iv) implies (iii) yields the desired conclusion.

(v) implies (ii). It is well known that extreme points have neighborhood bases consisting of "slices", so that we can find $f \in A(K)$ and $\alpha > 0$ such that

$$x \in S(f, \alpha) = \{y \in K : f(y) > \sup f(K) - \alpha\} \subseteqq V.$$

Let $M = \sup(-h)(K)$; by multiplying $f$ and $\alpha$ by a positive scalar we can assume that $[\alpha + f(x) - \sup f(K)] - M > 0$. Choose $\delta > 0$ so that $2\delta$ is smaller than this number and also smaller than

$$\inf\{h(y) : y \in V \cap \operatorname{ext} K, y \neq x\} - h(x).$$

Let

$$U = \{g \in A(K) : g(x) > f(x) - \delta, g < \min\{f(x), f\} + \delta\}.$$

This set is clearly open, and it can be shown to be nonempty by a standard strict separation argument, using the fact that $x$ is extreme and that the convex hull of the graph of $\min\{f(x), f\} + \delta$ is compact (hence closed) in $K \times \mathbb{R}$. We claim $\partial P_h(g) = \{x\}$ for each $g \in U$. Indeed, if $g \in U$ and if $y \in \operatorname{ext} \partial P_h(g)$, then

$$f(y) > g(y) - \delta \geqq g(x) - h(x) + h(y) - \delta > f(x) - M - 2\delta > \sup f(K) - \alpha,$$

so $y \in V \cap \operatorname{ext} K$. Moreover,

$$h(y) \leqq h(x) - g(x) + g(y) < h(x) - [f(x) - \delta] + [f(x) + \delta] = h(x) + 2\delta$$

so our second restriction on $\delta$ shows that $y = x$.

There are several easy consequences of the foregoing theorem. For instance, property (D) clearly negates assertion (v), hence assertion (i), so we obtain the following corollary.

3.5. COROLLARY. *If $K$ is a simplex and if $h \in \mathscr{H}$ is affine and satisfies property* (D), *then $P_h$ is nowhere Fréchet differentiable.*

It would be satisfying if the converse to this result were true, or if the existence of a nowhere Fréchet differentiable $P_h$ were to imply that $K$ is the Poulsen simplex. This, however, is not the case; equivalently, there can exist a $K$ which is not a Poulsen simplex and an $h \in \mathscr{H}$ for which assertion (v) of Theorem 3.4 fails, as shown in the following example.

3.6. EXAMPLE. Let $K_1, K_2$ be copies of the Poulsen simplex contained, say, in locally convex spaces $E_1, E_2$, respectively. By taking the convex hull of $(K_1 \times \{0\}) \cup (\{0\} \times K_2)$ in $E_1 \times E_2$ (and making the obvious identifications) we can obtain a simplex $K$ such that $K = \text{co}(K_1 \cup K_2)$ and each $x$ in $K$ has a unique representation

$$x = \lambda x_1 + (1-\lambda)x_2, \quad 0 \leqq \lambda \leqq 1, \; x_i \in K_i, \; i=1,2 .$$

In particular, the constant $\lambda$ is an affine continuous function of $x$, so if we let $h(x) = -\lambda$, then $h \in \mathscr{H}$ and equals $-1$ on $K_1$, 0 on $K_2$, respectively. It is easily checked that, since $\text{ext}\, K = \text{ext}\, K_1 \cup \text{ext}\, K_2$, assertion 3.4(v) fails for $h$. On the other hand, it is obvious that $K$ is not a Poulsen simplex; in particular, of course, property (D) also fails for $h$.

One *can* draw a density conclusion about $\text{ext}\, K$ from the existence of a nowhere Fréchet differentiable $P_h$ (where $h \in \mathscr{H}$ is affine), namely, that $\text{ext}\, K$ is dense in itself.

3.7. COROLLARY. *Suppose that $K$ is compact and convex (not necessarily a simplex) and that $h \in \mathscr{H}$ is affine. If* $\text{ext}\, K$ *contains a relative isolated point, then $P_h$ is Fréchet differentiable on some nonempty open set.*

PROOF. If $x \in \text{ext}\, K$ and if there exists a neighborhood $V$ of $x$ which misses $\text{ext}\, K \setminus \{x\}$, then the infimum in assertion 3.4(v) equals $+\infty$, so (v) is satisfied. The proof that (v) implies (ii) does not use the fact that $K$ is a simplex, and from (ii) we obtain the desired differentiability conclusion.

Note that if $K$ is a simplex, $h \in \mathscr{H}$ is affine and continuous and $\text{ext}\, K$ has no isolated points, then 3.4(v) fails for each $x \in \text{ext}\, K$, hence $P_h$ is nowhere Fréchet differentiable, that is, the converse of Corollary 3.7 is valid under these additional hypotheses.

It is not difficult to exhibit lower semicontinuous affine functions on the Poulsen (or any) metrizable simplex $K$ for which $P_h$ has points of Fréchet differentiability, hence which satisfy the various equivalent conditions of Theorem 3.4.

3.8. EXAMPLE. Fix $x_0 \in \operatorname{ext} K$ and for each $x \in K$ let $\mu_x$ denote the unique measure on $\operatorname{ext} K$ which represents $x$. Define

$$h(x) = -\mu_x(\{x_0\}) .$$

The map $x \to \mu_x$ is affine, so $h$ is affine. To see that $h$ is lower semicontinuous, we show that $-h$ is upper semicontinuous. To this end, let $A = \operatorname{ext} K \setminus \{x_0\}$ and suppose that $\{x_n\} \subseteqq K$, with $x_n \to x$. For each $f \in A(K)$ we have $\mu_{x_n}(f) = f(x_n) \to f(x) = \mu_x(f)$. Thus

$$\int_A f\, d\mu_{x_n} - f(x_0)h(x_n) \to \int_A f\, d\mu_x - f(x)h(x), \qquad f \in A(K) .$$

Given $\varepsilon > 0$, choose a compact subset $J$ of $A$ such that $\mu_x(A \setminus J) < \varepsilon$. By Edwards' separation theorem there exists $f \in A(K)$ such that $f = 0$ on $J$, $f(x_0) = 1$ and $0 \leqq f \leqq 1$. Thus,

$$-h(x_n) \leqq \int_A f\, d\mu_{x_n} - h(x_n) \to \int_A f\, d\mu_x - f(x)h(x) \leqq \mu_x(A \setminus J) - h(x) < -h(x) + \varepsilon .$$

We conclude that $\limsup -h(x_n) \leqq -h(x) + \varepsilon$ for all $\varepsilon > 0$, therefore $\limsup -h(x_n) \leqq -h(x)$ and hence $-h$ is upper semicontinuous. Now, since the supremum which defines $P_h$ is attained at an extreme point, we can write

$$P_h(f) = \sup\{f(x) + \mu_x(\{x_0\}) : x \in \operatorname{ext} K\} .$$

On $\operatorname{ext} K$, we know that $\mu_x(\{x_0\})$ is the characteristic function of $\{x_0,\}$, so this becomes

$$P_h(f) = \max[\sup\{f(A)\}, f(x_0) + 1] .$$

Consider, finally, the nonempty open subset $U$ of $A(K)$ defined by

$$U = \{f \in A(K) : \sup f(A) < f(x_0) + 1\} .$$

(This is nonempty since, for instance, it contains the constant functions.) Clearly, if $f \in U$, then $P_h(f) = f(x_0) + 1$, and this is obviously Fréchet differentiable on $U$.

*If $h$ is affine and lower semicontinuous on the simplex $K$, then $P_h$ cannot be everywhere Gateaux differentiable.* This assertion is a consequence of [6, Theorem V.2.2(a)] (after a translation into the present language), and is also an easy consequence of "(iv) implies (ii)" of Theorem 3.4. [Such a $P_h$ would necessarily have the form $f \to f(x) - h(x)$ for some $x \in \operatorname{ext} K$. By Edwards' theorem one could find $f \in A(K)$, $\|f\| > -h(x)$, such that $f \geqq 0$ and $f(x) = 0$, which would violate property (P3).] The above fact is also a consequence of the following easy corollary of Edwards' theorem, once one observes that the restriction of an affine function to the convex hull $F$ of finitely many extreme points is necessarily continuous. Since metrizability holds in the physical situation, the second assertion in this proposition gives a new characterization, in terms of continuity of the mean entropy, of those sets of states which can form the set of equilibrium states for some interaction.

3.9. PROPOSITION. *Suppose that $h \in \mathscr{H}$ is affine and that $K$ is a simplex. If $F$ is a nonempty face of $K$ and if $h|_F$ is continuous, then there exists $f \in A(K)$ such that $F \subseteqq \partial P_h(f)$. Suppose, further, that $K$ is metrizable; then the subdifferential sets $\partial P_h(f)$ of $P_h$ are precisely those closed faces $F$ of $K$ for which $h|_F$ is continuous.*

PROOF. Suppose that $F$ is a face of $K$ and that $h|_F$ is continuous. By a corollary to Edwards' theorem there exists $f \in A(K)$ which extends $h|_F$ and which satisfies $f \leqq h$ on $K$. Thus, $f-h \leqq 0$ on $K$ and $f-h=0$ on $F$, so $F \subseteqq \partial P_h(f)$. Suppose, now, that $K$ is metrizable and that $h$ is continuous on the closed face $F$. Choose $f$ as above. It is known (see [1, p. 121]) that there exists $g$ in $A(K)$ which vanishes on $F$ and is strictly negative on $K \setminus F$; it follows that $F = \partial P_h(f+g)$. Conversely, if $F = \partial P_h(f)$ for some $f \in A(K)$, then $F$ is a closed face of $K$ and $P_h(f) = f(x) - h(x)$ for $x \in F$, which implirs that $h|_F$ is continuous.

If $h$ is *strictly convex* and lower semicontinuous, then for each $f \in A(K)$ the function $f-h$ attains its maximum at a single point, that is, $P_h$ is *everywhere* Gateaux differentiable. Such strictly convex functions exist on any metrizable $K$; more generally, they exist if $K$ is affinely homeomorphic to a weakly compact convex subset of some Banach space (that is, if $K$ is an Eberlein compact). Such a space can be assumed to be weakly compactly generated, hence admits a strictly convex norm. The square of such a norm suffices, since it is weakly lower semicontinuous and strictly convex.

## 4. Extreme points of $\mathscr{P}$.

In this section we will identify some of the extreme points of $\mathscr{P}$, showing, in particular, that the physical pressure is such a point. We will also show that $\mathscr{P}$ is the closed convex hull of its extreme points, in the topology of pointwise convergence. This would be trivial if $\mathscr{P}$ were pointwise compact, but it is not bounded: If $P \in \mathscr{P}$, then $P + c \in \mathscr{P}$ for each constant $c > 0$, hence $\sup\{P(0): P \in \mathscr{P}\} = +\infty$. It *is* pointwise closed; since

$$|P(f) - P(g)| \leqq \|f-g\|, \quad f, g \in A(K),\ P \in \mathscr{P},$$

pointwise limits of elements of $\mathscr{P}$ are continuous. Such limits are also convex and satisfy properties (P1), (P2) and (P3). Moreover, $\mathscr{P}$ is the union of "large" compact convex subsets. To see this, define, for $t \geqq 0$,

$$\mathscr{P}_t = \{P \in \mathscr{P} : P(0) \leqq t\} .$$

Since $P \to P(0)$ is a pointwise continuous linear functional, each $\mathscr{P}_t$ is closed, convex and has convex complement in $\mathscr{P}$. Furthermore, if $P_h \in \mathscr{P}_t$ and $f \in A(K)$, then

$$P_0(f) = \sup f \leqq P_h(f) = \sup (f-h) \leqq P_0(f)+P_h(0) \leqq P_0(f)+t ,$$

which shows that

$$\mathscr{P}_t \subseteqq \prod \{[P_0(f), P_0(f)+t] : f \in A(K)\} ,$$

hence $\mathscr{P}_t$ is pointwise compact. (Thus, each $\mathscr{P}_t$ is a *cap* of $\mathscr{P}$ [8].)

4.1. Proposition. *The convex set $\mathscr{P}$ is well-capped, that is, $\mathscr{P} = \bigcup \{\mathscr{P}_t : t \geqq 0\}$.*

If $P \in \mathscr{P}$, then $P \in \mathscr{P}_{P(0)}$, so the proof of this proposition is immediate. It is the first step towards producing extreme points in $\mathscr{P}$: Each $\mathscr{P}_t$ is the closed convex hull of its extreme points, and it follows from elementary two-dimensional arguments that an extreme point $P$ of $\mathscr{P}_t$ is either an extreme point of $\mathscr{P}$ or "nearly" so; that is, it satisfies $P(0)=t$ and lies in the relative interior of a segment which is itself an extremal subset of $\mathscr{P}$. The segment could itself be part of a *ray* in $\mathscr{P}$, that is, a set of the form

$$\{P_h + \lambda(P_{h'} - P_h) : \lambda \geqq 0\}, \qquad P_h, P_{h'} \in \mathscr{P} .$$

We know that if $P_h \in \mathscr{P}$ and $c>0$, then $P_h + c = P_{h-c} = P_h + c(P_{h-1} - P_h) \in \mathscr{P}$, and it will be useful for us to know that these are the *only* rays contained in $\mathscr{P}$.

4.2. Proposition. *If $P_h, P_{h'} \in \mathscr{P}$ and if the ray*

$$P_h + \mathbf{R}^+(P_{h'} - P_h) \subseteqq \mathscr{P} ,$$

*then $h' = h - c$ for some constant $c>0$.*

Proof. We first show that if $P_0 + \mathbf{R}^+(P_h - P_0) \subseteqq \mathscr{P}$, then $h$ is constant. To this end note that for each $\lambda>0$ convexity of $P = P_0 + \lambda(P_h - P_0)$ implies that for all $f, g \in A(K)$

$$P(\tfrac{1}{2}f + \tfrac{1}{2}g) \leqq \tfrac{1}{2}P(f) + \tfrac{1}{2}P(g) ,$$

so that

$$(\lambda-1)[\tfrac{1}{2}P_0(f) + \tfrac{1}{2}P_0(g) - P_0(\tfrac{1}{2}f + \tfrac{1}{2}g)] \leqq \lambda[\tfrac{1}{2}P_h(f) + \tfrac{1}{2}P_h(g) - P_h(\tfrac{1}{2}f + \tfrac{1}{2}g)] .$$

Dividing both sides by $\lambda$, multiplying by 2 and letting $\lambda \to +\infty$ gives us

$$\sup f + \sup g - \sup (f+g) \leqq \sup (f-h) + \sup (g-h) - \sup (f+g-2h) .$$

If $h$ were not constant, we could apply the separation theorem to choose

$f \in A(K)$ such that $f \leqq h$ and $\sup f > \inf h$. Let $g = -f$; we would then have (since $f - h \leqq 0$)

$$\begin{aligned} \sup f - \inf f &\leqq \sup(f-h) + \sup(-f-h) + \sup(-2h) \\ &\leqq \sup(-f) + \sup(-h) - 2\sup(-h) \\ &= -\inf f + \inf h, \end{aligned}$$

contradicting our choice of $f$.

Suppose, next, that $h, h' \in \mathscr{H}$ and that

$$P_h + \mathrm{R}^+(P_{h'} - P_h) \subseteq \mathscr{P}.$$

Since $\mathscr{P}$ is closed and convex, the parallel ray

$$P_0 + \mathrm{R}^+(P_{h'} - P_h)$$

is also in $\mathscr{P}$. By what we have just proved, each point on this ray is of the form $P_{-c}$ for some constant $c \geqq 0$. In particular,

$$P_0 + P_{h'} - P_h = P_{-c}$$

for some $c \geqq 0$. This is equivalent to

$$\sup f + \sup(f-h') - \sup(f-h) = \sup(f+c) \quad \text{for all } f \in A(K),$$

that is,

$$\sup(f-h') - \sup(f-h) = c, \quad f \in A(K). \tag{1}$$

We will show that this implies that $h = h' + c$. Indeed, suppose that $h(x) > h'(x) + c$ for some $x \in K$. By the separation theorem we could choose $f \in A(K)$ such that $f \leqq h$ and $f(x) > h'(x) + c$. It would follow that

$$\sup(f-h') - \sup(f-h) \geqq f(x) - h'(x) > c,$$

contradicting (1). Another application of the separation theorem also shows that there is no point $x$ with $h(x) - c < h'(x)$, so the proof is complete.

It seems very difficult to decide directly whether an element of $\mathscr{P}$ is an extreme point, but the difficulties become fewer if one formulates the notion in terms of the members of $\mathscr{H}$. This requires the introduction of a modified version of the "inf-convolution" originated by Moreau [8].

DEFINITION. If $h_1, h_2 \in \mathscr{H}$ define $h_1 \square h_2$ on $K$ by setting

$$(h_1 \square h_2)(x) = \inf\{h_1(x_1) + h_2(x_2) : x_1, x_2 \in K, \ x = \tfrac{1}{2}(x_1 + x_2)\}$$

for each $x \in K$.

It is straightforward to prove that $h_1 \square h_2$ is convex, bounded, non-positive and (using the compactness of $K$) that it is lower semicontinuous; that is, $h_1 \square h_2 \in \mathscr{H}$.

DEFINITION. Say that a function $h \in \mathscr{H}$ is "*extreme*" *in* $\mathscr{H}$ provided $h=h_1=h_2$ whenever $h_1, h_2 \in \mathscr{H}$ and $h=\frac{1}{2}h_1 \square \frac{1}{2}h_2$.

4.3. LEMMA. *An element $P_h$ of $\mathscr{P}$ is an extreme point of $\mathscr{P}$ if and only if $h$ is "extreme" in $\mathscr{H}$.*

PROOF. We first show that if $h=\frac{1}{2}h_1 \square \frac{1}{2}h_2$, then $P_h=\frac{1}{2}P_{h_1}+\frac{1}{2}P_{h_2}$. Indeed, if $f \in A(K)$, then

$$\begin{aligned} P_h(f) &= \sup\{f(x)-h(x) : x \in K\} \\ &= \sup\{f(x)-\tfrac{1}{2}\inf\{h_1(x_1)+h_2(x_2) : x_1, x_2 \in K,\ x=\tfrac{1}{2}(x_1+x_2)\} : x \in K\} \\ &= \sup\{\tfrac{1}{2}[f(x_1)-h_1(x_1)]+\tfrac{1}{2}[f(x_2)-h_2(x_2)] : x_1, x_2 \in K\} \\ &= \tfrac{1}{2}P_{h_1}(f)+\tfrac{1}{2}P_{h_2}(f)\,. \end{aligned}$$

Conversely, if $P_h=\frac{1}{2}P_{h_1}+\frac{1}{2}P_{h_2}$, let $g=\frac{1}{2}h_1 \square \frac{1}{2}h_2$. By what we have just proved, $P_g=\frac{1}{2}P_{h_1}+\frac{1}{2}P_{h_2}=P_h$; since the correspondence $h \leftrightarrow P_h$ is one-one, $g=h$. The assertions of the Lemma are now immediate.

4.4. COROLLARY. *If $P_h$ is an extreme point of $\mathscr{P}$, then so is $P_{\lambda h}$, for each $\lambda \geqq 0$.*

PROOF. It is clear that $P_0$ is an extreme point of $\mathscr{P}$. If $\lambda > 0$ and $\lambda h=\frac{1}{2}h_1 \square \frac{1}{2}h_2$, then $h=\frac{1}{2}\lambda^{-1}h_1 \square \frac{1}{2}\lambda^{-1}h_2$, hence $h=\lambda^{-1}h_1=\lambda^{-1}h_2$.

Note the curious fact described by this corollary: Since the map $h \to P_h$ is continuous onto $\mathscr{P}$ with its supremum metric, every extreme point in $\mathscr{P}$ is joined to $P_0$ by a (uniformly continuous) arc of extreme points of $\mathscr{P}$. If $h \in \mathscr{H}$ is affine and continuous, with $\sup h=0>\inf h$, then it is elementary to verify that the image of the one-dimensional ray $\{\lambda h\colon \lambda \geqq 0\}m$ under the map $\lambda h \to P_{\lambda h}$ is an affinely independent set, hence spans an infinite-dimensional subset of $\mathscr{P}$. Such an $h$ exists whenever $\dim K \geqq 1$.

4.5. THEOREM. *For any compact convex $K$, the set $\mathscr{P}$ is the pointwise closed convex hull of* ext $\mathscr{P}$, *its set of extreme points.*

PROOF. Let $L$ be a pointwise continuous linear functional, so that there exist $a_1, \ldots, a_n$ in $\mathbf{R}$ and $f_1, \ldots, f_n$ in $A(K)$ such that

$$L(P) = \sum_{i=1}^{n} a_i P(f_i), \qquad P \in \mathscr{P} .$$

Given $\alpha \in \mathrm{R}$ it suffices to show that if $S = \{P \in \mathscr{P} : L(P) > \alpha\}$ is non-empty, then $S \cap \operatorname{ext} P \neq \emptyset$. Since $\mathscr{P}$ is the union of the sets $\mathscr{P}_t$, $t > 0$, there exists $t > 0$ such that $S \cap P_t \neq \emptyset$. Since $\mathscr{P}_t$ is compact and convex, there exists $P_{h'} \in S \cap \operatorname{ext} \mathscr{P}_t$. We are obviously finished if $P_{h'} \in \operatorname{ext} \mathscr{P}$, so assume otherwise. It follows that $P_{h'}$ lies in the relative interior of a segment $[P_h, P_{h''}] \subseteqq \mathscr{P}$ with $P_h, P_{h''} \in \operatorname{ext} \mathscr{P}$, or in the relative interior of a ray $P_h + \mathrm{R}^+(P_{h'} - P_h)$ in $\mathscr{P}$, where $P_h \in \operatorname{ext} \mathscr{P}$. In the first of these cases, one of $P_h, P_{h''}$ would lie in $S$ and we would be finished, so we assume that the second case holds and that $P_h \notin S$. This implies that $\sum a_i > 0$. Indeed, we have $L(P_h) \leqq \alpha < L(P_{h'})$. Moreover, by Proposition 4.2, we know that $h' = h - c$ for some constant $c > 0$. Thus,

$$\begin{aligned} L(P_h) &= \sum a_i \sup (f_i - h) \leqq \alpha < L(P_{h'}) = \sum a_i \sup (f_i - h + c) \\ &= \sum a_i (\sup f_i - h) + c \sum a_i , \end{aligned}$$

which proves our contention.

In order to complete the proof we first produce *some* extreme point $P_g$ of $\mathscr{P}$ with $g \neq 0$. The element $P_h$ (above) would do if $h \neq 0$, so assume that $h = 0$. Using the fact that $\mathscr{P}_t$ is obviously not one-dimensional (we ignore the trivial case when $K$ is a single point), we can choose $P_{h_1} \in \operatorname{ext} \mathscr{P}_t$ with $h_1 \neq h', 0$. Using the same geometric arguments we applied above to $P_{h'}$, we can either find $P_g \in \operatorname{ext} \mathscr{P}$ with $g \neq 0$ or $P_{h_1}$ would lie on the ray in $\mathscr{P}$ from $P_0$ through $P_{h_1}$. But by Proposition 4.2 there is only one such ray, the one through $P_{h'}$. Thus, an extreme point $P_g$ exists with $\sup(-g) > 0$. By Corollary 4.4, we have $P_{\lambda g} \in \operatorname{ext} \mathscr{P}$ for each $\lambda > 0$; we will show that $L(P_{\lambda g}) > \alpha$ for sufficiently large $\lambda$. To see this, note that for each $i$

$$-\sup(-f_i) + \lambda \sup(-g) \leqq \sup(f_i - \lambda g) \leqq \sup f_i + \lambda \sup(-g) .$$

Consequently,

$$\begin{aligned} L(P_{\lambda g}) &= \sum a_i \sup (f_i - \lambda g) \\ &\geqq \sum_{a_i > 0} a_i[-\sup(-f_i) + \lambda \sup(-g)] + \sum_{a_i < 0} a_i[\sup f_i + \lambda \sup(-g)] \\ &= \lambda (\sum a_i) \sup(-g) + \sigma , \end{aligned}$$

where $\sigma$ is a sum of terms not containing $\lambda$. Since $\sum a_i > 0$ and $\sup(-g) > 0$, the desired conclusion follows.

We show next that an important subclass of $\mathscr{H}$ (which contains the negative of the mean entropy) is made up of "extreme" points of $\mathscr{H}$, at least when $K$ is metrizable.

4.6. THEOREM. *If $K$ is metrizable and if $h \in \mathscr{H}$ is affine, with* $\sup h = 0$, *then $h$ is "extreme" in $\mathscr{H}$, so $P_h$ is an extreme point of $\mathscr{P}$.*

PROOF. Note first that if $h \in \mathscr{H}$ and $\sup h < 0$, then $P_h$ lies in the relative interior of the ray $P_{h'} + \mathbf{R}^+(P_h - P_{h'})$ in $\mathscr{P}$, where $h' = h - \sup h$, hence $P_h$ cannot be extreme in $\mathscr{P}$. Thus, $\sup h = 0$ is a necessary condition for $h$ to be "extreme". Suppose then, that $h \in \mathscr{H}$ is affine and that $h_1, h_2 \in \mathscr{H}$ with $h = \frac{1}{2}h_1 \,\square\, \frac{1}{2}h_2$. Let

$$J = \{x \in K : 2h(x) = h_1(x) + h_2(x)\};$$

it is immediate from the definition of the inf-convolution that $\operatorname{ext} K \subseteqq J$. We will show first that we must have $h_1 = h = h_2$ on $J$, then we will use metrizability to prove that $J = K$. Suppose, then, that $x, y \in J$. Since $h$ is affine we have

$$\text{(1)} \qquad \tfrac{1}{2}h(x) + \tfrac{1}{2}h(y) = h(\tfrac{1}{2}x + \tfrac{1}{2}y) \leqq \tfrac{1}{2}h_1(x) + \tfrac{1}{2}h_2(y)$$

and, similarly,

$$\text{(2)} \qquad \tfrac{1}{2}h(x) + \tfrac{1}{2}h(y) \leqq \tfrac{1}{2}h_1(y) + \tfrac{1}{2}h_2(x)\,.$$

If we add (1) and (2) and use the fact that $x, y \in J$ we obtain

$$h(x) + h(y) \leqq \tfrac{1}{2}h_1(x) + \tfrac{1}{2}h_2(x) + \tfrac{1}{2}h_1(y) + \tfrac{1}{2}h_2(y) = h(x) + h(y)\,,$$

which implies that equality must hold in both (1) and (2); that is,

$$h(x) - h_1(x) = h_2(y) - h(y) \quad \text{and} \quad h(x) - h_2(x) = h_1(y) - h(y)\,.$$

Let $c_2 = h(x) - h_1(x)$, $c_1 = h(x) - h_2(x)$; then since this holds for all $y \in J$, it says that on the set $J$

$$h_2 = h + c_2 \quad \text{and} \quad h_1 = h + c_1\,.$$

Moreover, $c_1 + c_2 = 2h(x) - h_1(x) - h_2(x) = 0$, so taking $c = c_1 = -c_2$ gives us $h_j = h \pm c$ on $J$. But

$$0 \geqq \sup h_j(\operatorname{ext} K) = \sup h(\operatorname{ext} K) \pm c = \pm c, \quad \text{so } c = 0\,.$$

Next, let $g = h_1 + h_2 - 2h$; this is convex and nonnegative, vanishing precisely on $J$. If $x \in K$ choose a maximal measure $\mu$ representing $x$; by the extended barycentric calculus,

$$h_i(x) \leqq \int h_i\, d\mu \quad \text{and} \quad h(x) = \int h\, d\mu\,,$$

so $0 \leqq g(x) \leqq \int g\, d\mu = 0$ since $g = 0$ on $\operatorname{ext} K$ and $K$ is metrizable (hence $\mu(\operatorname{ext} K) = 1$). Thus, $J = K$.

Note that we would not have needed metrizability in the above proof if $h$

were continuous: Since $J=\{x \in K: g(x) \leqq 0\}$, continuity of $h$ would imply that $g$ is lower semicontinuous, hence the convex set $J$ would be closed; since it contains ext $K$, it would be all of $K$. We used metrizability, of course, to guarantee that each point of $K$ can be represented by a Borel probability measure with support in ext $K$, and this is also true, for instance, when ext $K$ is an $F_\sigma$ set in $K$.

The simple example which follows shows that $\mathscr{P}$ need not be the closed convex hull of those $P_h$ with $h$ affine.

4.7. EXAMPLE. Let $K=[-1,1]$ and let $f_1(x)=-x-1$ and $f_2(x)=x-1$, $x \in [-1,1]$. Clearly, $f_1, f_2 \in A(K)$ and the function

$$g = \max(f_1, f_2)$$

is in $\mathscr{H}$. Define the continuous linear functional $L$ by $L(P)=P(f_1)+P(f_2) -P(0)$. It is straightforward to verify that $L(P_g) \leqq -1$ while $L(P_h) \geqq 0$ whenever $h \in \mathscr{H}$ is affine; this shows that $\mathscr{P}$ is not the closed convex hull of such $P_h$.

This example, when combined with Theorem 4.5, shows that—for $K= [-1,1]$ at least—*there exist "extreme" elements of $\mathscr{H}$ which are not affine.* (In fact, an elementary but character-building argument will show that the nonaffine function $g$ introduced in Example 4.7 is "extreme".) This observation is valid for any $K$ provided $K$ is a simplex, since one can prove that if $g$ is "extreme" in $\mathscr{H}(F)$, where $F$ is a closed face of $K$, then the function on $K$

$$h = \sup\{f : f \in A(K), f \leqq 0, f|_F \leqq g\}$$

extends $g$ and is "extreme' in $\mathscr{H}(K)$.

## 5. Concluding remarks.

The study of the set $\mathscr{P}$, when $K$ is the Poulsen simplex of invariant probability measure described in the introduction, can be viewed as an effort to see how much information about the pressure is contained in the variational principle, since $\mathscr{P}$ is the set of *all* convex continuous functions on $A(K)$ which satisfy that principle. It would be of interest to see what additional distinguishing characteristics of the pressure, as an element of $\mathscr{P}$, can be deduced by using its known properties. An example of what we mean by this is contained in Theorem 4.6; using known properties of the mean entropy, it shows that the pressure is an extreme point of $\mathscr{P}$ (but there are lots of these). Using property (D) one can obtain a considerable amount of information about differentiability of the pressure, but unfortunately this is shared by any extreme $P_h$ for which $h$ is continuous and affine. Finally, it is known (see e.g. [3 vol. II, p. 308] for the quantum case) that the physical pressure is Gateaux differentiable at the origin; equivalently, the lower semicontinuous affine function $h$ corresponding to (the negative of) the mean entropy attains its infimum on $K$ at a unique extreme point. Again, it is easy to produce many affine elements of $\mathscr{H}$ with this property.

## REFERENCES


1. E. Alfsen, *Compact convex sets and boundary integrals* (Ergeb. Math. Grenzgeb. 57), Springer-Verlag, Berlin - Heidelberg - New York, 1971.
2. L. Asimow and A. J. Ellis, *Convexity theory and its applications in functional analysis* (London Math. Soc. Monographs 16), Academic Press, New York, 1980.
3. O. Bratteli and D. W. Robinson, *Operator algebras and quantum statistical mechanics* I and II, Springer-Verlag, Berlin - Heidelberg - New York, 1979 and 1981.
4. H. A. M. Daniëls and A. C. D. Van Enter, *Differentiability properties of the pressure in lattice systems*, Comm. Math. Phys. 71 (1980), 65–76.
5. J. R. Giles, *Convex analysis with application in differentiation of convex functions* (Res. Notes in Math. 58), Pitman, Boston - London - Melbourne, 1982.
6. R. Israel, *Convexity in the theory of lattice gases* (Princeton Ser. Phys.), Pronceton University Press, Princeton, N.J., 1979.
7. R. Israel, *Generic triviality of phase diagrams in spaces of long-range interactions*, in preparation.
8. J. J. Moreau, Inf-*convolution des fonctions numériques sur un espace vectorielle*, C. R. Acad. Sci. Paris, 256 (1963), 5047-5049.
9. G. H. Olsen, *On simplices and the Poulsen simplex* in *Functional Analysis: Surveys and Recent Results* II (Proc. Second Conf. Functional Anal., Univ. Paderborn, Paderborn, 1979), eds. K.-D. Bierstedt and B. Fuchssteiner (North-Holland Math. Studies 38) pp. 31–52. North-Holland, Amsterdam - New York, 1980.
10. R. R. Phelps, *Lectures on Choquet's theorem*, Van Nostrand Co. Inc., Princeton, N.J. (1966).
11. D. Ruelle, *Statistical mechanics-rigorous results*, W. A. Benjamin, New York - Amsterdam, 1969.
12. D. Ruelle, *Thermodynamic formalism*, Encyclopedia Math. Appl., Addison–Wesley, Reading, Mass., 1978.



ROBERT B. ISRAEL
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF BRITISH COLUMBIA
VANCOUVER, B.C., CANADA V6T 1W5

AND

ROBERT R. PHELPS
DEPARTMENT OF MATHEMATICS GN-50
UNIVERSITY OF WASHINGTON
SEATTLE, WA 98195, U.S.A.