JOURNAL OF APPROXIMATION THEORY **34**, 1–11 (1982)

Upper Semicontinuous Functions and the Stone Approximation Theorem

GERALD BEER

Department of Mathematics, California State University, Los Angeles, California 90032

Communicated by Garrett Birkhoff

Received October 19, 1978

INTRODUCTION

In convex function theory it has long been recognized as useful to identify a convex function with its epigraph, the convex set of points on or above its graph. Similarly, a concave function is identified with its hypograph, the convex set of points on or below its graph. Analysis is then performed in the product space. We present two standard examples. First, the interior of the epigraph of a convex function f consists of those points (x, α) for which x lies in the interior of the domain of f and $\alpha > f(x)$. As a result f is upper semicontinuous on the interior of its domain. Second, a closed convex set is the intersection of the closed half spaces that contain it. As a result, a lower semicontinuous convex function is the pointwise supremum of the affine functions that it majorizes.

Such techniques have enjoyed only limited popularity in other branches of real analysis; that is, the topology and/or linear structure of the graph, epigraph, or hypograph of a real valued function are rarely used to define the fundamental concepts of analysis or to prove theorems. Their role has been essentially descriptive. It is the purpose of this paper to "perform analysis in the product" to gain a new understanding of the notion of uniform approximation. Our basic tool will be the Hausdorff metric on closed subsets of the product. Using this metric we present a generalization of the Stone Approximation Theorem to the space of upper semicontinuous functions defined on a compact metric space. In the process we extend Dini's theorem, characterizing those sequences of upper semicontinuous functions convergent pointwise from above to a continuous function that converge uniformly. Finally, since the topology on the continuous functions on a compact metric space induced by the Chebyshev norm coincides with the one induced by the Hausdorff metric when restricted to their graphs, we obtain a different view of equicontinuity and its place in the Arzela–Ascoli theorem.

0021-9045/82/010001-11$02.00/0

1. On the Stone Approximation Theorem

Let $C(X)$ be the vector space of real valued continuous functions on a compact Hausdorff space X equipped with the Chebyshev norm: $\|f\| = \sup_{x\in X} |f(x)|$. It is well known [9] that a sublattice Ω of $C(X)$ is dense if for each $\varepsilon > 0$, for each (x_1, x_2) in $X \times X$, and for each f in $C(X)$ there exists h in Ω such that $|h(x_1) - f(x_1)| < \varepsilon$ and $|h(x_2) - f(x_2)| < \varepsilon$. Following [5] we shall call this result the Stone Approximation Theorem. Often the hypotheses of the theorem are strengthened as follows: Ω is a lattice and for each two points in $X \times R$ with different first coordinates there exists a member of Ω whose graph contains them both. This condition is in turn satisfied if (1) for each f in Ω and each scalar α, both αf and $\alpha + f$ are in Ω, and (2) Ω separates points.

It is one purpose of this article to give a variant of the Stone Approximation Theorem for the space of upper semicontinuous functions on a compact metric space X. It will become clear that the condition on the graphs of members of Ω in the statement of the theorem might not be the essential one; instead what seems crucial is that points in $X \times R$ can be isolated from one another in the following sense.

Definition. Let X be a topological space and let Ω be a class of real valued functions on X. Ω is said to *isolate points* in $X \times R$ if, whenever (x_1, α_1) and (x_2, α_2) are points in $X \times R$ such that either $x_1 \neq x_2$ or $x_1 = x_2$ and $\alpha_1 < \alpha_2$, there exists f in Ω such that

$$(x_1, \alpha_1) \in \operatorname{int}\{(x, \alpha)\colon \alpha \leqslant f(x)\},$$

$$(x_2, \alpha_2) \notin \{(x, \alpha)\colon \alpha \leqslant f(x)\}.$$

This property is easy to visualize when Ω is a class of continuous functions: given a pair of distinct points (x_1, α_1) and (x_2, α_2) in $X \times R$, where (x_1, α_1) does not lie directly above (x_2, α_2), there exists a function in Ω such that (x_1, α_1) lies below its graph and (x_2, α_2) lies above. We next show that the condition on the graphs of the members of Ω in the Stone theorem implies that Ω isolates points.

Lemma 1.1. *Let X be a compact Hausdorff space and Ω a subset of $C(x)$ such that for each $\varepsilon > 0$, for each (x_1, x_2) in $X \times X$, and for each f in $C(x)$ there exists h in Ω such that $|h(x_1) - f(x_1)| < \varepsilon$ and $|h(x_2) - f(x_2)| < \varepsilon$. Then Ω isolates points in $X \times R$.*

Proof. Since each member of Ω is continuous, a member of Ω will isolate (x_1, α_1) from (x_2, α_2) if and only if (x_1, α_1) lies below its graph and (x_2, α_2) lies above. Now compact Hausdorff spaces are normal; so, by

Urysohn's lemma, for each pair of points in $X \times R$ with different first coordinates there is a continuous function whose graph passes through them both. If $x_1 = x_2$ and $\alpha_1 < \alpha_2$ there is a continuous f whose graph contains $(x_1, \frac{1}{2}(\alpha_1 + \alpha_2))$. We can find h in Ω such that $|h(x_1) - f(x_1)| < \frac{1}{2}(\alpha_2 - \alpha_1)$, and this function isolates (x_1, α_1) from (x_2, α_2). If $x_1 \neq x_2$, choose a continuous f whose graph passes through $(x_1, \alpha_1 + \varepsilon)$ and $(x_2, \alpha_2 - \varepsilon)$. There exists h in Ω that satisfies $|h(x_1) - f(x_1)| < \varepsilon$ and $|h(x_2) - f(x_2)| < \varepsilon$. Clearly, this h is the desired isolating function.

Using the Chebyshev norm it is impossible to approximate an arbitrary u.s.c. function by simple ones, e.g., continuous functions or u.s.c. step functions (when the domain is a rectangular parallelepiped). Instead, our approximation theorem in this context will be stated in terms of a different metric. Before describing this metric we recall the notion of Hausdorff distance between closed sets in a metric space.

Let Y be a metric space with metric d and for each y in Y let $B_\lambda[y]$ denote the closed λ-ball about y. If K is a closed subset of Y, then the *λ-parallel body* of K, $B_\lambda[K]$, is the set $\bigcup_{y \in K} B_\lambda[y]$. Parallel bodies need not be closed. For example, consider $[0, 1)$ as a subset of $[0, 1) \cup [2, 5]$ with the topology inherited as a subspace of the line. Clearly, $[0, 1)$ is relatively closed, yet its parallel body of radius two $[0, 1) \cup [2, 3)$ is not. If C and K are closed sets and there exists $\lambda > 0$ such that $B_\lambda[C] \supset K$ and $B_\lambda[K] \supset C$, then the *Hausdorff distance* of C from K is given by

$$D(C, K) = \inf\{\lambda : B_\lambda[C] \supset K \text{ and } B_\lambda[K] \supset C\}.$$

If no such λ exists, then we let $D(C, K)$ be infinity.

Now let X be a compact metric space with metric d. One of a number of ways to metrize $X \times R$ in a manner compatible with the product uniformity is to define the distance between (x_1, α_1) and (x_2, α_2) to be $\max\{d(x_1, x_2), |\alpha_2 - \alpha_1|\}$. To avoid excessive notation we will symbolize this distance in $X \times R$ by d, too. Let $U(X)$ denote the bounded u.s.c. functions on X. If f and g are in $U(X)$ the Chebyshev distance between them, $\sup_{x \in X} |f(x) - g(x)|$, will be represented by $d_1(f, g)$. Denote the closure of the graphs of f and g by $\bar{f}$ and $\bar{g}$. These are, or course, compact sets in $X \times R$, whence $D(\bar{f}, \bar{g}) < \infty$. We write $d_2(f, g) = \lambda$ if $D(\bar{f}, \bar{g}) = \lambda$, and the upper semicontinuity of the functions implies that d_2 is a metric, not just a pseudometric. The metric of special interest for this section requires one further definition.

DEFINITION. Let f be a real valued function on X. The *hypograph* of f, denoted by hypo f, is $\{(x, \alpha) : x \in X \text{ and } \alpha \leqslant f(x)\}$.

We finally let $d_3(f, g) = D(\text{hypo } f, \text{ hypo } g)$, a notion of distance analogous to that used by Mosco [3], Robert [6], and Salinetti and Wets [8] in their study of convex functions. Although both d_2 and d_3 are metrics on

$U(X)$, since upper semicontinuous functions are characterized by having closed hypographs, the metric d_3 seems more appropriate in this general context. On the other hand, since the continuous functions are precisely those members of $U(X)$ whose graphs are closed sets, the metric d_2 is a natural one for $C(X)$. Before stating our approximation theorem we need some basic facts about d_3 and its relationship to the other two metrics.

DEFINITION. Let X be a compact metric space and let f be u.s.c. For each positive λ the *upper λ-parallel function* of f is defined as follows:

$$f_\lambda^+(x) = \sup\{\alpha\colon (x, \alpha) \in B_\lambda[\bar{f}]\}.$$

LEMMA 1.2. *For each u.s.c. f and each positive λ, $B_\lambda[\text{hypo } f] =$* hypo f_λ^+.

Proof. Let (x, α) be in $B_\lambda[\text{hypo } f]$. Then there exists (y, β) in hypo f such that $d[(x, \alpha), (y, \beta)] \leqslant \lambda$. Since

$$d[(y, f(y)), (x, \alpha + f(y) - \beta)] = d[(y, \beta), (x, \alpha)],$$

we have $\alpha + f(y) - \beta \leqslant f_\lambda^+(x)$. Since $\alpha \leqslant \alpha + f(y) - \beta$, we conclude that (x, α) is in hypo f_λ^+. Conversely, suppose that (x, α) is in hypo f_λ^+. Now there is a sequence $\{y_n\}$ in X such that for each n, $d[(x, f_\lambda^+(x)), (y_n, f(y_n))] \leqslant \lambda + 1/n$. Since X is compact and $\{f(y_n)\}$ is bounded we can assume by passing to a subsequence that $\{(y_n, f(y_n))\}$ converges to a point (y, β), and since hypo f is closed, (y, β) will be in hypo f. Clearly, $d[(x, f_\lambda^+(x)), (y, \beta)] \leqslant \lambda$; so, $(x, f_\lambda^+(x))$ is in $B_\lambda[\text{hypo } f]$. Since $\alpha \leqslant f_\lambda^+(x)$, it is clear that (x, α) is in $B_\lambda[\text{hypo } f]$, too.

LEMMA 1.3. *Let X be a compact metric space and let f be u.s.c. Then for each $\lambda > 0$, the function f_λ^+ is in $U(X)$.*

Proof. It suffices to show that f_λ^+ is bounded below and is u.s.c. Suppose f_λ^+ were not bounded below. Then there exists a sequence $\{(z_n, \rho_n)\}$ in the complement of hypo f_λ^+ such that $\rho_n < -n$ for each n. By Lemma 1.2, (z_n, ρ_n) has distance at least λ from each point of hypo f. By passing to a subsequence we can assume that $\{z_n\}$ converges to some point z. Evidently, (z_n, ρ_n) can be made arbitrarily close to the half line $\{(z, \alpha)\colon \alpha \leqslant f(z)\}$ for all n sufficiently large. Since $\{(z, \alpha)\colon \alpha \leqslant f(z)\} \subset \text{hypo } f$, a contradiction ensues.

The upper semicontinuity of f_λ^+ follows from Lemma 1.2 and the fact that in a space where closed and bounded sets are compact, parallel bodies of closed sets are closed. Clearly, $X \times R$ is such a space.

As a result of Lemma 1.3, d_3 is a metric on the space of all u.s.c. functions and not just on $U(X)$. To see this we need only verify that the distance d_3 between two arbitrary u.s.c. functions is finite. To this end let f and g be u.s.c. By Lemma 1.3 there will exist $\alpha > -\infty$ such that $\inf\{f_1^+(x): x \in X\} > \alpha$. Let $\beta = \max\{g(x): x \in X\}$. We obtain the following inclusion:

$$B_{1+|\beta-\alpha|}[\text{hypo } f] \supset \text{hypo } g.$$

We have shown that hypo g is contained in some parallel body of hypo f. Similarly, hypo f is contained in some parallel body of hypo g, and as a result $d_3(f, g) < \infty$.

When f is continuous the function f_λ^+ need not be continuous. To see this let $X = [0, 1] \cup \{2\}$ and define $f: X \to R$ to be $2\chi_{\{2\}}$. Then

$$\begin{aligned} f_1^+(x) &= 1 \quad &&\text{if} \quad 0 \leqslant x < 1 \\ &= 3 \quad &&\text{if} \quad x = 1 \\ &= 3 \quad &&\text{if} \quad x = 2. \end{aligned}$$

The problem here seems to be that X is not connected. Actually, connectedness is neither necessary nor sufficient for the continuity of the upper parallel function. It can be shown, however, that if X is a compact convex subset of a normed linear space, then the upper parallel functions for each member of $C(X)$ are continuous.

LEMMA 1.4. *Let X be a compact metric space. If f and g are in $U(X)$, then $d_1(f, g) \geqslant d_2(f, g) \geqslant d_3(f, g)$.*

Proof. Suppose $d_1(f, g) = \lambda$. For each x in X we have $d[(x, f(x)), (x, g(x))] \leqslant \lambda$ so that the graph of f (resp. g) is a subset of $B_\lambda[\bar{g}]$ (resp. $B_\lambda[\bar{f}]$). Since $\bar{g}$ is compact, $B_\lambda[\bar{g}]$ is a closed set; so, $\bar{f} \subset B_\lambda[\bar{g}]$. Similarly, $\bar{g} \subset B_\lambda[\bar{f}]$, and we have shown that $d_2(f, g) \leqslant d_1(f, g)$.

To see that $d_2(f, g) \geqslant d_3(f, g)$ suppose that $B_\lambda[\bar{f}] \supset \bar{g}$ and $B_\lambda[\bar{g}] \supset \bar{f}$. By the first inclusion for each x in X, $f_\lambda^+(x) \geqslant g(x)$. Thus, by Lemma 1.2,

$$B_\lambda[\text{hypo } f] = \text{hypo } f_\lambda^+ \supset \text{hypo } g.$$

Similarly the second inclusion implies that $B_\lambda[\text{hypo } g] \supset \text{hypo } f$, and it follows that $d_2(f, g) \geqslant d_3(f, g)$.

In $U(X)$ convergence in d_2 does not force pointwise convergence, much less uniform convergence. Let $X = [0, 1]$ and let $f = \chi_{\{0\}}$. If n is even let $f_n = \chi_{\{1/n, 0\}}$, and if n is odd let $f_n = \chi_{\{1/n\}}$. Observe that $d_2(f_n, f) = 1/n$ but $\{f_n(0)\}$ does not converge. Convergence in d_3 does not force convergence in d_2 even when restricted to $C(X)$. To illustrate this fact for each positive integer n let $f_n: [0, 1] \to R$ be the function whose graph consists of the line segment joining $(0, -n)$ to $(1/n, 0)$ and the one joining $(1/n, 0)$ to $(1, 0)$. If f

is the zero function, then we have, for each n, $d_3(f_n, f) \leqslant 1/n$ but $d_2(f_n, f) = n$.

The next result presents a local characterization of d_3-convergence; it is a variant of a theorem for convex functions due to Mosco [3] and shows that d_3-convergence is dual to the *infimal convergence* of Wijsman [10] for l.s.c. functions.

LEMMA 1.5. *Let X be a compact metric space and let $\{f_n\}$ be a sequence of u.s.c. functions on X. Then $\{f_n\}$ converges to f in the metric d_3 if and only if*

(1) *For each x in X whenever $\{x_n\} \to x$ then* $\limsup_{n\to\infty} f_n(x_n) \leqslant f(x)$.

(2) *For each x in X there exists a sequence $\{x_n\}$ convergent to x for which* $\liminf_{n\to\infty} f_n(x_n) \geqslant f(x)$.

Proof. Suppose (1) holds. Let $\varepsilon > 0$. For each x in X there exists $\rho(x) \in (0, \varepsilon)$ and an integer N_x such that if $d(x, y) < \rho(x)$ and $n \geqslant N_x$ then $f_n(y) < f(x) + \varepsilon$. Choose $\{x_1, ..., x_k\} \subset X$ for which

$$X \subset \bigcup_{i=1}^{k} \{y: d(x_i, y) < \rho(x_i)\}.$$

Let $N = \max\{N_{x_1}, ..., N_{x_k}\}$ and let x in X be arbitrary. Choose x_i for which $d(x, x_i) < \rho(x_i) < \varepsilon$. By the definition of distance in $X \times R$, we have, for each n, $d[(x, f_n(x)), (x_i, f_n(x) - \varepsilon)] = \varepsilon$. Moreover if we choose $n \geqslant N$, then since $(x_i, f_n(x) - \varepsilon) \in \text{hypo} f$, the graph of f_n is a subset of $B_\varepsilon[\text{hypo} f]$. It follows that $\text{hypo} f_n \subset B_\varepsilon[\text{hypo} f]$.

Now suppose (2) holds. Again let $\varepsilon > 0$ be given. For each x choose $\rho(x) < \varepsilon/2$ such that if $d(y, x) < \rho(x)$ then $f(y) < f(x) + \varepsilon/2$. Pick $x_1, ..., x_k$ such that

$$X \subset \bigcup_{i=1}^{k} \{y: d(y, x_i) < \rho(x_i)\}.$$

By (2) we can choose N so large that for each $n \geqslant N$ there exists $\{x_{n1}, ..., x_{nk}\} \subset X$ such that for each $i = 1, ..., k$ both $f(x_i) - f_n(x_{ni}) < \varepsilon/2$ and $d(x_{ni}, x_i) < \rho(x_i)$. Let $x \in X$ be arbitrary. Choose x_i for which $d(x_i, x) < \rho(x_i)$. It follows that $d(x, x_{ni}) < \varepsilon$ and $f(x) < f_n(x_{ni}) + \varepsilon$. Thus $(x, f(x))$ is in $B_\varepsilon[\text{hypo} f_n]$ for each $n \geqslant N$ so that $\text{hypo} f \subset B_\varepsilon[\text{hypo} f_n]$. We have shown that (1) and (2) jointly imply that $\{f_n\}$ d_3-converges to f.

Conversely suppose $\{f_n\}$ d_3-converges to f and $\{x_n\} \to x$. There exists $\{(y_n, \beta_n)\} \subset \text{hypo} f$ such that $d[(y_n, \beta_n), (x_n, f_n(x_n))] \to 0$. Since $\{y_n\} \to x$ and f is u.s.c. at x,

$$\limsup_{n\to\infty} f_n(x_n) = \limsup_{n\to\infty} \beta_n \leqslant \limsup_{n\to\infty} f(y_n) \leqslant f(x).$$

We have established (1). Similarly there exist $\{(x_n, \lambda_n)\} \subset \text{hypo} f_n$ such that

$$d[(x_n, \lambda_n), (x, f(x))] \to 0.$$

Since $\{x_n\} \to x$ and $\{\lambda_n\} \to f(x)$ and for each $n, f_n(x_n) \geqslant \lambda_n$, we obtain (2):

$$f(x) = \lim_{n\to\infty} \lambda_n \leqslant \liminf_{n\to\infty} f_n(x_n).$$

Notice that if $\{f_n\}$ d_3-converges to f and $\{x_n\} \to x$ and (2) holds, then $\lim_{n\to\infty} f_n(x_n) = f(x)$. In particular if for each n we have $f_n(x) \geqslant f(x)$, then $\lim_{n\to\infty} f_n(x) = f(x)$. We also remark that the compactness of X not only guarantees the sufficiency of conditions (1) and (2) but is *necessary* for their sufficiency. In other words if X is not compact we can select u.s.c. $\{f_n\}$ and f that satisfy (1) and (2), but $\{f_n\}$ fails to d_3-converge to f. To see this let X be noncompact. Choose a sequence $\{y_n\}$ in X with no convergent subsequence. Let $f_n : X \to R$ be defined by

$$\begin{aligned} f_n(x) &= 1 \qquad \text{if} \quad x = y_n \\ &= 0 \qquad \text{otherwise.} \end{aligned}$$

Then $\{f_n\}$ satisfies (1) and (2) with respect to the zero function, but $\{f_n\}$ does not d_3-converge to zero.

Our main result, Theorem 1, and the Stone Approximation Theorem (which in light of Lemma 1.1 is an immediate corollary) are consequences of the following approximation theorem for compact sets. The superscript $\sim$ used below denotes set complement.

LEMMA 1.6. *Let Σ be a lattice of compact sets in a compact topological space X. Let C be a compact subset of X. Suppose for each (x, y) in $C \times \tilde{C}$ there exists $K(x, y)$ in Σ such that x is in* int $K(x, y)$ *but y is not in $K(x, y)$. Then if O is an open set containing C, there exists K in Σ such that $C \subset \text{int } K \subset O$.*

Proof. Fix y in $\tilde{O}$. For each x in C choose a set $K(x, y)$ as described above. By the compactness of C there exists a finite subset of $\{K(x, y): x \in C\}$ whose interiors cover C. Since Σ is a lattice, the union of this finite family K_y is in Σ. Repeating this construction for each y in $\tilde{O}$ we see that $\{\tilde{K}_y : y \in \tilde{O}\}$ is an open cover of $\tilde{O}$. Let $\{\tilde{K}_{y_i} : i = 1,..., n\}$ be a finite subcover. Then $C \subset \text{int} \bigcap_{i=1}^{n} K_{y_i} \subset O$. Since $\bigcap K_{y_i}$ is in Σ, we are done.

Our main theorem gives sufficient, but not necessary, conditions for a sublattice Ω of the u.s.c. functions to be d_3-dense. For example, the sublattice Θ of the u.s.c. functions consisting of the functions constant except at finitely many points does not isolate points but is nevertheless d_3-dense. To see this, note that the bounded u.s.c. functions (by virtue of including upper parallel functions) are d_3-dense, and using the total boundedness of X we can inscribe in the hypograph of each member of $U(X)$ the hypograph of a member of Θ to any desired degree of d_3-accuracy.

However, it is easy to see that the condition "Ω isolates points" actually characterizes sublattices Ω of the u.s.c. functions that are upper dense, i.e., for which each u.s.c. f is in the closure of $\{g: g \geqslant f \text{ and } g \in \Omega\}$.

THEOREM 1. *Let Ω be a lattice of u.s.c. functions on a compact metric space X that isolates points.*

(a) *If f is u.s.c. then there exists a sequence $\{h_n\}$ in Ω convergent to f from above in the metric d_3.*

(b) *If f is continuous and $\{h_n\}$ is d_3-convergent to f from above, then $\{h_n\}$ converges uniformly to f.*

Proof. (a) Since f can be d_3-approximated from above by its upper parallel functions, Lemma 1.3 allows us to assume w.l.o.g. that f is in $U(X)$. Next choose m and M such that for each x in X we have $m < f(x) < M$. For each h in Ω let $h^*: X \to R$ be defined by $h^*(x) = \min\{h(x), M\}$. Clearly, $\Omega^* = \{h^*: h \in \Omega\}$ is a lattice of u.s.c. functions whence $\{\text{hypo}\, h^*: h \in \Omega\}$ is a lattice of closed sets in $X \times R$. Since $X \times [m, M]$ is compact, for each h in Ω, $K_h = \{(x, \alpha): m \leqslant \alpha \leqslant h^*(x)\}$ is compact. Let $C = \{(x, \alpha): m \leqslant \alpha \leqslant f(x)\}$. We claim that $\{K_h: h \in \Omega\}$ satisfies the hypotheses of Lemma 1.6 relative to C in the compact space $X \times [m, M]$.

To see this let (x_1, α_1) be in C and let (x_2, α_2) be in $\tilde{C}$. It is not the case that $x_1 = x_2$ and $\alpha_2 \leqslant \alpha_1$. Since Ω isolates points there exists h in Ω such that $(x_1, \alpha_1) \in \text{int}(\text{hypo}\, h)$ and $\alpha_2 > h(x_2)$. Since $M \geqslant \alpha_2$ we have $h^*(x_2) = h(x_2)$ so that (x_2, α_2) is not in K_h. Choose $\lambda > 0$ and a neighborhood V of x_1 such that $V \times (\alpha_1 - \lambda, \ \alpha_1 + \lambda) \subset \text{hypo}\, h$. If $\alpha_1 < M$ set $\delta = \min\{M - \alpha_1, \lambda\}$. Clearly, $V \times (\alpha_1 - \delta, \ \alpha_1 + \delta) \subset \text{hypo}\, h^*$ so that (x_1, α_1) is an interior point of K_h relative to $X \times [m, M]$. If $\alpha_1 = M$ then h exceeds M throughout V so that $h^*(x) = M$ for each x in this neighborhood. Thus, $(x_1, \alpha_1) \in (V \times R) \cap (X \times [m, M]) \subset K_h$. Once again, (x_1, α_1) is in the interior of K_h relative to $X \times [m, M]$.

Continuing, choose N so large that for all x we have $f(x) + 1/N < M$. For each $n > N$ let O_n be the union in the subspace $X \times [m, M]$ of the open $1/n$ balls whose centers lie in C. By Lemma 1.6 there exists h_n in Ω such that $C \subset K_{h_n} \subset O_n$. Now the second coordinate of each point in O_n is less than M; so, $h_n^* = h_n$. Clearly, h_n majorizes f so that $B_{1/n}[\text{hypo}\, h_n] \supset \text{hypo}\, f$. Furthermore, since $O_n \subset B_{1/n}[\text{hypo}\, f]$ and $\{(x, \alpha): \alpha \leqslant m\} \subset (\text{hypo}\, f) \cap (\text{hypo}\, h_n)$, we have $\text{hypo}\, h_n \subset B_{1/n}[\text{hypo}\, f]$. Thus, $\{h_n\}$ converges to f in the metric d_3 from above. Notice that as a consequence of Lemma 1.5 the convergence is automatically pointwise.

(b) If $d_3(h_n, f) < 1/k$ then we have

$$\text{hypo}\, f \subset \text{hypo}\, h_n \subset B_{1/k}[\text{hypo}\, f].$$

By Lemma 1.2 for each x in X it follows that $f(x) \leqslant h_n(x) \leqslant f^+_{1/k}(x)$. By Lemma 1.3 and Lemma 1.5 the sequence $\{f^+_{1/k} - f\}$ satisfies the hypotheses of Dini's theorem [7]. Thus $d_1(f^+_{1/k}, f) \to 0$ so that $d_1(h_n, f) \to 0$.

Two remarks are in order. First, it is probably not possible to replace the condition "Ω isolates points" by "Ω separates points" in conjunction with a nice algebraic condition. Upper semicontinuous functions are not closed under multiplication by negative scalars, and if Ω were merely a cone-lattice of u.s.c. functions that separates points, then Ω need not be dense. For example, the cone-lattice generated by the increasing affine functions on $[0, 1]$ separates points but is not dense in $U[0, 1]$ because it consists solely of increasing functions. Second, the approximations described in the last theorem are not irrelevant. In particular for each u.s.c. function f on a closed interval there exists a decreasing sequence of u.s.c. step functions convergent to f pointwise [7]. Theorem 1 gives this result a metric interpretation, for u.s.c. step functions form a lattice that isolates points. More importantly, since the continuous functions on a compact metric space form a lattice that isolates points, they are dense in the u.s.c. functions relative to d_3.

We close this section by noting that a lattice of u.s.c. functions that isolates points need not be dense in $U(X)$ relative to d_2. In fact, $C(X)$ need not be dense relative to this metric. To see this let $X = [0, 1]$ and consider the u.s.c. function $\chi_{\{0\}}$. If $d_2(\chi_{\{0\}}, g) \leqslant \frac{1}{3}$, then the graph of g must lie in both of two disjoint rectangles, $[-\frac{1}{3}, \frac{1}{3}] \times [\frac{2}{3}, \frac{4}{3}]$ and $[-\frac{1}{3}, \frac{4}{3}] \times [-\frac{1}{3}, \frac{1}{3}]$. Clearly, these rectangles disconnect the graph of g, whence g cannot be continuous.

2. On the Ascoli Theorem

Lemma 2.1. *Let X be a compact metric space. Let $\{f_n\}$ be a sequence of u.s.c. functions and let f be a fixed member of $C(X)$. Then $d_1(f_n, f) \to 0$ if and only if $d_2(f, f_n) \to 0$.*

Proof. Since $d_1(f_n, f) \geqslant d_2(f_n, f)$, one direction is trivial. Suppose now that $d_2(f, f_n) \to 0$. For each positive λ we define the lower λ-parallel function of f as follows: $f^-_\lambda(x) = \inf\{\alpha\colon (x, \alpha) \in B_\lambda[\bar{f}]\}$. Since f is lower semicontinuous, Lemma 1.3 implies that each f^-_λ is l.s.c. By the convergence of $\{f_n\}$ in d_2 for each positive integer k there exists N such that $n > N$ implies that for all x, $f^-_{1/k}(x) \leqslant f_n(x) \leqslant f^+_{1/k}(x)$. Hence, to show that $d_1(f, f_n) \to 0$, it suffices to show that $\{f^+_{1/k} - f^-_{1/k}\}$ converges to zero uniformly. Since the graph of f is closed and $\bar{f} = \bigcap_{k=1}^\infty B_{1/k}[\bar{f}]$, it is clear that $\{f^+_{1/k} - f^-_{1/k}\}$ converges pointwise to the zero function. Moreover, since for each k,

$f^+_{1/k} - f^-_{1/k}$ is u.s.c., the sequence satisfies the hypotheses of Dini's theorem. Thus, the convergence is uniform and the lemma is proved.

Lemma 2.1 implies that d_2 and d_1 are equivalent metrics on $C(X)$, a fact noted by Naimpally [4]. The compactness of X is indispensable. To see this consider the spiked function f: $[0, \infty) \to R$ whose graph consists of line segments connecting the following points in succession:

$$(0, 0),\ (\tfrac{1}{2}, 0),\ (1, 1),\ (\tfrac{3}{2}, 0),\ (\tfrac{5}{3}, 0),\ (2, 1),\ (\tfrac{7}{3}, 0),\ (\tfrac{11}{4}, 0),\ (3, 1),\ (\tfrac{13}{4}, 0), \ldots .$$

For each positive integer n the upper $1/n$ parallel function $f^+_{1/n}$ is continuous, and it is evident that $d_2(f^+_{1/n}, f) \to 0$. However, the convergence fails to be uniform because for each $n > 1$ we have $f^+_{1/n}(n - 1 - 1/n) \geqslant 1$ whereas $f(n - 1 - 1/n) = 0$.

The equivalence of d_1 and d_2 on $C(X)$ when X is compact brings forth the geometric substance of the Ascoli theorem: a closed subset of $C(X)$ is compact in the metric d_1 if and only if it is equicontinuous and bounded.

Let $\{f_n\}$ be a bounded sequence of functions in $C(X)$. There exists $r > 0$ such that for each n, the graph $\bar{f}_n$ of f_n lies in $X \times [-r, r]$. Now the compact subsets of the compact space $X \times [-r, r]$ under the Hausdorff metric form a compact metric sace [1]; so, we can extract a subsequence $\{\bar{f}_{n_k}\}$ of $\{\bar{f}_n\}$ convergent in the Hausdorff metric to a compact subset C of $X \times [-r, r]$. If we knew that C were the graph of a function f, then automatically f would be continuous because its graph is a compact set, and by Lemma 2.1, $\{f_{n_k}\}$ must converge uniformly to f. The equicontinuity of $\{f_{n_k}\}$ not only guarantees that C is a graph of a function; the two notions are equivalent.

THEOREM 2. *Let $\{f_n\}$ be a sequence of continuous functions on a compact metric space X. Suppose there exists a compact subset C of $X \times R$ such that the graphs of the terms of the sequence converge in the Hausdorff metric to C, i.e., $D(\bar{f}_n, C) \to 0$. Then C is the graph of a function if and only if $\{f_n\}$ is an equicontinuous sequence.*

Proof. We recall some terminology. The upper (resp. lower) closed limit of a sequence of sets in a metric space is the set of points each neighborhood of which intersects infinitely many (resp. all but finitely many) terms of the sequence. Since $D(\bar{f}_n, C) \to 0$, the upper and lower closed limits of $\{\bar{f}_n\}$ both equal C [1]. In particular, since C equals the upper closed limit of $\{\bar{f}_n\}$ and for each x $\{f_n(x)\colon n = 1, 2, \ldots\}$ is a bounded set of numbers, there exists α_x such that (x, α_x) is in C.

First, suppose that C is not the graph of a function. By the comment at the end of the last paragraph, there must exist distinct points (x, α) and (x, β)

both contained in C. Since both points belong to the lower closed limit of $\{\bar{f}_n\}$ there exist sequences $\{(z_n, \alpha_n)\}$ and $\{(y_n, \beta_n)\}$ such that

$$\{(z_n, \alpha_n)\} \to (x, \alpha) \qquad \text{and} \qquad \{(y_n, \beta_n)\} \to (x, \beta)$$

and for each n,

$$\{(z_n, \alpha_n), (y_n, \beta_n)\} \subset \bar{f}_n.$$

For large n the number $d(z_n, y_n)$ can be made arbitrarily small whereas $|f_n(z_n) - f_n(y_n)|$ will exceed $\frac{1}{2}|\alpha - \beta|$. Thus, the sequence of functions is not equicontinuous.

Conversely, suppose that $\{f_n\}$ fails to be equicontinuous. Since each term of the sequence is uniformly continuous, there exists $\varepsilon > 0$, a subsequence $\{f_{n_k}\}$ of $\{f_n\}$ and points $\{z_k\}$ and $\{y_k\}$ in X such that for each k

$$d(z_k, y_k) < 1/k \qquad \text{but} \qquad |f_{n_k}(y_k) - f_{n_k}(z_k)| > \varepsilon.$$

Since C is the upper closed limit of $\{\bar{f}_n\}$, it is easy to verify that C contains two distinct points with the same first coordinate. Hence, C is not the graph of a function.

For completeness, we mention that the Ascoli theorem can be used to establish the compactness of the space of compact subsets of a compact metric space X under the Hausdorff metric, for the distance functions associated with the compact subsets of X are a closed, bounded, equicontinuous subset of $C(X)$ [2].

References

1. F. Hausdorff, "Set Theory," Chelsea, New York, 1962.
2. E. Heil, Über die Auswahlsatze von Blaschke und Arzela–Ascoli, *Math. Phys. Semesterber.* **14** (1967), 169–175.
3. U. Mosco, Convergence of convex sets and of solutions of variational inequalities, *Adv. in Math.* **3** (1969), 510–585.
4. S. A. Naimpally, Graph topology for function spaces, *Trans. Amer. Math. Soc.* **123** (1966), 267–272.
5. M. H. Protter and C. B. Morrey, "A First Course in Real Analysis," Springer–Verlag, New York, 1977.
6. R. Robert, Convergence de fonctionnelles convexes, *J. Math. Anal. Appl.* **45** (1974), 533–555.
7. H. L. Royden, "Real Analysis," Macmillan Co., New York, 1968.
8. G. Salinetti and R. Wets, On the relations between two types of convergence for convex functions, *J. Math. Anal. Appl.* **60** (1977), 211–226.
9. M. H. Stone, A generalized Weierstrass approximation theorem, *in* "Studies in Modern Analysis" (R. C. Buck, Ed.), Vol. 1, M.A.A. Studies in Mathematics, Prentice–Hall, Englewood Cliffs, N.J., 1962.
10. R. Wijsman, Convergence of sequences of convex sets, cones and functions, II, *Trans. Amer. Math. Soc.* **123** (1966), 32–45.