SETS WITH UNIQUE FARTHEST POINTS

BY
EDGAR ASPLUND

ABSTRACT
If a certain set in $\mathbb{R}^n$ has the property, that in some unsymmetric norm each point in $\mathbb{R}^n$ has a unique farthest point in this set, then it consists of exactly one point.

Let there be given a finite dimensional real vector space with a norm that satisfies all usual conditions except that it need not be symmetric, only homogeneous for multiplication with positive scalars. Suppose that, in the (unsymmetric) metric defined by this norm, a certain set has the property that *each* point in the space has a unique farthest point in this set, then we say that the set has unique farthest points. We shall prove that the sets with unique farthest points are exactly the one point subsets of the space. Since every one point set is a set with unique farthest points, we have only to prove the converse.

This problem was solved by V. L. Klee [3] for several types of special norms in finite dimensional space and also for some infinite dimensional cases under certain extra a priori chnditions on the set with unique farthest points. The coveted goal is to solve the problem for an infinite dimensional Hilbert space (c.f. Klee, loc. cit.). No counterexample is so for known in any normed real linear vector space.

Here we give a proof of the most general finite dimensional case. We will use the apparatus of "convexity calculus" developed by Brøndsted, Moreau, Rockafellar and others. In the finite dimensional case the notions are essentially due to Fenchel [2], except for that of subdifferential, which has appeared later. Furthermore, we will use a result of the author [1], which we refer to as the generalized Straszewics theorem. In the find final section we give a more detailed discussion of the relation of our results to Klee's, and also a proof for the infinite dimensional space $c_0(w)$.

1. **Conjugate convex functions and subdifferentials.** By a *lower semicontinuous proper convex function* we mean a function f on $\mathbb{R}^n$, with values in $\mathbb{R} \cup \{+\infty\}$ and not identically $+\infty$, such that

$$f(\lambda x + (1-\lambda)y) \leqq \lambda f(x) + (1-\lambda) f(y)$$

for all x, y in $\mathbb{R}^n$ and $0 \leqq \lambda \leqq 1$ $\quad f(x) = \liminf_{y \to x} f(y)$ for all x in $\mathbb{R}^n$

Received September 30, 1963; revised version, February 8, 1967.

For brevity, we will say simply convex function instead of lower semicontinuous proper convex function. Then each convex function f has a conjugate convex function f^*, defined by

$$f^*(x) = \sup \{\langle x, y\rangle - f(y) : y \in \mathbb{R}^n\} \tag{1.1}$$

and $(f^*)^* = f$. The set of points in $\mathbb{R}^n$ where f has a finite value is called the *effective domain* of f and denoted by

$$\operatorname{dom} f = \{x : x \in \mathbb{R}^n, f(x) < \infty\}$$

The *subdifferential* ∂f of f is a set valued function defined for each x in $\operatorname{dom} f$ by

$$\partial f(x) = \{y : f(z) \geqq f(x) + \langle y, z - x\rangle \text{ for all } z \in \mathbb{R}^n\} \tag{1.2}$$

This relation says, that in the "graph space"

$$\mathbb{R}^n \times \mathbb{R} = \{(x, a) : x \in \mathbb{R}^n, a \in \mathbb{R}\}$$

the set

$$\{(z, a) : a = f(x) + \langle y, z - x\rangle\} \tag{1.3}$$

is a supporting hyperplane to the *epigraph* grf of f, defined by

$$\operatorname{grf} = \{(z, a) : z \in \operatorname{dom} f, a \geqq f(z)\}$$

Actually, grf is a closed convex subset of the graph space, and each "non-vertical" supporting hyperplane (a hyperplane in $\mathbb{R}^n \times \mathbb{R}$ would be called "vertical" if its projection onto $\mathbb{R}^n$ is not all of $\mathbb{R}^n$) of grf can be written as (1.3) for some x in $\operatorname{dom} f$ and y in $\partial f(x)$.

From (1.1) and (1.2) we find that

$$\partial f(x) = \{y : f^*(y) = \langle y, x\rangle - f(x)\} \tag{1,4}$$

and from the symmetry of (1.4) it is clear that

$$y \in \partial f(x) \Leftrightarrow x \in \partial f^*(y \Leftrightarrow f(x) + f^*(y) = \langle y, x\rangle$$

Relation (1.4) can also be interpreted as stating: $\partial f(x)$ is the projection onto $\mathbb{R}^n$ of the intersection of grf* and its supporting hyperplane

$$\{(z, a) : a = \langle z, x\rangle - f(x) = f^*(y) + \langle z - y, x\rangle\}, \text{ c.f. (1.3).}$$

All statements in this section that are not completely evident are proved in Fenchel [2].

2. **The generalized Straszewics theorem.** The well known Straszewics theorem says that in a compact convex set C in $\mathbb{R}^n$, the set of all exposed points are dense in the set of all extreme points, a point of C being called exposed if

there is a hyperplane H such that $H \cup C$ contains this point alone. We say that a subset of C is a *face* of C if it is the intersection of C and some supporting hyperplane. Then a point is exposed if the smallest dimension of any face containing it is 0. To state a generalization of Straszewics theorem we define a point (in the boundary of C) to be *k-exposed* if the smallest dimension of any face containing it is at most k, and to be *k-extreme* if the largest dimension of any simplex contained in C, of which the point is the barycenter, is at most k. The following theorem is proved in [1]. For $k = 0$ it's the old Straszewics theorem.

THEOREM 1. (*Generalized Straszewics theorem*): *Each k-extreme point of the boundary of a compact convex set is the limit of k-exposed points.*

3. **Restatements of the farthest point problem.** We assume that we have in $\mathbb{R}^n$ an "unsymmetric norm", i.e. a function $x \to \| x \| : \mathbb{R}^n \to \mathbb{R}^+$ satisfying

(3.1) $\| \lambda x \| = \lambda \| x \|$ for $\lambda > 0$ and

(3.2) $B = \{x : \| x \| \leqq 1\}$is a bounded convex neighborhood of 0 in $\mathbb{R}^n$.

In general, however, $\| x \| \neq \| - x \|$ so one has to be careful with signs.

Now let S be a set that has unique farthest points with respect to this norm. In other words, given any point x in $\mathbb{R}^n$ we assume that there exists a point $q(x)$ in S (the "antiprojection" of x onto S) such that

$$q(x) \neq y \in S \text{ implies } \| y - x \| < \| q(x) - x \|$$

Also, we will let s denote one point in S, fixed once and for all. The theorem that we have set out to prove is the following.

THEOREM 2. *The set S consists of the point s alone.*

The proof will be indirect, assuming the counterhypothesis to Theorem 2 in the following version

(CH) $$q(x) \neq x \text{ for all } x \text{ in } \mathbb{R}^n$$

If (CH) is true then Theorem 2 must be false, since $q(q(x)) \neq q(x)$ and both are in S, whereas if Theorem 2 is true then (CE) is false because $q(s) = s$. Hence another way of stating Theorem 2 is: the antiprojection onto S has a fixpoint. It then follows immediately that the antiprojection is a constant.

Another reformulation of Theorem 2, which will not be used in the proof, is the following. Denote by $B(x, r)$ the (unsymmetric) ball with center x and radius r:

$$B(x, r) = \{y : \| y - x \| \leqq r\}$$

Then $S \subset B(x, \| q(x) - x \|)$ and $\bar{S} \cap bdB(x, \| q(x) - x \|) \neq \varnothing$ together define the number $\| q(x) - x \|$ in a way which would make sense for any bounded set S even if it did not have unique farthest points, whereas "for each x in $\mathbb{R}^n$,

$S \cap bdB(x, \| q(x) - x \|)$ has exactly one element'' expresses this unique farther point property. Now

$$\mathscr{B} = \{B(x\ r): x \in \mathbf{R}^n, r \geqq 0\}$$

is exactly the family of all (positive) homotheties of the unit ball B. Suppose that $\mathscr{C}$ were any family of closed convex sets. Define S to be *strictly convex with respect to* $\mathscr{C}$ if $S \subset C$ for some $C \in \mathscr{C}$ and

$$C \in \mathscr{C}\ \ S \subset C \text{ and } \bar{S} \cap bdC \neq \emptyset \text{ implies that } S \cap bdC \text{ has exactly one element}$$

Since B could have been a translate of any given compact convex body, we have the following reformulation of Theorem 2.

THEOREM 2′. *If the set S is strictly convex with respect to any family of all positive homotheties of some fixed compact convex body, then S consists of exactly one point.*

4. **Proof of Theorem 2.** We introduce the real-valued function g defined by

$$g(x) = \| q(x) - x \| = \sup \{ \| y - x \| : y \in S\}$$

The function g is the supremum of the elementary functions $x \to \| y - x \|$ that are clearly convex, so g itself is convex in the sense of Section 1. It also satisfies the following Lipschitz condition

$$-\| x - y \| \leqq g(x) - g(y) \leqq \| y - x \| \text{ for all } x, y \in \mathbf{R}^n \tag{4.1}$$

with equality on either side only if $q(x) = q(y)$. We will denote the polar body of the unit ball B by B^0:

$$\begin{aligned} B^0 &= \{x : \langle x, y \rangle \leqq 1 \text{ for all } y \in B\} \\ &= \{x : \langle x, y \rangle \leqq \| y \| \text{ for all } y \in \mathbf{R}^n\} \end{aligned}$$

We will also define two set-valued mappings D and D^* by

$$D(x) = \{y : \langle y, x \rangle = 1, y \in B^0\} \text{ for } x \in bdB (\text{i.e., } \| x \| = 1)$$

and

$$D^*(x) = \{y : \langle x, y \rangle = 1,\ y \in B\} \text{ for } x \in bdB^*$$

These are sometimes called the spherical mappings or duality mappings between the boundaries of the unit ball and its polar body. Obviously, $D(x) \subset bdB^0$ for each$x$ in bdB and $D^*(x) \subset bdB$ for each x in bdB^0. We assume (CH), so we may define the function b, with values in bdB, by

$$b(x) = \frac{q(x) - x}{\| q(x) - x \|}.$$

for all x in $\mathbf{R}^n$. Now we will prove the following lemma

LEMMA 1. $-D(b(x)) \subset \partial g(x)$

Proof. Take y in $-D(b(x))$. This means that

$$\langle -y, q(x)-x\rangle = \|q(x)-x\|, \langle -y,u\rangle \leqq \|u\| \text{ for all } u \text{ in } \mathbb{R}^n$$

Now compute

$$g(x) = \|q(x)-x\| = \|q(x)-z+(z-x)\| = \langle -y, q(x)-z+(z-x)\rangle$$
$$\leqq \|q(x)-z\| - \langle y, z-x\rangle \leqq g(z) - \langle y, z-x\rangle$$

Hence

$$g(z) \geqq g(x) + \langle y, z-x\rangle \text{ for all } z \text{ in } \mathbb{R}^n$$

so that y is also in $\partial g(x)$, as claimed.

We go on to interpret $\partial g(x)$ as the projection onto $\mathbb{R}^n$ of the intersection of the epigraph of g^*,

$$grg^* = \{(z,a): a \geqq g^*(z), z \in \operatorname{dom} g^*\}$$

and its supporting hyperplane

$$\{(z,a): a = \langle z,x\rangle - g(x)\}$$

In other words, $\partial g(x)$ is the projection onto $\mathbb{R}^n$ of a face of grg^*, in the sense of Section 2. We want of course to apply the generalized Straszewics theorem to grg^*. However, this is an unbounded set, so we must find out how to truncate it in a good way.

The following inequalities hold for g:

$$\|s-x\| \leqq g(x) \leqq g(0) + \|-x\| \text{ for all } x \text{ in } \mathbb{R}^n$$

From them, we derive the corresponding inequalities for g^*:

$$-g(0) \leqq g^*(x) \leqq \langle x,s\rangle \leqq \|-s\| \text{ for } x \text{ in } -B^0$$
$$g^*(x) = +\infty \text{ otherwise}$$

In other words, dom $g^* = -B^0$, and in its effective domain, g^* assumes values between $-g(0)$ and $\|-s\|$. Thus, we may consider grg^* to be truncated, by adding, say, $a \leqq \|-s\| + 2$ to the defining relations. We will be interested in that part of the boundary of grg^* which lies in the open set

$$\{(z,a): z \in \operatorname{int}(-B^0)\ a < \|-s\| + 1\}$$

For brevity, we will say simply that a boundary point of grg^* is "in int $(-B^0)$", when it lies in the above open set.

As a first application of Straszewics theorem (for this the classical version suffices) we show that grg^* has no extreme points in int $(-B^0)$. For if it were so, then some boundary point $(z, g^*(z))$ in int $(-B^0)$ would have to be exposed, i.e. possible to separate from the rest of grg^* by some hyperplane with gradient x. But then z would be the only element of $\partial g(x)$, whereas we know by Lemma 1 that $\partial g(x)$ contains points in the boundary of $-B^0$. Hence there are no extreme points of grg^* in the interior of $-B^0$.

Suppose now, inductively, that we have shown that grg^* has no $(k-1)$-extreme points in int $(-B^0)$, for some $k \leqq n-1$, and suppose that the point $(z, g^*(z))$ with z in int$(-B^0)$, were k-exposed. Then there is an x in $\mathbb{R}^n$ such that z is in $\partial g(x)$ and $\partial g(x)$ has dimension k. The relative boundary points of $\partial g(x)$ are $(k-1)$-extreme and so they all lie in the boundary of $-B^0$, by the induction hypothesis. Hence actually

$$\partial g(x) = (-B^0) \cap K$$

where K is a k-dimensional affine subspace of $\mathbb{R}^n$. By Lemma 1, the intersection of $-B^0$ and its supporting hyperplane

$$H = \{y : \langle y, b(x) \rangle = -1\}$$

which is $-D(b(x))$, is contained in $\partial g(x)$. Here we need a second lemma.

LEMMA 2. *If z in $\partial g(x)$ is a boundary point of $-B^0$, then any hyperplane supporting $-B^0$ at z intersects $-D(b(x)) = H \cap (-B^0)$.*

Proof. We have by definition that

$$\langle -z, u \rangle = \| u \| = 1 \text{ for all } u \text{ in } D^*(-z).$$

The hyperplane defined by u is the set

$$H_u = \{y : \langle -y, u \rangle = 1\}$$

and to prove Lemma 2 we have to show that $H \cap H_u \cap (-B^0)$ is nonempty. Since z is in $\partial g(x)$,

$$\langle z, y \rangle - g(y) \leqq \langle z\ x \rangle - g(x) \text{ for all } y \text{ in } \mathbb{R}^n$$

Thus, for $y = x - \lambda u$, $\lambda > 0$ we have

$$g(x) - g(y) \leqq -\langle -z, x-y \rangle = -\| x - y \|$$

so, by (4.1) we have equality, and $q(x) = q(y)$. We will now show that the midpoint of u and $b(x)$ is in bdB, i.e. $\| \frac{1}{2}(u + b(x)) \| = 1$. Take $\lambda = \| q(x) - x \|$ above. Then $\| \frac{1}{2}(u + b(x)) \| = \| q(x) - x + x - y \| / 2 \| q(x) - x \| = \| q(y) - y \| / 2 \| q(x) - x \| = (\| q(x) - x \| + \| x - y \|) / 2 \| q(x) - x \| = 1$ as asserted. Now let v be any element of $D(\frac{1}{2}(u + b(x))$, i.e. v is in B^0, and

$$\frac{1}{2} \langle v, u + b(x) \rangle = 1$$

Since u and $b(x)$ are both in B, we have that

$$\langle v, u\rangle = \langle v, b(x)\rangle = 1$$

Hence the element $-v$ is both in H and H_u, and since it is also in $-B^0$, we have completed the proof of Lemma 2.

Going back to the proof of Theorem 2, we see that $H \cap K$ is a supporting hyperplane of $\partial g(x)$ in K. Since $\partial g(x)$ is supposed to have interior points relative to K, there is in K a hyperplane supporting $\partial g(x)$ which is parallel to but different from H. This can then be extended to a hyperplane J supporting $-B^0$ in $\mathbb{R}^n$. But J meets $H \cap (-B^0)$ by Lemma 2, so J must contain $\partial g(x)$, which is therefore contained in the boundary of $-B^0$ a contradiction. We conclude that no point of grg^* in int $(-B^0)$ can be k-exposed.

Invoking Theorem 1, we now find by induction that no point of grg^* in int $(-B^0)$ is $(n-1)$-extreme. But that means, that if z is any point in the interior of $-B^0$ and x is in $\partial g^*(z)$, then $\partial g(x)$ is all of $-B^0$, since if z is k-extreme in $\partial g(x)$ then $(z, g^*(z))$ is k-extreme in qrq^*, for $k \leq n-1$. As in the proof of Lemma 2 we have that $q(x) = q(y)$ for every $y = x - \lambda u$ with $\lambda > 0$ and u in $D^*(-z)$ for some $-z$ in the boundary of B^0, which now coincides with $-\partial g(x)$. But this means that u can be any element of bdB so $q(x) = q(y)$ for all $y \in \mathbb{R}^n$ contradicting (CH). Thereby the proof of Theorem 2 becomes complete.

5. **Klee's results, an infinite dimensional case, and an open problem.** The special cases solved by Klee that were mentioned in the introduction are the following

1. The set S is closed.
2. The norm is rotund.
3. The norm is polyhedral, i.e. it is the maximum of a finite family of linear functions.

Each case has its own proof, and they are all very different from each other and from our proof here, which is closes related to Klee's proof of Case 3, but much more complicated. Klee states and proves Case 1 and 2 for general Banach spaces, with the additional assumption that S is compact and totally bounded, respectively. Klee does not extend Case 3 to any infinite dimensional case, but it is in fact possible to do so, by means of the following remark. Let E be a Banach space taking the role of $\mathbb{R}^n$ and let all other notations be as before. We have then

LEMMA 3. *The family $\{q^{-1}(x): x \in S\}$ is a cover of E consisting of closed sets that are pairwise either disjoint or identical. Hence the family has either all members equal to E or else uncountably many different members.*

Proof. The closedness of $q^{-1}(x) = \{y: q(y) = x\}$ is obvious from the continuity of the norm, and the disjointness or identity between any pair $q^{-1}(x), q^{-1}(y)$ is of course another way of stating the uniqueness of the farthest point. Furthermore, the existence of farthest points on all of E shows that E is covered by the family. Finally, the uncountability of a nontrivial closed disjoint cover of E is proved by reduction to the case $E = \mathbb{R}$ (take the trace of the whole configuration on a line contained in E) and there it is an easy consequence of the Cantor-Bendixon theorem on uncountability of perfect sets. Namely, the set of all interval endpoints of the complements of the closed sets in the family would be at the same time countable and perfect, which is impossible.

From Lemma 3 we got the following theorem.

THEOREM 3. *If the norm in E is the maximum of a countable family of linear functions (i.e. there is a sequence y_n in E^* such that, for each x in E, $\langle x, y_n\rangle \leqq \| x \|$ for all n and $\langle x, y_n\rangle = \| x \|$ for some n), then each set with unique farthest points consists of a single point.*

COROLLARY *Each set with unique farthest points in $c_0(\omega)$ consists of a single point.*

Proof. Say that a farthest point $q(x)$ corresponds to the element y_n in E^* if

$$\| q(x) - x \| = \langle (x) - x, y_n\rangle$$

Suppose another farthest point $q(z)$ also corresponds to y_n. From the following short computation

$$\langle q(z) - x, y_n\rangle \leqq \| q(z) - x \| \leqq \| q(x) - x \| = \langle q(x) - x, y_n\rangle \tag{5.1}$$

we deduce that $\langle q(z), y_n\rangle \leqq \langle q(x), y_n\rangle$, and the converse inequality is obtained by letting x and z change places in (5.1) so $\langle q(z), y_n\rangle = \langle q(x), y_n\rangle$. But then substitution into (5.1) shows that we have equality all along in (5.1), which implies $q(x) = q(z)$. Thus at most one farthest point corresponds to each y_n. It follows that $\{q^{-1}(x): x \in S\}$ is countable, hence by Lemma 3 trivial, and this proves Theorem 3.

Thus Klee's results extend in all cases to some infinite dimensional situation (note that our Theorem 3 extends Klee's Case 3 in some finite-dimensional cases too, but this has already been covered by our Theorem 2) and the extension of Case 3 gives the only proof known to the author of a case where S may be a priori non-precompact. In contrast to this, our method in the previous sections is impossible to extend to infinite dimensional cases, both because of its inductive nature and because of its use of the generalized Straszewics theorem.

The referee has pointed out that Klee's method in Case 1 works also if the norm function is replaced by an arbitrary continuous convex function that attains its minimum. Let f be such a function. Then

$$g(x) = \sup\{f(z - x): z \in S\} = f(q(x) - x) \tag{5.2}$$

together with the hypothesis that for each x there exists a unique such $q(x)$, serves to define this analog of the antiprojection function. By translation one may assume that $f(0) = 0 \leqq f(x)$ for all x. If S is compact it follows that q is continuous. and, as in [3] that it has a fixed point s, with $q(s) = s$. It then follows from (5.2) and the unicity that s is the only point in S. Thus, in particular, the method works as well for unsymmetric norms.

Klee's proof for Case 2, however, uses the homogeneity of the norm and so does not work for general convex function, although it will take a generally unsymmetric norm. The third case, again, has an appropriate extension to more general convex functions. We give here the corresponding generalization of our Theorem 3.

THEOREM 4. *If the* (*continuous*) *function f is defined on the Banach space E as the maximum* (*attained at each point of E*) *of a denumerable family of continuous affine functions, and f attains its minimum on E, then each subset S of E such that* (5.2) *has a unique so ution $q(x)$ for each x* in E (i.e. *each translate of f attains its maximum uniquely on S*), *consists of a single point.*

The proof of Theorem 4 consists of a repetition of the arguments in the proof of Theorem 3, and is omitted.

Again, it seems indicated by the success in Case 1 and Case 3 that the corresponding statement would be true, in finite dimensional spaces, for any continuous convex function that attains its minimum and for a priori arbitrary sets S. We have not been able to solve this and state it as an open problem.

PROBLEM. Suppose f is a continuous convex function on $\mathbb{R}^n$ that attains its minimum and that S is a subset of $\mathbb{R}^n$ such that each translate of f attains its maximum at a unique point in S. Must S consist of a single point?

REMARK. Some condition on f like attaining its minimum is needed, as shown by taking f to be a nonconstant linear function.

REFERENCES

1. E. Asplund, *A k-extreme point is the limit of k-exposed points.* Israel J. Math **1** (1963), 161–162.

2. W. Fenchel, *On conjugate convex functions.* Canad. J. Math. 1(1949), 73–77.

3. V. L. Klee, *Convexity of Chebyshev sets.* Math. Annalen **142** (1961), 292–304.

UNIVERSITY OF WASHINGTON,
SEATTLE, WASHINGTON