COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 538

ON THE TENDENCY TOWARD CONVEXITY OF THE VECTOR SUM OF SETS

Roger Howe

November 3, 1979

ON THE TENDENCY TOWARD CONVEXITY OF THE VECTOR SUM OF SETS*

by

Roger Howe

There are several instances ([Ni], p. 255, [H-K], p. 255) in mathematical economics where one is interested in the following question:

Given sets $\{X_i\}_{i=1}^n$ in a vector space V, how close is the vector sum

(1) $$\sum_{i=1}^n X_i = \{ \sum_{i=1}^n x_i : x_i \in X_i\}$$

to being convex?

The standard result on this topic is the Shapley-Folkman Theorem [H-K], [A-H]. In this note we offer some observations regarding the question, beginning with a proof of the Shapley-Folkman result and proceeding to various refinements.

As above, let $\{X_i\}_{i=1}^n$ be sets in a vector space V of dimension ℓ. We will assume that X_i are closed. (This is no loss of generality for Shapley-Folkman, and is convenient for our arguments.) Also for convenience and without loss of generality in what follows, we may assume that $0 \in X_i$ for all i. This is because if $x_i^0 \in X_i$, then $0 \in X_i - x_i^0$ and

*The research described in this paper was undertaken from grants from the National Science Foundation and from the Ford Foundation.

$$\sum_{i=1}^{n}(X_i - x_i^0) = \sum_{i=1}^{n} X_i - \sum_{i=1}^{n} x_i^0 .$$

So our assumption merely amounts to a translation of the whole situation.

Let $co(X_i)$ denote the convex hull of X_i . It is well known that $co(X_i)$ is closed. Further, the extreme points of $co(X_i)$ all belong to X_i . It is standard that

$$co(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} co(X_i) . \tag{2}$$

Thus any point z in $co(\sum_{i=1}^{n} X_i)$ can be written in the form

$$z = \sum_{i=1}^{n} x_i , \quad x_i \in co(X_i) . \tag{3}$$

Proposition 1 (Shapley-Folkman): In the representation (3) all but $\ell = \dim V$ of the x_i may be taken to belong to X_i (instead of only $co(X_i)$).

Proof. The proof proceeds by induction on n and ℓ . Put $Z^n = co(\sum_{i=1}^{n} X_i)$. A point $z \in Z^n$ can be written

$$z = y + x_n , \quad y \in Z^{n-1} , \quad x_n \in co(X_n) .$$

Since $0 \in X_n$, also $y \in Z^n$. Hence the whole line segment

$$y + tx_n , \quad 0 \leq t \leq 1$$

belongs to Z^n . There are two possibilities: either (i) $z \in Z^{n-1}$, or (ii) $z \notin Z^{n-1}$. In the first case, the result for Z^n follows from the result for Z^{n-1} . In the second case, there will be a largest t for which $y + tx_n$ belongs to Z^{n-1} . For this t , $y + tx_n = y'$ will be on the boundary ∂Z^{n-1} of Z^{n-1} . We may thus write

$$z = y' + x_n' , \quad y' \in \partial Z^{n-1} , \quad x' \in co(X_n) .$$

We come now to the crucial point. We may choose a supporting hyperplane to Z^{n-1} at y' . Thus, let λ be a linear functional on V such that

$$\max\{\lambda(u) : u \in Z^{n-1}\} = \lambda(y') .$$

Let

$$H(\lambda,s) = \{v \in V : \lambda(v) = s\}$$

be the set of parallel hyperplanes defined by λ . Set

$$s_i = \max\{\lambda(x_i) : x_i \in X_i\} .$$

Then by the way λ was chosen, we see that $\lambda(y') = \sum_{i=1}^{n-1} s_i$, and

$$H(\lambda, s_i) \cap co(X_i) = co(H(\lambda, s_i) \cap X_i)$$

and

$$Z^{n-1} \cap H(\lambda, \lambda(y')) = \sum_{i=1}^{n-1} (H(\lambda, s_i) \cap co(X_i)) . \tag{4}$$

See Figure 1 for a geometric picture of this situation when $n = 2$. Up to translation, we see that we are dealing with the situation of

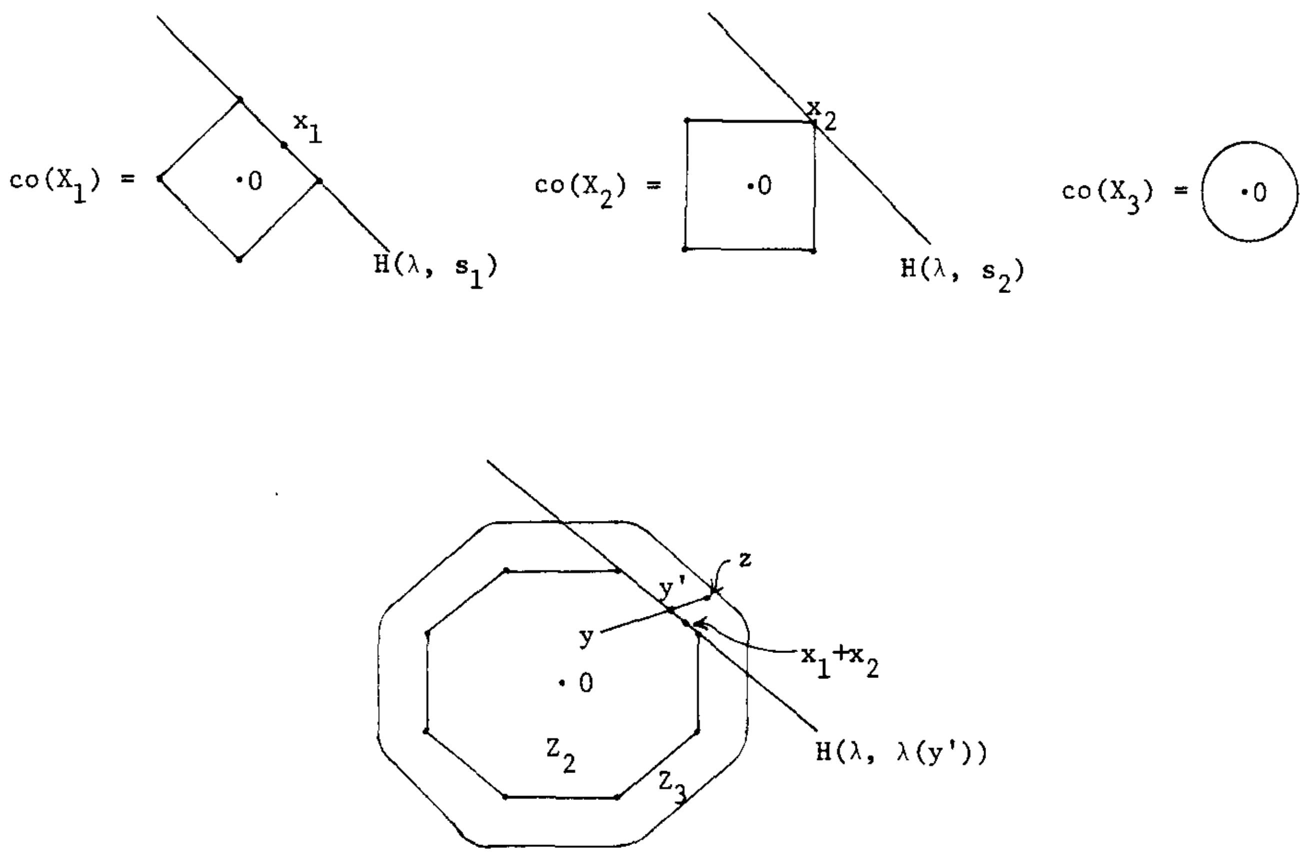

FIGURE 1

the theorem, but in dimension only $\ell-1$. Thus the result may be assumed for y' with $\ell-1$ in place of ℓ ; and then $z \in y' + x'_n$ is the desired representation for z .

We note a couple of corollaries. First, suppose that the only defect from convexity in the X_i is that they are hollow; precisely, assume that $\partial co(X_i) \subseteq X_i$. Then inductively one sees that $\partial Z^n \subseteq \sum_{i=1}^{n} X_i$. Hence, from proof we conclude

<u>Corollary 1.1</u>: If $\partial co(X_i) \subseteq X_i$ for all i , then an arbitrary $z \in Z^n = co(\sum_{i=1}^{n} X_i)$ can be written

$$z = y + x$$

where $y \in \sum_{i=1}^{n} X_i$ and $x \in co(X_j)$ for some j .

Corollary 1.1 obviously involves rather a special situation. However, the same basic aspect of Proposition 1 that it exploits can also be used in other ways. For example given a linear functional λ on V , let $e(\lambda)$ be the number of the sets X_i such that λ assumes its maximum on X_i at more than one point of X_i . For a given X_i , it can be shown that the typical λ (i.e., for λ in a set of full measure in V^* , the dual of V) assumes its maximum only once on X_i . Thus if the X_i are imagined to be chosen "at random," it is plausible that $e(\lambda)$ might be ≤ 1 for all λ . In any case, let us set

$$\tilde{e}(\{X_i^n\}_{i=1}) = \{\max e(\lambda) : \lambda \in V^*, \lambda \neq 0\} .$$

Then a repetition with slight changes of the proof of Proposition 1 above gives

Corollary 1.2: In the expansion (3), all but $\tilde{e}(\{X_i\}_{i=1}^n) + 1$ of the x_i's may be chosen to belong to X_i, rather than $co(X_i)$.

To refine Proposition 1 we introduce quantitative considerations. Suppose $\|\ \|$ is a norm on V. Define the diameter $\delta(X_i)$ of the sets X_i by

(5) $$\delta(X_i) = \sup\{\|x - x'\| : x, x' \in X_i\} .$$

Thus X_i is contained in the intersection of the balls of radius $\delta(X_i)$ centered at the points of X_i. It follows that

(6) $$\delta(co(X_i)) = \delta(X_i) .$$

If $v \in V$, and $X \subseteq V$ is a set, define the distance $d(v,X)$ from v to X by

(7) $$d(v,X) = \inf\{\|v-x\| : x \in X\} .$$

In terms of these definitions we immediately have the following "metrical" version of Proposition 1.

Corollary 1.3: For any $z \in co(\sum_{i=1}^n X_i)$, we have

$$d(z, \sum_{i=1}^n X_i) \leq (\dim V)(\max \delta(X_i)) .$$

To see that this estimate is of the correct order of magnitude, let $V = \mathbb{R}^\ell$ and let each of the X_i consist of the standard basis vectors together with 0. Let $\|\ \|$ be the 1-norm:

$$\|(x, \ldots, x_\ell)\|_1 = \sum_{j=1}^{\ell} |x_j| \ .$$

Then for very large n, precisely $n \geq \ell/2$, the vector

$$v = 1/2(1,1,1, \ldots, 1)$$

will be in Z^n, but the points of $\sum_{i=1}^{n} X_i$ all have integer coordinates. Thus for any x in $\sum_{i=1}^{n} X_i$, all the coordinates of $v-x$ are at least 1/2 in absolute value, whence $d(v, \sum_{i=1}^{n} X_i) \geq \ell/2$. On the other hand, $\delta(X_i) = 2$, so $d(v, \sum_{i=1}^{n} X_i) \geq 1/4\ \ell(\max \delta(X_i))$.

To get better results, we must make some further assumptions on the X_i. One obvious possibility is that some of the X_i have non-empty interiors. If this holds then $\sum_{i=1}^{n} X_i$ will tend to fill up the inside of its convex hull. Let $\nu(X_i)$ be the radius of the largest ball contained in X_i. Clearly

(8) $$\nu(\sum_{i=1}^{n} X_i) \geq \sum_{i=1}^{n} \nu(X_i) \ .$$

This fact lies at the base of our second result, which is also quite simple.

<u>Proposition 2</u>: Set

$$\max_i \delta(X_i) = \delta_0 \ .$$

Suppose we can find certain of the X_i, say $X_1, X_2, \ldots, X_m$, with $m \leq n$ such that $\sum_{i=1}^{m} \nu(X_i) \geq (\dim V)\delta_0$. Then $\sum_{i=1}^{n} X_i$ contains all points sufficiently far inside $Z^n = co(\sum_{i=1}^{n} X_i)$. Precisely, suppose $z \varepsilon Z^n$ and

$$d(z, \partial Z^n) \geq m\delta_0 .$$

Then $z \varepsilon \sum_{i=1}^{n} X_i$.

Proof. We may write

$$\sum_{i=1}^{n} X_i = (\sum_{i=1}^{m} X_i) + (\sum_{i=m+1}^{n} X_i) = Y_1 + Y_2 .$$

By our assumptions and (8), we have $\nu(Y_1) \geq \ell\delta_0$. Thus Y_1 contains a ball of radius $\ell\delta_0$. By a translation of Y_1, an operation which will not affect our argument, we may assume Y_1 contains the ball of radius $\ell\delta_0$ around the origin. Then also $\delta(Y_1) \leq m\delta_0$, so Y_1 is contained in the ball of radius $m\delta_0$ around the origin.

By Corollary 1.3, no point of $co(Y_2)$ is at distance greater than $\ell\delta_0$ from Y_2. Hence

$$co(Y_2) \subseteq Y_1 + Y_2 .$$

Further,

$$Z^n = co(Y_1 + Y_2) = co(Y_2) + co(Y_1) \subseteq co\ Y_2 + B_{m\delta_0}(0)$$

where $B_r(v)$ denotes the ball of radius r around v. It is intuitively clear from this that any point of Z^n at distance more than $m\delta_0$ from ∂Z^n must be in $co(Y_2)$, hence in $Y_1 + Y_2$. Here is the argument. Suppose $z \notin co(Y_2)$. Then we can find a linear functional

λ in V^*, the dual of V, separating z from $co(Y_2)$. We may assume (recalling that $co(Y_2)$ contains the origin)

$$\lambda(z) > \max\{\lambda(y) : y \in co(Y_2)\} \geq 0 .$$

By the Hahn-Banach Theorem, we can find $v \in B_{m\delta_0}(z)$ such that

$$\lambda(v) = \lambda(z) + \|\lambda\|^* m\delta_0$$

where $\|\lambda\|^*$ indicates the norm of λ in the norm on V^* dual to our given norm $\|\ \|$ on V. Thus clearly $v \notin Z^n$. Hence the distance from z to ∂Z^n is less than $m\delta_0$, as desired.

Thus, existence of interior points in the X_i causes $\sum_{i=1}^{n} X_i$ to fill up its convex hull. However, even if each X_i is the closure of its interior, there may be points in $co(\sum_{i=1}^{n} X_i)$ which stay a fixed distance away from $\sum_{i=1}^{n} X_i$, as the example in Figure 2 shows. Evidently the problem there is the angularity of X. Thus if we put some sort of smoothness condition on the X_i, we might expect to improve on Proposition 2. This is the case; here is one formulation of such a result.

We take $V = \mathbb{R}^{\ell}$ with the usual Euclidean norm.

<u>Proposition 3</u>: Suppose X_i is a union of balls of some fixed radius $r_i \geq 0$. Put $s = \sum_{i=1}^{n} r_i$, and put $\delta_0 = \max_i\{\delta(X_i) - 2r_i\}$. Assume $s \geq \ell\delta_0$. Then for any point $z \in Z^n = co(\sum_{i=1}^{n} X_i)$ we have

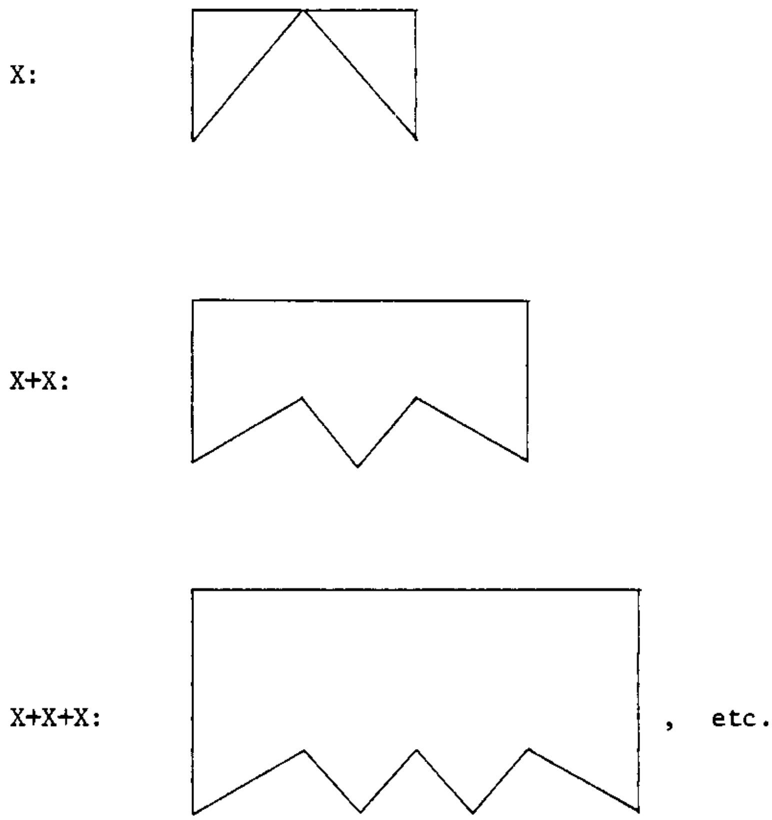

FIGURE 2

$$(9) \qquad d(z, \sum_{i=1}^{n} X_i) \leq \sqrt{s^2 + (\ell\delta_0)^2} - s = \frac{(\ell\delta_0)^2}{\sqrt{s^2 + (\ell\delta_0)^2} + s} \leq \frac{(\ell\delta_0)^2}{2s}$$

Remarks

a) Typically, it will take an infinite union of balls to fill out X_i . Thus, for example, if X_i is the interior of a smooth (twice continuously differentiable) hypersurface, then X_i will be the union of balls of some sufficiently small radius. What is important is that the radii be bounded away from zero.

b) By (9) we see that if a large number of the X_i are relatively rotund, so that s is large, then $\sum_{i=1}^{n} X_i$ fill most of Z^n , in the sense that all points of Z^n are fairly close to it.

Proof. Let $\tilde{X}_i$ be the set of centers of balls of radius r_i contained in X_i . Then

$$X_i = \tilde{X}_i + B_{r_i}(0)$$

and

$$\delta(X_i) = \delta(\tilde{X}_i) + 2r_i \ .$$

Hence

$$\sum_{i=1}^{n} X_i = \sum_{i=1}^{n} (\tilde{X}_i + B_{r_i}(0)) = \sum_{i=1}^{n} \tilde{X}_i + B_s \ .$$

Furthermore,

$$Z^n = \tilde{Z}^n + B_s \ .$$

Since we assume $s \geq \ell\delta_0$, we see by Corollary 1.3 that

$$\mathcal{Z}^n \subseteq \sum_{i=1}^{n} X_i .$$

Thus if $z \in Z^n - \sum_{i=1}^{n} X_i$, then in particular $z \in Z^n - \mathcal{Z}^n$. Let $\tilde{z}$ be the point of $\mathcal{Z}^n$ closest to z. Note $\tilde{z}$ is unique, since $\mathcal{Z}^n$ is convex and Euclidean balls are strictly convex. The hyperplane H passing through $\tilde{z}$ and perpendicular to $z - \tilde{z}$ will support $\mathcal{Z}^n$. As we saw in the proof of Lemma 1, specifically, if we combine Lemma 1 with the set equality (4), we can find y in $H \cap \sum_{i=1}^{n} \tilde{X}_i$, such that $\|y-\tilde{z}\| \leq \ell\delta_0$. (Actually, $(\ell-1)\delta_0$.) Since $y \in H$, we have that $y-\tilde{z}$ is perpendicular to $z-\tilde{z}$. Hence

$$\|z-y\|^2 = \|z-\tilde{z}\|^2 + \|y-\tilde{z}\|^2 \leq s^2 + (\ell\delta_0)^2 .$$

Since $B_s(y) \subseteq \sum_{i=1}^{n} X_i$, the inequality (9) follows.

To summarize, we have in Propositions 1, 2, and 3 the following progression. First if you take the vector sum of small sets, no point of the convex hull of the sum is far away from a point of the sum. Second, if an appreciable number of the summands have interior, then the sum tends to completely fill the inside of its convex hull. Third, if an appreciable number of the summands are relatively rotund, then the sum tends to fill almost all of its convex hull. To finish the discussion it seems appropriate to show that in some cases, even if the original X_i do not have interior, summing certain of them may create sets with interior, so that, after a partitioning of the sum

$\sum_{i=1}^{n} X_i$ into certain subsums, Proposition 2 or 3 may be applicable. First we remind the reader of the folklore fact that summing sets of positive measure creates sets with interior. Then we show that summing sets which are small, but still positive dimensional (e.g., sets containing continuous arcs) creates sets of positive measure. Finer results of this nature clearly exist, but these seem to exemplify fairly clearly our theme, the tendency to convexity of vector sums.

Suppose $X \subseteq V$ is a set of positive measure $\mu(X)$. A point $x \in X$ is called a <u>point of density</u> of X if

$$\lim_{r \to 0} \frac{\mu(X \cap B_r(x))}{\mu(B_r(x))} = 1 . \tag{10}$$

Here as before $B_r(x)$ is the ball of radius r around x in the given norm on V . A classical result [S] of Euclidean harmonic analysis guarantees that the set of points of X which are not points of density is of measure zero. This fact immediately implies the following known result:

<u>Proposition 4</u>: If X_1 and X_2 are sets of positive measure, then $X_1 + X_2$ has non-empty interior. More precisely, if $x_i \in X_i$ are points of density for $i = 1, 2$, then $X_1 + X_2$ contains a neighborhood of $x_1 + x_2$.

<u>Proof</u>. By a translation, we may just as well assume that $x_1 = x_2 = 0$. Then given $\varepsilon > 0$, we can find a radius $r > 0$ such that

$$\mu(X_i \cap B_r(0)) \geq (1-\varepsilon)\mu(B_r(0)) .$$

Choose $v \in V$, with $\|v\| = \alpha r$, with $0 < \alpha < 1$. Then

$$\begin{aligned}\mu(X_i \cap B_r(0) \cap B_r(v)) &\geq \mu(X_i \cap B_r(0)) - \mu(B_r(0) - B_r(v)) \\ &\geq (1-\varepsilon)\mu(B_r(0)) - (\mu(B_r(0)) - \mu(B_r(0) \cap B_r(v))) \\ &\geq -\varepsilon\mu(B_r(0)) + \mu(B_{(1-\alpha)r}(0)) \\ &= ((1-\alpha)^{\ell} - \varepsilon)\mu(B_r(0)) \\ \text{(where } \ell = \dim V\text{)} \quad &\geq ((1-\alpha)^{\ell} - \varepsilon)\mu(B_r(0) \cap B_r(v)) .\end{aligned}$$

We see that if α and ε are small enough, then $(1-\alpha)^{\ell} - \varepsilon$ is $> 1/2$. By symmetry, the same reasoning applies to $(X_i + v) \cap B_r(0) \cap B_r(v)$. Hence

$$X_1 \cap (X_2 + v) \cap B_r(0) \cap B_r(v)$$

is non-empty. Pick a point z in it. Then

$$z = x_1' = x_2' + v \quad \text{where} \quad x_i' \in X_i .$$

Thus

$$v = z - x_2' = x_1' - x_2' .$$

Repeating the above reasoning with $-X_2$ in place of X_2, which one checks is all right, yields $v \in X_1 + X_2$. Since v was arbitrary subject to being sufficiently close to 0, the proposition is proved.

Finally, we consider the situation when we have $\ell = \dim V$ continuous arcs inside certain of our sets. For purposes of the proposition, we may as well assume that the sets reduce to just the arcs.

Proposition 5: Let c_i , $0 \leq i \leq \ell = \dim V$ be ℓ continuous arcs in V . That is,

$$c_i : [0,1] \to V$$

are continuous maps. Assume for simplicity that $c_i(0) = 0$. Then the volume of the vector sum of the images of the c_i is at least as large as the volume of the parallopiped spanned by the points $c_i(1)$.

Proof. With no loss of generality we may assume that $V = \mathbb{R}^\ell$ and $c_i(1)$ is the i^{th} standard basis vector. We will in fact show that, for any point x in the unit cube $I^\ell = [0,1]^\ell \subseteq \mathbb{R}^\ell$, there is an integral translate $x+z$ of x , with $z \in \mathbb{Z}^\ell$ such that $x+z \in \sum_{i=1}^{\ell} \text{im } c_i = Y$. Here $\text{im } c_i$ denotes the image of the map c_i ; that is, $\text{im } c_i$ is the i^{th} arc. Then taking the intersections of of Y with the translates $I^\ell + z$ of the unit cube, and translating them by $-z$, we find they completely cover I^ℓ , and so have altogether measure at least 1, as desired.

Our assertion is equivalent to the statement that the translates $Y+z$, $z \in \mathbb{Z}^\ell$, cover $\mathbb{R}^\ell$. Define a mapping $C : \mathbb{R}^\ell \to \mathbb{R}^\ell$ by

$$C(t+z) = \sum_{i=1}^{\ell} c_i(t_i) + z , \quad t \in I^\ell , \quad z \in \mathbb{Z}^\ell .$$

It is easy to see that C is well-defined and continuous, and our assertion is seen to be equivalent to the statement that C is surjective.

The map C satisfies

$$C(x+z) = C(x) + z \ , \quad x \in \mathbb{R}^\ell \ , \quad z \in \mathbb{Z}^\ell \ .$$

In particular C is a mapping of bounded displacement, in the sense that $\|C(x) - x\| \leq M$, for some fixed number M. Mappings of bounded displacement of $\mathbb{R}^\ell$ to itself are known to be necessarily surjective. Here is why. Let ϕ be a continuous non-negative function on $\mathbb{R}$ such that $\phi(t) = 0$ for $t \leq 1$ and $\phi(t) = 1$ for $t \geq 2$. Pick a very large number r and define a map D_r on $\mathbb{R}^\ell$ by

$$D_r(x) = \left(1-\phi\left(\frac{\|x\|}{r}\right)\right)C(x) + \phi\left(\frac{\|x\|}{r}\right)x \ .$$

Then $D_r(x) = C(x)$ for $\|x\| \leq r$ and $D_r(x) = x$ for $x \geq 2r$. Also $D_r(x) - x = (1-\phi(\|x\|/r))(C(x)-x)$. Hence $\|D_r(x) - x\| \leq M$. Thus for $r > M$, D_r defines a map of B_{3r} to itself, and this map is the identity on the sphere ∂B_{3r}. By a well-known variant of the Brouwer fixed point theorem [Ni], D_r maps B_{3r} onto itself. In the limit as $r \to \infty$, we see that C maps $\mathbb{R}^\ell$ onto itself. This concludes Proposition 5.

Remark: The above proof combined with the Baire Category Theorem [Ry] fact implies that the set Y has non-empty interior.

REFERENCES

[A-H] K. Arrow and . Hahn, General Competitive Analysis. San Francisco: Holden Day, 1971.

[H-K] W. Hildenbrand and A. Kirman, Introduction to Equilibrium Analysis. Amsterdam: North-Holland Press, 1976.

[Ni] H. Nikaido, Convex Structures and Economic Theory. New York: Academic Press, 1968.

[Ry] H. Royden, Real Analysis. New York: Macmillan, 1963.

[S] E. Stein, Singular Integrals and Differentiability Properties of Functions. Princeton: Princeton University Press, 1970.