Reprinted from JOURNAL OF FUNCTIONAL ANALYSIS

Vol. 52, No. 3, July 1983
Printed in Belgium

Approximation by a Sum of Two Algebras. The Lightning Bolt Principle

DONALD E. MARSHALL*

Department of Mathematics, University of Washington, Seattle, Washington 98195

AND

ANTHONY G. O'FARRELL

Department of Mathematics, St. Patrick's College, Maynooth, County Kildare, Ireland

Communicated by the Editors

Received October 1982; revised May 3, 1983

The measures on a compact Hausdorff space X orthogonal to the sum $A_1 + A_2$ of two subalgebras of $C_{\mathbb{R}}(X)$, the real-valued continuous functions on X, are described. From this description, a geometric condition equivalent to the density of $C(x_1) + C(x_2)$ in $C_{\mathbb{R}}(X)$ is obtained, where $X \subset \mathbb{R}^2$ and where $C(x_j)$ denotes the continuous functions depending only on the jth coordinate function.

We are interested in finite sums of algebras of real-valued continuous functions. For example, if $X \subset \mathbb{R}^n$ is compact we seek a description of the linear space

$$C(x_1) + \cdots + C(x_n)$$

as a subset of $C_{\mathbb{R}}(X)$, the real-valued continuous functions on X. Here $x_1,..., x_n$ denote the coordinate functions on $\mathbb{R}^n$ and $C(x_j)$ the space of real-valued continuous functions that depend only on the jth coordinate function. Problems connected with this linear space have arisen in a number of contexts.

Perhaps the most celebrated result on sums of algebras is Kolmogorov's solution of Hilbert's thirteenth problem [14]. Expanding on the solution by Arnold [1, 2], Kolmogorov embedded the unit cube of $\mathbb{R}^n$ into the unit cube in $\mathbb{R}^{2n+1}$, $n \geqslant 2$, in such a way that on the image all continuous functions

* Work supported in part by an NSF grant.

belong to the space $C(y_1) + \cdots + C(y_{2n+1})$, where $y_1, \ldots, y_{2n+1}$ are the coordinate functions in $\mathbb{R}^{2n+1}$, and each y_j in turn is in $C(x_1) + \cdots + C(x_n)$. Writing

$$y_j = \sum_{k=1}^{n} \varphi_{k,j}(x_k), \qquad j = 1, \ldots, 2n+1,$$

it has been observed that each $\varphi_{k,j}$ is monotonic and Lipschitz continuous. Kahane [13] proved that almost any choice of monotonic, Lipschitz $\varphi_{k,j}$ will work. Vitushkin [23, 24], (see also Henkin [12]), however, proved that $\{\varphi_{k,j}\}$ cannot be continuously differentiable. See Vitushkin [22], for an excellent survey on this subject.

We are looking at this problem from the point of view of functional analysis. What geometric conditions on the image X of the unit cube of $\mathbb{R}^n$ guarantee that the linear space $C(y_1) + \cdots + C(y_{2n+1})$ is uniformly dense in $C_{\mathbb{R}}(X)$? The closure of $C(y_1) + \cdots + C(y_{2n+1})$ can, of course, be described in terms of the measures on X orthogonal to this space. While we cannot give a complete answer to this problem, we can describe the measures on a compact set $X \subset \mathbb{R}^2$ orthogonal to $C(x_1) + C(x_2)$ (see Theorem 2). From this description, we obtain a geometric condition equivalent to the density of $C(x_1) + C(x_2)$. Earlier [16] we described the measures orthogonal to $C(x_1) + C(x_2)$ provided orbits of X were closed. Havinson [11] used this description to characterize the compact sets $X \subset \mathbb{R}^2$ with

$$C(x_1) + C(x_2) = C_{\mathbb{R}}(X).$$

A similar necessary and sufficient condition was obtained by Sternfeld [21], who also obtained a necessary condition for $C(x_1) + \cdots + C(x_n)$ to equal $C_{\mathbb{R}}(X)$. Sproston and Strauss [20] have obtained a sufficient, but not necessary, condition for $C(x_1) + \cdots + C(x_n)$ to equal $C_{\mathbb{R}}(X)$.

The problem of approximating by elements of a sum of algebras has arisen in other contexts. Buck [5] studied the classical functional equation: Given $k, \beta \in C_{\mathbb{R}}[0,1]$, $\|k\|_\infty \leqslant 1$, $\|\beta\|_\infty \leqslant 1$, for which $u \in C_{\mathbb{R}}[0,1]$ does there exist $\varphi \in C_{\mathbb{R}}[0,1]$ such that $\varphi(x) = k(x)\,\varphi(\beta(x)) + u(x)$? The solution of this problem when $\|k\|_\infty < 1$ is classical and follows from a standard fixed-point theorem. The case when $k = 1$ is central to the study of when $\|k\|_\infty = 1$ (see Kuczma [15]). Buck proved that the set of all u, for which there is a $\varphi \in C_{\mathbb{R}}[0,1]$ with

$$\varphi(x) = \varphi(\beta(x)) + u(x), \tag{1}$$

is dense in $\{u \in C_{\mathbb{R}}[0,1] : u(x) = 0 \text{ whenever } \beta(x) = x\}$ if and only if $C(x_1) + C(x_2)$ is dense in $C_{\mathbb{R}}(K)$, where $K = \{(x_1, x_2) : x_2 = x_1 \text{ or } x_2 = \beta(x_1),\ 0 \leqslant x_1 \leqslant 1\}$. Corollary 2 gives a geometric means of deciding when $C(x_1) + C(x_2)$ is dense in $C_{\mathbb{R}}(K)$.

This type of approximation has also arisen, for example, in connection with the numerical solution of certain elliptic p.d.e. boundary value problems, e.g., Bank [3], the tabulation of functions of several variables, e.g., Diliberto and Straus [6], inner derivations on a C^* algebra, e.g., Sproston [19], and dimension theory, e.g., Sternfeld [21].

Approximation by polynomials in the *complex* variable $x_1 + ix_2$ usually deals with the topological and metric character of a set $X \subset \mathbb{R}^2$. However, the type of approximation we consider is more geometric in nature. For example, if X is the six point set $\{a_1 = (0,0),\ a_2 = (1,0),\ a_3 = (1,2),\ a_4 = (2,2),\ a_5 = (2,1),\ a_6 = (0,1)\}$, then $C(x_1) + C(x_2)$ is not dense in $C_{\mathbb{R}}(X)$. Indeed, if δ_a denotes point mass at a, the measure $\delta_{a_1} - \delta_{a_2} + \delta_{a_3} - \delta_{a_4} + \delta_{a_5} - \delta_{a_6}$ has zero mass on each vertical and horizontal line and hence is orthogonal to $C(x_1) + C(x_2)$. However, if we rotate our set slightly, the function x_1 will be one-to-one and hence $C(x_1) = C_{\mathbb{R}}(X)$. Consequently, a simple rotation drastically alters the answer to our density problem.

There are two ideas in the proof of our main result, Theorem 2. We call them the ergodic method and the stochastic method. For clarity of exposition we discuss in Section 2 a special case (where we obtain more information) in which the ergodic method alone suffices. In Section 3, we then combine this method with the stochastic method to prove our result. Essentially our approach is dynamical. A norm 1 annihilating measure for $C(x_1) + C(x_2)$ determines a discrete stochastic "flow" on X with two-dimensional "time," where the generators of the dynamical semigroup do not commute. The measure is extreme if and only if the flow is ergodic, in a reasonable sense. This is equivalent to the ergodicity of a single one-parameter subflow. In Section 1 we state our main definitions and theorem. In Section 4 we generalize our results to the sum of two subalgebras, $A_1 + A_2$, contained in $C_{\mathbb{R}}(X)$, where X is a compact Hausdorff space. Finally we indicate the difficulties with the sum of more than two algebras.

1. Bolts of Lightning

We begin with a definition. Let $\pi_1(x_1, x_2) = x_1$ and $\pi_2(x_1, x_2) = x_2$ denote the coordinate projections in $\mathbb{R}^2$. A *bolt of lightning* is a sequence $(a_1, a_2, a_3, \ldots)$ of points in $\mathbb{R}^2$ with either $\pi_1(a_{2n-1}) = \pi_1(a_{2n})$ and $\pi_2(a_{2n}) = \pi_2(a_{2n+1})$, for all $n = 1, 2, 3, \ldots$, or $\pi_2(a_{2n-1}) = \pi_2(a_{2n})$ and $\pi_1(a_{2n}) = \pi_1(a_{2n+1})$, for all $n = 1, 2, 3, \ldots$. The first case we will call a type I bolt and the second a type II bolt. A finite lightning bolt $(a_1, \ldots, a_n)$ with $a_k \neq a_{k+1}$, $k = 1, \ldots, n-1$, and $a_n = a_1$ is said to be a *closed* bolt. These objects have appeared, independently, in a number of papers, e.g., [1, 6, 8, 16–18] with several different names. The term *bolt of lightning* is the most common and problem [1, 2]. We define an equivalence relation on a compact subset X of $\mathbb{R}^2$ by: $a \sim b$ if and only if a and b belong to some (finite) lightning bolt contained in X. We call the equivalence classes *orbits*. Earlier [16], we proved the following result:

THEOREM 1. *If X is a compact subset of $\mathbb{R}^2$ with all its orbits closed, then $C(x_1) + C(x_2)$ is dense in $C_{\mathbb{R}}(X)$ if and only if X contains no closed lightning bolt.*

Bolts are explicit objects and give geometric means of deciding if $C(x_1) + C(x_2)$ is dense. For example, if X is the set $\{(0,0), (0,1), (-1,1), (-1,-\frac{1}{2}), (\frac{1}{2},-\frac{1}{2}), (\frac{1}{2},\frac{1}{4}), (-\frac{1}{4},\frac{1}{4}), (-\frac{1}{4},-\frac{1}{8}), (\frac{1}{8},-\frac{1}{8}),...\}$, then the theorem shows $C(x_1) + C(x_2)$ is dense. If X consists of three line segments, sufficiently long, then X contains a bolt with 6 distinct points and hence $C(x_1) + C(x_2)$ is not dense. However, if X consists of two parallel line segments, not parallel to a coordinate axis, then the theorem shows $C(x_1) + C(x_2)$ is dense. A more nontrivial example was constructed by Havinson [10] in his study of best approximation by elements of $C(x_1) + C(x_2)$. One version of his example can be described as follows. The set X will be a limit of sets X_n. Let X_1 consist of 4 disjoint line segments L_1, L_2, L_3, L_4 with slope 1 such that $\pi_1(L_2) = \pi_1(L_1)$, $\pi_1(L_3) = \pi_1(L_4)$, $\pi_2(L_1) = \pi_2(L_4)$, and $\pi_2(L_2) = \pi_2(L_3)$. To construct X_n from X_{n-1}, rotate one segment in X_{n-1} 90° about its center, then remove the middle one-third from each line segment. Clearly X_n has $4 \cdot 2^{n-1}$ line segments and every orbit in X_n consists of $4 \cdot 2^{n-1}$ points, one in each line segment. Hence all orbits in X are dense, so the theorem above does not apply. There are no closed bolts, but $C(x_1) + C(x_2)$ is not dense. Indeed, on X_n we can construct a measure μ_n orthogonal to $C(x_1) + C(x_2)$ by placing linear measure or (-1) times linear measure on each segment, normalized to have total variation 1. We can easily do this so that the sign of μ_n on a line segment L is the is the same sign as μ_{n+1} on the two line segments in X_{n+1} formed from L. If μ is a weak* cluster point of $\{\mu_n\}$, then μ is a nontrivial measure on X orthogonal to $C(x_1) + C(x_2)$.

The idea behind the proof of Theorem 1 is that each closed bolt $(a_1,..., a_n)$ in X determines a measure $\delta_{a_1} - \delta_{a_2} + \cdots + (-1)^n \delta_{a_{n-1}}$ which is orthogonal to $C(x_1) + C(x_2)$. On the other hand, if μ is an extreme point of the unit ball of the measures orthogonal to $C(x_1) + C(x_2)$, then μ is supported on a single orbit, since the orbits are closed. Out of this orbit we construct a closed lightning bolt. The curious part is that alternating-point-mass measure associated with this bolt does not have to be used in the measure μ. This follows from Proposition 1.

Let $(C(x_1) + C(x_2))_X^{\perp}$ denote the measures on X orthogonal to $C(x_1) + C(x_2)$ and let $\text{Ext}_X(\text{ball}(C(x_1) + C(x_2))^{\perp})$ denote the extreme points of the unit ball of $(C(x_1) + C(x_2))_X^{\perp}$.

ı a compact subset X of
ıe (finite) lightning bolt
ırbits. Earlier [16], we

ith all its orbits closed,
if X contains no closed

ıs of deciding if $C(x_1)+$
$0, 1), (-1, 1), (-1, -\frac{1}{2}),$
theorem shows $C(x_1)+$
sufficiently long, then X
$_1)+C(x_2)$ is not dense.
, not parallel to a coor-
$x_2)$ is dense. A more
0] in his study of best
sion of his example can
ets X_n. Let X_1 consist of
e 1 such that $\pi_1(L_2)=$
$=\pi_2(L_3)$. To construct
its center, then remove
ly X_n has $4 \cdot 2^{n-1}$ line
points, one in each line
heorem above does not
is not dense. Indeed, on
$(x_1)+C(x_2)$ by placing
segment, normalized to
the sign of μ_n on a line
line segments in X_{n+1}
}, then μ is a nontrivial

h closed bolt $(a_1, \ldots, a_n)$
$_{n-1}$ which is orthogonal
ne point of the unit ball
is supported on a single
we construct a closed
ng-point-mass measure
in the measure μ. This

orthogonal to $C(x_1)+$
extreme points of the

PROPOSITION 1. *Suppose X is a compact subset of the unit square S in $\mathbb{R}^2$ and suppose $\mu \in \mathrm{Ext}_X(\mathrm{ball}(C(x_1)+C(x_2))^\perp)$. Then $\mu \in \mathrm{Ext}_S(\mathrm{ball}(C(x_1)+C(x_2))^\perp)$.*

Proof. Let μ be supported on X and orthogonal to $C(x_1)+C(x_2)$ with $\|\mu\|=1$. If $\mu \notin \mathrm{Ext}_S(\mathrm{ball}(C(x_1)+C(x_2))^\perp)$, then

$$\mu = \alpha\nu + \beta\lambda,$$

where $0<\alpha,\ \beta<1,\ \alpha+\beta=1$, and $\nu,\ \lambda \in \mathrm{ball}(C(x_1)+C(x_2))_S^\perp$. Since the restrictions of ν and λ to X must have norm 1, they must be supported on X. Thus $\mu \notin \mathrm{Ext}_X(\mathrm{ball}(C(x_1)+C(x_2))^\perp)$.

In particular, we may take X to be Havinson's example described above. Then, as in the proof of Theorem 1, each $\mu \in \mathrm{Ext}_X(\mathrm{ball}(C(x_1)+C(x_2))^\perp)$ determines a closed lightning bolt in S, yet μ is singular to the measure determined by this bolt. The bolt, in fact, does not even lie in the support of μ. Hereafter we will fix a compact set X and all measures and bolts will be carried by X.

If $\mathbf{b}=(a_1, a_2, a_3, \ldots)$ is an infinitely long bolt of lightning, let $\mu_{n,\mathbf{b}}=(\delta_{a_1}-\delta_{a_2}+\cdots+(-1)^{n+1}\delta_{a_n})/n$. We say a measure μ is *generated* by the bolt $\mathbf{b}$ if $\mu_{n,\mathbf{b}}$ converges (weak*) to μ. In other words,

$$\lim_{n\to\infty} \int f\, d\mu_{n,\mathbf{b}} = \int f\, d\mu$$

for all $f \in C_{\mathbb{R}}(X)$.

THEOREM 2. *Each $\mu \in \mathrm{Ext}(\mathrm{ball}(C(x_1)+C(x_2))^\perp)$ is generated by some bolt in X.*

Clearly if μ is a weak* cluster point of $\{\mu_{n,\mathbf{b}}\}$ for some bolt $\mathbf{b}$, then $\mu \in (C(x_1)+C(x_2))^\perp$. By the Krein–Milman theorem it is easy to deduce

COROLLARY 1. *$C(x_1)+C(x_2)$ is dense in $C_{\mathbb{R}}(X)$ if and only if $\mu_{n,\mathbf{b}}$ converges to 0 (weak-*) for each bolt $\mathbf{b}$.*

Corollary 1 gives, for example, a geometric method for deciding if density occurs in the problem mentioned below (1), in the following sense. If U is an open set in $\mathbb{R}^2$, you can see how frequently the even numbered vertices in a bolt return to U, and hence if there is a nontrivial weak* cluster-point of $\mu_{n,\mathbf{b}}$.

COROLLARY 2. *If $f \in C_{\mathbb{R}}(X)$, then*

$$\mathrm{dist}(f, C(x_1)+C(x_2)) = \max\left\{\limsup \left|\int f\, d\mu_{n,\mathbf{b}}\right| : \mathbf{b} \text{ is a bolt}\right\}$$

Diliberto and Straus [6] proved a version of Corollary 2 in the special case when X is the unit square in $\mathbb{R}^2$, in connection with a problem posed by the RAND Corporation. Their version has "max" replaced by "sup." The proof in [6] requires, also, that the functions $g(x_1)=\frac{1}{2}(\max_{x_2} f(x_1,x_2)+\min_{x_2} f(x_1,x_2))$ and $h(x_2)=\frac{1}{2}(\max_{x_1} f(x_1,x_2)+\min_{x_1} f(x_1,x_2))$ be continuous for all $f\in C_{\mathbb{R}}(X)$. This is not always true for compact sets $X\subset\mathbb{R}^2$, as is the case in the example immediately following the statement of Theorem 1. The difficulty arises when X is not a product set. Overdeck [18] proved Corollary 2 when X is a certain kind of Jordan curve. Havinson [10] proved it for arbitrary compact sets $X\subset\mathbb{R}^2$ with the added hypothesis that f has a closest function in $C(x_1)+C(x_2)$. We can modify Havinson's example, described above, to show that it is possible for $C(x_1)+C(x_2)$ to not be dense in $C_{\mathbb{R}}(X)$ and yet there is an $f\in C_{\mathbb{R}}(X)$ that does not have a closest function in $C(x_1)+C(x_2)$. We simply begin with the 4 line segments in X_1, with L_2 and L_4 meeting at one point x_0. Arrange all rotations in the construction of each X_n to involve only segments outside some small neighborhood of x_0. By Havinson's theorem [10], each $f\in C_{\mathbb{R}}(X)$ which is not in $C(x_1)+C(x_2)$, does not have a closest function in $C(x_1)+C(x_2)$. We conclude that our corollary cannot be deduced from Havinson's theorem.

We include one final point of clarification. Of course if $\mu\in(C(x_1)+C(x_2))^{\perp}$, $\|\mu\|=1$ and $f\in C_{\mathbb{R}}(X)$, one can choose rational numbers r_n and $z_n\in X$ such that

$$\left|\int f\,d\mu-\sum_{n=1}^{N} r_n f(z_n)\right|<\varepsilon.$$

Repeating each $f(z_n)$ as many times as necessary, we may suppose $r_n=\pm 1/N$. However, one cannot assume, a priori, that $\sum_{n=1}^{N} r_n\, g(z_n)=0$ for all $g\in C(x_1)+C(x_2)$ or even that $\{z_n\}$ are the vertices of a bolt. This is the gap in the proof in Golomb [9]. (We mention this because it has also been cited as a proof elsewhere in the literature.)

2. The Ergodic Method

The special case we begin with is the "deterministic case." We say $X\subset\mathbb{R}^2$ has small fibers if $\pi_j^{-1}(\pi_j(a))$ consists of at most two points for all $a\in X$ and $j=1,2$. The example due to Havinson, described above, has small fibers. We define two bijections φ_j on X by $\varphi_j(a)=b$ if $\pi_j(a)=\pi_j(b)$ and $a\neq b$, or if $\pi_j^{-1}(\pi_j(a))=\{a\}$, then $\varphi_j(a)=a$. These maps "switch" the points in each fiber. They map X one-to-one and onto X although they might not be continuous.

LEMMA 1. *Each φ_j is a Borel map.*

Proof. Without loss of generality we assume $\varphi = \varphi_1$ and $\pi = \pi_1$. Let $X_n = \{a \in X : |a - b| \geqslant 1/n$ for some $b \in X$ with $\pi(a) = \pi(b)\}$. Since π is continuous and X is compact, X_n is closed. Let $\varphi_n(a) = b$ if $a \in X_n$ and $\pi(a) = \pi(b)$ with $a \neq b$ and let $\varphi_n(a) = a$ if $a \in X \backslash X_n$. Now if E is an open set in X and diameter $E < 1/2n$, then

$$\varphi_n^{-1}(E) = (X_n \cap \bar{E}^c \cap \pi^{-1}(\pi(E))) \cup (E \cap (X \backslash X_n)),$$

where $\bar{E}^c$ is the complement of the closure of E. We conclude φ_n is Borel and since $\varphi(x) = \lim_{n \to \infty} \varphi_n(x)$, φ is also Borel.

Let $\mathscr{M}(X) = C(X)^*$ denote the space of regular Borel measures on X. For a Borel function ψ, mapping X into $\mathbb{R}^1$ or $\mathbb{R}^2$ and a measure $\mu \in \mathscr{M}(X)$ we define $\psi_\# \mu \in \mathscr{M}(X)$ by

$$\psi_\# \mu(E) = \mu(\psi^{-1}(E))$$

for all Borel sets $E \subset \psi(X)$. Obviously $\mu \in (C(x_1) + C(x_2))^\perp$ if and only if $(\varphi_j)_\# \mu = -\mu$ for $j = 1$ and 2. For a positive measure $\mu \in \mathscr{M}(X)$ and a Borel map ψ of X into X, we say that ψ is μ-ergodic if

(a) $\psi_\# \mu = \mu$, and

(b) if E is Borel and $\psi(E) \subset E$, then $\mu(E) = 0$ or $\mu(E) = \mu(X)$.

In other words $\psi_\#$ preserves μ and the only ψ invariant sets either support μ or are μ-null. We let $T = \varphi_2 \circ \varphi_1$. Note that if $\mu \in (C(x_1) + C(x_2))^\perp$, then $T_\# \mu = \mu$.

LEMMA 2. *If $\mu \in \operatorname{ball}(C(x_1) + C(x_2))^\perp$, then μ is extreme (in this set) if and only if T is μ_+-ergodic, where $\mu = \mu_+ - \mu_-$ is the Haar decomposition of μ.*

Proof. Suppose μ is extreme. Since $(\varphi_j)_\# \mu_+ = \mu_-$ and $(\varphi_j)_\# \mu_- = \mu_+$, we conclude $T_\# \mu_+ = \mu_+$ and $T_\# \mu_- = \mu_-$. Suppose T is not μ_+ ergodic. Then there exist disjoint T-invariant Borel sets E and, F of positive μ_+ measure with $X = E \cup F$. Indeed, we can take E_1 Borel with $T(E_1) \subset E_1$ and $0 < \mu_+(E_1) < \mu_+(X)$. Let $E = \{a \in X : T^n a \in E_1$ for some $n\} = \bigcup_{n=0} T^{-n}(E_1)$, a Borel set. Since $T_\# \mu = \mu$ and since $T^{-n}(E_1) \subset T^{-(n+1)}(E_1)$, $\mu_+(E) = \mu_+(E_1)$. Let $F = X \backslash E$.

Let $\lambda_+ = \mu_+|_E$ and $\nu_+ = \mu_+|_F$ be the restriction measures. Define $\lambda_- = (\varphi_1)_\# \lambda_+$, $\nu_- = (\varphi_1)_\# \nu_+$, $\lambda = \lambda_+ - \lambda_-$, and $\nu = \nu_+ - \nu_-$. Since $T_\# \lambda_+ = \lambda_+$ and $T_\# \nu_+ = \nu_+$, a short computation shows $(\varphi_j)_\#(\lambda) = -\lambda$ and $(\varphi_j)_\# \nu = -\nu$ for $j = 1$, 2, and $\mu = \lambda + \nu$. Thus λ, $\nu \in (C(x_1) + C(x_2))^\perp$, and $1 = \|\mu\| = 2\|\mu_+\| = 2\|\lambda_+\| + 2\|\nu_+\| = \|\lambda\| + \|\nu\|$. So we may write $\mu = \|\lambda\| (\lambda/\|\lambda\|) +$

Conversely suppose T is μ_+-ergodic. Suppose μ is not extreme, and let $\mu = \alpha\lambda + (1-\alpha)\nu$, where $0 < \alpha < 1$, $\lambda \neq \nu$ and λ, $\nu \in (C(x_1) + C(x_2))^\perp$, $\|\lambda\| = \|\nu\| = 1$. Clearly $\mu_+ = \alpha\lambda_+ + (1-\alpha)\nu_+$ and $\|\lambda_+\| = \|\nu_+\| = \frac{1}{2}$. Also λ_+ and ν_+ must be absolutely continuous with respect to μ_+ and hence there exists nonnegative functions f_+ and g_+ in $L^1(\mu_+)$ such that

$$\lambda_+ = f_+ \mu_+ \qquad \text{and} \qquad \nu_+ = g_+ \mu_+$$

so that $\alpha f_+ + (1-\alpha) g_+ = 1$, μ_+ a.e. Since T preserves λ_+ and ν_+, we have

$$f_+ \circ T = f_+ \qquad \text{and} \qquad g_+ \circ T = g_+ \quad \text{a.e. } d\mu_+.$$

Let $E = \{a \in X: f_+(a) = g_+(a)\}$. If $\mu_+(E) > 0$, then since E is T-invariant, it has full measure and we conclude $\lambda_+ = \nu_+$. Thus $\lambda_- = (\varphi_j)_\# \lambda_+ = (\varphi_j)_\# \nu_+ = \nu_-$, contradicting $\lambda \neq \nu$. We conclude $\mu_+(E) = 0$. A similar argument shows exactly one of the sets $\{a \in X: f_a(a) > g_+(a)\}$, $\{x: f_+(a) < g_+(a)\}$ has full μ_+ measure. Without loss of generality, suppose that $f_+(a) > g_+(a)$ a.e. $d\mu_+$. Then

$$\tfrac{1}{2} = \int 1\, d\mu_+ = \alpha \int (f_+ - g_+)\, d\mu_+ + \int g_+\, d\mu_+ > \int 1\, d\nu_+ = \tfrac{1}{2}.$$

This contradiction shows μ is extreme.

The small fibers assumption implies that there are at most two bolts starting at each point of X, a type I bolt and a type II bolt.

THEOREM 3. *Suppose X has small fibers and then $\mu \in \mathscr{M}(X)$. $\mu \in \text{Ext}(\text{ball}(C(x_1) + C(x_2))^\perp)$ if and only if for μ_+ almost all $a \in X$, each bolt starting at a generates μ.*

Of course, by symmetry μ_- almost all points initiate a bolt generating $-\mu$ if μ is extreme, and conversely. We remark that it is not necessarily the case that all bolts generate μ. The point x_0 in the modification of Havison's example given after Corollary 2 has only the trivial bolt $\{x_0\}$ beginning at x_0.

Proof. Suppose μ is extreme. Then T is μ_+ ergodic by Lemma 2. By Birkhoff's ergodic theorem, e.g. [25], for each $f \in C_{\mathbb{R}}(X)$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(T^k a) = 2 \int f\, d\mu_+ \tag{3}$$

for μ_+ almost all $a \in X$. Since $C_{\mathbb{R}}(X)$ is separable, we can select a set G of full μ_+ measure such that (3) holds for all $a \in G$ and all g in a dense subset of $C_{\mathbb{R}}(X)$. By continuity (3) holds for all $a \in G$ and all $f \in C_{\mathbb{R}}(X)$. Clearly

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(\varphi_1 T^k a) = 2 \int f \circ \varphi_1 \, d\mu_+ = 2 \int f \, d\mu_- .$$

We conclude that for μ_+ almost all $a \in X$ the bolt $(a, \varphi_1 a, Ta, \varphi_1 Ta, T^2 a, \varphi_1 T^2 a, \ldots)$ generates μ. Clearly the same conclusion holds with T replaced by $T^{-1} = \varphi_1 \circ \varphi_2$.

Conversely suppose $\mu_{n,\mathbf{b}}$ converges weak* to μ for μ_+ almost all $a \in X$, where $\mathbf{b}$ is the bolt $\mathbf{b} = (a, \varphi_1 a, Ta, \varphi_1 Ta, \ldots)$. Clearly $\mu \in (C(x_1) + C(x_2))^{\perp}$ and so $T_{\#}\mu = \mu$. Suppose T is not μ_+-ergodic. As in the proof of Lemma 2 select a Borel set E with $0 < \mu_+(E) < \frac{1}{2}$ and $T^{-1}(E) = E$. The following argument was pointed out to us by Doug Lind, for which we are grateful. (See Walters [25, Corollaries 1.5, 1.6, p. 36] for another version.) Take $g \in C_{\mathbb{R}}(X)$ such that $g = \chi_E$ except on a set of μ measure less than ε and $\|g\|_\infty = 1$. Then by the dominated convergence theorem

$$\int g(a) \left[\int g(c) \, d(\mu_{n,\mathbf{b}})_+(c) - \int g(c) \, d\mu_+(c) \right] d\mu_+(a) \tag{4}$$

converges to 0 as $n \to \infty$. We may rewrite (4) when $n = 2m$ as

$$\frac{1}{2m} \sum_{k=1}^{m} \int g(a) \, g(T^k a) \, d\mu_+(a) - \left[\int g(a) \, d\mu_+(a) \right]^2 . \tag{5}$$

The set where $g(a) g(T^k a) \neq \chi_E(a) \chi_E(T^k a)$ has μ_+ measure less than 2ε, since $T_{\#}\mu_+ = \mu_+$. Since $a \in E$ if and only if $T^k a \in E$, (5) is within 4ε of

$$\frac{1}{2m} \sum_{k=1}^{m} \mu_+(E) - (\mu_+(E))^2 = \frac{1}{2} \mu_+(E) - (\mu_+(E))^2 .$$

We conclude $\frac{1}{2}\mu_+(E) = (\mu_+(E))^2$ and hence $\mu_+(E) = 0$ or $\mu_+(E) = \frac{1}{2} = \|\mu_+\|$. This contradiction implies T is μ_+-ergodic.

3. The stochastic Method

The general situation discussed in this section could be reduced to the small fibers case if it could be shown that each extreme annihilator for $C(x_1) + C(x_2)$ is supported on a compact set with small fibers. Unfortunately, this is not the case, as the following example shows. Let $\alpha < \frac{1}{2}$ be a positive irrational number and let X consist of the five line segments in $\mathbb{R}^2$,

$+ 2)\colon 0 \leqslant t \leqslant 1/\sqrt{2}\}$, $L_4 = \{(t+2, ((1-\alpha)/\sqrt{2}) + t)\colon 0 \leqslant t \leqslant \alpha/\sqrt{2}\}$, $L_5 = \{(t + 2 + (\alpha/\sqrt{2}), t)\colon 0 \leqslant t \leqslant (1-\alpha)/\sqrt{2}\}$. The points $(2+\alpha/\sqrt{2}, 0)$, $(2+\alpha/\sqrt{2}, 1)$, and $(2+\alpha/\sqrt{2}, 2+\alpha/\sqrt{2})$ all lie in one vertical fiber in X. If μ_n denotes arc length on the nth line segment, $\mu_1 - \mu_2 + \mu_3 - \mu_4 - \mu_5$ is orthogonal to $C(x_1) + C(x_2)$. Every bolt that does not meet the ends of L_4 or L_5, on the other hand, is dense in X. To see this suppose we start a type I bolt x_0 units along L_1, after 4 steps we return to L_1 at x_1 units along L_1, where $x_1 - x_0 + \alpha$ is an integer. Thus after $4n$ steps we return to L_1 at x_n units along L_1, where $x_n - x_0 + n\alpha$ is an integer. Since $\{[n\alpha]\colon n = 1, 2,...\}$ is dense in $[0, 1]$, where $[t]$ denotes the fractional part of t, the bolt must reach a dense set of points on L_1 and hence on X. By Theorem 2, there must be an extreme annihilating measure on X whose support does not have small fibers.

When $X \subset \mathbb{R}^2$ no longer has small fibers, we work instead on the space of bolts. Let $\mathscr{B} = \{\mathbf{b} = (a_1, a_2,...)\colon \mathbf{b}$ is a bolt$\}$ and give $\mathscr{B}$ the topology inherited from the product topology on the compact set $\prod_{n=1}^{\infty} X$. We prove the following version of Theorem 2, which clearly implies Theorem 2.

THEOREM 4. *To each* $\mu \in \mathrm{Ext}(\mathrm{ball}(C(x_1) + C(x_2))^{\perp})$ *there is associated a probability measure* $\mathscr{P}$ *on* $\mathscr{B}$ *such that* $\mathscr{P}$ *almost all type* I *bolts generate* μ.

We begin with Lemma 3. Let K and Y be compact Hausdorff spaces and let $\{\varphi_n\}$ be a collection of continuous maps from K into Y. Let $\mathscr{B} = \{(a_1, a_2, a_3,...) \in \prod_{n=1}^{\infty} K\colon \varphi_n(a_n) = \varphi_n(a_{n+1}),\quad n = 1, 2, 3,...\}$. Finally let $\{\mu_n\}_{n=1}^{\infty}$ be a collection of probability measures on K. As before $(\varphi)_{\#}\mu(E) = \mu(\varphi^{-1}(E))$.

LEMMA 3. *If* $(\varphi_n)_{\#}\mu_n = (\varphi_n)_{\#}\mu_{n+1}$, $n = 1, 2,...$, *then there is a probability measure* $\mathscr{P}$ *on* $\prod_{n=1}^{\infty} K$, *supported on* $\mathscr{B}$, *with marginal distribution* μ_n. *That is,*

$$\int_{\mathscr{B}} f(a_n)\, d\mathscr{P}(a_1, a_2,...) = \int_K f(x)\, d\mu_n(x), \qquad n = 1, 2,...,$$

for all $f \in C_{\mathbb{R}}(X)$.

Proof. Let $\mathscr{B}_K$ denote the Borel subsets of K and let $\mathscr{B}_Y$ denote the Borel subsets of Y. Define a map U_n from $L^1(\mathscr{B}_K, \mu_n)$ into $L^1(\mathscr{B}_K, \mu_{n+1})$ as follows. For $h \in L^{\infty}(\mathscr{B}_K, \mu_n)$ and $g \in L^1(\mathscr{B}_Y, (\varphi_n)_{\#}\mu_n)$ define a linear functional Λ_h by

$$\Lambda_h(g) = \int_K hg \circ \varphi_n\, d\mu_n. \tag{6}$$

Since Λ_h is bounded and linear on $L^1(\mathscr{B}_Y, (\varphi_n)_\# \mu_n)$, there is a unique $k \in L^\infty(\mathscr{B}_Y, (\varphi_n)_\# \mu_n)$ with

$$\Lambda_h(g) = \int_Y (kg)(\varphi_n)_\# \mu_n. \tag{7}$$

Note that

$$\int_K |k \circ \varphi_n| \, d\mu_{n+1} = \int_Y |k| \, (\varphi_n)_\# \mu_{n+1} = \int_Y |k| \, (\varphi_n)_\# \mu_n$$

$$= \int_K h \left(\frac{\bar{k}}{|k|} \circ \varphi_n \right) d\mu_n \leqslant \int_K |h| \, d\mu_n.$$

Define $U_n(h) = k \circ \varphi_n$. Since $\| U_n h \|_{L^1(\mathscr{B}_K, \mu_{n+1})} \leqslant \| h \|_{L^1(\mathscr{B}_K, \mu_{n+1})}$ for h in the dense subspace $L^\infty(\mathscr{B}_K, \mu_n)$, we can extend U_n to be a norm reducing linear map from $L^1(\mathscr{B}_K, \mu_n)$ into $L^1(\mathscr{B}_K, \mu_{n+1})$ as desired. Note that $\int_K h \, d\mu_n = \int_K U_n h \, d\mu_{n+1}$ by (6) and (7) with $g \equiv 1$, since $(\varphi_n)_\# \mu_n = (\varphi_n)_\# \mu_{n+1}$. Furthermore if $h \geqslant 0$, then $U_n h \geqslant 0$.

The map U_n is an averaging operator, and can be best understood in terms of disintegration of measures. We may write [4]

$$\mu_n(y) = \mu_n^x(y) \, \sigma_n(x),$$

where μ_n^x is a measure on the fiber $\varphi_n^{-1}(x)$ and σ_n is a measure on the space of fibers. Then $h \in L^1(\mathscr{B}_K, \mu_n)$ is replaced by a constant on each fiber, namely, $\int_{\varphi_n^{-1}(x)} h \, d\mu_n^x$, the conditional expectation. Since $(\varphi_n)_\# \mu_n = (\varphi_n)_\# \mu_{n+1}$, we can write $\mu_{n+1}(y) = \mu_{n+1}^x(y) \, \sigma_n(x)$, and so this defines $U_n(h)$ on each fiber for μ_{n+1}. To make this latter approach rigorous we encountered a number of difficult measure-theoretic problems and for that reason we adopt the former approach.

We will now define $\mathscr{P}$ on sets of the form $E_1 \times E_2 \times \cdots \times E_n \times \prod_{n=1}^\infty K$ by

$$\mathscr{P}\left(E_1 \times E_2 \times \cdots \times E_n \times \prod_{n=1}^\infty K\right)$$

$$= \int_K U_n(\chi_{E_n} U_{n-1}(\chi_{E_{n-1}} \cdots \chi_{E_2} U_1(\chi_{E_1}) \cdots) \, d\mu_{n+1}.$$

$\mathscr{P}$ is well defined on these sets because

$$\mathscr{P}\left(E_1 \times E_2 \times \cdots + E_n \times K \times \prod_{n=1}^\infty K\right) = \int_K U_{n+1}(\chi_K U_n(\chi_{E_n} \cdots)) \, d\mu_{n+2}$$

and it is additive on disjoint products. We may extend $\mathscr{P}$ as a bounded regular additive set function on the field Σ consisting of all finite disjoint unions of sets of the form $E_1 \times \cdots \times E_n \times \prod_{n+1}^{\infty} K$, $n = 1, 2, \ldots$. By Theorem III.5.14 Dunford and Schwartz [7, p. 138], $\mathscr{P}$ has a unique regular, countably additive extension to the σ field determined by Σ; call it $\mathscr{P}$ also.

Now let L be compact, $L \subset \prod_{n=1}^{\infty} K \backslash \mathscr{B}$. Let π_k be the projection from $\prod_{n=1}^{\infty} K$ onto the kth component, for $k = 1, 2, 3, \ldots$. We may cover L by a finite number of open sets E_α, where for some k

$$\varphi_k \pi_k(E_\alpha) \cap \varphi_k \pi_{k-1}(E_\alpha) = \varnothing$$

and

$$\pi_j(E_\alpha) = K \qquad \text{for} \quad j \neq k, \quad j \neq k-1.$$

Since $\chi_{\pi_k(E_\alpha)} U_k(\chi_{\pi_{k-1}(E_\alpha)}) = 0$ a.e. $d\mu_k$, $\mathscr{P}(E_\alpha) = 0$. Hence $\mathscr{P}(L) = 0$. Since $\mathscr{P}$ is regular, $\mathscr{P}$ is supported on $\mathscr{B}$. It is elementary to check that $\int_{\mathscr{B}} f(a_n)\, d\mathscr{P}(a_1, a_2, \ldots) = \int_K f\, d\mu_n$ as desired. Clearly $\mathscr{P} \geqslant 0$ since $U_n \geqslant 0$, and $\int 1 d\mathscr{P} = \int_K d\mu_1 = 1$, so $\mathscr{P}$ is a probability measure.

We apply this lemma with $K = X \subset \mathbb{R}^2$, $\varphi_{2n} = \pi_2$, $\varphi_{2n-1} = \pi_1$, $n = 1, 2, \ldots$, $Y = \pi_1(X) \cup \pi_2(X)$, and $\mu_{2n-1} = 2\mu_+$, $\mu_{2n} = 2\mu_-$, $n = 1, 2, \ldots$, where $\mu = \mu_+ - \mu_- \in \operatorname{Ext}(\operatorname{ball}(C(x_1) + C(x_2))^{\perp})$. Note that $\mathscr{B}$ will be the space of bolts of type I. Define a transformation T of $\prod_{n=1}^{\infty} X$ onto $\prod_{n=1}^{\infty} X$ by

$$T(a_1, a_2, a_3, \ldots) = (a_3, a_4, \ldots).$$

Note that T maps $\mathscr{B}$ onto $\mathscr{B}$. The map T is continuous for if E is a basic open set of the form $E = (E_1 \times E_2 \times \cdots \times E_n \times \prod_{k=n+1}^{\infty} X)$, then $T^{-1}(E) = X \times X \times E_1 \times E_2 \times \cdots \times E_n \times \prod_{k=n+1}^{\infty} X$. A $\mathscr{P}$-ergodic map is defined exactly as in Section 2.

LEMMA 4. *T is $\mathscr{P}$-ergodic.*

Proof. We first show T is measure preserving. Indeed, by the uniqueness of $\mathscr{P}$, it is enough to check it on sets of the form $E = E_1 \times E_2 \times \cdots \times E_n \times \prod_{n+1} X$, where $E_1, \ldots, E_n$ are Borel subsets of X. Note that $U_{2k} = U_2$ and $U_{2k-1} = U_1$ for $k = 1, 2, \ldots$, where U_k is the map defined in the proof of Lemma 3.

$$\mathscr{P}(T^{-1}(E)) = \mathscr{P}\left(X \times X \times E_1 \times \cdots \times E_n \times \prod_{n+3} X\right)$$

$$= \int_K U_{n+2}(\chi_{E_n} U_{n+1}(\cdots U_3(\chi_{E_1} U_2(\chi_X U_1(\chi_X) \cdots)))) \, d\mu_{n+2}$$

extend $\mathscr{P}$ as a bounded
ting of all finite disjoint
$[{}^{\infty}_{n+1} K$, $n = 1, 2, \ldots$. By
$\mathscr{P}$ has a unique regular,
ed by Σ; call it $\mathscr{P}$ also.
be the projection from
. We may cover L by a

$- 1$.

Hence $\mathscr{P}(L) = 0$. Since
mentary to check that
ly $\mathscr{P} \geqslant 0$ since $U_n \geqslant 0$,
sure.
$\varphi_{2n-1} = \pi_1$, $n = 1, 2, \ldots$,
$n = 1, 2, \ldots$, where $\mu =$
will be the space of bolts
o $\prod_{n=1}^{\infty} X$ by

uous for if E is a basic
$_{=n+1} X)$, then $T^{-1}(E) =$
godic map is defined

ndeed, by the uniqueness
$= E_1 \times E_2 \times \cdots \times E_n \times$
Note that $U_{2k} = U_2$ and
defined in the proof of

$$= \int_K U_n(\chi_{E_n}(U_{n-1} \cdots U_1(\chi_{E_1}) \cdots)) \, d\mu_n$$

$$= \mathscr{P}(E)$$

since $\mu_{n+2} = \mu_n$, and $U_2(\chi_X U_1(\chi_X)) = \chi_X$.

Now suppose E is Borel and $T(E) \subset E$, $0 < \mathscr{P}(E) < 1$. As in the proof of Lemma 2, we may suppose $T^{-1}(E) = E$. For $f \in C_{\mathbb{R}}(X)$ let $L_n(f) = \int_{\mathscr{B}} \chi_E(a_1, a_2, a_3, \ldots) f(a_n) \, d\mathscr{P}$. Clearly L_n is a bounded, linear, positive functional on $C_{\mathbb{R}}(X)$ and hence $L_n(f) = \int_X f \, d\sigma_n$ for some positive measure σ_n. Since $T_{\#}\mathscr{P} = \mathscr{P}$ and $T^{-1}(E) = E$ it is easy to check that $\sigma_{2n} = \sigma_2$ and $\sigma_{2n-1} = \sigma_1$, $n = 1, 2, \ldots$. Thus $\int_X f d(\sigma_1 - \sigma_2) = \int_{\mathscr{B}} [f(a_1) - f(a_2)] \, \chi_E \, d\mathscr{P} = \int_{\mathscr{B}} [f(a_3) - f(a_2)] \, \chi_E \, d\mathscr{P}$. Now if $f \in C(x_1)$ and $(a_1, a_2, \ldots) \in \mathscr{B}$, then $f(a_1) = f(a_2)$ and hence $d\sigma_1 - d\sigma_2 \in C(x_1)^{\perp}$. Likewise, if $f \in C(x_2)$ and $(a_1, a_2, \ldots) \in \mathscr{B}$, then $f(a_3) = f(a_2)$ and hence $d\sigma_1 - d\sigma_2 \in C(x_2)^{\perp}$. Thus $d\sigma_1 - d\sigma_2 \in (C(x_1) + C(x_2))^{\perp}$. Since for $f \in C_{\mathbb{R}}(X)$,

$$\int_X f(2d\mu_+ - d\sigma_1) = \int_{\mathscr{B}} \chi_{E^c} f(a_{2n-1}) \, d\mathscr{P}$$

and

$$\int_X f(2d\mu_- - d\sigma_2) = \int_{\mathscr{B}} \chi_{E^c} f(a_{2n}) \, d\mathscr{P}$$

we see that $\mu_+ - (\sigma_1/2)$ and $\mu_- - (\sigma_2/2)$ are also positive measures. Clearly $\mu_+ - (\sigma_1/2) - (\mu_- - (\sigma_2/2)) \in (C(x_1) + C(x_2))^{\perp}$. Hence if $\lambda = (\sigma_1 - \sigma_2)/2$

$$\mu = \|\lambda\| \frac{\lambda}{\|\lambda\|} + \|\mu - \lambda\| \frac{(\mu - \lambda)}{\|\mu - \lambda\|}$$

with

$$\|\lambda\| + \|\mu - \lambda\| = \int \frac{\sigma_1}{2} + \frac{\sigma_2}{2} + \int \left(\mu_+ - \frac{\sigma_1}{2}\right) + \int \left(\mu_- - \frac{\sigma_2}{2}\right) = 1.$$

Since μ is extreme, either $\sigma_1 = 0$ or $\sigma_1 = 2\mu_+$. But then $\mathscr{P}(E) = \int \chi_E \, d\mathscr{P} = \int d\sigma_1 = 0$ or 1. This contradiction establishes Lemma 4.

Now if $f \in C_{\mathbb{R}}(X)$, let $F(a_1, a_2, a_3, \ldots) = f(a_1)$. By Birkhoff's ergodic theorem

$$\lim \frac{1}{\ldots} \sum^{n} F \circ T^k(\mathbf{b}) = \int F \, d\mathscr{P}$$

for $\mathscr{P}$-almost all type I bolts $\mathbf{b}$. By Lemma 3, this may be rewritten as

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(a_{2k-1}) = 2\int_X f\, d\mu_+$$

for $\mathscr{P}$ almost type I bolts $\mathbf{b} = (a_1, a_2,...)$. Likewise

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(a_{2k}) = 2\int_X f\, d\mu_-$$

for $\mathscr{P}$ almost all type I bolts $\mathbf{b} = (a_1, a_2,...)$. Since $C_{\mathbb{R}}(X)$ is separable, as in the proof of Theorem 3, we conclude that $\mathscr{P}$-almost all type I bolts generate μ.

4. Subalgebras of $C_{\mathbb{R}}(X)$

Now suppose X is a compact Hausdorff space and that A_1 and A_2 are (real) subalgebras of $C_{\mathbb{R}}(X)$, the real-valued continuous functions on X, that together separate the points of X. We define equivalence relations $\underset{j}{\sim}$ by: $a \underset{j}{\sim} b$ if and only if $f(a) = f(b)$ for all $f \in A_j$, $l = 1, 2$. Let X_j denote the equivalence classes and let π_j be the usual quotient map of X onto X_j. A bolt of lightning $\mathbf{b}$, the associated measures $\mu_{n,\mathbf{b}}$ and the space of bolts $\mathscr{B}$ are defined as in Section 1.

THEOREM 5. *If X is a compact metric space, then to each $\mu \in$ Ext(ball$(A_1 + A_2)^\perp$) there is associated a probability measure $\mathscr{P}$ on $\mathscr{B}$ such that $\mathscr{P}$ almost all type* I *bolts generate μ.*

The proof of Theorem 5 is formally the same as the proof of Theorem 4. The only place where we require X to be a metric space is in the very last argument in the proof of Theorem 4, where we use that $C_{\mathbb{R}}(X)$ is separable. We can avoid this problem in the next two corollaries.

COROLLARY 3. *If X is a compact Hausdorff space, then $A_1 + A_2$ is dense in $C_{\mathbb{R}}(X)$ if and only if $\mu_{n,\mathbf{b}}$ converges weak* to 0 for each bolt $\mathbf{b}$.*

Proof. Indeed if $A_1 + A_2$ is not dense, there is a measure $\mu \in$ Ext(ball$(A_1 + A_2)^\perp$) and an $f \in C_{\mathbb{R}}(X)$ with $\int_X f\, d\mu \neq 0$. By the proof of Theorem 5, we can find a bolt $\mathbf{b}$ such that

$$\lim \int_X f\, d\mu_{n,\mathbf{b}} = \int_X f\, d\mu.$$

But if σ is any weak* cluster point of $\{\mu_{n,\mathbf{b}}\}$, then $\sigma \in (A_1 + A_2)^{\perp}$ and $\int f\, d\sigma = \int f\, d\mu \neq 0$. Thus $\{\mu_{n,\mathbf{b}}\}$ does not converge weak* to 0.

COROLLARY 4. *If $f \in C_{\mathbb{R}}(X)$, where X is a compact Hausdorff space, then*

$$\operatorname{dist}(f, A_1 + A_2) = \max \left\{ \limsup_{n \to \infty} \left| \int f\, d\mu_{n,\mathbf{b}} \right| : \mathbf{b} \text{ is a bolt} \right\}.$$

Finally, we mention that the description of measures orthogonal to $C(x_1) + \cdots + C(x_n)$ on a compact set $X \subset \mathbb{R}^n$, $n \geqslant 3$, seems to be beyond the scope of the methods discussed herein. A bolt of lightning beginning at a point $a_1 \in X \subset \mathbb{R}^2$ is built by first "cancelling" δ_{a_1} in one direction with $-\delta_{a_2}$, then cancelling $-\delta_{a_2}$ with a point mass in the other direction, etc. In $\mathbb{R}^3$, there are many directions in which we need to cancel δ_{a_1}. There does not seem to be a reasonable description of a sequence of points $\mathbf{b} = (a_1, a_2, ...)$ and weighted point mass measures $\mu_{n,\mathbf{b}}$ which generate extreme μ's such that any weak* cluster point of $\mu_{n,\mathbf{b}}$ is orthogonal to $C(x_1) + \cdots + C(x_n)$, $n \geqslant 3$. There is one such attempt in Diliberto and Straus [6], but there is a gap in the proof. Extreme annihilating measures are probably generated by trees, as used by Arnold in his proof of Hilbert's thirteenth problem, but these grow too fast for any sort of averaging process to work.

After this work was completed, the following article came to our attention: W. A. Light and E. W. Cheney. On the approximation of a bivariate function by the sum of univariate functions, *J. Approx. Theory* **29** (1980), 305–322. It contains another proof of the Diliberto and Straus theorem, as well as a number of references to related work.

ACKNOWLEDGMENTS

We wish to thank D. Lind for a useful discussion and to thank V. P. Havin and S. Hruscev for providing many of the references we have included.

REFERENCES

1. V. I. ARNOLD, On functions of three variables, *Dokl. Akad. Nauk. SSSR* **114** (1957), 679–681; English transl. *Amer. Math. Soc. Transl.* **28** (1963), 51–54.
2. V. I. ARNOLD, Representation of continuous functions of three variables by the superposition of continuous functions of two variables, *Mat. Sb. (N.S.)* **48** (1959), 3–74; English transl. *Amer. Math. Soc. Transl.* **28** (1963), 61–147.
3. R. E. BANK, An automatic scaling procedure for a D'Yakanov–Gunn iteration scheme,

4. N. Bourbaki, "Éléments de Mathématique," Livre VI, Integration, Chap. IX, Hermann, Paris 1969.
5. R. C. Buck, On approximation theory and functional equations, *J. Approx. Theory* 5 (1972), 228–237.
6. S. P. Diliberto and E. G. Straus, On the approximation of a function of several variables by the sum of functions of fewer variables, *Pacific J. Math.* **1** (1951), 195–210.
7. N. Dunford and J. Schwartz, "Linear Operators," Interscience, New York, 1958.
8. S. D. Fisher, The decomposition of $C_r(K)$ into the direct sum of subalgebras, *J. Funct. Anal.* **31** (1979), 218–223.
9. M. Golomb, Approximation of functions of fewer variables, *in* "On Numerical Approximation" (R. Langer, Ed.), especially pp. 312–313, Univ. Wisconsin Press, Madison, 1959.
10. S. Ja. Havinson, A Chebyshev theorem for the approximation of a function of two variables by sums of the type $\varphi(x) + \psi(y)$, *Izv. Akad. Nauk. SSSR Ser. Mat.* **33** (1969), 650–666; English transl. *Math. USSR-Izv.* **3** (1969), 617–632.
11. S. Ja. Havinson, private communication.
12. G. M. Henkin, On linear superposition of continuously differentiable functions, *Dokl. Akad. Nauk. SSSR* **157** (1964), 288–290; *Soviet Math. Dokl.* 5 (1964), 948–950.
13. J. P. Kahane, Sur le théorème de superposition de Kolmogorov, *J. Approx. Theory* **13** (1975), 229–234.
14. A. N. Kolmogorov, On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition, *Dokl. Akad. Nauk. SSSR* **114** (1957), 953–956; English transl. *Amer. Math. Soc. Transl.* **28** (1963), 55–59.
15. M. Kuczma, "Functional Equations," Polish Acad. Mono., No. 46, Warsaw, 1968.
16. D. E. Marshall and A. G. O'Farrell, Uniform approximation by real functions, *Fund. Math.* **104** (1979), 203–211.
17. Ju. P. Ofman, Best approximation of functions of two variables by functions of the form $\varphi(x) + \psi(y)$, *Izv. Akad. Nauk. SSSR Ser. Mat.* **25** (1961), 239–252; English transl. *Amer. Math. Soc. Transl.* **44** (1965), 12–28.
18. J. M. Overdeck, "Characterizations of Extreme Functionals on Function Spaces," Thesis, Univ. of Wisconsin, Madison, 1971.
19. J. P. Sproston, Derivations on some (possibly non-separable) C^* algebras, *Glasgow Math. J.* **22** (1981), 43–56.
20. J. P. Sproston and D. Strauss, Sums of subalgebras of $C(X)$, to appear.
21. Y. Sternfeld, Uniformly separating families of functions, *Israel J. Math.* **29** (1978), 61–91.
22. A. G. Vitushkin, "On Representation of Functions by Means of Superpositions and Related Topics," L'Enseignement Mathématique, Génève, 1978.
23. A. G. Vitushkin, Proof of the existence of analytic functions of several variables not representable by linear superpositions of continuously differentiable functions of fewer variables, *Dokl. Akad. Nauk. SSSR* **156** (1964), 1258–1261; *Soviet Math. Dokl.* 5 (1964), 793–796.
24. A. G. Vitushkin, Some properties of linear superposition of smooth functions, *Dokl. Akad. Nauk. SSSR* **156** (1964), 1003–1006; *Soviet Math. Dokl.* 5 (1964), 741–744.
25. P. Walters, "Ergodic Theory," Lecture Notes in Math., No. 458, Springer-Verlag, New York, 1975.

Printed by the St. Catherine Press Ltd., Tempelhof 41, Bruges, Belgium