Real Analysis Exchange Vol. 12 (1986-87)

R. M. Shortt, Department of Mathematics, Wesleyan University, Middletown, CT 06457.

THE SINGULARITY OF EXTREMAL MEASURES

0. Introduction.

Let λ be Lebesgue measure on R. A Borel measure μ on $I \times I$ is doubly-stochastic if $\mu(A \times I) = \mu(I \times A) = \lambda(A)$ for each Borel set $A \subseteq I$. The collection of all doubly-stochastic measures forms a convex, weakly compact set whose extreme points have been much studied: [2], [3], [4], [5], [6]. It was shown by Lindenstrauss [5] that every extreme doubly-stochastic measure is singular with respect to planar Lebesgue measure λ^2. It is our purpose to strengthen this result in a general context.

For example, suppose that $L_1,\ldots,L_m$ are lines through the origin in R^2 and that ν is a probability measure on R^2. Then one can consider the convex set of probabilities on R^2 whose projections onto $L_1,\ldots,L_m$ agree with those of ν. Theorem 2.1 infra will say that the extreme points of this set are singular with respect to Lebesgue product measure, no matter what the choice of ν! In the doubly-stochastic case, $m = 2$, L_1 and L_2 are the co-ordinate axes, and ν may be taken as λ^2 restricted to $I \times I$.

1. Preliminary results

A σ-algebra A of subsets of X is countably generated (c.g.)

if there is a sequence $A_1\ A_2\ \ldots$ of subsets of X such that $\mathcal{A}$ is the smallest σ-algebra containing the sets in the sequence. An $\mathcal{A}$-atom is a set A in $\mathcal{A}$ such that for any set $A_0 \subseteq A$ in $\mathcal{A}$ either $A_0 = A$ or $A_0 = \phi$. The σ-algebra $\mathcal{A}$ is atomic if X is a union of $\mathcal{A}$-atoms. If $\mathcal{A}$ is c.g., then $\mathcal{A}$ is atomic. The notation $\mathcal{B}(R^n)$ indicates the (c.g.) Borel σ-algebra on R^n.

1.1 Lemma: Let $\mathcal{A}$ and $\mathcal{A}_0$ be c.g. sub-σ-algebras of $\mathcal{B}(R^n)$ with the same atoms. Then $\mathcal{A} = \mathcal{A}_0$.

Indication: This is the so-called "strong Blackwell property" for R^n. See, for example, Proposition 6 on p. 21 of [1].

Suppose that $\mathcal{A}_1,\ldots,\mathcal{A}_m$ are sub-σ-algebras of $\mathcal{B}(R^n)$ and that μ is a Borel probability measure on R^n. Define $E(\mathcal{A}_1,\ldots,\mathcal{A}_m;\mu)$ to be the set of all Borel probabilities ν on R^n such that $\nu(A) = \mu(A)$ for each A in $\mathcal{A}_1 \cup\ldots\cup \mathcal{A}_m$. We assume that no $\mathcal{A}_i$ is one of the trivial σ-algebras $\{\phi,R^n\}$ or $\mathcal{B}(R^n)$. So $E = E(\mathcal{A}_1,\ldots,\mathcal{A}_m;\mu)$ is a convex set of measures containing μ.

Given $\mathcal{A}_1,\ldots,\mathcal{A}_m$ we let F be the linear space of all functions of the form $f_1+\cdots+f_m$, where $f_1,\ldots,f_m$ are bounded real functions on R^m which are respectively $\mathcal{A}_1,\ldots,\mathcal{A}_m$-measurable. Then a Borel probability ν belongs to E if and only if

$$\int f\, d\nu = \int f\, d\mu \quad \text{for all} \quad f \in F.$$

The extreme points of E are characterized in

1.2 Theorem (Douglas-Lindenstrauss): A Borel probability ν is an extreme point of $E(\mathcal{A}_1,\ldots,\mathcal{A}_m;\nu)$ if and only if F is dense in $L^1(\nu)$.

Indication: See Douglas [3:p. 243]. A special case is given in Lindenstrauss [5:p. 379].

We will prove that in the cases that occur naturally and geometrically, the extreme points of $E(\mathcal{A}_1,\ldots,\mathcal{A}_m;\mu)$ are singular with respect to n-dimensional Lebesgue measure λ^n.

A σ-algebra $\mathcal{A}$ of subsets of $\mathbf{R}^n$ is affine-invariant if $A \in \mathcal{A}$ implies $\alpha A + v \in \mathcal{A}$ for each non-zero scalar $\alpha \in \mathbf{R}$ and vector $v \in \mathbf{R}^n$. Let $f : \mathbf{R}^n \longrightarrow \mathbf{R}^m$ be Borel measurable. We say that f generates the sub-σ-algebra $\mathcal{A} \subseteq \mathcal{B}(\mathbf{R}^n)$, where $\mathcal{A} = \{f^{-1}(B) : B \in \mathcal{B}(\mathbf{R}^m)\}$.

1.3 Lemma: Let $\mathcal{A}$ be a c.g. sub-σ-algebra of $\mathcal{B}(\mathbf{R}^n)$. The following are equivalent:

1) $\mathcal{A}$ is generated by an orthogonal projection $T : \mathbf{R}^n \to \mathbf{R}^n$.

2) $\mathcal{A}$ is generated by a linear transformation $T : \mathbf{R}^n \to \mathbf{R}^n$.

3) $\mathcal{A}$ is affine-invariant.

Proof: $1 \Longrightarrow 2$: Immediate.

$2 \Longrightarrow 3$: If $T : \mathbf{R}^n \longrightarrow \mathbf{R}^n$ is linear, then for each scalar $\alpha \neq 0$ and $v \in \mathbf{R}^n$, we have $\alpha T^{-1}(B) + v = T^{-1}(\alpha B + Tv)$. So T generates an affine-invariant σ-algebra.

<u>3⟹1</u>: Let $K \subseteq R^n$ be the $\mathcal{A}$-atom containing the vector 0. For each $\alpha \neq 0$, the set αK is an $\mathcal{A}$-atom containing 0, so that $\alpha K = K$. Likewise, if $x \in K$, then $x + K$ is an $\mathcal{A}$-atom containing x, so that $x + K = K$. So K is a linear subspace of R^n.

Let $T : R^n \to R^n$ be orthogonal projection onto the orthocomplement $L = K^{\perp}$ and let T generate the σ-algebra $\mathcal{A}_0 \subseteq \mathcal{B}(R^n)$. Then $\mathcal{A}$ and $\mathcal{A}_0$ are c.g. sub-σ-algebras of $\mathcal{B}(R^n)$ with the same atoms. By lemma 1.1, $\mathcal{A}_0 = \mathcal{A}$.

Q.E.D.

The following geometric construction will facilitate the use of the Lebesgue density lemma in Theorem 2.1.

Let $|A|$ be the cardinality of the set A.

<u>1.4 Lemma</u>: Let $L_1,\ldots,L_m$ be non-trivial vector subspaces of R^n and let $\pi_i : R^n \to R^n$ be orthogonal projection onto L_i, $i = 1,\ldots,m$. Then there is a subset S of R^n such that

$$\sum_{i=1}^{m} |\pi_i(S)| < |S|.$$

<u>Proof</u>: Let $v_1,\ldots,v_m$ be unit vectors taken from the respective orthocomplements $L_1^{\perp},\ldots,L_m^{\perp}$. Let S_1 be a set of $m + 1$ points in R^n such that $|\pi_1(S_1)| = 1$. Let $d_1 = \mathrm{diam}(S_1)$. Put

$$S_2 = \bigcup_{k=0}^{m} (S_1 + 2kd_1v_2).$$

Then S_2 contains $(m + 1)^2$ points, $|\pi_1(S_2)| = (m + 1)|\pi_1(S_1)| = (m + 1)$, and $|\pi_2(S_2)| = |\pi_2(S_1)| \leq |S_1| = m + 1$.

In general, we suppose that S_p $(p < m)$ has been defined as a set of $(m+1)^p$ elements such that

$$|\pi_i(S_p)| \leq (m+1)^{p-1} \qquad i = 1,\ldots,p.$$

Let $d_p = \mathrm{diam}(S_p)$ and put

$$S_{p+1} = \bigcup_{k=0}^{m} (S_p + 2kd_pv_{p+1}).$$

Then S_{p+1} has $(m+1)^{p+1}$ elements, and

$$|\pi_i(S_{p+1})| \leq (m+1)(m+1)^{p-1} \qquad i = 1,\ldots,p$$

$$|\pi_{p+1}(S_{p+1})| \leq |S_p| = (m+1)^p$$

as desired. Finally, we take $S = S_m$ and check

$$\sum_{i=1}^{m}|\pi_i(S)| \leq \sum_{i=1}^{m} (m + 1)^{m-1} = m(m + 1)^{m-1} < (m + 1)^m = |S|.$$

Q.E.D.

2. <u>The main thoerem.</u>

Let $V_n(r)$ be the volume of a ball of radius r in R^n. Then

$$V_n(r) = \frac{\pi^{n/2}\, r^n}{\Gamma(1+n/2)}$$

is homogeneous of order n in the variable r.

2.1 Theorem: Let $A_1,\ldots,A_m$ be non-trivial c.g. affine-invariant sub-σ-algebras of $\mathcal{B}(R^n)$ and let μ be a Borel probability measure on R^n. If ν is an extreme point of $E(A_1,\ldots,A_m;\mu)$, then ν is singular (with respect to Lebesgue measure λ^n).

Proof: By lemma 1.3, the σ-algebras $A_1,\ldots,A_m$ are generated by orthogonal projections $\pi_1,\ldots,\pi_m$ of R^n onto subspaces $L_1,\ldots,L_m$. Let S be a finite subset of R^n as in lemma 1.4. For each $s \in S$, let $B(s)$ be a ball of radius r centered at s. We choose r small enough so that for each $i = 1,\ldots,m$ and any pair s,t in S, the projections $\pi_i(B(s))$ and $\pi_i(B(t))$ are either identical or disjoint. Select a large ball B of radius R containing all the sets $B(s)$. Set $k = |S|$ and put $\varepsilon = V_n(r)/[V_n(R)(k+1)]$.

Using the Lebesgue decomposition of ν into singular and absolutely continuous parts, we write $d\nu = d\nu_{\perp} + Fd\lambda^n$ for some $F \geq 0$ in $L^1(\lambda^n)$. Suppose, for the sake of argument, that ν is not singular. This means that for some positive δ, the set $P = \{x \in R^n : F(x) > \delta\}$ has positive λ^n-measure. We now appeal to the Lebesgue density theorem and choose a ball B_0 such that $\lambda^n(P \cap B_0) > (1 - \varepsilon)\lambda^n(B_0)$.

Let $M : R^n \to R^n$ be a mapping which is central (the composition of a translation and a central homothety) and takes B onto B_0. Let the image of S under M be $S_0 = \{s_1,\ldots,s_k\}$. If $M(s) = s_i$, define $B_0(s_i)$ to be the image of $B(s)$ under M. Then we claim that for

$i = 1,\dots,k,$

$$\lambda^n(P \cap B_0(s_i)) > \lambda^n(B_0(s_i)) \frac{k}{k+1} .$$

Otherwise,

$$\lambda^n(P \cap B_0) \le \lambda^n(B_0 \setminus B_0(s_i)) + \lambda^n(P \cap B(s_i))$$

$$\le \lambda^n(B_0) - \lambda^n(B_0(s_i)) + \lambda^n(B_0(s_i)) \frac{k}{k+1}$$

$$= \lambda^n(B_0) - \lambda^n(B_0(s_i)) \frac{1}{k+1}$$

and

$$\frac{\lambda^n(P \cap B_0)}{\lambda^n(B_0)} \le 1 - \frac{\lambda^n(B_0(s_i))}{\lambda^n(B_0)(k+1)} = 1 - \varepsilon,$$

a contradiction.

Now one may write $B_0(s_i) = s_i + C$, where C is a ball in R^n centered at the origin, $i = 1,\dots,k$. Define $P_0 \subseteq C$ by

$$P_0 = \bigcap_{i=1}^{k} [(P \cap B_0(s_i)) - s_i].$$

We claim that $\lambda^n(P_0) > 0$. Otherwise, we have for $j = 1,\dots,n$

$$\lambda^n(B_0(s_j)) = \lambda^n(B_0(s_j) \cap P) + \lambda^n(B_0(s_j) \cap P^c)$$

$$> \lambda^n(B_0(s_j))\frac{k}{k+1} + \lambda^n(B_0(s_j) \cap P^c),$$

so that

$$\lambda^n(B_0(s_j) \cap P^c) < \frac{\lambda^n(B_0(s_j))}{k+1},$$

and

$$\lambda^n(B_0(s_j)) = \lambda^n(C) = \lambda^n(C\setminus P_0)$$

$$= \lambda^n[\bigcup_{i=1}^{k} [(P \cap B_0(s_i))-s_i]^c \cap C]$$

$$\leq \sum_{i=1}^{k} \lambda^n[(P \cap B_0(s_i))^c \cap B_0(s_i)]$$

$$= \sum_{i=1}^{k} \lambda^n(B_0(s_i) \cap P^c)$$

$$< \sum_{i=1}^{k} \lambda^n(B_0(s_i)) \frac{1}{k+1}$$

$$= \lambda^n(B_0(s_j)) \frac{k}{k+1} < \lambda^n(B_0(s_j)),$$

a contradiction.

For each $i = 1,\ldots,k$, we define $A_i = P_0 + s_i$ and the linear functional $\ell_i : L^1(\nu) \to \mathbf{R}$ by

$$\ell_i(f) = \int_{A_i} f d\lambda^n.$$

Noting that $A_i \subseteq P$, we find

$$|\ell_i(f)| \leq \frac{1}{\delta}\int_{A_i}|f|Fd\lambda^n \leq \frac{1}{\delta}\int_{A_i}|f|d\nu \leq \frac{1}{\delta}\int|f|d\nu,$$

so that ℓ_i is continuous. Define a linear transformation $\ell : L^1(\nu) \to \mathbf{R}^k$ by setting $\ell = (\ell_1,\ldots,\ell_k)$.

Let F be the subspace of $L^1(\nu)$ comprising all functions of the form $f_1 \circ \pi_1 + \ldots + f_m \circ \pi_m$, where $f_1, \ldots, f_m$ are bounded Borel-measurable real functions on $L_1, \ldots, L_m$. Note that if $\pi_c(s_i) = \pi_c(s_j)$, then

$$\ell_i(f \circ \pi_c) = \int_{A_i} f \circ \pi_c d\lambda^n = \int_{P_0 + s_i} f \circ \pi_c d\lambda^n$$

$$= \int_{P_0} f(\pi_c(x) + \pi_c(s_i)) d\lambda^n(x)$$

$$= \int_{P_0} f(\pi_c(x) + \pi_c(s_j)) d\lambda^n(x) = \ell_j(f \circ \pi_c).$$

This fact allows a description of a set of spanning vectors for $\ell(F)$.

For each $c = 1, \ldots, m$ and each $p \in \pi_c(S_0)$, there is a k-vector $v = v(c,p)$ whose co-ordinates are given by

$$v_i = \begin{cases} 1 & \text{if } \pi_c(s_i) = p \\ 0 & \text{if } \pi_c(s_i) \neq p. \end{cases}$$

These vectors span $\ell(F)$. By lemma 1.4, there are fewer than k such vectors, so that $\ell(F)$ is a proper subspace of R^k.

However, the range of ℓ is all of R^k, as may be seen by taking linear combinations of indicator functions for the sets A_i. Now, the Douglas-Lindenstrauss Theorem implies that F is dense in $L^1(\nu)$. But this means that $R^k = \ell(\overline{F}) \subseteq \overline{\ell(F)} = \ell(F)$, a contradiction.

Q.E.D.

There seems to be no straightforward generalization of the theorem to the case $m = \infty$. For example, let L_1 L_2 ... be an enumeration of all lines in R^2 passing through the origin and having non-zero rational slope. Let $\pi_i : R^2 \to L_i$ be the projection maps generating the σ-algebras A_i, $i = 1, 2, \ldots$. Then $E = E(A_1, A_2, \ldots; \mu)$ is always a singleton set, even for absolutely continuous μ. To see this, let r_i be the slope of L_i. Then whenever $t_2 = r_i t_1$, we see that the Fourier-Stieltjes transform

$$\hat{\mu}(t_1, t_2) = \int_{R^2} e^{i(x_1 t_1 + x_2 t_2)} d\mu(x_1, x_2)$$

$$= \int e^{i(x_1 t_1 + x_2 r_i t_1)} d\mu(x_1, x_2)$$

$$= \int e^{i t_1 (x_1 + x_2 r_i)} d\mu(x_1, x_2)$$

depends only on the projection of μ on L_i. Therefore, these projections determine $\hat{\mu}$ on a dense set. So E is a singleton set.

3. References

[1] Bhaskara Rao, K. P. S. and Rao, B. V. Borel spaces. Dissertationes Mathematicae CXC (1981)

[2] Brown, J. R. and Shiflett, R. C. On extreme double stochastic measures. Michigan Math. Journal 17 (1970) 249-254

[3] Douglas, R. G. On extremal measures and subspace density. Michigan Math. J. 11 (1964) 243-246

[4] Doublas, R. G. On extremal measures and subspace density II. Proc. AMS 17 (1966) 1363-1365

[5] Lindenstrauss, J. A remark on doubly-stochastic measures. Amer. Math. Monthly 72 (1965) 379-382

[6] Losert, V. Counter-examples to some conjectures about doubly stochastic measures. Pacific J. of Math. 99 (1982) 387-397

Received November 15, 1985