ORBITS OF L^1-FUNCTIONS UNDER DOUBLY STOCHASTIC TRANSFORMATIONS(1)

BY
JOHN V. RYFF

1. **Introduction.** In their classic treatise on inequalities [2], Hardy, Littlewood and Pólya introduce a certain partial order $\prec$ in the real n-dimensional vector space E^n. If $\boldsymbol{x}$ and $\boldsymbol{y}$ are n-vectors and $\boldsymbol{y} \prec \boldsymbol{x}$, then $\boldsymbol{y}$ is said to be majorized by $\boldsymbol{x}$. Perhaps the most important result regarding this partial order is that $\boldsymbol{y} \prec \boldsymbol{x}$ if and only if $\boldsymbol{y}$ is an *average* of $\boldsymbol{x}$. That is, if and only if there is a doubly stochastic matrix T such that $\boldsymbol{y} = T\boldsymbol{x}$. (An n-square matrix is said to be *doubly stochastic* if it has non-negative elements with the sum of each row and each column equal to 1.) Equivalently, let X be the set of all vectors obtained by permuting the components of $\boldsymbol{x}$. Then $\boldsymbol{y} \prec \boldsymbol{x}$ if and only if $\boldsymbol{y}$ belongs to the convex hull of X (see [7]). Furthermore, the set of vertices of this convex hull is again X.

The purpose of this work is to present continuous analogues to the above results. Vectors will be replaced by integrable functions and matrices by linear operators. In particular, we shall be concerned with the class of *doubly stochastic operators* which have received some attention in current literature [6; 8; 9]. After certain modifications, the partial order of Hardy, Littlewood and Pólya carries over to the L^1-space and allows us to state the problem and its solution completely in terms of the partial order and doubly stochastic operators.

All variables encountered will be real and all functions measurable and finite-valued. $L^1 = L^1(0,1)$ is to be the space of Lebesgue integrable "functions" on the unit interval. It will be convenient to distinguish between elements of L^1 and their representatives. We shall do this by using boldface, so that $f \in \boldsymbol{f}$ is to mean that f is a representative of the class $\boldsymbol{f}$. The same convention will apply to elements of the dual space $(L^1)^* = L^\infty = L^\infty(0, 1)$ of essentially bounded "functions." Finally, we denote the unit interval $[0, 1]$ by I and Lebesgue measure by μ.

To define the partial order $\prec$ in E^n, suppose that $\boldsymbol{x} = (x_1, \cdots, x_n)$ and $\boldsymbol{y} = (y_1, \cdots, y_n)$. Take $\boldsymbol{x}^* = (x_1^*, \cdots, x_n^*)$ and $\boldsymbol{y}^* = (y_1^*, \cdots, y_n^*)$ to be the vectors obtained from $\boldsymbol{x}$ and $\boldsymbol{y}$ by rearranging their components in descending order. Then we define $\boldsymbol{y} \prec \boldsymbol{x}$ whenever

Presented to the Society, January 24, 1964; received by the editors December 19, 1963.
(1) This research was supported by the Army Research Office (Durham).

$$
\begin{aligned}
& y_1^* \leqq x_1^* \\
& \cdots\cdots\cdots \\
(1) \qquad & y_1^* + \cdots + y_k^* \leqq x_1^* + \cdots + x_k^* \qquad (1 \leqq k \leqq n) \\
& y_1 + \cdots + y_n = x_1 + \cdots + x_n \ .
\end{aligned}
$$

Now, if f is a measurable function on I, then the function

$$(2) \qquad m(y) = \mu\{s : f(s) > y\}$$

is right-continuous and nonincreasing. As such we may invert m according to the formula

$$(2') \qquad f^*(s) = \sup_{m(y)>s} y$$

obtaining again a right-continuous, nonincreasing function, known as the *decreasing rearrangement* of f (2). Moreover, f and f^* are equally integrable or nonintegrable, have the same distribution (2), and

$$
\begin{aligned}
\int_0^s f \leqq \int_0^s f^* & \qquad (0 \leqq s \leqq 1), \\
\int_0^1 f = \int_0^1 f^* & .
\end{aligned}
$$

Whenever f is integrable, we shall take $\boldsymbol{f}^*$ to be the element of L^1 containing f^*. If $\boldsymbol{f}$ and $\boldsymbol{g}$ are elements of L^1 we say that $\boldsymbol{f}$ *majorizes* $\boldsymbol{g}$ if

$$
\begin{aligned}
(3) \qquad \int_0^s \boldsymbol{g}^* \leqq \int_0^s \boldsymbol{f}^* & \qquad (0 \leqq s \leqq 1), \\
\int_0^1 \boldsymbol{g} = \int_0^1 \boldsymbol{f} & \ .
\end{aligned}
$$

and write $\boldsymbol{g} \prec \boldsymbol{f}$ (compare with (1)).

We intend to determine the necessary relationship between $\boldsymbol{f}$ and $\boldsymbol{g}$ in order that $\boldsymbol{g} \prec \boldsymbol{f}$. This question has been raised by Mirsky [6] and the author [9]. It will be shown that the relation $\boldsymbol{g} \prec \boldsymbol{f}$ holds if and only if there is a doubly stochastic operator which takes $\boldsymbol{f}$ onto $\boldsymbol{g}$. A linear transformation $T : L^1 \to L^1$ is said to be *doubly stochastic* if $T\boldsymbol{f} \prec \boldsymbol{f}$ for all $\boldsymbol{f} \in L^1$. This definition is suggested by the fact that a matrix D is doubly stochastic if and only if $D\boldsymbol{x} \prec \boldsymbol{x}$ for all vectors $\boldsymbol{x}$. Another equivalent (see [9]) definition is used by Rota in [8]. Doubly stochastic operators also carry L^∞ into itself and are contractions ($\| T \| \leqq 1$)

(2) This represents a departure from the convention of defining f^* as the decreasing rearrangement of $|f|$.

in both the L^1 and L^∞ operator norms. Because of this, we shall use the term *operator* hereafter and denote the class of doubly stochastic operators by $\mathscr{D}$. The set $\mathscr{D}$ forms a convex self-adjoint semigroup of operators (that is, $T \in \mathscr{D}$ implies $T^* \in \mathscr{D}$ and $T_1, T_2 \in \mathscr{D}$ then $T_1 T_2 \in \mathscr{D}$). This is not quite precise, since an operator and its adjoint act on different spaces. However, if we think of $T \in \mathscr{D}$ acting on L^1, then T^* (acting on L^∞) always admits a unique extension to an operator on L^1 which belongs to $\mathscr{D}$. This extension will tacitly be assumed hereafter.

In order to further fix our terminology, let $\mathscr{S}$ be a semigroup of transformations acting on a vector space V. To each $\boldsymbol{v} \in V$, the set of vectors $\{S\boldsymbol{v} : S \in \mathscr{S}\}$ will be called the *orbit* of $\boldsymbol{v}$ and will be written $\Omega(\boldsymbol{v})$. In particular, consider the case where $\mathscr{S}$ is the semigroup $\mathscr{D}_n$ of n-square doubly stochastic matrices. As $\mathscr{D}_n$ is a compact, convex subset of E^{n^2}, it follows that the orbit of any vector is likewise compact and convex. Thus, if $\boldsymbol{y} \prec \boldsymbol{x}$ but $\boldsymbol{y} \notin \Omega(\boldsymbol{x})$, we may separate $\Omega(\boldsymbol{x})$ and $\boldsymbol{y}$ by a hyperplane. That is, we can choose a vector $\boldsymbol{z}$ such that

$$(D\boldsymbol{x}, \boldsymbol{z}) < (\boldsymbol{y}, \boldsymbol{z})$$

for all $D \in \mathscr{D}_n$ (the parenthesis represents the scalar product in E^n). The addition of a constant vector $\boldsymbol{c} = (c, \cdots, c)$ to $\boldsymbol{z}$ will not destroy the inequality in view of (1). Hence, we may assume that the components of $\boldsymbol{z}$ are non-negative. Write $\boldsymbol{z} = P\boldsymbol{z}^*$ and $\boldsymbol{x} = Q\boldsymbol{x}^*$ where P and Q are permutation matrices. Then with D still at our disposal, set $D = PQ^*$, so that

$$(\boldsymbol{x}^*, \boldsymbol{z}^*) = (PQ^*Q\boldsymbol{x}^*, P\boldsymbol{z}^*) < (P^*\boldsymbol{y}, \boldsymbol{z}^*). \tag{4}$$

However, if $\boldsymbol{u}$ and $\boldsymbol{v}$ are n-vectors with $\boldsymbol{u} \prec \boldsymbol{v}$ and $\boldsymbol{w}$ is any vector with non-negative components, then

$$(\boldsymbol{u}, \boldsymbol{w}^*) \leqq (\boldsymbol{v}^*, \boldsymbol{w}^*). \tag{5}$$

As $\boldsymbol{y} \prec \boldsymbol{x}$ implies $P^*\boldsymbol{y} \prec \boldsymbol{x}$, we see that (4) is not possible. Incidentally, (5) is easily checked by observing that

$$0 \leqq \sum_{k=1}^{n} (w_k^* - w_{k+1}^*) \sum_{j=1}^{k} (v_j^* - u_j) = \sum_{k=1}^{n} w_k^*(v_k^* - u_k)$$

(where $w_{n+1}^* = 0$).

The preceding argument is essentially that of Rado [7] and represents a model we shall use for the continuous case. It is not at the moment clear that appropriate analogues to permutation matrices can always be found. Also, some kind of compactness will be necessary if we are to use separating hyperplanes.

2. **Equimeasurable functions.** Among the first examples of doubly stochastic operators encountered are those induced by measure preserving transformations of I into itself. A function $\sigma : I \to I$ is called *measure preserving* if, for each

measurable set $E \subset I$, $\sigma^{-1}(E)$ is measurable and $\mu(E) = \mu(\sigma^{-1}(E))$. Two such transformations will be identified if they differ on a set of measure zero. In keeping with our notation, the equivalence class containing σ will be $\boldsymbol{\sigma}$. If $\boldsymbol{f} \in L^1$, the transformation $T: \boldsymbol{f} \to \boldsymbol{f} \circ \boldsymbol{\sigma}$ is then doubly stochastic and is said to have been induced by $\boldsymbol{\sigma}$. Such operators exhibit behavior similar to that of permutation matrices although they are generally not invertible. While the permutation matrices constitute the extreme points of $\mathscr{D}_n$, those operators induced by measure preserving transformations, though extreme in $\mathscr{D}$, do not include all of the extreme points of $\mathscr{D}$ (see [9] for an example). A precise characterization of the extreme points of $\mathscr{D}$ does not seem to be known. Even so, we can still exibit enough operators in $\mathscr{D}$ to accomplish what the permutation matrices accomplish in E^n.

Two measurable functions $f \in \boldsymbol{f}$ and $g \in \boldsymbol{g}$ are said to be *equimeasurable* if, for each number y,

$$\mu\{s : f(s) > y\} = \mu\{s : g(s) > y\}.$$

In this case we write $\boldsymbol{f} \sim \boldsymbol{g}$. In particular, one always has $\boldsymbol{f} \sim \boldsymbol{f}^*$. The remainder of this section will be devoted to a proof of the following theorem.

THEOREM 1. *If $\boldsymbol{f}$ and $\boldsymbol{g}$ are in L^1 with $\boldsymbol{f} \sim \boldsymbol{g}$, then there exists a $T \in \mathscr{D}$ such that $\boldsymbol{g} = T\boldsymbol{f}$.*

An immediate corollary to this theorem is that $\Omega(\boldsymbol{f}) = \Omega(\boldsymbol{g})$ if and only if $\boldsymbol{f} \sim \boldsymbol{g}$. The sufficiency of the condition is clear, whereas necessity follows at once from the relation

$$\int_0^s \boldsymbol{f}^* = \int_0^s \boldsymbol{g}^*$$

which must be valid for each $s \in I$.

The proof will be broken down into three lemmas, the first of which is due to Lorentz [4]. The notational liberties $f^{-1}(A) = \{s : f(s) \in A\}$ and $f(B) = \{y = f(s) : s \in B\}$ will be employed throughout.

LEMMA 1 (LORENTZ). *Suppose f that and g are equimeasurable functions. If C is any set of real numbers for which $f^{-1}(C)$ is measurable then so is $g^{-1}(C)$ and both sets have the same measure.*

Roughly, this says that equimeasurable functions take on the same values equally often(3).

LEMMA 2. *To each $\boldsymbol{f} \in L^1$ there corresponds a measure preserving transformation σ such that $\boldsymbol{f} = \boldsymbol{f}^* \circ \boldsymbol{\sigma}$.*

(3) For other results along these lines, the reader is referred to the works of Lorentz [4] and Fan and Lorentz [1].

Proof. Choose $f \in \boldsymbol{f}$ and let f^* be given by (2′). Since $f^{-1}(f^*(I))$ has unit measure, we may assume (by suitably redefining f) that f and f^* have exactly the same range. Call the range C. To each $y \in C$ set $E_y = \{s : f(s) = y\}$. There is at most a denumerable set $J \subset C$ for which $\mu(E_y) > 0$, $y \in J$. Each such y corresponds to an interval of constancy in the graph of f^*. If $y \in C - J$, define, for each $s \in E_y$, $\sigma_y(s)$ to be that *unique* value $t \in I$ for which $f^*(t) = y$ (the monotonicity of f^* guarantees this).

The sets E_y, for $y \in J$, correspond to subintervals $[a, b)$ (or $[a, b]$) of I with $b - a = \mu(E_y)$. For such a y, consider E_y as a measure space in itself by restricting μ to the measurable subsets of E_y. Define $\sigma_y : E_y \to [a, b)$ by the equation

$$\mu([0, s] \cap E_y) = t - a,$$

where $s \in E_y$ and $a \leqq t \leqq b$. The relation $t = \sigma_y(s)$ is essentially a one-one correspondence between E_y and $[a, b)$(4). Furtheremore, σ_y is measure preserving. Indeed, if (t_1, t_2) is a subinterval of $[a, b)$ and $\sigma_y(s_i) = t_i$ $(i = 1, 2)$, then $\mu((s_1, s_2] \cap E_y) = t_2 - t_1$. But aside from a possible null set, $(s_1, s_2] \cap E_y$ is $\sigma^{-1}(t_1, t_2)$. It now follows easily that σ_y is measure preserving.

Compose the set of mappings σ_y, $y \in C$, to form a single map $\sigma : I \to I$. We intend to show that this composite is measure preserving. If E is a measurable subset of I, write

$$E = \bigcup_{y \in J} [E \cap \sigma(E_y)] \cup B,$$

a decomposition of E into mutually disjoint sets such that f^* is one-one on B. As

$$\sigma^{-1}(E) = \bigcup_{y \in J} [\sigma^{-1}(E) \cap E_y] \cup \sigma^{-1}(B)$$

and

$$\mu(\sigma^{-1}(E) \cap E_y) = \mu(E \cap \sigma(E_y))$$

for each $y \in J$, we shall be finished if we can show that $\mu(B) = \mu(\sigma^{-1}(B))$. But, if we set $D = f^*(B)$, then $f^{-1}(D) = \sigma^{-1}(B)$. Application of Lemma 1 now establishes the assertion.

In conclusion, we need only note that by the very definition of σ, it follows that $f(s) = f^*(\sigma(s))$ (a.e.) and consequently that $\boldsymbol{f} = \boldsymbol{f}^* \circ \boldsymbol{\sigma}$.

If T is the operator induced by $\boldsymbol{\sigma}$ in Lemma 2, then $T\boldsymbol{f}^* = \boldsymbol{f}$. We would like to show that one can go from $\boldsymbol{f}$ to $\boldsymbol{f}^*$ by means of an operator in $\mathscr{D}$. In those cases where σ is one-one, there is no problem—but this is not always the case. Nevertheless, it is still always possible to obtain $\boldsymbol{f}^*$ from $\boldsymbol{f}$ if we look at the *adjoint* of the operator which is given in Lemma 2.

LEMMA 3. *If T is the operator induced by $\boldsymbol{\sigma}$ then $T\boldsymbol{f}^* = \boldsymbol{f}$ and $T^*\boldsymbol{f} = \boldsymbol{f}^*$.*

(4) A denumerable number of sets of measure zero may go over to points of $[a, b)$.

Proof. Let χ_E be the characteristic function of the measurable set E. Then

$$\int_0^1 fT\chi_E = \int_0^1 (f^* \circ \sigma)(\chi_E \circ \sigma) = \int_0^1 \chi_E T^* f.$$

The middle term equals $\int_0^1 f^*\chi_E$. That is,

$$\int_E T^* f = \int_E f^*$$

or, $T^*f = f^*$.

That T^* is not always induced by a measure preserving transformation can readily be seen by considering the following example. Let $f(s) = 1-2s$ (mod 1) so that $f^*(s) = 1 - s$. Then $\sigma(s) = 2s$ (mod 1) and T^* is given by $(T^*g)(s) = \frac{1}{2}[g(s/2) + g((s+1)/2)]$ which cannot be written as $g \circ \tau$, τ measure preserving.

To prove Theorem 1 we simply observe that $f \sim g$ implies $f^* = g^*$ and, by composition of two operators from $\mathscr{D}$ $(f \to f^* \to g)$ we can go from f to g.

3. **Weakly compact orbits.** Since the operators in $\mathscr{D}$ act as contractions on L^∞ we may view $\mathscr{D}$ as a subset of the operator space of L^∞. From this point of view, we know, according to a general compactness theorem of Kadison [3], that $\mathscr{D}$ is compact in the weak*-operator topology. A sub-basic neighborhood of the null transformation in this topology is given by

$$N(f, g, \varepsilon) = \{T : \left| \int_0^1 fTu \right| < \varepsilon\},$$

where the operators T are taken from the operator space of L^∞ (not L^1) and $f \in L^1$, $u \in L^\infty$. Actually, it is necessary to show that $\mathscr{D}$ is closed in this topology. The following simple criterion helps to establish this: $T \in \mathscr{D}$ if and only if to each $\chi_E, 0 \leqq T\chi_E \leqq 1$ and $\int_0^1 T\chi_E = \mu(E)$ (see [9] for details). Exploiting the adjoint further, we now prove

THEOREM 2. *If $f \in L^1$, then $\Omega(f)$ is weakly compact.*

Proof. Let $\{T_\alpha f\}$ be a net in $\Omega(f)$. Then $\{T_\alpha^*\}$ is a net in $\mathscr{D}$. Choose a weak*-convergent subnet, $\{T_\beta^*\}$, and set $T_0^* = \lim_\beta T_\beta^*$ so that T_0 is the adjoint of this limit. Then $T \circ f = \lim_\beta T_\beta f$ (weakly). Indeed, if

$$N(T_0 f; u, \varepsilon) = \{g \in L^1 : \left| \int_0^1 (g - T_0 f)u \right| < \varepsilon\}$$

is a sub-basic neighborhood of $T_0 f$ ($u \in L^\infty$ and $\varepsilon > 0$) we may consider the corresponding weak*-neighborhood of T_0^*, $N(T_0^*; f, u, \varepsilon)$. The net $\{T_\beta^*\}$ is eventually in this neighborhood and, from

$$\left| \int_0^1 f(T_\beta^* - T_0^*)u \right| = \left| \int_0^1 u(T_\beta - T_0)f \right| ,$$

it follows that $\{T_\beta f\}$ is eventually in $N(T_0 f; u, \varepsilon)$. Similarly, one shows that $\Omega(f)$ is weakly closed; hence closed in the strong (norm) topology.

4. **Orbits and their extreme points.** The reader will recall that a weakly closed convex subset of a real Banach space B and any point not in that set may be (strictly) separated by a hyperplane: $\{x \in B : L(x) = \alpha\}$, where L is a linear functional continuous in the weak topology. Since the dual spaces of B, considered in the weak and strong topologies, coincide, the functional L is actually taken from B^*. This remark, together with the next lemma, combine to give our principal theorem.

LEMMA 4. *Suppose that f, g are in L^1 and that u is in L^∞. If $u \geqq 0$ and $g \prec f$, then*

$$\int_0^1 gu^* \leqq \int_0^1 g^* u^* \leqq \int_0^1 f^* u^* .$$

The proof of this lemma will be omitted . It may be established directly using integration by parts (see [4, p. 62]).

THEOREM 3. *If f and g are in L^1, then $g \prec f$ if and only if $g \in \Omega(f)$.*

Proof. The sufficiency of the condition remains to be shown. If $g \notin \Omega(f)$, then an element $u \in L^\infty$ exists such that for all $T \in \mathscr{D}$

$$\int_0^1 uTf < \int_0^1 ug .$$

Since $\int_0^1 Tf = \int_0^1 f = \int_0^1 g$, we may assume that $u \geqq 0$. Set $u = Pu^*$ and $f = Qf^*$, where $P, Q \in \mathscr{D}$. Then choose $T = PQ^*$. We have

$$\int_0^1 Pu^* \, PQ^* Qf^* \leqq \int_0^1 gPu^* ,$$

$$\int_0^1 u^* f^* \leqq \int_0^1 u^* P^* g .$$

But, as $P^* g \prec g \prec f$, we contradict Lemma 4; thus $g \in \Omega(f)$.

The question of extreme points is most interesting although it has yet to be completely settled. The next theorem give some indication of the situation.

THEOREM 4. *If $g \sim f$, then g is an extreme point of $\Omega(f)$.*

Proof(5). Suppose that $g = \frac{1}{2}(f_1 + f_2)$ with f_1 and f_2 in $\Omega(f)$. Then, for each $s \in I$

(5) Professor Doob kindly communicated this proof to the author.

$$\int_0^s f_i \leqq \int_0^s g^* , \qquad i = 1,2,$$

since $g^* = f^*$. Now it is well known (and easily verified) that if $v_1, v_2 \in L^1$, then

$$\int_0^s (v_1 + v_2)^* \leqq \int_0^s v_1^* + \int_0^s v_2^*$$

for each $s \in I$. Consequently,

$$\int_0^s g^* \leqq \frac{1}{2} \int_0^s f_1^* + f_2^* \leqq \int_0^s g^*$$

holds throughout I. This implies that $f_i \sim g$, $i = 1,2$. Since equimeasurable functions have the same L^1-norms,

$$\int_0^1 [|f_1 + f_2| - (|f_1| + |f_2|)] = 0.$$

The triangle inequality shows that f_1 and f_2 are positive and negative together. But for any constant c, $f_1 + c$ and $f_2 + c$ have the same properties as f_1 and f_2, so they too must be of like sign. From this follows $f_1 = f_2$.

It is somewhat disconcerting that the converse does not seem to lend itself to a simple solution. For all functions commonly encountered, the condition is also necessary. One should note that if g is an extreme point of $\Omega(f)$ then so is g^*. To see this, write $g^* = \frac{1}{2}(f_1 + f_2)$ with f_1 and f_2 in $\Omega(f)$. Next, determine σ, measure preserving such that $g = g^* \circ \sigma = \frac{1}{2}(f_1 \circ \sigma + f_2 \circ \sigma)$. It follows that $f_1 \circ \sigma = f_2 \circ \sigma$ and so, $f_1 = f_2$. The converse problem then reduces to one of determining whether or not $\Omega(f)$ can have an extreme point which is nonincreasing yet distinct from f^*. While this is still open we can still give an approximation theorem for functions in $\Omega(f)$ in terms of convex combinations of functions equimeasurable with f. We circumvent the problem of (possibly) not knowing all the extreme points of $\Omega(f)$. A point e of a convex set C is called an *exposed point* of C if it is possible to pass a supporting hyperplane through e which contains no other points of C. In terms of linear functionals L, this means that an inequality $L(v) < L(e)$, all $v \in C$, $v \neq e$, is satisfied. Of course, exposed points are always extreme points. To see how they are distinguished, consider a sphere set in the base of a truncated cone so that each ray on the surface of the cone is tangent to the sphere (in short, an "ice cream cone"). The points where the surfaces of the cone and sphere meet are extreme but not exposed. In [10, p. 96–97] Klee shows that weakly compact convex subsets of separable Banach spaces are equal to the closed convex hull of their exposed points (which, incidentally, are weakly dense in the set of extreme points). With regard to $\Omega(f)$ it is easy to show that each exposed point is equimeasurable with f. Let g be such a point. There exists $\phi \in L^\infty$ $(\phi \geqq 0)$ such that

$$\int_0^1 h\phi > \int_0^1 g\phi$$

for all $h \in \Omega(f)$ different from $g \circ$ Let $\phi = \phi^* \cdot \sigma$ where σ is measure preserving, and take $h = f^* \circ \sigma$. If $g = h$, then $g \sim f$. Otherwise, we have

$$\int_0^1 (f^* \circ \sigma)(\phi^* \circ \sigma) = \int_0^1 f^*\phi^* < \int_0^1 g \circ \phi \leqq \int_0^1 g^*\phi^* \leqq \int_0^1 f^*\phi^*$$

(since $g \prec f$) which is impossible. The approximation theorem then reads

THEOREM 5. *The convex combinations of all functions equimeasurable with f are L^1-dense in $\Omega(f)$.*

REFERENCES

1. K. Fan and G. G. Lorentz, *An integral inequality*, Amer. Math. Monthly **61** (1954), 626–631.

2. G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, Cambridge Univ. Press, Cambridge, 1952.

3. R. V. Kadison, *The trace in finite operator algebras*, Proc. Amer. Math. Soc. **12** (1961), 973–977.

4. G. G. Lorentz, *Bernstein polynomials*, Toronto Univ. Press, Toronto, 1953.

5. ———, *An inequality for rearrangements*, Amer. Math. Monthly **60** (1953), 176–179.

6. L. Mirsky, *Results and problems in the theory of doubly stochastic matrices*, Z. Wahr. **1** (1963), 319–334.

7. R. Rado, *An inequality*, J. London Math. Soc. **27** (1952), 1–6.

8. G.-C. Rota, *An "Alternierende Verfahren" for general positive operators*, Bull. Amer. Math. Soc. **68** (1962), 95–102.

9. J. V. Ryff, *On the representation of doubly stochastic operators*, Pacific J. Math. **13** (1963), 1379–1386.

10. V. L. Klee, Jr., *Extremal structure of convex sets*. II, Math. Z. **69** (1958), 90–104.

HARVARD UNIVERSITY,
CAMBRIDGE, MASSACHUSETTS