Can. J. Math., Vol. XXIX, No. 3, 1977, pp. 559–577

MATRICES DOUBLY STOCHASTIC BY BLOCKS

PAL FISCHER AND JOHN A. R. HOLBROOK

1. Introduction. The present work stems from the following classical result, due to G. H. Hardy, J. E. Littlewood, G. Pólya [**7**], and R. Rado [**10**].

THEOREM 1. *Concerning a pair of n-tuples $x, y \in \mathbf{R}^n$, the following four statements are equivalent:*

(a) *for every continuous, convex function $f : \mathbf{R} \to \mathbf{R}$*

$$(1) \qquad \sum_1^n f(y_i) \leqslant \sum_1^n f(x_i);$$

(b) *denoting the decreasing (non-increasing) rearrangement of the n-tuple x by x^*, we have*

$$(2) \qquad \sum_1^k y_i^* \leqslant \sum_1^k x_i^* \quad (k = 1, \ldots, n),$$

with equality for $k = n$;

(c) *for some $n \times n$ doubly stochastic matrix M (we recall that a square matrix is said to be doubly stochastic if all of its elements are non-negative and each row and column adds to one), we have $y = Mx$;*

(d) *the n-tuple y lies in the convex hull*

$$\operatorname{conv}\{\sigma(x) : \sigma \in S_n\}$$

where, for each permutation σ in the symmetric group S_n, $\sigma(x)$ denotes the n-tuple $(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)})$.

Theorem 1 has stimulated a profusion of related results, so that many generalizations and interpretations of this theorem are now available. However, we are not aware of any previous analysis of the conditions under which equality occurs in (1), except for the case of strictly convex f (I. Schur[**11**], K. M. Chong [**3**]). Here we shall characterize those (convex) f for which (1) becomes an equality, both in the classical one-dimensional situation of Theorem 1 and for the multi-dimensional generalizations.

To be specific, Section 3 of this paper contains a proof of the implication (b) $\Rightarrow$ (a) of Theorem 1 that is based on the integral representation of convex functions on the line, and that allows a complete description of those f for which (1) is an equality. In Section 4, these results are interpreted in terms of the doubly stochastic matrix M of Theorem 1 (c). Roughly speaking, equality in (1) for a non-trivial convex f corresponds to the case where M is doubly stochastic by blocks.

Received June 21, 1976 and in revised form, August 13, 1976. This research was supported in part by the National Research Council of Canada under grant A-8745.

To extend such results to the multi-dimensional case (where $f : \mathbf{R}^d \to \mathbf{R}$ and x, y are n-tuples of points in $\mathbf{R}^d$) seems to require a different approach. This analysis is carried out in our Section 5.

We shall continue the tradition of denoting by $x \succ y$ the relation between $x, y \in \mathbf{R}^n$ expressed in (b) of Theorem 1; likewise, we shall use the notation $x \gg y$ to indicate simply that the inequalities (2) hold (here the inequality may be strict when $k = n$). Pólya [**9**] and L. Mirsky [**8**] have given results analogous to Theorem 1 but with (b) replaced by the weaker relation $x \gg y$. Here, too, we are able to characterize those cases where (1) is an equality (see Sections 6 and 7 below).

Finally (Section 8) we discuss some results of G. F. D. Duff [**4**; **5**] concerning functions of the differences Δa_k obtained from an n-tuple a. Applying the techniques of the earlier sections, we are able to generalize Duff's inequalities and to analyse the conditions under which equality occurs. Recently we have found that similar work was done by K. M. Chong in [**2**]; the precise relation between our work and Chong's is spelled out in Section 8.

2. Lemmae.

For ease of reference we collect in this section several lemmae.

LEMMA 2. *Let $a_1, a_2, \ldots, a_n$ be a sequence of real numbers such that*

$$(3) \qquad s_k = \sum_{i=1}^{k} a_i > 0 \quad \text{for } 1 \leqq k \leqq n-1, \quad \text{and}$$

$$(4) \qquad s_n = \sum_{i=1}^{n} a_i = 0.$$

Let $\lambda_1 \geqq \lambda_2 \geqq \ldots \geqq \lambda_n$ be a sequence of reals. Then

$$\sum_{i=1}^{n} \lambda_i a_i \geqq 0$$

with equality if and only if $\lambda_1 = \lambda_2 = \ldots = \lambda_n$.

Proof. Clearly $0 = \sum_{i=1}^{n} \lambda_n a_i = \lambda_n a_n + \lambda_n s_{n-1}$. Since $s_{n-1} > 0$, $\lambda_{n-1}s_{n-1} \geqq \lambda_n s_{n-1}$ with equality if and only if $\lambda_n = \lambda_{n-1}$, i.e. $0 \leqq \lambda_{n-1}s_{n-2} + \sum_{i=n-1}^{n} \lambda_i a_i$ with equality if and only if $\lambda_n = \lambda_{n-1}$. By repeating the argument we obtain the proof of the lemma.

Lemma 2 can be generalized in the following way:

LEMMA 3. *Let $a_1, a_2, \ldots, a_n$ be a sequence of reals such that $s_k = \sum_{i=1}^{k} a_i \geqq 0$ for $1 \leqq k \leqq n$. Assume that $s_{l_1}, s_{l_2}, \ldots, s_{l_m}, s_n$ are the zero elements of the sequence $s_1, s_2, \ldots, s_n$, where $1 \leqq l_1 < l_2 < \ldots < l_m < n$. Let $\lambda_1 \geqq \lambda_2 \geqq \ldots \geqq \lambda_n$ be a sequence of reals. Then $\sum_{i=1}^{n} \lambda_i a_i \geqq 0$ with equality if and only if $\lambda_1 = \ldots = \lambda_{l_1}, \lambda_{l_1+1} = \ldots = \lambda_{l_2}, \ldots, \lambda_{l_m+1} = \ldots = \lambda_n$.*

Proof. From the proof of Lemma 2 it follows that

$$0 \leqq \lambda_{l_m+1} s_{l_m} + \sum_{i=l_m+1}^{n} \lambda_i a_i$$

with equality if and only if $\lambda_{l_m+1} = \ldots = \lambda_n$. Since $s_{l_m} = 0$, we have that

$$0 \leqq \lambda_{l_m} s_{l_m} + \sum_{i=l_m+1}^{n} \lambda_i a_i$$

with equality if and only if $\lambda_{l_m+1} = \ldots = \lambda_n$. By repeating the argument we obtain the proof of the lemma.

The following two lemmae can be proved similarly.

LEMMA 4. *Let $a_1, a_2, \ldots, a_n$ be a sequence of real numbers such that $s_k = \sum_{i=1}^{k} a_i > 0$ for $1 \leqq k \leqq n$. Let $\lambda_1 \geqq \lambda_2 \geqq \ldots \geqq \lambda_n \geqq 0$. Then $\sum_{i=1}^{n} \lambda_i a_i \geqq 0$ with equality if and only if $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$.*

LEMMA 5. *Let $a_1, a_2, \ldots, a_n$ be a sequence of reals such that*

$$(5) \qquad s_k = \sum_{i=1}^{k} a_i \geqq 0 \quad \text{for } 1 \leqq k \leqq n-1 \quad \text{and}$$

$$(6) \qquad s_n = \sum_{i=1}^{n} a_i > 0.$$

Assume that $s_{l_1}, s_{l_2}, \ldots, s_{l_m}$ are the zero elements of the sequence $s_1, s_2, \ldots, s_n$, where $1 \leqq l_1 < l_2 < \ldots < l_m < n$. Let $\lambda_1 \geqq \lambda_2 \geqq \ldots \geqq \lambda_n \geqq 0$ be a sequence of reals.

Then $\sum_{i=1}^{n} \lambda_i a_i \geqq 0$ with equality if and only if $\lambda_1 = \ldots = \lambda_{l_1}, \lambda_{l_1+1} = \ldots = \lambda_{l_2}, \ldots, \lambda_{l_m+1} = \ldots = \lambda_n = 0$.

3. The theorem of Hardy, Littlewood, and Pólya; Extremal functions.

THEOREM 6. *Let $y \in \mathbf{R}^n$ and $x \in \mathbf{R}^n$ be such that $y \prec x$ and assume that $f : [x_n^*, x_1^*] \to \mathbf{R}$ is a continuous convex function. Then*

$$(7) \qquad \sum_{i=1}^{n} f(y_i) \leqq \sum_{i=1}^{n} f(x_i).$$

If

$$(8) \qquad \sum_{1}^{k} x_i^* > \sum_{1}^{k} y_i^*$$

for all $k < n$, then equality occurs in (7) if and only if f is affine on $[x_n^*, x_1^*]$. If (8) holds for all $k < n$ except $l_1, l_2, \ldots, l_m$ where $l_1 < l_2 < \ldots < l_m$, then equality occurs in (7) if and only if f is affine on each of the intervals

$$[x_{l_1}^*, x_1^*], [x_{l_2}^*, x_{l_1+1}^*], \ldots, [x_n^*, x_{l_m+1}^*].$$

Remark 7. Suppose that $1 \leqq r < l \leqq n$. If $x_1^* = \ldots = x_r^* = x_{r+1}^* = y_1^* = \ldots = y_r^* > y_{r+1}^*$ and $x_l^* = \ldots = x_n^* = y_{l-1}^* = y_l^* = \ldots = y_n^* < x_{l-1}^*$ and $\sum_{i=1}^{k} x_i^* > \sum_{i=1}^{k} y_i^*$ for $r+1 \leqq k \leqq l-1$, then Theorem 6 states that equality occurs in (7) only for functions affine on $[x_n^*, x_1^*]$ (among the continuous convex functions).

Proof of theorem. Without loss of generality, we can assume that $x_1^* > y_1^*$ and $x_n^* < y_n^*$. Since f is continuous and convex on the interval $[x_n^*, x_1^*]$, there exists on the interval (x_n^*, x_1^*) an increasing function h such that

$$(9) \quad f(x) = c_0 + \int_{x_0}^{x} h(t)dt,$$

where $x_0 \in (x_n^*, x_1^*)$. It may be necessary to interpret (9) as an improper integral when x is an endpoint.

In order to prove (7) we have to show that

$$(10) \quad 0 \leqq \sum_{i=1}^{n} [f(x_i^*) - f(y_i^*)] = \sum_{i=1}^{n} \int_{y_i^*}^{x_i^*} h(t)dt;$$

equivalently, we have to show that

$$(11) \quad \sum_{x_i^* \geqq y_i^*} \int_{y_i^*}^{x_i^*} h(t)dt \geqq \sum_{x_i^* < y_i^*} \int_{x_i^*}^{y_i^*} h(t)dt.$$

Using the fact that h is monotone, we see that

$$(12) \quad \sum_{x_i^* < y_i^*} \int_{x_i^*}^{y_i^*} h(t)dt \leqq \sum_{x_i^* < y_i^*} (y_i^* - x_i^*)h(y_i^*) \quad \text{and}$$

$$(13) \quad \sum_{x_i^* \geqq y_i^*} \int_{y_i^*}^{x_i^*} h(t)dt \geqq \sum_{x_i^* \geqq y_i^*} (x_i^* - y_i^*)h(y_i^*).$$

In order to prove (7) it is enough to show that

$$(14) \quad \sum_{x_i^* \geqq y_i^*} (x_i^* - y_i^*)h(y_i^*) \geqq \sum_{x_i^* < y_i^*} (y_i^* - x_i^*)h(y_i^*);$$

equivalently, it is enough to establish that

$$(15) \quad \sum_{i=1}^{n} (x_i^* - y_i^*)h(y_i^*) \geqq 0.$$

This last relation follows immediately from Lemma 3 by letting $x_i^* - y_i^* = a_i$ and $h(y_i^*) = \lambda_i$ for $1 \leqq i \leqq n$.

To analyze the equality we shall distinguish two cases:

(i) $\sum_{i=1}^{k} x_i^* > \sum_{i=1}^{k} y_i^*$ for $1 \leqq k \leqq n - 1$.

It follows from Lemma 2 that in this case equality holds in (15) if and only if $h(y_1^*) = \ldots = h(y_n^*)$. Furthermore it is easy to see that in (12) and in (13) equality holds if and only if h is constant on the interval $(\min(x_i^*, y_i^*), \max(x_i^*, y_i^*)) \cup \{y_i^*\}$ for $1 \leqq i \leqq n$; that is, f is an affine function on the interval $[x_n^*, x_1^*]$.

(ii) $\sum_{i=1}^{k} x_i^* \geqq \sum_{i=1}^{k} y_i^*$ for $1 \leqq k \leqq n - 1$.

https://doi.org/10.4153/CJM-1977-058-8 Published online by Cambridge University Press

Assume that $l_1, l_2, \ldots, l_m, n$ is the sequence of those k's for which in (ii) the equality holds, and that $1 < l_1 < l_2 < \ldots < l_m < n$. Then by Lemma 3 equality holds in (15) if and only if $h(y_1^*) = \ldots = h(y_{l_1}^*), \ldots, h(y_{l_m+1}^*) = \ldots = h(y_n^*)$, and we see that h has to be constant in the intervals $(x_{l_1}^*, x_1^*)$, $(x_{l_1+1}^*, x_{l_2}^*), \ldots, (x_n^*, x_{l_m+1}^*)$. Therefore f is affine on those intervals. This completes the proof.

Remark 8. As a special case we obtain the following result of I. Schur [**11**]: *if $f''(x)$ exists for all x and is positive, then equality can occur in* (7) *only when the sequences $(x_1^*, \ldots, x_n^*)$ and $(y_1^*, \ldots, y_n^*)$ are identical.*

Remark 9. Since

$$(x_1, \ldots, x_n) \succ \left(\frac{1}{n}\sum_1^n x_i, \ldots, \frac{1}{n}\sum_1^n x_i\right),$$

the fact that a continuous convex function f satisfies

$$f\left(\frac{1}{n}\sum_1^n x_i\right) = \frac{1}{n}\sum_1^n f(x_i)$$

only when f is affine on some interval containing the x's may be regarded as an application of Theorem 6. See also Proposition 14 in Section 5.

4. Doubly stochastic matrices.

As a complement to Theorem 11, we mention first the following consequence of Theorem 6.

PROPOSITION 10. *Let $x = (x_1, x_2, x_3) \in \mathbf{R}^3$ and $y = (y_1, y_2, y_3) \in \mathbf{R}^3$ such that $y \prec x$, $x_1^* > y_1^*$, and $y_3^* > x_3^*$. If $f : [x_3^*, x_1^*] \to \mathbf{R}$ is a continuous convex function such that*

$$\sum_{i=1}^3 f(x_i) = \sum_{i=1}^3 f(y_i)$$

then f is affine on $[x_3^*, x_1^*]$.

We shall say that a matrix is *doubly stochastic by blocks* if it is a block diagonal matrix and if every (diagonal) block is doubly stochastic.

THEOREM 11. *Let n be a fixed positive integer, $n \geqq 4$. Let $x \in \mathbf{R}^n$ and $y \in \mathbf{R}^n$ be such that $y \prec x$ and $x_1^* > y_1^*$ and $x_n^* < y_n^*$. Then the following statements are equivalent:*

(i) *there exists a matrix T, doubly stochastic by blocks, such that $y^* = Tx^*$;*

(ii) *there exists a non-affine convex continuous function f defined on $[x_n^*, x_1^*]$ such that*

$$\sum_{i=1}^n f(y_i) = \sum_{i=1}^n f(x_i). \tag{15}$$

Remark 12. A third statement, equivalent to (i) and (ii) may be formulated as follows: *there exists a partition of $\{1, \ldots, n\}$ into disjoint sets $S_1, \ldots, S_m$ such that y lies in* conv $\{\sigma(x) : \sigma$ *belongs to the subgroup of permutations of $\{1, \ldots, n\}$ such that $\sigma(S_i) = S_i$ for $1 \leqq i \leqq m$*$\}$. This condition is analogous to (d) of Theorem 1, and the proof of its equivalence to (i) is evident in view of that theorem.

Proof of theorem. We shall first show that (i) $\Rightarrow$ (ii). Since T is a doubly stochastic matrix by blocks it is the direct sum of doubly stochastic matrices $T_1, T_2, \ldots, T_m$ with orders $l_1, l_2 - l_1, \ldots, n - l_m$. T_1 transforms $x_1^*, \ldots, x_{l_1}^*$ to $y_1^*, \ldots, y_{l_1}^*$. Since $x_1^* > y_1^*$ and $x_n^* < y_n^*$ we see that $2 \leqq l_1 \leqq n - 2$. It is evident that for any function f_1 which is affine on $[x_{l_1}^*, x_1^*]$, $\sum_{i=1}^{l_1} f_1(x_i^*) = \sum_{i=1}^{l_1} f_1(y_i^*)$. The transformation $U = T_2 \oplus \ldots \oplus T_m$ is doubly stochastic, and U maps $x_{l_1+1}^*, \ldots, x_n^*$ to $y_{l_1+1}^*, \ldots, y_n^*$. Hence for every function f_2 which is affine on $[x_n^*, x_{l_1+1}^*]$ we have

$$\sum_{i=l_1+1}^{n} f_2(x_i^*) = \sum_{i=l_1+1}^{n} f_2(y_i^*).$$

Since $l_1 + 1 < n$ and since $[x_n^*, x_{l_1+1}^*] \cap [x_{l_1}^*, x_1^*]$ contains at most one point we see that there exists a continuous convex non-affine function on $[x_n^*, x_1^*]$ satisfying (15).

(ii) $\Rightarrow$ (i). By Theorem 5 we see that there exists at least one k such that $2 \leqq k \leqq n - 2$ and $\sum_{i=1}^{k} x_i^* = \sum_{i=1}^{k} y_i^*$. Clearly $u_1 = (y_1^*, \ldots, y_k^*) \prec v_1 = (x_1^*, \ldots, x_k^*)$ and $u_2 = (y_{k+1}^*, \ldots, y_n^*) \prec v_2 = (x_{k+1}^*, \ldots, x_n^*)$. By Theorem 1 ((b) $\Rightarrow$ (c)) there exist doubly stochastic matrices T_1 and T_2 such that $T_1 v_1 = u_1$ and $T_2 v_2 = u_2$. Thus the matrix $T = T_1 \oplus T_2$ is doubly stochastic by blocks and $y^* = Tx^*$.

5. The multi-dimensional case. In this section x and y will denote n-tuples of points in $\mathbf{R}^d$. Certain of the results of the earlier sections concerning the case $d = 1$ can be extended so that they apply also when $d > 1$. We shall write $x \succ_c y$ when the analogue of statement (a) in Theorem 1 is true of x and y, i.e., when

$$(16) \quad \sum_{1}^{n} f(y_i) \leqq \sum_{1}^{n} f(x_i)$$

for every continuous convex $f : \mathbf{R}^d \to \mathbf{R}$.

In the multi-dimensional case there seems to be no simple "intrinsic" condition on x, y that serves as an analogue of statement (b) in Theorem 1, i.e., that is equivalent to $x \succ_c y$. However, the equivalence (a) $\Leftrightarrow$ (c) of Theorem 1 does carry over to the multi-dimensional situation. Thus we have the following theorem, due to S. Sherman [**12**] and C. Stein (see D. Blackwell [**1**]).

THEOREM 13. *Let $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$, where the x_k and y_k are elements of $\mathbf{R}^d$. Then $x \succ_c y$ is, and only if, there exists an $n \times n$ doubly stochastic matrix M such that $y = Mx$. Here we regard x and y as column vectors and interpret Mx formally as matrix multiplication; explicitly, if $M = [m_{ij}]$,*

$$y_i = \sum_{j=1}^{n} m_{ij} x_j \quad (i = 1, \ldots, n).$$

https://doi.org/10.4153/CJM-1977-058-8 Published online by Cambridge University Press

The information contained in Theorem 13 will help us to characterize the cases of equality in (16). We shall also need the two simple propositions presented below.

PROPOSITION 14. *Suppose that* $f : \mathbf{R}^d \to \mathbf{R}$ *is convex and continuous, and that* $y_0, x_1, \ldots, x_m$ *are points of* $\mathbf{R}^d$. *If* y_0 *is a convex combination* $\sum_1^m \lambda_k x_k$ *where each* $\lambda_k > 0$ *and* $f(y_0) = \sum_1^m \lambda_k f(x_k)$, *then* f *is affine on the convex hull* C *of* $x_1, \ldots, x_m$.

Proof. Let H be a supporting hyperplane to the convex set

$$\{(z, t) : z \in \mathbf{R}^d, t \geqq f(z)\}$$

at $(y_0, f(y_0))$. Then H is the graph of an affine function $F : \mathbf{R}^d \to \mathbf{R}$ such that $F(y_0) = f(y_0)$ and $F \leqq f$ on $\mathbf{R}^d$. Thus

$$\sum_1^m \lambda_k f(x_k) = f(y_0) = F(y_0) = \sum_1^m \lambda_k F(x_k)$$

and

$$f(x_k) \geqq F(x_k), \quad \lambda_k > 0 \quad (k = 1, \ldots, m).$$

Certainly, then, $f(x_k) = F(x_k)(k = 1, \ldots, m)$. Now for any $z \in C$, $z = \sum_1^m \mu_k x_k$ (convex combination) so that

$$F(z) \leqq f(z) \leqq \sum_1^m \mu_k f(x_k) = \sum_1^m \mu_k F(x_k) = F(z).$$

Hence $f \equiv F$ on C.

PROPOSITION 15. *Suppose that* $f : \mathbf{R}^d \to \mathbf{R}$ *is convex and continuous and, for two subsets* A, B, *of* $\mathbf{R}^d$, f *is affine on each of* A *and* B. *Then* f *is affine on* K, *the closed convex hull of* $A \cup B$ *provided that* $A \cap B$ *has non-empty interior.*

Proof. Let $F_A, F_B : \mathbf{R}^d \to \mathbf{R}$ be affine functions such that $F_A \equiv f$ on A and $F_B \equiv f$ on B. Since $A \cap B$ has interior, $F_A \equiv F_B$ on some (non-empty) open set U of $\mathbf{R}^d$, so that $F_A \equiv F_B$ on $\mathbf{R}^d$. We denote this common function simply by F. Now $f \geqq F$ on $\mathbf{R}^d$ since, if $f(z) < F(z)$ and $u \in U$, then $f(u) = F(u)$, F is affine and f convex on the line segment $[u, z]$; hence $f < F$ on (u, z), which is impossible since $f \equiv F$ on U and $U \cap (u, z) \neq \emptyset$ (U is open). From $f \geqq F$ is follows that the closed set

$$C = \{z : f(z) = F(z)\}$$

is also convex: if $z_1, z_2 \in C$ and $z = \lambda_1 z_1 + \lambda_2 z_2$ (convex combination) then

$$F(z) \leqq f(z) \leqq \lambda_1 f(z_1) + \lambda_2 f(z_2) = \lambda_1 F(z_1) + \lambda_2 F(z_2) = F(z).$$

But $C \supset A$ and $C \supset B$, so that $C \supset K$ as well.

The following theorem seems to be the appropriate multi-dimensional analogue of Theorem 6.

THEOREM 16. *Given* $x \succ_c y$ (*where* $x_k, y_k \in \mathbf{R}^d$, $k = 1, \ldots, n$), *there exist* *convex sets* $K_1, \ldots, K_p$ *in* $\mathbf{R}^d$ *such that: for convex, continuous* $f : \mathbf{R}^d \to \mathbf{R}$, *equality holds in* (16) *if, and only if, f is affine on each* $K_q (q = 1, \ldots, p)$.

Furthermore, the sets K_q *may be chosen to have the following properties. For certain* (*non-empty*) *subsets* I_q, J_q *of* $\{1, \ldots, n\}$ *such that*

$$\bigcup_1^p I_q = \bigcup_1^p J_q = \{1, \ldots, n\}$$

and the I_q *are pairwise disjoint* (*the* J_q *might not be*) *we have:*

(17) $K_q = \operatorname{conv} \{x_j : j \in J_q\} \quad (q = 1, \ldots, p)$

(18) $i \in I_q \Rightarrow y_i \in K_q \quad (q = 1, \ldots, p)$

(19) $q \neq q' \Rightarrow K_q \cap K_{q'}$ *has empty interior*,

i.e., *the* K_q*'s intersect at most their boundaries, and*

(20) $$\#\left(\bigcup_{q \in Q} I_q\right) \leqq \#\left(\bigcup_{q \in Q} J_q\right),$$

for any subset Q *of* $\{1, \ldots, p\}$.

In the expressions above we use the notations conv (S) and $\#(S)$ for the convex hull of the set S and the number of elements in S, respectively.

Proof. Since $x \succ_c y$, Theorem 13 ensures that $y = Mx$ for some $n \times n$ doubly stochastic matrix $M = [m_{ij}]$. Let

$$R_i = \{j : m_{ij} > 0\}.$$

We claim that for any partition of $\{1, \ldots, n\}$ into sets $I_1, \ldots, I_p$, (20) is satisfied if we set

(21) $$J_q = \bigcup_{i \in I_q} R_i \quad (q = 1, \ldots, p).$$

To see this note that, if $I = \bigcup_{q \in Q} I_q$,

$$\begin{aligned}
\#(I) &= \sum_{i \in I} \left(\sum_{j \in R_i} m_{ij}\right) \\
&\leqq \sum_{j \in \bigcup \{R_i : i \in I\}} \left(\sum_{i=1}^n m_{ij}\right) \\
&= \#(\bigcup \{R_i : i \in I\}) \\
&= \#\left(\bigcup_{q \in Q} \left(\bigcup_{i \in I_q} R_i\right)\right) = \#\left(\bigcup_{q \in Q} J_q\right).
\end{aligned}$$

Furthermore, since $y_i \in \operatorname{conv} \{x_j : j \in R_i\}$, the construction of the set K_q determined by (21) and (17) ensures that (18) holds.

We shall construct an appropriate partition $I_1, \ldots, I_p$ by successive amalgamations, starting with the finest partition:

$$I_1 = \{1\}, \ldots, I_n = \{n\}.$$

We must check that, for this initial partition, the sets K_q suffice to characterize those convex, continuous f for which equality holds:

$$(22) \quad \sum_1^n f(y_i) = \sum_1^n f(x_i).$$

But

$$\sum_{i=1}^n f(y_i) = \sum_{i=1}^n f\Big(\sum_{j\in R_i} m_{ij}x_j\Big)$$

and, by convexity,

$$f\Big(\sum_{j\in R_i} m_{ij}x_j\Big) \leqq \sum_{j\in R_i} m_{ij}f(x_j)$$

for each i. On the other hand

$$\sum_{i=1}^n \Big(\sum_{j\in R_i} m_{ij}f(x_j)\Big) = \sum_{j=1}^n \Big(\sum_{i=1}^n m_{ij}\Big)f(x_j) = \sum_1^n f(x_j),$$

so that (22) holds if, and only if,

$$f\Big(\sum_{j\in R_i} m_{ij}x_j\Big) = \sum_{j\in R_i} m_{ij}f(x_j) \quad (i = 1, \dots, n).$$

In view of Proposition 14 this occurs exactly when f is affine on each of the sets

$$\operatorname{conv} \{x_j : j \in R_i\}.$$

These are just the sets K_q for our initial partition. Now suppose that for a given partition $I_1, \dots, I_p$, the remaining condition (19) is not satisfied, i.e., for some $q \neq q'$, $K_q \cap K_{q'}$ has non-empty interior. Applying Proposition 15 (with $A = K_q$, $B = K_{q'}$) we see that, if f satisfies (22), then f must be affine on

$$\begin{aligned}\operatorname{conv}\ &(K_q \cup K_{q'}) \\ &= \operatorname{conv}\ (\operatorname{conv} \{x_j : j \in J_q\} \cup \operatorname{conv} \{x_j : j \in J_{q'}\}) \\ &= \operatorname{conv} \{x_j : j \in J_q \cup J_{q'}\} \\ &= \operatorname{conv} \{x_j : j \in \cup \{R_i : i \in I_q \cup I_{q'}\}\}.\end{aligned}$$

We may thus amalgamate I_q and $I_{q'}$ to obtain a new partition satisfying (21), (17), and (18). Continuing in this way we obtain (after less than n steps) a partition satisfying (19) as well.

As an application of Theorem 16, we give the following extension, to the multi-dimensional case, of results of I. Schur [**11**] and K. M. Chong [**3**].

THEOREM 17. *If $x \succ_c y$ and for some strictly convex function $f : \mathbf{R}^d \to \mathbf{R}$ we have*

$$\sum_1^n f(y_i) = \sum_1^n f(x_i),$$

then $y = \sigma(x)$ for some permutation σ.

https://doi.org/10.4153/CJM-1977-058-8 Published online by Cambridge University Press

Proof. Since f is strictly convex, it is not affine on any convex set containing more than one point. Hence the K_q of Theorem 16 must be individual points, so that, by (17) and (18),

$$i \in I_q, j \in J_q \Rightarrow y_i = x_j.$$

Let $z_1, \ldots, z_r$ be the *distinct* points among the y_i. For each $s \leqq r$, let

$$Q_s = \{q : i \in I_q \Rightarrow y_i = z_s\}.$$

Then, using (20),

$$\#\{i : y_i = z_s\} = \#\left(\bigcup_{q \in Q_s} I_q\right)$$
$$\leqq \#\left(\bigcup_{q \in Q_s} J_q\right) \leqq \#\{j : x_j = z_s\}.$$

Since

$$\sum_{s=1}^{r} \#\{i : y_i = z_s\} = \sum_{s=1}^{r} \#\{j : x_j = z_s\} = n,$$

we must have

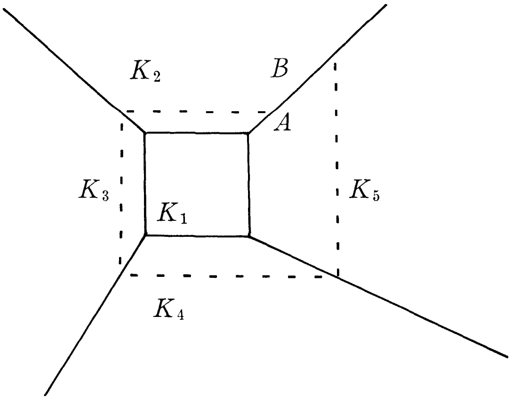
$$\#\{i : y_i = z_s\} = \#\{j : x_i = z_s\} \quad (s = 1, \ldots, r).$$

Remark 18. In the multi-dimensional setting, the existence of a convex, continuous f that is *not* affine on conv $\{x_1, \ldots, x_n\}$ but that satisfies (22) does not guarantee that $y = Mx$ for some matrix M doubly stochastic by blocks. It is not clear, then, what should be the extension of Theorem 11 of section 4 to the case $d > 1$. Consider this example with $d = 2$, $n = 3$; let x_1, x_2, x_3 be points in the place forming a non-degenerate triangle, and let y_1 ,y_2, y_3 be the midpoints of the three sides. For any given ordering of the y's, the doubly stochastic matrix M such that $Mx = y$ is uniquely determined and is never block diagonal. The set K_q of Theorem 16 are just the three sides of the triangle in this case. Thus it is easy to find a convex f that is affine on K_1, K_2, and K_3 but not on the triangle conv $\{x_1, x_2, x_3\}$. For example, assuming for convenience that the origin lies within the triangle, let $f(tz) = t$ for each $z \in K_1 \cup K_2 \cup K_2$, and each $t \geqq 0$.

Remark 19. The multi-dimensional situation is further complicated by another phenomenon. Given convex $K_1, \ldots, K_p$ in $\mathbf{R}^d$ and satisfying (19) it is clear that when $d = 1$ (so that the K_q are intervals overlapping at most at end-points) there exists a continuous convex function f affine on each K_q but not on $K_1 \cup \ldots \cup K_p$. For $d > 1$, this is not generally the case. Consider

https://doi.org/10.4153/CJM-1977-058-8 Published online by Cambridge University Press

the convex sets $K_1, \ldots, K_5$ of the figure:



Suppose f is a continuous function that is affine on each K_q. Subtracting that affine function that matches f on K_1 we may assume that $f \equiv 0$ on K_1. As the segments of the dotted line are parallel to sides of K_1, that line must be a level curve for f. Thus $f(A) = f(B)$ so that $f \equiv 0$ in K_2, and it follows that $f \equiv 0$ everywhere.

6. The theorem of Pólya; extremal functions.

THEOREM 20. *Let $x \in \mathbf{R}^n$ and $y \in \mathbf{R}^n$ be such that $y \ll x$, and assume that $f : [\min (x_n^*, y_n^*), x_1^*] \rightarrow \mathbf{R}$ is a continuous convex and increasing function. Then*

$$(23) \quad \sum_{i=1}^{n} f(y_i) \leqq \sum_{i=1}^{n} f(x_i).$$

If

$$(24) \quad \sum_{i=1}^{k} x_i^* > \sum_{i=1}^{k} y_i^*$$

for all $k \leqq n$, then equality occurs in (23) *if and only if f is constant on* $[\min (x_n^*, y_n^*), x_1^*]$. *If* (24) *holds for all $k \leqq n$ except $l_1, l_2, \ldots, l_m$ where $l_1 < l_2 < \ldots < l_m$ then equality holds in* (23) *if and only if f is affine on the intervals*

$$(x_{l_1}^*, x_1^*), \ldots, (x_{l_m}^*, x_{l_{m-1}+1}^*)$$

and is constant on $[\min (x_n^*, y_n^*), x_{l_m+1}^*]$.

Proof. Essentially we can repeat the proof of Theorem 6. In the present case there exists a function h, increasing and non-negative on the interval $[\min (x_n^*, y_n^*), x_1^*)$ such that

$$f(x) = c_0 + \int_{x_0}^{x} h(t)dt \quad \text{for } x \in [\min (x_n^*, y_n^*), x_1^*].$$

Without loss of generality we can assume that $x_1^* > y_1^*$. In order to prove (23) we can repeat (10), (11), (12), (13) aad (14) and we find that it is

https://doi.org/10.4153/CJM-1977-058-8 Published online by Cambridge University Press

enough to establish

$$(25) \quad \sum_{i=1}^{n} (x_i^* - y_i^*)h(y_i^*) \geqq 0.$$

This last relation follows immediately from Lemma 5 by letting $x_1^* - y_1^* = a_i$ and $h(y_i^*) = \lambda_i$ for $1 \leqq i \leqq n$.

To study the extremal functions, we shall distinguish two cases:

$$\text{(i)} \quad \sum_{i=1}^{k} x_i^* > \sum_{i=1}^{k} y_i^* \quad \text{for } 1 \leqq k \leqq n.$$

By Lemma 4 we see that in (25) equality holds if and only if $h(y_1^*) = \ldots = h(y_n^*) = 0$. In (12) and in (13) equality holds if and only if h is constant in the interval $(\min (x_i^*, y_i^*), \max (x_i^*, y_i^*)) \cup \{y_i\}$ for $1 \leqq i \leqq n$. Therefore equality occurs in (23) if and only if f is constant on $[\min (x_n^*, y_n^*), x_1^*]$.

$$\text{(ii)} \quad \sum_{i=1}^{k} x_i^* \geqq \sum_{i=1}^{k} y_i^* \quad \text{for } 1 \leqq k \leqq n - 1.$$

Let $l_1 < l_2 < \ldots < l_m$ be the sequence of those k's for which equality holds in (ii). By Lemma 5 we can see that in (25) equality holds if and only if $\lambda_1 = \ldots = \lambda_{l_1}, \ldots, \lambda_{l_{m-1}+1} = \ldots = \lambda_{l_m}, \lambda_{l_m+1} = \ldots = \lambda_n = 0$. As we mentioned earlier, in (12) and in (13) equality holds if and only if h is constant in the interval $(\min (x_i^*, y_i^*), \max (x_i^*, y_i^*)) \cup \{y_i^*\}$ for $1 \leqq i \leqq n$. Therefore equality occurs in (23) if and only if f is constant on $[\min (x_n^*, y_n^*), x_{l_m+1}^*]$ and f is affine on the intervals $(x_{l_1}^*, x_1^*), \ldots, (x_{l_m}^*, x_{l_{m-1}+1}^*)$.

Remark 21. Chong proved in [3] that if $x \gg y$ and if $f : \mathbf{R} \to \mathbf{R}$ is strictly increasing and convex, then $x > y$ whenever equality holds in (23). This result follows immediately from Theorem 20, because if $x \gg y$ and $\sum_{i=1}^{n} x_i^* > \sum_{i=1}^{n} y_i^*$ and if in (23) equality holds, then there exists an interval of positive measure on which f is constant.

Remark 22. A further result of Chong [**3**] states that if $x \gg y$ and if f is strictly convex and increasing, then equality holds in (23) if and only if $x \sim y$, i.e. the components of x form a permutation of those of y. This result, too, follows from the preceding because, if there exists k such that $\sum_{i=1}^{k} x_i^* > \sum_{i=1}^{k} y_i^*$, and equality holds in (23), then there exists an interval of positive measure on which f is affine (Theorem 6 or 20).

7. Doubly stochastic matrices again. For $x, y \in \mathbf{R}^n$ we write $x \geqq y$ whenever $x_k \geqq y_k$ $(k = 1, \ldots, n)$, and we write $x > y$ whenever $x_k > y_k$ $(k = 1, \ldots, n)$. We shall need the following proposition, which follows directly from an analogous result of Mirsky [**8**].

PROPOSITION 23. *For $x, y \in \mathbf{R}^n$, there exists a doubly stochastic matrix M such that $Mx > y$ if, and only if,*

$$(26) \quad \sum_{1}^{k} x_k^* > \sum_{1}^{k} y_i^* \quad (k = 1, \ldots, n).$$

https://doi.org/10.4153/CJM-1977-058-8 Published online by Cambridge University Press

Proof. If (26) holds then it holds as well when we replace each y_i by $z_i = y_i + \epsilon$ ($\epsilon > 0$), provided ϵ is sufficiently small. But Mirsky has proved that

$$(27)\quad \sum_1^k x_i^* \geqq \sum_1^k z_i^* \quad (k = 1, 2, \ldots, n).$$

if and only if there exists a doubly stochastic M such that $Mx \geqq z$. For this M, $Mx > y$, since $z > y$.

On the other hand, if M is doubly stochastic and $Mx > y$, then, by Theorem 1, (27) holds for x and z where $z = Mx$, and since $z > y$ it is clear that (27) implies (26).

From Theorem 20 and from Proposition 23 we can deduce the following theorem.

THEOREM 24. *Let $x \in \mathbf{R}^n$ and $y \in \mathbf{R}^n$ be such that $x_1^* > y_1^*$ and $y \ll x$. Then the following statements are equivalent:*

(i) *There exists a matrix T doubly stochastic such that $y < Tx$.*

(ii) *If f is any increasing continuous convex function defined on* $[\min(x_n^*, y_n^*), x_1^*]$ *such that*

$$\sum_{i=1}^n f(y_i) = \sum_{i=1}^n f(x_i),$$

then f is constant on $[\min(x_n^*, y_n^*), x_1^*)$.

As an introduction to Theorem 26, we note the following corollary of Theorem 20.

PROPOSITION 25. *Let $x = (x_1, x_2) \in \mathbf{R}^2$ and $y = (y_1, y_2) \in \mathbf{R}^2$ be such that $x_1^* > y_1^*$ and $x_1^* + x_2^* > y_1^* + y_2^*$. If $f : [\min(x_2^*, y_2^*), x_1^*] \to \mathbf{R}$ is any increasing continuous convex function such that $f(x_1) + f(x_2) = f(y_1) + f(y_2)$, then f is constant on* $[\min(x_2^*, y_2^*), x_1^*]$.

THEOREM 26. *Let n be a fixed positive integer, $n \geqq 3$. Let $x \in \mathbf{R}^n$ and $y \in \mathbf{R}^n$ be such that $y \ll x$, $x_1^* > y_1^*$ and $\sum_{i=1}^n x_i^* > \sum_{i=1}^n y_i^*$. Then the following statements are equivalent:*

(i) *There exists a matrix T, doubly stochastic by blocks of size l and $n - l$, such that*

$$(28)\quad \begin{aligned} (Tx^*)_i &= y_i^* \quad \text{for } 1 \leqq i \leqq l \text{ and} \\ (Tx^*)_i &> y_i^* \quad \text{for } l + 1 \leqq i \leqq n. \end{aligned}$$

(ii) *There exists a non-constant increasing continuous convex function f defined on* $[\min(x_n^*, y_n^*), x_1^*]$ *such that*

$$(29)\quad \sum_{i=1}^n f(y_i) = \sum_{i=1}^n f(x_i).$$

Proof. We shall show first that (i) $\Rightarrow$ (ii). Let $T = T_1 \oplus T_2$. T_1 maps $x_1^*, \ldots, x_l^*$ to $y_1^*, \ldots, y_l^*$. Therefore it is evident that for any function f_1, which is affine on $[x_l^*, x_1^*]$, $\sum_{i=1}^{l} f_1(x_i^*) = \sum_{i=1}^{l} f_1(y_i^*)$. Since $x_1^* > y_1^*$ we see that $x_l^* < x_1^*$, i.e. the interval $[x_l^*, x_1^*]$ has positive measure. It is evident that for any function f_2 which is constant on $[\min(x_n^*, y_n^*), x_{l+1}^*]$ we have

$$\sum_{i=l+1}^{n} f_2(x_i^*) = \sum_{i=l+1}^{n} f_2(y_i^*).$$

Using the fact $x_{l+1}^* > y_{l+1}^*$ we see that the interval $[\min(x_n^*, y_n^*), x_{l+1}^*]$ has positive measure. The set $[\min(x_n^*, y_n^*), x_{l+1}^*] \cap [x_l^*, x_1^*]$ contains at most one point, hence there exists a non-constant increasing continuous convex function on $[\min(x_n^*, y_n^*), x_1^*]$ satisfying (29).

(ii) $\Rightarrow$ (i). By Theorem 20 we see that there exists at least one k such that $2 \leqq k \leqq n - 1$ and $\sum_{i=1}^{k} x_i^* = \sum_{i=1}^{k} y_i^*$. Let l be the maximum of the k's with that property. Then $(x_1^*, \ldots, x_l^*) \succ (y_1^*, \ldots, y_l^*)$, so that by Theorem 1 there exists a doubly stochastic matrix T_1 such that $(T_1x^*)_i = y_i^*$ for $1 \leqq i \leqq l$. From the choice of l it follows that

$$\sum_{i=1}^{k} x_{l+k}^* > \sum_{i=1}^{k} y_{l+k}^* \quad \text{for } 1 \leqq k \leqq n - l.$$

Therefore from Proposition 23 it follows that there exists a doubly stochastic matrix T_2 such that $T_2u_2 > v_2$, where $u_2 = (x_{l+1}^*, \ldots, x_n^*)$ and $v_2 = (y_{l+1}^*, \ldots, y_n^*)$. Thus $T = T_1 \oplus T_2$ satisfies (28).

8. Applications. In this section we develop some of the techniques discussed above so that they may be applied to prove and extend results of G. F. D. Duff on differences of rearranged sequences. After the first version of this paper was written we found the article of K. M. Chong [**2**], where a similar programme is carried out. The discussion below may still be of interest, for two reasons. Firstly, our arguments are based on a very general construction of vectors $x, y \in \mathbf{R}^n$ such that $x \gg y$, with an analysis of those cases where $x \succ y$ or $x \sim y$ (see Theorem 28). Secondly, in Chong [**2**, Theorem 2.6] the description of the cases of equality is incomplete.

A simple example of equality not covered by Chong's formulation may be given as follows (we refer to the notation of [**2**, Theorem 2.6]). Let $a = (1, 0, 1)$. Here the intervals of (a_k) and (a_k^*) are, respectively: (0, 1) and (0, 1); (1, 1) and (0, 1). Hence the left and right sides of Chong's inequality are, respectively: $2\Phi(0) + 2\Phi(1)$; $\Phi(1) + \Phi(1)$. Of course, these are equal for *any* Φ such that $\Phi(0) = 0$. A full description of the extreme cases must be as in the original results of Duff, and is reflected in our Theorem 29 and in Theorem 28 (c).

The results of Duff that concern us here are taken from [**4**] and [**5**] and may be summarized as follows.

THEOREM 27. *Let $(a_k)_1^N$ be a sequence of real numbers and denote by Δa_k the difference $a_k - a_{k+1}$ $(k = 1, \ldots, N - 1)$. Similarly, let Δa_k^* denote the a_k^* −*

$a_{k+1}{}^*$ *where* $(a_k{}^*)_1{}^N$ *is the decreasing rearrangement of* $(a_k)_1{}^N$. *Then*

$$(30)\quad \sum_{k=1}^{N-1} (\Delta a_k{}^*)^p \leqq \sum_{k=1}^{N-1} (|\Delta a_k|)^p$$

for all $p \in [1, \infty)$. *Furthermore, if* n_k *denotes the number of indices* j *such that the interval*

$$(\min (a_j, a_{j+1}), \max (a_j, a_{j+1}))$$

contains the interval $(a_{k+1}{}^*, a_k{}^*)$, *then*

$$(31)\quad \sum_{k=1}^{N-1} n_k(\Delta a_k{}^*)^p \leqq \sum_{k=1}^{N-1} (|\Delta a_k|)^p$$

for all $p \in [1, \infty)$ *and the reverse inequality holds for all* $p \in (0, 1)$. *Equality holds in* (31) *if, and only if,* $p = 1$ *or, for each* j, *the values* a_j *and* a_{j+1} *are adjacent in the reordered set* $\{a_k{}^*\}_1{}^N$ (i.e., *no* a_k *lies strictly between* a_j *and* a_{j+1}).

We shall prove this theorem (and more) by constructing the appropriate m-tuples x, y such that $x \gg y$, and applying Theorems 6, 20. First we present the general construction of x, y promised above. Suppose we have a set function λ defined on a certain class $\mathscr{S}$ of subsets of a fixed set Ω, and assuming non-negative values. We assume that

$$(32)\quad \lambda(S) \geqq \sum_1^m \lambda(S_k)$$

whenever $S, S_1, \ldots, S_m \in \mathscr{S}$, S contains each S_k, and the S_k are disjoint. For example, λ could be any (non-negative) measure, or, as will be the case in our application to Theorem 27, λ might be the length function defined on the class $\mathscr{S}$ of all (bounded) real intervals.

THEOREM 28. *Let* $\lambda, \mathscr{S}$ *be as above and suppose that* $D_1, \ldots, D_p$ *are disjoint sets in* $\mathscr{S}$, *and that* $E_1, \ldots, E_q$ *are also sets in* $\mathscr{S}$. *Let* $R_i = \{j : E_j \supset D_i\}$, *and let* $n_i = \#(R_i)$ $(i = 1, \ldots, p)$. *If there are integers* m_i *such that* $0 \leqq m_i \leqq n_i$ *and,* $q \leqq m = \sum_i^p m_i$, *we define* $x, y \in \mathbf{R}^m$ *as follows:*

$$x = (\lambda(E_1), \ldots, \lambda(E_q), 0, \ldots, 0) \quad (\in \mathbf{R}^m) \text{ and}$$

$$y = (\lambda(D_1), \ldots, \lambda(D_1), \lambda(D_2), \ldots, \lambda(D_2), \lambda(D_3), \ldots),$$

where, in y, $\lambda(D_i)$ *is repeated* m_i *times* $(i = 1, \ldots, p)$. *Then:*

(a) *we always have the relation* $x \gg y$;

(b) $x \succ y$ *provided that*

$$(33)\quad m_i = n_i \quad \text{for each } i \text{ such that } \lambda(D_i) > 0,$$

and, for each j

$$(34)\quad \lambda(E_j) = \sum_{D_i \subset E_j} \lambda(D_i);$$

(c) $x \sim y$ (i.e., $y = \sigma(x)$ *for some permutation* σ) *if, and only if,* (33) *holds and*

(35) $\lambda(E_j) > 0 \Rightarrow \lambda(D_i) = \lambda(E_j)$ *for some* $D_i \subset E_j$.

Proof. (a) For each $k \leqq m$, $\sum_1^k y_i^*$ has the form $\sum_1^p b_i \lambda(D_i)$, where the integer b_i satisfies $0 \leqq b_i \leqq m_1$, and $\sum_1^p b_i = k$. Let B_i be a subset of R_i such that $\#(B_i) = b_i$ $(i = 1, \ldots, p)$, and denote $\bigcup_1^p B_i$ by B. Then $c = \#(B) \leqq k$, and since each $x_j \geqq 0$,

$$\sum_1^k x_j^* \geqq \sum_1^c x_j^* \geqq \sum_{j \in B} \lambda(E_j).$$

Moreover, by (32),

$$(36) \quad \sum_{j \in B} \lambda(E_j) \geqq \sum_{j \in B} \left(\sum_{\substack{i \\ j \in B_i}} \lambda(D_i) \right).$$

Regrouping the last expression we obtain

$$\sum_i \#\{j \in B : j \in B_i\} \lambda(D_i) = \sum_i b_i \lambda(D_i).$$

(b) We must verify that $\sum_1^m x_j = \sum_1^m y_j$ under the conditions (33) and (34). Taking $B_i = R_i$ in the argument above, (36) becomes equality by (34), so that

$$(37) \quad \sum_{j \in B} \lambda(E_j) = \sum_i n_i \lambda(D_i).$$

But (33) says that $m_i = n_i$ unless $\lambda(D_i) = 0$. Moreover $\lambda(E_j) > 0$ only if $j \in R_i$ for some i, in view of (34). Since $B = \bigcup_1^p R_i$ in (37), the verification is complete.

(c) Suppose first that we have (33) and (35). Consider any i such that $\lambda(D_i) > 0$. Then for each $j \in R_i$ we must have $\lambda(E_j) = \lambda(D_i)$ since otherwise (35) would require some $i' \neq i$ such that $D_{i'} \subset E_j$ and $\lambda(D_{i'}) = \lambda(E_j)$ while (32) would imply that

$$\lambda(E_j) \geqq \lambda(D_{i'}) + \lambda(D_i).$$

Moreover, for $i \neq i'$ such that $\lambda(D_i), \lambda(D_{i'}) > 0$, we must have $R_i \cap R_{i'} = \emptyset$ since $j \in R_i \cap R_{i'}$ implies, as we have just seen,

$$\lambda(E_j) = \lambda(D_i) = \lambda(D_{i'})$$

while, again by (32), we would also have

$$\lambda(E_j) \geqq \lambda(D_i) + \lambda(D_{i'}).$$

Hence whenever $\lambda(D_i) > 0$, the n_i occurrences of $\lambda(D_i)$ in y can be matched up with the $\lambda(E_j)$ in x corresponding to $j \in R_i$. The components of y that are left are all 0, and it only remains to show that the same is true of x, i.e., that

$$j \notin \bigcup \{R_i : \lambda(D_i) > 0\} \Rightarrow \lambda(E_j) = 0.$$

This is evident from (35).

Now assuming that $y = \sigma(x)$, we must verify (33) and (35). We may assume that the D_i are indexed so that

$$\lambda = \lambda(D_1) \geqq \lambda(D_2) \geqq \ldots \geqq \lambda(D_p).$$

Let us further suppose that

$$\lambda = \lambda(D_1) = \lambda(D_2) = \ldots = \lambda(D_u) > \lambda(D_{u+1}).$$

Now for each $i \leqq u$ we argue as follows. For each $j \in R_i$, $\lambda(E_j) \geqq \lambda(D_i) = \lambda$; but λ dominates all the components of $y = \sigma(x)$ so that $\lambda \geqq \lambda(E_j)$. Hence $j \in R_i$ implies $\lambda(E_j) = \lambda$. Furthermore, for any $i' \neq i$ such that $\lambda(D_{i'}) > 0$, $R_{i'} \cap R_i = \emptyset$, since (32) implies that

$$\lambda = \lambda(E_j) \geqq \lambda + \lambda(D_{i'}),$$

if $j \in R_{i'} \cap R_i$. In particular, there are $n_1 + n_2 + \ldots + n_u$ distinct j such that $\lambda(E_j) = \lambda$. Since $y = \sigma(x)$, we must also have this number of occurrences of λ in y; this requires that $m_i = n_i$ for each $i \leqq u$. Since there are at *most* $n_1 + \ldots + n_u$ occurrences of λ in y, the same is true of x and we conclude that:

$$\lambda(E_j) = \lambda \Rightarrow j \in \textstyle\bigcup_{i \leqq u} R_i \Rightarrow \text{(35) holds for } E_j.$$

Provided that $\lambda(D_{u+1}) > 0$, we can argue in a similar fashion concerning those i' such that $\lambda(D_{i'}) = \lambda(D_{u+1})$. Note that we have established above that $R_i \cap R_{i'} = \emptyset$ if $i \leqq u$. Continuing in this way we establish (33) and (35) for all i, j for which $\lambda(D_i), \lambda(E_j) > 0$.

Finally we show how a special case of Theorem 28 allows us to generalize the inequalities of Duff contained in Theorem 27. As explained at the beginning of this section, a variant of the following result has been proved by Chong in [**2**].

THEOREM 29. *Let Δa_k, Δa_k^* be as in Theorem 27. Then for any convex, increasing (non-decreasing) function $f : [0, \max_k |\Delta a_k|] \to \mathbf{R}$,*

$$\text{(38)} \quad \sum_{k=1}^{N-1} f(\Delta a_k^*) \leqq \sum_{k=1}^{N-1} f(|\Delta a_k|).$$

Furthermore, if n_k is defined as in Theorem 27, then

$$\text{(39)} \quad \sum_{k=1}^{N-1} n_k f(\Delta a_k^*) \leqq \sum_{k=1}^{N-1} f(|\Delta a_k|) + f(0) \sum_{k=1}^{N-1} (n_k - 1),$$

for any convex function $f : [0, \max_k |\Delta a_k|] \to \mathbf{R}$, and the reverse inequality holds if f is concave. If f is strictly convex, then equality holds in (39) if, and only if, no a_k lies strictly between a_j and a_{j+1} ($j = 1, 2, \ldots, N - 1$).

Proof. We begin with two simple remarks. The statement above concave f follows immediately from the convex case applied to $-f$. Secondly, the function f may be replaced by a function that is continuous and defined throughout

R without changing the values at 0, Δa_1, Δa_1^*, $\Delta a_2, \ldots$, the (strict) convexity, of the increasing property.

For the application of Theorem 28 we let $\lambda(I)$ denote the length of an interval I and let $D_1, \ldots, D_{N-1}$ be the (disjoint) intervals $(a_2^*, a_1^*), \ldots, (a_N^*, a_{N-1}^*)$. Furthermore, we let

$$E_j = (\min (a_j, a_{j+1}), \max (a_j, a_{j+1})) \quad (j = 1, \ldots, N-1).$$

Clearly $\lambda(D_i) = \Delta a_i^*$ and $\lambda(E_j) = |\Delta a_j|$. Evidently the definitions of n_k in Theorems 27 and 28 are in harmony with our notation. Now (38) follows by Theorem 20, because Theorem 28 (a) ensures that

$$(\Delta a_i^*, \ldots, \Delta a_{N-1}^*) \ll (|\Delta a_1), \ldots, |\Delta a_{N-1}|).$$

In this application of Theorem 28 we must take $m_i = 1$, so that it is important to note that $n_i \geqq 1$ $(i = 1, \ldots, N-1)$ in this example. In fact, if the end-points of $D_i = (a_{i+1}^*, a_i^*)$ are a_k and $a_{k'}$ with $k < k'$, then at least one of intervals E_j $(j = k, \ldots, k'-1)$ must "jump across" D_i, so that $j \in R_i$. It is also clear that (34) is satisfied for this example so that (39) is a direct application of Theorem 28 (b) and Theorem 6. Theorem 6 (or Theorem 17) also tells us that, when f is strictly convex, (39) is an equality only if (with the λ, D_i, E_j of this example and each $m_i = n_i$) $x \sim y$ in Theorem 28. Hence, by part (c) of that theorem, (35) must hold. Evidently this requires that no a_k lie strictly between a_j and a_{j+1}.

Remark 30. Elementary examples show that we cannot dispense with the convexity of f in the inequality (39). In the simpler inequality (39), however, the convexity condition enters only as an accidental result of our method of proof (via Theorem 28). In fact, for (38) it is only necessary for $f : [0, \max_k |\Delta a_k|] \to \mathbf{R}$ to be non-decreasing; this result may be derived, for example, from A. M. Garsia [**6**, Theorem 2.1].

References

1. D. Blackwell, *Equivalent comparisons of experiments*, Ann. Math. Stat. *24* (1953), 265–272.

2. K. M. Chong, *Variation reducing properties of decreasing rearrangements*, Can. J. Math. *27* (1975), 330–336.

3. ——— *An induction theorem for rearrangements*, Can. J. Math. *28* (1976), 154–160.

4. G. F. D. Duff, *Differences, derivatives, and decreasing rearrangements*, Can. J. Math. *19* (1967), 1153–1178.

5. ——— *Integral inequalities for equimeasurable rearrangements*, Can. J. Math. *22* (1970), 408–430.

6. A. M. Garsia, *Combinatorial inequalities and smoothness of functions*, Bull. Amer. Math. Soc. *82* (1976), 157–170.

7. G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, 2nd edition (Cambridge U.P., 1952).

8. L. Mirsky, *On a convex set of matrices*, Archiv. der Math. *10* (1959), 88–92.

9. G. Pólya, *Remark on Weyl's note*: *Inequalities between the two kinds of eigenvalues of a linear transformation*, Proc. Nat. Acad. Sci. *36* (1950), 49–51.

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

11. I. Schur, *Uber eine Klasse von Mittelbildungen mit Anwendungen auf die Determinantentheorie*, Sitzungsber. d. Berl. Math. Gesellsch. *22* (1923), 9–20.

12. S. Sherman, *On a theorem of Hardy, Littlewood, Pólya, and Blackwell*, Proc. Nat. Acad. Sci. *37* (1951), 826–831.

University of Guelph,
Guelph, Ontario