Can. J. Math.Vol. 44 (1), 1992 pp. 54–74

LOWER BOUNDS FOR MATRICES, II

GRAHAME BENNETT

ABSTRACT. Our main result is the following monotonicity property for moment sequences $\boldsymbol{\mu}$. Let p be fixed, $1 \le p < \infty$: then

$$\frac{1}{r}\sum_{n=r}^{\infty}\Big(\sum_{k=0}^{r-1}\binom{n}{k}\Delta^{n-k}\mu_k\Big)^p$$

is an increasing function of r ($r = 1, 2, \ldots$). From this we derive a sharp lower bound for an arbitrary Hausdorff matrix acting on ℓ^p. The corresponding upper bound problem was solved by Hardy.

1. **Introduction.** We shall be concerned with the spaces ℓ^p, $1 \le p < \infty$, of sequences of real numbers satisfying

$$\|\mathbf{x}\|_p = \Big(\sum_{k=0}^{\infty}|x_k|^p\Big)^{1/p} < \infty. \tag{1}$$

We seek lower bounds of the form

$$\|A\mathbf{x}\|_p \ge L\|\mathbf{x}\|_p, \tag{2}$$

valid for every $\mathbf{x} \in \ell^p$ with $x_0 \ge x_1 \ge \cdots \ge 0$. Here A is a matrix with non-negative entries, assumed to map ℓ^p into itself, and L is a constant not depending on $\mathbf{x}$. Results of this type may be found in [1], [6], [12], [13], [16], [17] and [20].

The general lower bound problem is solved in [1], where it is shown that the best possible value of L is given by

$$L = \inf\{f^{1/p}(r) : r = 0, 1, \ldots\}, \tag{3}$$

with

$$f(r) = \frac{1}{r+1}\sum_{n=0}^{\infty}\Big(\sum_{k=0}^{r}a_{n,k}\Big)^p. \tag{4}$$

The problem comes to life again, however, when we attempt to evaluate (3) for specific examples; indeed, the infimum may be intractable even for "nice" matrices A.

It turns out, in rather surprisingly many cases, that the function $f(r)$ actually increases with r. These are the easiest cases to handle, for the infimum in (3) is attained at $r = 0$,

Research supported in part by National Science Foundation grant NSF DMS–8702358
Received by the editors August 11, 1989; revised May 29, 1990.
AMS subject classification: 26D15, 40G05, 47A30, 47B37.

and the lower bound, L, is just the ℓ^p-norm of the first column of A. Most of our results are of this type.

The real challenge, then, is to study the monotonicity properties of the functions given by (4). We shall see that this leads to several intriguing elementary inequalities, quite unlike any that appear in the literature. The main tool used in providing such inequalities is the theory of majorization.

Only two lower bounds were worked out in detail in [1]. We restate these results here in a form not involving the monotonicity of $\mathbf{x}$. The first provides a natural complement to *Hardy's inequality* ([9], § 9.8), the second to *Hilbert's* ([9], Chapter IX).

Let $\mathbf{x}$ be a sequence of non-negative numbers and suppose that $p > 1$. If $\mathbf{x} \in \ell^p$, then

$$\sum_{n=0}^{\infty}\Big(\frac{1}{n+1}\sum_{k=0}^{n} x_k\Big)^p \geq \zeta(p)\sum_{n=0}^{\infty}\min_{k\leq n} x_k^p, \tag{5}$$

and

$$\sum_{n=0}^{\infty}\Big(\sum_{k=0}^{\infty}\frac{x_k}{n+k+1}\Big)^p \geq \zeta(p)\sum_{n=0}^{\infty}\min_{k\leq n} x_k^p. \tag{6}$$

In both cases the constant $\zeta(p)\big(=\sum_{k=1}^{\infty} k^{-p}\big)$ is best possible, and there is equality only when $x_1 = x_2 = \cdots = 0$.

The special case, $p = 2$, of inequality (5) was conjectured by Axler and Shields, and proved by Lyons in [13]. The general case was discovered simultaneously by Bennett [1] and Renaud [17]. A related result, for $0 < p < 1$, which provides a natural complement to the *Copson-Elliott inequalities* ([9], Theorems 338, 344 and 345), is given in [3].

In [1], page 90, we raised the problem of evaluating (3) for other "classical" matrices, A. This has been pursued by Rhoades [20], and by Lenard [12], whose results are described in more detail below.

The main purpose of this paper is to solve the lower bound problem for *Hausdorff matrices*. These have entries, $h_{n,k}$, $n, k = 0, 1, \ldots$, of the form

$$h_{n,k} = \begin{cases} \binom{n}{k}\Delta^{n-k}\mu_k & (k \leq n) \\ 0 & (k > n) \end{cases}, \tag{7}$$

where $\boldsymbol{\mu} = (\mu_k)_{k=0}^{\infty}$ is a sequence of real numbers, normalized so that $\mu_0 = 1$, and where Δ is the *forward difference operator*,

$$\Delta\mu_k = \mu_k - \mu_{k+1}. \tag{8}$$

The theory of Hausdorff matrices is described in [5], [8], [24], [25].

We are interested here only in matrices with non-negative entries and so we take $\boldsymbol{\mu}$ to be a *totally monotone sequence*, namely

$$\Delta^n\mu_k \geq 0 \quad (n, k = 0, 1, \ldots). \tag{9}$$

Now (9) holds, according to a fundamental theorem of Hausdorff, [10], precisely when $\boldsymbol{\mu}$ is a *moment sequence*

$$\mu_k = \int_0^1 \theta^k \, d\mu(\theta) \quad (k = 0, 1, \ldots) \tag{10}$$

associated with some (Borel) probability measure, $d\mu(\theta)$, on [0, 1]. Thus we may rewrite (7) in the equivalent form

$$(11)\qquad h_{n,k} = \begin{cases} \binom{n}{k}\int_0^1 \theta^k(1-\theta)^{n-k}\,d\mu(\theta) & (k \le n) \\ 0 & (k > n) \end{cases}.$$

Taking $d\mu(\theta)$ to be Lebesgue measure, we obtain the *Cesàro matrix* (of order 1), so that our main result, Theorem 1, is an extension of (5). Other choices (see [8], Chapter 11) lead to (C, α), the *Cesàro matrix of order* α ($\mu(\theta) = 1 - (1-\theta)^\alpha$), to (H, α), the *Hölder matrix of order* α ($\Gamma(\alpha)d\mu(\theta) = |\log\theta|^{\alpha-1}d\theta$), and to the *Euler matrices* (E, α) ($d\mu(\theta)$ = point evaluation at $\theta = \alpha$). Thus the lower bound problem is completely solved for these classes of matrices.

There are connections—already apparent in [10]—between Hausdorff matrices and probability theory (Bernoulli trials, exchangeable events), and our inequalities have applications in this direction too (see section 7 and, especially, [4]).

The main obstacle to computing lower bounds for Hausdorff matrices is the fact that no closed-form expression is known for the partial row-sum, $\sum_{k=0}^r h_{n,k}$, which appears in (4). This difficulty can be avoided, however, if p is a positive integer, for we may rewrite $f(r)$ as a p-fold sum and then interchange the order of summation. Lenard [12] adopts this approach to evaluate (3) for Euler matrices acting on $\ell^p (p = 1, 2, \ldots)$. His subsequent analysis is quite involved and requires skilful manipulations with generating functions. Our method has nothing in common with his, but we owe him a great deal just the same, for it was his *result* that led us to suspect the validity of the more powerful inequalities (Theorems 1 and 5) below. Lenard's analysis is given as an appendix to this paper.

Our main result is proved in sections 3 and 4 after some preliminary remarks on the method of proof in section 2. Quasi-Hausdorff matrices are studied in section 5, and weighted means in section 6. Some additional inequalities concerning moment sequences are given in section 7.

2. **Background on majorization.** We wish to show (see (4) and (11)) that the expression

$$\frac{1}{r+1}\sum_{n=0}^{\infty}\Big(\sum_{k=0}^{r}\int_0^1\binom{n}{k}\theta^k(1-\theta)^{n-k}\,d\mu(\theta)\Big)^p$$

increases with r. The first step is to eliminate p. We do this by appealing to the *theory of majorization*, which is described briefly below. This step, incidentally, allows us also to eliminate $d\mu(\theta)$, and leaves us with a "no-frills" inequality, (20), involving just the binomial probabilities: $\binom{n}{k}\theta^k(1-\theta)^{n-k}$. The inequality is proved by first establishing a rather mysterious polynomial identity, Lemma 1, and by then ejecting certain of its terms. (There ought to be a simple probabilistic proof of the identity, but I have been unable to discover one such.)

The theory of majorization is concerned with inequalities of the type

$$(12)\qquad \phi(x_1) + \cdots + \phi(x_N) \le \phi(y_1) + \cdots + \phi(y_N).$$

https://doi.org/10.4153/CJM-1992-003-9 Published online by Cambridge University Press

Here $\mathbf{x}$ and $\mathbf{y}$ are fixed N-tuples with non-negative entries, and the inequality holds for all continuous, convex functions ϕ (whose domain of definition includes the x's and the y's). If (12) holds, we say that $\mathbf{x}$ is *majorized* by $\mathbf{y}$, and we write $\mathbf{x} \preceq \mathbf{y}$. An excellent account of these ideas is given in the monograph [14].

The importance of the theory is due, in large part, to the following result of Hardy, Littlewood and Polya ([9], Theorem 108).

THEOREM HLP. *Let* $\mathbf{x}$ *and* $\mathbf{y}$ *be N-tuples with non-negative entries. Then* $\mathbf{x} \preceq \mathbf{y}$ *if both the following conditions hold.*

$$x_1^* + \cdots + x_k^* \leq y_1 + \cdots + y_k \quad (1 \leq k \leq N) \tag{13}$$

$$x_1 + \cdots + x_N = y_1 + \cdots + y_N, \tag{14}$$

where the x^**'s are the* x*'s arranged in descending order.*

Majorization is used in sections 6 and 7 (Lemma 8 and Theorem 6), but a less stringent condition is required for our main result. Following Marshall and Olkin, [14], we say that $\mathbf{x}$ is *weakly submajorized* by $\mathbf{y}$, and we write $\mathbf{x} \prec_W \mathbf{y}$, if condition (13) is satisfied. It then follows by a theorem of Tomic [22] (see also [14], page 66) that (12) holds for all increasing, continuous, convex functions ϕ. Polya, [15], has shown that Tomic's result is a consequence of Theorem HLP; in fact, the two results are equivalent.

3. **An identity.**

Let N and r be positive integers. Determine integers a and α by

$$N = a(r+1) + \alpha \quad (0 \leq \alpha < r+1) \tag{15}$$

and integers b and β by

$$N = br + \beta \quad (0 \leq \beta < r). \tag{16}$$

Then we have the following identity.

LEMMA 1.

$$(r+1)\sum_{n=0}^{a-1}\sum_{k=0}^{r-1} h_{n,k} + \alpha \sum_{k=0}^{r-1} h_{a,k} + r \sum_{n=a+1}^{b-1} (r+1-N/n)h_{n,r} + \frac{\beta}{b}(b-r)h_{b,r}$$
$$= r\sum_{n=0}^{b-1}\sum_{k=0}^{r} h_{n,k} + \beta \sum_{k=0}^{r} h_{b,k},$$

where the entries $h_{n,k}$ *are given by (11).*

It will be convenient to adopt the following notation

$$L = (r+1)\sum_{n=0}^{a-1}\sum_{k=0}^{r-1} h_{n,k} + \alpha \sum_{k=0}^{r-1} h_{a,k}, \tag{17}$$

$$R = r\sum_{n=0}^{b-1}\sum_{k=0}^{r} h_{n,k} + \beta \sum_{k=0}^{r} h_{b,k}, \tag{18}$$

$$D = r \sum_{n=a+1}^{b-1} (r+1-N/n)h_{n,r} + \frac{\beta}{b}(b-r)h_{b,r}. \tag{19}$$

https://doi.org/10.4153/CJM-1992-003-9 Published online by Cambridge University Press

(L stands for "left", R for "right", and D for "difference", the last term being justified by Lemma 1). Our ultimate goal is to prove the inequality

$$L \le R. \tag{20}$$

This follows from Lemma 1 since $D \ge 0$ (by (15), (16) and (11)).

Lemma 1 is proved by manipulating an easier identity, Lemma 3 below. The argument proceeds most efficiently if we agree to work with Euler matrices first, and then switch to Hausdorff matrices by averaging over $0 \le \theta \le 1$ with respect to $d\mu(\theta)$. We recall that the entries of an Euler matrix, $E = E(\theta)$, are given by

$$e_{n,k} = \begin{cases} \binom{n}{k}\theta^k(1-\theta)^{n-k} & (0 \le k \le n) \\ 0 & (k > n) \end{cases}. \tag{21}$$

Three of our identites, Lemmas 3–5, involve the parameter θ explicitly, and this must be eliminated before the averaging can be effected. For this step we need the following

LEMMA 2. *$(k+1)e_{n+1,k+1} = \theta(n+1)e_{n,k}$ whenever $n, k \ge 0$.*

PROOF. The result follows immediately from (21).

LEMMA 3. *$\sum_{k=0}^{r-1} e_{n,k} + \theta \sum_{j=0}^{n-1} e_{j,r-1} = 1$ whenever $n, r \ge 1$.*

PROOF. The rows of E are "binomial probability vectors" so that the first term above, $\sum_{k=0}^{r-1} e_{n,k}$, may be interpreted as the probability of "*less than r successes in n trials*". The columns of E, all with sums $1/\theta$, admit a similar interpretation if we first multiply each by θ. Indeed, the kth column of E (times θ) represents "the waiting time for $(k+1)$ successes" $(k = 0, 1, \ldots)$. Thus the second term above, $\theta \sum_{j=0}^{n-1} e_{j,r-1}$, is the probability that *the waiting time for r successes does not exceed n*, or, equivalently, that *at least r successes occur in n trials*. The two italicized events are complementary, and hence their probabilities add to one.

It would be most satisfying to have a similar probabilistic interpretation of Lemmas 1 and 6.

LEMMA 4. *$\sum_{k=0}^{r-1} e_{n,k} = \sum_{k=0}^{r-1} e_{m,k} + \theta \sum_{j=n}^{m-1} e_{j,r-1}$ whenever $n, r \ge 0$, and $m \ge n$.*

PROOF. Apply Lemma 3 twice; once as it stands, and once with n replaced by m.

LEMMA 5. *$\sum_{j=0}^{n-1} e_{j,r-1} \sum_{j=0}^{n-1} e_{j,r} + e_{n,r}/\theta$ whenever $n, r \ge 1$.*

PROOF. Apply Lemma 3 twice; once as it stands, and once with r replaced by $r+1$. (We remark that $e_{n,r}$ is divisible by θ, so that Lemma 5 holds even when $\theta = 0$.)

LEMMA 6. *$\sum_{n=0}^{s-1} \sum_{k=0}^{r-1} h_{n,k} = s \sum_{k=0}^{r-1} h_{s,k} + r \sum_{n=0}^{s} h_{n,r}$ whenever $r, s \ge 1$.*

PROOF. We prove this just for the Euler matrices, (21); the general result follows by integrating over $0 \le \theta \le 1$ with respect to $d\mu(\theta)$.

https://doi.org/10.4153/CJM-1992-003-9 Published online by Cambridge University Press

Applying Lemma 4 with m replaced by s, we have

$$\sum_{n=0}^{s-1}\sum_{k=0}^{r-1} e_{n,k} = \sum_{n=0}^{s-1}\Big(\sum_{k=0}^{r-1} e_{s,k} + \theta \sum_{j=n}^{s-1} e_{j,r-1}\Big)$$
$$= s\sum_{k=0}^{r-1} e_{s,k} + \theta \sum_{j=0}^{s-1}(j+1)e_{j,r-1},$$

and the result follows from Lemma 2.

Lemma 2 enables us to eliminate θ from Lemmas 3–5 as well, and thus to obtain versions of these results for Hausdorff matrices. The details are left to the reader.

PROOF OF LEMMA 1. Again, it suffices to consider Euler matrices. By applying Lemma 6 to both R and to L, and by recalling (15) and (16), we see that

$$R - L = N\sum_{k=0}^{r} e_{b,k} + r(r+1)\sum_{n=0}^{b} e_{n,r+1} - N\sum_{k=0}^{r-1} e_{a,k} - r(r+1)\sum_{n=0}^{a} e_{n,r}.$$

Applying Lemma 4, then Lemma 5, then Lemma 2 (twice), gives

$$R - L = Ne_{b,r} - N\theta\sum_{n=a}^{b-1} e_{n,r-1} + r(r+1)\Big(\sum_{n=0}^{b} e_{n,r+1} - \sum_{n=0}^{a} e_{n,r}\Big)$$
$$= Ne_{b,r} - N\theta\sum_{n=a}^{b-1} e_{n,r-1} + r(r+1)\sum_{n=a+1}^{b} e_{n,r} - \frac{r(r+1)}{\theta} e_{b+1,r+1}$$
$$= Ne_{b,r} - Nr\sum_{n=a+1}^{b} e_{n,r}/n + r(r+1)\sum_{n=a+1}^{b} e_{n,r} - (b+1)re_{b,r}.$$

Recalling (16) and (19), we see that $R - L = D$.

4. **Hausdorff matrices.** In this section we solve the lower bound problem for Hausdorff matrices. We recall that the entries, $h_{n,k}$, are given by (11).

It may be worthwhile to point out here that the corresponding *upper bound problem* is easily solved by means of a result of Hardy (see Theorem H, below). The problem is to determine the smallest constant, U, such that

$$\|H\mathbf{x}\|_p \le U\|\mathbf{x}\|_p \tag{22}$$

for every $\mathbf{x} \in \ell^p$ satisfying $x_0 \ge x_1 \ge \cdots \ge 0$. It turns out—and this is a consequence of Hardy's proof—that $U = \|H\|_{p,p}$, the (operator) *norm* of H. Thus the upper bound problem is solved by equation (23) below.

THEOREM H. *Let p be fixed, $1 \le p < \infty$, and let H be the Hausdorff matrix given by (11). Then H is bounded on ℓ^p if and only if $\int_0^1 \theta^{-1/p}\, d\mu(\theta) < \infty$, and we have*

$$\|H\|_{p,p} = \int_0^1 \theta^{-1/p}\, d\mu(\theta). \tag{23}$$

Theorem H is proved in [7] (see also Theorem 216 of [8]). The proof is based on the special version of (23) for Euler matrices, due to Bochner and Knopp (see [11], Satz II, and the footnote to page 19).

We now come to our main result.

https://doi.org/10.4153/CJM-1992-003-9 Published online by Cambridge University Press

THEOREM 1. *Let p be fixed, $1 \leq p < \infty$, and suppose that the Hausdorff matrix, H, is bounded on ℓ^p. Then*

$$\|H\mathbf{x}\|_p \geq L\|\mathbf{x}\|_p \tag{24}$$

for every $\mathbf{x} \in \ell^p$ satisfying $x_0 \geq x_1 \geq \cdots \geq 0$, where

$$L^p = \sum_{n=0}^{\infty}\left(\int_0^1 (1-\theta)^n \, d\mu(\theta)\right)^p. \tag{25}$$

There is equality in (24) only if $x_1 = x_2 = \cdots = 0$ or if $p = 1$ or if $d\mu(\theta)$ is the point mass at 1.

PROOF. The value of L determined by (25) is the ℓ^p-norm of the first column of H. Thus (24) follows form (3) and (4) once we show that the function $f(r)$ increases with r. To do this, we consider the more general inequality

$$(r+1)\sum_{n=0}^{\infty}\phi\left(\sum_{k=0}^{r-1} h_{n,k}\right) \leq r\sum_{n=0}^{\infty}\phi\left(\sum_{k=0}^{r} h_{n,k}\right), \tag{26}$$

for $r = 1, 2, \ldots$, where ϕ is any increasing convex function defined on $[0, \infty)$.

Fixing r, we denote by $\mathbf{x}$ the sequence, $\left(\sum_{k=0}^{r-1} h_{n,k}\right)_{n=0}^{\infty}$, with each term repeated $(r+1)$ times, and by $\mathbf{y}$ the sequence, $\left(\sum_{k=0}^{r} h_{n,k}\right)_{n=0}^{\infty}$, each term repeated r times. Then (26) is equivalent, via Tomic's theorem (see section 2), to the assertion that $\mathbf{x}$ be weakly submajorized by $\mathbf{y}$; in other words, to

$$\sum_{n=1}^{N} x_n^* \leq \sum_{n=1}^{N} y_n \quad (N = 1, 2, \ldots). \tag{27}$$

Given N, we determine integers a, α, b, β, as in (15) and (16), and note that (27) is then equivalent to (20). We have used here the fact that the expression $\sum_{k=0}^{r} h_{n,k}$ with r fixed, is a decreasing function of n. This may be deduced from Lemma 4 by dropping the "θ-term", and by integrating over $0 \leq \theta \leq 1$.

The value of the constant given by (25) is obviously the best possible, for we may take $\mathbf{x}$ in (24) to be the sequence $\mathbf{x} = (1, 0, 0, \ldots)$. The last clause of the theorem, concerning cases of equality, is a consequence of the results of [1]. We do not give the details here.

It is natural to assume, in Theorem 1, that H maps ℓ^p into itself, for then all the terms (24) are finite whenever $\mathbf{x} \in \ell^p$. This assumption however, is not necessary, provided that we agree to interpret (24) as follows: *if the left side is finite, for some* $\mathbf{x}$, *so is the right, and inequality (24) holds.* (Of course, we must assume that $L < \infty$ in order to get a meaningful result for any non-zero $\mathbf{x}$.)

We have already observed that inequality (5) is a special case of Theorem 1—obtained by taking H to be the $(C, 1)$ matrix. Rhoades ([20], Theorem 2) studies the $(C, 2)$ matrix, showing that its lower bound is also "attained in the first column". His analysis, however, is very complicated, and applies only to the case $p = 2$. He conjectures that the same result should hold true for all $p > 1$, and, indeed, for all the matrices (C, α), $\alpha > 0$. Theorem 1 confirms this conjecture.

5. **Quasi-Hausdorff matrices.** In this section we study the lower bound problem for *quasi-Hausdorff matrices*. These are just the transposes, H^t, of the Hausdorff matrices, (11). The terminology is that of Hardy ([8], section 11.19).

The ℓ^p-mapping properties of quasi-Hausdorff matrices may be determined from Hardy's result (Section 4), by using the familiar relationship

$$\|H^t\|_{p,p} = \|H\|_{p^*,p^*}, \tag{28}$$

where $p^* = p/(p-1)$ denotes the *conjugate* exponent to p.

There is no similar relationship between the lower bound of a matrix and that of its transpose. This can be seen readily by considering the Cesàro matrix, C. Inequality (5) shows that the lower bound is $\zeta(p)^{1/p}$, while Renaud ([17], Theorem 2) has shown that the lower bound for C^t is 1.

Renaud's proof depends in an essential way on the special structure of the Cesàro matrix. We give here a simple alternative proof—one that leads to a much more general result. For the statement of our theorem it will be convenient to call a matrix, A, a *quasi-summability matrix* if A is upper triangular and has column sums 1. This class includes the transposes of the Hausdorff matrices, of the *weighted means* (Section 6), and of the *Norlund means* (see [8], [25] for definitions).

THEOREM 2. *Let p be fixed, $1 \le p < \infty$, and let A be a quasi-summability matrix. If $\mathbf{x} \in \ell^p$ satisfies $x_0 \ge x_1 \ge \cdots \ge 0$, then*

$$\|A\mathbf{x}\|_p \ge \|\mathbf{x}\|_p. \tag{29}$$

There is equality in (29) when and only when (at least) one of the following conditions is satisfied: $p = 1$; $A = I$, the identity matrix; the first n columns of A coincide with those of I, and $x_n = x_{n+1} = \cdots = 0$.

PROOF. According to (4) and (5), we must show that

$$r+1 \le \sum_{n=0}^{r}\Big(\sum_{k=n}^{r} a_{n,k}\Big)^p \tag{30}$$

for every natural number r. We have

$$\begin{aligned} r+1 &= \sum_{k=0}^{r}\sum_{n=0}^{k} a_{n,k} \\ &= \sum_{n=0}^{r}\sum_{k=n}^{r} a_{n,k} \\ &\le (r+1)^{1/p^*}\Big(\sum_{n=0}^{r}\Big(\sum_{k=n}^{r} a_{n,k}\Big)^p\Big)^{1/p} \end{aligned}$$

by Hölder's inequality, and this is equivalent to (30).

The last sentence of the theorem follows from Proposition 1 and Theorem 2 of [1].

Theorem 2 shows that the infimum in (3) occurs at $r = 0$, but the proof says nothing of the monotonicity properties of the function (4). It can be shown, for quasi-Hausdorff matrices, that the function $f(r)$ does indeed increase with r. The proof is similar to that of Theorem 1—using majorization rather than submajorization—and will be omitted. (See also Theorem 6 in section 7.)

6. **Weighted means.** In this section we study the lower bound problem for *weighted means*. These are matrices with entries of the form

$$a_{n,k} = \begin{cases} a_k/A_n & (1 \leq k \leq n) \\ 0 & (k > n) \end{cases}, \tag{31}$$

where the a_k's are non-negative numbers, and $A_n = a_1 + \cdots + a_n$. (We take a_1 to be positive so that none of the A_n's vanishes. Note, also, that we now index the matrix entries from 1 instead of from 0.)

Weighted means arise naturally in summability theory, [8], [25], and have been studied extensively from this point of view. Moreover, their ℓ^p-mapping properties are described completely in [2].

Their simple structure makes them natural objects of study in the present context too. In particular, the partial sums appearing in (4), which proved so troublesome for Hausdorff matrices, are readily computable, and (3) takes on the following very simple form

$$L^p = 1 + \inf_{n \geq 1} A_n^p \sum_{k>n} A_k^{-p}. \tag{32}$$

Unfortunately, the infimum need not be attained when $n = 1$, and its evaluation, for general a_n's, appears to be intractable. Indeed, to make any progress at all, we find it necessary to impose a rather stringent monotonicity condition on the a_n's. Our proof involves the following elementary result, which may be of some interest in its own right.

LEMMA 7. *Let $\sum_{n=1}^{\infty} x_n$ be a convergent series of positive terms. If*

$$n\left(\frac{x_n - x_{n+1}}{x_{n+1}}\right) \text{ decreases with } n, \tag{33}$$

then

$$\frac{1}{nx_n} \sum_{k>n} x_k \text{ increases with } n. \tag{34}$$

PROOF. We first show that, if (33) holds, then

$$nx_n \longrightarrow 0 \text{ as } n \longrightarrow \infty. \tag{35}$$

To see this, we observe that

$$\frac{nx_n - (n+1)x_{n+1}}{x_{n+1}} = n\left(\frac{x_n - x_{n+1}}{x_{n+1}}\right) - 1$$

https://doi.org/10.4153/CJM-1992-003-9 Published online by Cambridge University Press

decreases with n. Consequently, if $nx_n < (n+1)x_{n+1}$ for some integer n, the same inequality persists for all larger values of n. But then $x_j > nx_n/j$ for all $j > n$, and this entails the divergence of $\sum x_j$. We conclude that nx_n decreases with n, and hence converges. The limit must be 0, else $\sum x_j$ diverges.

Consider now three positive integers, n, j, m, with $n \le j \le m$. We may rewrite (33) in the form

$$nx_nx_{j+1} \ge x_{n+1}\big(jx_j - (j+1)x_{j+1}\big) + (n+1)x_{n+1}x_{j+1}.$$

Summing from $j = n+1$ to m gives

$$nx_n \sum_{j=n+1}^{m} x_{j+1} \ge x_{n+1}\big((n+1)x_{n+1} - (m+1)x_{m+1}\big) + (n+1)x_{n+1} \sum_{j=n+1}^{m} x_{j+1}.$$

Letting $m \longrightarrow \infty$, and using (35), we see that

$$nx_n \sum_{j>n+1} x_j \ge (n+1)x_{n+1}^2 + (n+1)x_{n+1} \sum_{j>n+1} x_j,$$

and this is equivalent to (34).

REMARK. There is a companion result to Lemma 7 which we shall need later: "If the sequence (33) increases with n, then the sequence (34) decreases with n." The proof is similar to that of Lemma 7, but slightly easier, since we do not have to check that (35) holds.

There is another companion result, in which the summation "$k > n$" in (34) is replaced by "$k \ge n$". We leave the details to the reader.

THEOREM 3. *Fix p, $1 \le p < \infty$, and let A be a weighted mean matrix with entries given by (31). Suppose that*

(36) $$\sum_{n=1}^{\infty} A_n^{-p} < \infty,$$

and that

(37) $$n\left(\frac{A_{n+1}^p - A_n^p}{A_n^p}\right) \text{ decreases with } n.$$

Then

(38) $$\|A\mathbf{x}\|_p \ge L\|\mathbf{x}\|_p$$

for every $\mathbf{x} \in \ell^p$ satisfying $x_1 \ge x_2 \ge \cdots \ge 0$, where

$$L^p = a_1^p \sum_{n=1}^{\infty} A_n^{-p}.$$

There is equality in (38) only when $x_2 = x_3 = \cdots = 0$.

PROOF. We apply Lemma 7 with $x_n = A_n^{-p}$ to see that

$$\frac{A_n^p}{n} \sum_{k>n} A_k^{-p}$$

increases with n. Inequality (38) then follows from (32).

The statement about cases of equality in (38) is a consequence of Theorem 2 of [1].

Theorem 3 is a slight improvement of a result of Rhoades ([20], Theorem 1), but the proof given here is much simpler than his. Rhoades' version, based on the proof of (5) given in the original draft of [1], requires the additional assumptions "$p > 1$" and "a_n increasing in n".

The hypothesis (37), first formulated by Rhoades, is a rather curious one, and it deserves to be studied in some detail. Accordingly, we shall say that a sequence $\mathbf{x} = (x_n)_{n=1}^{\infty}$, of positive terms, satisfies the *Rhoades condition*, $R(p)$, and we shall write $\mathbf{x} \in R(p)$, if

(39) $$n\left(\frac{x_n^p - x_{n+1}^p}{x_{n+1}^p}\right) \text{ decreases with } n.$$

Now (39) may be written in a more suggestive form

(40) $$\left\{\frac{n}{n+1}(x_n x_{n+2})^p + \frac{1}{n+1}(x_{n+1}x_{n+2})^p\right\}^{1/p} \geq x_{n+1}^2.$$

We recognize the left-hand side as an L^p-norm on a two-point probability space, and it follows at once that $R(p) \subseteq R(q)$ if $p \leq q$. Thus the $R(p)$ conditions become less restrictive as p increases. The least stringent of all, say $R(\infty)$, is defined by taking the limiting condition, as $p \longrightarrow \infty$, in (40). If $\mathbf{x}$ is increasing, this condition is automatically satisfied, while if $\mathbf{x}$ is decreasing, the $R(\infty)$-condition is

(41) $$x_n x_{n+2} \geq x_{n+1}^2,$$

alias *logarithmic convexity*. In a similar fashion, we define the condition $R(-\infty)$, the strictest of all the $R(p)$'s, by taking the limit, as $p \longrightarrow -\infty$, in (40). If $\mathbf{x}$ is increasing, the $R(-\infty)$-condition is (41), while if $\mathbf{x}$ is decreasing, the condition is that x_n be constant for $n > 1$.

Thus we have a continuum of conditions, which, for increasing sequences, are all weaker than logarithmic convexity, and, for decreasing sequences, all stronger.

Rhoades applies his result to the *power means*, namely to the matrices (31) with

(42) $$a_n = n^{\alpha}$$

He notes that the sequence (42) belongs to $R(p)$, for $p > 1$, when $\alpha = 1, 2$ or 3 (though, of course, he does not use this terminology). Thus, for these values of α, he concludes that the lower bound is attained in the first column of A. He conjectures ([20], page 351) that the same result persists for all positive α. (The case $\alpha = 0$, of course, is just inequality (5)).

We show that the conjectured result is false for $0 < \alpha < 1$, true for $\alpha \geq 1$, and true even for $-1/p^* < \alpha \leq 0$ (Theorem 4). The restriction, $\alpha > -1/p^*$, is needed here, in order to get meaningful results, for it is equivalent to the first column of A having finite ℓ^p-norm (compare (36)). The same restriction, incidentally, guarantees that A maps ℓ^p into itself.

LEMMA 8. *Let α be a fixed real number. Then the sequence*

$$\frac{n(n+1)^\alpha}{1^\alpha+\cdots+n^\alpha} \quad (n=1,2,\ldots) \tag{43}$$

increases with n if $0<\alpha<1$, *and decreases otherwise.*

PROOF. The sequence (43) decreases with n precisely when

$$(n+1)(n+2)^\alpha(1^\alpha+\cdots+n^\alpha) \le n(n+1)^\alpha\big(1^\alpha+\cdots+(n+1)^\alpha\big), \tag{44}$$

for $n=1,2,\ldots$, and increases when (44) is reversed.

Let $\mathbf{x}$ be the $(n+1)n$-tuple formed by repeating $(n+1)$ times each term of the n-tuple: $\big(1(n+2),2(n+2),\ldots,n(n+2)\big)$. Similarly, let $\mathbf{y}$ be the $n(n+1)$-tuple formed by repeating n times each term of the $(n+1)$-tuple: $\big(1(n+1),2(n+1),\ldots,(n+1)(n+1)\big)$. A routine (but rather tedious) calculation shows that $\mathbf{x}$ is majorized by $\mathbf{y}$, so that (12) holds for all continuous convex functions $\phi:(0,\infty)\longrightarrow\mathbb{R}$. Taking $\phi(t)=t^\alpha$, in case $\alpha\ge 1$, or $\alpha\le 0$, shows that inequality (44) holds for these values of α. On the other hand, taking $\phi(t)=-t^\alpha$, shows that (44) is reversed when $0<\alpha<1$. This completes the proof of the lemma.

THEOREM 4. *Let α, p be fixed real numbers with $p\ge 1$ and $\alpha>-1/p^*$, and let A be the power mean with entries given by (42). Then A, acting on ℓ^p, attains its lower bound in its first column provided that $\alpha\le 0$ or $\alpha\ge 1$. The result fails if $0<\alpha<1$.*

PROOF. With $a_n=n^\alpha$ ($\alpha\le 0$ or $\alpha\ge 1$), we see that the sequence na_{n+1}/A_n decreases with n by Lemma 8. Therefore, the sequence A_n^{-1} belongs to $R(1)$, and hence to $R(p)$. The result now follows from Theorem 3.

The situation is a good deal more complicated when $0<\alpha<1$. Certainly, the theorem fails to hold, for all α's in this range, when $p=1$. To see this, we note that the sequence A_n^{-1} no longer belonges to $R(1)$; indeed, by Lemma 8, the defining inequality for $R(1)$ is *reversed.* It follows, from the remark after Lemma 7, that the infimum in (32) occurs when $r=\infty$, so that the lower bound for A is never attained.

On the other hand, it can be shown that the sequence A_n^{-1} is logarithmically convex, so that A_n^{-1} belongs to $R(\infty)$. This suggests, for each fixed α, $0<\alpha<1$, that the theorem might be true for suitably large p (depending on α). The theorem is certainly false for suitably small p (not just for $p=1$). It would be interesting to determine the exact "breakdown" point. At issue here is the following elementary

PROBLEM. Let α and p be real numbers with $p>1$ and $\alpha>-1/p^*$. Determine completely the monotonicity properties of the sequence

$$\frac{(1^\alpha+\cdots+n^\alpha)^p}{n}\sum_{k>n}(1^\alpha+\cdots+n^\alpha)^{-p}$$

(The outstanding case is $0<\alpha<1<p$.)

https://doi.org/10.4153/CJM-1992-003-9 Published online by Cambridge University Press

7. **Moment sequences.** In this section we describe some elementary inequalities satisfied by moment sequences. We recall that these are sequences having a representation of the form (10).

Our first inequality is merely a summary of the results of Sections 3 and 4.

THEOREM 5. *Let $\phi:[0,1]\longrightarrow\mathbb{R}$ be a non-negative, convex function, with $\phi(0)=0$, and let $\boldsymbol{\mu}$ be a moment sequence. Then*

$$\frac{1}{r+1}\sum_{n=0}^{\infty}\phi\Big(\sum_{k=0}^{r}\binom{n}{k}\Delta^{n-k}\mu_k\Big)$$

is an increasing function of r ($r=0,1,\ldots$).

Our next result is the analogue of Theorem 5 for quasi-Hausdorff matrices. The proof is omitted.

THEOREM 6. *Let $\phi:[0,\infty)\longrightarrow\mathbb{R}$ be a convex function and let $\boldsymbol{\mu}$ be a moment sequence. Then*

$$\frac{1}{r+1}\sum_{n=0}^{r}\phi\Big(\sum_{k=n}^{r}\binom{k}{n}\Delta^{k-n}\mu_n\Big)$$

is an increasing function of r ($r=0,1,\ldots$).

Our next three theorems are motivated by the papers [18], [19] and [21] of Rhoades. They are similar in spirit to the other results of this paper, but they do not have a direct bearing on the lower bound problem.

Rhoades ([21], Problem 1) asks whether the square root of every totally monotone sequence is again totally monotone (the square root being taken coordinatewise). The answer is no. We give here an indirect proof because our method, which provides rather more information than is needed, enables us to answer another question of Rhoades ([21], Problem 2).

Suppose, then, that the square root of every totally monontone sequence is totally monotone. If $\boldsymbol{\mu}$ is such a sequence, we must have

$$\Delta^n(\mu_k^p)\geq 0\quad (n,k=0,1,\ldots), \tag{45}$$

whenever p is of the form, $p=2^{-r}$ for some positive integer r. Now (45) may be rewritten in a more suggestive way

$$\Big(2^{1-n}\sum_{j\text{ even}}\binom{n}{j}\mu_{k+j}^p\Big)^{1/p}\geq\Big(2^{1-n}\sum_{j\text{ odd}}\binom{n}{j}\mu_{k+j}^p\Big)^{1/p}, \tag{46}$$

which we recognize as a comparison of L^p-norms on two probability spaces. Letting $p\longrightarrow 0$ (through the values $1/2, 1/4,\ldots$), and applying Theorem 3 of [9], gives

$$\prod_{j\text{ even}}\mu_{k+1}^{\binom{n}{j}}\geq\prod_{j\text{ odd}}\mu_{k+1}^{\binom{n}{j}}\quad (n,k=0,1,\ldots). \tag{47}$$

(The empty product, appearing on the right of (47), when $n = 0$, is to be interpreted as "zero". This is in accordance with the derivation of (47) as a limit of (46).) We call a sequence $\boldsymbol{\mu}$ satisfying (47) *logarithmically totally monotone.*

Now the first three inequalities ($n = 0, 1$ and 2) implied by (47) are satisfied by any totally monotone sequence. (For $n = 2$ see [23] and [19]). But the fourth inequality,

$$\mu_k \mu_{k+2}^3 \geq \mu_{k+1}^3 \mu_{k+3} \quad (k = 0, 1, \ldots), \tag{48}$$

is not, and this fact enables us to solve Rhoades' square root problem.

THEOREM 7. *There exists a totally monotone sequence that is not logarithmically totally monotone.*

PROOF. We give an example of a moment sequence that fails to satisfy (48). The example arises from a two-point probability space, with measure determined by

$$\int_0^1 f(\theta)\, d\mu(\theta) = af(x) + bf(y)$$

for f continuous on $0 \leq \theta \leq 1$. Here $b = 1 - a$, and the remaining numbers, a, x, y, will be determined later. Inequality (48) is

$$(a + b)(ax^2 + by^2)^3 - (ax + by)^3(ax^3 + by^3) \geq 0,$$

and this may be rewritten as

$$ab(x - y)^3(a^2x^3 - b^2y^3) \geq 0. \tag{49}$$

If $x = 1/3, y = 2/3, a = 3/4$, and $b = 1/4$, it is easily checked that (49) is violated. Thus $\boldsymbol{\mu}$, the moment sequence associated with $d\mu(\theta)$, fails to be logarithmically totally monotone. The corresponding "moments" are $\mu_0 = 1$, $\mu_1 = 5/12$, $\mu_2 = 7/36$, $\mu_3 = 11/108, \ldots$.

From the discussion preceding Theorem 7 we deduce the following

COROLLARY. *There exists a totally monotone sequence whose square root is not totally monotone.*

Our next result provides a natural complement to Theorem 7. We shall need the well-known *Leibnitz formula* for differences

$$\Delta^n(\mu_k \nu_k) = \sum_{j=0}^{n} \binom{n}{j} (\Delta^{n-j}\mu_{k+j})(\Delta^j \nu_k) \quad (n, k = 0, 1, \ldots). \tag{50}$$

From (50) it follows that the set of totally monotone sequences is closed under products. Being closed also under sums it is closed under *exponentiation*. In other words, if μ_k is totally monotone, then so is e^{μ_k}.

https://doi.org/10.4153/CJM-1992-003-9 Published online by Cambridge University Press

THEOREM 8. *A logarithmically totally monotone sequence, with positive terms, is totally monotone.*

PROOF. Suppose that $\boldsymbol{\mu}$ satisfies (47) and that $\mu_k > 0$ for all k. Taking logarithms in (47) gives

$$\Delta^n(\log \mu_k) \geq 0 \quad (n \geq 1, k \geq 0). \tag{51}$$

It follows that the sequence $\log(\mu_k/\mu_{k+1}) = \Delta(\log \mu_k)$ is totally monotone, and hence, by exponentiation, so is the sequence μ_k/μ_{k+1}. From (51), with $n = 1$, we see that $\mu_k/\mu_{k+1} \geq 1$, so that the sequence $(\mu_k/\mu_{k+1}) - 1 = \nu_k$, say, is totally monotone. Writing $\Delta\mu_k = \nu_k\mu_{k+1}$, and applying (50) inductively, we see that μ_k is totally monotone.

REMARK. Credit for Theorem 8 is due to Rhoades. He does not prove the theorem, nor does he even state it, but the ideas used in the proof are discussed in his paper [18].

Given a sequence $\boldsymbol{\mu}$, we denote by $\boldsymbol{\mu}^\alpha$ the sequence whose kth term is μ_k^α.

COROLLARY. *Let $\boldsymbol{\mu}$ be a totally monotone sequence. Then the following conditions are equivalent:*

(i) $\boldsymbol{\mu}^\alpha$ is totally monotone for all $\alpha > 0$,

(ii) $\boldsymbol{\mu}$ is logarithmically totally monotone.

PROOF. (i)$\Rightarrow$(ii) is just a restatement of the remarks made after Theorem 6.

(ii)$\Rightarrow$(i). If $\boldsymbol{\mu}$ is totally monotone, then either μ has the form $(c, 0, 0, \ldots)$, or $\mu_k > 0$ for all k (according as whether $\mu(0+) = \mu(1)$ or not). In the first case the desired implication is obvious; in the second we use Theorem 8—after noting that definition (47) holds for $\boldsymbol{\mu}^\alpha$ if it holds for $\boldsymbol{\mu}$.

The corollary contradicts Corollary 1 of [21]. The hypothesis, "$\boldsymbol{\mu} \in Q$", therein, should be replaced by "$\boldsymbol{\mu}$ is logarithmically totally monotone". (See also the corollary to Theorem 9.)

Our next theorem deals with products of Hausdorff matrices and enables us to solve another problem of Rhoades ([21], Problem 2). We recall ([8], Theorem 197) that the product, $H_\mu H_\nu$, of two Hausdorff matrices is again a Hausdorff matrix, and that the associated moment sequence is given by

$$\lambda_k = \mu_k\nu_k. \tag{52}$$

We shall need the following result, which, in conjunction with Lemma 4, shows that Hausdorff matrices "transform decreasing sequences into decreasing sequences".

LEMMA 9. *Let $A = (a_{n,k})_{n,k=1}^\infty$ be a matrix with non-negative entries, and consider the associated transformation, $\mathbf{x} \longrightarrow \mathbf{y}$, given by $y_n = \sum_{k=1}^\infty a_{n,k}x_k$. Then the following conditions are equivalent.*

(i) $y_1 \geq y_2 \geq \cdots \geq 0$ whenever $x_1 \geq x_2 \geq \cdots \geq 0$,

(ii) $\sum_{k=1}^r a_{n,k} \geq \sum_{k=1}^r a_{n+1,k}$ $(n, r = 1, 2 \ldots)$.

PROOF. (i)$\Rightarrow$(ii) follows by taking $\mathbf{x}$ to be the sequence $(1, \ldots, 1, 0, 0, \ldots)$ of r "ones" followed by "zeros".

(ii)$\Rightarrow$(i) follows by summation by parts.

THEOREM 9. *Fix $p \geq 1$, and let H_λ, H_μ, H_ν be Hausdorff matrices related by $H_\lambda = H_\mu H_\nu$. Then H_λ is bounded on ℓ^p if and only if both H_μ and H_ν are. Moreover, we have*

$$\|H_\lambda\|_{p,p} = \|H_\mu\|_{p,p}\|H_\nu\|_{p,p}. \tag{53}$$

PROOF. If H_μ and H_ν are bounded on ℓ^p, then so, clearly, is H_λ, and we have

$$\|H_\lambda\|_{p,p} \leq \|H_\mu\|_{p,p}\|H_\nu\|_{p,p}. \tag{54}$$

On the other hand, suppose that H_λ is bounded on ℓ^p. If $\mathbf{x}$ is a decreasing sequence with $\|\mathbf{x}\|_p = 1$, we have, by Lemma 9,

$$\begin{aligned}\|H_\lambda\|_{p,p} &\geq \|H_\lambda \mathbf{x}\|_p\\ &= \|H_\mu H_\nu \mathbf{x}\|_p\\ &\geq \text{(lower bound of } H_\mu) \cdot \|H_\nu \mathbf{x}\|_p\\ &\geq \|H_\nu \mathbf{x}\|_p.\end{aligned}$$

Taking the supremum over all $\mathbf{x}$, and applying (22), gives $\|H_\nu\|_{p,p} \leq \|H_\lambda\|_{p,p}$. Since Hausdorff matrices commute ([8], Theorem 197), we also have $\|H_\mu\|_{p,p} \leq \|H_\lambda\|_{p,p}$. This completes the proof of the first part of the theorem.

To prove (53) we follow an argument of Hardy ([7], page 48), with some slight modifications. We take $0 < \epsilon < \frac{1}{p}$, $x_n = (n+1)^{-\epsilon-1/p}$, and any positive δ, $0 < \delta < 1$. We choose α, N and ϵ, in that order, to satisfy

$$(1 + 1/\alpha)^{-2/p} > \delta$$

$$\int_{\sqrt{\alpha/N}}^1 \int_{\sqrt{\alpha/N}}^1 (\theta\phi)^{-1/p}\, d\mu(\theta)\, d\nu(\phi) > \delta \int_0^1\int_0^1 (\theta\phi)^{-1/p}\, d\mu(\theta)\, d\nu(\phi)$$

and

$$\sum_{n=N}^{\infty} x_n^p > \delta \sum_{n=0}^{\infty} x_n^p.$$

As in Hardy's proof, we have

$$\sum_{m=0}^{n} \binom{n}{m} (\theta\phi)^m (1-\theta\phi)^{n-m} x_m > \delta(\theta\phi)^{-1/p} x_n$$

whenever $\alpha/n < \theta\phi < 1$.

It follows from (52) that

$$(H_\lambda \mathbf{x})_n = \int_0^1\int_0^1 \sum_{m=0}^{n} \binom{n}{m} (\theta\phi)^m (1-\theta\phi)^{n-m} x_m\, d\mu(\theta)\, d\nu(\phi).$$

If $n \geq N$, we have

$$\begin{aligned}(H_\lambda \mathbf{x})_n &\geq \delta x_n \int_{\sqrt{\alpha/N}}^1 \int_{\sqrt{\alpha/N}}^1 (\theta\phi)^{-1/p}\, d\mu(\theta)\, d\nu(\phi)\\ &\geq \delta^2 x_n \int_0^1\int_0^1 (\theta\phi)^{-1/p}\, d\mu(\theta)\, d\nu(\phi),\end{aligned}$$

and it follows from (23) that

$$\|H_\lambda \mathbf{x}\|_p \geq \delta^{2+1/p} \|H_\mu\|_{p,p} \|H_\nu\|_{p,p} \|\mathbf{x}\|_p.$$

This gives

$$\|H_\lambda\|_{p,p} \geq \delta^{2+1/p} \|H_\mu\|_{p,p} \|H_\nu\|_{p,p},$$

and letting $\delta \longrightarrow 1$, we obtain (53).

COROLLARY. *Let p be fixed, $1 \leq p < \infty$, and let H_μ be a Hausdorff matrix bounded on ℓ^p. If $\boldsymbol{\mu}$ is logarithmically totally monotone then H_{μ^α} is bounded on ℓ^p, for each $\alpha > 0$, and*

$$\|H_{\mu^\alpha}\|_{p,p} = \|H_\mu\|_{p,p}^\alpha. \tag{55}$$

PROOF. If α is rational, say $\alpha = m/n$ with m and n positive integers, then

$$(H_{\boldsymbol{\mu}^\alpha})^n = H_{\mu^m} = (H_\mu)^m,$$

and (55) follows from Theorem 9. A simple continuity argument shows that equality in (55) persists even if α is irrational.

We remark that the corollary answers affirmatively a question raised by Rhoades ([22], page 296). Rhoades obtains (55), but only for $p = 2$ and for α of the form 2^{-k} ($k = 1, 2, \ldots$). (See also the remark following the corollary to Theorem 8 above).

We now return to our study of lower bounds. It turns out that there is an analogue, Lemma 10, of the familiar "submultiplicative" property, (54), of matrix norms. This analogue, when applied to products of Hausdorff matrices, leads to some intriguing inequalities for moment sequences.

LEMMA 10. *Fix p, $1 \leq p < \infty$. Let A and B be matrices, bounded on ℓ^p, and having non-negative entries. Suppose, further, that B transforms decreasing sequences into decreasing sequences. Then*

$$L(AB) \geq L(A)L(B).$$

PROOF. *This is an immediate consequence of the definition, (2), of lower bound.*

THEOREM 10. *Let $\boldsymbol{\mu}$, $\boldsymbol{\nu}$ be moment sequences. Then*

$$\sum_{m=0}^{\infty} (\Delta^m \mu_0)^p \sum_{n=0}^{\infty} (\Delta^n \nu_0)^p \leq \sum_{n=0}^{\infty} \big(\Delta^n (\mu\nu)_0\big)^p,$$

whenever $p \geq 1$.

PROOF. Consider the Hausdorff matrices, H_μ and H_ν, associated with the given moment sequences. It follows from Lemma 4 that these matrices satisfy condition (ii) of Lemma 9, and hence transform decreasing sequences into decreasing sequences. Moreover, the product matrix, $H_\mu H_\nu$, is the one generated by the product sequence, namely $H_{(\mu\nu)}$. We now apply Lemma 10, Theorem 1, and the representation (7).

Theorem 10 has several amusing consequences. We mention here one of the simplest of these, obtained by taking $\mu_n = \nu_n = 1/(n+1)$ (which is the moment sequence corresponding to Lebesgue measure).

COROLLARY. *$\zeta^2(p) \le \sum_{n=1}^{\infty}\left(\frac{H_n}{n}\right)^p$ $(p > 1)$, where $H_n = 1 + (1/2) + \cdots + (1/n)$ denotes the nth harmonic number.*

Based on our experience with the earlier results of this paper, it is natural to ask whether Theorem 10 can be extended to convex functions. That this is indeed the case is given by

THEOREM 11. *Let $\boldsymbol{\mu}$, $\boldsymbol{\nu}$ be moment sequences and let $\phi\colon[0,1] \longrightarrow \mathbb{R}$ be a non-negative convex function with $\phi(0) = 0$. Then*

$$\sum_{m,n=0}^{\infty} \phi(\Delta^m\mu_0 \cdot \Delta^n\nu_0) \le \sum_{n=0}^{\infty} \phi\left(\Delta^n(\mu\nu)_0\right).$$

Theorem 11, of course, depends upon a majorization inequality, and this is given below in its most basic form. The majorization, however, is different than any we have encountered in this paper. The sequence to be majorized (doubly-infinite, in this case) is not in descending order, and its decreasing rearrangement is not at all accessible. Theorem 12, then, requires a more sophisticated approach, and this will have to be described elsewhere (see [4]).

THEOREM 12. *Let x and y be fixed real numbers with $0 \le x,\ y \le 1$, and let N be any positive integer. Then the sum of any N terms from the set $\{x^m y^n : m, n = 0, 1, 2 \ldots\}$ does not exceed $1 + (x + y - xy) + \cdots + (x + y - xy)^{N-1}$.*

APPENDIX BY ANDREW LENARD†. In his studies on lower bounds of matrix operators Bennett proved a monotonicity property for the sequence

(1) $$S_r = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{r}\binom{n}{k}\int_0^1 \theta^k(1-\theta)^{n-k}\,d\mu\right)^p \quad (r = 0, 1, 2, \ldots)$$

namely

(2) $$S_0 \le S_1/2 \le S_2/3 \le \cdots$$

Here p is an arbitrary real number no less than 1, and μ is an arbitrary probability measure on the interval [0,1]. This surprising fact was at first conjectured from the much weaker statement

(3) $$S_0 \le S_r/(r+1) \quad (r = 1, 2, \ldots)$$

shown by the present writer for the particular cases when p is an integer and the measure μ has a single point of support. Let $0 \le x < 1$, and let μ be supported at $\theta = 1 - x$, so that (3) reads

(4) $$S_r \ge (r+1)(1-x^p)^{-1} \quad (r = 1, 2, \ldots)$$

† Department of Mathematics, Indiana University, Bloomington, Indiana, U. S. A.

where

$$S_r = \sum_{n=0}^{\infty}\Big(\sum_{k=0}^{r}\binom{n}{k}x^{n-k}(1-x)^k\Big)^p \tag{5}$$

In the application to lower bound problems the weaker inequality (3) rather than (2) is of importance. Furthermore, the method by which (4) was obtained is quite different from the one later designed by Bennett to prove his generalization. Thus it may be of some interest to have the proof of (4) on record, and this is the purpose of the present note.

Actually, a natural generalization, but in a different direction, is needed for our proof. Instead of one parameter x and the pth power of the truncated binomial sum we use p separate parameters and take the product of the corresponding sums. Thus we let $\mathbf{x} = (x_1, x_2, \ldots, x_p) \in [0,1)^p$ and

$$S_r = \sum_{n=0}^{\infty}\prod_{s=1}^{p}\sum_{k_s=0}^{r}\binom{n}{k_s}x_s^{n-k_s}(1-x_s)^{k_s}; \tag{6}$$

and we shall prove

$$S_r \geq (r+1)(1-x_1x_2\cdots x_p)^{-1} \quad (r=1,2,\ldots). \tag{7}$$

Let $t_1, t_2, \ldots, t_p$ be auxiliary variables, and consider the function

$$F(t_1,t_2,\ldots,t_p) = \Big(1 - \prod_{s=1}^{p}(x_s + t_s - x_st_s)\Big)^{-1} \tag{8}$$

If $c_{k_1k_x\cdots k_p}$ is the coefficient of $t^{k_1}t^{k_2}\cdots t^{k_p}$ in its multi-power series expansion, we find that

$$S_r = \sum_{k_1,k_2,\ldots,k_p\leq r} c_{k_1k_2\cdots k_p} \tag{9}$$

It is convenient to use the following abbreviations: $P = \{1,2,\ldots,p\}$ and, for $N \subseteq P$, $x_N = \prod_{s\subseteq N} x_s$, $(1-x)_N = \prod_{s\subseteq N}(1-x)_s$ and $t_N = \prod_{s\subseteq N} t_s$. Then

$$\prod_{s=1}^{p}(x_s + t_s - x_st_s) = \sum_{N\subseteq P} x_{P-N}(1-x)_Nt_N \tag{10}$$

When all the t's vanish we obtain from this the identity

$$\sum_{N\subseteq P} x_{P-N}(1-x)_N = 1. \tag{11}$$

We rewrite this in the form

$$\sum{}' \xi_N = 1, \tag{12}$$

where generally $\sum'$ denotes a sum over all *non-empty* subsets of P, and $\xi_N = x_{P-N}(1-x)_N(1-x_P)^{-1}$ (note the different use of the subscript N here). From (8) and (10) it follows then that

$$F(t_1,\ldots,t_p) = (1-x_P)^{-1}(1-\sum{}'\xi_Nt_N)^{-1} = (1-x_P)^{-1}\sum_{n=0}^{\infty}(\sum{}'\xi_Nt_N)^n \tag{13}$$

The sum $\sum'$ contains $q = 2^p - 1$ terms. We expand its nth power by the multinomial theorem. This yields

(14) $$F(t_1, \ldots, t_p) = (1 - x_p)^{-1} \sum_j (\textstyle\sum' j_N)! \prod' (\xi_N^{j_N} t^{j_N} / j_N!)$$

where $\prod'$ denotes a product over non-empty subsets N of P, and where $j = (j_{\{1\}}, j_{\{2\}}, \ldots, j_{\{1,2\}}, \ldots, j_P)$ is a multi-index with q components. Since

(15) $$\prod' t_N^{j_N} = \prod_{s=1}^{p} t^{(\sum_{s \subseteq N} j_N)}$$

one finds that the coefficient $c_{k_1 \cdots k_p}$ in the power series expansion of F can be written as the sum of all those terms

(16) $$(1 - x_P)^{-1} (\textstyle\sum' j_N)! \prod' (\xi_N^{j_N} / j_N!)$$

for which the conditions

(17) $$\sum_{1 \in N} j_N = k_1, \sum_{2 \in N} j_N = k_2, \ldots, \sum_{p \in N} j_p = k_p$$

hold. But then (9) shows that S_r is the sum of all those terms (16) for which the conditions

(18) $$\sum_{1 \in N} j_N, \sum_{2 \in N} j_N, \ldots, \sum_{p \in N} j_N \leq r$$

hold. Let us observe now that the single condition

(19) $$\textstyle\sum' j_N \leq r$$

implies all of the p conditions (18). Therefore if one sums the terms (16) under the sole restriction (19) one obtains a lower bound for S_r. But this latter sum can be easily carried out. It is

(20) $$(1 - x_P)^{-1} \sum_{i=0}^{r} i! \sum_{\sum' j_N = i} \prod' (\xi_N^{j_N} / j_N!) = (1 - x_P)^{-1} \sum_{i=0}^{r} (\textstyle\sum' \xi_N)^i$$

In view of (12) this is precisely the right hand side of (7), as required.

Knowing the validity of (4) for all positive integral values of p makes it, of course, tempting to speculate that it also holds for arbitrary real $p \geq 1$. Bennett's alternate proof, based on quite different ideas, not only shows that this is the case, but extends the conclusion to a vastly larger class of sequences. Still, if some usable argument could be found, to pass from integral to arbitrary real values of p, it would likely provide new insight.

The writer wishes to acknowledge the fortuitous but propitious circumstance that brought this problem to his attention, and his gratitude to Grahame Bennett for patiently explaining its origin and significance.

REFERENCES

1. G. Bennett, *Lower bounds for matrices*, Linear Algebra and Appl. **82**(1986), 81–98.
2. ———, *Some elementary inequalities*, Quart. Jour. Math. Oxford (2) **38**(1987), 401–425
3. ———, *Some elementary inequalities, II*, Quart. Jour. Math. Oxford (2) **39**(1988), 385–400.
4. ———, *Coin tossing and moment sequences*, Discrete Math., **84**(1990), 111–118.
5. H. L. Garabedian, E. Hille and H. S. Wall, *Formulations of the Hausdorff inclusion problem* Trans. Amer. Math. Soc. **8**(1941), 193–213.
6. P. Ghatage, *On the spectrum of the Bergman Hilbert matrix*, Linear Algebra and App. **97**(1987), 57–63.
7. G. H. Hardy, *An inequality for Hausdorff means*, Jour. London Math. Soc. **18**(1943), 46–50.
8. ———, *Divergent series*, Oxford University Press, 1949.
9. G. H. Hardy, J. E. Littlewood and G. Polya, *Inequalities*, 2nd edition, Cambridge University Press, 1967.
10. F. Hausdorff, *Summationsmethoden und Momentfolgen, I*, Math. Z. **9**(1921), 74–109.
11. K. Knopp, *Uber Reihen mit positiven Gliedern* (Zweite Mitteilung), Jour. London Math. Soc. **5**(1930), 13–21.
12. A. Lenard, Personal communication.
13. R. Lyons, *A lower bound on the Cesàro operator*, Proc. Amer. Math. Soc. **86**(1982), 694.
14. A. W. Marshall and W. Olkin, *Inequalities: Theory of Majorization and its Applications*, Academic Press, New York, 1979.
15. G. Polya, *Remark on Weyl's note "Inequalities between the two kinds of eigenvalues of a linear transformation"*, Proc. Nat. Acad. Sci. U.S.A. **36**(1950), 49–51.
16. J. M. Rassias, *On the generalized Cesàro operators*, Math. analysis, 32–34, Teubner Texte Math. **79**, Teubner, Leipzig, (1985).
17. P. F. Renaud, *A reversed Hardy inequality*, Bull. Australian Math. Soc. **34**(1986), 225–232.
18. B. E. Rhoades, *A sufficient condition for total monotonicity*, Trans. Amer. Math. Soc. **107**(1963), 309–319.
19. ———, *A sufficient condition for total monotonicity, II*, Jour. Indian Math. Soc. **41**(1977), 221–232.
20. ———, *Lower bounds for some matrices*, Linear and Multilinear Algebra **20**(1987), 347–352.
21. ———, *Square roots for Hausdorff operators*, Integral Equations and Operator Theory **11**(1988), 292–296.
22. M. Tomic, *Théorème de Gauss relatif au centre de gravité et son application*, Bull. Soc. Math. Phys. Serbie **1**(1949), 31–40.
23. H. S. Wall, *Continued fractions and totally monotone sequences*, Trans. Amer. Math. Soc. **48**(1940), 165–184.
24. D. V. Widder, *The Laplace Transform*, Princeton University Press, Princeton, N.J. 1946.
25. K. Zeller and W. Beekman, *Theorie der Limitierungsverfahren*, Ergebnisse der Math., **15**, Springer-Verlag, Berlin, 1970.

Department of Mathematics
Indiana University
Bloomington, Indiana 47405
U.S.A.