Mathematical
Inequalities
& Applications
Volume 14, Number 4 (2011), 747–775

MERCER'S INEQUALITY AND TOTALLY MONOTONIC SEQUENCES

GRAHAME BENNETT

This paper is dedicated
to Professor B.E. Rhoades,
my friend and colleague of forty years

(*Communicated by J. Pečarić*)

Abstract. A simple inequality of Mercer leads to some intriguing problems concerning moment sequences.

1. Introduction

Mercer [15] has shown that the following elegant inequality,

$$\frac{1}{2^n}\sum_{k=0}^{n}\binom{n}{k}x_k \leqslant \frac{1}{n+1}\sum_{k=0}^{n}x_k \qquad (n=0,1,2,\dots), \tag{1}$$

is valid for all convex sequences $\mathbf{x}=(x_0,x_1,x_2,\dots)$.

A remarkable feature of his result is that it leads quite naturally to delicate problems in the Theory of Moments. This becomes apparent when we attempt to extend (1) to arbitrary matrices:

$$\sum_k a_{n,k}x_k \leqslant \sum_k b_{n,k}x_k. \tag{2}$$

In order for the sums in (2) to be defined it is necessary to assume that the rows of A and B have only finitely many non-zero terms. Moreover, the row sums of A must match those of B,

$$\sum_k a_{n,k} = \sum_k b_{n,k} \qquad (n=0,1,2,\dots), \tag{3}$$

as is seen by taking $\mathbf{x}=\pm(1,1,1,\dots)$ in (2).

It is here that Summability Theory enters the picture, because that subject, more than any other, provides us with a rich supply of suitable matrices (lower triangular, non-negative entries, row sums $=1$). Working within this class, we say that

$$A \lhd B \tag{4}$$

Mathematics subject classification (2010): 44A60, 26D15, 40G05.
Keywords and phrases: Moment sequence, Hausdorff mean, convex sequence, unimodal sequence.

Paper MIA-14-63

if inequality (2) holds for all n and all convex sequences $\mathbf{x}$. Mercer's theorem gives the first example:

$$E\left(\tfrac{1}{2}\right) \lhd C(1). \tag{5}$$

(The matrix on the right in (1) is the famous arithmetic mean of Cesàro, while that on the left is familiar to summabilists as an Euler mean.)

The ordering (4) offers several surprises, not all pleasant ones. First, it is devoid of interest when applied to the simplest summability matrices, the *weighted means* (= forwards averages). And the same goes for the *Nørlund means* (= backwards averages). Interesting results, in fact, are available only when we come to consider Hausdorff means (section 2). But even then the possible comparisons are severely restricted (section 3). For example, the following two-sided estimate of the *iterated averages* of a convex sequence $\mathbf{x}$:

$$\frac{1}{\binom{n+3}{n}} \sum_{k=0}^{n} \binom{n-k+2}{n-k} x_k \leqslant \frac{1}{n+1} \sum_{j=0}^{n} \frac{1}{j+1} \sum_{k=0}^{j} x_k \tag{6}$$
$$\leqslant \frac{1}{\binom{n+\frac{1}{3}}{n}} \sum_{k=0}^{n} \binom{k-\frac{2}{3}}{k} x_k,$$

is pre-ordained. (The matrix on the left is a Cesàro mean (of order 3) and it cannot be replaced by any other Cesàro mean. Similarly, the matrix on the right, a Gamma mean, is the only one allowed there.)

Our main result, Theorem 1, determines all possible comparisons between two Hausdorff means. It turns out, rather surprisingly, that such a comparison is valid precisely when a certain sequence $\boldsymbol{\rho}$ is *totally monotonic,* i.e., when the *differences of* $\boldsymbol{\rho}$,

$$\Delta^n \rho_k = \sum_{j=0}^{n} (-1)^j \binom{n}{j} \rho_{j+k} \qquad (n,k = 0,1,2,\dots) \tag{7}$$

are all non-negative.

The connection between Mercer's inequality and moment sequences now becomes clear, courtesy of a fundamental result of Hausdorff.

THEOREM 0. ([9], Theorem 207) *A sequence* $\boldsymbol{\rho}$ *is totally monotonic if and only if it admits a representation of the form*

$$\rho_n = \int_0^1 \theta^n d\rho(\theta) \qquad (n = 0,1,2,\dots), \tag{8}$$

where $d\rho(\theta)$ *is a non-negative Borel measure on* $[0,1]$.

Hausdorff's theorem allows us to switch at will between total monotonicity and moment theory, and we shall see that it is advantageous to keep both viewpoints in mind.

Theorem 1 shows that Mercer-type inequalities are nothing more (and nothing less) than assertions of total monotonicity. Mercer's inequality (1), and those displayed in (6), are easy consequences of this observation.

But Theorem 1 marks the beginning of our analysis, rather than its end, because, despite Hausdorff's theorem, totally monotonic sequences are still not totally understood. When, for example, we seek the analogues of (6), for the r-fold iterated averages of a convex sequence, the algebra involved quickly becomes prohibitive as r $(= 1,2,3,\dots)$ increases. Similar obstacles arise when we compare other Hausdorff means and novel devices are called for at almost every turn (Lemmas 2–17). This, then, is the gist of our paper: *to seek new methods of identifying totally monotonic sequences.*

Mercer's inequality, of course, provides the motivation for our search, but considerable focus is added by the restrictions set out in section 3. There are, in fact, only nine possible comparisons between the most common Hausdorff means. These have all withstood extensive testing by computer and are listed as conjectures in section 3.

Theorem 1 is proved in section 4, the balance of our paper, sections 5–9, being devoted to a discussion of the nine conjectures.

2. Summability

The simplest summability methods are the *weighted means,* sequence transformations of the type

$$\mathbf{x} \longrightarrow \frac{a_0x_0+\cdots+a_nx_n}{a_0+\cdots+a_n}. \tag{9}$$

Here $\mathbf{a}$ is a sequence of positive "weights" and it is conventional to take $a_0 = 1$ and to denote the partial sums in upper case

$$A_n = a_0+a_1+\cdots+a_n. \tag{10}$$

PROPOSITION 1. *If two weighted means are comparable under the ordering* (4) *they must coincide.*

Proof. Let A and B be the two means, with respective weights $\mathbf{a}$ and $\mathbf{b}$, and suppose that $A \lhd B$. By setting

$$\mathbf{x} = \pm(n+1,n,n-1,\dots) \tag{11}$$

in the inequality

$$(A\mathbf{x})_n \leqslant (B\mathbf{x})_n, \tag{12}$$

we deduce that

$$\frac{A_0+\cdots+A_n}{A_n} = \frac{B_0+\cdots+B_n}{B_n}. \tag{13}$$

This system of equations has a unique solution (since $A_0 = B_0 = 1$) and the solution, of course, is

$$A_n = B_n. \tag{14}$$

Thus the two weighted means coincide. □

The backwards averages,

$$\mathbf{x} \longrightarrow \frac{a_n x_0 + a_{n-1}x_1 + \cdots + a_0 x_n}{A_n}, \tag{15}$$

are known as *Nørlund means.* They arise frequently in classical Analysis because of their simple action on power series.

PROPOSITION 2. *If two Nørlund means are comparable under the ordering* (4) *they must coincide.*

Proof. This is similar to the proof of Proposition 1. □

Propositions 1 and 2 are disappointing, to be sure, but there is a third class of summability methods, the Hausdorff means, wherein some action is guaranteed.

A *Hausdorff mean,* we recall, is a matrix of the type

$$(H_\mu)_{n,k} = \int_0^1 \binom{n}{k} \theta^k (1-\theta)^{n-k} d\mu(\theta), \tag{16}$$

where $d\mu(\theta)$ is a probability measure on $[0,1]$. Taking

$$d\mu(\theta) = d\theta, \tag{17}$$

we obtain the Cesàro matrix, while setting

$$d\mu(\theta) = \text{point evaluation at } \theta = \frac{1}{2} \tag{18}$$

leads to the Euler matrix $E(\frac{1}{2})$. Thus Mercer's theorem shows that there is some interest in comparing Hausdorff means. (This is fortunate because classical Summability Theory contains few concrete matrix results that lie beyond the pale of weighted, Nørlund or Hausdorff means.)

Our next two results are similar in spirit to Propositions 1 and 2. They again emphasize that interesting comparisons are to be found only when we restrict attention to Hausdorff means.

PROPOSITION 3. *If a weighted mean is comparable with a Hausdorff mean under the ordering* (4)*, then the weighted mean must, in fact, be Hausdorff.*

Proof. It follows from the identity

$$(n-k+1)\binom{n}{k} = (n+1)\binom{n}{k} - n\binom{n-1}{k-1} \tag{19}$$

that

$$\begin{aligned}
&\sum_{k=0}^{n}(n-k+1)\binom{n}{k}\theta^k(1-\theta)^{n-k} \qquad (20)\\
&= (n+1)\sum_{k=0}^{n}\binom{n}{k}\theta^k(1-\theta)^{n-k} - n\theta\sum_{k=1}^{n}\binom{n-1}{k-1}\theta^{k-1}(1-\theta)^{n-k}\\
&= (n+1)(\theta+(1-\theta))^n - n\theta(\theta+(1-\theta))^{n-1}\\
&= n+1-n\theta.
\end{aligned}$$

If H_μ is a Hausdorff mean, we deduce from (16) and (20) that

$$\sum_{k=0}^{n}(n-k+1)(H_\mu)_{n,k} = n+1-n\mu_1. \qquad (21)$$

Suppose now that A is a weighted mean that is comparable with H_μ. Applying both matrices to the sequences (11) shows that

$$\frac{A_0+A_1+\cdots+A_n}{A_n} = n+1-n\mu_1 \qquad (n=0,1,2,\ldots). \qquad (22)$$

Solving these equations for A_n in terms of μ_1 is a tedious exercise. But the solution is obviously unique, so it suffices to observe that (22) is satisfied by taking

$$A_n = \binom{n+\alpha}{n}, \quad \text{where } \alpha = \frac{\mu_1}{1-\mu_1} > 0. \qquad (23)$$

The weights $\mathbf{a}$ are then

$$a_n = A_n - A_{n-1} = \binom{n+\alpha-1}{n}, \qquad (24)$$

and the weighted mean takes the form

$$a_{n,k} = \frac{a_k}{A_n} = \frac{\binom{k+\alpha-1}{k}}{\binom{n+\alpha}{n}}. \qquad (25)$$

This turns out to be a Hausdorff mean because

$$a_{n,k} = \int_0^1 \binom{n}{k}\theta^k(1-\theta)^{n-k}\alpha\theta^{\alpha-1}d\theta. \qquad (26)$$

(The probability measure here is

$$d\mu(\theta) = \alpha\theta^{\alpha-1}d\theta \qquad (27)$$

and the matrix (25) is known as a Gamma mean.) □

PROPOSITION 4. *If a Nørlund mean is comparable with a Hausdorff mean under the ordering* (4)*, then the Nørlund mean must be Hausdorff.*

Proof. This is similar to the proof of Proposition 3 and we omit the details. It turns out that the Nørlund mean must be a Cesàro matrix of order α with

$$\alpha = \frac{1}{\mu_1} - 1 > 0. \tag{28}$$

The corresponding probability measure is

$$d\mu(\theta) = \alpha(1-\theta)^{\alpha-1} d\theta. \qquad \square \tag{29}$$

3. Hausdorff means

Mercer's theorem shows that non-trivial comparisons are possible between Hausdorff means. Our aim here is to point out that such comparisons are severely restricted.

The most commonly used Hausdorff means are those associated with the names of Cesàro, Euler and Hölder. We list them here, along with their measures and their first and second moments.

Cesàro means, $C(\alpha)$, $\alpha > 0$.

$$d\mu(\theta) = \alpha(1-\theta)^{\alpha-1} d\theta, \tag{30}$$
$$\mu_1 = \frac{1}{\alpha+1}, \qquad \mu_2 = \frac{2}{(\alpha+1)(\alpha+2)}.$$

These are also Nørlund means, and, in fact, the only Hausdorff means with this property.

Euler means, $E(\alpha)$, $0 \leqslant \alpha \leqslant 1$.

$$d\mu(\theta) = \text{point evaluation at } \theta = \alpha, \tag{31}$$
$$\mu_1 = \alpha, \qquad \mu_2 = \alpha^2.$$

Hölder means, $H(\alpha)$, $\alpha > 0$.

$$d\mu(\theta) = \frac{1}{\Gamma(\alpha)} |\log\theta|^{\alpha-1} d\theta, \tag{32}$$
$$\mu_1 = \frac{1}{2^\alpha}, \qquad \mu_2 = \frac{1}{3^\alpha}.$$

When α is a positive integer, these represent the α-fold iterated averages,

$$H(\alpha) = (C(1))^\alpha. \tag{33}$$

The "mass backwards" versions of the Cesàro means are also useful. These are called the Gamma means; their rows are the same as those of the corresponding Cesàro means, but are displayed in reverse order.

Gamma means, $G(\alpha)$, $\alpha > 0$.

$$d\mu(\theta) = \alpha\theta^{\alpha-1}d\theta, \tag{34}$$
$$\mu_1 = \frac{\alpha}{\alpha+1}, \qquad \mu_2 = \frac{\alpha}{\alpha+2}.$$

These are weighted means; the only Hausdorff means with this property.

Our next result gives a very simple necessary condition for two Hausdorff means to be comparable.

PROPOSITION 5. *Suppose that* H_μ *and* H_ν *are Hausdorff means and that*

$$H_\mu \lhd H_\nu. \tag{35}$$

Then

$$\mu_1 = \nu_1 \quad \textit{and} \quad \mu_2 \leqslant \nu_2. \tag{36}$$

Proof. The inequality

$$(H_\mu \mathbf{x})_1 \leqslant (H_\nu \mathbf{x})_1 \tag{37}$$

applied to the (convex) sequences $\mathbf{x} = \pm(0,1,2,\dots)$ guarantees that $\mu_1 = \nu_1$. Similarly, $\mu_2 \leqslant \nu_2$ follows from

$$(H_\mu \mathbf{x})_2 \leqslant (H_\nu \mathbf{x})_2 \tag{38}$$

by setting

$$\mathbf{x} = (0,0,1,2,3,\dots). \qquad \square \tag{39}$$

Proposition 5 is much more powerful than it appears at first sight; indeed, it allows only nine possible comparisons between the common Hausdorff means. These comparisons have all withstood extensive numerical testing and are listed below as conjectures. We explain the formulation of just the first two conjectures, leaving that of the others to the reader.

Conjectures (I) and (II) are concerned with comparisons between Cesàro means, $C(\beta)$, and Hölder means, $H(\alpha)$. No such comparison can be valid unless their first moments agree, and this forces us to accept

$$\frac{1}{\beta+1} = \frac{1}{2^\alpha}. \tag{40}$$

Moreover, the *direction* of any such comparison, $C(\beta) \lhd H(\alpha)$ (or $H(\alpha) \lhd C(\beta)$), is dictated by the second moments:

$$\frac{2}{(\beta+1)(\beta+2)} \leqslant \frac{1}{3^\alpha} \tag{41}$$

(or the reversal of (41)). Inequality (41), when combined with (40), may be rephrased as

$$2(3^\alpha) \leqslant 2^\alpha + 4^\alpha. \tag{42}$$

This estimate is valid, via convexity, whenever $\alpha \geqslant 1$ (and it switches direction whenever $0 < \alpha \leqslant 1$). It follows, then, that Conjecture (I) can only be true if $\alpha \geqslant 1$ (and Conjecture (II) only if $0 < \alpha \leqslant 1$).

The conjectures are:

(I)	$C(2^\alpha - 1) \lhd H(\alpha)$	if $\alpha \geqslant 1$;
(II)	$H(\alpha) \lhd C(2^\alpha - 1)$	if $0 < \alpha \leqslant 1$;
(III)	$H(\alpha) \lhd G\left(\frac{1}{2^\alpha - 1}\right)$	if $\alpha \geqslant 1$;
(IV)	$G\left(\frac{1}{2^\alpha - 1}\right) \lhd H(\alpha)$	if $0 < \alpha \leqslant 1$;
(V)	$C(\alpha) \lhd G\left(\frac{1}{\alpha}\right)$	if $\alpha \geqslant 1$;
(VI)	$G\left(\frac{1}{\alpha}\right) \lhd C(\alpha)$	if $0 < \alpha \leqslant 1$;
(VII)	$E\left(\frac{1}{\alpha+1}\right) \lhd C(\alpha)$	if $\alpha > 0$;
(VIII)	$E\left(\frac{1}{2^\alpha}\right) \lhd H(\alpha)$	if $\alpha > 0$;
(IX)	$E\left(\frac{\alpha}{\alpha+1}\right) \lhd G(\alpha)$	if $\alpha > 0$.

There are a couple of redundancies in the above list: (V) follows from (I) and (III), while (VI) follows from (II) and (IV). But, since I am unable to prove any of the conjectures (I)–(IV), these redundancies remain redundant. (Conjectures (V) and (VI), meanwhile, are confirmed in section 7.)

It is essential to point out here that Proposition 5 does not solve the comparison problem completely, else all would be trivial.

PROPOSITION 6. *Condition* (36) *is not sufficient for comparison* (35) *to hold.*

Proof. We exhibit a Hausdorff mean, H_ν, with

$$\frac{1}{2} = \nu_1 \quad \text{and} \quad \frac{1}{3} < \nu_2, \tag{43}$$

for which the comparison

$$C(1) \lhd H_\nu \tag{44}$$

fails to be valid.

To do this, we specify only the first four moments of $\boldsymbol{\nu}$:

$$\boldsymbol{\nu} = \left(1, \frac{1}{2}, \frac{11}{32}, \frac{23}{96}, \dots\right), \tag{45}$$

happy in the knowledge that there are infinitely many ways to fill in the remaining coordinates. (See the remark following Theorem 19 of [2].) It is clear that (43) is satisfied, whereas (44) is not, because

$$\frac{1}{4} > v_3. \tag{46}$$

The reverse inequality is needed for comparison (44) to hold, as is seen by replacing the sequence in (39) by

$$\mathbf{x} = (0,0,0,1,2,3,\dots). \qquad \square \tag{47}$$

4. Main result

In this section we obtain necessary and sufficient conditions for two Hausdorff means to be comparable. Our analysis is based upon a simple lemma of Levin and Stečkin ([14], D4). It is rather surprising that their result, which has been rediscovered by several authors, does not appear in the classical monograph [10]. A proof is given here because we have chosen to rephrase the statement of the lemma in a form that is more convenient for our applications.

LEMMA 1. ([14], page 2) *Suppose that $a_0, a_1, \dots, a_N$ are real numbers and that*

$$A_n := \sum_{k=0}^{n} a_k, \quad \mathscr{A}_n := \sum_{k=0}^{n} A_k \qquad (n = 0,1,\dots,N). \tag{48}$$

Then the inequality

$$\sum_{k=0}^{N} a_k x_k \geqslant 0 \tag{49}$$

is valid for all convex sequences $\mathbf{x}$ *if and only if*

$$A_N = \mathscr{A}_N = 0 \tag{50}$$

and

$$\mathscr{A}_n \geqslant 0 \qquad (n = 0,1,\dots,N). \tag{51}$$

Proof. (Necessity.) Taking $\mathbf{x} = \pm(1,1,1,\dots)$ or $\mathbf{x} = \pm(N+1, N, N-1, \dots)$ in (49) guarantees that (50) holds. (51) follows similarly by taking $\mathbf{x} = (n+1, n, \dots, 1, 0, \dots, 0)$.

(Sufficiency.) Summing by parts twice, we see that

$$\begin{aligned}\sum_{k=0}^{N} a_k x_k &= \sum_{k=0}^{N-1} A_k (x_k - x_{k+1}) + A_N x_N \\ &= \sum_{k=0}^{N-2} \mathscr{A}_k (x_k - 2x_{k+1} + x_{k+2}) + \boldsymbol{\mathscr{A}_{N-1}(x_{N-1} - x_N) + A_N x_N}.\end{aligned}$$

The boldface terms vanish because $\mathscr{A}_{N-1} = \mathscr{A}_N - A_N = 0$, and the remaining terms are non-negative since $\mathbf{x}$ is convex. $\square$

THEOREM 1. *Suppose that H_μ and H_ν are Hausdorff means. Then*

$$H_\mu \lhd H_\nu \tag{52}$$

if and only if

$$\mu_1 = \nu_1 \tag{53}$$

and the sequence $\boldsymbol{\rho}$ is totally monotonic, where

$$\rho_n = \frac{\nu_{n+2} - \mu_{n+2}}{(n+1)(n+2)} \qquad (n = 0, 1, 2, \ldots). \tag{54}$$

Proof. Let

$$H = H_\nu - H_\mu. \tag{55}$$

The lemma shows that comparison (52) holds precisely when

$$\sum_{k=0}^{n} h_{n,k} = 0, \quad \sum_{j=0}^{n}\sum_{k=0}^{j} h_{n,k} = 0 \qquad (n = 0, 1, 2, \ldots) \tag{56}$$

and

$$\sum_{j=0}^{m}\sum_{k=0}^{j} h_{n,k} \geqslant 0 \qquad (m = 0, 1, \ldots, n). \tag{57}$$

The first condition in (56) is satisfied automatically since H_μ and H_ν both have row sums 1. The second condition is equivalent to (53) because

$$\sum_{j=0}^{n}\sum_{k=0}^{j} h_{n,k} = n(\mu_1 - \nu_1) \tag{58}$$

by identity (21). The balance of the proof consists of showing the equivalence of (57) with (54). This entails a brute force evaluation of the sums in (57), a task made much shorter by the following result.

LEMMA 2. *If H_μ is a Hausdorff mean, its entries are expressible in terms of the associated moment sequence by the formula*

$$(H_\mu)_{n,k} = \sum_i (-1)^{i+k} \binom{n}{i} \mu_i \binom{i}{k}. \tag{59}$$

Proof. This is the familiar "$\delta\mu\delta$–representation" for Hausdorff matrices ([9], section 11.3). It is a simple exercise in binomialcoefficientology; we do not give the details here. □

Instead, we wish to emphasize that the summation in (59) runs over all integers i, rather than over the conventional interval, $k \leqslant i \leqslant n$. This feat is accomplished in the easiest possible way, *by fiat:* we simply extend Pascal's triangle to cover the whole plane of integer pairs by declaring that

$$\binom{n}{i} = \begin{cases} \frac{n(n-1)\cdots(n-i+1)}{i!} & \text{if } i \geqslant 0 \\ 0 & \text{if } i < 0. \end{cases} \tag{60}$$

The advantages of this convention are made clear in the wonderful text [8]: *interchanging the order of summation in certain double sums becomes a matter of routine.* Our subsequent proof of Theorem 1 provides an excellent illustration of this tactic.

It follows from Lemma 2 that

$$h_{n,k} = \sum_i (-1)^{i+k} \binom{n}{i} \binom{i}{k} \pi_i, \tag{61}$$

where

$$\pi_k = \nu_k - \mu_k. \tag{62}$$

We have therefore

$$\begin{aligned} \sum_{j=0}^{m} \sum_{k=0}^{j} h_{n,k} &= \sum_{k=0}^{m} (m-k+1) h_{n,k} \\ &= \sum_{k=0}^{m} (m-k+1) \sum_i (-1)^{i+k} \binom{n}{i} \binom{i}{k} \pi_i \\ &= \sum_i (-1)^i \binom{n}{i} \pi_i \sum_{k=0}^{m} (-1)^k (m-k+1) \binom{i}{k}. \end{aligned}$$

Using the identity

$$(-1)^k (m-k+1) = (-1)^m \binom{-2}{m-k}, \tag{63}$$

which is easily checked by expanding the binomial coefficient, the inner sum may be simplified as follows.

$$\begin{aligned} \sum_{k=0}^{m} (-1)^k (m-k+1) \binom{i}{k} &= (-1)^m \sum_{k=0}^{m} \binom{-2}{m-k} \binom{i}{k} \\ &= (-1)^m \binom{i-2}{m}, \end{aligned}$$

the last step courtesy of Vandermonde's formula ([8], page 174).

We have shown that

$$\sum_{j=0}^{m} \sum_{k=0}^{j} h_{n,k} = \sum_i (-1)^{i+m} \binom{n}{i} \binom{i-2}{m} \pi_i.$$

The summation on i is relevant only over the range, $0 \leqslant i \leqslant n$, because of the presence of the binomial coefficient $\binom{n}{i}$. It can, in fact, be further restricted, to $2 \leqslant i \leqslant n$, since

π_0 and π_1 are both zero by (62). But then $\binom{i-2}{m}$ vanishes as well, unless $i \geqslant m+2$, so that the summation on i actually registers only when $m+2 \leqslant i \leqslant n$.

Setting $h = i-m-2$, we have

$$\sum_{j=0}^{m}\sum_{k=0}^{j} h_{n,k} = \sum_{h=0}^{n-m-2}(-1)^h \binom{n}{h+m+2}\binom{h+m}{m}\pi_{h+m+2}.$$

The awful expression on the right becomes recognizable only after its binomial coefficients have been repackaged:

$$\binom{n}{h+m+2}\binom{h+m}{m} = \frac{n(n-1)}{(h+m+1)(h+m+2)}\binom{n-2}{m}\binom{n-m-2}{h}. \tag{64}$$

We then have

$$\begin{aligned}\sum_{j=0}^{m}\sum_{k=0}^{j} h_{n,k} &= n(n-1)\binom{n-2}{m}\sum_{h=0}^{n-m-2}(-1)^h\binom{n-m-2}{h}\frac{\pi_{h+m+2}}{(h+m+1)(h+m+2)}\\ &= n(n-1)\binom{n-2}{m}\sum_{h=0}^{n-m-2}(-1)^h\binom{n-m-2}{h}\rho_{h+m}\\ &= n(n-1)\binom{n-2}{m}\Delta^{n-m-2}\rho_m.\end{aligned}$$

The terms here are non-negative for all n and m, as demanded by (57), precisely when the sequence $\boldsymbol{\rho}$ is totally monotonic. □

The following simple corollaries of Theorem 1 are all dependent upon the well-known fact that

the product of two totally monotonic sequences is again totally monotonic. (65)

The proof of this assertion uses the discrete analogue of *Leibnitz's formula* for the derivatives of a product:

$$\Delta^n x_k y_k = \sum_{j=0}^{n}\binom{n}{j}\left(\Delta^j x_k\right)\left(\Delta^{n-j} y_{j+k}\right). \tag{66}$$

See Theorem 210 of [9].

Our first corollary is equivalent to the inequalities (6) mentioned in section 1.

COROLLARY 1. $C(3) \lhd H(2) \lhd G(\frac{1}{3})$.

Proof. We apply Theorem 1, noting first that the Hausdorff means $C(3)$, $H(2)$ and $G(\frac{1}{3})$ all have the same first moments $(=\frac{1}{4})$.

The left-hand comparison, $C(3) \lhd H(2)$, amounts to showing that the sequence

$$\frac{\frac{1}{(n+3)^2} - \frac{1}{\binom{n+5}{3}}}{(n+1)(n+2)} \qquad (n = 0,1,2,\ldots) \tag{67}$$

is totally monotonic. But this is obvious since (67) may be rephrased as

$$\frac{1}{(n+3)^2(n+4)(n+5)}, \tag{68}$$

a product of totally monotonic sequences.

The right-hand comparison, $H(2) \triangleleft G(\frac{1}{3})$, succumbs to a similar representation:

$$\frac{\frac{1}{3n+7} - \frac{1}{(n+3)^2}}{(n+1)(n+2)} = \frac{1}{(n+3)^2(3n+7)}. \qquad \square \tag{69}$$

COROLLARY 2. (Mercer's inequality)

$$E\left(\tfrac{1}{2}\right) \triangleleft C(1). \tag{70}$$

Proof. It suffices, in view of Theorem 1, to show that the sequence

$$\frac{\frac{1}{n+3} - \frac{1}{2^{n+2}}}{(n+1)(n+2)} \qquad (n = 0, 1, 2, \dots) \tag{71}$$

is totally monotonic. But this is obvious because (71) may be expressed as

$$\frac{1}{n+3} \cdot \mu_n, \tag{72}$$

where $\boldsymbol{\mu}$ is the moment sequence of the measure

$$d\mu(\theta) = \begin{cases} 0 & \text{if } 0 \leqslant \theta < \frac{1}{2} \\ (1-\theta)d\theta & \text{if } \frac{1}{2} \leqslant \theta \leqslant 1. \end{cases} \qquad \square \tag{73}$$

The analogue of Corollary 1 for 3-fold iterated averages may be proved in the same way.

COROLLARY 3. *If* $\mathbf{x}$ *is a convex sequence, then*

$$\frac{1}{\binom{n+7}{n}} \sum_{k=0}^{n} \binom{n-k+6}{n-k} x_k \leqslant \frac{1}{n+1} \sum_{i=0}^{n} \frac{1}{i+1} \sum_{j=0}^{i} \frac{1}{j+1} \sum_{k=0}^{j} x_k$$
$$\leqslant \frac{1}{\binom{n+\frac{1}{7}}{n}} \sum_{k=0}^{n} \binom{k-\frac{6}{7}}{k} x_k.$$

We do not give the proof here because the details are very tedious. An alternative approach to a more general result is adopted in sections 5 and 6.

5. Iterated averages of convex sequences

According to Conjectures (I) and (III), the analogue of (6) for r-fold iterated averages $(r = 1,2,3,\dots,)$ is

$$C(2^r - 1) \lhd H(r) \lhd G\left(\tfrac{1}{2^r-1}\right). \tag{74}$$

These comparisons, unfortunately, are difficult to derive directly from Theorem 1 because of the algebra involved. Our aim here is to circumvent these difficulties by adopting an entirely different approach, one that hinges upon the wonderful structural properties of Hausdorff means. We deal first with the comparison $H(r) \lhd G(\frac{1}{2^r-1})$, postponing treatment of $C(2^r-1) \lhd H(r)$ till section 6.

LEMMA 3. *The product of two Hausdorff means is again a Hausdorff mean. Moreover, the moment sequence of the product is the product of the moment sequences.*

Proof. See Chapter XI of [9]. □

LEMMA 4. *Hausdorff means preserve convexity.*

Proof. This is a special case of Theorem 1 of [3]. The proof is based upon the identity

$$y_n - 2y_{n+1} + y_{n+2} = \sum_{k=0}^{n} \binom{n}{k} \left(\Delta^{n-k}\mu_{k+2}\right)(x_k - 2x_{k+1} + x_{k+2}) \tag{75}$$

which is valid whenever

$$\mathbf{y} = H_\mu \mathbf{x}. \qquad \square \tag{76}$$

LEMMA 5. *Suppose that H_μ, H_ν and H_π are Hausdorff means. Then*

$$H_\mu \lhd H_\nu \Longrightarrow H_\mu H_\pi \lhd H_\nu H_\pi. \tag{77}$$

Proof. This useful observation follows at once from Lemma 4 when definition (4) is recalled. □

LEMMA 6. *If $\alpha > 0$, then*

$$G(\alpha)G(1) \lhd G\left(\frac{\alpha}{\alpha+2}\right). \tag{78}$$

Proof. This is a comparison of two Hausdorff means, thanks to Lemma 3, their respective moment sequences being

$$\mu_n := \frac{\alpha}{n+\alpha} \cdot \frac{1}{n+1} \quad \text{and} \quad \nu_n := \frac{\frac{\alpha}{\alpha+2}}{n + \frac{\alpha}{\alpha+2}}. \tag{79}$$

It is obvious that their first moments agree $(\mu_1 = \nu_1)$ and that the sequence

$$\frac{\nu_{n+2} - \mu_{n+2}}{(n+1)(n+2)} = \frac{\alpha}{(n+\alpha+2)(n+3)((n+2)(\alpha+2)+\alpha)} \tag{80}$$

is totally monotonic. Theorem 1, then, guarantees that comparison (78) is valid. □

THEOREM 2. *If* $r = 1, 2, 3, \ldots$, *then*

$$H(r) \lhd G\left(\frac{1}{2^r - 1}\right). \tag{81}$$

Proof. The comparison is trivial when $r = 1$, both matrices then coinciding with the arithmetic mean, $C(1)$. We proceed by induction on r $(= 2, 3, 4, \ldots)$, deducing comparison (81) directly from

$$H(r-1) \lhd G\left(\frac{1}{2^{r-1} - 1}\right). \tag{82}$$

We have

$$\begin{aligned} H(r) &= H(r-1)G(1) && \text{by Lemma 3} \\ &\lhd G\left(\frac{1}{2^{r-1}-1}\right)G(1) && \text{by (82) and Lemma 5} \\ &\lhd G\left(\frac{1}{2^r-1}\right) && \text{by Lemma 6.} \end{aligned}$$

□

6. Iterated averages (continued)

We now discuss the complement to Theorem 2, which is given by

THEOREM 3. *If* $r = 1, 2, 3, \ldots$, *then*

$$C(2^r - 1) \lhd H(r). \tag{83}$$

Proof. The basic idea is the same as that used in Theorem 2: we begin by observing that comparison (83) is trivial when $r = 1$, and then proceed by induction. The details, however, are now more troublesome because Lemma 6 must be replaced by Lemma 7 (listed below) and the latter result is by no means an easy consequence of Theorem 1.

Taking Lemma 7 for granted, let us suppose that comparison (83) is valid for some r. We then have

$$\begin{aligned} C(2^{r+1}-1) &\lhd C(2^r-1)C(1) && \text{by Lemma 7} \\ &\lhd H(r)C(1) && \text{by (83) and Lemma 5} \\ &= H(r+1) && \text{by Lemma 3.} \end{aligned}$$

□

LEMMA 7. *If $\alpha > 0$, then*

$$C(2\alpha+1) \lhd C(\alpha)C(1). \tag{84}$$

Proof. We begin by recalling that Hausdorff means commute ([9], Theorem 197). It suffices, therefore, by Lemma 5, to find a matrix, say Q, such that

$$C(2\alpha+1) = Q.C(\alpha) \tag{85}$$

and

$$Q \lhd C(1). \tag{86}$$

This seemingly difficult task turns out to have a straightforward solution, courtesy of Lemma 3. Q, in fact, must be a Hausdorff mean, the one associated with the moment sequence

$$\frac{\binom{n+\alpha}{n}}{\binom{n+2\alpha+1}{n}} \qquad (n = 0,1,2,\dots). \tag{87}$$

To see that (87) is indeed a moment sequence, we consider the probability measure on $[0,1]$:

$$d\mu(\theta) = \frac{\Gamma(\beta+\gamma)}{\Gamma(\beta)\Gamma(\gamma)}\,\theta^{\beta-1}(1-\theta)^{\gamma-1}d\theta, \tag{88}$$

where $\beta,\gamma > 0$ are fixed. This measure generates a Hausdorff mean, say $C(\beta,\gamma)$, with entries

$$C(\beta,\gamma)_{n,k} = \frac{\binom{k+\beta-1}{k}\binom{n-k+\gamma-1}{n-k}}{\binom{n+\beta+\gamma-1}{n}} \tag{89}$$

and moment sequence

$$\frac{\binom{n+\beta-1}{n}}{\binom{n+\beta+\gamma-1}{n}} \qquad (n = 0,1,2,\dots). \tag{90}$$

Q, evidently, is the Hausdorff mean $C(\alpha+1,\alpha+1)$, so that Lemma 7 follows at once from our next result. $\square$

LEMMA 8. *If $\alpha \geqslant 1$, then*

$$C(\alpha,\alpha) \lhd C(1). \tag{91}$$

Lemma 8 turns out to be an elusive result, the key to its proof being a brilliant observation of Laguerre [13].

LEMMA 9. *Suppose that $(a_0,a_1,\dots,a_N)$ is a sequence of real numbers, not identically zero, and that $a_0+a_1+\cdots+a_N = 0$. Then the sequence of partial sums, $(A_0,A_1,\dots,A_N)$, has strictly fewer sign changes than does $(a_0,a_1,\dots,a_N)$.*

Lemma 9 enables us to work directly from Lemma 1, thereby ignoring Theorem 1. The idea (in the context of Lemma 1) is to suppose that $(a_0, a_1, \dots, a_N)$ suffers at most two changes of sign; then $(A_0, A_1, \dots, A_N)$ suffers at most one; and $(\mathscr{A}_0, \mathscr{A}_1, \dots, \mathscr{A}_N)$ suffers none at all. This special case of Laguerre's result is, of course, a triviality, and it is presented as such, without proof, as Lemma 7.2 of [12]. A thorough exposition of Laguerre's work may be found in Part Five of [17]. See also Section 6.8 of [11].

We do not give a proof of Lemma 9 because only its trivial version is required here. Indeed, we shall be working with *unimodal sequences* (those that increase, then decrease) and it is clear that such sequences suffer no more than two changes of sign (from $-$ to $+$ to $-$).

LEMMA 10. *Suppose that H_μ and H_ν are Hausdorff means and that $\mu_1 = \nu_1$. If the rows of $H_\mu - H_\nu$ are unimodal, then*

$$H_\mu \lhd H_\nu. \tag{92}$$

Proof. The matrix H, defined in (55), obviously satisfies condition (56); at issue, then, is whether it also satisfies (57).

Unimodality, however, guarantees that the rows of H each have at most two sign changes (from $+$ to $-$ to $+$); that their partial sums have at most one (from $+$ to $-$); and that *their* partial sums have none at all $(+)$. $\square$

Unimodal sequences play a significant role in Combinatorics and in Probability Theory, and they have been studied extensively (and independently) in both these contexts. See, for example, [5] and [6]. The product of two unimodal sequences, unfortunately, need not be unimodal, and this defect forces us to consider a slightly more restrictive class:

$$\frac{a_0}{a_1} \leqslant \frac{a_1}{a_2} \leqslant \cdots \leqslant \frac{a_{N-1}}{a_N}, \tag{93}$$

namely, the positive *log-concave sequences.* This class is obviously closed under products, and each of its members is unimodal. [Think of placing 1, in its correct position, in the ordered list (93).]

LEMMA 11. *If $\alpha \geqslant 1$, the rows of the matrix*

$$C(\alpha,\alpha) - C(1) \tag{94}$$

are unimodal.

Proof. The n^{th} row $(n = 0, 1, 2, \dots)$ of the matrix (94) is given by

$$\frac{\binom{k+\alpha-1}{k}\binom{n-k+\alpha-1}{n-k}}{\binom{n+2\alpha-1}{n}} - \frac{1}{n+1} \qquad (k = 0, 1, \dots, n). \tag{95}$$

Removal of the constant terms (n being fixed) neither helps nor hinders our assertion, so it suffices to show that the sequence

$$\binom{k+\alpha-1}{k}\binom{n-k+\alpha-1}{n-k} \qquad (k = 0, 1, \dots, n) \tag{96}$$

is unimodal.

It is here that log-concavity becomes relevant. The sequence

$$\binom{k+\alpha-1}{k} \qquad (k=0,1,\dots,n) \tag{97}$$

is certainly (positive and) log-concave because

$$\begin{aligned}\binom{k+\alpha}{k+1}^2 - \binom{k+\alpha-1}{k}\binom{k+\alpha+1}{k+2} \qquad (0<k<n)\\ = \frac{\alpha-1}{(k+\alpha)(k+2)}\binom{k+\alpha}{k+1}^2\\ \geqslant 0.\end{aligned}$$

And the same goes for the sequence

$$\binom{n-k+\alpha-1}{n-k} \qquad (k=0,1,\dots,n) \tag{98}$$

since the system of inequalities (93) is *mass-backwards invariant:*

$$\frac{a_N}{a_{N-1}} \leqslant \frac{a_{N-1}}{a_{N-2}} \leqslant \dots \leqslant \frac{a_1}{a_0}. \tag{99}$$

It follows that the sequence (96) is log-concave, being the product of two log-concave sequences, (97) and (98). Being also positive, (96) must be unimodal. □

Lemmas 10 and 11 show that Lemma 8 is valid, and this observation completes our proof of Theorem 3.

7. Cesàro versus Gamma means

We have by now developed enough machinery to prove conjectures (V) and (VI). The key observation here is that the rows of the Cesàro means (and of the Gamma means) are especially well-behaved: *they are either all convex or all concave* (and which is which is determined solely by α, the order of the mean).

LEMMA 12. *The sequence*

$$\binom{k+\alpha-1}{k} \qquad (k=0,1,\dots,n) \tag{100}$$

is convex if $0 \leqslant \alpha \leqslant 1$ *or* $\alpha \geqslant 2$, *and concave if* $1 \leqslant \alpha \leqslant 2$.

Proof.

$$\begin{aligned}\binom{k+\alpha-1}{k} + \binom{k+\alpha+1}{k+2} - 2\binom{k+\alpha}{k+1}\\ = \frac{(\alpha-1)(\alpha-2)}{(k+2)(k+\alpha)}\binom{k+\alpha}{k+1}\end{aligned}$$

is non-negative if $0 \leqslant \alpha \leqslant 1$ or $\alpha \geqslant 2$, and non-positive if $1 \leqslant \alpha \leqslant 2$. □

It follows from Lemma 12 that the rows of $G(\alpha)$, $\alpha > 0$, are concave if $1 \leqslant \alpha \leqslant 2$, and convex otherwise. And the same goes for the rows of $C(\alpha)$ (which, after all, are merely those of $G(\alpha)$ displayed in reverse order). The relevance of these remarks is seen in the following result.

LEMMA 13. *A concave sequence of real numbers $(a_0, a_1, \ldots, a_N)$ is unimodal.*

Proof. The differences of $\mathbf{a}$ form an increasing sequence

$$a_0 - a_1 \leqslant a_1 - a_2 \leqslant \cdots \leqslant a_{N-1} - a_N. \tag{101}$$

By adding 0 to this ordered list, in its correct position, unimodality becomes obvious. □

LEMMA 14. *If $\alpha > 1$, then*

$$G\left(\frac{1}{\alpha-1}\right) G(\alpha) \lhd G\left(\frac{1}{\alpha}\right). \tag{102}$$

Proof. This is similar to Lemma 6. We are comparing here two Hausdorff means with moment sequences

$$\mu_n := \frac{1}{(\alpha-1)n+1} \cdot \frac{\alpha}{n+\alpha} \quad \text{and} \quad \nu_n := \frac{1}{\alpha n+1}. \tag{103}$$

Theorem 1 completes the proof because the first moments agree and the sequence

$$\frac{\nu_{n+2} - \mu_{n+2}}{(n+1)(n+2)} = \frac{\alpha - 1}{(\alpha(n+2)+1)((\alpha-1)(n+2)+1)(n+2+\alpha)} \tag{104}$$

is obviously totally monotonic. □

Our next result confirms Conjectures (V) and (VI).

THEOREM 4. *If $\alpha \geqslant 1$, then*

$$C(\alpha) \lhd G\left(\frac{1}{\alpha}\right), \tag{105}$$

and

$$G\left(\frac{1}{\alpha}\right) \lhd C(\alpha) \tag{106}$$

if $0 < \alpha \leqslant 1$.

Proof. Comparison (105) is certainly valid whenever $1 \leqslant \alpha \leqslant 2$, by Lemma 10, because the rows of $C(\alpha) - G(\frac{1}{\alpha})$ are unimodal. [To see this we note that the rows of $C(\alpha)$ are concave by Lemma 12, and that the rows of $G(\frac{1}{\alpha})$ are convex. It follows, then, that the rows of $C(\alpha) - G(\frac{1}{\alpha})$ are concave, and hence, by Lemma 13, unimodal.]

The proof of comparison (105) is completed by means of an induction argument. If $2 < \alpha \leqslant 3$ (then, similarly, if $3 < \alpha \leqslant 4$, etc.), we observe that

$$\begin{aligned} C(\alpha) &= C(\alpha-1)G(\alpha) \quad \text{by Lemma 3} \\ &\lhd G\left(\frac{1}{\alpha-1}\right)G(\alpha) \quad \text{by Lemma 5} \\ &\lhd G\left(\frac{1}{\alpha}\right) \qquad \text{by Lemma 14.} \end{aligned} \tag{107}$$

The reverse comparison, (106), follows from (105) by means of a simple mass-backwards argument. Indeed, (105) asserts that, for each fixed n $(n = 0,1,2,\dots)$ and for each $\alpha \geqslant 1$, the inequality

$$\frac{1}{\binom{n+\alpha}{n}} \sum_{k=0}^{n} \binom{n-k+\alpha-1}{n-k} x_k \leqslant \frac{1}{\binom{n+\frac{1}{\alpha}}{n}} \sum_{k=0}^{n} \binom{k+\frac{1}{\alpha}-1}{k} x_k \tag{108}$$

is valid for all convex sequences $\mathbf{x} = (x_0, x_1, \dots, x_n)$. Reversing the order of both summations in (108), and replacing $\mathbf{x}$ by $\mathbf{y}$ where

$$(y_0, y_1, \dots, y_n) = (x_n, x_{n-1}, \dots, x_0), \tag{109}$$

we deduce that

$$\frac{1}{\binom{n+\alpha}{n}} \sum_{k=0}^{n} \binom{k+\alpha-1}{k} y_k \leqslant \frac{1}{\binom{n+\frac{1}{\alpha}}{n}} \sum_{k=0}^{n} \binom{n-k+\frac{1}{\alpha}-a}{n-k} y_k. \tag{110}$$

The transformation (109) has no affect at all on convexity, so that inequality (110) is valid for all convex sequences $\mathbf{y}$. This shows that

$$G(\alpha) \lhd C\left(\frac{1}{\alpha}\right) \qquad (\alpha \geqslant 1), \tag{111}$$

which comparison is equivalent to (106). $\square$

It is perhaps worth pointing out that comparison (111) would have been difficult to establish without recourse to the mass-backwards argument used above. To be sure, there is no difficulty when $1 \leqslant \alpha \leqslant 2$ because the proof of (105), *concave-convex* $\Longrightarrow$ *unimodal,* applies here as well. But things go awry when $\alpha \geqslant 2$ because the analogue of (107), namely

$$G(\alpha) = G(\alpha-1)Q, \tag{112}$$

involves a matrix, Q, that is not a Hausdorff mean.

The plot thickens when Theorem 1 is added to the mix, because the equivalence of (105) with (111) then asserts that both the sequences

$$\frac{\frac{1}{\alpha(n+2)+1}-\frac{1}{\binom{n+2+\alpha}{n+2}}}{(n+1)(n+2)} \quad \text{and} \quad \frac{\frac{1}{\binom{n+2+\frac{1}{\alpha}}{n+2}}-\frac{\alpha}{n+2+\alpha}}{(n+1)(n+2)} \tag{113}$$

are totally monotonic if either one of them is. The reader is invited to prove this equivalence directly.

Fortunately, there is a simple explanation:

PROPOSITION 7. *Suppose that* $(\mu_0,\mu_1,\mu_2,\dots)$ *and* $(\nu_0,\nu_1,\nu_2,\dots)$ *are sequences of real numbers satisfying*

$$\mu_0=\nu_0 \quad \textit{and} \quad \mu_1=\nu_1. \tag{114}$$

Then the sequence

$$a_n := \frac{\nu_{n+2}-\mu_{n+2}}{(n+1)(n+2)} \tag{115}$$

is totally monotonic if and only if

$$b_n := \frac{\Delta^{n+2}\nu_0-\Delta^{n+2}\mu_0}{(n+1)(n+2)} \tag{116}$$

is too.

Proof. Use the fact that $b_n=\Delta^n a_0$ and $a_n=\Delta^n b_0$. □

8. Probability measures

Conjectures (VII), (VIII) and (IX) are perhaps the simplest ones in our list because they are direct generalizations of Mercer's inequality (to which they all collapse when $\alpha=1$). They may each be proved separately, but their common format (an Euler mean versus a Hausdorff mean) suggests that a common proof might be available.

CONJECTURE X. If H_μ is an arbitrary Hausdorff mean, then

$$E(\mu_1) \lhd H_\mu. \tag{117}$$

Here μ_1 is the first moment of the measure associated with H_μ, and it is plain that Conjectures (VII), (VIII) and (IX) are special cases of (117).

Theorem 1 is not of much help here because it demands that the sequence (119) be totally monotonic whenever $\boldsymbol{\mu}$ is the moment sequence of a Hausdorff mean. It does, however, radically change the nature of the conjecture from one concerned with Mercer's inequality/Summability Theory/Convexity to one that deals exclusively with probability measures.

CONJECTURE XI. If $d\mu(\theta)$ is an arbitrary probability measure on $[0,1]$, with moments

$$\mu_n = \int_0^1 \theta^n d\mu(\theta), \tag{118}$$

then the sequence

$$\frac{\mu_{n+2} - \mu_1^{n+2}}{(n+1)(n+2)} \qquad (n = 0, 1, 2, \dots) \tag{119}$$

is totally monotonic.

There is nothing like this assertion in the standard texts on Moment Theory ([2], [9], [19] and [20]); the remainder of this section is devoted to its proof.

LEMMA 15. *If* $\mathbf{a}$*,* $\mathbf{b}$ *and* $\mathbf{c}$ *are totally monotonic sequences, then so is* $\mathbf{d}$*, where*

$$d_n = \sum_{k=0}^{n} \binom{n}{k} a_k \left(\Delta^{n-k} b_k\right) c_{n-k} \qquad (n = 0, 1, 2, \dots). \tag{120}$$

Proof. The identity

$$\Delta^m d_n = \sum_{j=0}^{m} \sum_{k=0}^{n} \binom{m}{j} \binom{n}{k} \left(\Delta^j a_k\right) \left(\Delta^{m+n-j-k} b_{j+k}\right) \left(\Delta^{m-j} c_{n-k}\right) \tag{121}$$

shows that the differences of $\mathbf{d}$ are all non-negative.

The key to proving (121) is to recognize its purely algebraic nature; it has nothing at all to do with total monotonicity, being valid, in fact, for arbitrary sequences $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$.

When confronted with a novel identity such as (121), it makes sense to check things out by working first with *geometric sequences,* because differences are then easy to evaluate:

$$x_n = x^n \Longrightarrow \Delta^m x_n = (1-x)^m x^n. \tag{122}$$

Replacing a_k by a^k, b_k by b^k, and c_k by c^k, $\mathbf{d}$, itself, is a geometric sequence:

$$\begin{aligned} d_n &= \sum_{k=0}^{n} \binom{n}{k} a^k (1-b)^{n-k} b^k c^{n-k} \\ &= (ba + (1-b)c)^n. \end{aligned}$$

Its differences, according to (122), are given by

$$\begin{aligned} \Delta^m d_n &= (b(1-a) + (1-b)(1-c))^m (ba + (1-b)c)^n \\ &= \sum_{j=0}^{m} \sum_{k=0}^{n} \binom{m}{j} \binom{n}{k} (1-a)^j a^k (1-b)^{m+n-j-k} b^{j+k} (1-c)^{m-j} c^{n-k} \\ &= \sum_{j=0}^{m} \sum_{k=0}^{n} \binom{m}{j} \binom{n}{k} \left(\Delta^j a_k\right) \left(\Delta^{m+n-j-k} b_{j+k}\right) \left(\Delta^{m-j} c_{n-k}\right) \end{aligned}$$

in perfect agreement with (121).

The general version of (121) follows at once by applying L to both sides of the geometric version, where

$$L(a^i b^j c^k) = a_i b_j c_k, \tag{123}$$

and L is extended by linearity to all polynomials in a, b and c. □

Our next result is similar in spirit to Lemma 15, and it looks a good deal simpler, but I have been unable to find a purely algebraic proof. (The key step in proving Lemma 15—to look first at geometric sequences—fails miserably here!)

LEMMA 16. *If* $\mathbf{b}$ *is a totally monotonic sequence, then so is* $\mathbf{c}$*, where*

$$c_n = \frac{b_0 b_{n+2} - b_1 b_{n+1}}{n+1} \qquad (n = 0, 1, 2, \dots) \tag{124}$$

I know, in fact, of no direct proof of Lemma 16. The approach adopted below offers instead one more illustration of that old maxim:

if an assertion is too hard to prove, try making it harder.

"Harder" is interpreted here as adding one more sequence to the mix, Lemma 16 being obtained as a special case of Lemma 17 by setting $\mathbf{a} = \mathbf{b}$.

LEMMA 17. *If* $\mathbf{a}$ *and* $\mathbf{b}$ *are totally monotonic sequences, then so is* $\mathbf{c}$*, where*

$$c_n = \frac{a_0 b_{n+2} - a_1 b_{n+1} - a_{n+1} b_1 + a_{n+2} b_0}{n+1}. \tag{125}$$

Proof. Hausdorff's Theorem allows us to view $\mathbf{a}$ and $\mathbf{b}$ as moment sequences, say

$$a_n = \int_0^1 \theta^n d\alpha(\theta) \quad \text{and} \quad b_n = \int_0^1 \theta^n d\beta(\theta). \tag{126}$$

We then consider the functional,

$$L(f) = \int_0^1 \int_0^1 (y - x) \int_x^y f(t) dt\, d\alpha(x) d\beta(y), \tag{127}$$

defined on $\mathscr{C}[0,1]$, the Banach space of continuous functions on $[0,1]$.

It is clear that L is linear and continuous, so that its action is determined by a finite Borel measure, say $d\mu(\theta)$, on $[0,1]$:

$$L(f) = \int_0^1 f(\theta) d\mu(\theta). \tag{128}$$

But L is also non-negative $(f \geqslant 0 \Longrightarrow L(f) \geqslant 0)$ and the same, therefore, goes for $d\mu(\theta)$. Invoking Hausdorff's Theorem again (this time in the trivial direction) we see that the moment sequence $\boldsymbol{\mu}$ of $d\mu(\theta)$ must be totally monotonic. But

$$\begin{aligned}\mu_n &= \int_0^1 \theta^n d\mu(\theta)\\ &= \int_0^1\int_0^1 (y-x)\frac{y^{n+1}-x^{n+1}}{n+1}\,d\alpha(x)d\beta(y)\\ &= \frac{a_0b_{n+2}-a_1b_{n+1}-a_{n+1}b_1+a_{n+2}b_0}{n+1}. \qquad \Box\end{aligned}$$

Proof of Conjecture (XI). We apply Lemma 15 with

$$a_k = \frac{\mu_{k+2}-\mu_1\mu_{k+1}}{k+1}, \quad b_k = \frac{1}{k+2} \quad \text{and} \quad c_k = \mu_1^k. \tag{129}$$

(The total monotonicity of **a** is guaranteed by Lemma 16; that of **b** by the (easily checked) formula

$$\Delta^n b_k = \frac{n!(k+1)!}{(n+k+2)!}; \tag{130}$$

and that of **c** is obvious, since $0 \leqslant \mu_1 \leqslant 1$.)

Lemma 15 shows that **d** is totally monotonic, where

$$\begin{aligned}d_n &= \sum_{k=0}^{n}\binom{n}{k}\frac{\mu_{k+2}-\mu_1\mu_{k+1}}{k+1}\,\frac{(n-k)!(k+1)!}{(n+2)!}\,\mu_1^{n-k}\\ &= \sum_{k=0}^{n}\frac{\mu_1^{n-k}(\mu_{k+2}-\mu_1\mu_{k+1})}{(n+1)(n+2)}\\ &= \frac{\mu_{n+2}-\mu_1^{n+1}}{(n+1)(n+2)}\end{aligned}$$

(since the series telescopes). $\Box$

9. Some problems

PROBLEM 1. Show that a totally monotonic sequence, $(a_0, a_1, a_2, \dots)$, must be *log-convex,* i.e.,

$$a_{n+1}^2 \leqslant a_n a_{n+2} \qquad (n = 0, 1, 2, \dots). \tag{131}$$

When viewed as an assertion about moment sequences (courtesy of Hausdorff's Theorem) this is a trivial consequence of the Cauchy Schwarz inequality. The reader, however, is invited to solve this problem as stated, without recourse to Hausdorff's Theorem. This turns out to be a rewarding exercise on Inequalities, one that will surely appeal to any reader with interests in that field.

PROBLEM 2. Find other comparisons

$$H_\mu \lhd H_\nu \tag{132}$$

between Hausdorff means.

We have discussed only the most common examples (Cesàro, Euler, Gamma and Hölder) but our analysis forced us to consider also the Hausdorff means $C(\alpha,\beta)$. A partial solution to Problem 2, for this class, is given by

THEOREM 5. *Suppose that α, β, γ, δ are positive numbers. Then*

$$C(\alpha,\beta) \lhd C(\gamma,\delta) \tag{133}$$

if and only if

$$\alpha\delta = \beta\gamma \quad \textit{and} \quad \gamma \leqslant \alpha. \tag{134}$$

Proof. (Necessity.) According to Proposition 5, the first moments must coincide,

$$\frac{\alpha}{\alpha+\beta} = \frac{\gamma}{\gamma+\delta}, \tag{135}$$

and the second moments must satisfy

$$\frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} \leqslant \frac{\gamma(\gamma+1)}{(\gamma+\delta)(\gamma+\delta+1)}. \tag{136}$$

These restrictions are summarized succinctly in (134).

(Sufficiency.) Fixing n, we study the number of sign changes in the n^{th} row of the matrix

$$C(\gamma,\delta) - C(\alpha,\beta). \tag{137}$$

Our proof hinges upon the fact that there are only two possibilities: no sign changes at all, or exactly two.

We are dealing here with two Hausdorff means, whose first moments agree, so that the row sum and the iterated row sum (see (56)) both vanish. There may be no sign changes (as when $n=0$, or $n=1$, or $\alpha=\gamma$), but then the n^{th} row must be identically zero, in which case inequality (57) is automatically satisfied. Laguerre's argument (again the trivial version) shows that there cannot be exactly one sign change. Our goal is to show that there are *at most two,* and, if two sign changes are present, the signature has got to be (from $+$ to $-$ to $+$). A further application of Laguerre's argument then guarantees inequality (57), and thereby completes our proof.

At issue here is the inequality

$$\frac{\binom{k+\gamma-1}{k}\binom{n-k+\delta-1}{n-k}}{\binom{n+\gamma+\delta-1}{n}} \geqslant \frac{\binom{k+\alpha-1}{k}\binom{n-k+\beta-1}{n-k}}{\binom{n+\alpha+\beta-1}{n}}, \tag{138}$$

which we rephrase as

$$\frac{\Gamma(k+\gamma)\Gamma(n-k+\delta)}{\Gamma(k+\alpha)\Gamma(n-k+\beta)} \geqslant \frac{\Gamma(\gamma)\Gamma(\delta)\Gamma(\alpha+\beta)\Gamma(n+\gamma+\delta)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(\gamma+\delta)\Gamma(n+\alpha+\beta)}. \tag{139}$$

The right-hand side of (139) is constant, n being fixed, and we shall prove that the left-hand side, LHS, is a convex function of k $(k=0,1,\ldots,n)$.

To do this, we recall a basic property of the *trigamma function* ([1], §6.4),

$$\psi_1(x) := \frac{d^2}{dx^2}\ln\Gamma(x) \qquad (x>0), \tag{140}$$

namely,

$$\psi_1(x) = \sum_{n=0}^{\infty}\frac{1}{(x+n)^2}. \tag{141}$$

This shows that $\psi_1(x)$ is a decreasing function of x and it follows that

$$\begin{aligned}\frac{d^2}{dk^2}\ln(\text{LHS}) &= \psi_1(k+\gamma)-\psi_1(k+\alpha)+\psi_1(n-k+\delta)-\psi_1(n-k+\beta)\\ &\geqslant 0\end{aligned}$$

because $\alpha \geqslant \gamma$ and $\beta \geqslant \delta$.

The above argument confirms that LHS is a log-convex function of k; a standard application of the quotient rule

$$(\ln f)'' = \frac{ff''-(f')^2}{f^2} \tag{142}$$

then shows that LHS is convex in k. It follows that inequality (139) can suffer at most two "reversals of fortune", and if two are present, they must be (true to false to true). The same goes for the equivalent inequality (138) and these reversals correspond exactly to the signature $(+$ to $-$ to $+)$ in (137). $\square$

Theorem 4 is a special case of Theorem 5. This is seen by setting $\alpha=1$ and $\delta=1$ and by observing that

$$C(1,\beta)=C(\beta) \quad \text{and} \quad C(\gamma,1)=G(\gamma). \tag{143}$$

Theorem 5 guarantees that

$$C(1,\beta) \triangleleft C(\gamma,1) \tag{144}$$

precisely when $\gamma=\frac{1}{\beta}$ and $\gamma \leqslant 1$, in perfect agreement with (105). Comparison (106) follows similarly by taking $\beta=1$ and $\gamma=1$ in (133).

It is instructive also to look at certain limiting cases of Theorem 5. Taking $\gamma = 1$ and making $\alpha \to \infty$, we have

$$\begin{aligned}\frac{\binom{k+\alpha-1}{k}\binom{n-k+\beta-1}{n-k}}{\binom{n+\alpha+\beta-1}{n}} &\sim \frac{\frac{\alpha^k}{k!}\frac{(\alpha\delta)^{n-k}}{(n-k)!}}{\frac{(\alpha(1+\delta))^n}{n!}} \\ &= \binom{n}{k}\frac{1}{(1+\delta)^k}\left(\frac{\delta}{1+\delta}\right)^{n-k} \\ &= (n,k)^{th} \text{ entry of } \quad E\left(\frac{1}{1+\delta}\right).\end{aligned}$$

This gives an alternative confirmation of Conjecture (VII). Conjecture (IX) follows similarly by taking $\delta = 1$ and making $\beta \to \infty$ in Theorem 5.

COROLLARY 1. *If* $\mathbf{x}$ *is an arbitrary convex sequence, the expression*

$$\frac{1}{\binom{n+2\alpha-1}{n}}\sum_{k=0}^{n}\binom{k+\alpha-1}{k}\binom{n-k+\alpha-1}{n-k}x_k \tag{145}$$

decreases with α $(\alpha > 0)$.

Proof. Theorem 5 guarantees that

$$C(\alpha,\alpha) \lhd C(\gamma,\gamma) \tag{146}$$

whenever $\alpha \geqslant \gamma$. $\square$

The corollary confirms a conjecture made in [4], wherein $\mathbf{x}$ takes the special form

$$x_k = x^k y^{n-k} \qquad\qquad (k = 0,1,\dots,n). \tag{147}$$

COROLLARY 2. *If* $\mathbf{x}$ *is an arbitrary convex sequence and* $\alpha,\beta > 0$ *then*

$$\frac{1}{\binom{n+\alpha+\beta-1}{n}}\sum_{k=0}^{n}\binom{k+\alpha-1}{k}\binom{n-k+\beta-1}{n-k}x_k \leqslant \frac{\beta x_0 + \alpha x_n}{\alpha+\beta}. \tag{148}$$

Proof. Make $\gamma \to 0$ in Theorem 5. $\square$

We have had very little success in dealing with the Hólder means, $H(\alpha)$, except when $\alpha = 1,2,3,\dots$. (See sections 5 and 6.) Rephrasing Conjectures (I)—(IV) as explicit statements about total monotonicity may serve to make them more appealing. The reformulation, of course, comes courtesy of Theorem 1.

PROBLEM 3. Show that the following sequences are totally monotonic.

(I) $$\frac{\frac{1}{(n+3)^{\alpha}}-\frac{1}{\binom{n+2^{\alpha}+1}{n+2}}}{(n+1)(n+2)} \qquad \text{if } \alpha \geqslant 1;$$

(II) $$\frac{\frac{1}{\binom{n+2^{\alpha}+1}{n+2}}-\frac{1}{(n+3)^{\alpha}}}{(n+1)(n+2)} \qquad \text{if } 0<\alpha \leqslant 1;$$

(III) $$\frac{\frac{1}{(2^{\alpha}-1)(n+2)+1}-\frac{1}{(n+3)^{\alpha}}}{(n+1)(n+2)} \qquad \text{if } \alpha \geqslant 1;$$

(IV) $$\frac{\frac{1}{(n+3)^{\alpha}}-\frac{1}{(2^{\alpha}-1)(n+2)+1}}{(n+1)(n+2)} \qquad \text{if } 0<\alpha \leqslant 1.$$

Proposition 5 gives a very simple necessary condition, (36), for two Hausdorff means to be comparable, while Proposition 6 shows that said condition fails to be sufficient. It is an easy matter, however, to strengthen Proposition 5 in such a way as to vanquish Proposition 6 (or, at least, the idea behind its proof). Could the modified Proposition 5 then be sufficient to actually characterize the ordering (35)?

The modification consists of replacing the sequence (39) by

$$\mathbf{x}=(0,\ldots,0,1,2,3,\ldots), \tag{149}$$

which shows that

$$\mu_n \leqslant \nu_n \qquad (n=0,1,2,\ldots) \tag{150}$$

is a necessary condition for comparison (35) to hold. For good measure, the sequence

$$\mathbf{x}=(1,0,0,\ldots) \tag{151}$$

picks up another necessary condition

$$\Delta^n\mu_0 \leqslant \Delta^n\nu_0 \qquad (n=0,1,2,\ldots). \tag{152}$$

These ideas suggest the following

PROBLEM 4. Suppose that **a** and **b** are totally monotonic sequences satisfying

$$a_0=b_0 \quad \text{and} \quad a_1=b_1. \tag{153}$$

If

$$a_n \leqslant b_n \quad \text{and} \quad \Delta^n a_0 \leqslant \Delta^n b_0 \qquad (n=0,1,2,\ldots), \tag{154}$$

does it follow that the sequence

$$\frac{b_{n+2}-a_{n+2}}{(n+1)(n+2)} \tag{155}$$

is totally monotonic?

An affirmative solution to Problem 5 would cast everything we have said into the trivial bucket.

REFERENCES

[1] M. ABRAMOWITZ AND I. A. STEGUN, *Handbook of mathematical functions with formulas, graphs, and mathematical tables,* National Bureau of Standards Applied Mathematics Series, Vol. 55, National Bureau of Standards, Washington, D.C., 1966.

[2] N. I. AHIEZER AND M. KREIN, *Some questions in the theory of moments,* Translations of Mathematical Monographs, Vol. 2, Amer. Math. Soc., Providence, 1962.

[3] G. BENNETT, *Monotonicity-preserving matrices*, Analysis (Munich), **24** (2004), 317–327.

[4] G. BENNETT, *Hausdorff means and moment sequences*, Positivity, **15** (2011), 17–48.

[5] F. BRENTI, *Log-concave and unimodal sequences in algebra, combinatorics, and geometry: an update*, Jerusalem Combinatorics, **93**, 71–89, Contemp. Math. **178**, Amer. Math. Soc., Providence, 1994.

[6] S. DHARMADHIKARI AND J.-D. KUMAR, *Unimodality, convexity, and applications*, Probability and Mathematical Statistics, Academic Press, Boston, 1988.

[7] I. GAVREA, *Some remarks on a paper by A. McD. Mercer*, J. Inequal. Pure Appl. Math., **6**, 1 (2005), Art. 26, http://jipam.vu.edu.au/article.php?sid=495.

[8] R. L. GRAHAM, D. E. KNUTH AND O. PATASHNIK, *Concrete Mathematics,* Addison-Wesley, Reading, 1990.

[9] G. H. HARDY, *Divergent Series,* Chelsea Publ. Co., New York, 1991.

[10] G. H. HARDY, J. E. LITTLEWOOD AND G. PÓLYA, *Inequalities,* Cambridge University Press, Cambridge, 1967.

[11] S. KARLIN, *Total Positivity*, Vol. I, Stanford University Press, Stanford, 1968.

[12] S. KARLIN AND W. J. STUDDEN, *Tchebycheff systems: with applications in analysis and statistics*, Pure and Applied Mathematics, Vol. **XV**, Interscience Publishers, John Wiley and Sons, New York, 1966.

[13] E. LAGUERRE, *Oeuvres,* Chelsea Publ. Co., New York, 1973.

[14] V. I. LEVIN AND S. B. STEČKIN, *Inequalities*, Amer. Math. Soc. Transl. (2), **14** (1960), 1–29.

[15] A. MCD. MERCER, *A note on a paper by S. Haber*, Internat. J. Math. and Math. Sci., **6** (1983), 609–611.

[16] A. MCD. MERCER, *Polynomials and convex sequence inequalities*, J. Inequal. Pure Appl. Math., **6**, 1 (2005), Art. 8, http://jipam.vu.edu.au/article.php?sid=477.

[17] G. PÓLYA AND G. SZEGÖ, *Problems and Theorems in Analysis II,* Springer Verlag, New York, 1976.

[18] W. RUDIN, *Real and Complex Analysis,* McGraw-Hill, New York, 1966.

[19] T. A. SHOHAT AND J. D. TAMARKIN, *The Problem of Moments,* Amer. Math. Soc., Waverly Press, New York, 1943.

[20] D. V. WIDDER, *The Laplace Transform,* Princeton University Press, Princeton, 1941.

(Received November 20, 2009)

Grahame Bennett
Department of Mathematics, Indiana University,
Bloomington, Indiana 47405
U.S.A.
e-mail: bennettg@indiana.edu