Houston Journal of Mathematics

Volume 25, No. 4, 1999

AN INEQUALITY FOR HAUSDORFF MEANS

GRAHAME BENNETT
Communicated by Kenneth R. Davidson

Abstract. We show that skyscrapers are possible even in cities like Meanie–apolis. This lofty assertion leads to a new class of elementary inequalities, the simplest example being

$$(*) \qquad \sqrt{\frac{1}{m}\sum_{n=1}^{m}\left(\frac{1}{n}\sum_{k=1}^{n}x_k\right)^2} \le \frac{1}{m}\sum_{n=1}^{m}\sqrt{\frac{1}{n}\sum_{k=1}^{n}x_k^2} \qquad (m=1,2,\dots).$$

1. Introduction

We study inequalities of the type

$$(1.1) \qquad \left(\sum_n a_{m,n}\left(\sum_k b_{n,k}x_k\right)^p\right)^{\frac{1}{p}} \le \sum_n b_{m,n}\left(\sum_k a_{n,k}x_k^p\right)^{\frac{1}{p}} \qquad (p \ge 1),$$

wherein A, B are prescribed matrices with non-negative entries, and the estimate is to hold for all non-negative sequences $\mathbf{x}$. Such results arise naturally in the Theory of Moments [8], but they are pursued for their own sake here.

Inequality (1.1) has several striking features. It is a result that holds in considerable generality (valid, we shall see, whenever A, B are arbitrary Hausdorff means), yet it refuses, at first sight, to succumb to known ideas from the theory of ℓ^p–spaces. It turns out, instead, to be a consequence of the triangle inequality, and hence (1.1) holds for all norms. The resulting general inequality, with norm left unspecified, is not especially attractive: it becomes so only when we specialize to the ℓ^p–norm. We have here, then, a very curious phenomenon: a not–so–obvious ℓ^p–inequality, whose proof has nothing at all to do with ℓ^p.

Let us illuminate these comments by proving inequality $(*)$ in the case $m=3$. To do this, we consider six vectors in $\mathbb{R}^6$: (x,x,x,y,y,y), (x,x,y,x,y,y), plus two copies each of (x,x,x,x,x,x) and (x,x,y,y,z,z). The triangle inequality asserts that

$$\begin{aligned} &||(6x,6x,3(x+y),3(x+y),2(x+y+z),2(x+y+z)|| \\ &\qquad\qquad \leq ||(x,x,x,y,y,y)|| + ||(x,x,y,x,y,y)|| \\ &\qquad\qquad\qquad + 2||(x,x,x,x,x,x)|| + 2||(x,x,y,y,z,z)||, \end{aligned} \tag{1.2}$$

an estimate valid for all norms. Specializing to the ℓ^p-norm ($p \geq 1$) leads to

$$\left[\frac{(\frac{x}{1})^p + (\frac{x+y}{2})^p + (\frac{x+y+z}{3})^p}{3}\right]^{\frac{1}{p}} \leq \frac{(\frac{x^p}{1})^{\frac{1}{p}} + (\frac{x^p+y^p}{2})^{\frac{1}{p}} + (\frac{x^p+y^p+z^p}{3})^{\frac{1}{p}}}{3}, \tag{1.3}$$

which is $(*)$ when $p = 2$.

In order to complete the proof of $(*)$ it suffices to construct M vectors in $\mathbb{R}^M$ (M depending on m) with properties analogous to the ones above. This task leads to interesting combinatorial questions in Linear Algebra, which, we regret to say, are most easily tackled in the slick language of Urban Planning.

2. A complex real estate problem

Let us agree to view an $M \times M$ matrix with positive integer entries as a *city*. (M streets running East / West, M avenues running North / South, each lot (=entry) containing a $1-, 2-, \ldots,$ or m-story building.)

A new city, *Meanie-apolis*, is in the works, and it promises to be a large affair because of disagreements between the Developer, who likes high–rises, and the Planning Commission, which does not.

The PC-ers, recognizing that all large cities have some mean streets, have decided to make the phenomenon mandatory! So the Developer has to contend with the *mean streets conditions*:

(2.1) Any street containing s-story buildings ($s > 1$) must contain an equal number of $(s-1)$-story buildings.

(2.2) The number of *s-story streets* (those containing s-story buildings and none higher) must be independent of s ($s = 1, 2, \ldots, m$).

The PC-ers, of course, insist also that avenues and streets be treated equally, so there is a *mean city condition*:

(2.3) The city must be symmetric.

m-story cities (those containing m-story buildings and none higher) are easy to design when m is small, and even the PC-ers are able to come up with the following plans.

$$m=1,\ M=1 \qquad m=2,\ M=2 \qquad m=3,\ M=6$$

$$(1) \qquad \begin{pmatrix}1 & 1\\ 1 & 2\end{pmatrix} \qquad \begin{pmatrix}1&1&1&1&1&1\\1&1&1&1&1&1\\1&1&1&2&2&2\\1&1&2&1&2&2\\1&1&2&2&3&3\\1&1&2&2&3&3\end{pmatrix}$$

Figure 2.4. Teeny-weeny-Meanie-apolises

These plans are *minimal*, in the sense that M is as small as possible, and they are *unique* as well (up to symmetry–preserving permutations of avenues / streets). Any 2-story city, for instance, must certainly accomodate a "two–ish quarter," while the 3-story plan above is the one used in the Introduction (with 1, 2, 3 replaced by x, y, z) to discuss inequality $(*)$.

In order to prove $(*)$ in general it clearly suffices to produce a plan for an m-story city, m being arbitrarily large. This leads to the fundamental

Problem 2.1. *Are skyscrapers possible in Meanie-apolis?*

3. A Tale of Two Cities

When $m > 3$ we lose uniqueness, many plans being possible (for example) in cases $m = 4, \dots, 8$. This non–uniqueness is encouraging at first sight. If several plans are possible, our task of finding just one must surely be that much easier. But this is not the case at all (see below), and we make it our business here to reduce severely the number of available plans.

Consider the two 4-story cities in Figure 3.1.

111	111	111	111		111	111	111	111
111	111	111	111		111	111	111	111
111	111	111	111		111	111	111	111
111	111	222	222		111	112	122	222
111	111	222	222		111	121	221	222
111	111	222	222		111	211	212	222
111	222	123	333		111	122	223	333
111	222	231	333		111	221	232	333
111	222	312	333		111	212	322	333
111	222	333	444		111	222	333	444
111	222	333	444		111	222	333	444
111	222	333	444		111	222	333	444

Figure 3.1 Twin Cities: Meanie-apolis? / St. Paul?

Here $m = 4$, $M = 12$ and both plans are *minimal*. We have emphasized *subdivisions* of each of the cities by inserting spaces into their plans. The subdivisions, it will be noted, are square matrices, all the same size. In fact, they are symmetric matrices, the first street of each being constructed in *ascending order* (1-story buildings, then 2-story, etc.) with succeeding streets *filled in cyclically* (the cycles going left in order to guarantee symmetry).

These four devices (italicized) reduce greatly the number of possible plans, yet still we have non–uniqueness, as pictured above. The situation is much worse when $m = 5$, because then, even with the above devices, 60 distinct plans are available.

With so many possible choices it is difficult to see how we might proceed from stage m to $m+1$. Several of the 5-story plans have remarkable features, no doubt, but if these features are not viable when $m = 6, 7, \dots$, they are worthless. So how are we to decide which features are viable? This dilemma is present even at stage $m = 4$, when just two plans are available. Without a criterion for favoring one plan over another, we are compelled to raise

Problem 3.1. *Which of the twin cities pictured above is Meanie-apolis?*

4. The Developer's Solution

The Developer is aware of a fact, little known outside his profession, that the three most desirable features of any piece of real estate are: notation, notation, notation. So he chooses to view an m-story city as an $m \times m$ matrix, whose entries are m-dimensional probability vectors. The advantage is great, because m is generally very much smaller than M

For 4-story cities he offers two plans:

$$
\begin{array}{cccc}
1000 & 1000 & 1000 & 1000 \\
1000 & 1000 & 0100 & 0100 \\
1000 & 0100 & \frac{1}{3}\frac{1}{3}\frac{1}{3}0 & 0010 \\
1000 & 0100 & 0010 & 0001
\end{array}
\qquad\qquad
\begin{array}{cccc}
1000 & 1000 & 1000 & 1000 \\
1000 & \frac{2}{3}\frac{1}{3}00 & \frac{1}{3}\frac{2}{3}00 & 0100 \\
1000 & \frac{1}{3}\frac{2}{3}00 & 0\frac{2}{3}\frac{1}{3}0 & 0010 \\
1000 & 0100 & 0010 & 0001
\end{array}
$$

Figure 4.1: Meanie-apolis / St. Paul revisited

"Won't your cities be terribly congested?" asked the PCers suspiciously. "Not at all," replied the Developer, "You just have to understand my *building code*. Each probability vector is to be interpreted as a subdivision; $(0\ \frac{2}{3}\ \frac{1}{3}\ 0)$, for example, tells me that $\frac{2}{3}$ of the buildings in that subdivision must be 2-story, the remaining $\frac{1}{3}$, 3-story. I do not need to specify the size of my subdivisions, this being a mere landscaping problem, best left till later when I come to *bulldoze the denominators*. The plans I have given above, for instance, both involve denominators no worse than 3. Bulldozing them leads to subdivisions of size 3×3, and hence to cities of size 12×12. These correspond exactly to the twin cities of section 3."

Our Developer is smarter than others of his ilk, because, aside from the customary assortment of politicians, he has a mathematician in his pocket. He knows, therefrom, that the skyscraper Problem 2.1 has already been solved by means of the *urn matrices* introduced in [4]. These form a two-parameter family, $U^{m,n}$, with entries given by

$$u^{m,n}_{j,k} = \frac{\binom{n}{k}\binom{m-n}{j-k}}{\binom{m}{j}}. \tag{4.1}$$

The matrix $U^{m,n}$ is generated by sampling, without replacement, from an urn containing m balls, exactly n of which are good, and by setting

$$u_{j,k}^{m,n} = Pr\{k \text{ goods from } j \text{ samples}\}. \tag{4.2}$$

The urn matrices enjoy many applications ([4], [6]) owing to the fact that they are *sesqui-stochastic* [4] section 10, i.e., they have non-negative entries, their row sums are constant,

$$\sum_{k=0}^{n} u_{j,k}^{m,n} = 1 \qquad (0 \le j \le m), \tag{4.3}$$

and their column sums too,

$$\sum_{j=0}^{m} u_{j,k}^{m,n} = \frac{m+1}{n+1} \qquad (0 \le k \le n). \tag{4.4}$$

They may also be used for city planning, as we now show.

In order to design an $(m+1)$-story city $(m = 0, 1, 2, \dots)$, we have only to interpret $u_{j,k}^{m,n}$ as the fraction of the buildings in the $(n, j)^{th}$ subdivision $(0 \le j,\ n \le m)$ that are $(k+1)$-story $(0 \le k \le m)$. Our interpretation makes sense, it will be noticed, because

$$\sum_{k=0}^{m} u_{j,k}^{m,n} = 1 \qquad (0 \le j,\ n \le m). \tag{4.5}$$

The $(n, j)^{th}$ subdivision $(0 \le j \le n)$, by virtue of (4.3) and (4.5), contains only 1-story, $\dots$, $(n+1)$-story buildings. Moreover, (4.4) shows that buildings of each of these types occur in equal abundance in every one of the streets formed by the subdivisions $(n, 0), \dots, (n, m)$.

It follows that the mean streets conditions, (2.1) and (2.2), are both satisfied by our plan, $u_{j,k}^{m,n}$. The mean city condition (2.3),

$$u_{j,k}^{m,n} = u_{n,k}^{m,j} \qquad (0 \le j, k, n \le m), \tag{4.6}$$

is satisfied too, as may be seen by expanding the binomial coefficients in (4.1).

This construction shows that skyscrapers are indeed possible in Meanie-apolis, thereby answering Problem 2.1 affirmatively. Moreover, when $m = 3$, our plan reduces to the one displayed on the right-hand-side of Figure 3.1, so that Problem 3.1 is settled as well.

We have proved

Theorem 4.1. *If $p \geq 1$ and $\mathbf{x}$ is a sequence of non-negative numbers, then*

$$\left(\frac{1}{m}\sum_{n=1}^{m}\left(\frac{1}{n}\sum_{k=1}^{n} x_k\right)^p\right)^{\frac{1}{p}} \leq \frac{1}{m}\sum_{n=1}^{m}\left(\frac{1}{n}\sum_{k=1}^{n} x_k^p\right)^{\frac{1}{p}} \qquad (m = 1, 2, \ldots). \tag{4.7}$$

Inequality (4.7) reverses direction when $0 < p \leq 1$ and there is strict inequality (in either version) except when $p = 1$ or $x_1 = \ldots = x_m$ or $x_1 = \ldots = x_{m-1} = 0$. See Theorem 7.2.

5. The bureaucratic solution

The Planning Commission is not at all convinced by the foregoing arguments. Having been hoodwinked often by the Developer in the past, they want a "safe" solution this time around, certainly not one dependent upon publications as obscure as [4]. So they do what all bureaucrats are hired to do: they appoint a committee to study the problem.

Our Planning Commission, thank goodness, is smarter than others of its ilk, and they decide to appoint *many* committees, not just one. The numerator of $u_{j,k}^{m,n}$, they recognize, is the number of ways of appointing a committee of $j+1$ members from a pool of $m+1$ candidates, the $(n+1)^{th}$ oldest candidate becoming the $(k+1)^{th}$ oldest member. (Curious, is it not, how even PC-ers are able to ignore discrimination that favors the aged?)

Summing over k, $\sum_{k=0}^{m}\binom{n}{k}\binom{m-n}{j-k}$, records the committees that include a distinguished candidate (the $(n+1)^{th}$ oldest), and their number is $\binom{m}{j}$, proving (4.5). The sum over n, on the other hand, represents the number of committees having *any* (candidate as) $(k+1)^{th}$ member, so that

$$\sum_{n=0}^{m}\binom{n}{k}\binom{m-n}{j-k} = \begin{cases}\binom{m+1}{j+1} & \text{if } k \leq j\\ 0 & \text{if } k > j.\end{cases} \tag{5.1}$$

This may be rephrased as

$$\sum_{n=0}^{m} u_{j,k}^{m,n} = \frac{m+1}{j+1} \qquad (0 \leq k \leq j), \tag{5.2}$$

which is equivalent to (4.4), via (4.6).

PLAN APPROVED! concede the PC-ers.

6. Minimalism in Architecture.

Now that the plans for Meanie-apolis have been submitted (and checked), there arises the question as to whether they are as economical as possible. The issue is irrelevant for our inequalities: either of the plans,

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \qquad\qquad \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 2 & 2 \\ 1 & 1 & 2 & 2 \end{pmatrix}$$

Figure 6.1: Urban Sprawl

after all, may be used for proving $(*)$ when $m = 2$. There is, however, an obvious redundancy in the second plan, and the PCers find this to be alarming. They demand that the Developer keep his plans for Meanie-apolis as small as possible. This leads to the fundamental

Problem 6.1. *How big is Mini-apolis?*

Suppose that an m-story city has size $M \times M$. The city must have k-story streets ($k = 1, 2, \dots, m$), in accordance with (2.2), and any such street must possess 1-story, 2-story, ..., k-story buildings, all in equal abundance. A k-story street, then, must have length kd, for some positive integer d. It follows that M has got to be a multiple of k ($k = 1, 2, \dots, m$) so that

$$M \geq LCM\{1, 2, \dots, m\}. \tag{6.1}$$

On the other hand, our plan for an m-story city ($m = 1, 2, \dots$) is based upon an $m \times m$ matrix (cf. (4.1)) whose entries (indexed by n and j) are probability vectors (in k)

$$\frac{\binom{n}{k}\binom{m-1-n}{j-k}}{\binom{m-1}{j}} \qquad (0 \leq j, k, n < m). \tag{6.2}$$

It is obvious that the denominators in (6.2) may all be "bulldozed" by multiplication by $(m-1)!$ This leads to subdivisions of size $(m-1)! \times (m-1)!$, and hence to an $m! \times m!$ city plan. We have, therefore, the following estimates for M, the size of an m-story Mini-apolis:

$$LCM\{1, 2, \dots, m\} \leq M \leq m! \tag{6.3}$$

We close the gap in (6.3) by means of a "moment sequence argument" that is familiar to any student of Hausdorff matrices. If $0 \le j < m$, then

$$\frac{1}{m\binom{m-1}{j}} = \frac{\Gamma(j+1)\Gamma(m-j)}{\Gamma(m+1)}$$

$$= \int_0^1 \theta^j (1-\theta)^{m-1-j} d\theta$$

$$= \sum_{k=0}^{m-1-j} (-1)^k \binom{m-1-j}{k} \int_0^1 \theta^{j+k} d\theta$$

$$= \sum_{k=0}^{m-1-j} (-1)^k \binom{m-1-j}{k} \frac{1}{j+k+1}. \tag{6.4}$$

No denominator in (6.4) exceeds m, so that $LCM\{1,2,\dots,m\}$. $\frac{1}{m\binom{m-1}{j}}$ is a natural number. It follows that $\binom{m-1}{j}$ divides $\frac{LCM\{1,2,\dots,m\}}{m}$ $(0 \le j < m)$, and thus the denominators in (6.2) may all be bulldozed more delicately than suggested above, by multiplication by $\frac{LCM\{1,2,\dots,m\}}{m}$. This leads to a city plan of size $LCM\{1,2,\dots,m\}$, which, together with (6.3), solves Problem 6.1.

We may summarize all the results we have obtained so far in

Theorem 6.1. *Given $m = 1,2,\dots$, there exists a symmetric $M \times M$ matrix (M depending on m) with entries from $\{1,2,\dots,m\}$ such that $\frac{M}{m}$ rows contain $\frac{M}{k}$ "1's," $\frac{M}{k}$ "2's," ..., and $\frac{M}{k}$ "k's," $(k = 1,2,\dots,m)$. We may take $M = LCM\{1,2,\dots,m\}$, but no smaller value is possible.*

7. Further developments

We focus our attention now upon inequality (1.1). The Developer's notion of subdivisions, $u_{j,k}^{m,n}$, turns out to be a most profitable one, leading us to the following very general result.

Lemma 7.1. *Fix m $(m = 1,2,\dots)$ and suppose that A, B are $m \times m$ matrices with non-negative entries. Suppose, further, that there exist non-negative numbers*

$u_{j,k}^{m,n}$ $(1 \le j,k,n \le m)$ *satisfying*

$$\sum_{k=1}^{m} u_{j,k}^{m,n} = 1, \tag{7.1}$$

$$\sum_{j=1}^{m} u_{j,k}^{m,n} a_{m,j} = a_{n,k} \tag{7.2}$$

and

$$\sum_{n=1}^{m} u_{j,k}^{m,n} b_{m,n} = b_{j,k}. \tag{7.3}$$

Then

$$\left(\sum_n a_{m,n} \left(\sum_k b_{n,k} x_k \right)^p \right)^{\frac{1}{p}} \le \sum_n b_{m,n} \left(\sum_k a_{n,k} x_k^p \right)^{\frac{1}{p}} \qquad (p \ge 1) \tag{7.4}$$

whenever $x_1 \ge 0, \dots, x_m \ge 0$. *The inequality is reversed when* $0 < p \le 1$.

Proof. If $p \ge 1$, then

$$\begin{aligned}
&\left(\sum_j a_{m,j} \left(\sum_k b_{j,k} x_k \right)^p \right)^{\frac{1}{p}} \\
&= \left(\sum_j a_{m,j} \left(\sum_k \sum_n u_{j,k}^{m,n} b_{m,n} x_k \right)^p \right)^{\frac{1}{p}} && \text{by (7.3)} \\
&\le \sum_n b_{m,n} \left(\sum_j a_{m,j} \left(\sum_k u_{j,k}^{m,n} x_k \right)^p \right)^{\frac{1}{p}} && \text{by Minkowski} \\
&\le \sum_n b_{m,n} \left(\sum_j a_{m,j} \sum_k u_{j,k}^{m,n} x_k^p \right)^{\frac{1}{p}} && \text{by (7.1) and Jensen,} \\
&= \sum_n b_{m,n} \left(\sum_k a_{n,k} x_k^p \right)^{\frac{1}{p}} && \text{by (7.2)}
\end{aligned}$$

The above inequalities all reverse when $0 < p \le 1$. $\square$

An array $u^{m,n}_{j,k}$ $(1 \leq j,k,n \leq m)$ of non-negative numbers will be called an *(m-story) city plan for matrices* A, B if it satisfies the hypotheses of Lemma 7.1. Symmetry, unlike before, plays no role here, even when just one matrix is involved (i.e. $B = A$). Yet our new definition is consistent with the one previously given in as much as we may replace any city plan, say $u^{m,n}_{j,k}$, for the pair A, A, by a symmetric version

$$v^{m,n}_{j,k} = \frac{1}{2}u^{m,n}_{j,k} + \frac{1}{2}u^{m,j}_{n,k}. \tag{7.5}$$

The lemma is applicable with

$$u^{m,n}_{j,k} = \frac{\binom{n}{k}\binom{m-n}{j-k}}{\binom{m}{j}} \qquad (0 \leq j,k,n \leq m) \tag{7.6}$$

in case A and B are both replaced by the *Cesaro matrix*,

$$C = \begin{pmatrix} 1 & & & \\ \frac{1}{2} & \frac{1}{2} & & \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \\ . & . & . & . \end{pmatrix}. \tag{7.7}$$

This, of course, is the content of sections 1-5. But there is now a small bonus, because we have shown that inequality (4.7) holds for all $p > 0$ (the direction being reversed when $0 < p < 1$).

Lemma 7.1, however, applies in much greater generality. We choose here to keep the city plan (7.6) fixed, it being difficult to find an alternative, and to seek matrices A, B, other than (7.7), that satisfy the hypotheses of the lemma. Our choices of A, B are limited somewhat. If (7.1) and (7.2) hold then

$$\sum_k a_{n,k} = \sum_j \sum_k u^{m,n}_{j,k} a_{m,j} = \sum_j a_{m,j}, \tag{7.8}$$

i.e., the row sums of A have to be constant, and the same is true of B if (7.1) and (7.3) hold. Moreover, since inequality (7.4) is homogeneous in both A and B, we may as well restrict attention to matrices having row sums 1 (i.e., to *row stochastic* matrices).

A ready supply of such matrices is to be found in classical Summability Theory, and the following examples therefrom will be of special interest to us.

Hausdorff means:

$$a_{m,n} = \binom{m}{n} \int_0^1 \theta^n (1-\theta)^{m-n} d\mu(\theta) \qquad (m,n = 0,1,2,\dots), \tag{7.9}$$

where μ is any probability measure on [0,1];

Weighted means:

$$a_{m,n} = \begin{cases} \frac{a_n}{A_m} & \text{if } n \le m \\ 0 & \text{if } n > m \end{cases} \qquad (m, n = 1, 2, \dots), \tag{7.10}$$

where $\mathbf{a}$ is an arbitrary sequence of positive terms and $A_m = a_1 + \cdots a_m$;
Nørlund means:

$$a_{m,n} = \begin{cases} \frac{a_{m-n+1}}{A_m} & \text{if } n \le m \\ 0 & \text{if } n > m \end{cases} \qquad (m, n = 1, 2, \dots). \tag{7.11}$$

The matrices associated to these means are certainly *row stochastic*, as advertised above, but they share an additional property as well, being all *lower triangular*. These two properties are characteristic of the so–called *summability matrices* (see [7]). Any such matrix, A, will be said to be *positive* if $a_{m,n} > 0$ whenever $n \le m$. All weighted means and Nørlund means are positive summability matrices; all Hausdorff means are too, save for those that are convex combinations of $E(0)$ and $E(1)$ (see (7.14) below).

The remainder of this section is devoted to the Hausdorff matrices, treatment of the other means being postponed until sections 8 - 11. The entries of a Hausdorff matrix are completely determined by its diagonal:

$$A = \begin{bmatrix} \mu_0 & \cdot & \cdot & \cdot \\ \mu_0 - \mu_1 & \mu_1 & \cdot & \cdot \\ \mu_0 - 2\mu_1 + \mu_2 & 2(\mu_1 - \mu_2) & \mu_2 & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix}; \tag{7.12}$$

the diagonal terms,

$$\mu_n = \int_0^1 \theta^n d\mu(\theta) \qquad (n = 0, 1, \dots), \tag{7.13}$$

being the *moments* of the associated measure.

We show that Lemma 7.1 applies to arbitrary Hausdorff means, thereby justifying an assertion made in Section 1. The proof is easier if we begin by restricting attention to the *Euler means*, $E(\theta)$, $0 \le \theta \le 1$:

$$e_{m,n}(\theta) = \binom{m}{n} \theta^n (1-\theta)^{m-n} \qquad (m, n = 0, 1, 2, \dots). \tag{7.14}$$

These, of course, are Hausdorff means for which the associated measure is a point evaluation at some point, θ, in $[0, 1]$.

Theorem 7.2. *Let m be fixed, $m = 0, 1, 2, \dots$. If A and B are Hausdorff means, then*

$$\left(\sum_{n=0}^{m} a_{m,n}\left(\sum_{k=0}^{n} b_{n,k}x_k\right)^p\right)^{\frac{1}{p}} \le \sum_{n=0}^{m} b_{m,n}\left(\sum_{k=0}^{n} a_{n,k}x_k^p\right)^{\frac{1}{p}} \qquad (p \ge 1) \tag{7.15}$$

for every sequence $(x_0, x_1, \dots, x_m)$, of non-negative terms. The inequality is reversed when $0 < p \le 1$. There is strict inequality in (either version of) (7.15) except when at least one of the following conditions holds: (i) $p = 1$; (ii) $A = I$; (iii) $A = E(0)$; (iv) $B = I$; (v) $B = E(0)$; (vi) $x_0 = \cdots = x_m$; (vii) $x_0 = \cdots = x_{m-1} = 0$; (viii) A, B are proper convex combinations of $E(0)$, I and $x_1 = \cdots = x_{m-1} = 0$ and $x_0 = x_m$.

Proof. We apply Lemma 7.1 with $u_{j,k}^{m,n}$ given by (7.6), hypothesis (7.1) then being fulfilled automatically, courtesy of (4.3).

To show that hypothesis (7.2) is satisfied, with A an arbitrary Hausdorff mean, it suffices to restrict attention to the Euler matrices (7.14). Once the "Eulerian" version of (7.2), namely,

$$\sum_{j=0}^{m} \frac{\binom{n}{k}\binom{m-n}{j-k}}{\binom{n}{j}} e_{m,j}(\theta) = e_{n,k}(\theta), \tag{7.16}$$

has been established, the "Hausdorff" version follows by integrating over $0 \le \theta \le 1$ with respect to the measure associated to A. Identity (7.16), moreover, is easy to check. After the substitution (7.14), it reduces to

$$\sum_{j=0}^{m} \binom{m-n}{j-k} \theta^{j-k}(1-\theta)^{(m-n)-(j-k)} = 1, \tag{7.17}$$

and this is just the binomial theorem expansion of $[\theta + (1-\theta)]^{m-n}$.

The remaining hypothesis of the lemma, (7.3), is satisfied too, it being equivalent to (7.2) via the symmetry of the $u_{j,k}^{m,n}$'s, (4.6). This completes our proof of the inequality (7.15).

In order to verify the "cases of equality," we discuss first their sufficiency. It is routine to check that (7.15) reduces to an identity whenever (ii) or (iii) ... or (viii) holds. Case (i) is more subtle, the claimed identity, $A(B\mathbf{x}) = B(A\mathbf{x})$, being a consequence of the fact that any two Hausdorff matrices commute ([10], Theorem 197).

To show that these are the only cases of equality, we return to that point in the proof of Lemma 7.1 where Jensen's inequality,

$$\left(\sum_{k=0}^{m} u_k x_k\right)^p \le \sum_{k=0}^{m} u_k x_k^p, \tag{7.18}$$

was used, $\mathbf{u}$ being a probability vector. If $p \neq 1$, there is a strict inequality in (7.18) except when the x_k's are constant on the support of $\mathbf{u}$ ([11], Theorem 16).

Now Jensen's inequality is invoked $(m+1)^2$ times during the proof of Lemma 7.1, $\mathbf{u}$ being replaced by the probability vector

$$u_k = u_{j,k}^{m,n} \qquad (0 \le k \le m), \tag{7.19}$$

at the $(n,j)^{th}$ invocation $(0 \le j,n \le m)$. Strict inequality arising at *any* of these times, will manifest itself whenever both A and B are *positive* summability matrices.

Thus, if $p \neq 1$ and neither A nor B coincides with a convex combination of $E(0)$ and $E(1)$, there is a strict inequality except when

$$u_{j,k}^{m,n} \neq 0 \Longrightarrow x_k = c(j,n) \qquad (0 \le k \le m). \tag{7.20}$$

(The x_k's must be constant on the support of each probability vector, but the value assumed on each support may vary with j and $n (0 \le j,n \le m)$.) We complete the proof of this part of the theorem by showing that (7.20) forces $\mathbf{x}$ to assume one of the other forms (vi), (vii).

To do this, we fix $i (1 \le i < m)$ and observe, via (7.6), that

$$u_{j,k}^{m,n} \neq 0 \qquad (k = i-1 \text{ or } k = i) \tag{7.21}$$

when $j = n = i$. It follows from (7.20) and (7.21) that

$$x_{i-1} = x_i = c(i,i) \qquad (1 \le i < m) \tag{7.22}$$

and hence that $\mathbf{x}$ has the form

$$(x_0, \ldots, x_m) = (c, \ldots c, d). \tag{7.23}$$

Plugging such $\mathbf{x}$ into (7.15), the inequality reduces to

$$\begin{aligned} &\{(1-a)[(1-b)c+bc]^p + a[(1-b)c+bd]^p\}^{\frac{1}{p}} \\ &\qquad \le (1-b)[(1-a)c^p + ac^p]^{\frac{1}{p}} + b[(1-a)c^p + ad^p]^{\frac{1}{p}} \end{aligned} \tag{7.24}$$

where $a = a_{m,m}$ and $b = b_{m,m}$. Since $0 < a,b < 1$ and $p \neq 1$, there is strict inequality in (7.24) except when the matrix $\left(\begin{smallmatrix} c & c \\ c & d \end{smallmatrix}\right)$ has rank 1 (see [11], Theorem

26). This occurs precisely when $c = d$ or $c = 0$, in which cases $\mathbf{x}$ has the form specified by (vi) or (vii).

We have still to consider the cases when either A or B is a convex combination of $E(0)$ and $E(1)$. These are left to the reader. $\square$

8. Pedestrian thoughts

We have seen that any two Hausdorff means A, B must satisfy inequality (1.1), and that the estimate is reversed when $0 < p \leq 1$ (Theorem 7.2). The remainder of our paper is devoted to finding other pairs of matrices that have the same property. Such pairs, i.e. infinite matrices A, B satisfying (7.4) for all $m = 0, 1, 2, \ldots$, will be referred to as an *urbane couple*. A single matrix, A, will be called *urbane* if the couple A, A is.

We assume throughout that A, B have non-negative entries.

Proposition 8.1. *If A and B are diagonal matrices, then A, B form an urbane couple.*

Proposition 8.2. *A, I form an urbane couple.*

Proposition 8.3. *If A, B form an urbane couple, then so do B, A.*

Proof. We must show that

$$\left(\sum_n b_{m,n}\left(\sum_k a_{n,k}y_k\right)^q\right)^{\frac{1}{q}} \leq \sum_n a_{m,n}\left(\sum_k b_{n,k}y_k^q\right)^{\frac{1}{q}}, \tag{8.1}$$

for any non-negative sequence $\mathbf{y}$ ($q \geq 1$, reversal when $0 < q \leq 1$), whenever A, B form an urbane couple. But (8.1) reduces to (7.4) under the substitutions: $x_k = y_k^q$, $p = \frac{1}{q}$. $\square$

Proposition 8.4. *If A, B form an urbane couple, then A and B commute.*

Proof. Setting $p = 1$ in (7.4), the inequality must reduce to an identity, $A(B\mathbf{x}) = B(A\mathbf{x})$, valid for all non-negative sequences $\mathbf{x}$. $\square$

Our next three Propositions are less pedestrian. Each one presents a significant obstacle to showing that various pairs of matrices form an urbane couple.

Proposition 8.5. *If A, B form an urbane couple by virtue of satisfying the hypotheses of Lemma 7.1 with city plan (7.6), then both A and B are scalar multiples of Hausdorff means.*

Proof. We replace A by C (the Cesaro matrix, (7.7)) and we check the hypotheses of Lemma 7.1, working throughout with the city plan (7.6). Hypothesis (7.1) is certainly valid, since it holds for any city plan; (7.2) is too, having been verified in the proof of Theorem 7.2; hypothesis (7.3) is valid by assumption.

We deduce from Lemma 7.1 that C and B satisfy inequality (7.4), i.e., they form an urbane couple. Proposition 8.4 forces C and B to commute, and it follows from Theorem 2 of [12] and Theorem 198 of [10] that B must be a scalar multiple of a Hausdorff mean.

A similar argument applies to the matrix A (or we may invoke Proposition 8.3). $\square$

Proposition 8.5 is a converse to almost everything we have done thus far. It shows that no further urbane couples (beyond Hausdorff means) are to be found, so long as we refuse to consider alternatives to the plan (7.6). This is all a bit discouraging because new city plans are difficult to come by. Indeed, we have yet to design a city plan from scratch; (7.6), after all, was merely "lifted" by the Developer from [4].

Proposition 8.6. *Suppose that A and B are positive summability matrices. If $u_{j,k}^{m,n}$ $(1 \le j,k,n \le m)$ is any city plan for A, B, then*

$$u_{j,k}^{m,1} = \delta_k^1, \qquad u_{j,k}^{m,m} = \delta_k^j \tag{8.2}$$

and

$$u_{1,k}^{m,n} = \delta_k^1, \qquad u_{m,k}^{m,n} = \delta_k^n, \tag{8.3}$$

where δ_k^i is the Kronecker delta.

Proof. From (7.2) and (7.1) we have

$$\begin{aligned} 1 = \sum_{k=1}^{n} a_{n,k} &= \sum_{j=1}^{m} a_{m,j} \sum_{k=1}^{n} u_{j,k}^{m,n} \\ &\le \sum_{j=1}^{m} a_{m,j} \\ &= 1. \end{aligned}$$

The inequality must reduce to an identity, and we deduce that

$$\sum_{k=1}^{n} u_{j,k}^{m,n} = 1. \tag{8.4}$$

A few words of explanation are in order here, for the step just taken is more subtle than it might seem. From an equation of the type

$$\sum_{j=1}^{m} a_j u_j = \sum_{j=1}^{m} a_j \tag{8.5}$$

with $a_j \geq 0$, $0 \leq u_j \leq 1$, we wish to conclude that $u_j = 1$ $(j = 1, 2, \dots, m)$. Such an assertion cannot be made in general unless the a_j's are all positive. It is for this reason that we work with *positive* summability matrices: $a_{m,j} > 0(j = 1, 2, \dots, m)$: the proposition is, in fact, false without this assumption.

Comparing (8.4) with (7.1), we deduce that

$$u_{j,k}^{m,n} = 0 \qquad \text{if } k > n. \tag{8.6}$$

A similar argument, via (7.3) and (7.1), shows that

$$u_{j,k}^{m,n} = 0 \qquad \text{if } k > j. \tag{8.7}$$

The first parts of (8.2) and (8.3) follow from (8.7) and (8.6).

We turn now to the second part of (8.2), which, in view of (7.1), may be rephrased as

$$u_{k,k}^{m,m} = 1 \qquad (k = 1, 2, \dots, m). \tag{8.8}$$

This assertion will be established by induction on k, the initial step, $k = 1$, being covered already by the first part of (8.3). Let us assume, then, that $k < m$ and that

$$u_{j,j}^{m,m} = 1 \qquad (j = 1, 2, \dots, k). \tag{8.9}$$

From (7.2) and (8.6), we have

$$\begin{aligned} a_{m,i} &= \sum_{j=1}^{m} u_{j,i}^{m,m} a_{m,j} \\ &= \sum_{j=i}^{n} u_{j,i}^{m,m} a_{m,j}. \end{aligned}$$

If $i \leq k$, we deduce from (8.9) that

$$a_{m,i} = a_{m,i} + \sum_{j=i+1}^{n} u_{j,i}^{m,m} a_{m,j},$$

so that

$$u_{j,i}^{m,m} = 0 \qquad \text{if } j > i \text{ and } i \leq k. \tag{8.10}$$

Therefore

$$u^{m,m}_{k+1,i} = 0 \qquad \text{if } i \le k, \tag{8.11}$$

and this, together with (8.6), shows that

$$u^{m,m}_{k+1,k+1} = 0. \tag{8.12}$$

This completes the proof of the second part of (8.2); that of (8.3) is similar. □

Proposition 8.6 will be referred to as the *Suburban Principle*. It shows that, when dealing with positive summability matrices, *all* city plans have the same type of boundary. It is therefore only in the City Center that the architect has any scope for ingenuity.

$$\begin{array}{cccc} \underline{e}^1 & \underline{e}^1 & \cdots & \underline{e}^1 \\ & & & \\ \underline{e}^1 & \multicolumn{2}{c}{\text{DOWN}} & \underline{e}^2 \\ \vdots & \multicolumn{2}{c}{\text{TOWN}} & \vdots \\ & & & \\ \underline{e}^1 & \underline{e}^2 & \cdots & \underline{e}^m \end{array} \tag{8.13}$$

Figure 8.13: All suburbs are boring

The Suburban Principle suggests a possible restriction on urbane couples. To see this, let us agree to say that a matrix A has *decreasing diagonal* when

$$a_{k,k} \ge a_{k+1,k+1}. \tag{8.14}$$

Corollary 8.7. *Suppose that A and B are positive summability matrices. If A, B form an urbane couple by virtue of Lemma 7.1 (with any city plan), then both A and B have decreasing diagonals.*

Proof. From (7.2) and the second part of (8.3) we have, if $n \le m$,

$$\begin{aligned} a_{n,n} &= \sum_{j=1}^{m} u^{m,n}_{j,n} a_{m,j} \\ &= \sum_{j=1}^{m-1} u^{m,n}_{j,n} a_{m,j} + a_{m,m} \end{aligned}$$

It follows that

$$a_{n,n} \ge a_{m,m} \qquad (n = 1, 2, \dots, m), \tag{8.15}$$

and this is equivalent to (8.14). The proof for B is similar (or we may invoke Proposition 8.3). □

Hausdorff means and Nørlund means have decreasing diagonals, but not all weighted means do. This observation enables us to give our first example of a "bad" summability matrix, one that fails to be urbane. We take A to be the weighted mean matrix (see (7.10)) with weights $1, 2, 7, \dots,$ and we check that inequality (7.4) fails to hold when $B = A, m = 3, p = 2$ and $\mathbf{x} = (0, 1, 4)$.

A restriction that applies to Nørlund means is given by our next proposition. We shall say that a matrix B has *decreasing partial row sums* if

$$\sum_{k=1}^{r} b_{n,k} \geq \sum_{k=1}^{r} b_{n+1,k} \qquad (n, r = 1, 2, \dots). \tag{8.16}$$

Proposition 8.8. *Suppose that A and B are positive summability matrices and that A, B form an urbane couple (with, or without, a city plan). Then both matrices have decreasing partial row sums.*

Proof. It suffices to check that (8.16) holds, the corresponding statement for A being then a consequence of Proposition 8.3.

Making $p \to \infty$ in inequality (7.4), and using the fact that A is a positive summability matrix, we deduce from Theorem 4 of [11] that

$$\max_{1 \leq n \leq m} \sum_{k=1}^{n} b_{n,k} x_k \leq \sum_{n=1}^{m} b_{m,n} \big(\max_{1 \leq k \leq n} x_k \big) \tag{8.17}$$

for every non-negative sequence $\mathbf{x}$. Taking $\mathbf{x}$ to have the form

$$\mathbf{x} = (0, \dots, 0, 1, \dots, 1), \tag{8.18}$$

(8.17) reduces to

$$\max_{1 \leq n \leq m} \sum_{k=r}^{n} b_{n,k} \leq \sum_{n=r}^{m} b_{m,n}, \tag{8.19}$$

which is the same as (8.16), since B is a summability matrix. (It may be shown, by the so-called *Partial Sums Lemma* (Lemma 2.1 of [7]) that (8.17) and (8.16) are equivalent, so that nothing has been lost in making the specialization (8.18).) $\square$

The "decreasing partial row sums" condition arises in a different context in Lemma 9 of [3], where it is seen to hold for all Hausdorff means. Proposition 8.8, in conjunction with Theorem 7.2, provides a new proof of this fact. Weighted means satisfy (8.16) as well, but not all Nørlund means do. It is curious that the restrictions obtained on these two different classes of matrices, from Corollary 8.21

and Proposition 8.8, are the same, namely that

$$\frac{A_n}{A_{n+1}} \qquad \text{increase with } n. \tag{8.20}$$

Our next result is an analogue of Proposition 8.8 in which the exponent of inequality (7.4) is forced towards its other extreme, $p \to 0^+$. The ensuing estimate is not so easy to decipher as was (8.17), yet it must hold for arbitrary Hausdorff means, A, B, in view of Theorem 7.2.

Proposition 8.9. *If A, B form an urbane couple and A is a summability matrix, then*

$$\prod_{n=1}^{m}\left(\sum_{k=1}^{n} b_{n,k}x_k\right)^{a_{m,n}} \geq \sum_{n=1}^{m} b_{m,n}\prod_{k=1}^{n} x_k^{a_{n,k}} \tag{8.21}$$

for every non-negative sequence $\mathbf{x}$.

We have seen that there are non-urbane weighted means and Nørlund means, but it is much more interesting to seek out the urbane ones. Examples aplenty are to be found among the Hausdorff means.

The Gamma matrices, $\Gamma(\alpha)$, $\alpha > 0$. These are weighted/Hausdorff means generated by the measure $d\mu(\theta) = \alpha\theta^{\alpha-1}d\theta$ with weights

$$a_n = \binom{n+\alpha-1}{n} \tag{8.22}$$

and moments

$$\mu_n = \frac{\alpha}{n+\alpha} \qquad (n = 0, 1, 2, \dots). \tag{8.23}$$

The Cesaro matrices, $C(\alpha)$, $\alpha > 0$. There are Nørlund/Hausdorff means generated by the measure $d\mu(\theta) = \alpha(1-\theta)^{\alpha-1}d\theta$ with weights

$$a_n = \binom{n+\alpha-1}{n} \tag{8.24}$$

and moments

$$\mu_n = \frac{1}{\binom{n+\alpha}{n}} \qquad (n = 0, 1, 2, \dots). \tag{8.25}$$

Agnew [1] has shown that there are no other examples of weighted/Hausdorff or of Nørlund/Hausdorff means. This observation suggests the following

Problem 8.1. *(i) Are the Gamma matrices the only urbane weighted means? (ii) Are the Cesaro matrices the only urbane Nørlund means?*

9. Rural diversions

While tinkering with Problem 8.1 it is natural that we should direct our attention to the matrix

$$D = \begin{pmatrix} 1 & & & & \\ \frac{1}{3} & \frac{2}{3} & & & \\ \frac{1}{7} & \frac{2}{7} & \frac{4}{7} & & \\ \frac{1}{15} & \frac{2}{15} & \frac{4}{15} & \frac{8}{15} & \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} \tag{9.1}$$

Why? Because it is a Nørlund mean (with weights 2^{1-n}) as well as a weighted mean (with weights 2^{n-1}), yet D is not Hausdorff, so it is an ideal candidate for solving both parts of Problem 8.1 simultaneously. We shall call D the *doubling matrix*.

In attempting to show that D is urbane, we are forced to go back to basics (i.e. to city planning). When $m = 3$ the required inequality is

$$\left[\frac{(\frac{x}{1})^p + 2(\frac{x+2y}{3})^p + 4(\frac{x+2y+4z}{7})^p}{7}\right]^{\frac{1}{p}} \tag{9.2}$$
$$\leq \frac{(\frac{x^p}{1})^{\frac{1}{p}} + 2(\frac{x^p+2y^p}{3})^{\frac{1}{p}} + 4(\frac{x^p+2y^p+4z^p}{7})^{\frac{1}{p}}}{7}$$

($p \geq 1$, with reversal when $0 < p \leq 1$). This is proved by means of the following plan.

$$\begin{array}{c|ccc|ccc|ccc} & & \mathbf{1} & & & \mathbf{2} & & & \mathbf{4} & \\ \mathbf{1} & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ \mathbf{2} & 1 & 0 & 0 & \frac{2}{3} & \frac{1}{3} & 0 & 0 & 1 & 0 \\ \mathbf{4} & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}$$

Figure 9.3. A congested city.

The probability vectors are again interpreted as subdivisions, so that each, after bulldozing denominators, is of size 3×3. What is new here are the boldface numerals, indicating *repetitions* of the subdivisions. Thus the above plan represents a three-story city of size 21×21. There are 3 one-story streets, twice as many two-story streets (each with 7 one-story and 14 two-story buildings), and four times as many three-story streets (each with 3 one-story, 6 two-story and

12 three-story buildings). The triangle inequality gives (9.2) in the same way that (1.2) led to (1.3).

Unfortunately, it is difficult to proceed much further because we have nothing like the urn matrices to fall back upon here. In view of Proposition 8.5 a new idea is needed. This comes about by invoking a general technique, which we call the *Haystack Principle.*

It is easier to find a needle in a haystack when just one is hidden there, not many.

The principle is worth trying whenever we are confronted by a *search problem.* Whatever makes the search challenging is immaterial, because the idea is to impose additional constraints, making our task more difficult. The constraints may be too stringent, leaving no needles to be found, or too slack, leaving yet many, or they may be chosen just right. In the last case, with just one needle present, there may be a "uniqueness argument" that provides a completely new approach to our problem. The principle, far-fetched though it may seem, was used very effectively in [5], and it provides useful information here as well. The main difficulty, of course, lies in choosing the appropriate constraints.

We are here concerned (c.f. Lemma 7.1) with finding, for each $m = 0, 1, 2, \ldots,$ an array $u^{m,n}_{j,k}$ of *non-negative* numbers, with the properties

$$\sum_k u^{m,n}_{j,k} = 1, \tag{9.3}$$

$$u^{m,n}_{j,k} = u^{m,j}_{n,k} \tag{9.4}$$

and

$$\sum_j u^{m,n}_{j,k} d_{m,j} = d_{n,k}, \tag{9.5}$$

where $D = (d_{n,k})$ is the doubling matrix (9.1).

There are several solutions in each of the cases $m = 4, 5, 6$, but they are complicated, and it is not at all clear how we might proceed to higher cases. It is the Haystack Principle which comes to our rescue, courtesy of the additional constraint,

$$u^{m,m-n}_{m-j,k} = u^{m,n}_{j,n+j+k-m} \qquad (m \le n + j + k). \tag{9.6}$$

This constraint was suggested by a careful study of the urn matrices (4.1), it being clear that the u's defined by (7.6) satisfy (9.6). Our u's do too, *by fiat*, and, of course, they satisfy (9.3), (9.4), and (9.5) as well. It transpires that these four equations suffice to determine our u's completely, and this means we can compute city plans for D at will.

When $m = 5$, for instance, we obtain the following plan, in which we have chosen not to display the suburbs (c.f. (8.13)).

$$\begin{matrix} \frac{14}{15} & \frac{1}{15} & 0 & 0 & 0 \\ \frac{4}{5} & \frac{1}{5} & 0 & 0 & 0 \\ \frac{8}{15} & \frac{7}{15} & 0 & 0 & 0 \end{matrix} \qquad \begin{matrix} \frac{4}{5} & \frac{1}{5} & 0 & 0 & 0 \\ \frac{16}{35} & \frac{18}{35} & \frac{1}{35} & 0 & 0 \\ 0 & \frac{4}{5} & \frac{1}{5} & 0 & 0 \end{matrix} \qquad \begin{matrix} \frac{8}{15} & \frac{7}{15} & 0 & 0 & 0 \\ 0 & \frac{4}{5} & \frac{1}{5} & 0 & 0 \\ 0 & 0 & \frac{14}{15} & \frac{1}{15} & 0 \end{matrix}$$

Figure 9.8. An inner city.

More generally, the Haystack Principle, via (9.6), provides an m-story plan for D of size $M \times M$ where

m	1	2	3	4	5	6	7	8
M	1	3	21	105	3,255	9,765	1,240,155	21,082,635

All that is missing from our proof that D is urbane is the non-negativity of the u's. This is readily seen by inspection when m is small, as above, but it cannot be guaranteed generally until we have an explicit formula for the $u_{j,k}^{m,n}$'s.

10. Urban commuters

The results of Section 9 provide strong evidence that D is indeed an urbane matrix, but they fall short of proving this fact. Our next step is the decisive one: we look at the *commutant* of D.

It turns out that a matrix B commutes with D if and only if it has the form

$$B = \begin{bmatrix} \mu_0 & & & \\ \mu_0 - \mu_1 & \mu_1 & & \\ \mu_0 - 3\mu_1 + 2\mu_2 & 3(\mu_1 - \mu_2) & \mu_2 & \\ \mu_0 - 7\mu_1 + 14\mu_2 - 8\mu_3 & 7(\mu_1 - 3\mu_2 + 2\mu_3) & 7(\mu_2 - \mu_3) & \mu_3 \\ . & . & . & . \end{bmatrix} \tag{10.1}$$

In other words, B has got to be a lower triangular matrix whose entries are determined completely by its diagonal in accordance with (10.1).

The resemblance between (10.1) and the structure of Hausdorff matrices, (7.9), is as irresistible as it is surprising. We have here, after all, differences of some kind displayed *consistently* down the sub-diagonals, and multiples of some kind,

$$\begin{array}{llll} 1 & & & \\ 1 & 1 & & \\ 1 & 3 & 1 & \\ 1 & 7 & 7 & 1\,, \end{array} \tag{10.2}$$

displayed out front. The multiples, in fact, are recognizable to anyone who has followed the *Theory of Partitions*: they are the 2-binomial coefficients. There is thus the suggestion, and it is a powerful one, that we ought to be working with q-binomial coefficients ($q > 0$, not just $q = 2$), and that we ought to investigate the possibility of developing a theory of q-Hausdorff matrices.

We initiate such a theory by defining the *q-difference operator*, Δ_q $(q > 0)$, on the space of numerical sequences, by setting

$$\Delta_q \boldsymbol{\mu} = (\mu_0 - \mu_1, q(\mu_1 - \mu_2), q^2(\mu_2 - \mu_3), \dots), \tag{10.3}$$

where $\boldsymbol{\mu} = (\mu_0, \mu_1, \mu_2, \dots)$. This operator seems not to have been considered before, which is curious, because it leads directly to the *q–binomial coefficients* via iteration, $\Delta_q^n = \Delta_q(\Delta_q^{n-1})$:

$$(\Delta_q^n \boldsymbol{\mu})_j = q^{nj} \sum_{k=0}^{n} (-1)^k q^{\binom{k}{2}} \begin{bmatrix} n \\ k \end{bmatrix}_q \mu_{k+j}. \tag{10.4}$$

Formula (10.4) may be used to define the q-binomial coefficients; their basic properties (e.g. (10.6)–(10.9) below) then all follow quickly from this definition.

We prefer, however, to derive formula (10.4) from the existing theory, and so we back up a bit and make a few general comments about *q-Mathematics*. The entire subject is based upon the fiction that any real number r may be written as

$$r = \frac{q^r - 1}{q - 1} \qquad (q \neq 1). \tag{10.5}$$

Fiction becomes fact, of course, in the limit as $q \to 1$, so that identities in q-Mathematics, however strange they may seem, are “checkable” by making $q \to 1$. Equation (10.4), it will be noticed, reduces in this fashion to the correct formula for the familiar n^{th} *difference operator*, Δ^n.

The *binomial coefficients* in q-Mathematics are defined by

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \begin{cases} \frac{n!_q}{k!_q (n-k)!_q} & \text{if k=0,1, \dots , n} \\ 0 & \text{otherwise,} \end{cases} \tag{10.6}$$

factorials being given by

$$n!_q = \begin{cases} \frac{q-1}{q-1}\frac{q^2-1}{q-1}\cdots\frac{q^n-1}{q-1} & \text{if n = 1, 2, } \dots \\ 1 & \text{if n=0} \end{cases} \tag{10.7}$$

We have

$$\begin{bmatrix} n \\ 0 \end{bmatrix}_q = 1, \qquad \begin{bmatrix} n \\ k \end{bmatrix}_q = \begin{bmatrix} n \\ n-k \end{bmatrix}_q \tag{10.8}$$

and

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = \begin{bmatrix} n-1 \\ k \end{bmatrix}_q + q^{n-k}\begin{bmatrix} n-1 \\ k-1 \end{bmatrix}_q \quad . \tag{10.9}$$

It is worthwhile to note that the associated *Pascal triangle* agrees, when $q = 2$, with the array (10.2) that we encountered when discussing the commutant of the doubling matrix.

There is also a *q-Binomial Theorem*:

$$(a+b)_q^n = \sum_{j=0}^{n} q^{\binom{j}{2}} \begin{bmatrix} n \\ j \end{bmatrix}_q a^{n-j}b^j, \tag{10.10}$$

wherein

$$(a+b)_q^n = \begin{cases} (a+b)(a+bq)\dots(a+bq^{n-1}) & \text{if n=1,2, } \dots \\ 1 & \text{if n=0.} \end{cases} \tag{10.11}$$

This, just like the classical theorem, may be proved via induction on n, using recursion (10.9) and the identity

$$(a+b)_q^{n+1} = (a+bq^n)(a+b)_q^n. \tag{10.12}$$

In order to prove formula (10.4) we invoke a reduction argument called the *Method of Moriarty* (see [4], section 5). The idea, a commonplace in the Theory of Moments, is this: to prove an algebraic identity, such as (10.4), that is linear in $\boldsymbol{\mu}$, it suffices to do so just when $\boldsymbol{\mu}$ is a *geometric sequence*

$$\mu_n = r^n \qquad (n = 0, 1, 2, \dots). \tag{10.13}$$

The reduction is especially valuable when dealing with difference operators, for $\Delta_q^n\boldsymbol{\mu}$ then takes on a very simple form. Indeed, with $\boldsymbol{\mu}$ as in (10.13), we have

$$\Delta_q\boldsymbol{\mu} = (1-r)(1, rq, (rq)^2, \dots), \tag{10.14}$$

which is again geometric, so that

$$(\Delta_q^n\boldsymbol{\mu})_k = (1-r)_q^n(rq^n)^k. \tag{10.15}$$

Thus formula (10.4), in case $\boldsymbol{\mu}$ is of the form (10.13), reduces to

$$(1-r)_q^n = \sum_{j=0}^{n} q^{\binom{j}{2}} \begin{bmatrix} n \\ j \end{bmatrix}_q (-r)^j, \tag{10.16}$$

and this is a special case of (10.11).

Central to the theory of (ordinary) Hausdorff means is a certain matrix δ (see [10], Chapter 11). We define here an analogue, δ_q, via its action on numerical sequences

$$(\delta_q \mathbf{x})_n = (\Delta_q^n \mathbf{x})_0 \qquad (n = 0, 1, 2, \dots), \tag{10.17}$$

so that its entries are given by

$$(\delta_q)_{n,k} = (-1)^k q^{\binom{k}{2}} \begin{bmatrix} n \\ k \end{bmatrix}_q . \tag{10.18}$$

δ_q, being lower triangular with non-zero diagonal entries, is obviously invertible.

We say that a matrix A is *q-Hausdorff* if it admits a factorization of the form

$$A = \delta_q \{\text{diag } \boldsymbol{\mu}\} \delta_q^{-1}, \tag{10.19}$$

for some sequence $\boldsymbol{\mu} = (\mu_0, \mu_1, \dots)$. It is clear that any two q-Hausdorff matrices commute. The relevance of the definition for us lies in the fact that the doubling matrix, D, is 2-Hausdorff. This is an obvious consequence of the following result, D arising therein when $q = 2$.

Proposition 10.1. *If $q > 0$, the weighted mean / Nørlund mean matrix A given by*

$$a_{m,n} = \begin{cases} \frac{q^n}{1+q+\cdots+q^m} & \text{if } n \le m \\ & \text{if } n > m \end{cases} \tag{10.20}$$

($m, n = 0, 1, 2, \dots$) is q-Hausdorff.

Proof. We show that

$$A\delta_q = \delta_q \{\text{diag } \boldsymbol{\mu}\}, \tag{10.21}$$

where

$$\mu_n = \frac{q^n}{1+q+\cdots+q^n} \qquad (n = 0, 1, 2, \dots), \tag{10.22}$$

which assertion obviously implies (10.19).

Equation (10.21),

$$\sum_{n=k}^{m} \frac{q^n}{1+q+\cdots+q^m}(-1)^k q^{\binom{k}{2}} \begin{bmatrix} n \\ k \end{bmatrix}_q = (-1)^k q^{\binom{k}{2}} \begin{bmatrix} m \\ k \end{bmatrix}_q \frac{q^k}{1+q+\cdots+q^k}, \tag{10.23}$$

reduces to

$$\sum_{n=k}^{m} q^n \begin{bmatrix} n \\ k \end{bmatrix}_q = q^k \begin{bmatrix} m+1 \\ k+1 \end{bmatrix}_q, \tag{10.24}$$

because

$$\begin{bmatrix} m+1 \\ k+1 \end{bmatrix}_q = \frac{1+q+\cdots+q^m}{1+q+\cdots+q^k} \begin{bmatrix} m \\ k \end{bmatrix}_q. \tag{10.25}$$

Identity (10.25) is an immediate consequence of the definition, (10.6), of the q-binomial coefficients. Identity (10.24) follows from (10.9) by repeated application of the recursion, expanding always the first binomial coefficient on the right-side of (10.9). □

Our next result leads to a complete description of the commutant of the doubling matrix D. (See also Proposition 12.1).

Proposition 10.2. *Fix $q > 0$. Suppose that A is a q-Hausdorff matrix with distinct diagonal entries and that B is any lower triangular matrix that commutes with A. Then B is q-Hausdorff.*

Proof. The argument, Theorem 198 of [10], used to prove the classical result ($q = 1$) applies here as well. □

Proposition 10.3. *Fix $q > 0$. If A is the weighted mean matrix given by (10.20), the commutant of A is the set of q-Hausdorff matrices.*

Proof. If B is any matrix that commutes with A, then B must be lower triangular (see [12]) and it follows from Proposition 10.2 that B has got to be q-Hausdorff.

The converse statement is a consequence of Proposition 10.1 since any two q-Hausdorff matrices must commute. □

11. Doublin's Fair City

We are now able to provide a city plan for the doubling matrix (9.1). Our design, in fact, covers all the matrices of Proposition 10.1 (i.e. weighted means having *geometric weights*).

Theorem 11.1. *If $q > 0$, $p \geq 1$ and $m = 0, 1, 2, \ldots,$ then*

$$\left(\frac{q-1}{q^{m+1}-1}\sum_{n=0}^{m} q^n \left(\frac{q-1}{q^{n+1}-1}\sum_{k=0}^{n} q^k x_k\right)^p\right)^{\frac{1}{p}} \leq \frac{q-1}{q^{m+1}-1}\sum_{n=0}^{m} q^n \left(\frac{q-1}{q^{n+1}-1}\sum_{k=0}^{n} q^k x_k^p\right)^{\frac{1}{p}} \tag{11.1}$$

for every non-negative sequence $\mathbf{x}$. *The inequality reverses direction when* $0 < p \leq 1$.

Proof. Fixing $q > 0$, we take $B = A$ in Lemma 7.1, A being the weighted mean (10.20). For city plan $u_{j,k}^{m,n}$ $(0 \leq j, k, \leq m)$ we set

$$u_{j,k}^{m,n} = q^{(j-k)(n-k)} \frac{\begin{bmatrix} n \\ k \end{bmatrix}_q \begin{bmatrix} m-n \\ j-k \end{bmatrix}_q}{\begin{bmatrix} m \\ j \end{bmatrix}_q} \tag{11.2}$$

Our plan is obviously *symmetric*,

$$u_{j,k}^{m,n} = u_{n,k}^{m,j}, \tag{11.3}$$

as may be seen by expanding the q-binomial coefficients of (11.3) in accordance with definition (10.6). It is clear, too, that $u_{j,k}^{m,n} \geq 0$, as is required by Lemma 7.1.

Hypothesis (7.1) of the lemma is a well-known formula in q-Mathematics,

$$\sum_k q^{(j-k)(n-k)} \begin{bmatrix} n \\ k \end{bmatrix}_q \begin{bmatrix} m-n \\ j-k \end{bmatrix}_q = \begin{bmatrix} m \\ j \end{bmatrix}_q, \tag{11.4}$$

the so-called *q-Vandermonde convolution* [2], [9]. Hypotheses (7.2) and (7.3) are equivalent by symmetry, (11.3), so that only (7.3) needs to be checked here. The required identity is

$$\sum_n u_{j,k}^{m,n} \frac{q^n}{1+q+\cdots+q^m} = \frac{q^k}{1+q+\cdots+q^j}, \tag{11.5}$$

and this reduces to

$$\sum_n q^{(j-k+1)(n-k)} \begin{bmatrix} n \\ k \end{bmatrix}_q \begin{bmatrix} m-n \\ j-k \end{bmatrix}_q = \begin{bmatrix} m+1 \\ j+1 \end{bmatrix}_q, \tag{11.6}$$

courtesy of (10.25).

In order to prove (11.6) we use the observation,

$$\begin{bmatrix} r \\ k \end{bmatrix}_q = (-1)^k q^{rk-\binom{k}{2}} \begin{bmatrix} k-r-1 \\ k \end{bmatrix}_q, \tag{11.7}$$

which follows at once from the definition (10.6). This simple device enables us to re-locate the parameters of a q-binomial coefficient. It is used below to transform the convolution (11.7), with summation index "upstairs," to the familiar Vandermonde, (11.4), with index "downstairs."

We have

$$\sum_n q^{(j-k+1)(n-k)} \begin{bmatrix} n \\ k \end{bmatrix}_q \begin{bmatrix} m-n \\ j-k \end{bmatrix}_q$$

$$= \qquad [\ (10.8)\ \text{twice}\ ]$$

$$\sum_n q^{(j-k+1)(n-k)} \begin{bmatrix} n \\ n-k \end{bmatrix}_q \begin{bmatrix} m-n \\ m-n-j+k \end{bmatrix}_q$$

$$= \qquad [\ (11.7)\ \text{twice}\ ]$$

$$(-1)^{m+j} q^{(m+1)(m-j)-\binom{m-j}{2}} \times \sum_n q^{-(m-n-j+k)(n+1)} \begin{bmatrix} -k-1 \\ n-k \end{bmatrix}_q \begin{bmatrix} k-j-1 \\ m-n-j+k \end{bmatrix}_q$$

$$= \qquad [\ (11.4)]$$

$$(-1)^{m+j} q^{(m+1)(m-j)-\binom{m-j}{2}} \begin{bmatrix} -j-2 \\ m-j \end{bmatrix}_q$$

$$= \qquad [\ (11.7)]$$

$$\begin{bmatrix} m+1 \\ m-j \end{bmatrix}_q$$

$$= \qquad [\ (10.8)]$$

$$\begin{bmatrix} m+1 \\ j+1 \end{bmatrix}_q.$$

Inequality (11.1) thus follows from Lemma 7.1. $\square$

When $q = 1$, Theorem 11.1 is to be interpreted via our convention (10.5), i.e., by taking limits as $q \to 1$. This special case, it will be noticed, coincides with Theorem 4.1.

Theorem 11.1 solves both parts of Problem 8.1 because the matrix (10.20) is a Nørlund mean with weights $(q^{-n}, n = 0, 1, 2, \ldots)$ as well as a weighted mean.

12. q-Hausdorff means

We close with a few remarks on q-Hausdorff matrices. Our first result shows how the entries of such a matrix are determined by its diagonal $\boldsymbol{\mu}$.

Proposition 12.1. *Fix $q > 0$. A matrix, A, is q-Hausdorff if and only if it has entries of the form*

$$a_{m,k} = q^{-k(m-k)} \begin{bmatrix} m \\ k \end{bmatrix}_q \Delta_q^{m-k} \mu_k \qquad (m, k = 0, 1, 2, \ldots) \tag{12.1}$$

for some sequence $\boldsymbol{\mu} = (\mu_0, \mu_1, \ldots)$.

Proof. We shall use the identities

$$\begin{bmatrix} n \\ k \end{bmatrix}_q = q^{k(n-k)} \begin{bmatrix} n \\ k \end{bmatrix}_{\frac{1}{q}} \tag{12.2}$$

and

$$\begin{bmatrix} m \\ n \end{bmatrix}_q \begin{bmatrix} n \\ k \end{bmatrix}_q = \begin{bmatrix} m \\ k \end{bmatrix}_q \begin{bmatrix} m-k \\ n-k \end{bmatrix}_q, \tag{12.3}$$

both of which follow from definition (10.6).

The $(m, k)^{th}$ entry of the matrix $\delta_q\{\text{diag}\,\boldsymbol{\mu}\}\delta_{\frac{1}{q}}$ is

$$\begin{aligned}
&\sum_{n=k}^{m} (-1)^n q^{\binom{n}{2}} \begin{bmatrix} m \\ n \end{bmatrix}_q \mu_n (-1)^k q^{-\binom{k}{2}} \begin{bmatrix} n \\ k \end{bmatrix}_{\frac{1}{q}} \\
&= \sum_{n=k}^{m} (-1)^{n+k} q^{\binom{n}{2} - \binom{k}{2} - k(n-k)} \begin{bmatrix} m \\ n \end{bmatrix}_q \begin{bmatrix} n \\ k \end{bmatrix}_q \mu_n \\
&= \begin{bmatrix} m \\ k \end{bmatrix}_q \sum_{n=k}^{m} (-1)^{n+k} q^{\binom{n}{2} - \binom{k}{2} - k(n-k)} \begin{bmatrix} m-k \\ n-k \end{bmatrix}_q \mu_n \\
&= \begin{bmatrix} m \\ k \end{bmatrix}_q \sum_{j=0}^{m-k} (-1)^j q^{\binom{j}{2}} \begin{bmatrix} m-k \\ j \end{bmatrix}_q \mu_{k+j}.
\end{aligned}$$

Thus, by (10.4),

$$(\delta_q\{\text{diag }\boldsymbol{\mu}\}\delta_{\frac{1}{q}})_{m,k} = q^{-k(m-k)} \begin{bmatrix} m \\ k \end{bmatrix}_q \Delta_q^{m-k}\mu_k. \tag{12.4}$$

Taking $\boldsymbol{\mu} = (1, 1, \dots)$, it is clear from definition (10.3) that $\Delta_q^n \boldsymbol{\mu} = (1, 0, 0, \dots)$ whenever $n = 1, 2, \dots$, so that

$$(\Delta_q^{m-k}\boldsymbol{\mu})_k = \begin{cases} 1 & \text{if } m = k \\ 0 & \text{otherwise.} \end{cases} \tag{12.5}$$

In this case, (12.4) reduces to

$$\delta_q \delta_{\frac{1}{q}} = I, \tag{12.6}$$

so that

$$\delta_q^{-1} = \delta_{\frac{1}{q}}. \tag{12.7}$$

(12.1) is now seen to be the same as (10.19) in view of (12.4) and (12.6). □

We remark that formula (12.1) agrees, when $q = 2$, with the explicit description of the commutant of D given in (10.1).

Corollary 12.2. *A q-Hausdorff matrix ($q > 0$) has constant row sums.*

Proof. According to (10.19) and (12.6), any such matrix has the form

$$\delta_q\{\text{diag }\boldsymbol{\mu}\}\delta_{\frac{1}{q}} \tag{12.8}$$

From definition (10.3) we see that

$$\begin{aligned} &\delta_q\{\text{diag }\boldsymbol{\mu}\}\delta_{\frac{1}{q}}(1, 1, \dots)^T \\ &= \delta_q\{\text{diag }\boldsymbol{\mu}\}(1, 0, 0, \dots)^T \\ &= \delta_q(\mu_0, 0, 0, \dots)^T \\ &= (\mu_0, \mu_0, \dots)^T, \end{aligned}$$

so that the sum of every row is μ_0. □

Henceforth we shall take μ_0 to be 1. Some examples of q-Hausdorff matrices are listed below.

q-Euler matrices, $E_q(\theta)$, $0 \le \theta \le 1$. These are given by

$$E_q(\theta)_{n,k} = \begin{bmatrix} n \\ k \end{bmatrix}_q \theta^k (1-\theta)_q^{n-k} \quad (n, k = 0, 1, 2, \dots), \tag{12.9}$$

the definition being based directly upon the classical one, (7.14). It is easy to check that (12.9) is indeed a q-Hausdorff matrix, in fact, the one generated by the geometric sequence

$$\boldsymbol{\mu} = (1, \theta, \theta^2, \dots). \tag{12.10}$$

We shall refer to $\boldsymbol{\mu}$ as the *q-moment sequence* associated with $E_q(\theta)$, even though we are unable to identify a probability measure on $[0, 1]$ that produces these q-moments. The problem is that we must now perform our integrations q-wise, and the existing theory seems not to be up to our task.

The *q-integral* ($q > 1$) of a continuous function on $[0, 1]$ is defined by

$$\int_0^1 f(\theta) d_q(\theta) = \frac{q-1}{q} \sum_{j=0}^{\infty} f(q^{-j}) q^{-j} \tag{12.11}$$

and there is a similar formula, with q replaced by $\frac{1}{q}$, when $0 < q < 1$. The right side of (12.11) is the Riemann sum of f evaluated at the right-hand endpoints of the partition $\dots \frac{1}{q^2} < \frac{1}{q} < 1$, of $[0, 1]$. Thus we obtain the correct result,

$$\int_0^1 f(\theta) d_q(\theta) \to \int_0^1 f(\theta) d\theta, \tag{12.12}$$

as $q \to 1^+$. The q-integral was first used by Fermat in his ingenious proof of the fact that

$$\int_0^1 \theta^n d\theta = \frac{1}{n+1} \qquad (n = 0, 1, 2, \dots). \tag{12.13}$$

The q-Cesaro matrix, C_q. The ordinary Cesaro matrix, C, arises from the formula (7.9) when $d\mu(\theta)$ is the uniform probability distribution on $[0, 1]$:

$$\binom{m}{n} \int_0^1 \theta^n (1-\theta)^{m-n} d\theta = \begin{cases} \frac{1}{m+1} & \text{if } n \le m \\ 0 & ifn > m \end{cases} \tag{12.14}$$

$(m, n = 0, 1, 2, \dots)$. The q-analogue of this formula ($q > 1$) is

$$\begin{bmatrix} m \\ n \end{bmatrix}_q \int_0^1 \theta^n (1-\theta)_q^{m-n} d_q(\theta) = \begin{cases} \frac{q^n}{1+q+\cdots+q^m} & \text{if } n \le m \\ 0 & \text{if } n > m \end{cases} \tag{12.15}$$

$(m, n = 0, 1, 2, \dots)$. We omit the proof. This shows, in particular, that the doubling matrix, D, is none other than the 2-Cesaro matrix.

When $0 < q < 1$, however, things go awry.

$$\begin{bmatrix} m \\ n \end{bmatrix}_q \int_0^1 \theta^n (1-\theta)_q^{m-n} d_q(\theta) = \begin{cases} \frac{q^{n+1}}{1+q+\cdots+q^m} & \text{if } n < m \\ \frac{1}{1+q+\cdots+q^m} & \text{if } n = m \\ 0 & \text{if } n > m \end{cases} \tag{12.16}$$

$(m, n = 0, 1, 2, \dots)$.

The matrix arising on the right in (12.16) is certainly q-Hausdorff (for all $q > 0$), its entries having been obtained "linearly" from those of a q-Hausdorff matrix, $E_q(\theta)$. We prefer, for reasons that will become clear below, to take the matrix displayed in (12.15) as *the q-Cesaro matrix* $(q > 0)$. By favoring (12.15) over (12.16) we are conceding the fact that q-integration may not be an appropriate tool for studying q-Hausdorff matrices.

q-Cesaro matrices, $C_q(\alpha), \alpha > 0$. The ordinary Cesaro matrices, $C(\alpha)$ (see (8.24) and (8.25)) form what is perhaps the most important family in all of classical summability theory and it is natural to seek out their q-analogues. We refuse to mess with q-integration, working instead with the most salient feature of the $C(\alpha)$'s: they are Nørlund means as well as Hausdorff.

Among the q-Hausdorff matrices ($q > 0$, fixed) we find a similar family of Nørlund means. The family may be conveniently parametrized by a single variable, and, after "lining it up" with α, we arrive at the following definition for the *q-Cesaro matrix of order* $\alpha (\alpha > 0)$.

$$C_q(\alpha)_{m,n} = \begin{cases} \frac{a_{m-n}}{A_m} & \text{if } n \le m \\ 0 & \text{if } n > m \end{cases} \tag{12.17}$$

$(m, n = 0, 1, 2, \dots)$, the sequence of weights being given by

$$\mathbf{a} = 1, \frac{\alpha}{q}, \frac{\alpha(\alpha+q)}{q^2(1+q)}, \frac{\alpha(\alpha+q)(\alpha+q+q^2)}{q^3(1+q)(1+q+q^2)}, \dots \tag{12.18}$$

We omit the proof that the $C_q(\alpha)$'s are indeed q-Hausdorff matrices, and that they are the only such Nørlund means. The q-moments of $C_q(\alpha)$, i.e. the diagonal entries, are given by

$$\begin{aligned} \boldsymbol{\mu} = 1,\ \frac{q}{\alpha+q},\ \frac{q^2(1+q)}{(\alpha+q)(\alpha+q+q^2)}&, \\ &\frac{q^3(1+q)(1+q+q^2)}{(\alpha+q)(\alpha+q+q^2)(\alpha+q+q^2+q^3)}, \dots \end{aligned} \tag{12.19}$$

It will be noticed that the weights and moments tend to their classical values (8.24), (8.25), when $q \to 1$.

q-Gamma matrices, $\Gamma_q(\alpha), \alpha > 0$. These are defined to be the q-Hausdorff matrices that are weighted means. The *q-Gamma matrix of order* α ($\alpha > 0$) is given by

$$\Gamma_q(\alpha)_{m,n} = \begin{cases} \frac{a_n}{A_m} & \text{if } n \le m \\ 0 & \text{if } n > m \end{cases} \tag{12.20}$$

($m, n = 0, 1, 2, \dots$). The weights / moments are given now by

$$\mathbf{a} = 1, q\alpha, q^2 \frac{\alpha(\alpha + \frac{1}{q})}{1 + \frac{1}{q}}, q^3 \frac{\alpha(\alpha + \frac{1}{q})(\alpha + \frac{1}{q} + \frac{1}{q^2})}{(1 + \frac{1}{q})(1 + \frac{1}{q} + \frac{1}{q^2})}, \dots \tag{12.21}$$

and

$$\boldsymbol{\mu} = \frac{\alpha}{\alpha}, \frac{\alpha}{\alpha + \frac{1}{q}}, \frac{\alpha}{\alpha + \frac{1}{q} + \frac{1}{q^2}}, \frac{\alpha}{\alpha + \frac{1}{q} + \frac{1}{q^2} + \frac{1}{q^3}}, \dots, \tag{12.22}$$

and these, too, agree with their classical values (8.22), (8.23) when $q \to 1$.

We come now to our final theorem which contains all the preceding inequalities as special cases: Theorem 11.1 by taking $A = B = C_q$, Theorem 7.2 by making $q \to 1$ and Theorem 4.1 by doing both.

Theorem 12.3. *Fix $q > 0$. If A and B are q-Hausdorff matrices with non-negative entries, then A, B form an urbane couple.*

Proof. We apply Lemma 7.1 with city plan (11.2). As in the proof of Theorem 7.2 it suffices to check just hypothesis (7.2), namely,

$$\sum_{j=0}^{m} u_{j,k}^{m,n} a_{m,j} = a_{n,k} \quad (0 \le k, n \le m). \tag{12.23}$$

We shall show, in fact, that (12.23) holds for all q-Hausdorff matrices, non-negative or not.

To do this, we begin with the "Eulerian" version of (12.23), the special case in which $A = E_q(\theta)$ for some $\theta \in [0, 1]$. The required identity, in view of (11.2) and (12.9), is

$$\sum_{j=k}^{m-n+k} q^{(j-k)(m-k)} \begin{bmatrix} m-n \\ j-k \end{bmatrix}_q \theta^j (1-\theta)_q^{m-j} = \theta^k (1-\theta)_q^{n-k} \tag{12.24}$$

($0 \le k \le n \le m$). Dividing the left side by the right and recalling definition (10.11), (12.24) may be rewritten as

$$\sum_{j=k}^{m-n+k} q^{(j-k)(m-k)} \begin{bmatrix} m-n \\ j-k \end{bmatrix}_q \theta^{j-k}(1-\theta q^{n-k})_q^{m-n-j+k} = 1$$

or

$$\sum_{i=0}^{m-n} \begin{bmatrix} m-n \\ i \end{bmatrix}_q (\theta q^{n-k})^i (1-\theta q^{n-k})^{m-n-i} = 1. \tag{12.25}$$

Identity (12.25) follows from Corollary 12.2 once we recognize its left side as the $(m-n)^{th}$ row sum of the q-Euler matrix $E_q(\alpha)$, with $\alpha = \theta q^{n-k}$. The Eulerian version of (12.23) is thus correct.

To show that (12.23) holds for all q-Hausdorff matrices, A, we apply the Method of Moriarty. This entails consideration of the linear operator $L : \mathbb{R}^{m+1} \to \mathbb{R}^{m+1}$, determined by

$$L(\theta^k) = a_{k,k} \qquad (k = 0, 1, \ldots, m),$$

θ being an indeterminate. The general version of (12.23) follows by applying L to both sides of the Eulerian version, (12.24). □

13. Future Plans

Machine calculations indicate that inequality (1.1) holds in far greater generality than has been discussed above. This is exciting because it suggests that new combinatorial identities of Vandermonde type are possible. (If A is an urbane matrix that is not q-Hausdorff, then any city plan for A must differ from (11.2). To see this we have only to observe that the q-analogue of Proposition 8.5 holds.)

Problem 13.1. *(i) Is the odd weighted mean urbane? (ii) How about the odd Nørlund mean?*

Both matrices are generated by the sequence of weights $\mathbf{a} = 1, 3, 5, 7, \ldots$.

References

[1] R.P. Agnew, A genesis for Cesaro methods, *Bull. Amer. Math. Soc.* **51** (1945), 90–94.

[2] G.E. Andrews, *The Theory of Partitions*, Encyclopedia of Mathematics and its Applications **2** (1976), Addison–Wesley Reading, Mass

[3] G. Bennett, Lower bounds for matrices, II, *Canadian Journ. Math.* **44 (1)** (1992), 54–74.

[4] G. Bennett, Double dipping: the case of the missing binomial coefficient identities, *Theoretical Computer Science* **123** (1994), 351–375.

[5] G. Bennett, From coin tossing to the Jacobi Polynomials, *Jour. Combinatorial Theory*, Series A **73 (2)** (1996), 248–272.

[6] G. Bennett, Factorizing the Classical Inequalities, *Memoirs Amer. Math. Soc.* **576** (1996), 1–130.

[7] G. Bennett, Inequalities complimentary to Hardy, *Quart. J. Math. Oxford (2)* **49** (1998), 395–432.

[8] G. Bennett, *Summability Matrices and Random Walk*, in preparation.

[9] G. Gasper and M. Rahman, *Basic Hypergeometric Series*, Encyclopedia of Mathematics and its Applications **35** (1990), Cambridge Univ. Press.

[10] G.H. Hardy, *Divergent Series*, (1949), Oxford University Press.

[11] G.H. Hardy, J.E. Littlewood and E. Polya, *Inequalities*, 2nd edition (1967), Cambridge University Press.

[12] B.E. Rhoades, Commutants of some Hausdorff matrices, *Pacific Jour. Math.* **42** (1972), 715–719.

Received September 15, 1998

Department of Mathematics, Indiana University, Bloomington, IN 47405-7106
E-mail address: bennettg@indiana.edu