A DEVICE FOR STUDYING HAUSDORFF MOMENTS

BY

CHANDLER DAVIS

1. **Introduction.** The *Hausdorff moment problem* [9, pp. 1, 8–9] asks for necessary and sufficient conditions on the numbers μ_n in order that there exist a distribution function Φ on $[0, 1]$ such that

$$\text{(1)} \qquad \text{for all } n \in I, \qquad \mu_n = \int_0^1 t^n d\Phi(t).$$

Here $I = \{0, 1, 2, \cdots\}$. The *reduced Hausdorff moment problem* [9, p. 77] asks the same question where I is a proper subset of $\{0, 1, 2, \cdots\}$, usually a finite subset. If I is allowed to be any subset, this includes the first problem.

It is known (cf. [11, Theorem 10.30]) and easy to prove that one condition is the existence of a matrix $A = ((A_{ij}))_{i,j=0}^{\infty}$, such that $0 \leqq A \leqq 1$ and

$$\text{(2)} \qquad \text{for all } n \in I, \qquad \mu_n = (A^n)_{00}.$$

The matrix A may be chosen to be a Jacobi matrix, that is, $A_{ij} = 0$ for $|i-j| > 1$; and $A_{n,n+1} \geqq 0$ may be required. If I is $\{0, 1, 2, \cdots\}$, A is then determined uniquely, assuming the convention that any invariant subspace of A orthogonal to the 0th coordinate subspace will be ignored. (Assuming, that is, that if $A_{k,k+1} = A_{k+1,k} = 0$ then in $A = ((A_{ij}))$ indices will be let run up to k only.) For finite I, A is not in general determined uniquely.

This paper gives a convenient canonical form for A. There is little trouble in including in this result the generalized sort of moment problem introduced by Nagy [6], where μ_n above are in $\mathfrak{B}$, the set of bounded self-adjoint operators on a Hilbert space $\mathfrak{H}$. One virtue of the generalized problem is its application to the classical problem; see Proposition 4 below. Accordingly some of the facts outlined in the preceding paragraph may as well be proved for the generalized problem; this is done in §2, which in fact is essentially a recitation in the wider setting of the proof for the classical case. The main result is in §3. The following sections mostly examine the classical problem in light of it. I have found the canonical form handy for getting numerical bounds on moments, but will not discuss this use further.

Throughout the paper I use the notation $\tilde{\alpha} \equiv 1 - \alpha$, where α may be a number or an operator. If it is an operator this requires some understanding as to the space on which it operates, but I think I have avoided ambiguity.

Presented to the Society February 25, 1956; received by the editors September 10, 1956.

2. Jacobi matrices.

THEOREM 1. *Let* $\mu_n \in \mathcal{B}$ *for* $n \in I$, *where* $I \subseteq \{0, 1, 2, \cdots\}$. *The following are equivalent*:

(i) *There exists a function* Φ *on*(1) $[0, 1]$ *to* $\mathcal{B}$ *such that* $\Phi(0^-) = 0 \leqq \Phi(t_1) \leqq \Phi(t_2) \leqq \Phi(1) = 1$ *for* $0 \leqq t_1 \leqq t_2 \leqq 1$, *and such that* (1) *holds*.

(ii) *There exists a matrix* $A = ((A_{ij}))_{i,j=0}^{\infty}$, *with* $A_{ij} \in \mathcal{B}$, $A_{ij} \geqq 0$, *and* $A_{ij} = 0$ *for* $|i-j| > 1$; *such that* $0 \leqq A \leqq 1$ *and* (2) *holds*.

Proof. Take $I = \{0, 1, 2, \cdots\}$, because from this the result for any subset of indices will follow.

Neumark's theorem [7](2) says that $\Phi(t)$ with the properties described in (i) can be expressed as $\Phi(t) = PE(t)P$, where $E(t)$ is a resolution of the identity in a Hilbert space $\mathcal{K}$ containing $\mathcal{H}$, and P is the projection on $\mathcal{K}$ onto $\mathcal{H}$. That is, (i) is equivalent to the existence of such $\mathcal{K}$ and an operator A on $\mathcal{K}$, $0 \leqq A \leqq 1$, such that for $n \in I$, $\mu_n = PA^nP$. Now it is straightforward to define orthogonal projections $P_1, P_2, \cdots$ such that relative to the subspaces $P_i\mathcal{K}$ the matrix of A has the desired form; as follows.

The P_i and A_{ij} will be defined inductively together with operators E_i such that E_i maps $P_i\mathcal{K}$ isometrically onto $\mathcal{H}$, and $E_iP_j = 0$ for $i \neq j$, $P_i = E_i^*E_i$. Let $E_0 = P_0 = P$ above, that is, $\mathcal{H}$ is identified by the identity mapping with $P_0\mathcal{K}$; $A_{00} = P_0AP_0$. Suppose $A_{ij} \in \mathcal{B}$ and E_i, satisfying all requirements, have been defined whenever $i, j \leqq k$. Let $A_{k+1,k} \in \mathcal{B}$ and E_{k+1} satisfy the following:

$$\tilde{P}_k A P_k = E_{k+1}^* A_{k+1,k} E_k; \tag{3}$$

$A_{k+1,k} \geqq 0$; $E_{k+1}^*E_k$ an isometry on the range of $E_k^*A_{k+1,k}E_k \geqq 0$. (This is a polar resolution, for which see e.g. [8, §110].) As noted above, E_{k+1}^* is zero on $\mathcal{H}^\perp$. But also (3) may fail to define E_{k+1}^* because $E_k^*A_{k+1,k}E_k$ as an operator on $P_k\mathcal{K}$ may have a nullspace. If so, let $E_{k+1}^*E_k$ map it isometrically onto a subspace of the nullspace of A orthogonal to $\sum_0^k P_i\mathcal{K}$; this is always possible because $\mathcal{K}$ is a space being constructed and can be augmented if desired by a new orthogonal subspace on which A is defined to be 0. Last, of course, define $P_{k+1} = E_{k+1}^*E_{k+1}$ and $P_{k+1}AP_{k+1} = E_{k+1}^*A_{k+1,k+1}E_{k+1}$.

It is now clear that (i) implies (ii). The converse deduction is the same only easier.

I left the facts on uniqueness out of the statement of the theorem. The operator A given by Neumark's theorem is unique up to isomorphism if we require (as we clearly may) that there is no subspace of $\mathcal{K}$ invariant under A and orthogonal to $\mathcal{H}$ [7]. But the construction above introduced such in-

(1) With the usual understanding that $\Phi(0^-)$ is defined. We may as well assume $\Phi(t) = \Phi(t^+)$.

(2) This is an appropriate occasion to acknowledge in print my mistake in announcing this theorem as new [1]. My method, it happened, was different from Neumark's, being an extended version of a device of E. A. Michael [4, Thm. 2].

essential subspaces of $\mathcal{K}$. Before uniqueness can be asserted $\mathcal{K}$ must be pared down to that subspace which actually plays a part. This is done (at the cost of some clumsiness of statement) in Theorem 3 below.

3. **The canonical form.**

THEOREM 2. Let [3] $A=((A_{ij}))_{i,j=0}^{\infty}$, *$A_{ij}$ bounded operators on* $\mathcal{K}$; $A=A^*$; *and* $A_{ij}=0$ *for* $|i-j|>1$. *If A has any invariant subspace orthogonal to the* 0*th coordinate Hilbert space, assume A annihilates it. In order that* $0\leqq A\leqq 1$, *it is necessary and sufficient that there exist* $\eta_i\in\mathcal{B}$ $(i=1, 2, \cdots)$ *with* $0\leqq\eta_i\leqq 1$, *and partial isometries ζ_i defined on $\mathcal{K}$ into $\mathcal{K}$, such that*

$$A_{ii} = \zeta_i^*((\eta_{2i})^{1/2}\tilde{\eta}_{2i-1}(\eta_{2i})^{1/2} + (\tilde{\eta}_{2i})^{1/2}\eta_{2i+1}(\tilde{\eta}_{2i})^{1/2})\zeta_i, \tag{4}$$

$$A_{i-1,i} = A_{i,i-1}^* = \zeta_{i-1}^*(\tilde{\eta}_{2i-2})^{1/2}(\eta_{2i-1}\tilde{\eta}_{2i-1})^{1/2}(\eta_{2i})^{1/2}\zeta_i. \tag{5}$$

Here the convention is $\eta_0=0$, $\zeta_0=1$.

Some general facts about matrices will be given as lemmas. In the lemmas, *let each $\mathcal{K}_i$ be a Hilbert space; let each a_{ij} be a bounded operator on $\mathcal{K}_j$ to $\mathcal{K}_i$; let* $a_{ji}=a_{ij}^*$ *and* $a_{ii}\geqq 0$.

LEMMA 1. *In order that*

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \geqq 0,$$

it is necessary and sufficient that $a_{21}a_{11}^{-1}$ be bounded on the range of a_{11} and that $b\equiv a_{22}-a_{21}a_{11}^{-1}a_{12}\geqq 0$.

Proof. *Case* I. $a_{21}a_{11}^{-1}$ is densely defined but unbounded.

Then there is a sequence $\{x_\nu\}$ of unit elements of $\mathcal{K}_1$ such that $\|a_{11}x_\nu\|$ approaches zero but $\|a_{21}x_\nu\|$ does not. Selecting a subsequence if necessary gives that for some $\delta>0$ and for all ν, $\|a_{21}x_\nu\|>\delta$. As a temporary convenience, identify $\mathcal{K}_1$ with $\mathcal{K}_2$, in such a way that $a_{21}\geqq 0$. (This may always be done, by enlarging one or the other $\mathcal{K}_i$ if needed.) Then, by expanding $\|(\|a_{21}\|-a_{21})x_\nu\|^2 \leqq\|a_{21}\|^2$, one easily computes that $(a_{21}x_\nu, x_\nu)>\delta'\equiv\delta^2/2\|a_{21}\|>0$. Now consider

$$P(x, y) \equiv (a_{11}x, x) + (a_{12}y, x) + (a_{21}x, y) + (a_{22}y, y). \tag{6}$$

I must show that, for some choice of $x, y\in\mathcal{K}_1$, $P(x, y)<0$. This is accomplished by letting ν be so large that $\|a_{11}x_\nu\|<\delta'^2/\|a_{22}\|$; whence

$$P(\|a_{22}\|x_\nu, -\delta'x_\nu)<0$$

may be readily verified.

[3] The assumption $A_{ij}\geqq 0$ has been dropped. Keeping it would not obviate the nuisance of mentioning "phases" ζ_i anyhow.

Case II. a_{11} has a nullspace. Then a_{21} may be assumed to be zero there, and $a_{21}a_{11}^{-1}$ may be defined to be zero there. (The proof is like the preceding one, but simpler.) Granted this, Case III applies.

Case III. $a_{21}a_{11}^{-1}$ exists as a bounded operator on $\mathcal{H}_1$ to $\mathcal{H}_2$. From this follow the existence and boundedness of $a_{11}^{-1}a_{12}$, $a_{21}a_{11}^{-1/2}$, and $a_{11}^{-1/2}a_{12}$.

Now (6) may be rewritten, by making the substitution $a_{22}=a_{21}a_{11}^{-1}a_{12}+b$ and simplifying, as

$$P(x, y) = \left\|a_{11}^{1/2}x + a_{11}^{-1/2}a_{12}y\right\|^2 + (by, y).$$

Clearly $b\geqq 0$ implies $P(x, y)\geqq 0$; half of the lemma is proved. In case $b\ngeqq 0$, choose $y\in\mathcal{H}_2$ so $(by, y)<0$, and let $x=-a_{11}^{-1}a_{12}y$. For this choice, $P(x, y)<0$.

LEMMA 2. *In order that*

$$\begin{pmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & a_{23} \\ 0 & a_{32} & a_{33} \end{pmatrix} \geqq 0, \tag{7}$$

it is necessary and sufficient that

$$a_{22} = b + c, \text{ with } \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & b \end{pmatrix} \geqq 0, \qquad \begin{pmatrix} c & a_{23} \\ a_{32} & a_{33} \end{pmatrix} \geqq 0. \tag{8}$$

The definite matrix is being expressed as a sum of simpler definite matrices.

Proof. (8) is equivalent to

$$a_{22} \geqq a_{21}a_{11}^{-1}a_{12} + a_{23}a_{33}^{-1}a_{32} \geqq 0. \tag{9}$$

This results by applying Lemma 1 to $\mathcal{H}_1$ and $\mathcal{H}_2$, then to $\mathcal{H}_3$ and $\mathcal{H}_2$.

But Lemma 1 may also be applied to the pair of spaces $\mathcal{H}_1\oplus\mathcal{H}_3$ and $\mathcal{H}_2$. This says (7) is equivalent to

$$a_{22} \geqq (a_{21} \ \ a_{23})\begin{pmatrix} a_{11} & 0 \\ 0 & a_{33} \end{pmatrix}^{-1}\begin{pmatrix} a_{12} \\ a_{32} \end{pmatrix} \geqq 0,$$

which is the same as (9).

LEMMA 3. *In order that there exist a_{22} such that*

$$0 \leqq \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \leqq 1,$$

it is necessary and sufficient that $0\leqq a_{11}\leqq 1$ and $a_{21}a_{11}^{-1}\tilde{a}_{11}^{-1}a_{12}\leqq 1$.

Proof. Although I have chosen to state Lemma 3 in this simple form, a slight generalization is needed below. What I will prove is the generalization. Impose as upper bound on the matrix

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

not 1, but instead

$$\begin{pmatrix} b & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{with} \quad 0 \leqq b.$$

This requires $0 \leqq a_{11} \leqq b$, or, what is equivalent, the existence of c such that $0 \leqq c \leqq 1$ and $a_{11} = b^{1/2} c b^{1/2}$. The problem is to find a condition which, together with $0 \leqq c \leqq 1$, is necessary and sufficient for the matrix to satisfy the inequalities.

Use Lemma 1 twice: for the matrix to be $\geqq 0$, $a_{22} \geqq a_{21} a_{11}^{-1} a_{12}$; and for the matrix to be

$$\leqq \begin{pmatrix} b & 0 \\ 0 & 1 \end{pmatrix}, \qquad \tilde{a}_{22} \geqq (-a_{21})(b - a_{11})^{-1}(-a_{12}).$$

Such a_{22} can exist if and only if

$$\begin{aligned} 1 = a_{22} + \tilde{a}_{22} &\geqq a_{21}(a_{11}^{-1} + (b - a_{11})^{-1}) a_{12} \\ &= a_{21} b^{-1/2} (c^{-1} + \tilde{c}^{-1}) b^{-1/2} a_{12} = a_{21} b^{-1/2} c^{-1} \tilde{c}^{-1} b^{-1/2} a_{12}. \end{aligned}$$

(As in Lemma 1, the formula makes sense even though some inverses may fail to exist everywhere.) When $b=1$, $c=a_{11}$, giving Lemma 3 as stated.

Proof of Theorem 2. To say that $0 \leqq A \leqq 1$ is to say that $A \geqq 0$ and $\tilde{A} \geqq 0$. Both A and $\tilde{A}$ are given as Jacobi matrices with operator elements; $(\tilde{A})_{ii} = (A_{ii})^{\sim}$, $(\tilde{A})_{i-1,i} = -A_{i-1,i}$.

According to (4), $A_{00} = \eta_1$; so all the theorem says about A_{00} is $0 \leqq A_{00} \leqq 1$. This is evidently necessary and sufficient for the existence of $A_{01}, A_{11}, \cdots$ such that $0 \leqq A \leqq 1$.

Proceed by induction on k. Inductive hypothesis: That $A_{00}, A_{01}, \cdots, A_{k-2,k-1}, A_{k-1,k-1}$ satisfy (4) and (5), with η_i and ζ_i as described, is necessary and sufficient for the existence of $A_{k-1,k}, A_{kk}, \cdots$, such that the A in the statement of the theorem will satisfy $0 \leqq A \leqq 1$.

The theorem will have been proved(4) once the following is deduced from the inductive hypothesis: Given $A_{00}, A_{01}, \cdots, A_{k-2,k-1}, A_{k-1,k-1}$ expressed in the prescribed form. That $A_{k-1,k}$ (and of course its adjoint $A_{k,k-1}$) and A_{kk} be expressible in the prescribed form, consistently with the preceding A_{ij}, is necessary and sufficient for the existence of $A_{k,k+1}, A_{k+1,k+1}, \cdots$ such that $0 \leqq A \leqq 1$.

Furthermore we may simplify the discussion by assuming in the proof that $A_{k,k+1} = A_{k+1,k+1} = \cdots = 0$, or that the matrix expression of A is

(4) To speak strictly, I must invoke the second sentence of the statement of Theorem 2 in order to guarantee that the induction's succeeding proves the Theorem.

$(k+1)\times(k+1)$. To see this, use P_i, the projections on the coordinate Hilbert spaces, as in the proof of Theorem 1. In general $0 \leqq A \leqq 1$ implies $0 \leqq (P_0 + \cdots + P_k)A(P_0 + \cdots + P_k) \leqq 1$; while this relation, though it does not imply $0 \leqq A \leqq 1$, does imply the existence of *some* choice of $A_{k,k+1}$, $A_{k+1,k+1}, \cdots$ such that $0 \leqq A \leqq 1$—namely, the choice that all be zero. So take $P_{k+1} = P_{k+2} = \cdots 0$.

As another simplification, we may assume $\zeta_0 = \cdots = \zeta_{k-1} = 1$ hereafter. (The most that could possibly be involved here is replacing one replica of $P_0\mathcal{K}$ by another. If ζ_i^* annihilates some of the range of an operator it premultiplies, either that operator or ζ_i may be redefined.)

Consider $\mathcal{K}$, the Hilbert space where A is defined, as the sum of the following three spaces: $\mathcal{K}_1 = (P_0 + \cdots + P_{k-2})\mathcal{K}$, $\mathcal{K}_2 = P_{k-1}\mathcal{K}$, $\mathcal{K}_3 = P_k\mathcal{K}$. Decompose A in this way: $a_{22} = A_{k-1,k-1}$, etc.

Now to derive conditions on $A_{k-1,k}$ and $A_{k,k-1} = (A_{k-1,k})^*$, apply Lemma 2.

By the inductive hypothesis, the g.l.b. (in the poset of bounded Hermitian operators) of allowable values for $A_{k-1,k-1}$ is $\eta_{2k-2}^{1/2}\tilde{\eta}_{2k-3}\eta_{2k-2}^{1/2}$. Hence, for $A \geqq 0$ the requirement on a_{23} is

$$\begin{pmatrix} d & a_{23} \\ a_{32} & a_{33} \end{pmatrix} \geqq 0,$$

with $d = A_{k-1,k-1} - \eta_{2k-2}^{1/2}\tilde{\eta}_{2k-3}\eta_{2k-2}^{1/2} = \tilde{\eta}_{2k-2}^{1/2}\eta_{2k-1}\tilde{\eta}_{2k-2}^{1/2}$. The same reasoning derives a requirement from the condition $\tilde{A} \geqq 0$. The g.l.b. of allowable values for $\tilde{A}_{k-1,k-1}$ is $\eta_{2k-2}^{1/2}\eta_{2k-3}\eta_{2k-2}^{1/2}$, so the condition on a_{23} is

$$\begin{pmatrix} d' & -a_{23} \\ -a_{32} & \tilde{a}_{33} \end{pmatrix} \geqq 0,$$

with $d' = \tilde{A}_{k-1,k-1} - \eta_{2k-2}^{1/2}\eta_{2k-3}\eta_{2k-2}^{1/2} = \tilde{\eta}_{2k-2}^{1/2}\tilde{\eta}_{2k-1}\tilde{\eta}_{2k-2}^{1/2}$. Combining the two conditions on a_{23},

$$0 \leqq \begin{pmatrix} d & a_{23} \\ a_{32} & a_{33} \end{pmatrix} \leqq \begin{pmatrix} \eta_{2k-2} & 0 \\ 0 & 1 \end{pmatrix},$$

because $d + d' = \eta_{2k-2}$. By the generalization of Lemma 3 (or, when $k=1$, Lemma 3 itself), this is equivalent to

$$1 \geqq A_{k,k-1}\eta_{2k-2}^{-1/2}\eta_{2k-1}^{-1}\tilde{\eta}_{2k-1}^{-1}\eta_{2k-2}^{-1/2}A_{k-1,k}. \tag{10}$$

(Remember that $a_{23} = A_{k-1,k}$.)

$A_{k-1,k}$ may be written in the form (5) for some bounded $\eta_{2k} \geqq 0$ and partial isometry ζ_k, just by virtue of the fact, contained in (10), that

$$\eta_{2k-1}^{-1/2}\tilde{\eta}_{2k-1}^{-1/2}\eta_{2k-2}^{-1/2}A_{k-1,k}$$

is bounded. The closure of the range of η_{2k} may be assumed contained in that of ζ_k. But when (5) is substituted in, (10) reduces readily to $\zeta_k\zeta_k^* \geqq \zeta_k\zeta_k^*\eta_{2k}\zeta_k\zeta_k^*$, which is equivalent to $\eta_{2k} \leqq 1$.

It remains to consider $A_{kk}=a_{33}$.

By Lemmas 2 and 1, the condition equivalent to $A\geqq 0$ is this:

$$a_{22} = b + c, \quad \text{with} \quad \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & b \end{pmatrix} \geqq 0, \quad \text{and} \quad a_{33} \geqq a_{32}c^{-1}a_{23}.$$

The existence of some choice for c which will allow a solution, is ensured by the argument just concluded about a_{23}. The g.l.b. of acceptable values for $a_{32}c^{-1}a_{23}$, hence for a_{33}, is attained when c attains its l.u.b., hence when b attains its g.l.b.; this is not hard to prove. Also any a_{33} satisfying the resulting inequality will be consistent with $A\geqq 0$. Now the g.l.b. of b under the restriction

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & b \end{pmatrix} \geqq 0$$

is, by the inductive hypothesis, $\eta_{2k-2}^{1/2}\tilde{\eta}_{2k-3}\eta_{2k-2}^{1/2}$; the corresponding value for c is $\tilde{\eta}_{2k-2}^{1/2}\eta_{2k-1}\tilde{\eta}_{2k-2}^{1/2}$. Using this, and using (5) for a_{23} and a_{32}, the condition on a_{33} equivalent to $A\geqq 0$ becomes

$$a_{33} \geqq \overset{*}{\zeta}_k(\eta_{2k})^{1/2}\tilde{\eta}_{2k-1}(\eta_{2k})^{1/2}\zeta_k. \tag{11}$$

By similar reasoning, the condition on a_{33} equivalent to $\tilde{A}\geqq 0$ becomes

$$\tilde{a}_{33} \geqq \overset{*}{\zeta}_k(\eta_{2k})^{1/2}\tilde{\eta}_{2k-1}(\eta_{2k})^{1/2}\zeta_k. \tag{12}$$

Assuming[5] as in analogous situations above that $a_{33}=\zeta_k\zeta_k^*a_{33}$, we see that the only effect of $\zeta_k\neq 1$ is to oblige us to deal with $\zeta_k^*a_{33}\zeta_k$ in the rest of this paragraph, so assume $\zeta_k=1$. Now $a_{33}-\eta_{2k}^{1/2}\tilde{\eta}_{2k-1}\eta_{2k}^{1/2}$ is an operator which whenever (11) holds is $\geqq 0$ and which whenever (12) holds is $\leqq\tilde{\eta}_{2k}$. Hence it may be written in the form $\tilde{\eta}_{2k}^{1/2}\eta_{2k+1}\tilde{\eta}_{2k}^{1/2}$ for some η_{2k+1} with $0\leqq\eta_{2k+1}\leqq 1$. This expresses $a_{33}=A_{kk}$ in the required form (4).

The proof is complete.

Next I combine Theorems 1 and 2, in a somewhat superior formulation.

DEFINITION. *Let* $\mathcal{H}=\mathcal{H}_0\supseteq\mathcal{H}_1\supseteq\mathcal{H}_2\supseteq\cdots$ *be Hilbert spaces; let* A_{ij} *be a bounded operator on* $\mathcal{H}_j$ *to* $\mathcal{H}_i$, *for* $i, j=0, 1, 2, \cdots$. *Then* $A=((A_{ij}))$ *is an operator with domain in* $\mathcal{K}=\sum_i\oplus\mathcal{K}_i$, *where* $\mathcal{K}_i$ *is a replica of* $\mathcal{H}_i$, *defined in the natural way: denoting* $x=\sum\oplus x_i$ *and* $Ax=\sum\oplus(Ax)_i$, *then* $(Ax)_i=\sum_j A_{ij}x_j$. Evidently A need not thereby be everywhere defined. *This is called an expression of* A *as a pruned matrix* with respect to the given descending sequence of subspaces. The usual rule for matrix multiplication of course applies to pruned matrices.

THEOREM 3. *Let* $\mu_n\in\mathcal{B}$ *for* $n\in I$, *where* $I\subseteq\{0, 1, 2, \cdots\}$. *The following are equivalent*:

(5) See footnote 4.

(i) *There exists a function* Φ *on*[1] $[0, 1]$ *to* $\mathcal{B}$ *such that* $\Phi(0^-)=0\leqq\Phi(t_1)\leqq\Phi(t_2)\leqq\Phi(1)=1$ *for* $0\leqq t_1\leqq t_2\leqq 1$, *and such that*

$$\text{(1)} \qquad \text{for all } n\in I, \qquad \mu_n = \int_0^1 t^n d\Phi(t).$$

(ii) *For* $i=1, 2, \cdots$, *there exist subspaces* $\mathcal{H}_i$, $\mathcal{H}_{i+1}\subseteq\mathcal{H}_i\subseteq\mathcal{H}_0=\mathcal{H}$; *and there exist* $\eta_n\in\mathcal{B}$, *with* $0\leqq\eta_n\leqq 1$, *with the closure of the range of* η_{2i} *being* $\mathcal{H}_i$, *the range of* η_{2i+1} *being* $\subseteq\mathcal{H}_i$, *and the range of* $\eta_{2i-1}\tilde{\eta}_{2i-1}$ *being* $\supseteq\mathcal{H}_i$; *such that the pruned matrix* $A=((A_{ij}))$ *defined by*

$$\text{(4')} \qquad A_{ii} = (\eta_{2i})^{1/2}\tilde{\eta}_{2i-1}(\eta_{2i})^{1/2} + (\tilde{\eta}_{2i})^{1/2}\eta_{2i+1}(\tilde{\eta}_{2i})^{1/2},$$

$$\text{(5')} \qquad A_{i-1,i} = A^*_{i,i-1} = (\tilde{\eta}_{2i-2})^{1/2}(\eta_{2i-1}\tilde{\eta}_{2i-1})^{1/2}(\eta_{2i})^{1/2},$$

$A_{ij}=0$ *for* $|i-j|>1$, *satisfies*

$$\text{(2)} \qquad \text{for all } n\in I, \qquad \mu_n = (A^n)_{00}.$$

Here the convention is $\eta_0=0$. *Necessarily* $0\leqq A\leqq 1$.

If $I=\{0, 1, 2, \cdots\}$, *then the* η_n *and the* μ_n *uniquely determine each other. In fact,* η_n *is determined by those* μ_p *with* $p\leqq n$, *and conversely,* μ_n *is determined by those* η_p *with* $p\leqq n$.

No detailed proof need be given. The unsightly conditions put on the $\mathcal{H}_i$ and the η_n bar the introduction of inessential subspaces to $\mathcal{H}$, and the uniqueness assertion of Neumark's theorem can be invoked[6]. The strong uniqueness statement of the last sentence of Theorem 3 now involves no ideas not already enlisted in the proof of Theorems 1 and 2. A full proof may be supplied by the reader immune to tedium.

Note. Since η_{2i} is always zero on $\mathcal{H}_i^\perp$, it might be suggested that it and $\tilde{\eta}_{2i}$ be defined only on $\mathcal{H}_i$. However, it turns out to be more natural to define $\eta_{2i}=0$ and $\tilde{\eta}_{2i}=1$ on $\mathcal{H}_i^\perp$. Odd subscripts are different: η_{2i+1} and $\tilde{\eta}_{2i+1}$ enter just as symmetrically as A and $\tilde{A}$ do. We may reasonably stick to $\eta_{2i+1}=\tilde{\eta}_{2i+1}=0$ on $\mathcal{H}_i^\perp$.

4. **The classical case.** Here the Hilbert space $\mathcal{H}$ is 1-dimensional, so members of $\mathcal{B}$ are real numbers. In particular, $\mu_n\in[0, 1]$. In (4′) and (5′), everything commutes, and $A_{i-1,i}\geqq 0$. It may be that $\mathcal{H}_i=\mathcal{H}$ for all i, in which case A has rows and columns indexed $0, 1, 2, \cdots$; or, from $i=k$ on, $\mathcal{H}_i$ may be zero, in which case A is $k\times k$.

Theorem 1 is more closely related than it might appear to the standard solution [9, Theorem 1.5] of the Hausdorff moment problem. One considers the mapping M on polynomials defined by

$$M(a_nt^n+\cdots+a_0) = a_n\mu_n+\cdots+a_0.$$

The condition put on the μ_n is that kth order differences all be non-negative

(6) See the last paragraph of §2.

for all k. This is proved equivalent to the positivity of M as a mapping on polynomials on [0, 1]. Neumark's theorem [7] may be regarded as asserting that such M is obtainable as a homomorphism to another C^*-algebra followed by a projection [10, §1]. But any C^*-algebra which is a homomorphic image of the C^*-algebra of continuous functions on [0, 1], is generated by a single operator A, the image of the polynomial t. There must be an operator A and a projection P such that $\mu_n P = M(t^n)P = PA^nP$. As above, if $E(t)$ is the spectral resolution of A and $\Phi(t) = PE(t)P$, then $\mu_n = \int_0^1 t^n d\Phi(t)$. Tools used in [6] extend this to the case where the μ_n are operators.

The foregoing remarks are not advanced as an improved or even an alternate proof of the Hausdorff moment theorem. They avoid no difficulty of the standard proof, and they entail new ones. The aim is merely to make the relation explicit.

But now, the operator homomorphism having been introduced, the canonical form of Theorem 3 above is made available. I will apply this to the classical case.

5. **Relation between the parameters and the distribution.** *Throughout the rest of the paper*, except where otherwise stated, μ_n *are numbers*, A *is the associated matrix in the sense of Theorem* 3, *with numerical parameters* η_n, *and* $\Phi(t)$ *is the associated distribution.*

PROPOSITION 1. *The set of points of nonconstancy of* $\Phi(t)$ *is exactly the spectrum of* A.

This is obvious from Theorem 1.

DEFINITION (cf. [12; 5, §9]). *A distribution on* [0, 1] *is of degree* m, *with* $m = 1/2, 3/2, \cdots$, *provided it is concentrated in* $m+1/2$ *distinct points* $\alpha_0 < \alpha_1 < \cdots < \alpha_{m-1/2}$, *and either* $\alpha_0 = 0$ *or* $\alpha_{m-1/2} = 1$ (*but not both*). *In the former case the degree may be written* m_*, *in the latter case* m^*.

A distribution on [0, 1] *is of degree* m, *with* $m = 1, 2, \cdots$, *provided either*

it is concentrated in m *distinct points* α_i, $0 < \alpha_0 < \alpha_1 < \cdots < \alpha_{m-1} < 1$, *in which case the degree may be written* m^*; *or*

it is concentrated in $m+1$ *distinct points* α_i, $0 = \alpha_0 < \alpha_1 < \cdots < \alpha_m = 1$, *in which case the degree may be written* m_*.

PROPOSITION 2. *Let the sequence* $\{\eta_n\}$ *terminate at* η_{2m}, *where* $m = 1/2, 1, 3/2, 2, \cdots$. *That is, let* $\eta_1\bar{\eta}_1 > 0, \cdots, \eta_{2m-1}\bar{\eta}_{2m-1} > 0$, $\eta_{2m}\bar{\eta}_{2m} = 0$. *If* $\eta_{2m} = 0$, Φ *is of degree* m_*; *if* $\eta_{2m} = 1$, Φ *is of degree* m^*. *Conversely, all* Φ *of finite degree arise in this way.*

Proof. Let $k = 1, 2, \cdots$.

If $\eta_{2k} = 0$, A is $k \times k$ and can have at most k distinct eigenvalues. Then by Proposition 1, Φ has at most k points of nonconstancy. Similarly if $\eta_{2k-1} = 0$, $\eta_{2k-1} = 1$, or $\eta_{2k-2} = 1$.

Conversely, suppose the distribution Φ concentrated in exactly k distinct points. The construction of Neumark's theorem in this case(7) yields an A which is $k\times k$. By the last paragraph of Theorem 3, no different A will do. Therefore $\{\eta_n\}$ terminates at $\eta_{2k}=0$, $\eta_{2k-1}=0$, $\eta_{2k-1}=1$, or $\eta_{2k-2}=1$.

If $\eta_{2k-1}=0$, $A_{k-1,k-1}$ is $\tilde{\eta}_{2k-3}\eta_{2k-2}$, and, by Theorem 3, diminishing $A_{k-1,k-1}$ while leaving other matrix elements unchanged would make $A\geqq 0$ cease to hold. But if there was a positive lower bound to the spectrum of A, a sufficiently small positive multiple of *any* positive semidefinite matrix could be subtracted from A keeping the result $\geqq 0$. Therefore 0 is an eigenvalue of A. Similarly, if $\eta_{2k-1}=1$, 1 is an eigenvalue of A; and if $\eta_{2k-2}=1$, 0 and 1 are eigenvalues of A.

Conversely, let Φ as above have 0 as a point of nonconstancy. Now it may seem that the corresponding eigenspace of A could be orthogonal to the $(k-1)$th coordinate space; suppose this, and let the jth coordinate space be the highest to which it is *not* orthogonal. Then reduction of A_{jj} would be inconsistent with $A\geqq 0$. It follows that either $\eta_{2j+1}=0$ or $\eta_{2j}=1$ (otherwise diminishing of η_{2j+1} would be an available way of diminishing A_{jj}). If $j<k-1$, this is a contradiction. For $j=k-1$, it gives the desired conclusion that either $\eta_{2k-1}=0$ or $\eta_{2k-2}=1$. Similarly, if Φ has 1 as a point of nonconstancy, either $\eta_{2k-1}=1$ or $\eta_{2k-2}=1$.

The statement of Proposition 2 is exactly the expression of the facts just proved in terms of the preceding definition.

In this proof I bypassed the question of whether in general an eigenvector of A can be orthogonal to the last coordinate space. The question is worth settling, though. To begin with, no eigenvector of A can be orthogonal to the 0th coordinate space, as already remarked. But also any admissible finite-dimensional A remains admissible when the order of rows and columns is reversed. (The proof of this fact is omitted; it involves retracing some of Theorem 2, in the classical case. Of course, reversing the order of rows and columns does *not* merely reverse the sequence of parameters η_n, but replaces them by an entirely new sequence.) Therefore no eigenvector of A can be orthogonal to the last coordinate space. This is used in Proposition 4 below.

PROPOSITION 3. *The other η_p being held constant, μ_n is a strictly increasing linear function of η_n.*

Proof. I mean to imply by granting the existence of η_n that the earlier parameters do not have extreme values. The later parameters, on the other hand, are without effect on μ_n, so $\eta_{n+1}=0$ may be assumed.

Let B differ from A only in having $\eta_n=0$.

Case I. $n=2i+1$. Then A is $(i+1)\times(i+1)$ with last entry $A_{ii}=\tilde{\eta}_{2i-1}\eta_{2i}+\tilde{\eta}_{2i}\eta_{2i+1}$. The only entry in which B differs is $B_{ii}=\tilde{\eta}_{2i-1}\eta_{2i}$. Therefore

(7) Alternative constructions exist in this simple case and even somewhat more generally, e.g., [3].

$$\begin{aligned}\mu_{2i+1} &= (A^{2i+1})_{00} = ((B + (A - B))^{2i+1})_{00} \\ &= (B^{2i+1})_{00} + B_{01}B_{12} \cdots B_{i-1,i}(A - B)_{ii}B_{i,i-1} \cdots B_{21}B_{10} \\ &= (B^{2i+1})_{00} + \eta_1\tilde{\eta}_1\eta_2\tilde{\eta}_2 \cdots \tilde{\eta}_{2i-1}\eta_{2i}\tilde{\eta}_{2i}\eta_{2i+1},\end{aligned}$$

and only the last term involves η_{2i+1}.

Case II. $n=2i$. Since A is $(i+1)\times(i+1)$ and B is $i\times i$, adjoin to B a zero last row and column.

$$A - B = \begin{Bmatrix} 0 & 0 & 0 \\ 0 & 0 & (\tilde{\eta}_{2i-2}\eta_{2i-1}\tilde{\eta}_{2i-1}\eta_{2i})^{1/2} \\ 0 & (\tilde{\eta}_{2i-2}\eta_{2i-1}\tilde{\eta}_{2i-1}\eta_{2i})^{1/2} & \tilde{\eta}_{2i-1}\eta_{2i} \end{Bmatrix},$$

where the zero initial row stands for $i-1$ such, and likewise for columns.

$$\begin{aligned}(A^{2i-1})_{01} &= ((B + (A - B))^{2i-1})_{01} \\ &= (B^{2i-1})_{01} + B_{01}B_{12} \cdots B_{i-2,i-1}(A - B)_{i-1,i}(A - B)_{i,i-1}B_{i-1,i-2} \cdots B_{21} \\ &= (B^{2i-1})_{01} + (\eta_1\tilde{\eta}_1\eta_2)^{1/2}\tilde{\eta}_2\eta_3 \cdots \tilde{\eta}_{2i-3}\eta_{2i-2}\tilde{\eta}_{2i-2}\eta_{2i-1}\tilde{\eta}_{2i-1}\eta_{2i},\end{aligned}$$

and by a similar argument $(A^{2i-1})_{00}=(B^{2i-1})_{00}$. Therefore

$$\begin{aligned}\mu_{2i} &= (A^{2i})_{00} = (B^{2i})_{00} + (A^{2i-1} - B^{2i-1})_{01}B_{10} \\ &= (B^{2i})_{00} + \eta_1\tilde{\eta}_1\eta_2\tilde{\eta}_2 \cdots \eta_{2i-1}\tilde{\eta}_{2i-1}\eta_{2i},\end{aligned}$$

and only the last term involves η_{2i}.

The following proposition, suggested naturally by the last two, is much harder to prove.

PROPOSITION 4. *Let the sequence $\{\eta_n\}$ terminate at η_{2m}, where $m=1/2$, 1, 3/2, 2, $\cdots$. Let $\alpha_0<\alpha_1<\cdots$ be the eigenvalues of A, and let i be such that $0<\alpha_i<1$. If $\eta_{2m}=0$, α_i is a monotone strictly increasing function of η_{2m-1}; if $\eta_{2m}=1$, α_i is a monotone strictly decreasing function of η_{2m-1}.*

Proof. Let $k=1, 2, \cdots$.

Case I. $\eta_{2k+2}=0$. Then A is $(k+1)\times(k+1)$, and η_{2k+1} appears only in $A_{kk}=\tilde{\eta}_{2k-1}\eta_{2k}+\tilde{\eta}_{2k}\eta_{2k+1}$. Therefore to increase η_{2k+1} is to add to A a positive semidefinite matrix having a strictly positive eigenvalue in a subspace not orthogonal to any eigenspace of A (see the remark following Proposition 2).

The conclusion follows in this case, by virtue of the following general property of matrices [8, p. 236].

LEMMA 4. *Let $\alpha_0<\alpha_1<\cdots<\alpha_n$ be the eigenvalues of the $(n+1)\times(n+1)$ Hermitian matrix A; let $\alpha_0'<\alpha_1'<\cdots<\alpha_n'$ be the eigenvalues of $A+B$, where $B\geqq 0$. Then $\alpha_i\leqq\alpha_i'$, with equality if and only if the eigenvector corresponding to α_i is annihilated by B.*

Now the remaining cases of Proposition 4 will be proved by reducing them similarly to Lemma 4.

Case II. $\eta_{2k+1}=0$. Again A is $(k+1)\times(k+1)$. For any value of η_{2k} it has 0 as a simple eigenvalue, so it is natural to change basis and restrict attention to the complementary k-dimensional invariant subspace.

Now in order to do this I will in a formal sense reduce the case of arbitrary k to the first case, $\eta_3=0$. Keep the given A, but regard it as a 2×2 matrix, whose top left entry A_{00} is itself a $k\times k$ matrix (in contrast to the former A_{00}), and whose bottom right entry A_{11} is the former A_{kk}. In the terminology of Theorem 3, $\mathfrak{H}_0$ is k-dimensional, $\mathfrak{H}_1$ 1-dimensional. Also by η_1 we shall now mean the *matrix* A_{00}; and $A_{01}=A_{10}^*=(\eta_1\tilde{\eta}_1)^{1/2}\eta_2^{1/2}$, $A_{11}=\eta_2^{1/2}\tilde{\eta}_1\eta_2^{1/2}$. The new η_2 is the former η_{2k}—on the subspace $\mathfrak{H}_1$, not on the whole space; η_2 is zero on $\mathfrak{H}_1^{\perp}$, so η_2 does not commute with η_1.

So regarded, A has as its domain 2-vectors with components $x_0\in\mathfrak{H}_0$ and $x_1\in\mathfrak{H}_1$. The range consists of those of the form

$$y=\begin{pmatrix}(\eta_1)^{1/2}x_0\\(\eta_2)^{1/2}(\tilde{\eta}_1)^{1/2}x_0\end{pmatrix},\qquad x_0\in\mathfrak{H}_0.$$

I have to consider the eigenvalues of A acting on such y, that is, to consider minimax values of $(Ay,y)/\|y\|^2$. This can be replaced by a minimax problem on $\mathfrak{H}_0$. Set $G=\eta_1+\tilde{\eta}_1^{1/2}\eta_2\tilde{\eta}_1^{1/2}$. Then $\|y\|^2$ reduces to (Gx_0,x_0); or $\|y\|=\|w_0\|$ if $w_0=G^{1/2}x_0$. A further computation gives $(Ay,y)=(G^2x_0,x_0)=(Gw_0,w_0)$. The problem is therefore to find the dependence on η_2 of each eigenvalue of G.

To increase η_2 to a new value η_2' is to add to $\eta_1+\tilde{\eta}_1^{1/2}\eta_2\tilde{\eta}_1^{1/2}$ a positive semi-definite matrix, hence no eigenvalue can be decreased, by Lemma 4. But can the ith eigenvalue be unchanged? Only if the ith eigenvector w_{0i} is annihilated by $\tilde{\eta}_1^{1/2}(\eta_2'-\eta_2)\tilde{\eta}_1^{1/2}$. This operator is a numerical multiple of $\tilde{\eta}_1^{1/2}\eta_2\tilde{\eta}_1^{1/2}$; hence w_{0i} would have to be an eigenvector of η_1, hence of $\tilde{\eta}_1^{1/2}$; hence w_{0i} would have to be annihilated by η_2. Suppose this is the case. Adjoining to w_{0i} a zero last component gives an eigenvector of A which is orthogonal to the last component subspace—a contradiction (see again the remark following Proposition 3). Hence all eigenvalues are strictly increasing.

Case III. $\eta_{2k+1}=1$. Apply Case II to $\tilde{A}$.

Case IV. $\eta_{2k+2}=1$. This time A is $(k+2)\times(k+2)$ and has 0 and 1 as simple eigenvalues. The proof copies Case II.

Rewrite A, letting η_1 now be $k\times k$ ($\mathfrak{H}_0$ k-dimensional), letting η_2 be the former η_{2k} (on the 1-dimensional $\mathfrak{H}_1$ only), and letting η_3 be the former η_{2k+1} (on $\mathfrak{H}_1$). Since $\eta_2\leftrightarrow\eta_3$,

$$A=\begin{pmatrix}\eta_1 & (\eta_1\tilde{\eta}_1)^{1/2}(\eta_2)^{1/2} & 0\\ (\eta_2)^{1/2}(\eta_1\tilde{\eta}_1)^{1/2} & (\eta_2)^{1/2}\tilde{\eta}_1(\eta_2)^{1/2}+\tilde{\eta}_2\eta_3 & (\tilde{\eta}_2\eta_3\tilde{\eta}_3)^{1/2}\\ 0 & (\tilde{\eta}_2\eta_3\tilde{\eta}_3)^{1/2} & \eta_3\end{pmatrix}.$$

The range of $A\tilde{A}$ consists of those vectors of the form[8]

[8] This expression is justified only by using the convention in the last paragraph of §3: $\tilde{\eta}_2$ acts on all $\mathfrak{H}$. It is nonsingular there, for $\eta_2<1$, and $\tilde{\eta}_2$ is 1 on $\mathfrak{H}_1^{\perp}$.

$$y = \begin{pmatrix} (\eta_1\tilde{\eta}_1)^{1/2}(\tilde{\eta}_2)^{1/2}x_0 \\ (\eta_2)^{1/2}(\tilde{\eta}_1 - \eta_3)(\tilde{\eta}_2)^{1/2}x_0 \\ -(\eta_2\eta_3\tilde{\eta}_3)^{1/2}x_0 \end{pmatrix}, \qquad x_0 \in \mathcal{H}_0.$$

I have to consider $(Ay, y)/\|y\|^2$ for such y. Set $H = \tilde{\eta}_2^{1/2}\eta_1\tilde{\eta}_2^{1/2} + \eta_2\tilde{\eta}_3$. Then[8] $\|y\|^2 = (H\tilde{H}x_0, x_0) = \|w_0\|^2$ if $w_0 = (H\tilde{H})^{1/2}x_0$. A longer computation yields[8] $(Ay, y) = (H^2\tilde{H}x_0, x_0) = (Hw_0, w_0)$. The problem is therefore to find the dependence on η_3 of each eigenvalue of H.

To increase η_3 to a new value η_3' is to subtract from $\tilde{\eta}_2^{1/2}\eta_1\tilde{\eta}_2^{1/2} + \eta_2\tilde{\eta}_3$ a positive semidefinite matrix which is 0 on $\mathcal{H}_1^{\perp}$. The ith eigenvalue must be strictly decreased; justifying the word "strictly" is even easier here than at the end of Case II.

From Theorem 3 and Propositions 1–4, involving the η_n, some familiar relationships between the μ_n and Φ are immediate. (In the following statements, $m = 1/2, 1, 3/2, 2, \cdots$.)

From Theorem 3 and Propositions 1 and 2—*If $\mu_1, \cdots, \mu_{2m-1}$ are the first $2m-1$ moments of any distribution, they are the first $2m-1$ moments of a distribution of degree at most m.*

From Theorem 3 and Propositions 2 and 3—*Consider the set of all distributions having $\mu_1, \cdots, \mu_{2m-1}$ as the first $2m-1$ moments. Assume the set contains no distribution of degree less than m. Then the maximum value of the 2mth moment within the set is attained for just one distribution: the unique one with degree m^*. Similarly the minimum is attained just for the unique disribution with degree m_*.*

From the foregoing with Proposition 4—*If Φ_1 and Φ_2 are different distributions, both of degree m^* (or less), or else both of degree m_* (or less), and having the same first $2m-2$ moments, then their points of nonconstancy are interlocking sets of real numbers; except that they may have 0 and-or 1 in common.*

6. **Relation between the parameters and associated determinants.** If $a_{ij} = \mu_{i+j}$ for $i, j = 0, 1, \cdots, k$, then define det $(a_{ij}) = \Delta_{*2k}$. If $a_{ij} = \mu_{i+j+1}$ for $i, j = 0, 1, \cdots, k$, then define det $(a_{ij}) = \Delta_{*2k+1}$. If $a_{ij} = \mu_{i+j+1} - \mu_{i+j+2}$ for $i, j = 0, 1, \cdots, k-1$, then define det $(a_{ij}) = \Delta_{2k}^*$. If $a_{ij} = \mu_{i+j} - \mu_{i+j+1}$ for $i, j = 0, 1, \cdots, k$, then define det $(a_{ij}) = \Delta_{2k+1}^*$.

It is known (see e.g. [5, Chap. 4]) that a finite sequence of numbers μ_n, $\mu_0 = 1$, can be the beginning of a nonextreme Hausdorff moment sequence if and only if all the Δ_n which can be formed from them are positive. Also it is clear from the definitions that each Δ_n depends only on $\mu_1, \cdots, \mu_n$; and that Δ_{*n} is linear strictly increasing in μ_n but Δ_n^* is linear strictly decreasing in μ_n. The facts concerning the η_n proved in Theorem 3 and Proposition 3 show a strong resemblance. Indeed it is a matter of elementary algebra (which I omit) to deduce from the cited facts that

$$\eta_1\tilde{\eta}_1\eta_2 \cdots \eta_{n-1}\tilde{\eta}_{n-1}\eta_n = \Delta_{*n}/\Delta_{*n-2}, \qquad \eta_1\tilde{\eta}_1\eta_2 \cdots \eta_{n-1}\tilde{\eta}_{n-1}\tilde{\eta}_n = \Delta_n^*/\Delta_{n-2}^*,$$

which may alternatively be expressed

$$\frac{\eta_n}{\tilde{\eta}_n} = \frac{\Delta_{*n}\Delta^{*}_{n-2}}{\Delta^{*}_{n}\Delta_{*n-2}} .$$

This is simple enough, yet I do not know how to prove it any more directly.

7. **Subsidiary remarks.**

REMARK 1. The representation of operators in Theorem 2 is in a sense in close analogy to their representation by the spectral theorem. Consider only the classical case (i.e., the P_i of Theorem 1 are 1-dimensional). The spectral theorem puts operator A, if it has only point spectrum, in the form of a sum of multiples of orthogonal projections. Theorem 2 puts A in the form of a sum of multiples of 1-dimensional projections Q_i, $i=1, 2, 3, \cdots$, where the Q_iQ_j are required to be zero(9), not for $i\neq j$, but only for $|i-j|>1$. I specify the Q_i: $Q_iP_j=0$ unless $j=i-1$ or $j=i$, while in the coordinate system adapted to P_{i-1} and P_i,

$$(\tilde{\eta}_{2i-2}\eta_{2i-1} + \tilde{\eta}_{2i-1}\eta_{2i})Q_i = \begin{pmatrix} \tilde{\eta}_{2i-2}\eta_{2i-1} & (\tilde{\eta}_{2i-2}\eta_{2i-1}\tilde{\eta}_{2i-1}\eta_{2i})^{1/2} \\ (\tilde{\eta}_{2i-2}\eta_{2i-1}\tilde{\eta}_{2i-1}\eta_{2i})^{1/2} & \tilde{\eta}_{2i-1}\eta_{2i} \end{pmatrix}.$$

Then $A = \sum_i (\tilde{\eta}_{2i-2}\eta_{2i-1} + \tilde{\eta}_{2i-1}\eta_{2i})Q_i$.

From this point of view Theorem 2 is not in its most general form, to be sure, because of its second sentence. Removing this limitation is no problem.

REMARK 2. Suppose the questions taken up in this paper are asked for distributions on some other set than $[0, 1]$? In order for the same considerations to apply immediately, that set must be a finite interval $[a, b]$. The parametrization of a Jacobi matrix B, $a\leqq B\leqq b$, deserves to be given explicitly. It is in terms of $\xi_1, \eta_2, \xi_3, \eta_4, \cdots$, with $a\leqq \xi_{2i+1}\leqq b$, $0\leqq \eta_{2i}\leqq 1$. These parameters may be operators in $\mathfrak{B}$, that is, this is not restricted to the classical case. Define $\breve{\xi}=a+b-\xi$. Instead of (4′), (5′), we have

$$B_{ii} = (\eta_{2i})^{1/2}\breve{\xi}_{2i-1}(\eta_{2i})^{1/2} + (\tilde{\eta}_{2i})^{1/2}\xi_{2i+1}(\tilde{\eta}_{2i})^{1/2}, \tag{4''}$$

$$B_{i-1,i} = B^{*}_{i,i-1} = (\tilde{\eta}_{2i-2})^{1/2}(\xi_{2i-1}\breve{\xi}_{2i-1} - ab)^{1/2}(\eta_{2i})^{1/2}, \tag{5''}$$

$i, j=0, 1, 2, \cdots$; and $B_{ij}=0$ for $|i-j|>1$. (Again, $\eta_0=0$ is the convention.) For $a=0$, $b=1$, this reduces to (4′), (5′), with the obvious notational equivalences.

REMARK 3. The canonical form of Theorem 3 suggests as a byproduct still another parametrization of all Hausdorff moment sequences. Define $x_1=\eta_1-\tilde{\eta}_1$, and for $n>1$,

$$x_n = 2^{n-1}\{\eta_1\tilde{\eta}_1 \cdots \eta_{n-1}\tilde{\eta}_{n-1}\}^{1/2}(\eta_n - \eta_n). \tag{13}$$

Alternatively, this could be written in terms of angles $\theta_n\in[0, \pi]$ defined by $\cos\theta_n=\eta_n-\tilde{\eta}_n$; the advantage would be the more "geometric" formulas $x_1=\cos\theta_1$, and for $n>1$, $x_n=\sin\theta_1 \cdots \sin\theta_{n-1}\cos\theta_n$.

(9) In the unconventional terminology of Guttman [2], the Q_i (or rather vectors in their respective ranges) must form a "perfect simplex."

In any case, the facts are easy to prove: By Theorem 3, Hausdorff moment sequences correspond 1-1 to sequences of numbers $\eta_n \in [0, 1]$, $n=1, 2, \cdots$, provided any sequence is regarded as terminating at the first η_n which is equal to 0 or 1. But (13) gives for each such sequence a unique sequence of real numbers x_n, $n=1, 2, \cdots$. It is easy to verify from the identity $(\eta-\tilde{\eta})^2+4\eta\tilde{\eta}=1$, that $\sum x_n^2 \leqq 1$. Conversely, x_n real and $\sum x_n^2 \leqq 1$ imply that (13) may be solved successively for η_n until some η_n is 0 or 1. The unit ball (unit sphere and its interior) in real Hilbert space of countably infinite dimensionality has been put in 1-1 correspondence with Hausdorff moment sequences, hence (see e.g. [13, Theorem 6.1]) with distribution functions on [0, 1].

$\eta_n=0$ or 1 if and only if $x_{n+1}=x_{n+2}= \cdots =0$. Such points lie on, but are far from exhausting, the unit sphere in the space. The corresponding distributions are exactly those concentrated in a finite number of points.

The distribution Φ corresponding to the center of the sphere may be shown (using [5, §25]) to be given by

$$d\Phi(t) = \frac{dt}{\pi(t(1-t))^{1/2}} \cdot$$

Bibliography

1. C. Davis, *Theorems on projections in Hilbert space*, Bull. Amer. Math. Soc. Abstract 60-2-218.

2. L. Guttman, *A new approach to factor analysis: the radex*, Mathematical Thinking in the Social Sciences, ed. by P. Lazarsfeld, Glencoe, The Free Press, 1954, pp. 258–348.

3. H. Hadwiger, *Über ausgezeichnete Vektorsterne und reguläre Polytope*, Comment. Math. Helv. vol. 13 (1940) pp. 90–107.

4. P. R. Halmos, *Normal dilations and extensions of operators*, Summa Brasiliensis Math. vol. 2 (1950) pp. 125–134.

5. S. Karlin and L. S. Shapley, *Geometry of moment spaces*, Memoirs of the American Mathematical Society, no. 12, Providence, 1953.

6. B. Nagy, *A moment problem for self-adjoint operators*, Acta Mathematica Academiae Scientiarum Hungaricae vol. 3 (1952) pp. 285–293.

7. M. A. Neumark, *On a representation of additive operator set functions*, Doklady Akademii Nauk SSSR. vol. 41 (1943) pp. 359–361.

8. F. Riesz and B. Nagy, *Leçons d'analyse fonctionelle*, 2d ed., Budapest, 1953.

9. J. A. Shohat and J. D. Tamarkin, *The problem of moments*, Mathematical Surveys, no. 1, New York, 1943.

10. W. F. Stinespring, *Positive functions on C*-algebras*, Proc. Amer. Math. Soc. vol. 6 (1955) pp. 211–216.

11. M. H. Stone, *Linear transformations in Hilbert space and their applications to analysis*, Amer. Math. Soc. Colloquium Publications, vol. 15, New York, 1932.

12. A. Wald, *Limits of a distribution function determined by absolute moments and inequalities satisfied by absolute moments*, Trans. Amer. Math. Soc. vol. 46 (1939) pp. 280–306.

13. D. V. Widder, *The Laplace transform*, Princeton, 1946.

Columbia University,
New York, N. Y.