The Annals of Probability
1993, Vol. 21, No. 3, 1295–1309

A NORMAL LIMIT THEOREM FOR MOMENT SEQUENCES

By Fu-Chuen Chang, J. H. B. Kemperman[1] and W. J. Studden[2]

Purdue University, Rutgers University and Purdue University

Let Λ be the set of probability measures λ on $[0,1]$. Let $M_n = \{(c_1, \dots, c_n) | \lambda \in \Lambda\}$, where $c_k = c_k(\lambda) = \int_0^1 x^k \, d\lambda$, $k = 1, 2, \dots$ are the ordinary moments, and assign to the moment space M_n the uniform probability measure P_n. We show that, as $n \to \infty$, the fixed section $(c_1, \dots, c_k)$, properly normalized, is asymptotically normally distributed. That is, $\sqrt{n}[(c_1, \dots, c_k) - (c_1^0, \dots, c_k^0)]$ converges to MVN$(0, \Sigma)$, where c_i^0 correspond to the arc sine law λ_0 on $[0,1]$. Properties of the $k \times k$ matrix Σ are given as well as some further discussion.

1. Introduction and main theorem. The set of probability measures on $[0,1]$ is denoted as Λ. Let further

$$M_n = \{(c_1, \dots, c_n) | \lambda \in \Lambda\}, \tag{1.1}$$

where $c_k = c_k(\lambda) = \int_0^1 x^k \lambda(dx)$, $k = 0, 1, 2, \dots,$ $c_0 = 1$. This so-called moment space M_n is the convex hull of the curve $\{(x, x^2, \dots, x^n)\colon 0 \le x \le 1\}$ in $\mathbb{R}_n$ and is a very small compact subset of the unit cube $[0,1]^n$. For instance, it is known that

$$V_n = \text{Vol } M_n = \prod_{k=1}^{n} B(k,k) = \prod_{k=1}^{n} \frac{\Gamma(k)\Gamma(k)}{\Gamma(2k)} \tag{1.2}$$

[see Karlin and Studden (1966), page 129, Theorem 6.2] (another proof is given below). Thus V_n is roughly of size 2^{-n^2}, more precisely, $\log V_n \approx -n^2 \log 2$ as $n \to \infty$.

Our investigations stem from an attempt to understand more fully the shape and structure of M_n by looking, in some sense, at a typical point of M_n. Let P_n be the uniform probability measure on M_n, that is, $dP_n = dx/V_n$ is n-dimensional Lebesgue measure on M_n normalized by the volume of M_n. In this way $(c_1, \dots, c_n) \in M_n$ can now be viewed as a random vector. The symbol E_n will indicate expected values relative to P_n.

For example, M_2 is determined by the inequalities $c_1^2 \le c_2 \le c_1 \le 1$ and has volume $V_2 = 1/6$, thus $dP_2 = 6\, dc_1\, dc_2$ on M_2. The marginal densities of c_1, c_2 are $6(c_1 - c_1^2)$, $0 < c_1 < 1$, and $6(\sqrt{c_2} - c_2)$, $0 < c_2 < 1$, respectively. The means are $E_2[c_1] = 1/2$ and $E_2[c_2] = 2/5$ and the squared correlation is

Received June 1991; revised April 1992.
[1]Partially supported by NSF Grant DMS-90-0-2856.
[2]Partially supported by NSF Grant DMS-91-01730.
AMS 1991 *subject classifications*. Primary 60F05, 30E05; secondary 60D05, 60J15, 33C45.
Key words and phrases. Moment spaces, canonical moments, normal limit, random walk.

35/38. General closed form expressions even for, say, the means $E_n[c_k]$ seem difficult to obtain.

The so-called center $(c_1^0, \dots, c_n^0)$ of the moment space M_n is given by

$$c_k^0 = \int_0^1 x^k f_0(x)\,dx = 2^{-2k}\binom{2k}{k} \approx \frac{1}{\sqrt{\pi k}} \quad \text{as } k \to \infty. \tag{1.3}$$

Here, $f_0(x) = \pi^{-1}x^{-1/2}(1-x)^{-1/2}$, $0 < x < 1$, is the density of the arc sine probability measure λ_0 on $[0,1]$. The word "center" will become clearer below. Our main result is the following.

THEOREM 1.1. *As $n \to \infty$, the distribution of $\sqrt{n}[(c_1, \dots, c_k) - (c_1^0, \dots, c_k^0)]$ relative to P_n converges to a multivariate normal distribution* MVN$(0, \Sigma_k)$. *Here, $\Sigma_k = (1/2)A_k A_k'$ with A_k as the lower triangular $k \times k$ matrix defined by*

$$a_{ij} = \begin{cases} 2^{-2i+1}\binom{2i}{i-j}, & \text{if } 1 \le j \le i, \\ 0, & \text{if } j > i; \end{cases} \tag{1.4}$$

thus $a_{ii} = 2^{-2i+1}$. In particular, if c_k is governed by P_n and $n \to \infty$, then $c_k \to c_k^0$ in probability.

By $A = (a_{ij};\ 1 \le i, j < \infty)$ we will denote the corresponding infinite lower triangular matrix, having A_k as its left upper $k \times k$ submatrix. The proof of the theorem is, in essence, quite simple and, at the same time, illuminating. The boundary of M_n has P_n-measure zero and thus can be ignored. Note that $(c_1, \dots, c_n) \in \operatorname{int} M_n$ implies that $(c_1, \dots, c_k) \in \operatorname{int} M_k$ for all $k \le n$.

It will be convenient to employ the canonical coordinates p_k, $k = 1, 2, \dots$, introduced by Skibinsky (1967). For each $k = 1, \dots, n$, the kth canonical coordinate p_k of a moment point $(c_1, \dots, c_n) \in \operatorname{int} M_n$ is well defined, satisfies $0 < p_k < 1$, and depends only on $c_1, \dots, c_k$. The associated function $p_k = f_k(c_1, \dots, c_k)$ is independent of n. Conversely, c_k is fully determined by $p_1, \dots, p_k$.

Given $(c_1, \dots, c_{k-1}) \in M_{k-1}$, let $c_k^+ = c_k^+(c_1, \dots, c_{k-1})$ and $c_k^- = c_k^-(c_1, \dots, c_{k-1})$, respectively, denote the largest and smallest possible value of c_k which is compatible with $(c_1, \dots, c_{k-1}, c_k) \in M_k$. Thus, $c_k^- \le c_k \le c_k^+$ when $(c_1, \dots, c_k) \in M_k$. In particular, $c_1^- = 0$; $c_1^+ = 1$ and $c_2^- = c_1^2$; $c_2^+ = c_1$. As is easily seen, $(c_1, \dots, c_k) \in \operatorname{int} M_k$ if and only if $c_j^- < c_j < c_j^+$, $j = 1, \dots, k$. Put

$$\Delta_k = \Delta_k(c_1, \dots, c_{k-1}) = c_k^+(c_1, \dots, c_{k-1}) - c_k^-(c_1, \dots, c_{k-1}).$$

Here, $\Delta_k > 0$ for all $(c_1, \dots, c_{k-1}) \in \operatorname{int} M_{k-1}$. For $k = 1, \dots, n$, the kth canonical coordinate (or moment) of a moment point $(c_1, \dots, c_n) \in \operatorname{int} M_n$ is defined by

$$p_k = \frac{c_k - c_k^-}{c_k^+ - c_k^-}; \quad \text{thus } c_k = c_k^-(c_1, \dots, c_{k-1}) + \Delta_k(c_1, \dots, c_{k-1})p_k. \tag{1.5}$$

Note that $0 < p_k < 1$. It follows by induction that, for all $k \geq 1$, there is a 1 : 1 correspondence between points $(c_1, \dots, c_k) \in \operatorname{int} M_k$ and points $(p_1, \dots, p_k) \in (0,1)^k$. Thus c_k^-, c_k^+ and $\Delta_k = c_k^+ - c_k^-$ can also be regarded as functions of $p_1, \dots, p_{k-1}$; these functions happen to be polynomial [as is clear from (3.6) or (3.19)]. Similarly, c_k is a polynomial in $p_1, \dots, p_k$ which is linear in the variable p_k with coefficient Δ_k [see (1.5)]. The canonical moments p_k for the Beta(α, β) distribution on [0, 1] are given in Skibinsky [(1969), page 1759]. The above arc sine distribution λ_0 corresponds to $\alpha = \beta = 1/2$ and has canonical moments $p_k^0 = 1/2$ for all $k \geq 1$. This partially explains why the corresponding moment point $(c_1^0, \dots, c_n^0)$ may be regarded as the center of M_n. Here, the c_k^0 are as in (1.3).

REMARK. The canonical coordinates p_k admit a more general interpretation and as such are quite robust. Namely, consider any nondegenerate compact interval $[a, b]$ and let $\{W_j(x)\}_{j=1}^{\infty}$ be a given system of polynomials of the form $W_j(x) = \sum_{m=0}^{j} d_{jm} x^m$ with $d_{jj} > 0$. For example, $W_j(x) = x^j$. Next consider all moment sequences $\{w_j\}_{j=1}^{\infty}$ of the form $w_j = \int W_j(x)\lambda(dx)$, $j = 1, 2, \dots,$ with λ as a probability measure on $[a, b]$. Given the moments $w_1, \dots, w_{n-1}$, let w_n^-, w_n^+ denote the smallest and largest possible value of w_n. Provided $\Delta_n = w_n^+ - w_n^- > 0$, define $p_n = (w_n - w_n^-)/\Delta_n$; thus $0 < p_n < 1$. As is easily seen, the resulting sequence $\{p_n\}$ of (generalized) canonical coordinates is independent of the particular choice of the system of polynomials $\{W_j(x)\}$. In addition, as was already observed by Skibinsky [(1969), page 1763, Theorem 5] if the probability measure λ on $[a, b]$ is linearly transformed (with positive slope) to a measure μ on another interval $[\alpha, \beta]$ then λ and μ have exactly the same canonical coordinates p_n $n \geq 1$. Here, $\mu(F) = \lambda(g^{-1}F)$, where $g(x) = \alpha + (\beta - \alpha)(x - a)/(b - a)$.

Let us return to the above (Hausdorff) sequences $\{c_n\}$ of the special form $c_n = \int x^n \lambda(dx)$, with λ as a probability measure on [0, 1]. Using (1.5), one finds that

$$\frac{\partial c_k}{\partial p_j} = \begin{cases} 0, & \text{if } j > k, \\ \Delta_k = c_k^+ - c_k^- = \prod_{r=1}^{k-1} p_r q_r, & \text{if } j = k. \end{cases} \tag{1.6}$$

Here and from now on, $q_r = 1 - p_r$. The latter elegant formula for Δ_k was established by Skibinsky (1967). A different proof is given below; see (3.4). It follows from (1.6) that

$$\frac{\partial(c_1, \dots, c_n)}{\partial(p_1, \dots, p_n)} = \prod_{k=1}^{n} \frac{\partial c_k}{\partial p_k} = \prod_{r=1}^{n-1} (p_r q_r)^{n-r}. \tag{1.7}$$

Transforming the integral $V_n = \int_{M_n} dc_1 \cdots dc_n$ to an integral over $(0,1)^n$ relative to the p_j, we see that formula (1.2) is an immediate consequence of (1.7). Both (1.2) and (1.7) are special cases of the following result (namely, with $m = 0$ and $m = n - 1$, respectively).

THEOREM 1.2. *Let* $0 \le m < n$ *and* $(c_1, \dots, c_m) \in \operatorname{int} M_m$. *Then the set* $M_n(c_1, \dots, c_m)$ *of all* $(c_{m+1}, \dots, c_n)$ *such that* $(c_1, \dots, c_n) \in M_n$ *has* $(n - m)$-*dimensional volume*

$$\text{(1.8)} \qquad \operatorname{Vol} M_n(c_1, \dots, c_m) = \prod_{r=1}^{m} (p_r q_r)^{n-m} \prod_{k=2}^{n-m} \frac{\Gamma(k)\Gamma(k)}{\Gamma(2k)}.$$

The latter is maximal when $p_r = 1/2$, $r = 1, \dots, m$. *Note that under* P_n *the conditional distribution of* $(c_{m+1}, \dots, c_n)$ *given* $(c_1, \dots, c_m)$ *is the uniform distribution* $dc_{m+1} \cdots dc_n / \operatorname{Vol} M_n(c_1, \dots, c_m)$ *on* $M_n(c_1, \dots, c_m)$.

In the sequel, for each fixed n, when we assign to M_n the uniform distribution P_n, functions on M_n such as $c_1, \dots, c_k$ or $p_1, \dots, p_k$, $k \le n$, can be regarded as random variables. But note that the resulting joint distribution will depend on n.

PROOF. Prescribing $(c_1, \dots, c_m) \in \operatorname{int} M_m$ is the same as prescribing the parameters $0 < p_r < 1$, $r = 1, \dots, m$. Further note, using (1.6), that

$$\frac{\partial(c_{m+1}, \dots, c_n)}{\partial(p_{m+1}, \dots, p_n)} = \prod_{s=m+1}^{n} \prod_{r=1}^{s-1} p_r q_r = \prod_{r=1}^{m} (p_r q_r)^{n-m} \prod_{r=m+1}^{n-1} (p_r q_r)^{n-r}.$$

The volume on hand is equal to the integral of $dc_{m+1} \cdots dc_n$ over $M_n(c_1, \dots, c_m)$. Transforming that integral to an integral with respect to the variables $p_{m+1}, \dots, p_n$ over the unit cube $(0,1)^{n-m}$, one obtains (1.8). □

THEOREM 1.3. *The uniform probability measure* P_n *on* M_n *is equivalent to the first n canonical coordinates* $p_1, \dots, p_n$ *being independent random variables in such a way that* p_k *has a symmetric* $\text{Beta}(\alpha_k, \alpha_k)$ *distribution with* $\alpha_k = n - k + 1$, $k = 1, \dots, n$.

PROOF. Simply transform the integral

$$E_n f(p_1, \dots, p_n) = \int_{M_n} f(p_1, \dots, p_n)\, dc_1 \cdots dc_n / V_n,$$

where f is arbitrary, to the variables $p_1, \dots, p_n$, again using (1.7). □

The symmetric distribution $\text{Beta}(\alpha, \alpha)$, $\alpha > 0$, has mean $1/2$ and variance $1/(8\alpha + 4)$. Hence, for $k = 1, \dots, n$, letting $\alpha = n - k + 1$,

$$\text{(1.9)} \quad E_n[p_k] = \frac{1}{2}, \qquad \operatorname{Var}[p_k] = \frac{1}{8(n - k + 3/2)} = \frac{1}{8n} + O\left(\frac{1}{n^2}\right),$$

as $n \to \infty$. Moreover, as is well known and easily seen, $\sqrt{n}\,[p_k - 1/2] \to N(0, 1/8)$ in distribution under P_n as $n \to \infty$. Two proofs of the following central lemma are given in Section 3.

LEMMA 1.4. *The first order Taylor expansion of* $c_k = c_k(p_1, \dots, p_k)$ *about the center* $(p_1^0, \dots, p_k^0)$ *with* $p_j^0 = 1/2$ *is given by*

$$(1.10) \qquad c_k = c_k^0 + 2 \sum_{m=1}^{k} a_{km}(p_m - \tfrac{1}{2}) + O\left(\sum_{m=1}^{k} |p_m - \tfrac{1}{2}|^2 \right).$$

Here, the a_{km} *are as in* (1.4). *In particular* $a_{km} = 2^{-2k+1}\binom{2k}{k-m}$ *if* $m \le k$.

PROOF OF THEOREM 1.1. Let k be fixed and $j, m = 1, \dots, k$. With $n \ge k$ and relative to P_n as the underlying measure, consider the random variables $X_{nj} = \sqrt{n}(c_j - c_j^0)$ and $Z_{nm} = 2\sqrt{n}(p_m - 1/2)$. Here, $Z_{n1}, \dots, Z_{nk}$ are independent, for each fixed n, while $Z_{nm} \to N(0, 1/2)$ when m is fixed and $n \to \infty$. Writing (1.10) as

$$X_{nj} = \sum_{m=1}^{k} a_{jm} Z_{nm} + O\left(\frac{1}{\sqrt{n}} \sum_{m=1}^{k} Z_{nm}^2 \right), \qquad j = 1, \dots, k,$$

Theorem 1.1 becomes an immediate consequence. □

2. Further discussion. Let Σ be the infinite symmetric matrix $\Sigma = (\sigma_{ij}) = (1/2)AA'$ having $\Sigma_k = (1/2)A_k A_k'$ as its left upper $k \times k$ submatrix. Recall that Σ_k is the covariance matrix of the asymptotic MVN$(0, \Sigma_k)$ distribution as $n \to \infty$ of $\sqrt{n}[(c_1, \dots, c_k) - (c_1^0, \dots, c_k^0)]$, when the latter is governed by the uniform measure P_n on M_n. Thus asymptotically, as $n \to \infty$, the c_i have means $c_i^0 + o(1)$ and covariances $(\sigma_{ij}/n)(1 + o(1))$. Let further

$$\rho_{ij} = \sigma_{ij} / \sqrt{\sigma_{ii}\sigma_{jj}}.$$

Thus ρ_{ij} is the limiting value as $n \to \infty$ of the correlation coefficient under P_n between the moments c_i and c_j. The following result is proved in Section 4.

LEMMA 2.1. *One has*

$$(2.1) \qquad \sigma_{ij} = c_{i+j}^0 - c_i^0 c_j^0,$$

where the c_k^0 *are as in* (1.3). *Hence,* $\sigma_{ij} \to 0$ *as* $i, j \to \infty$. *If* s *is fixed then* $\rho_{s,s+r} \to 0$ *as* $r \to \infty$. *If* r *is fixed, then* $\rho_{s,s+r} \to 1$ *as* $s \to \infty$. *More generally, for any fixed* $k \ge 0$,

$$(2.2) \qquad \rho_{ij} \to \left(\frac{4K}{(K+1)^2} \right)^{1/4} \quad \text{when } i, j \to \infty, \ \frac{j}{i} \to K.$$

Let $k \ge 1$ be fixed. It is natural to inquire into the diagonalization of the symmetric $k \times k$ matrix Σ_k and corresponding linear transformations of $(c_1, \dots, c_k)$. In view of the usual Gram–Schmidt orthogonalization procedure, it suffices to determine the essentially unique linear combinations $t_i = b_{i1}c_1 + \cdots + b_{ii}c_i$, $1 \le i \le k$, with $b_{ii} \ne 0$ that are asymptotically uncorrelated under P_n as $n \to \infty$. Equivalently, letting $b_{im} = 0$ when $m > i$, we want

$B_k = (b_{im};\ i, m = 1, \dots, k)$ to be a nonsingular lower triangular $k \times k$ matrix such that $D_k = B_k \Sigma_k B_k'$ is diagonal. Adding suitable constants b_{i0}, one can further achieve that

$$t_i = \sum_{m=0}^{i} b_{im} c_m, \qquad i = 1, \dots, k;\ c_0 = 1 \tag{2.3}$$

are asymptotically uncorrelated and of mean 0. Equivalently, letting $t_0 = c_0 = 1$, we want $t_0, t_1, \dots, t_k$ to be asymptotically orthogonal under P_n as $n \to \infty$.

The preceding diagonalization process happens to be intimately connected with the usual Chebyshev polynomials. Namely, consider the probability space Ω_0 consisting of the interval $[0, 1]$ together with the arc sine measure λ_0 as the underlying probability measure. The functions $x \to x^i$ on Ω_0 can then be regarded as random variables Z_i. We see from (1.3) that $EZ_i = c_i^0$ and $EZ_i Z_j = c_{i+j}^0$. Therefore,

$$\operatorname{Cov}(Z_i, Z_j) = \sigma_{ij} \quad \text{for all } i, j \ge 1, \tag{2.4}$$

with σ_{ij} exactly as in (2.1). Hence, the means and covariances of $\sqrt{n}\,(c_i - c_i^0)$, $i = 1, \dots, k$, under P_n coincide asymptotically (as $n \to \infty$) with the means and covariances of $Z_i - c_i^0$, $i = 1, \dots, k$. Thus the above diagonalization is equivalent to finding $k + 1$ linear combinations of the form $T_i^{\#} = \sum_{m=0}^{i} b_{im} Z_m$, $i = 0, 1, \dots, k$, with $b_{ii} \ne 0$; $b_{00} = 1$, that are orthogonal as random variables on Ω_0. But that simply means that the corresponding polynomials

$$T_i^*(x) = \sum_{m=0}^{i} b_{im} x^m, \qquad i = 0, 1, 2, \dots, \tag{2.5}$$

one of each degree, are orthogonal with respect to the arc sine measure λ_0. Choosing the leading coefficient b_{ii} appropriately, we may as well assume that the $T_i^*(x)$ are precisely the Chebyshev polynomials, adapted to the interval $[0, 1]$. And then the resulting coefficients b_{im} are independent of k [where $k \ge \max(i, m)$].

The functions $\cos i\theta$, $i = 0, 1, 2, \dots$ are clearly orthogonal with respect to the uniform measure on $[0, \pi]$. Letting $y = \cos\theta$, $\cos i\theta = T_i(y)$ one arrives at the system $\{T_i(y)\}_{i=0}^{\infty}$ of ordinary Chebyshev polynomials, orthogonal with respect to the measure $dy/\sqrt{1 - y^2}$ on $(-1, 1)$. Letting $x = (1 + y)/2 = (1 + \cos\theta)/2 = (\cos\theta/2)^2$ leads to the desired system

$$T_i^*(x) = T_i(2x - 1), \qquad i = 0, 1, \dots \tag{2.6}$$

as in (2.5) of orthogonal polynomials with respect to the measure λ_0 on $(0, 1)$. Here, $T_i^*(x)$ is of exact degree i, while $T_0^*(x) \equiv 1$. The coefficients in (2.5) are given by $b_{i0} = (-1)^i$ and

$$\begin{aligned} b_{im} &= (-1)^{i+m} 2^{2m-1} \frac{i}{m} \binom{i + m - 1}{i - m}, \\ &= (-1)^{i+m} 2^{2m} \frac{i}{i + m} \binom{i + m}{i - m}, \quad \text{if } 1 \le m \le i. \end{aligned} \tag{2.7}$$

Thus $b_{ii} = 2^{2i-1}$ if $i \geq 1$. Further, from now on, $b_{im} = 0$ if $m > i$. Formula (2.7) easily follows from the known result that $T_n(2x - 1) = (-1)^n F(-n, n; 1/2, x)$ [see Abramowitz and Stegun (1965), page 795 and Henrici (1977), page 176]. For the sake of completeness, an independent proof of (2.7) is included in Section 4. Further note that

$$\int_0^1 T_j^*(x)^2 \lambda_0(dx) = \int_0^\pi (\cos j\theta)^2 \frac{d\theta}{\pi} = \frac{1}{2}. \tag{2.8}$$

THEOREM 2.2. *Consider the linear combinations*

$$t_i = \sum_{m=0}^{i} b_{im} c_m = \sum_{m=1}^{i} b_{im}\left(c_m - c_m^0\right), \qquad i = 1, 2, \dots; c_0 = 1. \tag{2.9}$$

Here the b_{im} are as in (2.5) *and* (2.7). *Then, for any fixed $k \geq 1$ and $n \to \infty$, the distribution of $\sqrt{n}\,(t_1, \dots, t_k)$ relative to P_n converges in distribution to the multivariate normal distribution $MVN(0, (1/2)I_k)$. Here, I_k denotes the $k \times k$ identity matrix.*

PROOF. The second equality sign in (2.9) follows from $c_0 = c_0^0 = 1$ and

$$t_i^0 = \sum_{m=0}^{i} b_{im} c_m^0 = \int_0^1 T_i^*(x) \lambda_0(dx) = 0 \quad \text{if } i \geq 1. \tag{2.10}$$

In view of Theorem 1.1, it suffices to show that $B_k \Sigma_k B_k' = (1/2)I_k$. In some sense this already follows from the previous discussion. As a direct proof, if $1 \leq i, j \leq k$, then

$$\begin{aligned}\sum_{r=0}^{k} \sum_{s=0}^{k} b_{ir} b_{js}\left(c_{i+j}^0 - c_i^0 c_j^0\right) &= \sum_{r=0}^{k} \sum_{s=0}^{k} b_{ir} b_{js} c_{i+j}^0 \\ &= \int_0^1 T_i^*(x) T_j^*(x) \lambda_0(dx) = \tfrac{1}{2}\delta_i^j.\end{aligned}$$

Here, we used (2.5), (2.8) and (2.10) as well as the orthogonality of the $T_i^*(x)$ with respect to λ_0. Note that $c_{i+j}^0 - c_i^0 c_j^0 = 0$ when either $i = 0$ or $j = 0$. In view of (2.1), it follows that $B_k \Sigma_k B_k' = (1/2)I_k$. □

THEOREM 2.3. *The lower triangular matrices $A = (a_{ij}, i, j \geq 1)$ and $B = (b_{ij}; i, j \geq 1)$ are each other's inverse. Similarly for A_k and B_k (any $k \geq 1$). Moreover, for $m \geq 1$,*

$$x^m = c_m^0 + \sum_{r=1}^{m} a_{mr} T_r^*(x). \tag{2.11}$$

COROLLARY 2.4. *We have for all $m, r \geq 1$ that*

$$\int_0^1 x^m T_r^*(x) \lambda_0(dx) = \tfrac{1}{2} a_{mr}. \tag{2.12}$$

Moreover,

$$c_m = c_m^0 + \sum_{r=1}^{m} a_{mr} t_r. \tag{2.13}$$

Here, the t_r *are as in* (2.3); *thus* $t_r = \int T_r^*(x)\lambda(dx)$.

We will present several proofs. Note that (2.12) is an immediate consequence of (2.8), (2.11) and the orthogonality of the $T_r^*(x)$ with respect to λ_0. Further, (2.13) follows from (2.11) from an integration relative to any $\lambda \in \Lambda$ having the moments $c_0 = 1, c_1, \dots, c_m$. Choosing $\lambda = \lambda_0$, one has $c_m = c_m^0$, $m \ge 0$ and $t_r = 0$, $r \ge 1$; $t_0 = 1$. This explains the constant term c_m^0 in (2.11) and (2.13). Finally observe that (2.11) is actually *equivalent* to A, B being each other's inverse, as can be seen by substituting formula (2.5) for the $T_r^*(x)$ into (2.11) and equating coefficients.

A first proof of Theorem 2.3 amounts to a direct verification of (2.11); see Section 4. A second proof is to directly verify the property $AB = I$; see Section 4. As still another demonstration, recall that, in the proof of Theorem 2.2, we already established that $B\Sigma B' = (1/2)I$ where $\Sigma = (1/2)AA'$. Hence, the *lower triangular* matrix $C = BA$ satisfies $CC' = I$, in particular, the rows of C are mutually orthogonal. Also using that $c_{ii} = a_{ii}b_{ii} = (2^{-2i+1})(2^{2i-1}) = 1$, we conclude that C must be the identity matrix.

3. Proof of Lemma 1.4. We will present two different proofs. The first one exploits an important relation between the Hausdorff moment problem and a certain random walk. This relation, which one of us plans to discuss in more detail in a subsequent paper, is implicit in the work of Karlin and McGregor (1959).

Let $\{X_n\}_{n=0}^{\infty}$ be a stationary discrete time Markov chain (also called random walk) on the nonnegative integers Z_+ which is determined by the transition probabilities

$$P(X_{n+1} = j | X_n = i) = \begin{cases} p_i, & \text{if } j = i - 1, \\ q_i, & \text{if } j = i + 1, \\ 0, & \text{otherwise.} \end{cases} \tag{3.1}$$

Here, $q_i = 1 - p_i$. Further $0 < p_i < 1$ for $i \ge 1$, while $p_0 = 0$; $q_0 = 1$. The corresponding n-step probabilities are denoted as $P_{ij}^{(n)} = P(X_n = j | X_0 = i)$. It was shown by Karlin and McGregor [(1959), page 69] that there exists a necessarily unique probability measure λ of infinite support on $[0, 1]$ such that

$$P_{00}^{(2n)} = P(X_{2n} = 0 | X_0 = 0) = \int_0^1 x^n \lambda(dx), \quad \text{for all } n \ge 0. \tag{3.2}$$

In other words,

$$c_n = P_{00}^{(2n)}, \qquad n = 0, 1, \dots \tag{3.3}$$

always defines a Hausdorff moment sequence having $c_0 = 1$; $(c_1, \dots, c_n) \in \operatorname{int} M_n$ for all $n \ge 1$. In fact, (3.2) establishes a $1:1$ correspondence between all such Hausdorff moment sequences $\{c_n\}$ on the one hand and all random walks $\{X_n\}$ on the other hand, each random walk being determined as above by a sequence $\{p_n\}_{n=1}^{\infty}$ of canonical coordinates, $0 < p_n < 1$.

Consider a random walk $\{X_n\}$ as above and define c_n as in (3.3). Conditional on $X_0 = 0$, the conditional probability $c_k = P_{00}^{(2k)}$ (to be back in state 0 after $2k$ steps, not necessarily for the first time), obviously depends only on the parameters $p_1, \dots, p_k$. Fixing $c_1, \dots, c_k$ is equivalent to fixing $p_1, \dots, p_k$. Hence, for given $c_1, \dots, c_{n-1}$, the smallest and largest possible value c_n^- and c_n^+ of $c_n = P_{00}^{(2n)}$ is realized by choosing $p_n = 0$ or $p_n = 1$, respectively. In fact, c_n^- represents the (common) part of the return probability $c_n = P_{00}^{(2n)}$ arising from paths of length $2n$ (from 0 back to 0 in $2n$ steps) which never reach state n, and thus have their probability as a function of $p_1, \dots, p_{n-1}$, independent of p_n. Similarly, $c_n^+ - c_n^-$ is equal to the probability $q_1 q_2 \cdots q_{n-1} p_n p_{n-1} \cdots p_1$ of the *single* path which leads from 0 to 0 in $2n$ steps which does reach state n. Maximizing c_n given $p_1, \dots, p_{n-1}$, that is, choosing $p_n = 1$, this reduces to

$$(3.4) \quad \Delta_n := c_n^+ - c_n^- = q_1 q_2 \cdots q_{n-1} p_{n-1} p_{n-2} \cdots p_1 = \prod_{r=1}^{n-1} p_r q_r > 0.$$

Finally note that $c_n = P_{00}^{(2n)} = c_n^- + p_n(c_n^+ - c_n^-)$. Comparing the latter with (1.5), we conclude that, for all $n \geq 1$, the random walk parameter p_n *coincides* with the nth canonical coordinate of the moment point $(c_1, \dots, c_n) \in \operatorname{int}(M_n)$.

First proof of Lemma 1.4. Let $\{c_n\}_{n=0}^{\infty}$ be a Hausdorff moment sequence and $\{p_r\}_{r=1}^{\infty}$ be the associated sequence of canonical coordinates. Let $r \geq 1$ be fixed and

$$(3.5) \qquad C_n(r) = \left[\frac{\partial}{\partial p_r} c_n\right]_0 = \left[\frac{\partial}{\partial p_r} P_{00}^{(2n)}\right]_0.$$

The subscript zero here indicates that $p_k = p_k^0 = 1/2$, for all $k \geq 1$. We want to show that $C_n(r) = 2a_{nr}$ with a_{nr} as in (1.4). In the present proof, we exploit the above random walk interpretation. Hence,

$$(3.6) \qquad c_n = P_{00}^{(2n)} = \sum p_1^{m_1} p_2^{m_2} \cdots q_0^{n_0} q_1^{n_1} \cdots,$$

where we sum over all paths $x = (x_0, x_1, \dots, x_{2n})$ with $x_k - x_{k-1} = \pm 1$, $k = 1, \dots, 2n$, and such that $x_0 = 0$; $x_{2n} = 0$; (thus c_n is a polynomial of degree $2n - 1$ in terms of $p_1, \dots, p_n$). Further, for each such path, m_j, $j \geq 1$ and n_j, $j \geq 0$, respectively, will denote the number of transitions $x_{k-1} \to x_k$, $k = 1, \dots, 2n$, of type $j \to j - 1$ and $j \to j + 1$, respectively. Differentiating the latter sum with respect to p_r causes an extra factor $m_r/p_r - n_r/q_r$. Setting afterwards $p_k = 1/2$ for all $k \geq 1$, we find that

$$(3.7) \qquad C_n(r) = 2E((m_r - n_r) I_0(X_{2n}) | X_0 = 0),$$

where $I_0(x)$ is the indicator function on the set $\{0\}$. Here, and from now on in the present proof, $\{X_n\}$ will be the simple random walk on Z_+ having one-step probabilities $p_k = q_k = 1/2$ for all $k \geq 1$ (while $p_0 = 0$; $q_0 = 1$). Moreover, since the path $\{X_0, X_1, \dots, X_{2n}\}$ is random, so are the associated transition numbers m_j and n_j.

Let further $\{Y_n\}_{n=0}^{\infty}$ be the classical random walk on $Z = \{0, \pm 1, \pm 2, \dots\}$ with independent increments such that $P(Y_n - Y_{n-1} = -1) = P(y_n - Y_{n-1} = +1) = 1/2$. For each $s \in Z$, let

$$D_n(s) = E\big[(m_s - n_s) I_0(Y_{2n}) | Y_0 = 0\big]. \tag{3.8}$$

Here, m_s and n_s, respectively, denote the (random) number of transitions $Y_{k-1} \to Y_k$, $k = 1, \dots, 2n$, of the forms $s \to s-1$ and $s \to s+1$, respectively.

Identifying the states j and $-j$ (for all j), the process $\{Y_n\}$ reduces precisely to the above simple random walk $\{X_n\}$. And it easily follows from (3.7) that

$$\tfrac{1}{2} C_n(r) = D_n(r) - D_n(-r) = 2D_n(r). \tag{3.9}$$

We further claim that

$$D_n(r) = P(Y_{2n} = 0;\ Y_k = r \text{ for some } 0 \leq k \leq 2n | Y_0 = 0). \tag{3.10}$$

After all, consider any fixed path $y = (y_0, y_1, \dots, y_{2n})$ with $y_k - y_{k-1} = \pm 1$, $k = 1, \dots, 2n$ and $y_0 = 0$; $y_{2n} = 0$. Since $r \geq 1$ such a path y can contribute to $D_n(r)$ only when $y_k = r$ for *some* $0 \leq k \leq n$. Let k_1 and k_2 be the minimal and maximal such index k. Thus, $0 < k_1 \leq k_2 < 2n$ and further $y_{k_1} = y_{k_2} = r$; $y_{k_1 - 1} = y_{k_2 + 1} = r - 1$. Given such a path y, consider the associated (partially reflected) path y^* obtained from y by replacing y_k by $y_k^* = 2r - y_k$ for all $k_1 < k < k_2$ (leaving the other coordinates y_k unchanged). Thus $(y^*)^* = y$, while $y^* = y$ if and only if $k_1 = k_2$.

For each fixed index k with $k_1 \leq k < k_2$, a possible contribution ± 1 to the value $(m_r - n_r)(y^*)$ (for the reflected path y^*), due to a pair $y_k = r$, $y_{k+1} = r \pm 1$, is exactly opposite in sign to the corresponding contribution to the value $(m_r - n_r)(y)$ (for the original path y). Hence, since y and y^* have the same probability 2^{-2n}, one may as well ignore all such contributions, in which case there only remains the single contribution $+1$ to $(m_r - n_r)(y)$ due to the single pair $y_{k_2} = r$; $y_{k_2+1} = r - 1$. This completes the proof of (3.10).

It now follows from (3.9), (3.10) and (1.4) that

$$C_n(r) = 4D_n(r) = 4P(Y_{2n} = 2r | Y_0 = 0) = 4\binom{2n}{n-r} 2^{-2n} = 2a_{nr}.$$

Here, we also used the standard André reflection principle. Namely, associate to each path y as above, of length $2n$ which begins and ends at 0 and meets state r at least once, the path y^* having $y_k^* = 2r - y_k$ when $k \geq k_1$ while $y_k^* = y_k$, otherwise. This sets up a 1 : 1 correspondence with the set of paths y^* of length $2n$ which begin at 0 and end at $2r$. This completes the proof of Lemma 1.4. □

SECOND PROOF OF LEMMA 1.4. Skibinsky (1968, 1969) showed that the mapping from the canonical moments p_i to the power moments c_i is given by the following formulae. Here $q_i = 1 - p_i$, $i \geq 1$, $\zeta_i = p_i q_{i-1}$, $i \geq 1$; thus

$\zeta_1 = p_1$. Define $S_{ij} = 0$ unless $0 \le i \le j$. Further S_{ij}, $0 \le i \le j$, is recursively defined by $S_{0j} \equiv 1$, $j \ge 0$, and

$$S_{ij} = S_{i,j-1} + \zeta_{j-i+1} S_{i-1,j} \quad \text{if } 1 \le i \le j. \tag{3.11}$$

Thus the case $j = i$ reduces to $S_{ii} = \zeta_1 S_{i-1,i}$. The moments c_n themselves are finally given by $c_n = S_{nn}$, $n \ge 0$. Note that S_{ij} is independent of the p_r with $r > j$.

For j and n as integers and $n \ge 0$, define

$$Q_j^n = 2^{-n}\binom{n}{m} \quad \text{if } n = |j| + 2m \text{ with } m = 0, 1, 2, \dots \tag{3.12}$$

and $Q_j^n = 0$ in all other cases. Note from (1.4) that $a_{nr} = 2Q_{2r}^{2n}$. As is easily seen,

$$Q_j^n = \tfrac{1}{2}\left(Q_{j-1}^{n-1} + Q_{j+1}^{n-1}\right) \quad \text{and} \quad Q_{-j}^n = Q_j^n \quad \text{thus } Q_0^n = Q_1^{n-1}. \tag{3.13}$$

Let further S_{ij}^0 denote the value S_{ij} in the special case that $p_k = 1/2$ for all $k \ge 1$. Using (3.13), it follows from (3.11) by induction that

$$S_{ij}^0 = 2^{j-i} Q_{j-i}^{i+j} \quad \text{if } 0 \le i \le j. \tag{3.14}$$

For instance, $S_{ii}^0 = Q_0^{2i} = Q_1^{2i-1} = \zeta_1 S_{i-1;i}^0$ with $\zeta_1 = p_1 = 1/2$.

Let $r \ge 1$ be fixed, and introduce

$$U_{ij} = 2^{i-j-1}\left[\frac{\partial}{\partial p_r} S_{ij}\right]_0 \quad \text{for } k \ge 1.$$

Thus $U_{ij} = 0$ unless $0 \le i \le j$ and $r \le j$. Moreover, $U_{0j} \equiv 0$ since $S_{0j} \equiv 1$. We want to show that $[(\partial/\partial p_r)c_n]_0 = 2a_{nr}$. In view of $c_n = S_{nn}$ and $a_{nr} = 2Q_{2r}^{2n}$, this is equivalent to $U_{nn} = 2Q_{2r}^{2n}$. More generally, we will show that, for all $0 \le i \le j$,

$$U_{ij} = \begin{cases} Q_{j-i+2r}^{i+j}, & \text{if } j - i \ge r \ge 1, \\ Q_{j-i+2r}^{i+j} + Q_{i-j+2r}^{i+j}, & \text{if } 0 \le j - i < r. \end{cases} \tag{3.15}$$

For instance $U_{ii} = 2Q_{2r}^{2i}$ and $U_{i-1,i} = Q_{2r+1}^{2i-1} + Q_{2r-1}^{2i-1}$, $r \ge 2$; $U_{i-1,i} = Q_3^{2i-1}$ if $r = 1$.

Differentiating the recursion formula (3.11) with respect to p_r at $p_k = 1/2$ (all $k \ge 1$) and using (3.14), one finds that the U_{ij} satisfy the recursion relation

$$U_{ij} - \tfrac{1}{2}\left(U_{i,j-1} + U_{i-1,j}\right) = \begin{cases} -\tfrac{1}{2} Q_{r+1}^{i+j-1}, & \text{if } j - i = r, \\ -\tfrac{1}{2} Q_r^{i+j-1}, & \text{if } j - i = r - 1, \\ 0, & \text{otherwise,} \end{cases} \tag{3.16a}$$

as long as $1 \le i < j$. The case $j = i$ is of the form

$$U_{i,i} = U_{i-1,i} + \delta_r^1 Q_1^{2i-1}. \tag{3.16b}$$

The recursion (3.16) and boundary condition $U_{0j} \equiv 0$ together completely determine the U_{ij}. Using (3.13), one easily verifies that U_{ij}, $0 \le i \le j$, as defined by the right-hand side of (3.15), does indeed satisfy (3.16) and $U_{0j} \equiv 0$. This establishes (3.15) and completes the second proof of Lemma 1.4. □

REMARKS. Formula (3.11) for the S_{ij}, which furnishes a recursive calculation of $c_n = S_{nn}$ from the canonical coordinates p_i, also follows from a simple random walk argument. In fact, the S_{ij} have the simple probabilistic interpretation (3.18).

Namely, let $\{X_n\}$ be the random walk on Z_+ described by (3.1), with the p_j as the usual canonical coordinates. We know that $c_n = P_{00}^{(2n)}$, for all $n \ge 0$. Clearly, $P_{0j}^{(n)} = P(X_n = j | X_0 = 0)$ satisfy $P_{0j}^{(0)} = \delta_j^0$ and

$$P_{0k}^{(n)} = P_{0;k-1}^{(n-1)} q_{k-1} + P_{0;k+1}^{(n-1)} p_{k+1}, \tag{3.17}$$

$n \ge 1$; $k \ge 0$; $q_{-1} = 0$. This allows us to calculate the $P_{0k}^{(n)}$ in a recursive manner. For instance, $c_n = P_{00}^{(2n)} = p_1 P_{01}^{(2n-1)}$. Since $P_{0k}^{(n)} = 0$ if $n < k$, (3.17) is trivially satisfied when $n < k$. Also note that $P_{0k}^{(k)} = q_0 q_1 \cdots q_{k-1}$. All terms in (3.17) vanish unless $n = k + 2i$ with $i \in Z_+$, in which case $n = i + j$; $k = j - i$ with $0 \le i \le j$ as integers. It follows from (3.17) that the S_{ij} defined by

$$S_{ij} = \frac{1}{q_0 q_1 q_2 \cdots q_{j-i-1}} P_{0;j-i}^{(i+j)} \quad \text{for } 0 \le i \le j, \tag{3.18}$$

$q_0 = 1$, satisfy the recursion relation (3.11). Moreover, $S_{0k} = P_{0k}^{(k)} / q_0 q_1 \cdots q_{k-1} = 1$, for all $k \ge 0$. Finally, $c_n = P_{00}^{(2n)} = S_{nn}$.

In view of the interpretation (3.18) of the S_{ij}, formula (3.15) can also be regarded as an explicit formula for the quantities $[(\partial/\partial p_r) P_{0j}^{(n)}]_0$, equivalently, as an explicit formula for $E[(m_r - n_r)(X_n = j) | X_0 = 0]$, with m_r, n_r as in (3.7).

Theorem 2 in Skibinsky (1968) also has a simple probabilistic proof. It states that

$$c_n = \sum_{0 \le i \le n/2} (S_{i,n-i})^2 \prod_{j=1}^{n-2i} \zeta_j. \tag{3.19}$$

In fact, paying attention to the value $X_n = k$ (say),

$$c_n = P(X_{2n} = 0 | X_0 = 0) = \sum_k P_{0k}^{(n)} P_{k0}^{(n)} = \sum_k \frac{1}{\pi_k} (P_{0k}^{(n)})^2. \tag{3.20}$$

Here, $\pi_k = q_0 q_1 \cdots q_{k-1} / p_1 p_2 \cdots p_k$, $\pi_0 = 1$. We also used the well known relation $\pi_i P_{ij}^{(n)} = \pi_j P_{ji}^{(n)}$ [all i, j, n; see, e.g., Karlin and McGregor (1959), page 68]. Noting that $P_{0k}^{(n)}$ vanishes unless $k = n - 2i$ with $0 \le i \le n/2$, and using (3.18), one easily verifies that (3.19) and (3.20) are equivalent.

4. Further proofs.

PROOF OF LEMMA 2.1. Let $i, j \geq 1$. From $\Sigma = (1/2)AA'$ and $a_{kr} = 0$ for $r > k$, one has

$$\sigma_{ij} = \frac{1}{2} \sum_{r=1}^{\min(i,j)} a_{ir} a_{jr} = 2^{-2i-2j+1} \sum_{r=1}^{\min(i,j)} \binom{2i}{i-r} \binom{2j}{j-r}$$

$$= -c_i^0 c_j^0 + \sum_{r=-\min(i,j)}^{\min(i,j)} 2^{-2i} \binom{2i}{i-r} 2^{-2j} \binom{2j}{j+r} = -c_i^0 c_j^0 + c_{i+j}^0,$$

proving (2.1). After all, the latter sum is equal to the coefficient of z^{i+j} in the expansion of $((1+z)/2)^{2i}((1+z)/2)^{2j}$.

Recall that $c_k^0 \approx 1/\sqrt{\pi k}$ as $k \to \infty$. Hence, $\sigma_{jj} = c_{2j}^0 - (c_j^0)^2 \approx (2\pi j)^{-1/2}$ and $\sigma_{ij} = c_{i+j}^0 - c_i^0 c_j^0 \approx (1 - c_i^0)(\pi j)^{-1/2}$ as $j \to \infty$. Thus, for i fixed and $j \to \infty$,

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}}\sqrt{\sigma_{jj}}} \approx D_i j^{-1/4}, \quad \text{where } D_i = \left(\frac{\pi}{2}\right)^{-1/4} (1 - c_i^0)(\sigma_{ii})^{-1/2}.$$

In particular, $\rho_{s,s+r} \to 0$ as $r \to \infty$. If both i and j tend to infinity, then

$$\sigma_{ij} = c_{i+j}^0\left(1 - c_i^0 c_j^0 / c_{i+j}^0\right) \approx c_{i+j}^0 \approx 1/\sqrt{\pi(i+j)}.$$

Here we used that $c_i^0 c_j^0 / c_{i+j}^0 \approx [1/\pi((1/i) + (1/j))]^{1/2} \to 0$. Hence, if $i, j \to \infty$ in such a way that $j/i \to K$ then

$$\rho_{ij} \approx \left[\frac{4ij}{(i+j)^2}\right]^{1/4} \to \left(\frac{4K}{(K+1)^2}\right)^{1/4}. \qquad \square$$

PROOF OF (2.7). We want to prove that the coefficients b_{im} in (2.5) are given by (2.7). Letting $y = \cos\theta = 2x - 1$, one has $\cos n\theta = T_n(y) = T_n^*(x)$; thus

$$\sum_{n=0}^{\infty} T_n^*(x) u^n = \sum_{n=0}^{\infty} \cos n\theta\, u^n = \operatorname{Re}\left[\sum_{n=0}^{\infty} (e^{i\theta} u)^n\right] = \operatorname{Re} \frac{1}{1 - ue^{i\theta}}$$

$$= \frac{1 - u\cos\theta}{1 + u^2 - 2u\cos\theta} = \frac{1 + u - 2ux}{(1+u)^2 - 4ux}$$

$$= (1 + u - 2ux) \sum_{r=0}^{\infty} (4ux)^r (1+u)^{-2r-2}.$$

The coefficient of x^m is found to be $2^{2m-1} u^m (1-u)(1+u)^{-2m-1}$. Expanding the latter in powers of u, we find that the coefficient of u^n is precisely b_{nm} as given by (2.7). $\square$

PROOF OF THE IDENTITY (2.11). This identity must be known. Recall that $T_r^*(x) = \cos r\theta$ when $x = (\cos \theta/2)^2$. If $m \geq 1$, then

$$x^m = \left(\cos \frac{\theta}{2}\right)^{2m} = 2^{-2m}\left(e^{i\theta/2} + e^{-i\theta/2}\right)^{2m} = 2^{-2m} \sum_{j=0}^{2m} \binom{2m}{j} \cos(m - j)\theta.$$

The term with $j = m$ gives rise to $2^{-2m}\binom{2m}{m} = c_m^0$. Further, for $r = 1, \dots, m$, the two terms with $j = m \pm r$ together give rise to $2^{-2m+1}\binom{2m}{m-r}\cos r\theta = a_{mr}T_r^*(x)$, in view of (1.4). This proves (2.11). □

PROOF THAT $AB = I$ (see Theorem 2.3). Here A, B are lower triangular, hence also $C = AB$. Further $c_{ii} = a_{ii}b_{ii} = 1$ thus it suffices to show that $c_{im} = 0$ when $1 \leq m < i$. From (1.4) and (2.7),

$$c_{im} = \sum_{j=m}^{i} a_{ij}b_{jm} = \sum_{j=m}^{i} 2^{-2i+1}\binom{2i}{i+j}(-1)^{j+m}2^{2m-1}\frac{j}{m}\binom{j+m-1}{2m-1}.$$

This can be written as $c_{im} = \sum_{j=m}^{i}(-1)^j\binom{2i}{i+1}g(j)$, where

$$g(x) = \alpha x\binom{x+m-1}{2m-1} = \frac{\alpha x^2}{(2m-1)!}\prod_{r=1}^{m-1}(x+r)(x-r),$$

with $\alpha = \alpha_{im}$ as a constant factor. Note that $g(x)$ is an *even* polynomial of degree $2m$ such that $g(r) = 0$ for $r = 0, \pm 1, \dots, \pm (m-1)$. Hence, letting $i + j = s$,

$$2c_{im} = \sum_{j=-i}^{i}(-1)^j\binom{2i}{i+j}g(j) = \sum_{s=0}^{2i}(-1)^{s-i}\binom{2i}{s}g(s-i)$$

$$= (-1)^i\Delta^{2i}g(-i) = 0,$$

since g is of degree $2m < 2i$. Here $\Delta = E - 1$ is the usual difference operator; thus $(Eg)(x) = g(x+1)$. □

Acknowledgments. Many thanks to Burgess Davis for simplifying our original random walk proof of Lemma 1.4.

REFERENCES

ABRAMOWITZ, M. and STEGUN, I. A. (1965). *Handbook of Mathematical Functions*. Dover, New York.

HENRICI, P. (1977). *Applied and Computational Complex Analysis* **2**. Wiley, New York.

KARLIN, S. and MCGREGOR, J. (1959). Random walks. *Illinois J. Math* **3** 66–81.

KARLIN, S. and STUDDEN, W. J. (1966). *Tchebycheff Systems: With Applications in Analysis and Statistics*. Wiley, New York.

LAU, T. S. (1983). Theory of canonical moments and its application in polynomial regression, Parts I and II. Technical Reports 83-23, 83-24, Dept. Statistics, Purdue Univ.

RIVLIN, T. J. (1974). *The Chebyshev Polynomials*. Wiley, New York.

SKIBINSKY, M. (1967). The range of the $(n + 1)$th moment for distributions on $[0, 1]$. *J. Appl. Probab.* **4** 543–552.

SKIBINSKY, M. (1968). Extreme n-th moments for distributions on $[0, 1]$ and the inverse of a moment space map. *J. Appl. Probab.* **5** 693–701.

SKIBINSKY, M. (1969). Some striking properties of binomial and beta moments. *Ann. Math. Statist.* **40** 1753–1764.

F.-C. CHANG
W. J. STUDDEN
DEPARTMENT OF STATISTICS
PURDUE UNIVERSITY
WEST LAFAYETTE, INDIANA 47907

J. H. B. KEMPERMAN
DEPARTMENT OF STATISTICS
RUTGERS UNIVERSITY
NEW BRUNSWICK, NEW JERSEY 08903