JOURNAL OF APPROXIMATION THEORY 65, 32–45 (1991)

Some Properties of Markov Systems*

R. A. ZALIK

Department of Algebra, Combinatorics, and Analysis, Auburn University, Auburn, Alabama 36849-5307, U.S.A.

AND

D. ZWICK*

Department of Mathematics and Statistics, University of Vermont, Burlington, Vermont 05401, U.S.A.

Communicated by D. S. Lubinsky

Received May 30, 1989; revised January 20, 1990

We demonstrate some new properties of Markov systems, involving generalized divided differences, relative differentiation, and weak nondegeneracy.

0. INTRODUCTION AND STATEMENT OF MAIN RESULTS

In this paper, A denotes a subset of the real line with at least $n+2$ elements, and $I(A)$ denotes the convex hull of A. A is said to satisfy *property B* if between any two distinct points of A there is another point of A. If, in addition, A contains neither a first nor a last element (i.e., $\inf A \notin A$, $\sup A \notin A$) then A is said to satisfy *property D*. The numbers $\inf A$ and $\sup A$ are called the *endpoints* of A.

A sequence of functions $Z_n = \{z_0, ..., z_n\}$ defined on A is called a (weak) *Tchebycheff system* if it is linearly independent and for all points $x_0 < \cdots < x_n$ in A, $\det\{z_i(x_j)\}_{i,j=0}^{n} > 0$ ($\geqslant 0$). If Z_k is a (weak) Tchebycheff system for $k = 0, ..., n$, we say that Z_n is a (weak) *Markov system.* Note that, in this case, $z_0 > 0$ ($z_0 \geqslant 0$). If $z_0 \equiv 1$, we say that Z_n is *normalized.* In the following definitions, when we say that a basis $U_n = \{u_0, ..., u_n\}$ is obtained from Z_n by a triangular linear transformation, we mean that $u_0 = z_0$ and $u_k - z_k \in \operatorname{span}\{z_0, ..., z_{k-1}\}$ ($k = 1, ..., n$).

* The second author was supported in part by an NSF grant to VT-EPSCoR. The authors thank the Centro Atómico Bariloche, Argentina, for their hospitality during the completion of this paper, and the referees for their careful reading of the paper and suggestions for improvements.

0021-9045/91 $3.00

DEFINITION 1. Z_n is said to satisfy *condition E* if for all $c \in I(A)$ the following two requirements are satisfied:

(a) If Z_n is linearly independent on $[c, \infty) \cap A$ then there exists a basis $\{u_0, ..., u_n\}$ for span(Z_n), obtained by a triangular linear transformation, such that for any sequence of integers $0 \leqslant k(0) < \cdots < k(m) \leqslant n$, $\{u_{k(r)}\}_{r=0}^{m}$ is a weak Markov system on $A \cap [c, \infty)$.

(b) If Z_n is linearly independent on $(-\infty, c] \cap A$ then there exists a basis $\{v_0, ..., v_n\}$ for span(Z_n), obtained by a triangular linear transformation, such that for any sequence of integers $0 \leqslant k(0) < \cdots < k(m) \leqslant n$, $\{(-1)^{r-k(r)} v_{k(r)}\}_{r=0}^{m}$ is a weak Markov system on $(-\infty, c] \cap A$.

DEFINITION 2. Z_n is said to satisfy *condition I* if for every real number c, Z_n is linearly independent on at least one of the sets $(-\infty, c) \cap A$ and $A \cap (c, \infty)$.

DEFINITION 3. Z_n is called *weakly nondegenerate* if it satisfies both conditions I and E.

DEFINITION 4. Z_n is *representable* if and only if, for all $c \in A$, there is a basis U_n, obtained from Z_n by a triangular linear transformation (hence, $u_0(x) = z_0(x)$); a strictly increasing function h (an "embedding function") defined on A, with $h(c) = c$; and a set $W_n = \{w_1, ..., w_n\}$ of continuous, increasing functions defined on $I(h(A))$, such that

$$u_1(x) = u_0(x) \int_c^{h(x)} dw_1(t_1)$$

$$\vdots$$

$$u_n(x) = u_0(x) \int_c^{h(x)} \int_c^{t_1} \cdots \int_c^{t_{n-1}} dw_n(t_n) \cdots dw_1(t_1).$$

In this case we say that (h, c, W_n, U_n) is a *representation* of Z_n.

We note that [2, Theorem 5.22] proves that a normalized weak Markov system is representable if and only if it satisfies condition E.

In the sequel, given a set W_n as above, we will define functions

$$p_0(x) \equiv 1,$$

and (1)

$$p_i(x) = \int_c^{x} \int_c^{t_1} \cdots \int_c^{t_{i-1}} dw_i(t_i) \cdots dw_1(t_1) \; (i = 1, ..., n).$$

The purpose of this paper is to solve a number of problems motivated by the previous work of one of the authors. For example, it is shown in [5] that every weakly nondegenerate weak Markov system is representable. However, an example is also given of a representable system that is not weakly nondegenerate. Thus, the question naturally arises as to what conditions, in addition to representability, must be imposed to obtain a necessary and sufficient condition. In Theorem 1 we answer this question for Markov systems, but first we introduce an important definition.

DEFINITION 5. Let $W_n = \{w_1, ..., w_n\}$ be a sequence of real-valued functions defined on (a, b), let h be a real-valued function defined on an $A \subset \mathbb{R}$ with $h(A) \subset (a, b)$, and let $x_0 < \cdots < x_n$ be points of $h(A)$. We say that W_n satisfies *property M* with respect to h at $x_0 < \cdots < x_n$ if there is a sequence $\{t_{i,j}: i = 0, ..., n;\ j = 0, ..., n-i\}$ in $h(A)$ such that

(a) $x_j = t_{0,j} (j = 0, ... n)$;

(b) $t_{i,j} < t_{i+1,j} < t_{i,j+1}$ $(i = 0, ..., n-1;\ j = 0, ..., n-i)$;

(c) For $i = 1, ..., n$, $w_i(x)$ is not constant at $t_{i,j}$ $(j = 0, ..., n-i)$.

To say that a function f is not constant at a point $c \in (a, b)$ is to say that for every $\varepsilon > 0$ there are points $x_1, x_2 \in (a, b)$ with $c - \varepsilon < x_1 < c < x_2 < c + \varepsilon$, such that $f(x_1) \neq f(x_2)$.

If W_n satisfies property M for every choice of points $x_0 < \cdots < x_n$ in $h(A)$ then we simply say that W_n satisfies property M with respect to h on A.

THEOREM 1. *Suppose that A has neither a first nor a last element. Then the following statements are equivalent*:

(a) *Z_n is a weakly nondegenerate normalized Markov system*;

(b) *Z_n is representable, and for every representation (h, c, W_n, U_n) and any $d \in \mathbb{R}$ there is a sequence $x_0 < \cdots < x_n$ in A for which W_n satisfies property M with respect to h on $(-\infty, d) \cap A$, or else there is a sequence $y_0 < \cdots < y_n$ in A for which W_n satisfies property M with respect to h on $A \cap (d, \infty)$.*

Remark 1. It follows from (b) $\Rightarrow$ (a) in Theorem 1 that if (b) is satisfied for some representation then it must be satisfied for all representations.

DEFINITION 6. Let $w_0, ..., w_n$ be continuous on an open interval I, with $w_0 > 0$ and $w_1, ..., w_n$ strictly increasing in I. Let f be a real-valued function defined on I. For $x \in I$, set

$$D_0 f(x) = \frac{f(x)}{w_0(x)}$$

and, provided the limits exist,

$$D_k f(x) = \lim_{h \to 0} \frac{D_{k-1} f(x+h) - D_{k-1} f(x)}{w_k(x+h) - w_k(x)} \qquad (k = 1, ..., n).$$

Set $W_k^0 = \{w_0, ..., w_k\}$ and let $D(W_k^0, I)$ denote the set of functions f for which $D_0 f, ..., D_k f$ exist in I. If $f \in D(W_k^0, I)$ we say that f is *relatively differentiable* with respect to W_k^0.

Remark 2. f is an element of $D(W_1^0, I)$ if and only if $(f/w_0) \circ w_1^{-1}$ is differentiable in $w_1(I)$ and $((f/w_0) \circ w_1^{-1})'(w_1(x)) = D_1 f(x)$. Also note that if $w_0(x) \equiv 1$ and the functions $p_i(x)$ are given by (1), then $D_{i-1}\, p_i(x) = w_i(x) - w_i(c)$ $(i = 1, ..., n)$. This is an easy consequence of Lemma 1, below.

DEFINITION 7. Let Z_n be a Tchebycheff system on A, let $x_0, ..., x_n$ be distinct points of A, and let f be a real-valued function defined on A. The (generalized) divided difference of f of order n with respect to Z_n is defined as

$$\begin{bmatrix} z_0, ..., z_n \\ x_0, ..., x_n \end{bmatrix} f = \frac{\begin{vmatrix} z_0(x_0) & \cdots & z_0(x_n) \\ \vdots & & \vdots \\ z_{n-1}(x_0) & \cdots & z_{n-1}(x_n) \\ f(x_0) & \cdots & f(x_n) \end{vmatrix}}{\begin{vmatrix} z_0(x_0) & \cdots & z_0(x_n) \\ \vdots & & \vdots \\ z_{n-1}(x_0) & \cdots & z_{n-1}(x_n) \\ z_n(x_0) & \cdots & z_n(x_n) \end{vmatrix}}$$

(for $n = 0$ this reduces to $\left[\begin{smallmatrix} z_0 \\ x_0 \end{smallmatrix}\right] f = f(x_0)/z_0(x_0)$).

The next theorem generalizes a result that is well known for extended complete Tchebycheff systems (see [1]).

THEOREM 2. *Let $W_k^0 = \{w_0, ..., w_k\}$ be as in Definition* 6 *and set $v_i = w_0 \cdot p_i$, with p_i given by* (1). *If $f \in D(W_k^0, I)$, then for all $x_0 < \cdots < x_k$ in I,*

$$\begin{bmatrix} v_0, ..., v_i \\ x_0, ..., x_i \end{bmatrix} f = D_i f(\xi_i) \qquad (i = 0, ..., k),$$

where $\xi_0 = x_0$, and $x_0 < \xi_i < x_i$ $(i = 1, ..., k)$.

COROLLARY 1. *Under the assumptions of Theorem* 2,

$$\lim_{x_0, \ldots, x_{i-1} \to \xi} \begin{bmatrix} v_0, \ldots, v_{i-1} \\ x_0, \ldots, x_{i-1} \end{bmatrix} v_i = D_{i-1} v_i(\xi) = w_i(\xi) - w_i(c) \qquad (i = 1, \ldots, n).$$

In [6] the following theorem is proved:

THEOREM A. *If A contains neither a first nor a last element, then Z_n is a Markov system if and only if it has a representation (h, c, W_n, U_n) such that W_n satisfies property M with respect to h on A.*

We say that the span of a Markov system Z_n defined on a set A can be *continued to the left* if there is an n-dimensional linear space U defined on a set of the form $(d, a) \cup A$, $d < a$ (where $a = \inf A$), such that $U|_A = \operatorname{span}(Z_n)$ and U has a basis U_n that is a Markov system (i.e., U_n is a *Markov space*).

Remark 3. Z_n is automatically weakly nondegenerate if it is a Markov system and if A has no first or last element: Condition I is satisfied and condition E follows from the possibility of extending Z_n both to the left and to the right of any $c \in A$ (see the proof of Theorem 1 for details).

The situation is different if A has a first or a last element. Our next result is based on the concepts of generalized divided difference and relative differentiation.

THEOREM 3. *Let Z_n be a Markov system on a set A with property B, and assume that if* $\inf A \in A$ *or* $\sup A \in A$, *then they are accumulation points of A and all the z_i are continuous there. Then the following statements are equivalent*:

(a) *Z_n is a weakly nondegenerate Markov system*;

(b) *If* $\inf A \in A$, *then Z_n can be extended to the left, and if* $\sup A \in A$, *then Z_n can be extended to the right*;

(c) *If d is an endpoint of A such that $d \in A$, then*

$$\limsup_{x_0, \ldots, x_{i-1} \to d} \left| \begin{bmatrix} z_0, \ldots, z_{i-1} \\ x_0, \ldots, x_{i-1} \end{bmatrix} z_i \right| < \infty \qquad (i = 1, \ldots, n);$$

(d) *If* $\inf A \in A$, *then Z_n has a representation (h, c, W_n, U_n) such that the w_i are bounded from below on $h(A)$, and if* $\sup A \in A$, *then Z_n has a representation (h, c, W_n, U_n) such that the w_i are bounded from above on $h(A)$.*

(e) *If (h, c, W_n, U_n) is a representation for Z_n on $A' = A \setminus \{\inf A, \sup A\}$, $w_0 \equiv 1$, and p_i are defined as in* (1), *then for any endpoint d of $h(A)$ such that $d \in h(A)$,* $\lim_{x \to d} D_{i-1} p_i(x)$ *is finite.*

In [3, Theorems 2.2 and 2.6], part (d) of Theorem 3 was shown under hypotheses similar to those in (a).

1. *E*-Systems and Proof of Theorem 1

Definition 8. We say that Z_n is a (weak) *E-system* if for any integers $0 \leqslant r(0) < \cdots < r(m) \leqslant n$, $\{z_{r(k)}\}_{k=0,\dots,m}$ is a (weak) Markov system on A. The linear span of a (weak) E-system will be called a (weak) *E-space*.

E-systems were utilized in [3] and in [7] to give a necessary and sufficient condition for extending the linear span of a Markov system beyond its domain of definition. For example, the following theorem is proved in [7]:

Theorem B. *Let Z_n be a Markov system on a bounded set A with property B. Assume, further, that if an endpoint of A belongs to A then it is a point of accumulation of A and all the elements of Z_n are continuous there. Then* $\operatorname{span}(Z_n)$ *can be continued to the left if and only if it has a basis that is an E-system.*

Assume that Z_n is a Markov system defined on a set A containing both of its endpoints, and that all its elements are continuous at these endpoints. Assume, moreover, that the linear span S_n of Z_n contains a basis that is an E-system. From Theorem B we know that S_n can be continued to a Markov space U_n defined on a set of the form $(d, a) \cup A$, $d < a$. Thus, the restriction of U_n to any set of the form $(d', a) \cup A$, with $d < d' < a$, can be continued to the left, and by a second application of Theorem B we conclude that U_n is an E-space. Thus, if U_n denotes the linear space obtained from U_n by making the change of variable $t \to -t$, from [7, Remark 2] we readily conclude that U_n^- is an E-space. Applying Theorem B to S_n, we conclude that U_n can be continued to the right (cf. [7, Corollary 2]). The foregoing discussion demonstrates that under the conditions of Theorem B, an E-system can be simultaneously continued both to the left and to the right. Conclusions similar to these can also be found in [3].

Proof of Theorem 1. (a) $\Rightarrow$ (b) From [5, Theorem 1], Z_n is representable. Let (h, c, W_n, U_n) be a representation for Z_n. Then $u_i = p_i \circ h$ $(i = 0, \dots, n)$, where p_i are defined as in (1). Define $P_n = \{p_0, \dots, p_n\}$, and let d be an arbitrary real number. P_n is linearly independent either on $h(A) \cap (d, \infty)$ or on $(-\infty, d) \cap h(A)$, and we will assume the latter. By [6, Lemma] it must satisfy property M (with respect to the identity function) at some $x_0 < \cdots < x_n$ in $(-\infty, d) \cap h(A)$, hence (b) follows.

(b) $\Rightarrow$ (a) Let (h, c, W_n, U_n) be a representation for Z_n and let $d \in \mathbb{R}$

be given. If property M is satisfied with respect to h for, say, some choice of points in $(-\infty, d) \cap A$, then from [6, Lemma] U_n(and, hence, Z_n) is linearly independent on $(-\infty, d) \cap A$. Thus, condition I is satisfied. An argument similar to the one given in [7, Theorem 2] shows that U_n is a weak Markov system. Since U_n is obtained from Z_n by a triangular linear transformation, it is clearly normalized.

If A has neither a first nor a last element, then on $A \cap (c, \infty)$, U_n can be continued to the left, whence from [7, Theorem 2], it satisfies part (a) of condition E. Moreover, since on $(-\infty, c) \cap A$, U_n can be continued to the right, from [7, Corollary 3] we deduce that is also satisfies part (b) of condition E. ∎

Remark 4. Suppose that if $\inf A \in A$, it is an accumulation point of A, and v_1/v_0 is continuous at $\inf A$, where v_0 and v_1 are as in Theorem 2. Since

$$\frac{v_1}{v_0}(x) = w_1(h(x)) - w_1(c),$$

we have

$$w_1(h(x)) = \frac{v_1}{v_0}(x) + w_1(c).$$

Thus,

$$h(x) = w_1^{-1}(w_1(h(x))) = w_1^{-1}\left(\frac{v_1}{v_0}(x) + w_1(c)\right),$$

hence h is continuous at $\inf A$. A similar statement holds for $\sup A$.

2. Relative Differentiation and Proof of Theorem 2

The proof of Theorem 2 is based on the following two lemmas.

Lemma 1. *Let f and g be relatively differentiable with respect to W_1^0, where $w_0 \equiv 1$, and assume that $D_1 g \neq 0$ on (a, b). Then*

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{D_1 f(\xi)}{D_1 g(\xi)}$$

for some $a < \xi < b$, and

$$D_1 f(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{w_1(x+h) - w_1(x)}.$$

Proof. Let $f_1 = f \circ w_1^{-1}$, $c = w_1(a)$, and $d = w_1(b)$. Then

$$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(\xi)}{g'(\xi)} = \frac{D_1 f_1(\xi_0)}{D_1 g_1(\xi_0)},$$

where $\xi_0 = w_1^{-1}(\xi)$. ∎

Remark 5. Lemma 1 is false if $w_0 \not\equiv 1$ (e.g., if $w_0(x) = w_1(x) = f(x) = x$, $g(x) = x^2$).

LEMMA 2. *Let f be relatively differentiable with respect to W_1^0. Then for $k \geq 1$, there are points $x_{i-1} < \xi_i < x_i$ $(i = 1, \ldots, k)$ such that*

$$\begin{bmatrix} v_0, \ldots, v_k \\ x_0, \ldots, x_k \end{bmatrix} f = \begin{bmatrix} D_1 v_1, \ldots, D_1 v_k \\ \xi_1, \ldots, \xi_k \end{bmatrix} D_1 f.$$

Proof. In the representation of $\left[\begin{smallmatrix} v_0, \ldots, v_k \\ x_0, \ldots, x_k \end{smallmatrix}\right] f$ as a ratio of determinants, first divide each column in both numerator and denominator by the value of $v_0(t)$ that corresponds to that column (this has the effect of making $w_0 \equiv 1$), then subtract from the last column in both numerator and denominator the preceding one. The resulting quotient has the same value as the original divided difference. Thus, the divided difference may be expressed in the form

$$\frac{F(x_k) - F(x_{k-1})}{G(x_k) - G(x_{k-1})},$$

where F and G are relatively differentiable with respect to W_1.

By Lemma 1 this equals $D_1 F(\xi_k)/D_1 G(\xi_k)$ for some $x_{k-1} < \xi_k < x_k$; i.e., the last column in both numerator and denominator is replaced by the corresponding relative derivative at ξ_k. We now perform a similar operation on the second to the last column, and so on, until columns 2 through k have been replaced. Since the first element in each of these columns is zero, we finally end up with $\left[\begin{smallmatrix} D_1 v_1, \ldots, D_1 v_k \\ \xi_1, \ldots, \xi_k \end{smallmatrix}\right] D_1 f$. ∎

Proof of Theorem 2. The proof is by induction on k. For $k = 0$ we have

$$\begin{bmatrix} v_0 \\ x_0 \end{bmatrix} f = \frac{f(x_0)}{v_0(x_0)} = \frac{f(x_0)}{w_0(x_0)} = D_0 f(x_0).$$

For $k = 1$, from Lemma 2 we have

$$\begin{bmatrix} D_1 v_1 \\ \xi_1 \end{bmatrix} D_1 f = \frac{D_1 f(\xi_1)}{D_1 v_1(\xi_1)} = D_1 f(\xi_1),$$

because $D_1 v_1 \equiv 1$.

For $k \geqslant 2$, we use Lemma 2 to get

$$\begin{bmatrix} v_0, ..., v_k \\ x_0, ..., x_k \end{bmatrix} f = \begin{bmatrix} D_1 v_1, ..., D_1 v_k \\ \xi_1, ..., \xi_k \end{bmatrix} D_1 f$$

for $x_{i-1} < \xi_i < x_i$ $(i = 1, ..., k)$. Note that $D_1 v_1(x) \equiv 1$. For $i = 2, ..., k$, the Mean Value Theorem for Stieltjes integrals [4] yields

$$D_1 v_i(x) = \lim_{h \to 0} \frac{\int_x^{x+h} \int_c^{t_1} \cdots \int_c^{t_{i-1}} dw_1(t_i) \cdots dw_1(t_1)}{\int_x^{x+h} dw_1(t_1)}$$

$$= \lim_{h \to 0} \int_c^{\xi_h} \int_c^{t_2} \cdots \int_c^{t_{i-1}} dw_i(t_i) \cdots dw_2(t_2),$$

with $x < \xi_h < x + h$; hence

$$D_1 v_i(x) = \int_c^x \int_c^{t_2} \cdots \int_c^{t_{i-1}} dw_i(t_i) \cdots dw_2(t_2) \qquad (i = 2, ..., k).$$

Setting $\tilde{p}_0 \equiv 1$ and $\tilde{p}_i = D_1 v_{i+1}$ $(i = 1, ..., k-1)$, we see that $\{\tilde{p}_0, \tilde{p}_1, ..., \tilde{p}_{k-1}\}$ is a Markov system defined in the same way as $\{p_0, ..., p_k\}$, but using $\{w_2, ..., w_k\}$. Let the corresponding relative differentiation operators be denoted by $\tilde{D}_i$. By the induction hypothesis we then have

$$\begin{bmatrix} v_0, ..., v_k \\ x_0, ..., x_k \end{bmatrix} f = \begin{bmatrix} D_1 v_1, ..., D_1 v_k \\ \xi_1, ..., \xi_k \end{bmatrix} D_1 f$$

$$= \tilde{D}_{k-1}(D_1 f)(\zeta_k) = D_k f(\zeta_k),$$

for $\xi_1 < \zeta_k < \xi_k$; i.e., $x_0 < \zeta_k < x_k$. This completes the proof of Theorem 2. ∎

3. Divided Differences and Proof of Theorem 3

The proof of Theorem 3 is based on several propositions of some independent interest.

LEMMA 3. *Let f, g, and w be continuous in $[a, b]$, with w strictly increasing and g positive in (a, b). Then for some $\xi \in (a, b)$,*

$$\frac{\int_a^b f(t)\, dw(t)}{\int_a^b g(t)\, dw(t)} = \frac{f(\xi)}{g(\xi)}.$$

Proof. By the Mean Value Theorem for Stieltjes integrals, for all $a \leqslant c < d \leqslant b$, there is an $\eta_1 \in [c, d]$ such that

$$f(\eta_1) = \frac{1}{w(d) - w(c)} \int_c^d f(t)\, dw(t). \tag{2}$$

Similarly, for all $a \leqslant c < d \leqslant b$, there is an $\eta_2 \in [c, d]$ such that

$$g(\eta_2) = \frac{1}{w(d) - w(c)} \int_c^d g(t)\, dw(t). \tag{3}$$

Let

$$F(x) = \int_a^x f(t)\, dw(t), \qquad G(x) = \int_a^x g(t)\, dw(t),$$

and set $Q(x) = F(x)\, G(b) - F(b)\, G(x)$. Since $Q(a) = Q(b) = 0$, Q has a relative extremum $\xi \in (a, b)$. Thus, in a neighborhood of ξ,

$$\delta \cdot \operatorname{sgn}(x - \xi) \cdot \operatorname{sgn} \frac{Q(x) - Q(\xi)}{w(x) - w(\xi)} = \eta, \tag{4}$$

where $\delta = \pm 1$ and η depends on x and equals zero or one. However,

$$\frac{Q(x) - Q(\xi)}{w(x) - w(\xi)} = \frac{F(x) - F(\xi)}{w(x) - w(\xi)}\, G(b) - F(b)\, \frac{G(x) - G(\xi)}{w(x) - w(\xi)}. \tag{5}$$

From (2), (3), and (5) we have

$$\lim_{x \to \xi} \frac{Q(x) - Q(\xi)}{w(x) - w(\xi)} = f(\xi)\, G(b) - F(b)\, g(\xi).$$

On the other hand, (4) implies that

$$\lim_{x \to \xi} \frac{Q(x) - Q(\xi)}{w(x) - w(\xi)} = 0,$$

hence

$$\frac{F(b)}{G(b)} = \frac{f(\xi)}{g(\xi)}. \qquad \blacksquare$$

LEMMA 4. *Let* $w_1, ..., w_n$ *be continuous, strictly increasing functions* *defined on an interval I and let $p_i(x)$ be given by* (1). *Then for all* $x_0 < \cdots < x_{n-1}$ *in I,*

$$\begin{bmatrix} p_0, ..., p_{i-1} \\ x_0, ..., x_{i-1} \end{bmatrix} p_i = w_i(\xi_{i-1}) - w_i(c) \qquad (i = 1, ..., n),$$

where $\xi_0 = x_0$ *and* $x_0 < \xi_i < x_i$ $(i = 1, ..., n-1)$.

Proof. The proof is by induction on n. For $n = 1$, we have

$$\begin{bmatrix} p_0 \\ x_0 \end{bmatrix} p_1 = p_1(x_0) = \int_c^{x_0} dw_1(t_1) = w_1(x_0) - w_1(c) = w_1(\xi_0) - w_1(c).$$

For $n = 2$, we have

$$\begin{bmatrix} p_0, p_1 \\ x_0, x_1 \end{bmatrix} p_2 = \frac{p_2(x_1) - p_2(x_0)}{p_1(x_1) - p_1(x_0)} = \frac{\int_{x_0}^{x_1} (\int_c^{t_1} dw_2(t_2)\, dw_1(t_1))}{\int_{x_0}^{x_1} dw_1(t_1)}$$

$$= \int_c^{\xi_1} dw_2(t_2) = w_2(\xi_1) - w_2(c)$$

for some $x_0 < \xi_1 < x_1$, by Lemma 3.

Define $q_0(x) \equiv 1$ and

$$q_i(x) = \int_c^x \int_c^{t_2} \cdots \int_c^{t_i} dw_{i+1}(t_{i+1}) \cdots dw_2(t_2) \qquad (i = 1, ..., n-1).$$

Then

$$p_i(x) = \int_c^x q_{i-1}(t_1)\, dw_1(t_1) \qquad (i = 1, ..., n).$$

Proceeding as in [1, XI, Lemma 2.1] we obtain

$$\det\{p_i(x_j)\}_{i,j=0}^k = \int_{x_0}^{x_1} \cdots \int_{x_{k-1}}^{x_k} \det\{q_i(t_j)\}_{i,j=0}^{k-1} dw_1(t_{k-1}) \cdots dw_1(t_0).$$

It follows by a straightforward inductive procedure that $\{p_i\}_{i=0}^n$ is a Markov system. Moreover, for $k \geqslant 2$, we have

$$\begin{bmatrix} p_0, ..., p_{k-1} \\ x_0, ..., x_{k-1} \end{bmatrix} p_k$$

$$= \frac{\det\{p_i(x_j)\}_{(i=0, ..., k-2, k;\, j=0, ..., k-1)}}{\det\{p_i(x_j)\}_{(i, j=0, ..., k-1)}}$$

$$= \frac{\int_{x_0}^{x_1} \cdots \int_{x_{k-2}}^{x_{k-1}} \det\{q_i(t_j)\}_{(i=0, ..., k-3, k-1;\, j=0, ..., k-2)}\, dw_1(t_{k-2}) \cdots dw_1(t_0)}{\int_{x_0}^{x_1} \cdots \int_{x_{k-2}}^{x_{k-1}} \det\{q_i(t_j)\}_{(i, j=0, ..., k-2)}\, dw_1(t_{k-2}) \cdots dw_1(t_0)}.$$

Using Lemma 3 repeatedly we get

$$\begin{bmatrix} p_0, ..., p_{k-1} \\ x_0, ..., x_{k-1} \end{bmatrix} p_k = \begin{bmatrix} q_0, ..., q_{k-2} \\ \eta_0, ..., \eta_{k-2} \end{bmatrix} q_{k-1},$$

with $x_i < \eta_i < x_{i+1}$ $(i = 0, ..., k-2)$.

By the inductive hypothesis, this equals $w_k(\xi_{k-1}) - w_k(c)$, for $x_{k-2} < \xi_{k-1} < x_{k-1}$. ▮

Lemma 5. *Let A have property D and let Z_n be a Markov system on A with a representation (h, c, W_n, U_n) such that the w_i are strictly increasing on $(\inf h(A), \sup h(A))$. Then for any $x_0 < \cdots < x_{n-1}$ in A,*

$$\begin{bmatrix} z_0, ..., z_{i-1} \\ x_0, ..., x_{i-1} \end{bmatrix} z_i = \begin{bmatrix} u_0, ..., u_{i-1} \\ x_0, ..., x_{i-1} \end{bmatrix} u_i = w_i(\xi_{i-1}) - w_i(c),$$

where $\xi_0 = h(x_0)$ and $h(x_0) < \xi_i < h(x_i)$ $(i = 1, ..., n-1)$.

Remark 6. By [8, Corollary] such a representation with w_i strictly increasing exists.

Proof of Lemma 5. The first equality is a consequence of the fact that $\{u_0, ..., u_n\}$ is obtained from $\{z_0, ..., z_n\}$ by a triangular linear transformation. Let $q_i = u_i/u_0$ $(i = 0, ..., n)$; then $q_0(x) \equiv 1$ and

$$\begin{bmatrix} u_0, ..., u_{i-1} \\ x_0, ..., x_{i-1} \end{bmatrix} u_i = \begin{bmatrix} q_0, ..., q_{i-1} \\ x_0, ..., x_{i-1} \end{bmatrix} q_i \qquad (i = 1, ..., n).$$

The functions q_i can be written as $q_i = p_i \circ h$ for p_i satisfying the hypotheses of Lemma 4. Since

$$\begin{bmatrix} q_0, ..., q_{i-1} \\ x_0, ..., x_{i-1} \end{bmatrix} q_i = \begin{bmatrix} p_0, ..., p_{i-1} \\ h(x_0), ..., h(x_{i-1}) \end{bmatrix} p_i,$$

the assertion follows from Lemma 4. ▮

Before proving Theorem 3, we introduce an additional definition.

Definition 9. A set Z_n defined on a set A is called *endpoint non-degenerate* (END) if for any point $c \in A$ the restrictions of its linear span S to $(-\infty, c) \cap A$ and $A \cap (c, \infty)$ have the same dimension as S. S is then called an END space.

Proof of Theorem 3. (b) ⇒ (a) Suppose that $\inf A \in A$ and $\sup A \notin A$. Let c be a point of A that is not an endpoint. From [5, Theorem 1], Z_n has a representation (h, c, W_n, U_n). As in the proof of [7, Theorem 2], it follows that both parts (a) and (b) of Definition 1 are satisfied. If $\inf A \in A$ and $c = \inf A$ then part (a) of Definition 1 is satisfied (for the same reason), and part (b) is vacuously satsfied since Z_n is linearly dependent on $(-\infty, c] \cap A$. A similar argument holds if $\sup A \in A$ and $\inf A \notin A$.

(d)$\Rightarrow$(c) and (d)$\Rightarrow$(e) These are immediate consequences of Lemma 5 and the second part of Remark 2, respectively.

(a)$\Rightarrow$(d) We may assume that Z_n is normalized (otherwise, divide first by z_0). Suppose that $\inf A \in A$. Since Z_n is weakly nondegenerate, it has a representation (h, c, W_n, U_n) [5, Theorem 1]. Moreover, since $\inf A \in A$ we may select $c = \inf A$. By an argument similar to the one given in the proof of (b)$\Rightarrow$(a) above, it follows that U_n is a normalized weak E-system on A. From [7, Theorem 2], U_n can be continued to the left to an END normalized weak Markov space, say, to $(-\infty, \inf A) \cup A$. By [8, Theorem 3], U_n has an integral representation $(\tilde{h}, \tilde{c}, \tilde{W}_n, \tilde{U}_n)$, and we may assume that $\tilde{c} \in A$. This implies that the $\tilde{w}_i$ are bounded from below on $\tilde{h}(A)$. Moreover, as in the proof of [5, Theorem 2], since A satisfies property B, the $\tilde{w}_i$ must be strictly increasing on $(\inf \tilde{h}(A), \sup \tilde{h}(A))$. If $\sup A \in A$, a change of variables $t \to -t$ leads to a similar proof.

(d)$\Rightarrow$(b) Assume that $\inf A \in A$; then by hypothesis Z_n has a representation (h, c, W_n, U_n) such that the w_i are bounded from below on $h(A)$, and we may assume $c = \inf A$. Thus, U_n can be extended to the left to a Markov system by setting $h(t) = t - c$ and $w_i(t) = (t - c) + w_i(c)$, for $t < c$. A similar argument works for the case $\sup A \in A$.

(c)$\Rightarrow$(d) Let $A' = A \setminus \{\inf A, \sup A\}$. By Lemma 5 and Remark 5, Z_n has a representation (h, c, W_n, U_n) on A' such that

$$\begin{bmatrix} z_0, \ldots, z_{i-1} \\ x_0, \ldots, x_{i-1} \end{bmatrix} z_i = w_i(\xi_{i-1}) - w_i(c),$$

where $\xi_0 = h(x_0)$ and $h(x_0) < \xi_i < h(x_i)$ for $i = 1, \ldots, n-1$. Suppose that $a = \inf A \in A$ ($b = \sup A \in A$). By Remark 4, h is continuous at a (at b), hence by (c),

$$\lim_{x \to a} w_i(h(x)) - w_i(c) = \liminf_{x_0, \ldots, x_{i-1} \to a} \begin{bmatrix} z_0, \ldots, z_{i-1} \\ x_0, \ldots, x_{i-1} \end{bmatrix} z_i > -\infty$$

$$\left(\lim_{x \to b} w_i(h(x)) - w_i(c) = \limsup_{x_0, \ldots, x_{i-1} \to b} \begin{bmatrix} z_0, \ldots, z_{i-1} \\ x_0, \ldots, x_{i-1} \end{bmatrix} z_i < \infty \right)$$

for $i = 1, \ldots, n$. Thus, (d) is valid.

(e)$\Rightarrow$(d) As in the proof of [5, Theorem 2], the elements of W_n must be strictly increasing. The assertion now follows from Lemma 5 and the second part of Remark 2. ∎

References

1. S. Karlin and W. J. Studden, "Tchebycheff Systems: With Applications in Analysis and Statistics," Wiley–Interscience, New York, 1966.
2. F. Schwenker, "Integraldarstellung schwacher Markov-Systeme," Dissertation, Universität Osnabrück, 1988.
3. M. Sommer and H. Strauss, A characterization of Descartes systems in Haar subspaces, preprint.
4. D. V. Widder, "The Laplace Transform," Princeton Univ. Press, Princeton, NJ, 1946.
5. R. A. Zalik, Integral representation and embedding of weak Markov systems, *J. Approx. Theory* **58** (1989), 1–11.
6. R. A. Zalik, Integral representation of Markov systems and the existence of adjoined functions for Haar spaces, *J. Approx. Theory* **65** (1991), 22–31.
7. R. A. Zalik and D. Zwick, On extending the domain of definition of Čebyšev and weak Čebyšev systems, *J. Approx. Theory* **57** (1989), 202–210.
8. R. Zielke, Relative differentiability and integral representation of a class of weak Markov systems, *J. Approx. Theory* **44** (1985), 30–42.