AN OPERATOR IDENTITY

GLEN BAXTER

1. Introduction. Recently, some combinatorial results by Andersen [1, 2], Spitzer [5], and others have been applied quite successfully to problems in probability theory. Many of these applications have given rise to results which are entirely analytical in nature. For example, Spitzer used a combinatorial theorem to find the distribution function for the maximum of the partial sums $S_1, S_2, \cdots, S_n$ for a sequence $\{X_k\}$ of independent, identically distributed random variables. His final result is a functional identity,

$$\sum_{n=0}^{\infty} \varphi_n(t) s^n = \exp\left\{\sum_{k=1}^{\infty} \frac{s^k}{k} \psi_k(t)\right\}, \tag{1.1}$$

where $\varphi_n(t)$ is the characteristic function of $\max(0, S_1, \cdots, S_n)$ and where $\psi_k(t)$ is the characteristic function of $\max(0, S_k)$. One of our purposes in this paper is to generalize (1.1) to an identity involving operators. Our proofs involve more or less analytical methods and thus show that the combinatorial methods hitherto employed can be avoided. We also obtain certain results concerning $\max(X_0, X_1, \cdots, X_n)$ when $\{X_k, k \geq 0\}$ forms a stationary Markov process.

To illustrate the results we consider a simple example. Let N be an $n \times n$ matrix and let N^+ be the matrix formed from N by replacing with zeros all elements of N which are either on or below the diagonal. Let $N^- = N - N^+$, and suppose that N^+ and N^- commute. Now consider the matrix equation

$$PQ = e^N = I + N + N^2/2\,! + \cdots \tag{1.2}$$

where P-I (I is the identity matrix) has non-zero terms only above the diagonal and where $Q - I$ has non-zero terms only on or below the diagonal. The properties of N^+ and N^- imply that

$$\begin{aligned} P &= e^{N^+} = I + N^+ + (N^+)^2/2\,! + \cdots, \\ Q &= e^{N^-} = I + N^- + (N^-)^2/2\,! + \cdots \end{aligned} \tag{1.3}$$

satisfy (1.2) and have the proper form for P and Q. In particular, $\exp(N^+)$ has the proper form for P by virtue of the fact that the product of two matrices with non-zero elements only above the diagonal is a

This paper was originally accepted by the Trans. Amer. Math. Soc. Presented to the American Mathematical Society August 28, 1957; received by the editors of the Trans. Amer. Math. Soc., July 29, 1957. This work was supported by the United States Air Force under Contract No. AF18(603)-30 monitored by the Office of Scientific Research.

matrix of the same type. A similar statement holds for $\exp(N^-)$. It is not hard to see that P and Q are uniquely determined by (1.2). Thus (1.3) is the unique solution of (1.2).

Suppose further that in some neighborhood of $s = 0$, $N = N_1 s + N_2 s^2 + \cdots$, where convergence of the infinite series of $(n \times n)$ matrices is equivalent to convergence of the series of ijth elements for all fixed i and j. Relations (1.3) may be rewritten as power series in s

$$P = \sum_{n=0}^{\infty} P_n s^n \,, \qquad Q = \sum_{n=0}^{\infty} Q_n s^n \tag{1.4}$$

which converge in some neighborhood of $s = 0$. It follows from the form of P and Q that $P_1, P_2, \cdots$ have non-zero elements only above the diagonal while $Q_1, Q_2, \cdots$ have non-zero elements only on or below the diagonal. Certain problems will lead directly to an equation of the form (1.2) where P and Q have the form (1.4). For example, in one case we will have

$$PQ = (I - sM)^{-1} = \exp\left\{\sum_{k=1}^{\infty} \frac{M^k}{k} s^k\right\} . \tag{1.5}$$

Under the appropriate commutativity conditions it will follow that

$$P = \exp\left\{\sum_{k=1}^{\infty} \frac{(M^k)^+}{k} s^k\right\} , \qquad Q = \exp\left\{\sum_{k=1}^{\infty} \frac{(M^k)^-}{k} s^k\right\} . \tag{1.6}$$

We see later that (1.6) is the operator analogue of Spitzer's identity (1.1) whenever the operator M has a special form.

Equation (1.5) is of particular importance in finding the distribution of $\max(X_0, X_1, \cdots, X_n)$ when $\{X_k, k \geq 0\}$ is a Markov process with a stationary transition probability matrix M. In this case the matrix M in (1.5) is identified (see § 4) with the stationary transition probability matrix M. Unfortunately, in the general Markov chain, the commutativity conditions which give (1.6) as the solution of (1.5) are not satisfied. Some information can be obtained directly from (1.5).

In the next section we give general definitions and a few preliminary results. The main theorems are proved in § 3 and illustrated in § 5. A probabilistic interpretation of the theorems is contained in § 4.

2. Definitions and preliminaries. Let L_0 be the space of bounded Baire functions (real-valued and Borel measurable) $f(x)$ on the infinite interval $-\infty < x < \infty$. We will deal with bounded linear operators M defined over L_0 which have the form

$$Mf \equiv \int_{-\infty}^{\infty} f(y) m(x\,; dy) \tag{2.1}$$

where $m(x; A)$ is a function of a real number x and a linear Borel measurable set A such that

(2.2) (i) for each fixed set A, $m(x; A)$ is a Baire function of x,
(ii) for each fixed x, $m(x; A)$ is a signed measure in A on the linear Borel sets.

The norm of the operator M is defined in the usual way in terms of the norm $\|f\| = \max |f(x)|$ in the Banach space L_0. Let $\mu(x; A)$ and $v(x; A)$ be, respectivery, the upper variation and the lower variation of the signed measure $m(x; A)$ (see [4, page 122]) The boundedness of M in (2.1) implies that

$$\int_{-\infty}^{\infty} [\mu(x; dy) + \nu(x; dy)] \leq \max_{-\infty<x<\infty} \int_{-\infty}^{\infty} [\mu(x; dy) + \nu(x; dy)] = \| M \| < \infty . \tag{2.3}$$

We call $m(x; A)$ the kernel of the operator M. The notation which will be used for integration with respect to a given measure is indicated in (2.1). From now on when we call M a bounded linear operater of the form (2.1), we imply that (2.2) is also satisfied. As a matter of fact, with proper understanding of the notation, (2.2) follows directly from (2.1). If M_1 and M_2 are bounded linear operators of the form (2.1) with kernels $m_1(x; A)$ and $m_2(x; A)$, respectively, then M_1M_2 is also of the form (2.1) with kernel

$$m(x; A) = \int_{-\infty}^{\infty} m_2(y; A)m_1(x; dy) . \tag{2.4}$$

We now let $[x]$ be the greatest integer less than or equal to x.

DEFINITION 2.1. Set $B_n(x) = \{y : y > [2^n x + 1]/2^n\}$. For any bounded linear operator M of the form (2.1) with kernel $m(x; A)$, define

$$m^+(x; A) \equiv \lim_{n\to\infty} m(x; B_n(x)A) , \tag{2.5}$$

and let M^+ be the operator of form (2.1) with kernel $m^+(x; A)$. Finally, set, $M^- = M - M^+$.

Almost directly from the definition of M^+ follow certain useful facts which we list below. The bounded, linear operators M, M_1, M_2, etc. are all of the form (2.1); I denotes the identity operator, which is also of the form (2.1); and s, α, and β, are real numbers:

(i) $I^- = I$,
(ii) $(M^+)^+ = M^+$, (iii) $(M^-)^- = M^-$,
(iv) $(M_1^+M_2^+)^+ = M_1^+M_2^+$, (v) $(M_1^-M_2^-)^- = M_1^-M_2^-$,

(vi) $\|M^+\| \leq \|M\|$, (vii) $\|M^-\| \leq \|M\|$,

(viii) $(\alpha M_1 + \beta M_2)^+ = \alpha M_1^+ + \beta M_2^+$,

(2.6) (ix) if $M_0 + M_2 + \cdots$ is a strongly convergent series of bounded, linear operators of the form (2.1), i.e. if $\|M_n + \cdots + M_m\| \to 0$ as $n, m \to \infty$, then $T = M_0 + M_1 + M_2 + \cdots$ is of the form (2.1), and $M_0^+ + M_1^+ + M_2^+ + \cdots$ and $M_0^- + M_1^- + M_2^- + \cdots$ are both convergent in the strong sense. Moreover, $T^+ = M_0^+ + M_1^+ + M_2^+ + \cdots$ and $T^- = M_0^- + M_1^- + M_2^- + \cdots$.

We prove only (ix) of (2.6). Let $T_n = M_0 + \cdots + M_n$, let $t_n(x\,;A)$ be the kernel of T_n, and let χ_A be the characteristic function of a measurable set A. If $T = \lim T_n$, we note that $\|T\|$ is finite. Now

$$|t_n(x\,;\,A) - t_m(x\,;\,A)| = |(T_n - T_m)\chi_A| \leq \|T_n - T_m\|,$$

so that $\lim t_n(x\,;\,A) = t(x\,;A)$ exists uniformly in A. If $A = \Sigma A_k$ where the A_k are disjoint, then by the Moore double-limit theorem

$$\sum_{k=1}^{\infty} t(x\,;A_k) = \lim_{N\to\infty}\lim_{n\to\infty}\sum_{k=1}^{N} t_n(x\,;A_k) = \lim_{n\to\infty} t_n(x\,;A) = t(x\,;A)\,. \tag{2.7}$$

This shows that $t(x\,;A)$ is a signed measure. Since $T\chi_A = t(x\,;A)$, a simple argument shows that $t(x\,;A)$ is the kernel of T. Finally, since $\|T^+ - T_n^+\| \leq \|T - T_n\|$, it follows that $T^+ = \lim T_n^+$. In terms of M_n this means $T^+ = M_0^+ + M_1^+ + M_2^+ + \cdots$. A similar argument gives $T^- = M_0^- + M_1^- + M_2^- + \cdots$.

It is interesting to note that the proofs of the main theorems will depend only on the facts listed in (2.6). Before proceeding to the next section we mention two special subclasses of operators which have the form (2.1).

Case 1. Let $M = (m_{ij})$ be a matrix for which uniformly in i

$$\sum_{(j)} |m_{ij}| < C \tag{2.8}$$

for some constant C. For any Borel measurable set A and any real number x define

$$m(x\,;A) = \begin{cases} \sum_{j\in A} m_{ij} & x = i \text{ (an integer)} \\ 0 & x \neq [x]. \end{cases} \tag{2.9}$$

Condition (2.8) insures the existence of a bounded linear operator of form (2.1) with the kernel $m(x\,;A)$ of (2.9). Certainly the operator given by (2.1) in this case and the original matrix M can be identified. In fact, L_0 could be replaced here by the class of bounded, doubly infinite sequences $\{a_k\}$, that is $a_k = f(k)\,(-\infty < k < \infty)$ where $f(x) \in L_0$. It will be convenient whenever possible to think of the matrix M rather than the operator M. Note that the matrix M^+ is formed from the matrix M by replacing with zeros all elements of M either on or below the diagonal. Moreover, the matrix M^+ satisfies (2.6).

Case 2. Let $m(x, y)$ be Borel measurable and integrable over the plane and such that for some constant C

$$\int_{-\infty}^{\infty} |\, m(x, y)\, |\, dy < C \tag{2.10}$$

uniformly in x. For any Borel measurable set A and any real number x, define

$$m(x\,;\,A) \equiv \int_A m(x, y) dy\ . \tag{2.11}$$

Then, (2.1) gives a bounded, linear operator M which has the form

$$M\, \cdot\, = \int_{-\infty}^{\infty} \cdot\, m(x, y) dy\ , \tag{2.12}$$

and M^+ becomes simply

$$M^+\, \cdot\, = \int_{x}^{\infty} \cdot\, m(x, y) dy \tag{2.13}$$

with a similar formula for M^-.

3. The theorems. When we say a sequence of operators $\{M_n\}$ converges to an operator M, we mean it converges in the strong sense, that is $\|\, M_n - M \,\| \to 0$ as n becomes infinite.

LEMMA 3.1. *Let $\{K_k\}$, $\{P_k\}$, and $\{Q_k\}$, $k = 1, 2, 3, \cdots$, be sequences of bounded, linear operators of the form* (2.1) *for which $P_k^+ = P_k$ and $Q_k^- = Q_k$. For any $|\, s\, | < s_0$, let*

$$\begin{aligned} P &= I + P_1 s + P_2 s^2 + \cdots\ , \\ Q &= I + Q_1 s + Q_2 s^2 + \cdots\ , \\ K &= I + K_1 s + K_2 s^2 + \cdots \end{aligned} \tag{3.1}$$

converge. If $PQ = K$ for all $|\, s\, | < s_0$, then $\{P_k\}$ and $\{Q_k\}$ are uniquely determined by $\{K_k\}$.

Proof. Equating coefficients of like powers of s on the two sides of the equation $PQ = K$ we obtain

$$\sum_{k=0}^{n} P_k Q_{n-k} = K_n\ . \tag{3.2}$$

If $P_1, P_2, \cdots, P_{n-1}$ and $Q_1, Q_2, \cdots, Q_{n-1}$ have been uniquely determined by $K_1, K_2, \cdots, K_{n-1}$, then we may write (3.2) as

$$P_n + Q_n = J_n \tag{3.3}$$

where J_n is determined uniquely by $K_1, K_2, \cdots, K_n$. Since $P_n^- = Q_n^+ = 0$, we have $P_n = J_n^+$ and $Q_n = J_n^-$ and the proof follows by induction.

The next theorems give results in the direction of solving equations which involve the operation " $+$ ". Later we give a probabilistic interpretation of these equations. As we will see, in certain cases the equations may be solved completely in terms of the known operator M.

THEOREM 3.1. *Let M be a bounded, linear operator of the form* (2.1). *Define the sequences $\{P_k\}$, $\{Q_k\}$, $\{R_k\}$, and $\{T_k\}$ by*

$$\begin{aligned} P_0 = Q_0 = I\ , \qquad & R_0 = T_0 = 0\ , \\ P_{n+1} = (MP_n)^+\ , \qquad & Q_{n+1} = (Q_nM)^-\ , \\ T_{n+1} = (MP_n)^-\ , \qquad & R_{n+1} = (Q_nM)^+\ , \end{aligned} \tag{3.4}$$

and let the generating functions of these sequences be

$$\begin{aligned} P = \sum_{n=0}^{\infty} P_n s^n\ , \qquad & Q = \sum_{n=0}^{\infty} Q_n s^n\ , \\ R = \sum_{n=0}^{\infty} R_n s^n\ , \qquad & T = \sum_{n=0}^{\infty} T_n s^n\ , \end{aligned} \tag{3.5}$$

Then, the series' in (3.5) *all converge for $|s| < 1/\|M\|$, and, moreover, they are the unique bounded, linear operators of the form* (2.1) *which satisfy.*

$$\begin{aligned} P = I + s(MP)^+\ , \qquad & T = s(MP)^-\ , \\ Q = I + s(QM)^-\ , \qquad & R = s(QM)^+\ . \end{aligned} \tag{3.6}$$

Proof. Let P be a bounded, linear operator of the form (2.1) which satisfies the first equation of (3.6). By iteration we may write $P = I + P_1 s + P_2 s^2 + \cdots + P_n s^n + L_n$, where $L_0 = s(MP)^+$ and $L_n = s(ML_{n-1})^+$ and where $P_1, P_2, \cdots, P_n$ are determined in (3.4). Property (vi) of (2.6) implies that $\|L_n\| \le |s|^n \|M\|^n \|P\|$ which approaches zero as n becomes infinite for all $|s| < 1/\|M\|$. Thus, the solution (if it exists) of the first equation of (3.6) is unique. Let $\{P_k\}$ satisfy the conditions of (3.4). By property (vi) of (2.6), it follows that $\|P_n\| \le \|M\|^n$. For $|s| < 1/\|M\|$, the power series in (3.5) for P converges and by property (ix) of (2.6)

$$P - I = \sum_{n=0}^{\infty} P_{n+1} s^{n+1} = \sum_{n=0}^{\infty} (MP_n)^+ s^{n+1} = \Big(\sum_{n=0}^{\infty} MP_n s^{n+1}\Big)^+ = s(MP)^+\ . \tag{3.7}$$

The proofs of the other parts of the theorem follow similarly.

THEOREM 3.2. *Let* $|s| < 1/\|M\|$ *and let* P *and* Q *be the bounded, linear operators of the form* (2.1) *which satisfy the equations of* (3.6). *Then,*

$$
(3.8) \qquad \begin{aligned} PQ &= (I - sM)^{-1} \\ sP' &= P(QP - I)^{+}\ , \qquad sQ' = (QP - I)^{-}Q\ , \end{aligned}
$$

where $'$ *indicates derivative with respect to* s.

Proof. From (3.6) we find that $\|Q\| \leq 1/(1 - |s|\|M\|)$ and

$$\|R\| \leq |s|\|M\|\|Q\| \leq |s|\|M\|/(1 - |s|\|M\|)\ .$$

Thus, for $|s| < (1 - |s|\|M\|)/\|M\|$, the operator $(I - R)^{-1}$ is a bounded linear operator of the form (2.1) and has a convergent power series expansion in s. But $Q = I - R + sQM$, or equivalently, $(I - R)^{-1}Q = (I - sM)^{-1}$. Similarly we show that $(I - T)^{-1}$ is a bounded linear operator of the form (2.1) which has a convergent power series expansion in s for $|s| < (1 - |s|\|M\|)/\|M\|$, and that $P(I - T)^{-1} = (I - sM)^{-1}$. Applying Lemma 3.1 in the common interval of convergence of P, Q, $(I - T)^{-1}$ and $(I - R)^{-1}$, we deduce that

$$(3.9) \qquad P = (I - R)^{-1}\ , \qquad Q = (I - T)^{-1}\ .$$

and hence that $PQ = (I - sM)^{-1}$. Since P, Q, and $(I - sM)^{-1}$ all converge for $|s| < 1/\|M\|$, we have finally $PQ = (I - sM)^{-1}$ for all $|s| < 1/\|M\|$. To show the second half of (3.8), we consider $(PQ)' = P'Q + PQ' = (I - sM)^{-2}M$. It follows that

$$(3.10) \qquad (PQ)^2 - s(PQ)' = (I - sM)^{-2}(I - sM) = PQ\ .$$

Multiplying on the left of (3.10) by P^{-1} and on the right by Q^{-1} (take $|s| < (1 - |s|\|M\|)/\|M\|$) we obtain

$$(3.11) \qquad QP - s(P^{-1}P' + Q'Q^{-1}) = I\ .$$

By properties (iv), (v), and (ix) of (2.6), it is not hard to see that $(P^{-1}P')^{+} = P^{-1}P'$ and $(Q'Q^{-1})^{-} = Q'Q^{-1}$. From (3.11) we find $sP' = P(QP - I)^{+}$ and $sQ' = (QP - I)^{-}Q$. These latter equations can certainly be extended to hold for all $|s| < 1/\|M\|$, and the theorem is proved.

THEOREM 3.3. *Let* $\{a_k\}$ *be a sequence of real numbers such that* $a_1s + a_2s^2 + a_3s^3 + \cdots$ *has a positive radius of convergence. Let* M *be a bounded, linear operator of the form* (2.1) *such that* $(M^k)^{+}M = M(M^k)^{+}$ *for all* $k = 1, 2, 3, \cdots$. *Then for* $|s|$ *such that*

$$(3.12) \qquad \sum_{k=1}^{\infty} |a_k|\|M\|^k|s|^k < 1\ ,$$

there is a unique pair of bounded linear operators P and Q of the form (2.1) *which satisfy*

$$P = I + \left[\sum_{k=1}^{\infty}(a_k M^k s^k)P\right]^+ ,\qquad (3.13)$$
$$Q = I + \left[Q\sum_{k=1}^{\infty}(a_k M^k s^k)\right]^- .$$

Moreover, the solution of (3.13) *is*

$$P = \exp\left\{\left[-\log\left(I - \sum_{k=1}^{\infty} a_k M^k s^k\right)\right]^+\right\} ,\qquad (3.14)$$
$$Q = \exp\left\{\left[-\log\left(I - \sum_{k=1}^{\infty} a_k M^k s^k\right)\right]^-\right\} .$$

Before proving Theorem 3.3 we mention a result of particular interest which occurs when both Theorems 3.1 and 3.3 apply, i.e. when $a_1 = 1$ and $a_2 = a_3 = \cdots = 0$.

COROLLARY 3.1. *Let M be a bounded linear operator of the form* (2.1) *such that $(M^k)^+ M = M(M^k)^+$ for all $k = 1, 2, 3, \cdots$, and let the sequences $\{P_k\}$ and $\{Q_k\}$ be defined as in* (3.4). *Then, for all $|s| < 1/\|M\|$, the P and Q of* (3.5) *have the form*

$$P = \exp\left\{\sum_{k=1}^{\infty}\frac{(M^k)^+}{k}s^k\right\} ,\qquad Q = \exp\left\{\sum_{k=1}^{\infty}\frac{(M^k)^-}{k}s^k\right\} .\qquad (3.15)$$

Proof of Theorem 3.3. Let $|s|$ satisfy the condition of (3.12), and let

$$L = \sum_{k=1}^{\infty} a_k M^k s^k ,\qquad (3.16)$$
$$N = \log\left(I - \sum_{k=1}^{\infty} a_k M^k s^k\right) = \sum_{k=1}^{\infty} L^k/k .$$

Both L and N are bounded linear operators of the form (2.1). The commutativity of $(M^k)^+$ and M together with property (ix) of (2.6) implies that $L^+L = LL^+$. Again by property (ix) of (2.6) and the second relation of (3.16), we deduce that $N^+N = NN^+$. In terms of N the first equation in (3.13) may be written in the form

$$P = I + [(I - e^N)P]^+ .\qquad (3.17)$$

Using that $(\exp(-N^+))^+ = \exp(-N^+) - I$ and that $(\exp(N^-))^+ = 0$, it is easy to show by substitution that $P = \exp(-N^+)$ is a solution of (3.17). To show that this solution is unique we apply Theorem 3.1, where the operator "M" of Theorem 3.1 is now

$$\sum_{k=1}^{\infty} a_k M^k s^k \tag{3.18}$$

and the number "s" of Theorem 3.1 is now 1. In a similar manner we can show that the Q of (3.14) is the unique solution of the second equation in (3.13). This finishes the proof.

Before proceeding into the next section, we point out some implications of the theorems above. In Theorem 3.3, the operators P, Q, M, M^+, and M^- all commute. Thus, the order of the factors Q and M^k or of P and M^k in (3.13) is unimportant. In the s interval determined by (3.12), there is a power series expansion in s for the solutions of (3.13). The coefficients in this power series satisfy

$$\begin{aligned} &P_0 = Q_0 = I\ , \\ &P_{n+1} = (a_1 M P_n + a_2 M^2 P_{n-1} + \cdots + a_{n+1} M^{n+1})^+\ , \\ &Q_{n+1} = (a_1 Q_n M + a_2 Q_{n-1} M^2 + \cdots + a_{n+1} M^{n+1})^-\ . \end{aligned} \tag{3.19}$$

If the M in Theorem 3.1 is a matrix of finite order, the P and Q of (3.5) can be conveniently evaluated in terms of subdeterminants of the matrix $I - sM$ (See example 3, § 5).

4. Probabilistic interpretation. In this section we give a probabilistic interpretation of the sequences $\{P_k\}$, $\{Q_k\}$, $\{R_k\}$, and $\{T_k\}$ of Theorem 3.1. Let $m(x\,;A)$ be a function of a real number x and a linear Borel measurable set A such that

(4.1)
(i) for each fixed set A, $m(x\,;A)$ is a Baire function of x,
(ii) for each fixed x, $m(x\,;A)$ is a probability measure in A on the linear Borel measurable sets.

Let $\{X_k, k \geq 0\}$ be a stationary Markov process for which $m(x\,;A) = P\{X_{k+1} \in A \mid X_k = x\}$ is defined and satisfies the conditions of (4.1) (see [3, pp. 18, 26–27]). We deal here only with processes of this type. By (2.1) and (2.3) each Markov process under consideration has associated with it a bounded linear operator M, with $\|M\| = 1$. We call this the transition probability operator of the process.

Two subcases of special interest may be mentioned. The first one is that of a discrete Markov *chain* (countable state space). In this case the transition probabilities form a matrix $M = (m_{ij})$. The connection between the matrix M and the function $m(x\,;A)$ has already been discussed in § 2, case 1. The second type process of interest is the one for which the joint distributions have densities. In this latter case, there exists a transition probability density function $m(x, y)$, and the connection with $m(x\,;A)$ is given in § 2, case 2.

For convenience in stating the next theorem we introduce a random variable L_n.

$$L_n: \text{ the index } k\,(=0, 1, 2, \cdots) \text{ for which } \max(X_0, X_1, \cdots, X_n) = X_k \text{ and } \max(X_0, X_1, \cdots, X_{k-1}) < X_k\,. \tag{4.2}$$

Note in particular the meaning of the statements $L_n = n$ and $L_n = 0$. In Theorem 4.1 and thereafter we will have occasion to refer to the kernel associated with a given operator of the form (2.1). If the operator is denoted by some capital letter, the kernel will be denoted by the corresponding small letter.

THEOREM 4.1. *Let $\{X_k, k \geq 0\}$ be a stationary Markov process with transition probability operator M, and let $\{P_k\}$, $\{Q_k\}$, $\{R_k\}$ and $\{T_k\}$ be defined as in* (3.4). *Then, if the right hand members of* (4.3) *are defined and satisfy* (2.2), *we have*

$$\begin{aligned} p_n(x\,;A) &= P\{L_n = n, X_n \in A \mid X_0 = x\}\,, \\ q_n(x\,;A) &= P\{L_n = 0, X_n \in A \mid X_0 = x\}\,, \\ r_n(x\,;A) &= P\{L_n = n, L_{n-1} = 0, X_n \in A \mid X_0 = x\}\,, \\ t_n(x\,;A) &= P\{L_n = 0, \max(X_1, \cdots, X_{n-1}) < X_n, X_n \in A \mid X_0 = x\}\,. \end{aligned} \tag{4.3}$$

Proof. We prove only the first one of the relations in (4.3). Our proof is by induction. Since $P_0 = I$, it follows that

$$p_0(x\,;A) = P\{X_0 \in A \mid X_0 = x\} = \begin{cases} 1 & x \in A \\ 0 & x \notin A\,. \end{cases} \tag{4.4}$$

Now assume the first relation of (4.3) is true for the case n and set $B_N(x) = \{y : y > [2^N x + 1]/2^N\}$ for $N = 1, 2, 3, \cdots$. Then,

$$\begin{aligned} &P\{L_{n+1} = n + 1, X_{n+1} \in B_N(x)A \mid X_0 = x\} \\ &= \int_{-\infty}^{\infty} P\{\max(X_1, \cdots, X_n) < X_{n+1}, X_{n+1} \in B_N(x)A \mid X_1 = z\} \\ &\qquad\qquad \cdot P\{X_1 \in dz \mid X_0 = x\} \\ &= \int_{-\infty}^{\infty} \{L_n = n, X_n \in B_N(x)A \mid X_0 = z\} P\{X_1 \in dz \mid X_0 = x\} \\ &= \int_{-\infty}^{\infty} p_n(z\,;B_N(x)A) m(x\,;dz)\,. \end{aligned} \tag{4.5}$$

From (2.4) we see that the last term of (4.5) is the kernel of MP_n evaluated at x and $B_N(x)A$. Set $A_x = A \cap (x, \infty)$, and note that for any $n > 0$,

$$P\{L_n = n, X \in A \mid X_0 = x\} = P\{L_n = n, X_n \in A_x \mid X_0 = x\}\,.$$

Thus, by Definition 2.1 and (4.5)

$$
\begin{aligned}
p_{n+1}(x\,;A) &= \lim_{N\to\infty}\int_{-\infty}^{\infty} p_n(z\,;B_N(x)A,\ m(x\,;dz)) \\
&= \lim_{N\to\infty} P\{L_{n+1}=n+1,\ X_{n+1}\in B_N(x)A \mid X_0=x\} \\
&= P\{L_{n+1}=n+1,\ X_{n+1}\in A_x \mid X_0=x\} \\
&= P\{L_{n+1}=n+1,\ X_{n+1}\in A \mid X_0=x\}\ ,
\end{aligned} \tag{4.6}
$$

and the proof follows by induction.

Combining the first and second of the relations in (4.3) we get certain additional information about $\max(X_0, \cdots, X_n)$. In fact, we can evaluate the generating function

$$
\sum_{n=0}^{\infty} P\{\max(X_0, \cdots, X_n)\in A \mid X_0=x\}s^n \tag{4.6}
$$

in terms of the kernels of P and Q. Let $S=(-\infty,\infty)$. Then, by Theorem 4.1

$$
\begin{aligned}
&P\{\max(X_0, \cdots, X_n)\in A \mid X_0=x\} \\
&\quad= \sum_{k=0}^{n} P\{L_n=k,\ \max(X_0, \cdots, X_n)\in A \mid X_0=x\} \\
&\quad= \sum_{k=0}^{n}\int_A P\{L_{n-k}=0 \mid X_0=y\}P\{L_k=k,\ X_k\in dy \mid X_0=x\} \\
&\quad= \sum_{k=0}^{n}\int_A q_{n-k}(y\,;S)p_k(x\,;dy)\ .
\end{aligned} \tag{4.7}
$$

Multiplying through (4.7) by s^n and summing over $s=0, 1, 2, \cdots$ we obtain

$$
\int_A q(y\,;S)p(x\,;dy) = \sum_{n=0}^{\infty} P\{\max(X_0, \cdots, X_n)\in A \mid X_0=x\}s^n\ . \tag{4.8}
$$

Relation (4.8) takes on a particularly simple form if $q(y\,;S)$ is independent of y (See example 2, § 5). In fact, in this special case we have the following Corollary to Theorem 4.1:

COROLLARY 4.1. *Let $\{X_k, k\ge 0\}$ be a stationary Markov process with transition probability operator M and let P and Q be defined as in Theorem 3.1. Furthermore, let $q(x\,;A)$ be the kernel of Q, and let Φ be the bounded, linear operator of the form* (2.1) *determined by*

$$
\varphi(x\,;A) = \sum_{n=0}^{\infty} P\{\max(X_0, \cdots, X_n)\in A \mid X_0=x\}s^n\ . \tag{4.9}
$$

Then, if $q(x\,;S)=q$ is independent of x,

$$
\Phi = qP\ . \tag{4.10}
$$

Relation (4.10) is an operator analogue of Spitzer's identity (1.1).

5. Examples. We now give applications of the theorems to some particular examples.

EXAMPLE 1. Let the operator of form (2.1) be (See case 1, § 2)

$$M = \begin{bmatrix} a & 0 & b \\ (a-c)d/b & c & d \\ 0 & 0 & a \end{bmatrix}, \tag{5.1}$$

so that for $k = 1, 2, 3, \cdots$

$$M^k = \begin{bmatrix} a & 0 & ka^{k-1}b \\ (a^k - c^k)d/b & c^k & ka^{k-1}d \\ 0 & 0 & a^k \end{bmatrix}. \tag{5.2}$$

It is not hard to see that $(M^k)^+M = M(M^k)^+$ in this case so Corollary 3.1 applies here. The solution of $P = I + s(MP)^+$ for $|s| < 1/\|M\| < 1/|a|$ is

$$\begin{aligned} P &= \exp\left\{\sum_{k=1}^{\infty} \frac{s^k}{k}(M^k)^+\right\} = \exp\left\{\sum_{k=1}^{\infty} s^k \begin{bmatrix} 0 & 0 & a^{k-1}b \\ 0 & 0 & a^{k-1}d \\ 0 & 0 & 0 \end{bmatrix}\right\} \\ &= \exp\left\{\begin{bmatrix} 0 & 0 & bs/(1-as) \\ 0 & 0 & ds/(1-as) \\ 0 & 0 & 0 \end{bmatrix}\right\} = \begin{bmatrix} 1 & 0 & bs/(1-as) \\ 0 & 1 & ds/(1-as) \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned} \tag{5.3}$$

In a similar manner it follows that the solution of $Q = I + s(QM)^-$ for $|s| < 1/\|M\| < 1/\min(|a|, |c|)$ is

$$Q = \begin{bmatrix} 1/(1-as) & 0 & 0 \\ (a-c)ds/b(1-as)(1-cs) & 1/(1-cs) & 0 \\ 0 & 0 & 1/(1-as) \end{bmatrix}. \tag{5.4}$$

These solutions are easily checked by substitution.

EXAMPLE 2. Let $\{X_k\}(k = 1, 2, 3, \cdots)$ be a sequence of independent, identically distributed random variables with a common density function $f(x)$, and let $S_n = X_1 + \cdots + X_n$. If T_0 is any random variable independent of $\{X_k\}$, and if we set $T_n = S_n + T_0\,(n = 1, 2, 3, \cdots)$, then $\{T_n, n \geq 0\}$ is a stationary Markov process with transition probability

$$m(x\,; A) \equiv P\{T_{k+1} \in A \mid T_k = x\} = \int_A f(y-x)dy\,. \tag{5.5}$$

The conditions (4.1) are satisfied by $m(x\,; A)$ (as well as by the right hand members of (4.3)) in this case so we so may talk about the transition probability operator M associated with $\{T_n, n \geq 0\}$. This operator has the form

$$M \cdot = \int_{-\infty}^{\infty} \cdot f(y - x)dy \; . \tag{5.6}$$

Using (2.4) and (5.6) it is not hard to deduce that M^k also has a kernel with a density. In fact,

$$M^k \cdot = \int_{-\infty}^{\infty} \cdot f_k(y - x)dy \; , \tag{5.7}$$

where $f_k(x)$ is the k-fold convolution of $f(x)$ with itself.

By (5.6), (5.7), and (2.4) we see that the kernel of $(M^k)^+ M$ has a density of the form

$$\int_{-\infty}^{\infty} f(y - w)f_k^+(w - x)dw = \int_{x}^{\infty} f(y - w)f_k(w - x)dw \; . \tag{5.8}$$

We now make the change of variable $z = y + x - w$ in the second integral of (5.8) to get

$$\int_{y}^{\infty} f_k(y - z)f(z - x)dz = \int_{-\infty}^{\infty} f_k^+(y - z)f(z - x)dz \; . \tag{5.9}$$

The second term of (5.9) is the density of the kernel of $M(M^k)^+$. Thus, $(M^k)^+ M = M(M^k)^+$ in this case and Corollary 3.1 applies. If P and Q are as defined in Theorem 3.1, then for $|s| < 1$ (that is $\|M\| = 1$)

$$P = \exp\left\{\sum_{k=1}^{\infty} \frac{(M^k)^+ s^k}{k}\right\} , \qquad Q = \exp\left\{\sum_{k=1}^{\infty} \frac{(M^k)^- s^k}{k}\right\} . \tag{5.10}$$

Since $(M^k)^-$ has a kernel with a density of the form $f_k(y - x)$, we deduce that Q must have a kernel with a density of the form $q(y - x)$. This means

$$q(x \, ; S) = \int_{-\infty}^{\infty} q(y - x)dy = \exp\left[\sum_{k=1}^{\infty} \frac{P\{S_k \leq 0\}}{k} s_k\right] \tag{5.11}$$

is independent of x and Corollary 4.1 applies. Spitzer's identity (1.1) is found in this case from (4.10) by operating with each side on the function $g(y) = \exp(ity)$. In fact, in the notation of (1.1)

$$\begin{aligned} \Phi g &= e^{itx} \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} e^{it(y-x)} P\{\max(T_0, \cdots, T_n) \in dy \mid T_0 = x\} s^n \\ &= e^{itx} \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} e^{ity} P\{\max(0, S_1, \cdots, S_n) \in dy\} s^n \\ &= e^{itx} \sum_{n=0}^{\infty} \phi_n(t) s^n \, . \end{aligned} \tag{5.12}$$

Now in the special case of the exponential function $g(y) = e^{ity}$,

$$(M^k)^+(M^n)^+ge^{-itx} = [(M^k)^+ge^{-itx}][(M^n)^+ge^{-itx}] . \tag{5.13}$$

From (6.10), we find[1]

$$\begin{aligned} Pg &= e^{itx} \exp\left\{\sum_{k=1}^{\infty} \frac{(M^k)^+ge^{-itx}}{k} s^n\right\} \\ &= e^{itx} \exp\left[\sum_{k=1}^{\infty} \frac{s^k}{k} \int_0^{\infty} e^{ity} P\{S_k \in dy\}\right] . \end{aligned} \tag{5.14}$$

Putting (5.11), (5.12), and (5.14) into (4.10), it follows that

$$\sum_{n=1}^{\infty} \varphi_n(t)s^n = \exp\left[\sum_{k=1}^{\infty} \frac{s^k}{k} \int_0^{\infty} e^{ity} P\{\max(0, S_k) \in dy\}\right] . \tag{1.1}$$

In passing we note that the existence of a density is convenient but not necessary for the derivation of (1.1) from (4.10). In general, we can replace (5.5) by

$$m(x ; A) = P\{(X_1 + x) \in A\} , \tag{5.15}$$

which is Borel measurable in x for each fixed set A. The conditions (4.1) are satisfied and the derivation continues in the obvious manner.

EXAMPLE 3. Let M be a matrix of finite order. We denote by D_k the subdeterminant formed from the determinant of $I - sM$ by crossing out all but the first k rows and columns. Moreover, $D_k(i ; j)$ $(1 \leq i, j \leq k)$ will denote the cofactor of the ijth element in D_k. Finally, for any matrix N, let $N(k)$ denote the matrix formed from N by crossing out all but the first k rows and columns.

Let $\{P_n\}$, $\{Q_n\}$, $P = (p_{ij})$, and $Q = (q_{ij})$ denote the matrices defined by (3.4) and (3.5) when Theorem 3.1 is applied to M. We may also apply Theorem 3.1 to $M(k)$. It is not hard to show by induction that $\{P_n(k)\}$, $\{Q_n(k)\}$, $P(k)$, and $Q(k)$ are the matrices defined by (3.4) and (3.5) when Theorem 3.1 is applied to $M(k)$. Thus, by (3.8)

$$P(k)Q(k) = [I(k) - sM(k)]^{-1} . \tag{5.16}$$

Equating elements of the last row (the kth row) in the matrix product of (5.16), we find

$$q_{kj} = D_k(j ; k)/D_k , \qquad j = 1, 2, \cdots, k . \tag{5.17}$$

Using (5.17) and the elements of the last column of the product in (5.16), it follows that

$$p_{ik} = D_k(k ; i)/D_{k-1} , \qquad i = 1, 2, \cdots, k . \tag{5.18}$$

[1] The referee points out that (5.14) holds if and only if g is the exponential function.

Let M be the transition probability matrix of a stationary Markov chain $\{X_k, k \geq 0\}$ with states $\alpha_1 < \alpha_2 < \cdots < \alpha_N$. From (4.3), we find

$$\begin{aligned} P\{L_n = n, X_n = \alpha_j \mid X_0 = \alpha_i\} &= D_j(j\,;\,i)/D_{j-1}\,, \qquad (i \leq j)\,, \\ P\{L_n = 0, X_n = \alpha_j \mid X_0 = \alpha_i\} &= D_i(j\,;\,i)/D_i\,, \qquad (i \leq j)\,. \end{aligned} \tag{5.19}$$

References

1. E. Sparre Andersen, *On the fluctuations of sums of random variables*, Math. Scand. **1** (1953), 263-285.
2. ———, *On the fluctuations of sums of random variables* II, Math. Scand. **2** (1954), 195-223.
3. J. L. Doob, *Stochastic process*, New York 1953.
4. P. R. Halmos, *Measure theory*, New York 1958.
5. Frank Spitzer, *A combinatorial lemma and its application to probability theory*, Trans. Amer. Math. Soc. **82** (1956), 323-339.

UNIVERSITY OF MINNESOTA