ON A FUNDAMENTAL IDENTITY IN THE THEORY OF SEMI-MARKOV PROCESSES

BY

E. ARJAS

Reprinted from the
ADVANCES IN APPLIED PROBABILITY
Vol. 4, No. 2, pp. 258-270
August 1972

Adv. Appl. Prob. **4**, 258–270 (1972)
Printed in Israel

ON A FUNDAMENTAL IDENTITY IN THE THEORY OF SEMI-MARKOV PROCESSES

E. ARJAS, *The Academy of Finland, Helsinki*

Abstract

A fundamental identity, due to Miller (1961a), (1962a, b) and Kemperman (1961), is generalized to semi-Markov processes. Thus the identity applies to processes defined on a Markov chain with discrete state space and random walks with Markov dependent steps (Section 2). Wald's identity is discussed briefly in Section 3. Section 4 is a study of the maxima of partial sums, and Section 5 of maxima in a semi-Markov process.

SEMI-MARKOV KERNEL; STOPPING TIME; FUNDAMENTAL IDENTITY; WALD'S IDENTITY; MAXIMUM OF PARTIAL SUMS; MAXIMAL STATE

1. Introduction

Let E be an arbitrary set, $\mathscr{E}$ a σ-algebra of subsets of E, and let $\mathscr{B}$ be the Borel sets of $R^1 = (-\infty, +\infty)$. Denote $R^+ = [0, +\infty)$ and $R^- = (-\infty, 0]$.

We consider a semi-Markov kernel $q(\cdot,\cdot,\cdot)$ and make the following specifications for it (cf. Çinlar (1969a)):

(i) q is a mapping $E \times \mathscr{L} \to [0,1]$, where $\mathscr{L}$ is the σ-algebra of subsets of $E \times R^1$, generated by the rectangles $A \times B$ with $A \in \mathscr{E}$ and $B \in \mathscr{B}$;

(ii) for fixed $x \in E$, $q(x,\cdot,\cdot)$ is a measure on $\mathscr{L}$;

(iii) for fixed $C \in \mathscr{L}$, $q(\cdot, C)$ is $\mathscr{E}$-measurable.

For convenience we assume throughout that $q(x, E, R^1) = 1$ for all $x \in E$.

We can define $X_0, X_1, \cdots$ and $Y_0, Y_1, \cdots$, two sequences of random variables such that

$$P(X_{n+1} \in A, Y_{n+1} \in B \mid X_n) = q(X_n, A, B) \tag{1.1}$$

for any $n = 0, 1, 2, \cdots$; $A \in \mathscr{E}$ and $B \in \mathscr{B}$. Also we define the sum S_n by

$$S_0 = 0,\ S_n = Y_1 + \cdots + Y_n,\ n = 1, 2, \cdots. \tag{1.2}$$

Remark. The family $\mathscr{B}$ could well be the Borel sets of a d-dimensional real space R^d. This would lead to no additional difficulties; we would only have to interpret the random variables Y_n and S_n as random vectors.

Received in revised form 31 January 1972. Research done while the author was in the Department of Probability and Statistics, University of Sheffield.

The process $\{X_n\}$ is often referred to as an underlying Markov chain. Given the transitions in the chain, the random variables $\{Y_n\}$ are conditionally mutually independent. On the other hand, if S_n is understood to be the time, which is required for the n transitions, we have a semi-Markov process with state variable X_n. If S_n is a d-dimensional random vector, we can interpret one of its coordinates in the same way.

Without repeating the constructive definitions we assume N to be a stopping time, taking values in $\{0, 1, 2, \cdots\}$. In Section 4 we choose N to be the first index n for which $\{S_n\}$ hits a given region, and in Section 5 N corresponds to the absorption of $\{X_n\}$.

2. The fundamental identity

As a basis for all subsequent analysis we use a fundamental identity. This identity was originally given for independent and identically distributed random variables by Miller (1961a) and Kemperman (1961). Our result is a generalization of this, as it is also of a later and more general version derived by Miller (1962a).

The argument in the following is of forward type. We want to study the distribution of $\{X_n, S_n\}$ at the random stopping time N. This is done by tracing the process first for values $n < N$ and then looking at the transition at $n = N$.

We define the probability distributions

(2.1a) $$g^{(n)}(x, A, B) = P(X_n \in A, S_n \in B, N > n \mid X_0 = x)$$

and

(2.1b) $$h^{(n)}(x, A, B) = P(X_n \in A, S_n \in B, N = n \mid X_0 = x),$$

$n = 0, 1, 2, \cdots$. Here $P(C \mid X_0 = x)$ stands for the conditional probability $P(C \mid X_0)$ evaluated on the set $\{X_0 = x\}$. We assume that this can always be done in such a way that for fixed C the resulting conditional probability is measurable as a function of x. For $n = 0$ it is convenient to write

(2.2) $$g^{(0)}(x, A, B) + h^{(0)}(x, A, B) = I(x, A)I(0, B),$$

where $I(z, C) = 1$ if $z \in C$ and $I(z, C) = 0$ otherwise.

Assume that $N > n$ holds for some fixed n. Then either $N > n+1$ or $N = n + 1$ is necessarily true. Consequently, we may write the Chapman-Kolmogorov equation

(2.3) $$\int_E \int_{R^1} g^{(n)}(x, dx', dy) q(x', A, B - y) = g^{(n+1)}(x, A, B) + h^{(n+1)}(x, A, B),$$

$n = 0, 1, 2, \cdots$. The integral on the left hand side is defined for all $A \in \mathscr{E}$ and $B \in \mathscr{B}$, and it is a convolution with respect to the third argument. We now define

$$\hat{q}_\beta(x', A) = \int_{R^1} e^{\beta y} q(x', A, dy), \tag{2.4}$$

$$\hat{g}_{\alpha\beta}(x, A) = \sum_n \alpha^n \int_{R^1} e^{\beta y} g^{(n)}(x, A, dy) \tag{2.5a}$$

and

$$\hat{h}_{\alpha\beta}(x, A) = \sum_n \alpha^n \int_{R^1} e^{\beta y} h^{(n)}(x, A, dy) \tag{2.5b}$$

$$= E(e^{\beta S_N} \alpha^N;\ X_N \in A \,|\, X_0 = x),$$

whenever the right hand sides exist. (Here we have used the notation $E(\cdot\,; C) = E(\cdot \,|\, C)P(C)$.) The function $\hat{q}_\beta$ is a moment generating function, abbreviated as m.g.f., whereas $\hat{g}_{\alpha\beta}$ and $\hat{h}_{\alpha\beta}$ are also probability generating functions. For simplicity we call them all transforms. Without more knowledge on the kernel q and the stopping time N it is not possible to give convergence regions of these transforms; in any case they converge for purely imaginary β and $|\alpha| < 1$.

The m.g.f. transformation on both sides of Equation (2.3), multiplication by α^{n+1} and summation over n yields the following proposition.

Proposition 2.1.

$$\hat{g}_{\alpha\beta}(x, A) + \hat{h}_{\alpha\beta}(x, A) = I(x, A) + \alpha \int_E \hat{g}_{\alpha\beta}(x, dx')\hat{q}_\beta(x', A).$$

Remark. If the random walk $\{S_n\}$ has a non-zero starting point, $S_0 = y$, this will only result in the multiplication of $I(x, A)$ by $e^{\beta y}$.

Proposition 2.1. is called the fundamental identity. It is closely related to a forward Markov renewal equation. To see this we can think that N is infinitely large, hence $\hat{h}_{\alpha\beta}$ vanishes, and then set $\alpha = 1$.

As a rule the transformed kernel $\hat{q}_\beta$ is assumed to be known, whereas the transforms $\hat{g}_{\alpha\beta}$ and $\hat{h}_{\alpha\beta}$ are unknown. We think mainly of the fundamental identity as a relation between the transforms $\hat{q}_\beta$ and $\hat{h}_{\alpha\beta}$. Later in the paper we consider two choices of the stopping time N, which make $\hat{h}_{\alpha\beta}$ correspond to a "ladder" point. Then the fundamental identity can be used as a factorization theorem in the sense of Kingman (1966). In another paper, Arjas (1972b), we consider the various consequences of the fundamental identity in queuing theory. There we also make an attempt to determine the unknown transforms $\hat{g}_{\alpha\beta}$ and $\hat{h}_{\alpha\beta}$ in a special case.

Processes defined on a Markov chain with a discrete state space. We now consider a special case of Proposition 2.1. Let $\{X_n\}$ be a Markov chain with a discrete state space. Denote the states by $i, j, \cdots \in E$. The semi-Markov kernel $q(\cdot\,, \cdot\,, \cdot)$ is replaced by a transition matrix $Q(\cdot) = \{q_{ij}(\cdot)\}$, where the elements are

(2.6) $$q_{ij}(B) = P(X_{n+1} = j, Y_{n+1} \in B \mid X_n = i)$$

for any $i, j \in E$, $B \in \mathscr{B}$ and $n = 0, 1, 2, \cdots$. Let the matrices $\boldsymbol{G}^{(n)}(B) = \{g_{ij}^{(n)}(B)\}$ and $\boldsymbol{H}^{(n)}(B) = \{h_{ij}^{(n)}(B)\}$ have entries

(2.7a) $$g_{ij}^{(n)}(B) = P(X_n = j, S_n \in B, N > n \mid X_0 = i)$$

and

(2.7b) $$h_{ij}^{(n)}(B) = P(X_n = j, S_n \in B, N = n \mid X_0 = i),$$

and denote the corresponding matrix transforms by

(2.8) $$\hat{\boldsymbol{Q}}_\beta = \left\{\int_{R^1} e^{\beta y} q_{ij}(dy)\right\},$$

(2.9a) $$\hat{\boldsymbol{G}}_{\alpha\beta} = \left\{\sum_n \alpha^n \int_{R^1} e^{\beta y} g_{ij}^{(n)}(dy)\right\}$$

and

(2.9b) $$\hat{\boldsymbol{H}}_{\alpha\beta} = \left\{\sum_n \alpha^n \int_{R^1} e^{\beta y} h_{ij}^{(n)}(dy)\right\}.$$

It is then obvious that the integral in Proposition 2.1 is replaced by a product of two matrices. Thus the fundamental identity becomes:

(2.10) $$\hat{\boldsymbol{G}}_{\alpha\beta} + \hat{\boldsymbol{H}}_{\alpha\beta} = \boldsymbol{I} + \alpha \hat{\boldsymbol{G}}_{\alpha\beta} \hat{\boldsymbol{Q}}_\beta,$$

where $\boldsymbol{I}$ is the unit matrix.

Random walks with Markov dependent steps. We consider another special case of Proposition 2.1. This is obtained by identifying the random variables X_n and Y_n. We then assume that $E = R^1$ and $\mathscr{E} = \mathscr{B}$. If Y_n is a d-dimensional vector we may identify X_n with one of the coordinates of Y_n. The process $\{S_n\}$ is now a random walk with Markov dependent steps. We have here Markovian dependence between the increments. Let $S_0 = Y_0 = x$ and $S_n = x + Y_1 + Y_2 + \cdots + Y_n$, $n = 1, 2, \cdots$.

For convenience we denote the Markov kernel by $q(x', A)$ with $x' \in R^1$ and $A \in \mathscr{B}$, and write

(2.11) $$\hat{q}_\beta(x', A) = \int_A e^{\beta y} q(x', dy),$$

(2.12a) $$\hat{g}_{\alpha\beta}(x, A) = \sum_n \alpha^n \int_{R^1} e^{\beta y} P(Y_n \in A, S_n \in dy, N > n \mid Y_0 = x),$$

(2.12b) $$\hat{h}_{\alpha\beta}(x, A) = \sum_n \alpha^n \int_{R^1} e^{\beta y} P(Y_n \in A, S_n \in dy, N = n \mid Y_0 = x)$$

for the corresponding transforms. Then the fundamental identity has the form

$$\hat{g}_{\alpha\beta}(x, A) + \hat{h}_{\alpha\beta}(x, A) = I(x, A)e^{\beta x} + \alpha\int_{R^1} \hat{g}_{\alpha\beta}(x, dx')\hat{q}_\beta(x', A). \tag{2.13}$$

3. Wald's identity in the general case

Miller used the formula (2.10) in the derivation of Wald's identity. Essentially the same procedure applies here in the more general case. A detailed study on Wald's identity and its uses in this context has been made by Matthews (1971), where also more references are given.

Assume that the fundamental identity is valid with $\hat{g}_{\alpha\beta}$ and $\hat{h}_{\alpha\beta}$ finite in a certain region. Assume also that there exists in this region a maximal eigenvalue λ_β of $\hat{q}_\beta$ with the corresponding eigenfunction u_β :

$$\int_E \hat{q}_\beta(x, dx')u_\beta(x') = \lambda_\beta u_\beta(x). \tag{3.1}$$

Finally assume that

$$\int_E\int_E \left| \hat{g}_{\alpha\beta}(x, dx')q_\beta(x', dz)\; u_\beta(z)\right| < \infty \tag{3.2}$$

at $\alpha = \lambda_\beta^{-1}$.

Proposition 3.1.

$$E(e^{\beta S_N}\lambda_\beta^{-N}u_\beta(X_N);\, N < \infty \,|\, X_0 = x) = u_\beta(x).$$

The verification of this is as follows. In Proposition 2.1 let $A = dz$. Both sides of the identity are then multiplied by $u_\beta(z)$. Integration over $z \in E$ yields

$$\int_E \hat{g}_{\alpha\beta}(x, dz)u_\beta(z) - \alpha\int_E\int_E \hat{g}_{\alpha\beta}(x, dx')\hat{q}_\beta(x', dz)u_\beta(z) = u_\beta(x) - \int_E \hat{h}_{\alpha\beta}(x, dz)u_\beta(z). \tag{3.3}$$

The change of order of integration in the double integral (which is justified by (3.2)) and the use of Property (3.1) yields

$$\alpha\int_E \hat{g}_{\alpha\beta}(x, dx')\int_E \hat{q}_\beta(x', dz)u_\beta(z) = \alpha\lambda_\beta\int_E \hat{g}_{\alpha\beta}(x, dx')u_\beta(x').$$

Consequently, when evaluating (3.3) at $\alpha = \lambda_\beta^{-1}$ the left hand side vanishes and Proposition 3.1. follows by Definition (2.5b).

4. Maximum of partial sums

We consider the maximum of partial sums

$$M_n = \max(0, S_1, S_2, \cdots, S_n), \qquad n = 1, 2, \cdots. \tag{4.1}$$

To demonstrate the role of the fundamental identity as a factorization theorem, we first study the distribution of M_n, assuming that $\{S_n\}$ is a random walk defined on a finite state space Markov chain. The consequent easy matrix formalism makes it natural to apply the algebraic argument of Kingman (1966). After this, the distribution of the maximum is derived anew, now assuming the Markov chain to have a general state space and using a method based on ladder indices. In this latter derivation the fundamental identity is not needed explicitly; it merely connects the associated ladder process to the original random walk.

First derivation. Assume that the state space of the Markov chain $\{X_n\}$ is finite. Let

$$\psi_{ij}^{(n)}(B) = P(X_n = j, M_n \in B \mid X_0 = i) \tag{4.2}$$

and write $\boldsymbol{\Psi}^{(n)}(B) = \{\psi_{ij}^{(n)}(B)\}$, $n = 0, 1, 2, \cdots$. Having assumed that $M_0 = S_0 = 0$ we find that $\boldsymbol{\Psi}^{(0)}(B) = \boldsymbol{I}$ if and only if $0 \in B$. Finally, denote the corresponding transform by

$$\hat{\boldsymbol{\Psi}}_{\alpha\beta} = \sum_n \alpha^n \int_{R^+} e^{\beta y} \boldsymbol{\Psi}^{(n)}(dy), \tag{4.3}$$

whenever the right hand side converges.

To enable the algebraic argument of Kingman to be used we introduce a mapping as defined below.

Definition. Let μ be any finite signed measure on the Borel sets $\mathscr{B}$ of R^1. Then $\Pi\mu$ is the measure defined by

$$(\Pi\mu)(B) = \mu(x : x^+ \in B) \tag{4.4a}$$

for any $B \in \mathscr{B}$.

Using Kingman's description Π "sweeps the (probability) mass in the negative half-line up to the origin". Kingman showed that Π was a Wendel projection, by which he meant the following: let μ_+ and ν_+ be any two measures belonging to the range of Π, i.e., $\Pi\mu_+ = \mu_+$ and $\Pi\nu_+ = \nu_+$. Then $\Pi(\mu_+ * \nu_+) = \mu_+ * \nu_+$, where $*$ stands for convolution. Also, if μ_- and ν_- belong to the null-space of Π, i.e., $\Pi\mu_- = \Pi\nu_- = 0$, then $\Pi(\mu_- * \nu_-) = 0$.

We extend (4.4a) to include matrix arguments by writing for any matrix of measures $\boldsymbol{A} = \{a_{ij}\}$

$$\Pi\boldsymbol{A} = \{\Pi a_{ij}\}. \tag{4.4b}$$

It is obvious that the Wendel-property of Π is preserved if the ordinary convolution operation between measures is replaced by matrix product convolution.

We now study what effect one additional step in the random walk has upon the maximum of partial sums. This effect is seen most clearly, if the new step is added to the beginning of the random walk, i.e., as a new initial step. Let $X_n^* = X_{n+1}$ and $M_n^* = \max(0, S_2 - Y_1, \cdots, S_{n+1} - Y_1)$, $n = 0, 1, 2, \cdots$. By the definition of M_n

$$M_{n+1} = \max(0, Y_1 + M_n^*) = (Y_1 + M_n^*)^+. \tag{4.5a}$$

Consequently, the following Chapman-Kolmogorov equation holds:

$$P(X_{n+1} = j, M_{n+1} \in B \mid X_0 = i)$$

$$= \Pi\left(\sum_k \int_{R^1} P(X_1 = k, Y_1 \in dy \mid X_0 = i) P(X_n^* = j, M_n^* \in B - y \mid X_0^* = k)\right).$$

For equal initial states X_0 and X_0^* the random variables M_n and M_n^* have identical distributions. Hence the above equation can be written more compactly as

$$\Psi^{(n+1)}(B) = \Pi(Q * \Psi^{(n)})(B), \qquad n = 0, 1, 2, \cdots. \tag{4.5b}$$

The conversion to moment generating functions, multiplication of both sides by α^{n+1} and summation over n yields

$$\hat{\Psi}_{\alpha\beta} = I + \alpha\hat{\Pi}(\hat{Q}_\beta \hat{\Psi}_{\alpha\beta}). \tag{4.5c}$$

Here $\hat{\Pi}$ is the direct counterpart of Π in the space of transforms, as used by Kingman (1966).

Remark. It follows directly from the definition of Π that for any measure A, which is concentrated on the non-positive half-line, ΠA is the total mass of the measure placed at the origin. Hence in the corresponding transform, the operation $\hat{\Pi}$ on $\hat{A}_\beta$ is realised by setting $\beta = 0$.

In the following the fundamental identity has the role of a factorization theorem. We define the stopping time by

$$N^+ = \inf\{n : S_n > 0\}. \tag{4.6}$$

For the sake of clarity we write $\hat{G}_{\alpha\beta}^-$ and $\hat{H}_{\alpha\beta}^+$ for the transforms (2.9a) and (2.9b). This notation indicates the half-lines on which the corresponding measures are concentrated. By re-ordering terms in (2.10) we write the fundamental identity in the form

$$\hat{G}_{\alpha\beta}^-(I - \alpha\hat{Q}_\beta) = I - \hat{H}_{\alpha\beta}^+. \tag{4.7}$$

We now proceed to find the solution of (4.5c). Because every $\psi_{ij}^{(n)}$ is only non-zero on the non-negative half-line, and hence $\hat{\Pi}\hat{\Psi}_{\alpha\beta} = \hat{\Psi}_{\alpha\beta}$, we can write (4.5c) as

$$\hat{\Pi}((\boldsymbol{I} - \alpha\hat{\boldsymbol{Q}}_\beta)\hat{\boldsymbol{\Psi}}_{\alpha\beta}) = \boldsymbol{I}.$$

Clearly this implies

$$(\boldsymbol{I} - \alpha\hat{\boldsymbol{Q}}_\beta)\hat{\boldsymbol{\Psi}}_{\alpha\beta} = \boldsymbol{I} + \hat{\boldsymbol{K}}^-_{\alpha\beta}, \tag{4.5d}$$

where $\hat{\boldsymbol{K}}^-_{\alpha\beta}$ has the property $\hat{\Pi}\hat{\boldsymbol{K}}^-_{\alpha\beta} = \boldsymbol{0}$. Hence $\hat{\boldsymbol{K}}^-_{\alpha\beta}$ must be the transform of a matrix measure, which is concentrated on the non-positive half-line. Multiplication of (4.5d) by $\hat{\boldsymbol{G}}^-_{\alpha\beta}$ and the use of (4.7) gives

$$\hat{\boldsymbol{G}}^-_{\alpha\beta}(\boldsymbol{I} - \alpha\hat{\boldsymbol{Q}}_\beta)\hat{\boldsymbol{\Psi}}_{\alpha\beta} = \hat{\boldsymbol{G}}^-_{\alpha\beta}(\boldsymbol{I} + \hat{\boldsymbol{K}}^-_{\alpha\beta}) = (\boldsymbol{I} - \hat{\boldsymbol{H}}^+_{\alpha\beta})\hat{\boldsymbol{\Psi}}_{\alpha\beta}.$$

Now, $\hat{\boldsymbol{H}}^+_{\alpha\beta}$ and $\hat{\boldsymbol{\Psi}}_{\alpha\beta}$ belong to the range of $\hat{\Pi}$. Using the Wendel property we find, by considering the right hand term that

$$\hat{\Pi}((\boldsymbol{I} - \hat{\boldsymbol{H}}^+_{\alpha\beta})\hat{\boldsymbol{\Psi}}_{\alpha\beta}) = (\boldsymbol{I} - \hat{\boldsymbol{H}}^+_{\alpha\beta})\hat{\boldsymbol{\Psi}}_{\alpha\beta}.$$

In a similar manner we find that $\hat{\boldsymbol{G}}^-_{\alpha\beta}(\boldsymbol{I} + \hat{\boldsymbol{K}}^-_{\alpha\beta})$ is the transform of a measure, which is concentrated on the non-positive half-line. Because of this fact $\hat{\Pi}(\hat{\boldsymbol{G}}^-_{\alpha\beta}(\boldsymbol{I} + \hat{\boldsymbol{K}}^-_{\alpha\beta}))$ is obtained by putting $\beta = 0$ (see the remark after (4.5c)), and we denote the result by simply dropping the subscript β. Thus

$$\hat{\Pi}(\hat{\boldsymbol{G}}^-_{\alpha\beta}(\boldsymbol{I} + \hat{\boldsymbol{K}}^-_{\alpha\beta})) = \hat{\boldsymbol{G}}^-_{\alpha},$$

since $\hat{\boldsymbol{K}}^-_{\alpha} = \boldsymbol{0}$. Finally, the last two equations give $(\boldsymbol{I} - \hat{\boldsymbol{H}}^+_{\alpha\beta})\hat{\boldsymbol{\Psi}}_{\alpha\beta} = \hat{\boldsymbol{G}}^-_{\alpha}$, and by inversion of $\boldsymbol{I} - \hat{\boldsymbol{H}}^+_{\alpha\beta}$ Theorem 4.1 is obtained.

Theorem 4.1.

$$\left\{\sum_{n=0}^{\infty} \alpha^n E(e^{\beta M_n}, X_n = j \mid X_0 = i)\right\} = (\boldsymbol{I} - \hat{\boldsymbol{H}}^+_{\alpha\beta})^{-1}\hat{\boldsymbol{G}}^-_{\alpha}.$$

The region of validity of the above theorem has been discussed in detail by H. D. Miller (1962a, b) and is not considered here. The result corresponds to Spitzer's famous identity, which holds for independent and identically distributed summands. In this case, let $f(\cdot)$ be the common distribution of the random variables $\{Y_n\}$, and denote by $\hat{f}_\beta$ the corresponding m.g.f. Then as pointed out by Miller,

$$\exp\left(\sum_{n=1}^{\infty} \frac{\alpha^n}{n} \int_{-\infty}^{0+} e^{\beta y} f^{n*}(dy)\right)(1 - \alpha\hat{f}_\beta) = \exp\left(-\sum_{n=1}^{\infty} \frac{\alpha^n}{n} \int_{0+}^{\infty} e^{\beta y} f^{n*}(dy)\right)$$

corresponds to the factorization (4.7). Spitzer's identity has the consequent form

$$\sum_{n=0}^{\infty} \alpha^n E(e^{\beta M_n}) = \exp\left(\sum_{n=1}^{\infty} \frac{\alpha^n}{n} \int_{0+}^{\infty} e^{\beta y} f^{n*}(dy)\right)\exp\left(\sum_{n=1}^{\infty} \frac{\alpha^n}{n} P(S_n \leq 0)\right).$$

In the present construction, where the distributions of the random variables $\{Y_n\}$ may depend on a Markov chain, it seems no longer possible to obtain similar explicit expressions except in special cases. For a different formulation of the result of Theorem 4.1 see Presman (1969).

Second derivation. We now give a more direct derivation of the result given in Theorem 4.1. The Markov chain $\{X_n\}$ is also allowed to have a general state space. Instead of considering moment generating functions we here consider the actual measures.

Let $N^+ = \inf\{n: S_n > 0\}$ be as above, and define the generating functions

$$g_\alpha(x, A, B) = \sum_n \alpha^n g^{(n)}(x, A, B), \tag{4.8a}$$

and

$$h_\alpha(x, A, B) = \sum_n \alpha^n h^{(n)}(x, A, B). \tag{4.8b}$$

The former function converges at least for $|\alpha| < 1$ and the latter for $|\alpha| \leqq 1$.

Iterated convolutions of h_α can be defined as

$$h_\alpha^{n*}(x, A, B) = \int_E \int_{R^1} h_\alpha^{(n-1)*}(x, dx', dy) h_\alpha(x', A, B - y), \tag{4.9}$$

where for convenience $h_\alpha^{0*}(x, A, B) = I(x, A)I(0, B)$. A resolvent is then formed by

$$r_\alpha(x, A, B) = \sum_n h_\alpha^{n*}(x, A, B); \tag{4.10}$$

r_α converges apparently at least for $|\alpha| < 1$.

It follows immediately from the definition of N^+ that h_α corresponds to the first ascending strict ladder point of $\{S_n\}$, and as a consequence, h_α^{n*} corresponds to the nth ladder point. On the other hand, g_α is the generating function of the probabilities of not obtaining a ladder point. A necessary and sufficient condition for S_m to be maximal among n first partial sums is that

(a) S_m is a ladder point,

(b) there are no ladder points among $S_{m+1}, \cdots, S_n$.

For any given value of X_m the two requirements are independent of each other. After this observation the following is immediate:

$$\sum_{n=0}^{\infty} \alpha^n P(X_n \in A, M_n \in B \mid X_0 = x)$$

$$= \sum_{n=0}^{\infty} \alpha^n \sum_{m=0}^{n} \int_E P(X_m \in dx', S_m \in B, S_m \geqq 0, S_m > S_1, \cdots, S_m > S_{m-1} \mid X_0 = x)$$

$$\cdot P(X_n \in A, S_{m+1} - S_m \leqq 0, \cdots, S_n - S_m \leqq 0 \mid X_m = x')$$

$$= \int_E \sum_{n=0}^{\infty} h_\alpha^{n*}(x, dx', B) g_\alpha(x', A, R^-) = \int_E r_\alpha(x, dx', B) g_\alpha(x', A, R^-).$$

We have thus the following theorem.

Theorem 4.2. For all $|\alpha| < 1$

$$\sum_{n=0}^{\infty} \alpha^n P(X_n \in A, M_n \in B \mid X_0 = x) = \int_E r_\alpha(x, dx', B) g_\alpha(x', A, R^-).$$

It is easy to see that Theorem 4.1. constitutes a transformed version of this in a special case.

Next, we consider the distribution of the limiting variable

$$M = \max(0, M_1, M_2, \cdots). \tag{4.11}$$

The particular form of our result requires that the resolvent $r_\alpha(x, A, B)$ is finite for all $A \in \mathscr{E}$ and $B \in \mathscr{B}$ at $\alpha = 1$, which clearly implies that the ascending ladder process is a terminating Markov renewal process with finite lifetime, cf. Çinlar (1969a). This condition is satisfied if either

(i) the Markov chain $\{X_n\}$ is irreducible with a finite state space and $h(x, E, R^+) < 1$ holds for some x, or

(ii) the condition $h(x, E, R^+) \leqq b < 1$ holds for all $x \in E$. (Here we have indicated the evaluation at $\alpha = 1$ by dropping α as a subscript.)

We note that $1 - h(x, E, R^+)$ is the defect of the ladder process, given that $\{X_n\}$ is in state x. Let X be the value of $\{X_n\}$ at the time the ultimate maximum M is reached for the first time. Then it follows immediately (cf. Feller (1971), p. 396) that the following theorem is true.

Theorem 4.3.

$$P(X \in A, M \in B \mid X_0 = x) = \int_A r(x, dx', B)(1 - h(x', E, R^+)).$$

Remark. The above proof could have been replaced by a more formal one. This would use Theorem 4.2 as a starting point, then a factorization similar to (4.7) and finally an Abelian argument. Theorem 4.3 has a similar relation to the asymptotic formula of Täcklind (1942) and Spitzer (1956),

$$E(e^{\beta M}) = \exp\left(\sum_{n=1}^{\infty} \frac{1}{n} \int_0^{\infty} (e^{\beta y} - 1) f^{n*}(dy)\right),$$

as Theorem 4.2 to Spitzer's identity.

5. Maximum of a semi-Markov process

In the preceding section the maximum of partial sums S_n was studied, where the random walk $\{S_n\}$ was defined on a Markov chain $\{X_n\}$. Here we turn to a slightly different topic, and consider the maximal states of the Markov chain itself. Thus we define:

(5.1) $$M_n = \max(X_0, X_1, \cdots, X_n).$$

The maximal state process has been investigated by Baxter (1958) and Stone ((1968) and (1969)). See also Dinges (1969).

Despite the changed topic, the above methods of analysis apply here as well, and the notation of Section 4 is retained. For simplicity we assume that the state space of $\{X_n\}$ is discrete and totally ordered. It seems, however, that similar proofs go through for a general totally ordered state space. The distributions we want to determine are

(5.2) $$\Psi^{(n)}(B) = \{P(M_n = j, S_n \in B \mid X_0 = i)\}, \qquad n = 0, 1, 2, \cdots.$$

The corresponding transform is defined as in (4.3). The semi-Markov transition matrix $\boldsymbol{Q}(\cdot) = \{q_{ij}(\cdot)\}$ and its m.g.f. transformation $\hat{\boldsymbol{Q}}_\beta$ are defined as in (2.6) and (2.8). The sum

(5.3) $$S_0 = 0,\ S_n = Y_1 + Y_2 + \cdots + Y_n, \qquad n = 1, 2, \cdots,$$

is here most naturally interpreted as the time which is needed for n transitions in the semi-Markov process. It thus seems reasonable to assume that $\boldsymbol{Q}(0) = \boldsymbol{0}$.

We now use the idea of an ascending ladder process, and define a stopping time N^+ by

(5.4) $$N^+ = \inf\{n : X_n > X_0\}.$$

The matrices $\boldsymbol{G}^{(n)}(\cdot)$ and $\boldsymbol{H}^{(n)}(\cdot)$, and their transforms $\hat{\boldsymbol{G}}_{\alpha\beta}$ and $\hat{\boldsymbol{H}}_{\alpha\beta}$ are then defined exactly as in (2.7a), (2.7b) and (2.9a), (2.9b), respectively, by only choosing the stopping time to be N^+.

In line with the choice of the stopping time we give the following definition (cf. Stone (1968)) for the mapping Π. The relation between N^+ and Π corresponds exactly with that in Section 4.

Definition. Let $\boldsymbol{A} = \{a_{ij}\}$ be any matrix. Then Π performs the operation $\Pi\boldsymbol{A} = \{a'_{ij}\}$, where

(5.5)
$$a'_{ij} = a_{ij} \quad \text{if } j > i,$$
$$a'_{ij} = \sum_{k \le i} a_{ik} \quad \text{if } j = i,$$
and
$$a'_{ij} = 0 \quad \text{if } j < i.$$

It is now possible to obtain the distribution of the maximal state process. This is essentially a restatement of a result due to Baxter (1958).

Theorem 5.1.

$$\left\{\sum_{n=0}^{\infty} \alpha^n E(e^{\beta S_n}; M_n = j, X_0 = i)\right\} = (\boldsymbol{I} - \hat{\boldsymbol{H}}_{\alpha\beta})^{-1}\, \Pi \hat{\boldsymbol{G}}_{\alpha\beta}.$$

The equation holds at least if $|\alpha| < 1$ and $Re(\beta) \leqq 0$ or $|\alpha| \leqq 1$ and $Re(\beta) < 0$.

The proof of the theorem can be based, as in Section 4, either on an algebraic argument or on the direct use of the ladder process. For the former type of proof we note first that

$$M_{n+1} = \max(X_0, M_n^*),$$

where $M_n^* = \max(X_1, X_2, \cdots, X_{n+1})$. Thus, as in Section 4, we obtain the relation

$$\Psi^{(n+1)}(B) = \Pi(Q * \Psi^{(n)})(B). \tag{5.6}$$

Then, it only remains to verify that Π is a Wendel projection, and the proof follows. As before, the fundamental identity is used as a factorization theorem.

To complete the analogy between Section 4 and the present one, we consider the distribution of the limiting variable

$$M = \max(X_0, M_1, M_2, \cdots). \tag{5.7}$$

Again, it is assumed that the ascending ladder process in $\{X_n\}$ terminates within a finite lifetime. For this to be the case it is sufficient to assume that Condition (ii) of Section 4 holds. (In this context Condition (i) seems to be contradictory.) We write S for the value of $\{S_n\}$ when the ultimate maximum M is reached for the first time. Then a similar argument as was used in Section 4 yields the following theorem.

Theorem 5.2.

$$P(M = j, S \in B \mid X_0 = i) = r(i, j, B)(1 - h(j, E, R^+)).$$

Acknowledgements

The author wishes to thank the Director, Professor J. Gani, and the staff of the Manchester-Sheffield School of Probability and Statistics, for their advice and co-operation. Special thanks are due to Dr. T. P. Speed for many valuable discussions. The numerous constructive remarks of the referee, especially in context of Theorems 4.3 and 5.2, are gratefully acknowledged.

References

ARJAS, E. (1972a) On the asymptotic behaviour of a generalization of Markov renewal processes. *Soc. Sci. Fenn. Comment. Phys.-Math.* **42**, 17–25.

ARJAS, E. (1972b) On the use of a fundamental identity in the theory of semi-Markov queues. *Adv. Appl. Prob.* **4**, 271–284.

BAXTER, G. (1958) An operator identity. *Pacific J. Math.* **8**, 649–663.

ÇINLAR, E. (1969a) On semi-Markov processes on arbitrary spaces. *Proc. Camb. Philos. Soc.* **66**, 381–392.

ÇINLAR, E. (1969b) Markov renewal theory. *Adv. Appl. Prob.* **1**, 123–187.

DINGES, H. (1969) Wiener-Hopf-Faktorisierung für substochastische Übergangsfunktionen in angeordneten Räumen. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **11,** 152–164.

FELLER, W. (1971) *An Introduction to Probability Theory and its Applications*. Vol. 2, Second Ed. Wiley, New York.

KEMPERMAN, J. H. B. (1961) *The Passage Problem for a Stationary Markov Chain.* The University of Chicago Press.

KINGMAN, J. F. C. (1966) On the algebra of queues. *J. Appl. Prob.* **3,** 285–326.

MATTHEWS, J. P. (1971) A study of processes associated with a finite Markov chain. Ph.D. thesis, University of Sheffield (unpublished).

MILLER, H. D. (1961a) A generalization of Wald's identity with applications to random walks. *Ann. Math. Statist.* **32,** 549–560.

MILLER, H. D. (1961b) A convexity property in the theory of random variables defined on a finite Markov chain. *Ann. Math. Statist.* **32,** 1260–1270.

MILLER, H. D. (1962a) A matrix factorization problem in the theory of random variables defined on a finite Markov chain. *Proc. Camb. Philos. Soc.* **58,** 268–285.

MILLER, H. D. (1962b) Absorption probabilities for sums of random variables defined on a finite Markov chain. *Proc. Camb. Philos. Soc.* **58,** 286–298.

PRESMAN, E. L. (1969) Factorization methods and boundary problems for sums of random variables given on Markov chains. *Math. USSR Izv.* **3,** 815–852. (English translation.)

SPITZER, F. (1956) A combinatorial lemma and its application to probability theory. *Trans. Amer. Math. Soc.* **82,** 323–339.

STONE, L. D. (1968) On the distribution of the maximum of a semi-Markov process. *Ann. Math. Statist.* **39,** 947–956.

STONE, L. D. (1969) On the distribution of the supremum functional for semi-Markov processes with continuous state space. *Ann. Math. Statist.* **40,** 844–853.

TÄCKLIND, S. (1942) Sur le risque de ruine dans des jeux inéquitables. *Skand. Aktuarietidsk.* **25,** 1–42.