DISCUSSION PAPER NO. 53

MARKOV RENEWAL THEORY: A SURVEY

by

ERHAN ÇINLAR*

September 30, 1973

* Northwestern University, Evanston, Illinois

Research Supported under National Science Foundation Grant GK-36432

MARKOV RENEWAL THEORY: A SURVEY

by

ERHAN ÇINLAR

Northwestern University, Evanston, Illinois

Markov renewal theory is the union of renewal theory and Markov chains, or, better put, it is the result of the union of Markov chains with renewal processes. The answer to whether it is the Markov chain or the renewal aspect which is dominant has always eluded me however long I thought about it. I find its renewal aspect dominant on questions concerning the existence of limits, and its Markovian aspect on the evaluation of those limits, and quite often it is neither or both.

In its female aspect a Markov renewal process appears as a family of renewal processes, possibly infinite in number, and each of a different colour, intertwined by being superposed together in such a fashion as to arrange their colours in a harmonious sequence pleasing to observe: the colour of the first point depending on that of the original one, and the colour of the next point selected in accordance with that of the first, and this rule being repeated forever. This invests in the Markov renewal process an awesome regenerative power, each point along with its colour holding the key to the secrets of all the future beyond.

In its male aspect a Markov renewal process may be likened to an account of the shapes assumed by Zeus in his varied amorous pursuits. Initially, having banished his father, he pursues his mother and violates her in the shape of a serpent; then he courts his twin sister Hera in his own shape, and this failing, in the disguise of a bedraggled cuckoo; afterwards he rapes Maia, daughter of Atlas, while in his own shape; and then

---

Research supported by the National Science Foundation under Grant GK-36432.

Leto, daughter of Titans Coeus and Phoebe, transforming himself and her into quails when they couple; and later, after the creation of man, he disguises himself as a mortal and has a love affair with Semele, daughter of King Cadmus of Thebes; and still later, begats Eros on his own daughter Aphrodite in his own shape; etc. The successive shapes he assumes form a Markov chain (though it was then called something else), and the lengths of the intervals between the moments of heavenly union are conditionally independent of each other once the successive shapes are known.

Having thus established the divine importance of our subject I shall next put down a crude introduction into its nature, along with a description of the major rituals performed at its temples. The already initiated will find very little which is new. My present purposes lead me to keep the notation and formalities to a minimum, and, partly because of this, to limit myself to the finite state space case and merely to point out the nature of the difference from the infinite state space case if any. The literature, starting with about 5 titles in 1960, has grown to over 300 by 1972. This rapid growth, no doubt spurred by PYKE [53],[54] and the recognition of many applications in a large number of fields, makes it difficult for me to give due credit to everyone concerned. I apologize beforehand to all those who might feel left out. Instead of presenting merely the results, I have tried to capture the flavor of the field and to give a highly stylized presentation in the unified framework of ÇINLAR [11], which is, essentially, the analog of FELLER's treatment of renewal theory. As a result, most of the results I will be listing are presented in the manner which seems to me the most natural in view of the present day state of the field. Further, instead of merely presenting results, I have endeavored to point out the main lines of attack and the reasons for it. For example, in the applications, I felt it better to tell where an imbedded process is and how it can be exploited best rather than showing the details of computations.

1. MARKOV RENEWAL PROCESSES

Let E be a finite set, $\mathbb{N}$ the set of non-negative integers and $\mathbb{R}_+$ the non-negative real numbers. Suppose we have, on a probability space $(\Omega,\mathcal{H},P)$, random variables

$$X_n : \Omega \to E$$
$$T_n : \Omega \to \mathbb{R}_+$$

defined for each $n \in \mathbb{N}$ so that

$$0 = T_0 \leq T_1 \leq T_2 \leq \cdots.$$

These elements are said to form a <u>Markov renewal process</u> (X,T) with state space E provided that

$$(1.1) \qquad P\{X_{n+1} = j,\ T_{n+1} - T_n \leq t \mid X_0,\ldots,X_n;T_0,\ldots,T_n\} = P\{X_{n+1} = j,\ T_{n+1} - T_n \leq t \mid X_n\}$$

for all $n \in \mathbb{N}$, $j \in E$, $t \in \mathbb{R}_+$. We will always assume that the process is temporally homogeneous: that is,

$$(1.2) \qquad P\{X_{n+1} = j,\ T_{n+1} - T_n \leq t \mid X_n = i\} = Q(i,j,t)$$

independent of n. The family of probabilities

$$Q = \{Q(i,j,t);\ i,j \in E,\ t \in \mathbb{R}_+\}$$

is called a <u>semi-Markovian kernel</u>. For our present purposes we assume $Q(i,j,0) = 0$ for all i,j in E.

For each pair (i,j) the function $t \to Q(i,j,t)$ is right continuous non-decreasing and bounded. Defining

$$(1.3) \qquad P(i,j) = \lim_{t\to\infty} Q(i,j,t)$$

we see that

$$(1.4) \qquad P(i,j) \geq 0, \quad \sum_{j \in E} P(i,j) = 1; \quad i,j \in E;$$

that is, the $P(i,j)$ are the transition probabilities for some Markov chain with state space E. It follows from (1.1) and (1.3) that, in fact,

$$(1.5) \qquad P\{X_{n+1} = j \mid X_0, \ldots, X_n; T_0, \ldots, T_n\} = P(X_n, j)$$

for all $n \in \mathbb{N}$ and $j \in E$. This implies, in particular, the following

(1.6) PROPOSITION. $X = (X_n)_{n \in \mathbb{N}}$ is a Markov chain with state space E and transition probability matrix P.

If $P(i,j) = 0$ for some pair (i,j) then $Q(i,j,t) = 0$ for all $t \in \mathbb{R}_+$ and we define the ratio $Q(i,j,t)/P(i,j)$ to be unity. With this convention we define

$$(1.7) \qquad G(i,j,t) = Q(i,j,t)/P(i,j), \quad i,j \in E, \ t \in \mathbb{R}_+.$$

Then, for each pair (i,j), the function $G(i,j,\cdot)$ is a distribution function. From (1.2) and (1.5) we see that

$$(1.8) \qquad G(i,j,t) = P\{T_{n+1} - T_n \leq t \mid X_n = i, \ X_{n+1} = j\}.$$

Using this interpretation together with (1.1) we can show by induction the following

(1.9) PROPOSITION. For any integer $n \geq 1$ and numbers $u_1, \ldots, u_n \in \mathbb{R}_+$,

$$P\{T_1 - T_0 \leq u_1, \ T_2 - T_1 \leq u_2, \ldots, T_n - T_{n-1} \leq u_n \mid X_0, X_1, \ldots\}$$
$$= G(X_0, X_1; u_1) G(X_1, X_2; u_2) \cdots G(X_{n-1}, X_n; u_n);$$

in words, the increments $T_1 - T_0$, $T_2 - T_1, \ldots$ are conditionally independent given the Markov chain $X_0, X_1, \ldots$.

In particular, if the state space E consists of a single point, then the increments are independent and identically distributed non-negative random variables; namely, we have

(1.10) COROLLARY. If E consists of a single point then $(T_n)_{n \in \mathbb{N}}$ is a renewal process.

This result together with Proposition (1.6) justify the term Markov renewal process, somewhat, by exhibiting it as a generalization of Markov chains and renewal processes. The full justification however is contained in Proposition (1.6) and the following very important

(1.11) PROPOSITION. Let $i \in E$ be fixed and define $S_0^i, S_1^i, \ldots$ to be the successive T_n for which $X_n = i$. Then $S^i = (S_n^i)_{n \in \mathbb{N}}$ is a (possibly delayed) renewal process.

Thus to each state i there corresponds a renewal process (S_n^i); the superposition of all these renewal processes gives the points T_n, $n \in \mathbb{N}$; the renewal process which contributed the point T_n is the i^{th} one if and only if $X_n = i$; the types of the successive points, namely $X_0, X_1, \ldots$, form a Markov chain. This is the Markov renewal process in its female aspect. Next we describe its male aspect.

Since (X_n) is a Markov chain with a finite state space, not all the states are transient and the chain, starting from any initial state, eventually will reach some recurrent state with probability one. For any recurrent state i, X_n becomes i infinitely often, thus causing the renewal process (S_n^i) to be non-terminating which in turn implies that $\sup_n S_n^i = +\infty$ with probability one. Since the supremum of the T_n is greater than or equal to $\sup_n S_n^i$ for any i, this argument proves the following

(1.12) PROPOSITION. If E is finite, then $\sup_n T_n = +\infty$ almost surely.

By weeding out those $\omega \in \Omega$ for which $\sup_n T_n(\omega) < \infty$ we may assume, and we do, that $\sup_n T_n(\omega) = +\infty$ for all ω. Then, for any $\omega \in \Omega$ and $t \in \mathbb{R}_+$ there is some integer n such that $T_n(\omega) \leq t < T_{n+1}(\omega)$. We can therefore define a continuous-time parameter process $Y = (Y_t)_{t \in \mathbb{R}_+}$ with state space E by putting

$$Y_t = X_n \quad \text{on } \{T_n \le t < T_{n+1}\}. \tag{1.13}$$

The process $Y = (Y_t)_{t \in \mathbb{R}_+}$ so defined is called a <u>semi-Markov process</u> with state space E and semi-Markovian transition kernel $Q = \{Q(i,j,t)\}$.

The semi-Markov process Y provides a picture which is convenient in describing the Markov renewal process underlying it. We may think of Y_t as the state at time t of some system or particle (or deity) which moves from one state to another with random sojourn times in between. The length of a sojourn interval $[T_n, T_{n+1})$ is a random variable whose distribution depends both on the state X_n being visited and the state X_{n+1} to be entered next. The successive states visited form a Markov chain and, conditional on that sequence, the successive sojourn times are independent. This is the Markov renewal process in its male aspect.

Using the description above we may obtain a realization of the semi-Markov process Y in the following manner. Suppose we have the transition probabilities $P(i,j)$ satisfying (1.4) prescribed along with a distribution function $G(i,j,t)$ for each pair (i,j). Suppose the initial state is to be i. First the next state to be entered is sampled from the distribution $P(i,\cdot)$; if the outcome is j, then a sojourn time u is sampled from the distribution $G(i,j,\cdot)$; the function y_t is set to be i for all $t < u$ and y_u is set to be j. The second step starts by sampling the next state to be entered from the distribution $P(j,\cdot)$; if the outcome is k, then a sojourn time v is sampled from the distribution $G(j,k,\cdot)$; the function y_t is now set to be j for $t \in [u, u+v)$ and y_{u+v} is set to be k. The third step .... The resulting function y_t is a realization of the semi-Markov process Y with the semi-Markovian kernel Q given as $Q(i,j,t) = P(i,j)G(i,j,t)$.

A different causal relationship which leads to the same mathematical model is as follows. Suppose we are given, for each $i \in E$, a distribution function $t \to H(i,t)$ on $\mathbb{R}_+$ and, for each $i \in E$ and $t \in \mathbb{R}_+$, probabilities $K(i,t,j)$ so that $\Sigma_j K(i,t,j) = 1$. Suppose y_0 is to be i. First a sojourn time is sampled from the distribution $H(i,\cdot)$; if the outcome is u, then y_t is set to be i for all $t < u$, and the state to be entered at time u is sampled from the distribution $K(i,u,\cdot)$. Supposing the outcome for y_u to be j, the second step starts with sampling a sojourn time from the distribution $H(j,\cdot)$; if this outcome is v then y_t is set to be j for all $t \in [u,u+v)$, and the state to be entered next at time $u+v$ is sampled from the distribution $K(j,v,\cdot)$. Supposing the outcome y_{u+v} to be $k,\ldots$; and so on. The resulting function y_t is the realization of a semi-Markov process with the semi-Markovian kernel Q defined by

$$Q(i,j,t) = \int_0^t H(i,du)K(i,u,j)$$

for all $i,j \in E$ and $t \in \mathbb{R}_+$.

Finally, a word of justification for the name semi-Markov process. If the semi-Markovian kernel Q is of the form

$$Q(i,j,t) = P(i,j)(1 - e^{-\lambda(i)t}), \tag{1.14}$$

then one can show that

$$P\{Y_{t+s} = j \mid Y_u;\ u \leq t\} = P\{Y_{t+s} = j \mid Y_t\} \tag{1.15}$$

for all $t,s \in \mathbb{R}_+$ and $j \in E$; and furthermore,

$$P\{Y_{t+s} = j \mid Y_t = i\} = P_s(i,j) \tag{1.16}$$

is independent of t. In other words, under (1.14) the semi-Markov process becomes a (temporally homogeneous) Markov process. Conversely, any Markov process $Y = (Y_t)_{t \in \mathbb{R}_+}$ with a finite state space E is automatically a semi-

Markov process; if $A(i,j)$ denotes the right-hand derivative at $t = 0$ of its transition function $P_t(i,j)$, then the corresponding semi-Markov kernel has the form (1.14) with

$$\lambda(i) = -A(i,i),\ P(i,i) = 0,\ P(i,j) = A(i,j)/\lambda(i)$$

for all i and $j \neq i$ with $A(i,i) \neq 0$, and if $A(i,i) = 0$,

$$\lambda(i) = 1,\ P(i,i) = 1,\ P(i,j) = 0$$

for all $j \neq i$.

Thus, with respect to Markov processes, the novel feature of the semi-Markov process is the freedom allowed in the choice of the distributions of the sojourn times; this freedom, however, is achieved at the expense of the Markov property (1.15) which, instead of holding for all t, holds now only for the jump times T_n.

Semi-Markov processes were introduced by LÉVY [42] and SMITH [60] independently and simultaneously. At about the same time TÁKACS [63],[64] had studied a process which is in many respects equivalent to it. The Markov renewal processes were studied in detail by PYKE [53],[54]. Our terminology follows [11] which presents the theory from the point of view of "semi-regenerative" processes and the analytic techniques of Markov renewal equations paralleling renewal theory.

2. PRELIMINARY RESULTS

Throughout this section $(X,T) = (X_n,T_n)_{n \in \mathbb{N}}$ will be a Markov renewal process with a semi-Markovian kernel Q over a finite space E. Then $X = (X_n)_{n \in \mathbb{N}}$ is a Markov chain on E with transition probabilities $P(i,j) = Q(i,j,+\infty)$.

We will be denoting the conditional probability $P\{\cdot \mid X_0 = i\}$ simply by $P_i\{\cdot\}$ and will write E_i for the expectation corresponding to it.

Let us define

(2.1) $$Q^n(i,j,t) = P_i\{X_n = j,\ T_n \le t\}, \qquad i,j \in E,\ t \in \mathbb{R}_+$$

for all $n \in \mathbb{N}$. Then,

(2.2) $$Q^0(i,j,t) = I(i,j)$$

for all $t \geq 0$ where $I(i,j)$ is 1 or 0 according as $i = j$ or $i \neq j$ (namely, $I(i,j)$ is the (i,j)-entry of the identity matrix I); and for $n \geq 0$ we have the recursive relation

(2.3) $$Q^{n+1}(i,k,t) = \sum_{j \in E} \int_0^t Q(i,j,du)Q^n(j,k,t-u);$$

(for which can be given the following plausibility argument: in order for $\{X_{n+1} = k,\ T_{n+1} \le t\}$ to happen, the first transition must occur at some time $u \le t$ into some state j and then, starting from j, the remaining n transitions should take no longer $t-u$ and end with state k).

We had noted, by Proposition (1.11), that the times T_n for which $X_n = j$ is a (possibly delayed) renewal process for each fixed j. The number of renewals, in $[0,t]$, belonging to this renewal process is

(2.4) $$\sum_{n=0}^{\infty} I_{\{X_n = j,\ T_n \le t\}},$$

(we denote by I_A the indicator function of the event A). Since the number of renewals in a finite interval has a finite expected value, this random number has the <u>finite</u> expectation

(2.5) $$R(i,j,t) = E_i\Big[\sum_{n=0}^{\infty} I_{\{X_n = j,\ T_n \le t\}}\Big] = \sum_{n=0}^{\infty} Q^n(i,j,t),$$

for any initial state i.

The $R(i,j,\cdot)$ are called <u>Markov renewal functions</u>, and $R = \{R(i,j,t);\ i,j \in E,\ t \in \mathbb{R}_+\}$ is called the Markov renewal kernel corresponding to Q. Because R plays a very important role in the theory, we stress once more the fact that for fixed $i,j \in E$ the function $t \to R(i,j,t)$

is a renewal function.* If the initial state is j then the entrances to j form an ordinary renewal process and

$$R(j,j,t) = \sum_{n=0}^{\infty} F^n(j,j,t) \tag{2.6}$$

where $F^n(j,j,\cdot)$ is the n-fold convolution of the function $F(j,j,\cdot)$ which is the distribution of the time between two occurrences of state j. On the other hand, if the initial state is $i \neq j$, the time until the first visit to j has a distribution $F(i,j,\cdot)$ which might be different from $F(j,j,\cdot)$. In this case, namely when $i \neq j$, we have

$$R(i,j,t) = \int_0^t F(i,j,du)R(j,j,t-u). \tag{2.7}$$

Once the $R(i,j,\cdot)$ are obtained by the formula (2.5) the expressions (2.6) and (2.7) may be used to solve for the first passage time distributions $F(i,j,\cdot)$. Computationally, especially in the present case of finite state space E, it is convenient to use Laplace transforms.

We define, for $\lambda \geq 0$,

$$Q_\lambda(i,j) = \int_0^\infty e^{-\lambda t}\, Q(i,j,dt), \tag{2.8}$$

$$F_\lambda(i,j) = \int_0^\infty e^{-\lambda t}\, F(i,j,dt), \tag{2.9}$$

and for $\lambda > 0$,

$$R_\lambda(i,j) = \int_0^\infty e^{-\lambda t}\, R(i,j,dt). \tag{2.10}$$

Then, it follows from (2.3) that the Laplace transform $Q_\lambda^n(i,j)$ of the mass $Q^n(i,j,\cdot)$ satisfies

$$Q_\lambda^{n+1}(i,k) = \sum_{j \in E} Q_\lambda(i,j)Q_\lambda^n(j,k) \tag{2.11}$$

for all $n \in \mathbb{N}$. Hence, the matrix Q_λ^n is precisely the n^{th} power of the matrix Q_λ; that is,

---

*This fact sometimes goes unnoticed by authors who are too busy computing to have any time left for reflection.

(2.12) $$Q_\lambda^n = (Q_\lambda)^n$$

for all $n \geq 1$ and also for $n = 0$ with the usual convention that the zeroth power of any matrix is the identity matrix.

Using the matrix notation, the expression (2.5) implies

(2.13) $$R_\lambda = I + Q_\lambda + Q_\lambda^2 + \cdots, \qquad \lambda > 0.$$

By virtue of the interpretation (2.12) we can write

(2.14) $$R_\lambda(I - Q_\lambda) = (I - Q_\lambda)R_\lambda = I$$

which means that, for $\lambda > 0$, R_λ is the inverse of $I - Q_\lambda$:

(2.15) $$R_\lambda = (I - Q_\lambda)^{-1}.$$

Next, considering the Laplace transforms (2.9) of the first passage distributions, we obtain from (2.6) and (2.7) that

(2.16) $$R_\lambda(i,j) = \begin{cases} F_\lambda(i,j)R_\lambda(j,j) & \text{if } i \neq j, \\ [1 - F_\lambda(j,j)]^{-1} & \text{if } i = j. \end{cases}$$

Conversely,

(2.17) $$F_\lambda(i,j) = \begin{cases} R_\lambda(i,j)/R_\lambda(j,j) & \text{if } i \neq j, \\ 1 - 1/R_\lambda(j,j) & \text{if } i = j. \end{cases}$$

These show that the matrices Q_λ, F_λ, R_λ define each other uniquely.

If the state space E were infinite, all the expressions we have above would hold except possibly (2.14) and (2.15) which do not even make sense in that case. From (2.13) we could write $R_\lambda = I + Q_\lambda R_\lambda$ which is an infinite system of linear equations. However, in general, this system may have more than one solution. Then, R_λ turns out to be the minimal solution of this system $M = I + Q_\lambda M$. We refer to [11] for a complete discussion.

Aside from the computations given here, a large number of quantities of possible interest are computed in [10],[27],[41],[45].

3. CLASSIFICATION OF STATES

Let (X,T) be a Markov renewal process and let S^i be the renewal process formed by all the T_n for which $X_n = i$. Suppose the initial state is i so that S^i is an ordinary renewal process. We define state i to be recurrent (transient, aperiodic, periodic with period λ respectively) according as the renewal process S^i is recurrent (transient, aperiodic, periodic with period λ respectively).

State i is recurrent if and only if the recurrence time distribution $F(i,i,\cdot)$ is honest: that is, if and only if $F(i,i) = F(i,i,+\infty) = 1$. Then, $R(i,i) = R(i,i,\infty) = +\infty$ and since $R(i,i) = \sum_n P^n(i,i)$ we see that i is recurrent if and only if i is recurrent in the Markov chain X. Similarly, i is transient if and only if it is transient in the Markov chain X; then, $F(i,i) = F(i,i,+\infty) < 1$ and $R(i,i) = R(i,i,\infty) = 1/(1 - F(i,i)) < \infty$ and these quantities $F(i,i)$ and $R(i,i)$ are computed directly from the Markov chain X. So, recurrence and transientness are just as in Markov chains.

Periodicity of a state i in (X,T), however, has nothing to do with the periodicity of i in the Markov chain X. State i is periodic in (X,T) if and only if the distribution $F(i,i,\cdot)$ is arithmetic. In view of the fact that the $F(i,i,\cdot)$ are not easy to compute, the following solidarity theorem and the theorem after that are pleasant results.

(3.1) PROPOSITION. If i and j can be reached from each other, then either they are both aperiodic or else they are both periodic. In the latter case they both have the same period.

For the proof of this, as well as the results we are about to mention, we refer to [1] and [17].

In view of the Proposition (3.1) all states belonging to an irreducible closed set C of the Markov chain X behave similarly: either all states in C are aperiodic, or they are all periodic with the same period. To determine periodicity we need not compute the distributions $F(i,j,\cdot)$. We introduce

(3.2) $$\lambda_{ij} = \inf\{t \geq 0: Q(i,j,t) > 0\} \qquad i,j \in E$$

with the usual convention that if the set in question is empty then its infimum is $+\infty$.

(3.3) PROPOSITION. Let $C \subset E$ be an irreducible closed set of periodic states with the common period λ. Then, for any $i,j \in E$ for which $P(i,j) = Q(i,j,\infty) > 0$, all the jumps of the function $t \to Q(i,j,t)$ belong to the set

(3.4) $$\{\lambda_{ij}, \lambda_{ij} + \lambda, \lambda_{ij} + 2\lambda, \ldots\}.$$

Furthermore, for any $n \in \mathbb{N}$ and any sequence of states $j_0, j_1, \ldots, j_n \in C$ with $j_n = j_0$, we have

(3.5) $$\lambda_{j_0 j_1} + \lambda_{j_1 j_2} + \cdots + \lambda_{j_{n-1}, j_n} = k\lambda$$

for some integer k (possibly infinite).

In particular, (3.5) implies that λ_{ii} is a constant multiple of λ. Moreover, (3.5) shows that λ is the greatest common divisor of all finite sums $\lambda_{j_0 j_1} + \cdots + \lambda_{j_{n-1} j_n}$ with $j_0 = j_n$. In view of the first statement that all the $Q(i,j,\cdot)$ are to be step functions with all jumps occurring at points of form $\lambda_{ij} + k\lambda$, the aperiodicity is not hard to recognize. In particular, if any one of the $Q(i,j,\cdot)$ has a derivative, then none of the states in the same recurrent class as i and j can be periodic.

4. MARKOV RENEWAL EQUATIONS

Consider a Markov renewal process (X,T) with state space E and semi-Markovian kernel Q, and let $R = \Sigma\, Q^n$ be the corresponding Markov renewal kernel. The equations we will introduce form a system of integral equations which generalize the renewal equation in a natural manner. The class of functions which we will be working with, to be denoted by $\mathbb{B}$, are functions

$$f: E \times \mathbb{R}_+ \to \mathbb{R}$$

such that, for every $i \in E$, the mapping $t \to f(i,t)$ is Borel measurable and bounded over finite intervals. In particular, all functions f on $E \times \mathbb{R}_+$ which are continuous in the second variable belong to $\mathbb{B}$ as well as all those which are right continuous and monotone in the second variable. As special examples we note that, for any fixed $j \in E$, the mappings $(i,t) \to Q^n(i,j,t)$ and $(i,t) \to R(i,j,t)$ both belong to $\mathbb{B}$.

A function $f \in \mathbb{B}$ is said to satisfy a Markov renewal equation if

$$(4.1) \qquad f(i,t) = g(i,t) + \sum_{j \in E} \int_0^t Q(i,j,du) f(j,t-u), \quad i \in E,\ t \in \mathbb{R}_+$$

for some function $g \in \mathbb{B}$. Here the point of view is that g and Q are known and the problem is to solve for f and to study the limiting behavior of $t \to f(i,t)$ as t gets large.

If E consists of a single point, then the equation (4.1) is nothing more than a renewal equation. Just as in that simpler case, the system (4.1) has a unique solution with the Markov renewal kernel R playing the role of the renewal function. Here is the main result.

(4.2) PROPOSITION. The Markov renewal equation (4.1) has one and only one solution; it is

$$(4.3) \qquad f(i,t) = \sum_{j \in E} \int_0^t R(i,j,ds) g(j,t-s), \quad i \in E,\ t \in \mathbb{R}_+.$$

NOTATION. For any $g \in \mathbb{B}$ the function $Q * g$ defined by

$$(4.4) \qquad Q * g(i,t) = \sum_{j \in E} \int_0^t Q(i,j,ds) g(j,t-s), \quad i \in E,\ t \in \mathbb{R}_+$$

is well defined and belongs to $\mathbb{B}$. Hence, the iterates of this operator on $\mathbb{B}$ are all well defined and the n^{th} iterate $Q^n * g$ is given by (4.4) with Q there replaced by Q^n which was defined recursively by (2.1)-(2.3). We can replace Q on the right side of (4.4) by R and still have a well defined function in $\mathbb{B}$ which we will denote by $R * g$. Further, in (4.4), if g is $R(\cdot,k,\cdot)$ for some fixed k, we write $Q * R(i,k,t)$ for the left side.

PROOF of (4.2). In the notation of (4.4), the Markov renewal function (4.1) becomes

$$f = g + Q * f, \tag{4.5}$$

and the assertion is that

$$f = R * g \tag{4.6}$$

is the only solution.

That $R * g$ is a solution follows easily from the relation $R = I + Q * R$ which is the same, in the present notation, as the equation (2.13).

To show that $R * g$ is the only solution, first note that, if f is any other solution, then $h = f - R * g$ must satisfy

$$h = Q * h, \qquad h \in \mathbb{B}, \tag{4.7}$$

and we will show that this implies $h = 0$.

By iteration (4.7) gives $h = Q * h = Q^2 * h = \cdots = Q^n * h = \cdots$. For fixed t, by the fact that $h \in \mathbb{B}$, $|h(i,s)| \leq c$ for all $i \in E$ and $s \leq t$ for some constant c. Thus,

$$\begin{aligned} |h(i,t)| &= |Q^n * h(i,t)| \\ &\leq \sum_j \int_0^t Q^n(i,j,du)\,|h(j,t-u)| \leq c \sum_j Q^n(i,j,t) = cP_i\{T_n \leq t\}. \end{aligned} \tag{4.8}$$

But E is finite and Proposition (1.12) applies to the effect that $\sup T_n = \infty$. Hence, as $n \to \infty$,

$$P_i\{T_n \leq t\} \to 0.$$

This shows that, by (4.8), $h = 0$. □

The finiteness of the state space E enters into the preceding theorem at two points: existence of $R * g$ in $\mathbb{B}$ and non-finiteness of $\sup T_n$. All the above results hold therefore for countably infinite spaces E also provided that these two particular points have been satisfied. However, especially the second point is not easy to settle in that more general case. We refer to Chapters III and IV of [11] for a full treatment in the general case.

Markov renewal equations were introduced in [11] without the knowledge of the earlier paper by FABENS and KARLIN [23]. They introduce the same equations in the finite state space case and obtain the solution under conditions meant to insure the finiteness of $R(i,j,t)$ (which always is).

5. LIMIT THEOREMS

We are using the same notation as in the preceding section. In particular, Q is a semi-Markovian kernel on E and R is the corresponding Markov renewal kernel. Since we assume E to be finite, most of our main results will follow from renewal theory.

For fixed $j \in E$, the function $R(j,j,\cdot)$ is an ordinary renewal function. If j is transient, then the limit as $t \to \infty$ of $R(j,j,t)$ is simply

(5.1) $$\lim_{t\to\infty} R(j,j,t) = R(j,j,+\infty) = R(j,j)$$

which is also the expected number of visits to j by the Markov chain X starting at j. This quantity, as well as the probability $F(i,j) = F(i,j,\infty)$ of ever reaching j from i, can be computed directly by the well known methods of Markov chains.

If j is recurrent then $R(j,j) = +\infty$ and the relevant limit theorem states that

(5.2) $$\lim_{t\to\infty}[R(j,j,t) - R(j,j,t-\tau)] = \tau\, \eta(j)$$

for all $\tau > 0$ if j is aperiodic and for τ of form $\tau = n\lambda$ if j is periodic with period λ. Here $\eta(j)$ is the inverse of the mean recurrence time of j, that is,

$$(5.3) \qquad \eta(j) = 1/\int_0^\infty t\, F(j,j,dt)$$

with the convention that $1/\infty = 0$. We refer to FELLER [25] for a proof of (5.2) and of the next result which is in fact equivalent to (5.2). If j is recurrent aperiodic, then for any directly Riemann integrable function g we have

$$(5.4) \qquad \lim_{t\to\infty} \int_0^t R(j,j,ds)g(t-s) = \eta(j)\int_0^\infty g(u)du,$$

and a similar result holds in the periodic case. For the concept of direct Riemann integrability we again refer to FELLER [25, p. 348].

We write $\mathbb{D}$ for the class of all directly Riemann integrable functions. These are non-negative (finite) functions defined on $\mathbb{R}_+$ and satisfying a certain summability condition. Every function in $\mathbb{D}$ is Riemann integrable in the ordinary sense. Any continuous function vanishing outside a finite interval is in $\mathbb{D}$. Every Riemann integrable decreasing function is in $\mathbb{D}$. ~~Every Riemann integrable decreasing function is in $\mathbb{D}$~~. Any right continuous function dominated by a function in $\mathbb{D}$ is again in $\mathbb{D}$. The convolution $\phi * g$ of a function g in $\mathbb{D}$ with a distribution ϕ is in $\mathbb{D}$.

The Following is on computing $\eta(j)$ directly from Q. As far as computing $\eta(j)$ is concerned, it is sufficient to consider only the recurrent irreducible class j belongs to. Therefore, the next proposition should be applied to each irreducible class separately.

(5.5) PROPOSITION. Suppose X is irreducible recurrent and let π be an invariant measure for X, that is, let π be a solution of

$$(5.6) \qquad \sum_{i\in E} \pi(i)P(i,j) = \pi(j), \qquad j \in E.$$

Further, let $m(i)$ be the mean sojourn time in state i, that is,

$$(5.7)\qquad m(i) = E[T_{n+1} - T_n | X_n = i] = \int_0^\infty [1 - \sum_k Q(i,k,t)]dt.$$

Then,

$$(5.8)\qquad \eta(j) = \pi(j)/\sum_{i \in E} \pi(i)m(i), \qquad j \in E.$$

PROOF will be given later. First, a few remarks are in order. Since X is irreducible recurrent, the system of linear equations (5.6) has a solution, and the solution is unique up to multiplication by a constant. Because of the form of (5.8) this constant factor plays no role. In particular, we may take π to be the limiting distribution by normalizing it to have $\sum \pi(j) = 1$. Inverting (5.8) we see that the mean recurrence time of state j is $\sum_i m(i)[\pi(i)/\pi(j)]$ which has a good intuitive explanation once we note that $\pi(i)/\pi(j)$ is the expected number of visits to i in between two visits to j, and that each visit to i lasts a time whose expectation is $m(i)$.

The following is the main limit theorem.

(5.9) THEOREM. Suppose X is irreducible recurrent and $g(j,\cdot)$ is directly Riemann integrable for every $j \in E$. Then, if the states are aperiodic in (X,T),

$$(5.10)\qquad \lim_{t\to\infty} R * g(i,t) = \sum_j \pi(j)n(j)/\sum_j \pi(j)m(j)$$

where π and m are as in (5.5) and

$$(5.11)\qquad n(j) = \int_0^\infty g(j,s)ds, \qquad j \in E.$$

If the states are periodic with period λ, then (5.10) holds for $t = x + k\lambda$, $k \in \mathbb{N}$, $k \to \infty$, and $x \in [0,\lambda]$, with

$$(5.12)\qquad n(j) = \lambda \sum_{k=1}^\infty g(j, x + k\lambda - \lambda_{ij})$$

where $\lambda_{ij} = \inf\{t: F(i,j,t) > 0\}$ (modulo λ).

PROOF. Since E is finite the limit in question is

$$(5.13)\qquad \lim_{t\to\infty} R*g(i,t) = \sum_j \lim_{t\to\infty} \int_0^t R(i,j,ds)g(j,t-s).$$

To obtain the limit as $t\to\infty$ of the j-term, we note that $R(i,j,\cdot)$ is the convolution of $F(i,j,\cdot)$ with $R(j,j,\cdot)$. Thus, the j-term is the convolution of $R(j,j,\cdot)$ with the function $F(i,j,\cdot)*g(j,\cdot)$. Since $g\in\mathbf{D}$, so is $F(i,j,\cdot)*g(j,\cdot)$ and

$$\int_0^\infty dt \int_0^t F(i,j,ds)g(j,t-s) = F(i,j,\infty)\int_0^\infty g(j,s)ds = n(j)$$

with $n(j)$ defined by (5.11) since $F(i,j,\infty) = 1$ by irreducibility. Now applying (5.4) in the aperiodic case we obtain

$$(5.14)\qquad \lim_t \int_0^t R(i,j,ds)g(j,t-s) = \eta(j)n(j), \qquad j\in E.$$

Putting (5.14) into (5.13) and then using Proposition (5.5) to evaluate the $\eta(j)$ we obtain the desired result. We omit the proof in the periodic case. □

If X is reducible and i is recurrent, then the limit of $R*g(i,t)$ may be obtained by considering only the recurrent class i belongs to. However, if i is transient, and especially if there are periodic classes which can be reached from i, the limit of $R*g(i,t)$ becomes harder to obtain. For a thorough treatment of these and related matters we refer to [17].

PROOF of (5.5). We give this in the aperiodic case as an illustration of the use of (5.13) and (5.14). If $m(j) = \infty$ for any j, the mean recurrence time of j being greater than the mean sojourn time $m(j)$ in j, we must have $\eta(j) = 1/\infty = 0$. Thus (5.8) holds trivially if $m(j) = \infty$ for any j. Suppose now that all the $m(j)$ are finite.

Let $h(j,t) = 1 - \sum_k Q(j,k,t)$. Clearly $h(j,\cdot)$ is non-increasing and its integral is defined to be $m(j)$ in (5.7). Since $m(j) < \infty$, $h(j,\cdot)$ is directly integrable. Applying (5.13),(5.14) with g there replaced by h we obtain

$$\lim_{t\to\infty} R*h(i,t) = \sum_j \eta(j)m(j).$$

But $h = 1 - Q*1$ and $R = I + Q*R$; thus, $R*h = 1$ identically and we must have

$$\sum_j \eta(j)m(j) = 1. \tag{5.15}$$

Next, for fixed $\tau > 0$ and fixed $j \in E$ define g by

$$g(i,t) = \begin{cases} 1 & \text{if } i = j,\ t \leq \tau, \\ 0 & \text{otherwise.} \end{cases}$$

Note that $g(i,\cdot)$ is directly integrable, and that $Q*g(i,\cdot)$ is directly integrable with

$$\begin{aligned} \int_0^\infty Q*g(i,s)ds &= \int_0^\infty ds[Q(i,j,s) - Q(i,j,s-\tau)] \\ &= \tau Q(i,j,+\infty) = \tau P(i,j). \end{aligned}$$

Since $R = I + R*Q$, $R*g = g + R*(Q*g)$ and we have

$$\lim_t R*g(i,t) = \lim_t R*Q*g(i,t).$$

Applying (5.13) and (5.14) to both sides separately we obtain

$$\eta(j)\tau = \tau \sum_i \eta(i)P(i,j).$$

In other words, η is a solution of (5.6). From the uniqueness up to multiplication of such solutions it follows that

$$\eta(j) = c\,\pi(j), \qquad j \in E \tag{5.16}$$

for some constant c. In view of (5.15) that constant must be

$$c = 1/\Sigma\, \pi(i)m(i). \tag{5.17}$$

The desired result now follows from (5.16) and (5.17) □

We end this section with the following two results which complement the preceding one. We omit the proofs.

(5.18) LEMMA. Let ϕ be a non-decreasing function defined on $\mathbb{R}_+$ and with $\phi(t+b) - \phi(t) \leq c$ for all t for some constants c and b > 0. Then, for any non-decreasing right continuous function g we have

$$\lim_{t\to\infty} \frac{\phi * g(t)}{\phi(t)} = g(+\infty).$$

Note that each $R(i,j,\cdot)$ satisfies the hypothesis of (5.18) regarding ϕ. Thus, letting $\phi = R(j,j,\cdot)$ and $g = F(i,j,\cdot)$ we have in particular that, as $t \to \infty$,

(5.19) $$R(i,j,t)/R(j,j,t) \to F(i,j).$$

The following is a more powerful result.

(5.20) PROPOSITION. Let X be irreducible recurrent and let π be a solution of (5.6). If $g(k,\cdot)$ is right continuous and non-decreasing, then

(5.21) $$\lim_{t\to\infty} \frac{1}{R(h,i,t)} \int_0^t R(j,k,ds)g(k,t-s) = \frac{1}{\pi(i)}\, \pi(k)g(k,\infty)$$

for any $h,i,j,k \in E$. In particular,

(5.22) $$\lim_{t\to\infty} R(j,k,t)/R(h,i,t) = \pi(k)/\pi(i).$$

In the infinite state space case the computational result (5.5) still holds; but (5.9) is not true in general since, then, we cannot pass the limit inside the summation and obtain (5.13) without some further conditions on the function g or the process (X,T). A reasonably complete solution is contained in [17].

Since the $R(i,j,\cdot)$ are renewal functions, their asymptotic behaviors follow from renewal theory and we only need to compute the constants involved. These constants are the first, second, etc. moments of the recurrence time distributions $F(j,j,\cdot)$ and the object is to compute them directly from Q. The following papers (and the references contained therein) are primarily devoted to this task: GUPTA [27], HUNTER [31],[32], and KEILSON [38]; see also [11, p. 165].

6. SEMI-MARKOV PROCESSES

Let (X,T) be a Markov renewal process with a finite state space E and semi-Markovian kernel Q, and consider the semi-Markov process Y defined by (1.13). Define

(6.1) $$P_t(i,j) = P_i\{Y_t = j\}, \qquad i,j \in E,\ t \in \mathbb{R}_+,$$

where, as before, we write P_i for the conditional probability given that $X_0 = Y_0 = i$. Throughout the following R is the Markov renewal kernel corresponding to Q and h is defined by

(6.2) $$h(j,t) = 1 - \sum_{k \in E} Q(j,k,t), \qquad j \in E,\ t \in \mathbb{R}_+.$$

(6.3) PROPOSITION. For all $i,j \in E$ and $t \in \mathbb{R}_+$ we have

$$P_t(i,j) = \int_0^t R(i,j,ds)h(j,t-s).$$

PROOF. Consider the time T_1 of the first jump. If $T_1 > t$ then $Y_t = Y_0$; if $T_1 = s \leq t$ then Y_t has the same distribution as Y_{t-s} with initial value X_1. This renewal argument shows that

(6.4) $$P_t(i,j) = I(i,j)h(i,t) + \sum_k \int_0^t Q(i,k,ds)P_{t-s}(k,j).$$

For fixed j, if we define $f(i,t) = P_t(i,j)$ and $g(i,t) = I(i,j)h(i,t)$, then (6.4) becomes $f = g + Q * f$. Solving this Markov renewal equation by using Proposition (4.2) we obtain $f = R * g$ which, in view of the definition of g here, is precisely the desired result. □

Let $U(i,j,t)$ be the expected time spent in j during $[0,t]$ by the process Y starting at i, that is, let

(6.5) $$U(i,j,t) = E_i\left[\int_0^t 1_j(Y_s)ds\right], \qquad i,j \in E,\ t \in \mathbb{R}_+$$

where $1_j(k) = 1$ or 0 according as $k = j$ or $k \neq j$. Passing the expectation inside the integral, noting that $E_i[1_j(Y_s)] = P_i\{Y_s = j\} = P_s(i,j)$, and using the preceding proposition we obtain the following.

(6.6) PROPOSITION. We have

$$U(i,j,t) = \int_0^t R(i,j,ds) \int_0^{t-s} h(j,u)du. \qquad \square$$

We will call U the <u>potential kernel</u> of Y. It can be used to compute quantities such as

$$(6.7) \qquad Uf(i,t) = E_i\left[\int_0^t f(Y_s,s)ds\right]$$

for functions f in $\mathbb{B}$ (see Section 4 for the description of $\mathbb{B}$). If we interpret f(j,s) as the rate (per unit time) of rewards being received at an instant s at which the semi-Markov process Y is in state $Y_s = j$, then Uf(i,t) becomes the expected value of the total reward received during [0,t] starting at state i. In particular, when $f(j,s) = e^{-\alpha s} g(j)$ for some $\alpha > 0$, we may interpret Uf as the expected value of the total discounted reward received during [0,t], discounted at the rate of α, where the rate of reward is g(j) in state j. Using Proposition (6.6) we compute Uf defined by (6.7) to be

$$(6.8) \qquad Uf(i,t) = \sum_j \int_0^t R(i,j,du) \int_0^{t-u} h(j,s)f(j,u+s)ds.$$

Next consider the total amount of time spent in j by the process Y. If the initial state is i, then the expected value of it is U(i,j) = $U(i,j,+\infty) = \lim_{t\to\infty} U(i,j,t)$. Note that the expression given by (6.6) for U(i,j,t) is of form $\phi * g$ with $\phi = R(i,j,\cdot)$ and $g(t) = \int_0^t h(j,u)du$. Thus, using Lemma (5.18), and noting that $g(\infty) = m(j)$ is the mean sojourn time in state j we obtain the following

(6.9) PROPOSITION. For any $i,j \in E$

$$\lim_{t\to\infty} U(i,j,t)/R(i,j,t) = m(j). \qquad \square$$

In particular, if j is transient and can be reached from i, then the preceding result implies that $U(i,j) = U(i,j,\infty) = R(i,j)m(j)$. On the other hand, if j is recurrent and can be reached from i, then $U(i,j) = +\infty$. Combining Proposition (6.9) with (5.22) we also obtain

$$(6.10) \qquad \lim_{t\to\infty} U(i,j,t)/U(h,k,t) = \pi(j)m(j)/\pi(k)m(k)$$

for all i,j,h,k in the same irreducible closed set; here $\pi(j)$ is the limiting probability that $X_n = j$ as $n \to \infty$. If we sum $U(i,j,t)$ over all j we obtain exactly t; hence, for k recurrent and reachable from h, (6.10) implies that

$$(6.11) \qquad \lim_{t\to\infty} U(h,k,t)/t = \pi(k)m(k)/\sum_j \pi(j)m(j).$$

This provides an intuitive explanation for the following.

(6.12) PROPOSITION. Suppose X is irreducible recurrent and $m(k) < \infty$. Then, for any $i \in E$,

$$(6.13) \qquad \nu(j) = \lim_{t\to\infty} P_i\{Y_t = j\} = \pi(j)m(j)/\sum_k \pi(k)m(k)$$

in the aperiodic case, and

$$(6.14) \qquad \lim_{n\to\infty} P_i\{Y_{x+n\lambda} = j\} = \lambda\pi(j) \sum_n h(j,x+n\lambda-\lambda_{ij})/\sum_k \pi(k)m(k)$$

in the periodic case if the period is λ; here λ_{ij} is the first atom of the distribution $F(i,j,\cdot)$ and the summation over n is for n with $x + n\lambda \geq \lambda_{ij}$. □

In the particular case where the sojourn distributions are all exponential, namely when (1.14) holds, Y becomes a Markov process. In that case all states are aperiodic and

$$m(j) = \int_0^\infty h(j,s)ds = \int_0^\infty e^{-\lambda(j)s}\, ds = 1/\lambda(j).$$

Putting this in (6.13) gives the well known result that the limiting distribution of an irreducible Markov process is given by the solution of

$$(6.15) \qquad \sum_j \nu(j)A(j,k) = 0, \qquad \sum_j \nu(j) = 1$$

where $A(j,k)$ is the derivative, at $t = 0$, of the transition function $P_t(j,k)$.

In the case of infinitely many states, it is possible to define semi-Markov processes of much greater complexity. All of the results above hold, even in the infinite E case, provided that there be no instantaneous states and that the definitions of $R(i,j,\cdot)$ and $\pi(j)$ be altered slightly so that $R(i,j,t)$ becomes the expected number of visits to j by Y in $[0,t]$ (rather than the defining sum (2.5)), and $\pi(j)/\pi(i)$ becomes the expected number of visits to j in between two visits to i. For such more general processes we refer to PYKE and SCHAUFELE [55],[56]. Results such as (6.8) generalize [44] and [11, p. 167].

7. SEMI-REGENERATIVE PROCESSES

These are processes which are, in general, non-Markovian and yet possess the strong Markov property at certain selected random times. Then, imbedded at such instants, one finds a Markov renewal process. Using that property, one ends up with Markov renewal equations whose behavior we have already studied.

Let $Z = (Z_t)_{t \in \mathbb{R}_+}$ be a stochastic process with a topological state space F, and suppose that the function $t \to Z_t(\omega)$ is right-continuous and has left-hand limits for almost all ω. A random variable $T:\Omega \to [0,\infty]$ is called a stopping time for Z provided that, for any $t \in \mathbb{R}_+$, the occurrence or non-occurrence of the event $\{T \leq t\}$ can be determined once the history $\mathcal{H}_t = \sigma(Z_u;\ u \leq t)$ of Z before t is known.* If T is a stopping time

*Somewhat better put, $\mathcal{H}_t$ is the σ-algebra generated by $\{Z_u,\ u \leq t\}$ which, further, includes all the negligible sets of $\mathcal{H}$.

for Z, then we denote by $\mathcal{H}_T$ the history of Z before T.*

The process Z is said to be semi-regenerative if there exists a Markov renewal process (X,T) with a finite state space E such that

a) for each $n \in \mathbb{N}$, T_n is a stopping time for Z;

b) for each $n \in \mathbb{N}$, X_n is determined by $\mathcal{H}_{T_n}$;

c) for each $n \in \mathbb{N}$, $m \geq 1$, $0 \leq t_1 < \cdots < t_m$ and bounded function f defined on F^m,

$$(7.1) \qquad E_i[f(Z_{T_n+t_1},\ldots,Z_{T_n+t_m})|\mathcal{H}_{T_n}] = E_j[f(Z_{t_1},\ldots,Z_{t_m})] \quad \text{on } \{X_n = j\}.$$

Here E_i and E_j refer to expectations given the initial state for the Markov chain X. Condition (a) is that an observer who has watched Z until time t can tell whether T_n is less than or equal t or not. Condition (b) is that the observer who has watched Z until T_n can tell what X_n is. Condition (c) is the most important: it considers a random variable $W_n = f(Z_{T_n+t_1},\ldots,Z_{T_n+t_m})$ which is a function of the values Z takes at the future instants $T_n + t_1,\ldots,T_n + t_m$. The left-side of (7.1) is the conditional expectation of W_n by an observer who has watched the process until T_n (and therefore knows T_n and X_n in addition to all the other information this gives him). To a second observer who takes T_n as the time origin, the variable W_n appears as $W_0 = f(Z_{t_1},\ldots,Z_{t_m})$ and therefore, if the present state is $X_n = j$, his estimate of W_0 is $E_j(W_0)$, which is the right-hand side of (7.1). The process is semi-regenerative if the two observers come up with the same answer; in other words, if the extra information that the first observer had concerning the past of Z is worthless as far as predicting the future.

All Markov processes are semi-regenerative; all semi-Markov processes

---

* $\mathcal{H}_T = \{A \in M\colon A \cap \{T \leq t\} \in \mathcal{H}_t$ for all $t\}$.

are semi-regenerative; the queue size process in an M/G/1 queue is neither a Markov process nor a semi-Markov process but is semi-regenerative; similarly, the queue size process in a G/M/1 queue is semi-regenerative. Following are the main results.

(7.2) PROPOSITION. Let Z be a semi-regenerative process with state space F and let (X,T) be the Markov renewal process imbedded in Z. Let Q be the semi-Markovian kernel for (X,T) and let R be the corresponding Markov renewal kernel. For any open set $A \subset F$ let

$$K_t(i,A) = P_i\{Z_t \in A,\ T_1 > t\}, \qquad P_t(i,A) = P_i\{Z_t \in A\} \tag{7.3}$$

for $t \in \mathbb{R}_+$, $i \in E$. Then,

$$P_t(i,A) = \sum_{j \in E} \int_0^t R(i,j,ds)K_{t-s}(j,A). \tag{7.4}$$

PROOF. It is clear that

$$P_t(i,A) = K_t(i,A) + P_i\{Z_t \in A,\ T_1 \le t\}.$$

Now, using the definition of semi-regeneration, the second term can be written as

$$E_i[I_{[0,t]}(T_1)P_i\{Z_t \in A|\mathcal{H}_{T_1}\}]$$
$$= E_i[I_{[0,t]}(T_1)P_{t-T_1}(X_1,A)]$$
$$= \sum_j \int_0^\infty Q(i,j,ds)\ I_{[0,t]}(s)P_{t-s}(j,A).$$

Hence,

$$P_t(i,A) = K_t(i,A) + \sum_j \int_0^t Q(i,j,ds)P_{t-s}(j,A). \tag{7.5}$$

For fixed A, this is a Markov renewal equation of form $f = g + Q * f$ with $f(i,t) = P_t(i,A)$ and $g(i,t) = K_t(i,A)$. Now the proposition follows from Proposition (4.2).

Concerning the limiting behavior we have the following. The notation is that of Theorem (5.9) and is immediate from the preceding proposition (7.2) and theorem (5.9) upon taking $g(i,t) = K_t(i,A)$ in (5.9).

(7.6) PROPOSITION. Under the hypotheses and the notations of (7.2), further assuming that (X,T) is irreducible recurrent aperiodic and $m(j) < \infty$ for all j, we have

$$\lim_{t\to\infty} P_t(i,A) = \sum_j \pi(j)n(j,A)/\sum_j \pi(j)m(j)$$

with

$$n(j,A) = \int_0^\infty K_t(j,A)dt. \tag{7.7}$$

We end this section with a number of applications of the preceding two results to processes such as the "backward recurrence time," "forward recurrence time," etc. in semi-Markov processes. Many other applications and examples may be found in the next section.

Let (X,T) be a Markov renewal process with a finite state space E, and let $Y = (Y_t)_{t \in \mathbb{R}_+}$ be the semi-Markov process associated with it. For each t and ω, there is an $n \in \mathbb{N}$ such that $T_n(\omega) \leq t < T_{n+1}(\omega)$ and then we put

$$Y_t(\omega) = X_n(\omega),\ V_t(\omega) = T_{n+1}(\omega) - t,\ U_t(\omega) = t - T_n(\omega). \tag{7.8}$$

Then, V_t is called the time until the next transition after t, and U_t is called the time since the last transition before t. A little reflection shows that any one of the four processes

$$\begin{gathered} Y = (Y_t), \quad (Y,V) = (Y_t,V_t)_{t \in \mathbb{R}_+}, \quad (Y,U) = (Y_t,U_t)_{t \in \mathbb{R}_+}, \\ (Y,U,V) = (Y_t,U_t,V_t)_{t \in \mathbb{R}_+} \end{gathered} \tag{7.9}$$

is semi-regenerative admitting (X,T) as an imbedded Markov renewal process. We had studied Y extensively in the last section. We now consider the process (Y,U,V) from which results about the remaining two are easy to obtain.

For the semi-regenerative process $Z = (Y,U,V)$ with state space $F = E \times \mathbb{R}_+ \times \mathbb{R}_+$, if we take $A = \{k\} \times (x,\infty) \times (y,\infty)$ for some $x,y \in \mathbb{R}_+$ and $k \in E$, the function defined by (7.3) becomes

$$(7.10)\qquad \begin{aligned} K_t(i,A) &= P_i\{T_1 > t;\ Y_t = k,\ U_t > x,\ V_t > y\} \\ &= I(i,k)I_{(x,\infty)}(t)h(i,t+y) \end{aligned}$$

where h is as defined by (6.2). Thus, using Proposition (7.2) we have

$$(7.11)\qquad P_i\{Y_t = k,\ U_t > x,\ V_t > y\} = \int_0^{t-x} R(i,k,ds)h(k,t+y-s).$$

When (X,T) is recurrent aperiodic with $m(j) < \infty$ for all j we may apply Proposition (7.6) to (Y,U,V) to obtain

$$(7.12)\qquad \lim_{t\to\infty} P_i\{Y_t = k,\ U_t > x,\ V_t > y\} = \frac{1}{\pi m}\,\pi(k) \int_{x+y}^{\infty} h(j,s)ds$$

where $\pi m = \Sigma\ \pi(j)m(j)$.

For a large number of other results on the process (Y,U,V) we refer to [56]. Some of the results above may also be found in SCHÄL [57] with some further results in [11]; see also KEILSON [37] and KEILSON and WISHART [40].

8. APPLICATIONS

The importance of Markov renewal theory lies in its large domain of applicability rather than the inner richness of Markov renewal processes. In this respect, the situation is the same as that with renewal theory: a renewal process is a trite object but renewal theory is the most important tool in elementary probability theory. It is from such a point of view that we have stressed the theory of Markov renewal equations and the manner in which they arise in studying semi-regenerative processes. Following is a rapid review of some of the more important applications. For other applications we refer to the bibliography [6].

(8.1) M/G/1 Queueing Systems. This is a single server queueing system subject to a Poisson process of arrivals and arbitrary service time distributions. Let T_n be the time of the n^{th} departure, X_n the queue size just after the n^{th} departure, and Z_t the queue size at time t. Then, (X,T) is a Markov renewal process, and Z is a semi-regenerative process having (X,T) imbedded in it. If the queue size is allowed to be infinite, then the state space E will be infinite. Otherwise, if only finite queues are allowed, E is finite. In either case Propositions (7.2) and (7.6) hold. Therefore, the time-dependent behavior of the queue size process Z reduces simply to the derivation of the probabilities Q(i,j,t) and $K_t(i,j)$, and the computation of R and π. For the results we refer to [11].

(8.2) M/G/1 Queues with Bulk Service or Queue Dependent Service Times. The main properties of the system described in (8.1) are preserved even when the number of customers served during a service time, instead of being one, is a random variable depending on the number of customers in the system just at the start of that service. Thus, defining T_n to be the time of departure for the n^{th} batch, and X_n to be the number of customers just after T_n, we again have a Markov renewal process (X,T). Again, the behavior of the queue size process Z may be obtained by direct applications of Propositions (7.2) and (7.6) once the semi-Markovian kernel Q and the probabilities $K_t(i,j)$ are computed. We refer to NEUTS [46],[48] and LAMBOTTE, TEGHEM, LORIS-TEGHEM [65] for further details.

Similar remarks hold in the case of service times depending on the queue size. With T_n and X_n defined as in (8.1) we still have a Markov renewal process and all the remarks made in (8.1) apply. For detailed treatments we refer to HARRIS [29],[30], SCHÄL [59], and SUZUKI [62].

(8.3) G/M/1 Queues. Entirely similar remarks hold for this system as for the M/G/1 systems. Some of the computations are carried out in [11]. See also FABENS [22] for earlier treatments of queues and inventories.

(8.4) Markov Renewal Branching Processes. This is a Markov renewal process (X,T) where X is a branching process. In other words, the size X_{n+1} of the $(n+1)^{th}$ generation depends not only on the size X_n of the n^{th} generation but also on the lifetime $T_{n+1} - T_n$ of the n^{th} generation. Such processes were studied and applied to a number of queueing systems by NEUTS [49],[50].

(8.5) Queues with Semi-Markovian Arrivals or Services. We concentrate on the system with semi-Markovian arrivals and exponential services. Let T_n be the time of the n^{th} arrival and let X'_n denote the "type" of the n^{th} arrival. It is assumed that (X',T) is a Markov renewal process and that the service times are exponential. Then, letting X''_n be the number of customers in the system just before the n^{th} arrival, we obtain a Markov renewal process ((X',X"),T) with state space $E = \{(i,j): i \in F,\ j \in \mathbb{N}\}$ where F is the finite set of "types." Then, the queue size process $Z = (Z_t)$ becomes a semi-regenerative process with ((X',X"),T) imbedded in it. These processes were studied by NEUTS [47] and ÇINLAR [8],[9] by rather poor techniques necessitating superfluous assumptions. A straightforward analysis using Propositions (7.2) and (7.6) would yield the same results quicker and in greater generality.

(8.6) Machine Repair Problem. Consider a machine with an arbitrary lifetime distribution. There are a number of spares available so that, when the machine in operation fails, it is replaced by one of the spares and the failed one is sent to be repaired. The repair times are exponential and a repaired machine is as good as new. Let Z_t be the number of machines working or waiting as spares at t; then the total number of machines minus Z_t is the number of them in the repair shop. Let T_n be the time of the n^{th} failure and let $X_n = Z_{T_n}$. Then, (X,T) is a Markov renewal process and Z is a semi-regenerative process with (X,T) imbedded in it. The first passage time distribution $F(i,0,\cdot)$ is the distribution of the time until the first breakdown of the system, namely, the first time a failure occurs with no available spares for replacement. The probability that the system is working at time T is $P_i\{Z_t \neq 0\} = 1 - \int_0^t R(i,0,ds)e^{-\lambda(t-s)}$ with λ as the "rate" of repairs.

(8.7) System Reliability. Consider a finite number of components in series (so that if any one fails the system fails also). When a failure occurs, the component which has failed is repaired and all the other components are re-adjusted. A repaired or readjusted component has a lifetime distribution $\phi(i,\cdot)$ depending only on the type i of the component itself, and the repair-time of the i^{th} component plus the readjustment time for all the others has the distribution $\psi(i,\cdot)$. Let X_n be the type of the component which has caused the n^{th} failure and let T_n be the time of that n^{th} failure. Then (X,T) is a Markov renewal process with

$$Q(i,j,t) = \int_0^t \psi(i,ds) \int_0^{t-s} \phi(j,du) \prod_{k \neq j} (1 - \phi(k,u)).$$

In addition to asymptotic values of the number of times j fails in [0,t] (whose expectation is R(i,j,t) if the initial failure was due to i), a quantity of interest is the system reliability at t, namely the probability that the system is working at time t. This is equal to

$$1 - \sum_j \int_0^t R(i,j,ds)(1 - \psi(j,t-s)).$$

Further quantities of interest such as the ratio of times the system is under repair to times the system is working can be obtained by easy integrations. We refer to [11] for a somewhat more extensive treatment in a special case.

(8.8) Clinical Trials. Markov renewal processes turn out to be useful as models for the behavior of diseases such as cancer and leukemia. There the Markov chain X models the successive phases of the disease and the duration $T_{n+1} - T_n$ becomes the length of time the phase X_n persists. Then, the first passage distributions $F(i,j,\cdot)$ yield useful information concerning the development of the disease. WEISS and ZELEN [69] have applied this model to actual clinical data of acute leukemia using a model with six states (death, initial relapse, first partial remission, second partial remission, first complete remission, and second complete remission) and obtained the distributions of the sojourn times directly. See also [70], [24].

(8.9) An Insurance Problem. A similar problem occurs in constructing a model to represent the evolution of the "degree of disability" after an accident in order to evaluate the expected payments by the insurance company. JANSSEN [34] has modeled this process by a semi-Markov process Y where Y_t is the degree of disability t units of time after the accident. Relevant quantities of interest are the limits as $t \to \infty$ of $P\{Y_t = j\}$.

(8.10) Counters of Type I. Particles arriving at a counter can be classified (with respect to their energy levels, velocities, or physical types) into a number of "types." Let T_n be the time of the n^{th} arrival and X_n the type of the particle arriving then. It is supposed that (X,T) is a Markov renewal process. An arriving particle which finds the counter free gets registered and locks the counter for some random time whose distribution depends on the type of the particle causing it. Arrivals during a locked period have no effects whatsoever. Supposing that the time origin is a time of registration, and the first locked period has length L, then the time of the next registration is $T_1' = \inf\{T_n : T_n > L\}$. Thus, if X_1' denotes the type of the next particle to get registered, we have

$$\begin{aligned} Q'(i,j,t) &= P\{X_1' = j,\ T_1' \le t \mid X_0 = i\} \\ &= E_i[P\{X_1' = j,\ T_1' \le t \mid X_0, L\}] \\ &= E_i\Big[\sum_k \int_0^L R(X_0,k,ds)[Q(k,j,t-s) - Q(k,j,L-s)]\Big] \\ &= \int_0^t \psi(i,du) \sum_k \int_0^u R(i,k,ds)[Q(k,j,t-s) - Q(k,j,u-s)] \end{aligned}$$

if the distribution of the locked period caused by an i-type particle is $\psi(i,\cdot)$. A little reflection shows that the "registration process" (X',T') (where X_n' is the type and T_n' is the time of the n^{th} registration) is also a Markov renewal process and Q' is the corresponding semi-Markovian kernel.

For counter problems of this type and the type to be discussed next we refer to [11, p. 178], BARLOW [2] and VANDEVIELE [68].

(8.11) Counters of Type II. Arrival process is again a Markov renewal process (X,T); an arriving particle which finds the counter free gets registered and locks it for some time; a particle which arrives to find the counter locked does not get registered, erases the influence of the past arrivals, and locks the counter for some random time. Again, let X_n' and T_n' be the type and time of the n^{th} registration and let $\psi(i,\cdot)$ be the distribution of a locked period caused by an i-type particle (if it does not get erased).

Now using a Markov renewal argument at the time T_1 of the first arrival (we assume $X_0 = X_0'$, $T_0 = T_0' = 0$) we can write

$$Q'(i,j,t) = \int_0^t Q(i,j,ds)\psi(i,s) + \sum_k \int_0^t Q(i,k,ds)(1 - \psi(i,s))Q'(k,j,t-s).$$

Letting

$$f(i,t) = Q'(i,j,t), \quad g(i,t) = \int_0^t \psi(i,s)Q(i,j,ds),$$

and $$\hat{Q}(i,k,ds) = Q(i,k,ds)(1 - \psi(i,s))$$

we can put the above equation in the form

$$f = g + \hat{Q} * f.$$

This is a Markov renewal equation again, except that the semi-Markovian kernel $\hat{Q}$ here is defective:

$$\sum_k \hat{Q}(i,k,\infty) < 1.$$

It turns out that Proposition (4.2) remains true provided that R there is replaced by $\hat{R} = \Sigma\, \hat{Q}^n$ where the $\hat{Q}^n$ are defined by (2.2),(2.3) by replacing the Q's there by $\hat{Q}$. For the results we refer to [11, p. 183].

It was this applied problem as well as the one to be discussed next which led to the more general theory of Markov renewal processes allowing "deaths" after finitely many transitions in [11, Definition (2.5)].

(8.12) Pedestrian Delay Problem. Let T_n be the time the n^{th} vehicle crosses a fixed point on a highway and let X_n be its "type" (this "type" may stand for velocity or shape or whether the vehicle is a car or a truck, etc.). It seems plausible that this process (X,T) be a Markov renewal process; cf. JEWELL [36].

At time $0 = T_0$ a pedestrian arrives to cross the highway; considering various factors involved, he "accepts" a gap of size x with probability $\psi(j,x)$ if the oncoming vehicle is of type j. We are interested in finding the distribution of the delay to the pedestrian. Letting L denote this delay we have

$$P_i\{L \leq t\} = \sum_j \int_0^\infty Q(i,j,ds)\psi(j,s) + \sum_j \int_0^t Q(i,j,ds)(1-\psi(j,s))P_j\{L \leq t-s\}.$$

Again, letting $f(i,t) = P_i\{L \leq t\}$ and $\hat{Q}(i,j,ds)(1 - \psi(j,s))$ we can put this equation in the form

$$f = g + \hat{Q} * f$$

with a suitable definition for g. This is a "generalized" Markov renewal equation and (4.2) holds with R there replaced by $\hat{R} = \sum_n \hat{Q}^n$. This problem is based on MARADUDIN and WEISS [43] and may be found in [11, p. 177].

For other interesting applications to traffic theory we refer to BULLEN [3] and WIENER et al. [71].

References to a number of other applications which are of somewhat more routine nature may be found in the bibliography [6].

9. RESTRICTION TO A SUBSET

Let (X,T) be a Markov renewal process with state space E, and let D be a subset of E. Suppose the initial state i is in D, and consider the states visited by X at its subsequent entrances to D along with the times of those entrances to D. In other words, letting

(9.1) $$N_0 = 0,\ N_1 = \inf\{n > 0: X_n \in D\},\ N_2 = \inf\{n > N_1: X_n \in D\}, \dots$$

we are interested in the process $(\hat{X},\hat{T})$ where

(9.2) $$\hat{X}_n = X_{N_n},\quad \hat{T}_n = T_{N_n}, \qquad n \in \mathbb{N}.$$

The process $(\hat{X},\hat{T})$, then, is called the restriction of (X,T) to D except that we have a minor problem to settle. It could happen that, for some $\omega \in \Omega$, the process X enters D only finitely many times; then, if the total number of entrances to D after 0 is n, we have $N_0(\omega) = 0$, $N_1(\omega) < \infty, \dots,$ $N_n(\omega) < \infty$ but $N_{n+1}(\omega) = N_{n+2}(\omega) = \cdots = +\infty$. In such a case, (9.2) does not make sense since we have not defined X_∞ or T_∞. To settle the matter, we adjoin a distinguished state Δ to the state space E, and define $X_\infty(\omega) = \Delta$, $T_\infty(\omega) = +\infty$ for all ω. Thus, the process $(\hat{X},\hat{T})$ has the state space $D_\Delta = D \cup \{\Delta\}$ for $\hat{X}$ and $[0,\infty]$ for $\hat{T}$.

We notice that the process $(\hat{X},\hat{T})$ defined above is a Markov renewal process with state space $D_\Delta = D \cup \{\Delta\}$. In particular, if D is a set of recurrent states, then $N_1, N_2, \ldots$ are all finite and $\hat{X}_n$ is in D for all $n < \infty$; that is, then $\hat{X}$ has the state space D. Otherwise, if D contains transient states from which X may leave and never return to D again, then $\hat{X}$ has the state space D_Δ where Δ is an absorbing state and $\hat{X}_n = \Delta$ means that X has left D after at most n-1 visits to D never to return.

Computations regarding $(\hat{X},\hat{T})$ are easy. We observe that the number of times the semi-Markov process Y visits a state $j \in D$ during $(0,t]$ is the same whether it be computed with E as the state space or with D; that is,

$$\sum_n I_{\{X_n = j,\ T_n \le t\}} = \sum_m I_{\{\hat{X}_m = j,\ \hat{T}_m \le t\}} \tag{9.3}$$

for any $j \in D$. Thus, if $\hat{R}$ is the Markov renewal kernel corresponding to $(\hat{X},\hat{T})$ and R is the one corresponding to (X,T), we have

$$\hat{R}(i,j,t) = R(i,j,t), \qquad i,j \in D,\ t \ge 0. \tag{9.4}$$

Once $\hat{R}$ is known, the semi-Markov kernel $\hat{Q}$ and the first passage distributions $\hat{F}$ for the process $(\hat{X},\hat{T})$ can all be computed by using the relations in Section 2. Note that $\hat{Q}(i,j,t) \neq Q(i,j,t)$ even for $i,j \in D$; in fact, we have

$$\hat{Q}(i,j,t) = Q(i,j,t) + \sum_{k,k' \notin D} \int_0^t Q(i,k,du) \int_0^{t-u} R(k,k',ds)Q(k',j,t-u-s). \tag{9.5}$$

In many applications, this idea of restricting (X,T) to a subset D of E leads to considerable conceptual simplifications. For example, if one is interested in the relative behavior of one state j with respect to another state k, one could restrict one's attention to the Markov renewal process $(\hat{X},\hat{T})$ with $D = \{j,k\}$. Similarly, to obtain a ratio limit theorem such as (6.10), one needs to consider only a four state Markov renewal process $(\hat{X},\hat{D})$ with $D = \{h,i,j,k\}$.

10. HITTING DISTRIBUTIONS AND THE MAXIMUM OF A SEMI-MARKOV PROCESS

Let (X,T) be a Markov renewal process with state space $E = \{0,1,2,...\}$ and semi-Markovian kernel Q, let Y be the associated semi-Markov process, and define

(10.1) $$Z_t = \sup\{Y_s : s \leq t\},$$

in other words, Z_t is the maximum level the process Y has ever attained during $[0,t]$. We are interested in the distribution

(10.2) $$\hat{P}_t(i,j) = P_i\{Z_t = j\}.$$

We adjoin a distinguished state Δ to the state space and define $X_\infty = \Delta$, $T_\infty = +\infty$. For a subset A of E let us define N to be the smallest integer $n \geq 1$ for which $X_n \in A$ if there is any such n; otherwise we set $N = \infty$. Then,

(10.3) $$F_A(i,j,t) = P_i\{X_N = j,\ T_N \leq t\}$$

is the probability that, starting at i, the process Y enters A for the first time at the state $j \in A$ and this happens at or before t.

Assuming we have F_A computed we now finish the computation of $\hat{P}_t(i,\cdot)$. For $i \in E$ let $A_i = \{i+1,i+2,...\}$. Then,

(10.4) $$\hat{Q}(i,j,t) = F_{A_i}(i,j,t), \qquad j \in A_i,\ t \in \mathbb{R}_+$$

is the probability that the first transition of Z is to state j and this happens before t. A little reflection will show that Z itself is a semi-Markov process (all of whose jumps are upward), and that $\hat{Q}$ is the corresponding semi-Markovian kernel. Hence, $\hat{P}$ can be computed as in Section 6 with $\hat{Q}$ replacing Q.

Returning to the computation of (10.3), we use the standard "Markov renewal" argument: either the first transition of Y is to j and this happens before t, or else the first transition is to some state $k \notin A$ at time $s < t$ and from k the process Y hits A at j during (s,t]; in other words, for $j \in A$,

$$F_A(i,j,t) = Q(i,j,t) + \sum_k \int_0^t Q_A(i,k,ds)F_A(k,j,t-s) \tag{10.5}$$

where

$$Q_A(i,k,t) = \begin{cases} 0 & \text{if } k \in A, \\ Q(i,k,t) & \text{if } k \notin A. \end{cases} \tag{10.6}$$

For fixed $j \in A$, the equation (10.5) can be put in the form $f = g + Q_A * f$ by defining $f(i,t) = F_A(i,j,t)$ and $g(i,t) = Q(i,j,t)$. This is a Markov renewal equation with a defective semi-Markov kernel Q_A. As we have remarked earlier in (8.11) and (8.12), Proposition (4.2) remains true with R there replaced by $R_A = \Sigma\, Q_A^n$ where Q_A^n are computed through (2.2) and (2.3) with Q therein replaced by Q_A.

Maximum of a semi-Markov process is of interest in level crossing problems and queueing theory. Our results above cover those of [61] who arrived at these results by computational techniques. Noticing the semi-Markovian character of Z reduces the task immensely. Moreover, the hitting distributions $F_A(i,j,\cdot)$ are of interest in themselves.

11. GEOMETRIC ERGODICITY

Consider a Markov renewal process (X,T) with Markov renewal kernel R and first passage distributions F. The results we are now concerned with are sharper versions of the limit results (5.1) and (5.2). We will merely sketch the main results in this area; for detailed treatments we refer to TEUGELS [66],[67], CHEONG [4],[5], and SCHÄL [58].

For fixed $\lambda \geq 0$, let $F^\lambda(i,j,dt) = e^{\lambda t} F(i,j,dt)$; this defines a mass function $F^\lambda(i,j,\cdot)$. We denote by $F^\lambda(i,j) = F^\lambda(i,j,+\infty)$ its total mass. Note that $F^0(i,j) = F(i,j) = F(i,j,+\infty)$ and $F^{-\lambda}(i,j) = F_\lambda(i,j)$ is the Laplace transform of $F(i,j,\cdot)$. In analogy with the definitions of transience, etc. we now define a state j to be λ-recurrent if $F^\lambda(j,j) = 1$ and λ-transient if $F^\lambda(j,j) < 1$. A λ-recurrent state j is said to be null if $\int tF^\lambda(j,j,dt)$ diverges and non-null otherwise. For $\lambda = 0$ these concepts coincide with the ordinary definitions of recurrence and transience. Just as in that case, if X is irreducible, then either all states are λ-recurrent or else they are all λ-transient. If they are all λ-recurrent, either they are all null or all non-null.

Suppose X is irreducible and let $\lambda \geq 0$ be fixed. If the states are all λ-transient, then

(11.1) $$R^\lambda(j,j) = \int_0^\infty e^{\lambda t} R(i,j,dt) = F^\lambda(i,j)R^\lambda(j,j)$$

for $i \neq j$, and for $i = j$,

(11.2) $$R^\lambda(j,j) = 1/[1 - F^\lambda(j,j)] < \infty.$$

On the other hand, if the states are λ-recurrent, then

(11.3) $$R^\lambda(i,j) = +\infty$$

for all i,j. In this case, if the states are null, then

(11.4) $$\lim_{t\to\infty} \frac{1}{t} \int_0^t e^{\lambda t} R(i,j,dt) = 0,$$

and if they are non-null,

(11.5) $$\lim_{t\to\infty} \frac{1}{t} \int_0^t e^{\lambda t} R(i,j,dt) = F^\lambda(i,j)\eta^\lambda(j)$$

where

(11.6) $$\eta^\lambda(j) = 1/\int_0^\infty te^{\lambda t} F(j,j,dt).$$

These results have further implications concerning the limiting behavior of the semi-Markov process Y. Let $P_t(i,j)$ be the transition probability as defined by (6.1). Then, assuming irreducibility, we have

$$\lim_{t\to\infty} e^{\lambda t} P_t(i,j) = 0 \tag{11.7}$$

if the states are λ-transient or λ-recurrent and null. Otherwise, if the states are λ-recurrent and non-null, then

$$\lim_{t\to\infty} e^{\lambda t} P_t(i,j) = F^\lambda(i,j)\eta^\lambda(j)m^\lambda(j) \tag{11.8}$$

where

$$m^\lambda(j) = \int_0^\infty e^{\lambda t}(1 - \sum_k Q(j,k,t))dt \tag{11.9}$$

and $\eta^\lambda(j)$ is as defined by (11.6).

12. STRONG LAWS AND THE CENTRAL LIMIT THEOREM FOR CUMULATIVE PROCESSES

Let (X,T) be a Markov renewal process with X irreducible. Let $W_1, W_2, \ldots$ be random variables taking values in $\mathbb{R} = (-\infty, +\infty)$ and such that $W_1, W_2, \ldots$ are conditionally independent given (X,T) and that the distribution of W_n given (X,T) depends only on $X_{n-1}, X_n, T_n - T_{n-1}$. That is,

$$\begin{aligned} &P\{W_n \leq y \mid X_m, T_m, W_k;\ m \in \mathbb{N},\ k \in \mathbb{N},\ k \neq n\} \\ &= P\{W_n \leq y \mid X_{n-1}, X_n, T_n - T_{n-1}\} \\ &= K(i,j,t;y) \end{aligned} \tag{12.1}$$

on $\{X_{n-1} = i,\ X_n = j,\ T_n - T_{n-1} = t\}$ for some kernel K. Define the partial sums

$$S_0 = 0;\ S_{n+1} = S_n + W_{n+1}, \qquad n \in \mathbb{N} \tag{12.2}$$

and let $N_t = \sum_n I_{(0,t]}(T_n)$ be the number of transitions in (0,t]. We are interested in the asymptotic behavior of S_{N_t} as $t \to \infty$.

Let the initial state i be fixed and consider the successive values $\hat{S}_0, \hat{S}_1, \ldots$ of the sequence $\{S_n : X_n = i\}$. The important point on which all limit theorems are based is that this sequence $(\hat{S}_n)$ is a random walk in $\mathbb{R}$, in other words, $\hat{S}_1 - \hat{S}_0$, $\hat{S}_2 - \hat{S}_1, \ldots$ are independent and identically distributed. Let M_t be the number of entrances to i in (0,t]; this is the counting process associated with the renewal process formed by the times T_n for which $X_n = i$. The second important point is that we can write

(12.3) $$S_{N_t} = \hat{S}_{M_t} + R_t$$

where the remainder term R_t is, roughly speaking, stationary in time. Therefore, the limiting behavior of S_{N_t} is the same as that of $\hat{S}_{M_t}$ and the latter is fairly easy to analyze.

Here are the arguments leading up to a strong law of large numbers (we leave out the adornment "almost surely" in the statements below). Let S_n^+ be computed from the sequence $|W_1|, |W_2|, \ldots$ just as $\hat{S}_n$ is computed from $W_1, W_2, \ldots$; then, (S_n^+) is an ordinary renewal process and in particular

(12.4) $$S_1^+ = |W_1| + \cdots + |W_n| \quad \text{on } \{X_1 \neq i, \ldots, X_{n-1} \neq i,\ X_n = i\}.$$

We assume (see (12.11) below for the computation) that

(12.5) $$a^+(i) = E_i[S_1^+] = E_i[S_{n+1}^+ - S_n^+] < \infty.$$

(See (12.11) below for the computation, either $a^+(i) < \infty$ for all i or for none; so the selected initial state plays very little role.) Then, clearly,

(12.6) $$(S_{n+1}^+ - S_n^+)/n \to 0, \qquad n \to \infty.$$

Since M_t is the number of renewals in (0,t] in a renewal process, as $t \to \infty$, $M_t \to \infty$ and $M_t/t \to \eta(i)$. Thus, (12.6) implies that

(12.7) $$(S_{M_t+1}^+ - S_{M_t}^+)/t \to 0, \qquad t \to \infty.$$

Now, the remainder term R_t in (12.3) obviously satisfies $|R_t| \leq S_{M_t+1}^+ - S_{M_t}^+$ so that, by (12.7), (12.8) $$R_t/t \to 0, \qquad t \to \infty.$$

It follows from (12.3) and (12.8) that

(12.9)
$$\begin{aligned}\lim_t S_{N_t}/t &= \lim_t \hat{S}_{M_t}/t \\ &= \lim_t(\hat{S}_{M_t}/M_t)\lim_t(M_t/t) \\ &= \lim(\hat{S}_n/n)\lim_t(M_t/t) = a(i)\eta(i)\end{aligned}$$

where

(12.10) $$a(i) = E_i[\hat{S}_1]$$

and $\eta(i)$ is the quantity computed in (5.8).

The dependence of the limit in (12.9) on the state i is in appearance only. Assumption (12.5) implies that a(i) is well defined, and a computation yields

(12.11) $$a(i) = \sum_j \pi(j)n(j)/\pi(i)$$

with

(12.12) $$n(j) = \sum_k \int_0^\infty Q(j,k,dt) \int_{-\infty}^\infty K(j,k,t,dy)y,$$

where K is as in (12.1). (Incidentally, $a^+(i)$ of (12.5) is obtained from (12.11) by replacing n there by n^+ where n^+ is obtained from (12.12) by replacing y there by its absolute value $|y|$.) It follows from (12.11) and (5.8) that the limit in (12.9) is

(12.13) $$a = a(i)\eta(i) = \sum_j \pi(j)n(j)/\sum_j \pi(j)m(j)$$

independent of i.

The strong law (12.9) is a generalized version of the one given in [55]. Next, we will discuss the central limit theorem for the process S_{N_t}. This also will be a generalized version of the one in [55] where W_n are assumed to be of the form $W_n = k(X_{n-1},X_n,T_{n+1} - T_n)$ for some function k. Further, [39] contains a central limit theorem for the sequence S_n. Here are the main steps of the central limit theorem.

We now need to assume that the variance of S_1^+ is finite. Then, in (12.3),

(12.14) $$R_t/\sqrt{t} \to 0, \qquad t \to \infty.$$

Thus we need concern ourselves with only the term $\hat{S}_{M_t}$. Now, $\hat{S}$ is a random walk and M is a renewal counting process. However, $\hat{S}$ and M are dependent and a certain amount of care is necessary. The applicable result is the central limit theorem due to SMITH [60] in whose terminology $(\hat{S}_{M_t})$ is a cumulative process defined on (M_t). We have,

(12.15) $$\lim_{t\to\infty} P\left\{\frac{1}{b\sqrt{t}}\left(\hat{S}_{M_t} - at\right) \le x\right\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2}\, dy$$

where a is the constant computed in (12.13) and

(12.16) $$b^2 = \mathrm{Var}(\hat{S}_1 - a\hat{T}_1)$$

with $\hat{T}_1$ denoting the first T_n for which $X_n = i$, $n \ge 1$ provided that the variance of $\hat{T}_1$ is finite. In view of (12.14),(12.15) yields the central limit theorem we were looking for: under the assumptions mentioned,

(12.17) $$\lim_{t\to\infty} P\left\{\frac{1}{b\sqrt{t}}\left(S_{N_t} - at\right) \le x\right\} = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-y^2/2}\, dy.$$

13. EXTENSIONS

We have been considering Markov renewal processes $(X_n, T_n)_{n \in \mathbb{N}}$ with finite state space E (except for some comments regarding the case of countably infinite E). The most obvious extension concerns E. When E is countably infinite the relevant results are in [11] and the references contained there.

In the case where the state space E is arbitrary, one can still obtain the same results as in the countable case but with some adjustments. Such a theory was introduced in ÇINLAR [12]; one of the open problems left in [12] was solved by JACOD [33] in a comprehensive study. A theorem of [12], due to an overanxious estimation, is indeed faulty; this was discovered and corrected by KESTEN (his paper is to appear shortly). This theory finds applications in continuous storage theory, (see [13],[18], [19]) and superpositions, etc. in [7].

The second extension concerns the space within which T_n take values. Instead of requiring T_n to be increasing and non-negative, one can keep the definition (1.1) with T_n taking values in $\mathbb{R}$ or $\mathbb{R}^m$ without further restriction. This was done by ANDERSON [1], JANNSEN [35], and JACOD [33]. The resulting theory is the generalization of the theory of random walks (rather than renewal theory).

Finally, an extension can be made by replacing the parameter set $\mathbb{N}$ of $(X_n,T_n)_{n \in \mathbb{N}}$ by the parameter set $\mathbb{R}_+$ to obtain a genuine continuous-time process $(X_t,T_t)_{t \in \mathbb{R}_+}$. Here, $(X_t)_{t \in \mathbb{R}_+}$ is a Markov process and $(T_t)_{t \in \mathbb{R}_+}$ has independent increments given (X_t). The defining property now becomes

$$P\{X_{t+s} \in A;\ T_{t+s} - T_t \in B \mid X_u, T_u;\ u \le t\} = Q_s(X_t,A,B)$$

for all $t,s \in \mathbb{R}_+$, where Q is a "semi-Markov kernel" satisfying

$$Q_{t+s}(x,A,B) = \iint Q_t(x,dy,du)Q_s(y,A,B-u).$$

The resulting theory is a generalization of the theory of processes with stationary independent increments, and of the theory of additive functionals of Markov processes. These processes were introduced by ÇINLAR [14],[15] and EZHOV and SKOROKHOD [20],[21]. In addition to generalizing the modern potential theory of Markov processes, this finds applications in processes in random environments [16], signal detection problems in the presence of noise, etc. In the case X takes values in a finite set, we also have some results in NEVEU [51], and central limit theorems in FUKUSHIMA and HITSUDA [26] and PINSKY [52].

REFERENCES

[1] L. B. ANDERSON. Markov renewal theory on the whole line. Unpublished doctoral dissertation, Northwestern University, 1970.

[2] R. E. BARLOW. Applications of semi-Markov processes to counter problems. Studies in Applied Probability and Management Sciences, eds. Arrow, Karlin, Scarf. Stanford University Press, Stanford, 1962.

[3] A. G. R. BULLEN. A Markov renewal model of multi-lane traffic.

[4] C. K. CHEONG. Geometric convergence of semi-Markov transition probabilities. Z. Wahrscheinlichkeitstheorie verw. Geb. **7** (1967), 122-130.

[5] C. K. CHEONG. Ergodic and ratio limit theorems for α-recurrent semi-Markov processes. Z. Wahrscheinlichkeitstheorie verw. Geb. **9** (1968), 270-286.

[6] C. K. CHEONG, J. H. A. DE SMIT and J. L. TEUGELS, Notes on semi-Markov processes, part II: bibliography. Discussion paper No. 7207, CORE, Université Catholique de Louvain.

[7] W. P. CHERRY. Superposition of two independent Markov renewal processes. Tech. Report 72-10, Dept. Industr. Engg., University of Michigan, 1972.

[8] E. ÇINLAR. Time dependence of queues with semi-Markovian arrivals. J. Appl. Prob. **4** (1967), 356-364.

[9] E. ÇINLAR. Queues with semi-Markovian arrivals. J. Appl. Prob. **4** (1967), 365-379.

[10] E. ÇINLAR. Some joint distributions for Markov renewal processes. Austral. J. Statist. **10** (1968), 8-20.

[11] E. ÇINLAR. Markov renewal theory. Adv. Appl. Prob. **1** (1969), 123-187.

[12] E. ÇINLAR. On semi-Markov processes on arbitrary spaces. Proc. Cambridge Phil. Soc. **66** (1969), 381-392.

[13] E. ÇINLAR. On dams with continuous semi-Markovian inputs. J. Math. Anal. Appl. 35 (1971), 434-448.

[14] E. ÇINLAR. Markov additive processes, I. Z. Wahrscheinlichkeitstheorie verw. Geb. 24 (1972), 85-93.

[15] E. ÇINLAR. Markov additive processes, II. Z. Wahrscheinlichkeitstheorie verw. Geb. 24 (1972), 94-121.

[16] E. ÇINLAR. Theory of continuous storage with Markov additive inputs and a general release rule. J. Math. Anal. Appl. 43 (1973),

[17] E. ÇINLAR. Periodicity in Markov renewal theory. J. Appl. Prob. (to appear).

[18] E. ÇINLAR and M. PINSKY. A stochastic integral in storage theory. Z. Wahrscheinlichkeitstheorie verw. Geb. 17 (1971), 227-240.

[19] E. ÇINLAR and M. PINSKY. On dams with additive inputs and a general release rule. J. Appl. Prob. 9 (1972), 422-429.

[20] I. I. EZHOV and A. V. SKOROHOD. Markov processes with homogeneous second component, I. Theor. Prob. Appl. 14 (1969), 3-14.

[21] I. I. EZHOV and A. V. SKOROHOD. Markov processes with homogeneous second component, II. Theor. Prob. Appl. 14 (1969), 679-692.

[22] A. J. FABENS. The solution of queueing and inventory models by semi-Markov processes. J. Roy. Statist. Soc. Ser. B 23 (1961), 113-127. (Correction, same journal, 25 (1962), 455-456.)

[23] A. J. FABENS and S. KARLIN. Generalized renewal functions and stationary inventory models. J. Math. Anal. Appl. 5 (1962), 461-487.

[24] M. FEINLEIB and M. ZELEN. On the theory of screening for chronic diseases. Biometrika 56 (1969), 601-604.

[25] W. FELLER. An Introduction to Probability Theory and Its Applications, 2, Wiley, New York, 1966.

[26] M. FUKUSHIMA and M. HITSUDA. On a class of Markov processes taking values on lines and the central limit theorem. Nagoya Math. J. 30 (1967), 47-56.

[27] Y. P. GUPTA. Joint distributions and asymptotic values of the moments

for Markov renewal processes in space. J. Indian Statist. Assoc. 10 (1972),

[28] Y. P. GUPTA and A. M. KSHIRSAGAR. Mean and variance of the number of renewals in certain Markov renewal processes. J. Indian Statist. Assoc. 6 (1968),

[29] C. M. HARRIS. Queues with state-dependent stochastic service rates. Oper. Res. 15 (1967), 117-130.

[30] C. M. HARRIS. Queues with stochastic service rates. Naval Res. Logist. Quart. 14 (1967), 219-230.

[31] J. J. HUNTER. On the moments of Markov renewal processes. Adv. Appl. Prob. 1 (1969), 188-210.

[32] J. J. HUNTER. On the renewal density matrix of a semi-Markov process. Sankhyā, A. 31 (1969), 281-308.

[33] J. JACOD. Théorème de renouvellement et classification pour les chaînes semi-Markoviennes. Ann. Inst. H. Poincaré, Sect. B, 7 (1971), 83-129.

[34] J. JANSSEN. Application des processus semi-markoviens à un problème d'invalidité. Bull. A. R. A. B. 63 (1966), 35-52.

[35] J. JANSSEN. Sur une généralisation du concept de promenade aléatoire sur la droite réelle. Ann. Inst. H. Poincaré, Sect. B, 6 (1970), 249-269.

[36] W. S. JEWELL. Markov renewal models in traffic flow theory. In Vehicular Traffic Science; Proc. Third Int. Symp. Theory of Traffic Flow; eds. Edie, Herman, Rothery; Elsevier, New York, 1965.

[37] J. KEILSON. A limit theorem for passage times in ergodic regenerative processes. Ann. Math. Statist. 37 (1966), 866-870.

[38] J. KEILSON. On the matrix renewal function for Markov renewal processes. Ann. Math. Statist. 40 (1969), 1901-1907.

[39] J. KEILSON and D. M. G. WISHART. A central limit theorem for processes defined on a finite Markov chain. Proc. Cambridge Phil. Soc. 60 (1964), 547-567.

[40] J. KEILSON and D. M. G. WISHART. Boundary problems for processes defined on a finite Markov chain. Proc. Cambridge Phil. Soc. 61 (1965), 173-190.

[41] A. M. KSHIRSAGAR and R. WYSOCKI. Some distribution and moment formulae for the Markov renewal process. Proc. Cambridge Phil. Soc. 68 (1970), 159-166.

[42] P. LÉVY. Processus semi-Markoviens. Proc. Int. Congr. Math. (Amsterdam) 3 (1954), 416-426.

[43] A. A. MARADUDIN and G. WEISS. Some problems in traffic delay. Oper. Res. 10 (1967), 74-104.

[44] R. A. McLEAN and M. F. NEUTS. The integral of a step function defined on a semi-Markov process. SIAM J. Appl. Math. 15 (1967), 726-738.

[45] M. F. NEUTS. Generating functions for Markov renewal processes. Ann. Math. Statist. 35 (1964), 431-434.

[46] M. F. NEUTS. The busy period of a queue with batch service. Oper. Res. 13 (1965), 815-819.

[47] M. F. NEUTS. The single server queue with Poisson input and semi-Markov service times. J. Appl. Prob. 3 (1966), 202-230.

[48] M. F. NEUTS. Semi-Markov analysis of a bulk queue. Bull. Soc. Math. Belg. 18 (1966), 28-42.

[49] M. F. NEUTS. The queue with Poisson input and general service times treated as a branching process. Duke Math. J. 36 (1969), 215-231.

[50] M. F. NEUTS. Two servers in series studied in terms of a Markov renewal branching process. Adv. Appl. Prob. 2 (1970), 110-149.

[51] J. NEVEU. Une généralisation des processus à accroisements positifs indépendants. Abhandl. Math. Sem. Univ. Hamburg 25 (1961), 36-61.

[52] M. PINSKY. Differential equations with a small parameter and the central limit theorem for functions defined on a finite Markov chain. Z. Wahrscheinlichkeitstheorie verw. Geb. 9 (1968), 101-111.

[53] R. PYKE. Markov renewal processes: definitions and preliminary properties. Ann. Math. Statist. 32 (1961), 1231-1242.

[54] R. PYKE. Markov renewal processes with finitely many states. Ann. Math. Statist. 32 (1961), 1243-1259.

[55] R. PYKE and R. SCHAUFELE. Limit theorems for Markov renewal processes. Ann. Math. Statist. 35 (1964), 1746-1764.