REGENERATIVE ASPECTS OF THE
STEADY-STATE SIMULATION PROBLEM FOR MARKOV CHAINS

by

Peter W. Glynn

TECHNICAL REPORT NO. 61

July 1982

Prepared under Contract NO0014-76-C-0578 (NR 042-343)*

for the

Office of Naval Research

Approved for public release: distribution unlimited.

Reproduction in Whole or in Part is Permitted for any
Purpose of the United States Government

DEPARTMENT OF OPERATIONS RESEARCH
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

*This research was also partially supported under
National Science Foundation Grant MCS79-09139.
REGENERATIVE ASPECTS OF THE
STEADY-STATE SIMULATION PROBLEM FOR MARKOV CHAINS

Peter W. Glynn

1. Introduction

Let \( \{X_n: n \geq 0\} \) be a Markov chain taking values in a countable state space \( E = \{0, 1, 2, \ldots\} \). If \( \{X_n\} \) is irreducible and positive recurrent, then \( \{X_n\} \) possesses an invariant probability \( \pi \) such that

\[
\frac{1}{n} \sum_{j=0}^{n} f(X_j) + r(f) = \int f(y) \pi(dy)
\]

P1 a.s., for all \( \pi \)-integrable \( f \), and each \( i \in E \). Given that \( \{X_n\} \) represents the output of the simulation of a stochastic system, the simulator is often interested in obtaining point estimates and confidence intervals for the parameter \( r(f) \). The problem of constructing such estimates and intervals is known, in the simulation literature, as the steady-state simulation problem for Markov chains.

One approach to dealing with this problem is to exploit the fact that \( \{X_n: n \geq 0\} \) is a regenerative process; the process "regenerates" itself at return times to a fixed state \( i \). When the sample path is "blocked" according to consecutive regeneration times, the resulting regenerative intervals are independent and identically distributed (i.i.d.). CRANE and IGLEHART (1974) and FISHMAN (1973).
proposed techniques for the steady-state simulation problem that exploit the i.i.d. structure of the above regenerative intervals to produce a methodology that bears close resemblance to the classical statistical procedure for i.i.d. sequences. This regenerative method for analyzing the steady-state of countable state Markov chains has attracted a considerable amount of interest since its introduction; see FISHMAN (1978), RUBINSTEIN (1981), and LAW and KELTON (1982), for example.

It has generally been assumed, however, that the regenerative method has only limited applicability, due to the belief that many stochastic systems that arise in simulation are not regenerative. In this chapter, we shall show that in fact, virtually any well-behaved simulation has a regenerative-type structure.

As has been mentioned by WHITT (1980), the class of generalized semi-Markov processes (GSMP's) appears to be an attractive mathematical model for the general discrete-event simulation. These processes can be regarded as Markov chains taking values in a complete, separable metric space (see [24] for details on how the state space of GSMP can be topologized to be such a space). After dealing with some Markov chain preliminaries in Section 2, we shall introduce the concept, in Section 3, of a "well-posed" steady-state simulation problem. We shall prove that if the simulation problem is well-posed for a Markov chain, then the Markov chain is Harris recurrent with an invariant probability; the converse also holds. In Section 4, we shall show that any Harris chain can be embedded in a weakly regenerative environment. A weakly regenerative process is a generalization
of the notion of regenerative process that allows the process blocks
to be w-dependent, as opposed to independent as in the regenerative
case.

In [11], a steady-state simulation methodology for weakly regen-
erative processes is developed that retains all the attractive
features of the classical regenerative method.

In Section 5, we prove that the assumption of a weak regenerative
embedding for a Markov chain is in fact equivalent to assuming that
the steady-state simulation problem is well-posed for the chain. This
result shows that a simulation is well-behaved, in a certain sense, if
and only if the process is weakly regenerative. Regenerative struc-
ture is therefore the norm, rather than the exception, in simulation
problems. This has important consequences for a wide variety of simu-
lation related questions: Quantile estimation (IGLEHART (1976)),
extreme values (IGLEHART (1977)), and sequential stopping rules for
simulations (LAVENBERG and SAUER (1977)) are among the examples.

In Section 6, we study two examples which show how a Markov chain
can fail to have weak regenerative structure. Section 7 considers a
necessary and sufficient condition for the simulation problem to be
well-posed, and shows that the concept of well-posedness is equivalent
to assuming a certain smoothness on the transition kernel.

2. General State Space Markov Chains

Let $E$ be a complete, separable metric space, and $E$ its
associated Borel sets. A function $P: E \times E \rightarrow [0,1]$ is called a
probability transition kernel if:
1) $P(x,\cdot)$ is a probability measure on $(\mathbb{E},\mathcal{E})$, for each $x \in \mathbb{E}$.

2) $P(\cdot,B)$ is an $\mathbb{E}$-measurable function for every $B \in \mathcal{E}$.

The $n$-step transition probabilities are defined through the relations

\[
P^n(x,B) = P(x,B)
\]

\[
P^{n+1}(x,B) = \int_{\mathbb{E}} P^n(y,B) P(x,dy)
\]

Let $\Omega$ be the product space $\mathbb{E} \times \mathbb{E} \times \cdots$ and $\mathcal{F}$ the associated product $\sigma$-field. For $\omega = (\omega_0, \omega_1, \ldots) \in \Omega$, put $X_i(\omega) = \omega_i$. Then, for each probability $\mu$ on $(\mathbb{E},\mathcal{E})$, there exists a unique probability measure $P_\mu$ on $(\Omega,\mathcal{F})$ (see Ionescu-Tulcea (1949)) such that for any $n \geq 0$ and $B_0, \ldots, B_n \in \mathcal{E}$,

\[
P_\mu\{X_0 \in B_0, \ldots, X_n \in B_n\}
= \int_{B_0} \mu(dx_0) \int_{B_1} P(x_0, dx_1) \cdots \int_{B_n} P(x_{n-1}, dx_n).
\]

We shall commonly write $P_x$ to denote the measure $P_\mu$, where $\mu = \delta_x$ ($\delta_x(B) = 1$ or 0 depending on whether or not $x \in B$). If $E_\mu(\cdot)$ corresponds to expectation under $P_\mu$, then it can be shown (see Revuz (1957), p. 16) that the measurability of the kernel $P$ allows one to write

\[
E_{\mu_Z} = \int_{\mathbb{E}} E_{\mu_Z} \mu(dx)
\]

for any bounded $\mathcal{F}$-measurable $Z$. 
The process \( \{X_n\} \), constructed relative to \( (\mathcal{Q}, \mathbb{Q}, P) \), possesses the Markov property. To be precise, let \( \theta_n: \mathbb{Q} \to \mathbb{Q} \) be the shift given by

\[
\theta_n(\omega) = (\omega_n, \omega_{n+1}, \ldots).
\]

Then, for any bounded \( F \)-measurable \( Z \),

\[
(2.2) \quad E \left( Z \cdot \theta_n \mid F_n \right) = E_{X_n} Z \quad \text{for } P \mu \text{-a.s.}
\]

where \( F_n = \mathcal{B}(X_0, ..., X_n) \), the \( \sigma \)-field generates by \( X_0, ..., X_n \). Equation (2.2) is a statement of the Markov property and justifies calling \( \{X_n\} \) the general state space Markov chain associated with kernel \( P \).

3. Well-Posed Steady State Simulation Problems

In problems where the decision time horizon is either infinite or very large, a simulator is often concerned with the long-run behavior of the sample time averages

\[
r_n(f) = \frac{1}{n+1} \sum_{k=0}^{n} f(X_k)
\]
where the function $f$ corresponds to some performance criterion for the stochastic system $(X_n)$ under study. The simulator expects that $r_n(f)$ converges in some sense to a number $r(f)$, and desires a confidence interval for $r(f)$.

We specialize now to the Markov chain context and suppose that $f \in b\mathbb{E}$, the class of real-valued bounded $\mathbb{E}$-measurable $f$. Observe that for each $x \in \mathbb{E}$, the sequence $(E_x r_n(f); n \geq 0)$ is then bounded. Hence, the Bolzano-Weierstrass theorem implies that all subsequences chosen from $(E_x r_n(f); n \geq 0)$ contain a further subsequence which is convergent. The simulator's expectation of the system "settling down" then requires that all limit points must coincide. Hence, in any "reasonable" simulation of a Markov chain, we must have

\[(3.1) \quad E_x r_n(f) \to r(f,x)\]

$r(f,x)$ being the common limit point constructed above.

Furthermore, in a well-behaved simulation, the limit in (3.1) should be independent of the initial position $x$, as one would hope that the simulator's initial condition plays no role in the convergence issue. This motivates the following definition.

\[(3.2) \quad \text{DEFINITION. The steady-state simulation problem is said to be well-posed for the chain } (X_n) \text{ if for all } f \in b\mathbb{E}, \text{ there exists a number } r(f) \text{ such that} \]

\[(3.3) \quad E_x r_n(f) \to r(f),\]
the convergence holding for all $x$.

One point requires further elaboration. We require that (3.3) occur for all $f \in bE$, because otherwise the simulator would have to check that his performance criterion $f$ lies in the class for which convergence holds, and this is, in general, unrealistic to expect. We return to this question later.

Our first theorem shows that chains that are well-behaved from a simulation standpoint are endowed with some important properties.

(3.4) **THEOREM.** Suppose that the steady-state simulation problem is well-posed for $\{X_n\}$. Then, there exists a $P$-invariant probability measure $\pi$ on $(E,E)$ such that for all $x \in E$, we have

$$r_n(f) = r(f) = \pi f = \int f(y) \pi(dy)$$

$P_x$ a.s., provided $\|f\| < \infty$.

**Proof.** Letting $f$ range over the class of indicator functions in (3.3), we get

\begin{equation}
\frac{1}{n+1} \sum_{k=0}^{n} P_x(X_k \in A) = r(A)
\end{equation}
for all $A \in \mathcal{E}$. Since the set function on the left-hand side of (3.5) is a probability measure for all $n$, the Vitali-Hahn-Saks theorem (RENYI (1970)) shows that the set function $r(I_A)$ is also a probability measure, say $\pi(\cdot)$. A simple argument then proves that $r(f) = \pi f$, allowing us to re-write (3.3) as

$$\frac{1}{n+1} \sum_{k=0}^{n} (p^k f)(x) + \pi f,$$

where $(p^k f)(x) = \int f(y) p^k(x, dy)$. It follows, from substituting the bounded function $g(y) = (Pf)(y)$ in (3.6), that $\pi$ must satisfy

$$\pi f = \pi Pf$$

for all $f \in \mathcal{B}\mathcal{E}$, and hence $\pi$ must be $P$-invariant.

To obtain almost sure convergence, we first observe that the shift $\theta$ is a measure-preserving transformation under $P_\pi$, and thus Birkhoff's ergodic theorem (see LAMPERTI (1977), p. 87-95 for definitions and results) applies, yielding

$$r_n(f) \rightarrow Z(f) \quad P_\pi \text{ a.s.}$$

provided $|f| < (Z(f)$ is the bounded invariant r.v. $E_\pi fg(X_0) I$), where $I$ is the class of shift invariant events). Furthermore, $Z(f) = \pi f$, $P_\pi$ a.s. in the case that $I$ is a $P_\pi$-trival $\sigma$-field (i.e., all events in $I$ have either $P_\pi$ probability 0 or 1).
The triviality of \( I \) follows by studying the P-harmonic functions. A function \( g \) is said to be P-harmonic if

\[
g(x) = (Pg)(x)
\]

for all \( x \in \mathbb{E} \). Thus, if \( g \) is a bounded P-harmonic function, (3.6) proves that \( g(x) = \pi g \), for all \( x \in \mathbb{E} \), and thus all bounded P-harmonic functions are constants. A basic result in Markov chain theory (see e.g., OREY (1971), p. 18) yields the triviality of \( I \) as a consequence.

We complete the proof by setting

\[
A = \{\omega: r_n(f,\omega) + \pi f\}
\]

and observe that \( Y = I_A \) is a shift invariant event. It follows easily that \( h(x) = E_X Y \) is a bounded P-harmonic function, and thus \( h(x) = P_\pi(A) \). But (3.7) proves that \( P_\pi(A) = 1 \), yielding \( P_{\pi}(A) = 1 \) for all \( x \in \mathbb{E} \). \( \square \)

This theorem shows that convergence of expectations turn into almost sure convergence provided that the chain is well-behaved from a simulation standpoint. This demonstrates that well-posedness is a rather strong requirement to impose.

In fact, it turns out that the concept of well-posedness is equivalent to one of the best known recurrence concepts for general state space Markov chains. A chain \( \{X_n\} \) is said to be Harris
v-recurrent (see HARRIS (1956)) if there exists a σ-finite measure \( v \) on \((E, \mathcal{E})\) such that \( v(A) > 0 \) implies

\[
P_x\left( \sum_{k=1}^{\infty} I_A(X_k) = \infty \right) = 1
\]

for all \( x \in E \). Theorem 3.4 shows that a chain for which the simulation problem is well-posed is Harris \( \pi \)-recurrent. A converse is also available.

(3.8) Proposition. The steady-state simulation problem is well-posed for \( \{X_n\} \) if and only if \( \{X_n\} \) possesses an invariant probability \( \pi \) and is Harris \( \pi \)-recurrent.

Proof. Only the converse needs explanation. By Corollary 4.21, the process \( \{X_n\} \) obeys the strong law

\[
\sum_{k=0}^{n} f(X_k)/(n+1) + r(f)
\]

\( P_x \) a.s., for all \( f \in \mathcal{B}_E \) and all \( x \in E \). The bounded convergence theorem, applied to (3.9), then proves the result. \( \Box \)

4. Weakly Regenerative Structure of Well-Posed Problems

We will show, in this section, that Harris chains can be embedded in a weakly regenerative environment. This will be accomplished by
using a "splitting" technique due to ATHREYA and NEY (1978) and NUMMELIN (1978). Our first task is to obtain existence of a "splitting" measure $\phi$.

(4.1) PROPOSITION. Let $M_m$ denote the set of pairs $(\phi, \lambda)$, where:

1) $\lambda(x)$ is an $E$-measurable function,

2) $P^m(x,*) \geq \lambda(x) \phi(*)$ for all $x \in E$,

3) $\int \lambda(x) \nu(dx) > 0$.

Then, $M_m$ is non-empty for some $m$, provided $P$ is a Harris kernel, and the corresponding Markov chain $(X_n)$ is Harris $\nu$-recurrent.

Proof. The process $(X_n)$ is Harris $\nu$-recurrent, taking values in a separable metric space $E$. Since the $\sigma$-field $E$ is therefore countably generated, Theorem 2.1 of [20] guarantees existence of a $C$-set $C$, and an integer $m$, such that $\nu(C) > 0$ and

\[
\inf_{(x,y) \in C \times C} p^m(x,y) = \alpha > 0 ,
\]

where $p^m(x,*)$ is the derivative of the $\nu$-absolutely continuous part of $P^m(x,*)$ with respect to $\nu$. Then, (4.1)(i) to (iii) are satisfied by $\phi(*) = \nu(C \cap *)/\nu(C)$, and $\lambda(x) = \alpha I_C(x) \nu(C)$. 

The following result will also be necessary.
PROPOSITION. Let $(\phi, \lambda(x)) \in \mathbb{N}$.

Then, there exists a set $A \in E$ with $\nu(A) > 0$ such that:

1) $\inf(x(x): x \in A) = \lambda > 0$,

2) $P_x(X_{\infty} \in A \text{ infinitely often}) = 1$ for all $x \in A$.

If, in addition, $(X_n)$ possesses an invariant probability $\pi$, then $A$ contains a subset $A_1 \in E$ such that $\nu(A_1) > 0$, and

3) $\sup_{x \in A_1} E_T(A_1) = \omega$, where $T_x(A_1) = \inf\{k: X_{km} \in A_1\}$

Proof. Let $A_n = \{x: \lambda(x) > 1/n\}$, and observe that $\nu(A_n) > 0$ for $n$ sufficiently large, since $\int \lambda(x) \nu(dx) > 0$. We now recall that all Harris chains possess a cyclic decomposition $C_1, C_2, \ldots, C_d, F$ such that (see Theorem 1.3 of [22], p. 162):

1) $C_1, \ldots, C_d, F$ are $E$-measurable subsets partitioning $E$, with $\nu(F) = 0$,

2) $P(\cdot, C_{i+1}) = 1$ on $C_i$ for $i = 1, \ldots, d-1$ and $P(\cdot, C_1) = 1$ on $D_d$,

3) $d$ is the period of the chain.

Let $A_n$ be such that $\nu(A_n) > 0$, and observe that $\nu(A_n \cap C_1) > 0$ for some $i$. Setting $A = A_n \cap C_i$, it is easy to verify that $(X_{nd})$ is an aperiodic $\tau$-recurrent Harris chain on $C_i$, where $\tau(\cdot) = \nu(C_i \cap \cdot)$. The chain $(X_{nd})$ on $C_i$ therefore has a unique $\sigma$-finite invariant measure $\pi_d$, and $\tau \ll \pi_d$ ($\ll$ denotes absolutely continuous with respect to); see Theorem 2.7 of [22]. One can now apply Lemma 2.1 of [19], which shows that the skeleton $(X_{ndm}: m \geq 0)$ on $C_i$ is again Harris recurrent, with invariant
measure \( \pi_{dm} \). But since \( \pi_d \) is invariant for \( P^d \), and hence \( P^{dm} \), it must be that \( \pi_d = \pi_{dm} \) by unicity of invariant measures for Harris chains. Now, \( \tau \ll \pi_{dm} \) so \( \pi_{dm}(A) > 0 \), and therefore

\[
P_x \{ X_{ndm} \in A \text{ infinitely often} \} = 1
\]

for all \( x \in C_i \), proving (4.3)(i) and (ii).

Now, suppose \( \{X_n\} \) possesses an invariant probability \( \pi \). Then, \( \pi_d = \pi(\cdot \cap C_i)/\pi(C_i) \) is easily shown to be an invariant probability of \( \{X_{nd}\} \) on \( C_i \), and consequently an invariant probability for \( \{X_{ndm}\} \). For a Harris recurrent chain with an invariant probability \( \pi \), the set of 1-regular states has \( \pi \)-measure 1 (see [19], Proposition 3.5(i), for definitions and details), and hence, by Theorem 3.1 of COGBURN (1975), the chain possesses a strongly uniform subset \( D \). Applying this result to \( \{X_{ndm}\} \), we obtain existence of a set \( D \) such that \( \pi_d(D) > 0 \) and

\[
\sup_{x \in D} E_x T_D^{(md)} < \infty.
\]

Now, the set of 1-regular states is given by (Proposition 3.4, [5]),

\[
R = \{ x : E_x^t T_D < \infty \}
\]

and thus \( \pi_d(D_n) \downarrow \pi_d(A) \) where
\[ D_n = \{ x \in A : \mathbb{E}_x T_d \leq n \} . \]

By Lemma 3.2 of [5], it follows that for \( n \) sufficiently large, \( D_n \) is strongly uniform for \( \{ X_{nmd} \} \) with \( x_d(D_n) > 0 \). This \( D_n \) plays the role of \( A_1 \) in (4.3)(iii).

Let \((\phi, \lambda) \in \mathcal{M}_m\), and observe that

\[ (4.4) \quad P^\lambda(x, \cdot) = \lambda(x) \phi(\cdot) + (1-\lambda(x)) Q(x, \cdot) \]

where \( Q(x, \cdot) \) is defined via (4.4). Of course, \( 0 \leq \lambda(x) \leq P^\lambda(x, E) = 1 \), and hence (4.4) proves that the signed measures \( Q(x, \cdot) \) are, in fact, probability measures.

In the case \( m = 1 \), Athreya, Ney and Nummelin have applied the decomposition (4.4) to show that the chain \( \{ X_n \} \) may be regarded as a regenerative stochastic process. This allowed them to simplify the proofs of a variety of classical results in the theory of Harris chains.

However, the parameter \( m \) is, in general, greater than one. In that event, it may be that only a weakly regenerative embedding of \( \{ X_n \} \) is possible. To describe the notion of a weakly regenerative process \( \{ X_n \} \), defined relative to a probability space \( (Q, \mathcal{G}, P) \), we need some definitions.

\[ (4.5) \quad \text{DEFINITION. A random time } T \text{ is a } Q\text{-measurable function from } Q \text{ into the non-negative integers.} \]
For any two random times $T_1, T_2$ with $T_1 \leq T_2$, we define $\mathcal{F}_{T_2}^{T_1}$ to be the $\sigma$-field generated by all events of the form $A \cap (T_2 - T_1 - k)$, where $A \in \mathcal{F}_{T_1}^{T_1+k} = \mathcal{B}(X_{T_1}, \ldots, X_{T_1+k})$.

(4.6) DEFINITION. The process $(X_n)$ is said to be weakly regenerative of order $\alpha$ if there exist random times $0 = T_0 < T_1 < T_2 < \cdots$ such that:

i) the $\sigma$-field $\mathcal{F}_{T_n}^{T_{n+1}-1}$ are $\alpha$-dependent for some $\alpha$, i.e., the $\sigma$-fields $\mathcal{F}_{T_n}^{T_{n+1}-1}$ and $\mathcal{F}_{T_{n+j}}^{T_{n+j+1}}$ are independent for $j > \alpha$,

ii) for $l, n \geq 1$ and $j_0 < \cdots < j_l$,

$$
\mathbb{P}((X_{T_n}, \ldots, X_{T_{n+k}}) \in B; \tau_n = j_0, \ldots, \tau_{n+l} = j_l) = \mathbb{P}((X_{T_1}, \ldots, X_{T_{1+k}}) \in B; \tau_1 = j_0, \ldots, \tau_{l+1} = j_l)
$$

for each $B$ in the product $\sigma$-field $\mathcal{F}^k$, where $\tau_n = T_{n+1} - T_1$.

We now turn to construction of a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, which supports the chain $(X_n)$ under initial distribution $\mu$, such that the Harris process $(X_n)$ is weakly regenerative of order 1. Recall that $\mathcal{E}$ is a complete, separable metric space, and thus, so is the product space $\mathcal{E}^m$ under the product topology. Furthermore, the separability of $\mathcal{E}^m$ implies that the Borel sets under the product topology coincide with the product $\sigma$-field $\mathcal{E}^m$ (see Dellacherie and Meyer (1978), p. 9). Thus, there exist regular conditional distributions $P_x(y,dz)$ (see Breiman (1968), p. 79, 401) such that
\[
P_x(B) = \int \int I_B(z,y) \, P_x(y, dz) \, P^m(x, dy)
\]

for all \( B \in \mathbb{E}^m \).

Suppose now that \( H_\mu \) is non-empty, and that \((\phi, \lambda) \in H_\mu\).

Let \( \mathbb{E}^k = \bigcup_{j=0}^{m} E_j^k \times A \), \( E_\mu = E^m \times \{0,1\} \), \( \hat{Q} = \mathbb{E}^k \times E_\mu \times E_\mu \times \ldots \), and take \( \hat{F} \) as the corresponding Borel product \( \sigma \)-field (treat \( \mathbb{E}^k \) as the topological sum of \( E_j^k \times A \), as \( j \) ranges from 0 to \( m \)).

Denote the coordinate projections on \( \hat{Q} \) by \( Y_0, (Y_1, \delta_1), (Y_2, \delta_2), \ldots \)

and construct a probability measure \( \hat{P}_\mu \) on \( (\hat{Q}, \hat{F}) \) as follows.

Writing \( Y_0 = (Y_0(0), \ldots, Y_0(N)) \), \( Y_k = (Y_k(1), \ldots, Y_k(m)) \), we set, for \( B \in \mathbb{E}^{k+1} \)

\[
\hat{P}_\mu(Y_0 \in B; N = k) = P_\mu((X_0, \ldots, X_k) \in B; S_A = k)
\]

where \( S_A = \inf\{k \geq 0: X_k \in A\} \), and \( A \) is the \( \mathbb{E} \)-measurable subset of \((4.3)(i) \) and \((ii) \) corresponding to \((\phi, \lambda) \). Recall that \( \nu(A) > 0 \) so that \( P_\mu(S_A = \infty) = 1 \), and hence (4.7) defines \( \hat{P}_\mu \) as a probability on \( \mathbb{E}^k \). Also, for \( x \in \mathbb{E} \) and \( B \in \mathbb{E}^m \), let

\[
\hat{P}(x, B \times \{0\}) = (1 - \lambda(x)) \int \int I_B(v, u) \, P_x(u, dv) \, Q(x, du),
\]

(4.8)

\[
\hat{P}(x, B \times \{1\}) = \lambda(x) \int \int I_B(v, u) \, P_x(u, dv) \, \phi(du).
\]

For any \( x \in \mathbb{E} \), we can then define a probability measure \( \hat{P}_x \) on \( \mathbb{E}_m \times \mathbb{E}_m \times \ldots \) by the formula.
existence of such a measure follows by the theorem of Ionescu-Tulcea. The measure \( \hat{\mathbb{P}}_\mu \) is then given by

\[
\hat{\mathbb{P}}_\mu \{ Y_0 < B_0, N = k, (Y_1, \delta_1, \ldots, Y_n, \delta_n) \in B_1 \} = \int \hat{\mathbb{P}}_\mu \{ Y_0 \in dy_0; N_1 = k \} \hat{\mathbb{P}}_\nu(k) \{ (Y_1, \delta_1, \ldots, Y_n, \delta_n) \in B_1 \}.
\]

Finally, we define the process \( \{ X_n \} \), on the probability space \( (Q, \mathbb{P}, \mathbb{P}_\mu) \), by the formula

\[
X_0 = Y_0(k); \quad k \leq N
\]

\[
X_{N+k+n+1} = Y_{k+1}(i); \quad 1 \leq i \leq m.
\]

(4.10) PROPOSITION. The process \( \{ X_n \} \) has marginal distribution \( \mathbb{P}_\mu \) on \( (Q, \mathbb{F}) = (E^\infty, \mathbb{F}^\infty) \).

Proof. For \( B_1 \subseteq E^{k+1}, B_2 \subseteq E^n \), it is sufficient to prove that

\[
\hat{\mathbb{P}}_\mu \{ (X_0, \ldots, X_n) \in B_1, (X_{k+1}, \ldots, X_{k+n}) \in B_2, N = k \}
\]

\[
= \mathbb{P}_\mu \{ (X_0, \ldots, X_n) \in B_1, (X_{k+1}, \ldots, X_{k+n}) \in B_2, S_A = k \}.
\]
But this follows from (4.7) and (4.8), upon observing that for \( B \in \mathcal{F}^\infty \),

\[
\mathbb{P}(x, B) = \int \int I_B(v, u) \mathbb{P}_x(u, dv) \mathbb{P}(x, du)
= \mathbb{P}_x(\{X_1, \ldots, X_n\} \in B) \cdot 1
\]

Let

\[
T_1 = \inf\{k \geq 1: \delta_k = 1\}
\]

\[
T_n = \inf\{k > T_{n-1}: \delta_k = 1\}
\]

(4.11) PROPOSITION. The process \( \{(X_n, \delta_n): n \geq 1\} \) is weakly regenerative of order 1 with respect to the random times \( \{T_j: j \geq 2\} \), under the probability \( \mathbb{P}_x \), for all \( x \in A \).

Proof. We first need to shown that the \( T_n \)'s are finite \( \mathbb{P}_x \) a.s.

Now,

(4.12) \[ \mathbb{P}_x(T_1 = \infty) \leq \mathbb{P}_x(\delta_{R_1+1} = 0, \ldots, \delta_{R_n+1} = 0) \]

where \( R_0 = 0, R_{n+1} = \inf\{k > R_n: Y_k(m) \in A\} \). Now, recall that \( X_{km} \in A \) infinitely often, provided \( X_0 \in A \). Hence, by Proposition 4.8, the \( R_k \)'s are all finite \( \mathbb{P}_x \) a.s., and thus, by the strong Markov property applied at time \( R_n \), we can bound (4.12) by
since \((1-\lambda(y)) \leq 1-\lambda\) for \(y \in A\). Repeating this process \(n-1\) more times shows that

\[
P_X(T_1 = \infty) \leq (1-\lambda)^n
\]

for all \(n\), and hence \(T_1 < \infty, P_X\) a.s. A similar proof shows that \(T_n < \infty, P_X\) a.s. for all \(n\).

Now, let \(G_{U_2}^{U_1}\) be the \(\sigma\)-field generated by events of the form \(A \cap (U_2 - U_1 = k)\), where

\[
A \in G_{U_1}^{U_2} = B((Y_{U_1}, \delta_{U_1}), \ldots, (Y_{U_1+k}, \delta_{U_1+k})),
\]

and \(U_1 \leq U_2\) are random times. Let \(F\) be the \(\sigma\)-field \(E_m \times E_m \times \cdots\) and let \(\tilde{\theta}\) be the shift on \(E_m \times E_m \times \cdots\). Then \((Y, \tau) \circ \tilde{\theta} = ((Y_2, \tau_2), \ldots)\) where \((Y, \tau) = ((Y_1, \tau_1), (Y_2, \tau_2), \ldots)\). For \(B \in F\),
(4.14) \[
\mathbb{P}(\mathbb{Y}_{n+1} \in \mathbb{B} | \mathbb{G}_n^T) \\
= \mathbb{E}(\mathbb{P}(\mathbb{Y}_{n+1} \in \mathbb{B} | \mathbb{G}_n^T) | \mathbb{G}_n^T) \\
= \mathbb{E}(\mathbb{P}(\mathbb{Y}_{n+1} \in \mathbb{B} | \mathbb{G}_n^T) | \mathbb{G}_n^T) \\
= \int_{\mathbb{P}(\mathbb{Y}_{n+1} \in \mathbb{B})} \phi(dx)
\]

provided that one can show that

\[
\mathbb{P}(\mathbb{Y}_{n+1} \in \mathbb{B} | \mathbb{G}_n^T) = \phi(\mathbb{A}) \text{ a.s.}
\]

But this follows from the fact that

\[
\mathbb{P}(\mathbb{Y}_0, \ldots, \mathbb{Y}_{k-1} \in \mathbb{B}, \mathbb{Y}_k(m) \in \mathbb{A}, T_n = k) \\
= \mathbb{E}(\phi(\mathbb{A}) \lambda(\mathbb{Y}_{k-1}(m)); \mathbb{Y}_0, \ldots, \mathbb{Y}_{k-1} \in \mathbb{B}; T_{n-1} < k) \\
= \phi(\mathbb{A}) \mathbb{P}(\mathbb{Y}_0, \ldots, \mathbb{Y}_{k-1} \in \mathbb{B}; T = k).
\]

Equation (4.14) shows that \( \mathbb{E}_{n-1} \) and \( \mathbb{E}_{n+1} \) are independent, and have the "shift" property (4.6)(ii), proving the result.
Because of the form of $P_\mu$, it is now a simple matter to verify that $Y_0, (Y_1, \delta_1), (Y_2, \delta_2), \ldots$ is itself weakly regenerative of order 1, under the probability $P_\mu$.

A weakly regenerative process $(X_n)$ is said to be **positive recurrent** if $E \tau_2 < \infty$. We will now extend Proposition 4.11 so as to address the recurrence nature of the weakly regenerative embedding of a Harris chain.

(4.15) **PROPOSITION.** Let $(X_n)$ be a Harris chain with invariant probability $\pi$. Then, $(X_n)$ can be embedded in a probability space $(\Omega, \mathcal{F}, \hat{P}_\mu)$ on which $(X_n)$ is positive recurrent as a weakly regenerative process.

**Proof.** We shall prove that for any $(\phi, \lambda) \in M_\mu$, the weak regenerative embedding of (4.7) through (4.8) yields a positive recurrent weakly regenerative process.

Let $A_1 \subseteq A$ be as in (4.3)(iii), and set $U_0 = 0$,

$$U_k = \inf\{j > U_{k-1} : Y_j(m) \in A_1\} .$$

Observe that for $x \in A_1$,

$$\bar{E}_x U_k \leq k \sup_{x \in A_1} E_\mu \tau_{A_1}(m) \equiv k\beta < \infty .$$

Then, in particular,
\[ \mathbb{E}_x ( \mathbb{E}_{Y_1(m)} U_1 ) \leq \mathbb{E}_x U_2 \]

so that

\[ \int \mathbb{E}_y U_1 \phi(dy) \leq \frac{\mathbb{E}_x U_2}{\lambda} < = \]

for \( x \in A_1 \). Hence, for \( x \in A \), we have

\[ (4.16) \quad \mathbb{E}_x T_2 \leq \mathbb{E}_x ( \mathbb{E}_{Y_1(m)} (T_1) ) \leq \int \mathbb{E}_x T_1 \phi(dx) \]

\[ \leq \frac{\mathbb{E}_y U_2}{\lambda} + \int \mathbb{E}_x ( \mathbb{E}_{Y_1(m)} (T_1) ) \phi(dx) . \]

where \( y \in A_1 \) is fixed. Now, given \( x \in A_1 \), it is evident that

\[ (4.17) \quad \mathbb{E}_x T_1 \leq \sum_{k=1}^{\infty} \mathbb{E}_x (U_k; \delta_{U_1} = 0, \ldots, \delta_{U_{k-1}} = 0, \delta_{U_{k+1}} = 1) + 1 \]

\[ \leq \sum_{k=1}^{\infty} \mathbb{E}_x (U_k; \delta_{U_1} = 0, \ldots, \delta_{U_{k-1}} = 0) + 1 . \]

Applying the strong Markov property at time \( U_{k-1} \) shows that the \( k \)'th summand in the last expression is just
\[ E_x[U_{k-1}; \delta_1 = 0; \delta_{U_1+1} = 0, \ldots, \delta_{U_{k-2}+1} = 0] \]

\[ + E_x[E_{U_{k-1}}(m); \delta_1 = 0; \delta_{U_1+1} = 0, \ldots, \delta_{U_{k-2}+1} = 0] \]

which is clearly bounded by

\[ (1-\lambda) E_x[U_{k-1}; \delta_{U_1+1} = 0, \ldots, \delta_{U_{k-2}+1} = 0] \]

\[ + \beta P_x[\delta_{U_1+1} = 0, \ldots, \delta_{U_{k-2}+1} = 0] \]

\[ \leq (1-\lambda) E_x[U_{k-1}; \delta_{U_1+1} = 0, \ldots, \delta_{U_{k-2}+1} = 0] \]

\[ + \beta (1-\lambda)^{k-2} ; \]

for the second inequality, see (4.13). Repeating this process \( k-1 \) more times shows that the \( k \)th summand of (4.17) is bounded by \( \beta (1-\lambda)^{k-2} \), and hence

\[ E_x T_1 \leq 2\beta + \sum_{0}^{\infty} (k+2) \beta (1-\lambda)^k + 1 . \]

Since this bound is uniform over \( x \in A \), relation (4.16) proves that \( \bar{E}_x \tau_2 < \infty \) is uniformly bounded as well, on \( x \in A \). The form of \( \bar{P}_\mu \) then dictates that \( \bar{E}_\mu \tau_2 < \infty \), proving the positive recurrence.
Given Proposition 4.15, we immediately obtain the following corollary, the proof of which is based on the strong law for weakly regenerative positive recurrent processes (Theorem 3.1, [10]).

(4.18) COROLLARY. If \( \{X_n: n \geq 0\} \) is a Harris chain with invariant probability, then

\[
\sum_{k=1}^{n} \frac{f(X_k)}{n} + \mathbb{E}_\mu \frac{Y_n(f)}{\mu_2} \mathbb{E}_\mu \tau_2 \quad \text{P}_x \text{ a.s.}
\]

for all \( f \in \mathcal{M} \), where

\[
Y_n(f) = \sum_{k=T_n}^{T_{n+1} - 1} f(X_k)
\]

5. Harris Chains and the Weak Regenerative Property

In Section 4, we showed that Harris chains with invariant probability can be embedded in a positive recurrent weakly regenerative environment. In this section, we will show that the converse holds.

(5.1) DEFINITION. A subset \( A \subset \mathcal{X} \) is said to be **stochastically closed** if \( P(x, A) = 1 \) for all \( x \in A \).

For stochastically closed \( A \), the restriction to \( A \) is the chain with transition kernel \( \hat{P}(x, E) = P(x, E \cap A) \), where \( x \) ranges over \( A \). The following theorem is a precise statement of the converse.
(5.2) **THEOREM.** Let \( P \) be a transition kernel and \( \mu \) a probability on \((\mathbb{E}, \mathcal{E})\). Suppose that there exists a probability space \((Q, \mathcal{G}, \hat{P})\), supporting a sequence \( \{X_n\} \), such that:

i) the marginal distribution of \( \{X_n\} \) is \( P_\mu \),

ii) the process \( (X_n) \) is positive recurrent and weakly regenerative of order \( m \).

Then, \( P \) possesses a stochastically closed subset \( A \) such that the restriction to \( A \) is a Harris chain with invariant probability.

The main tool required in proving this result is the following proposition.

(5.3) **PROPOSITION.** Under the hypotheses of Theorem 5.2, \( P \) possesses an invariant probability \( \pi \) and a set \( B \in \mathcal{E} \), with \( \pi(B) = 1 \), such that

\[
(5.4) \quad \mathbb{E}_{X_n} (f) \to \pi f, \quad \text{for all } x \in B.
\]

Leaving the proof of this result aside for the moment, we can prove Theorem 5.2.

**Proof of Theorem 5.2.** Let \( B \) be as in Proposition 5.3 and put \( B_0 = B \). We then define the decreasing sequence of sets \( B_k \) by the iteration

\[
B_{k+1} = \{ x \in B_k : P(x, B_k) = 1 \}.
\]
Observe that since

\[ \pi(B) = \int P(x,B) \pi(dx) , \]

\[ P(x,B_0) = 1 \text{ for } \pi \text{ a.e. } x, \text{ and hence } \pi(B_1) = 1. \] A simple induction proceeds to show that \( \pi(B_k) = 1 \) for all \( k \). Let \( A = \bigcap_{k=0}^{\infty} B_k \). Then \( \pi(A) = 1 \) and \( P(x,A) = 1 \) for all \( x \in A \), proving that \( A \) is stochastically closed. Furthermore, (5.4) implies that the steady-state simulation problem is well-posed for the restriction on \( A \), and thus, the theorem follows by Proposition 3.8.

For the proof of Proposition 5.3, we will reduce the analysis from that of a weakly regenerative process to that of a family of regenerative processes, i.e., weakly regenerative of order 0. Assuming that \( \{X_n\} \) is weakly regenerative of order \( m \), let \( \{\nu_n: n \geq 0\} \) be processes defined as follows:

\[ \nu^0_i = X_i , \quad \text{if } i < T_1(k) + 1 \]

\[ \nu^1_i = X_i , \quad \text{if } i \leq k, j = 1, \ldots, m \]

\[ \nu^j_{T_1(j)+1} = X_{i+T_1(j)+1} , \quad 0 \leq i < \eta_j(j) \]

where

\[ \eta_j(j) \]
\[ \gamma_r(0) = T_{x(k)+1} + \sum_{s=0}^{r} \gamma_{s}(0) \]

\[ \gamma_r(j) = k + \sum_{s=0}^{r} \gamma_{s}(j), \quad j = 1, \ldots, m \]

\[ \eta_{s}(j) = I(k) + 1 + s(m+1) + j. \]

Basically, the process \( W_n^{j} \) records the process values of \( X \) over weak regenerative epochs with associated indices of the form \( s(m+1)+j \), modulo a common portion of length \( k+1 \) at the beginning. Intuitively then, the individual processes \( W_n^{j} \) are regenerative.

(5.5) **Lemma.** For \( j = 0, 1, \ldots, m \), the processes \( \{ W_n^{j}; n \geq 0 \} \) are regenerative, with respect to the random times \( \{ \gamma_r(j); r \geq 1 \} \).

**Proof.** We prove only the case \( j = 0 \), the other processes being similar. Let \( A \in \mathcal{E}^n \), \( B \in \mathcal{E}^p \), and consider for \( n \geq k \),

\[ \hat{P}(W_0^0, \ldots, W_{n-1}^0) \in A; \gamma_1(0) = n; \]

\[ (W_n^0, \ldots, W_{n+p-1}^0) \in B; \gamma_2(0) = p) \]

\[ = \hat{P}(X_0^0, \ldots, X_{n-1}^0) \in A; T_{x(k)+2} = n; \]

\[ (X_T^0, \ldots, X_T^0) \in B; \tau_{\eta_1}^0 = p) \]

\[ = \sum_{j=0}^{k} \sum_{r_1 < k} \hat{P}(A; r_1, r_2 \cap B; r_1, r_2) \]

27
where

\[ A_{j}, r_1, r_2 = \{(X_0, \ldots, X_{n-1}) \in A; T_j = r_1, T_{j+1} = r_2, T_{j+2} = n\} \]

\[ B_{j}, r_1, r_2 = \{\{(X_{j+l}, \ldots, X_{j+l+(m+1)}) \in B; \tau_{j+l+(m+1)} = p\} \]. \]

Of course, \( A_{j}, r_1, r_2 \in \bigcup_{q=0}^{j+1} B_q \), the \( \sigma \)-field generated by \( B_0, \ldots, B_{j+1} \), whereas \( B_{j}, r_1, r_2 \in B_{j+l+(m+1)} \), and hence the two events are independent by m-dependence. It easily follows that the generating sets of the regenerative \( \sigma \)-fields associated with \( \{\eta_n^0\} \) are independent; property (4.5)(ii) follows through a similar argument.

It is straightforward to show that the regenerative processes \( \{\eta_n^j\} \) are positive recurrent if the process \( \{X_n\} \) is positive recurrent as a weakly regenerative process. Consequently, Proposition 5.9 of [10] can be applied to show that if \( f \in B_\mathbb{E} \), then

\[ \left| \mathbb{E}\left[ \frac{1}{n}\sum_{r=k+1}^{n+k} f(W^j_r)/n \right] - \mathbb{E}Y_1(f; j)/\mathbb{E}Y_1(j) \right| \leq \|f\| Z_n(j) \]

where \( \|f\| = \sup\{|f(x)|: x \in \mathbb{E}\} \), \( Z_n(j) \) is a random variable, independent of \( f \), such that \( Z_n(j) \to 0 \) a.s., and

\[ Y_1(f; j) = \sum_{n=\gamma_1(j)}^{\gamma_1(j+1)-1} f(W^j_n) = \sum_{n=T_{\lambda(k)+1+j+1}}^{T_{\lambda(k)+1+j+1}-1} f(X_n). \]

\[ \tau_1(j) = Y_1(1; j). \]
The representation (5.7), in terms of $X_n$, implies that $E_{Y_1}(f;j)/E_{R_1}(j)$ is independent both of $j$ and $k$; denote the common value by $\pi_f$. We are now ready to prove Proposition 5.3.

Proof of Proposition (5.3). Let $I_j(n)$ be the number of $X_r$, from the collection $\{X_{k+1}, \ldots, X_{k+n}\}$, recorded by $W_j$. Then,

$$\sum_{r=k+1}^{n+k} f(X_r)/n = \sum_{j=0}^{m} \sum_{r=k+1}^{n+k} f(W_r)/n$$

so that

$$E_{I_j}(n+1) = \sum_{j=0}^{m} E_{I_j}(n+1) = \sum_{j=0}^{m} \left\{ \sum_{r=k+1}^{n+k} f(W_r)/n \right\} - \pi_f.$$

Now, it is easy to prove, using $m$-dependence, that

$$\Gamma_n(j) = \left| I_j(n)/n - 1/(m+1) \right| \rightarrow 0$$

$P$ a.s. We can therefore find constants $a_n(j) \rightarrow 0$ such that
see Lemma 5.14 for a proof. Then, by Theorem 9.4.8 of CHUNG (1974),

\[(5.9) \quad \hat{P}(\Gamma_n(j) > a_n(j)|w_0^j, \ldots, w_k^j) \rightarrow 0 \quad \hat{P} \text{ a.s.} \]

Thus, we have that

\[
\frac{t_j(n) + k}{\sum_{r=k+1}^{n} \hat{E}(w_r^j)/n \mid w_0^j, \ldots, w_k^j} - \frac{n}{m+1}
\]

\[
\leq \frac{t_j(n) + k}{\sum_{r=k+1}^{n} \hat{E}(w_r^j)/n \mid w_0^j, \ldots, w_k^j} - \frac{n}{m+1} + a_n(j) \quad \text{if} \quad \hat{P}(\Gamma_n(j) > a_n(j)|w_0^j, \ldots, w_k^j)
\]

\[
\leq \text{if} \quad (z_n(j) + a_n(j) + \hat{P}(\Gamma_n(j) > a_n(j)|w_0^j, \ldots, w_k^j))
\]

by (5.6) and (5.9), where \( b_n \) is the greatest integer less than or equal to \( n/(m+1) \). It therefore follows from (5.8) that there exists a sequence \( R_n \) of random variables, independent of \( f \), such that

\[(5.10) \quad \frac{n}{m+1} \leq \frac{t_j(n) + k}{\sum_{r=k+1}^{n} \hat{E}(x_r^j)/n \mid x_0^j, \ldots, x_k^j} - \frac{n}{m+1} \leq \text{if} \quad R_n \quad \hat{P} \text{ a.s.}
\]

where \( R_n \rightarrow 0 \) \( \hat{P} \) a.s. Let

\[
\tau_n(x,*) = \sum_{r=1}^{n} p^r(x,*)/n
\]
and observe that (5.10) implies that

\[
|\tau_n(x_k, A) - \pi(A)| \leq R_n
\]

\( P \) a.s. Now, \( E \) is the Borel field of a separable metric space, and therefore generated by a countable algebra, call it \( A \). By (5.11), the set \( A_k \) has \( P \) measure 1, where

\[
A_k = \{ \omega : |\tau_n(x_k(\omega), A) - \pi(A)| \leq R_n(\omega), \text{ for all } A \in A, n \geq 1 \}.
\]

Hence, \( P(x_k \in D) = 1 \), where \( D \) is the \( E \)-measurable set

\[
D = \{ x : \sup_{A \in A} |\tau_n(x, A) - \pi(A)| = 0 \}.
\]

We can then apply Lemma 5.15 to any \( x \in D \), to obtain the result that

\[
\int f(y) \tau_n(x, dy) + \pi f
\]

over all \( f \in bE \) simultaneously, provided \( x \in D \). Now, relation (5.12) can be re-written as

\[
E \pi_n f + \pi f,
\]

and hence, the proof of the result is complete, provided that we show \( \pi(D) = 1 \). By the strong law for positive recurrent weakly regenerative processes (Theorem 3.1, [10]), and bounded convergence,
\[
\frac{1}{n} \sum_{k=1}^{n} \hat{P}(X_k \in A)/n + \pi(A)
\]

and therefore it is clear that \( \pi(\cdot) \) is absolutely continuous with respect to

\[
(5.13) \quad \frac{1}{n} \sum_{k=0}^{n} \hat{P}(X_k \in \cdot) 2^{-(k+1)}.
\]

But \( \hat{P}(X_k \in D \text{ for all } k) = 1 \), and hence \( D \) has full measure with respect to the probability given by (5.13).

(5.14) **Lemma.** Suppose that a sequence \( \Gamma_n \to 0 \) a.s. Then, there exists constants \( a_n \to 0 \) such that

\[
P(\Gamma_n > a_n \text{ infinitely often}) = 0.
\]

**Proof.** Let \( m_0 = 0 \) and define \( m_k \) by

\[
m_k = \inf\{m > m_{k-1} : P(\Gamma_n \leq 1/k \text{ for all } n \geq m) \geq 1 - 2^{-k}\}.
\]

For \( m_k \leq j < m_{k+1} \), put \( a_j = 1/k \). Then

\[
P(\Gamma_n > a_n \text{ infinitely often}) \leq \sum_{k=1}^{\infty} P(E_k^c) \leq \sum_{k=1}^{\infty} 2^{-k} < \infty,
\]

where \( E_k = \{\Gamma_n \leq 1/k \text{ for } n \geq m_k\} \). Then, the result follows from the Borel-Cantelli lemma.
(5.15) **LEMMA.** Let \( \tau_n, \tau \) be probabilities on a measurable space \((\mathcal{E}, \mathcal{F})\), and let \( \mathcal{A} \) be an algebra which generates \( \mathcal{F} \). Then, if

\[
\sup_{A \in \mathcal{A}} |\tau_n(A) - \tau(A)| \to 0
\]

one also has

\[
\tau_n f \to \tau f \quad \text{for all } f \in b\mathcal{F}.
\]

**Proof.** Let \( \Delta \) represent the operation of symmetric set difference. We first need to show that if \( B \in \mathcal{F} \), then, for all \( \epsilon > 0 \), there exists \( A \in \mathcal{A} \) such that \( \nu_i(A \Delta B) < \epsilon \) (i.e., \( B \) can be approximated by sets \( \in \mathcal{A} \)), where \( \nu_i \) \( (i = 1,2) \) are probabilities on \((\mathcal{E}, \mathcal{F})\).

We proceed via a "monotone class" argument. Let \( \mathcal{M} = \{B \in \mathcal{F}: B \) can be approximated by sets \( \in \mathcal{A} \}, \) and note that \( \mathcal{A} \supset \mathcal{M}, \) and that \( \mathcal{M} \) is closed under increasing unions and decreasing intersections. The monotone class theorem ([17], p. 13) then shows that \( \mathcal{M} = \mathcal{F} \).

With this result in hand, observe that for \( B \in \mathcal{F}, \) there exists \( B_n \in \mathcal{A} \) such that \( \tau_n(B \Delta B_n) < \epsilon, \tau(B \Delta B_n) < \epsilon. \) Then,

\[
|\tau_n(B) - \tau(B)|
\]

\[
= |\tau_n(B) - \tau_n(B_n)| + |\tau_n(B_n) - \tau(B_n)| + |\tau(B) - \tau(B_n)|
\]

\[
\leq 2\epsilon + \sup_{A \in \mathcal{A}} |\tau_n(A) - \tau(A)| + 2\epsilon.
\]
and hence $\tau_n(A) + \tau(A)$ for all $A \in \mathcal{F}$. For $f \in \mathbf{b}\mathcal{F}$, approximate $f$ uniformly by indicators $f_k$ such that $|f - f_k| < 1/k$. Then, since we already have established convergence for simple functions,

$$|\tau_n f - \tau f| \leq 2|f - f_k| + |\tau_n f_k - \tau f_k| + 2|f_k|$$

which proves our result. $\dagger$

Theorem 5.2 shows that if a chain can be embedded in a weakly regenerative environment, then the chain is Harris recurrent. On the other hand, in Section 4, it was shown that Harris chains can be made weakly regenerative with respect to random times that are nonanticipating with respect to the chain. Hence, we may conclude that if a weak regenerative embedding is at all possible, then it can be done via random times that are nonanticipating.

6. More on the Steady-State Simulation Problem

In the last three sections, we have shown that the assumption of well-posedness is essentially equivalent to either of the following:

i) assuming that the chain is Harris recurrent,

ii) assuming that the chain has a weakly regenerative embedding.

In view of the strength of the well-posedness assumption, it is incumbent upon us to investigate the concept somewhat further. Before proceeding, a little reflection shows that a minimal condition for
discussion of the steady-state simulation problem is existence of an initial probability \( \mu \) and a limiting probability \( v \) such that

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} P^k(x, \cdot) \mu(dx)/n \to v(\cdot)
\]

where \( \to \) denotes weak convergence of probabilities (recall that \( E \) is a metric space so this makes sense). Now, virtually all chains that arise through applied probability considerations have kernels \( P \) that are Feller, i.e., if \( f \in bC \) (bounded \( E \) continuous functions), then \( Pf \in bC \). The following result is well known; we give a proof for the sake of completeness.

(6.2) **LEMMA.** If \( P \) is a Feller kernel and (6.1) holds, then \( v \) is \( P \)-invariant.

**Proof.** For \( f \in bC \), (6.1) implies that

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} (P^k f(x)) \mu(dx)/n + vf \to vPf
\]

But \( Pf \in bC \) as well, so we also get

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} (P^k f(x)) \mu(dx)/n + vPf \to vPf
\]

Hence \( vf = vPf \) for all \( f \in bC \), proving that \( v \) is \( P \)-invariant.
The above discussion suggests that we limit our study of the
simulation problem to chains that have a unique invariant probability.
In the remainder of this section, we consider two examples of chains
that illustrate how the well-posedness assumption can be violated,
even in the presence of such a probability.

Consider the chain \( \{X_n: n \geq 0\} \) on state space \( E = \{0,1,2,\ldots\} \)
associated with transition matrix given by

\[
\begin{align*}
P_{00} &= 1 \\
P_{ij} &= \delta_{i+1,j} \quad \text{for } i \geq 1.
\end{align*}
\]

The simulation problem is not well posed for \( \{X_n\} \) since for
\( f(y) = \delta_{0y} \), we have

\[
E_i r_n(f) = 0 \neq 1 = E_0 r_n(f)
\]

when \( i \geq 1 \), violating (3.3). The difficulty here is that part of the
state space is transient. The situation is saved by restricting the
chain to the subset \( \{0\} \) — the simulation problem is then well-posed
for the restricted process.

One might hope that this can be done in general. Specifically,
the existence of a unique invariant probability \( \pi \) forces the shift
\( \theta \) to be a measure-preserving ergodic transformation on \( (\Omega,F,P) \)
(see ASH & GARDNER, p. 141) for a proof of the ergodicity). Application
of Birkhoff's ergodic theorem proves that for each \( A \in F \) with
\( \pi(A) > 0 \), there exists a set \( N_A \in F \) for which \( \pi(N_A) = 0 \) and

36
for $x \notin N_A$. If one could find a common $\pi$-null set $N$ such that for $x \in N$, (6.3) holds for all $A \subset E$ with $\pi(A) > 0$, then one could hope to argue that the chain restricted to the complement of $N$ is Harris recurrent. Such a result would therefore prove that a chain with unique invariant probability always possesses a restriction for which the simulation problem is well-posed.

However, this is too much to expect, in general. Consider the autoregressive process on $E = [-1,1]$ defined by

$$X_{n+1} = \frac{1}{2} X_n + \epsilon_{n+1}$$

where $\{\epsilon_n : n \geq 1\}$ is independent and identically distributed with $P(\epsilon_n = -1/2) = P(\epsilon_n = 1/2) = 1/2$. Solving the recursion (6.4) yields

$$X_{n+1} = \sum_{k=0}^{n} \left(\frac{1}{2}\right)^k \epsilon_{n+1-k} + \left(\frac{1}{2}\right)^n X_0 .$$

Hence, for $-1 \leq x \leq 1$,

$$E_x \exp(itX_{n+1})$$

$$= \exp(2^{-n} itx) \cdot \prod_{k=0}^{n} E \exp(2^{-k} it \epsilon_{n+1-k})$$

$$= \exp(2^{-n} itx) \cdot \prod_{k=0}^{n} E \exp(2^{-k} it \epsilon_k)$$

37
where the exchangeability of the sequence \( \{e_k \mid k \geq 1\} \) was used in the last equality. Taking limits in (6.5) gives

\[
\lim_{n \to \infty} E_x \exp(it\xi_n) = \prod_{k=0}^{\infty} E \exp(-k it c_k) = \prod_{k=0}^{\infty} \cos(2^{-k-1} \xi) = \sin t/t
\]

(see [4], p. 165, for the infinite product closed form). The continuity theorem for characteristic functions can now be applied to show that

\[
(6.6) \quad X_n \xrightarrow{\text{P}_x-\text{weakly}} U
\]

for all \( x \in [-1,1] \), where \( U \) is a uniform random variable on \([-1,1]\). By Lemma 6.2, it follows that \( \pi(dx) = dx/2 \) is invariant for the chain. Furthermore, since the weak convergence in (6.6) occurs for all \( x \), \( \pi(dx) \) is the unique invariant probability for \( \{X_n\} \).

Suppose now that \( \{X_n\} \), when appropriately restricted, defines a well-posed simulation problem. Then, there must exist a point \( x \in E \) such that

\[
(6.7) \quad \frac{1}{n+1} \sum_{k=0}^{n} (P^k f)(x) + \frac{1}{2} \int_{-1}^{1} f(y) dy/2, \quad \text{for all } f \in \mathbb{B}_E.
\]

Now, observe that when the chain starts at \( x \), the process \( \{X_n\} \) thereafter takes values in the countable set.
A = \{2^{-j}x + j2^{-k}; j, k \geq 0\}.

Thus, putting \( f = I_A \), we see that

\[
\frac{1}{n+1} \sum_{k=0}^{n} P(X_k \in A) = 1 \neq 0 = \int_A \frac{dy}{2},
\]

contradicting (6.7). Hence, the process has no well-behaved restriction. Note that in light of Theorem 5.2, this also implies that \( \{X_n\} \) possesses no initial distribution \( \mu \) under which \( \{X_n\} \) can be embedded in a weakly regenerative environment.

The difficulty here lies in the class of functions for which convergence holds. In the well-posed case, it is assumed that

\[
(6.8) \quad \frac{1}{n+1} \sum_{k=0}^{n} (p^k f)(x) + \pi f
\]

for all \( f \in \mathcal{B}_E \), whereas the autoregressive example above is typical of the case where (6.8) holds only for \( f \in \mathcal{B}_C \). One might call such a case weakly well-posed since (6.8) then is equivalent to the weak convergence statement

\[
\frac{1}{n+1} \sum_{k=0}^{n} p^k(x, \cdot) \Rightarrow \pi(\cdot)
\]

The conclusion is therefore that assumption of only weak convergence of the averaged transition probabilities can lead to processes that
exhibit quite different behavior from that which one might expect from "well-behaved" chains.

From a simulation standpoint, a weakly well-posed problem has several deficiencies. First of all, convergence in (6.8) is guaranteed only for continuous functions $f$, and therefore problems may arise with indicators (indicators arise frequently, in practice). Secondly, for discontinuous $f$, it may be that

$$r_n(f) \to r(f) \quad P_x \text{ a.s.}$$

and yet $r(f) \neq \pi f$. For such an $f$, a simulator would have a tendency to incorrectly interpret $r_n(f)$ as $\pi f$, while the simulation would give no indication of bad behavior.

7. Smoothness of the Kernel and its Relation to Well-posedness

In Section 6, we showed that Markov chains with unique invariant probability can not necessarily be restricted in such a way so as to produce a chain for which the simulation problem is well-posed. It is therefore of interest to produce necessary and sufficient conditions which guarantee existence of such a restriction. Our next theorem addresses this question.

(7.1) THEOREM. Suppose that $\{X_n\}$ has a unique invariant probability $\pi$. Then, the following are equivalent:

i) the chain $\{X_n\}$ possesses a stochastically closed set $A$, with $\pi(A) = 1$, such that the steady-state simulation problem is well-posed for the process restricted to $A$. 

40
there exists a non-trivial \(\sigma\)-finite measure \(\nu\) such that the set

\[
B = \{x: \nu(-) \leq R(x,\cdot)\}
\]

has positive \(\nu\)-measure (\(\leq\) denotes "absolutely continuous with respect to"), where \(R(x,\cdot) = \sum_{k=1}^{\infty} 2^{-k} P^k(x,\cdot)\).

Before proving the result, we remark that \(B\) is an \(\mathcal{F}\)-measurable subset of \(E\) (see REVUZ (1975), p. 78).

**Proof.** Under (i), the process restricted to \(A\) is Harris \(\pi\)-recurrent (Proposition 3.8), where \(\hat{\pi}\) is an invariant probability for the process restricted to \(A\). In fact, because \(A\) is stochastically closed, \(\hat{\pi}\) must also be invariant for the unrestricted chain, and hence \(\hat{\pi} = \pi\). In any case, for all \(x \in A\), \(\pi(D) > 0\) implies that

\[
\sum_{k=1}^{\infty} I_D(X_k) = \infty \quad \mathbb{P}_x \text{ a.s.}
\]

Taking expectations in (7.3) proves that \(R(x,D) > 0\) and thus (ii) holds with \(\nu = \pi\) and \(B = A\).

For the converse, observe that the shift \(\theta\) is a measure-preserving ergodic transformation on \((Q,\mathcal{F},\mathbb{P}_x)\) (the ergodicity follows from the uniqueness of \(\pi\); see [1], p. 141), and hence Birkhoff's ergodic theorem applies. Thus, \(\pi(D) > 0\) implies

\[
\sum_{k=1}^{\infty} I_D(X_k) = \infty \quad \mathbb{P}_x \text{ a.s.}
\]
for \( x \) a.e. \( x \). Taking expectations in (7.4) shows that \( \pi(D) > 0 \) forces \( R(\cdot, D) > 0 \) a.e., and so \( \{X_n\} \) is \( \pi \)-essentially irreducible (see [22], p. 78, for definitions and results). Now, observe that

\[
\nu(\cdot) \leq \int_B R(x, \cdot) \pi(dx)
\]

by (7.2). On the other hand,

\[
\int_B R(x, \cdot) \pi(dx) \leq \int E R(x, \cdot) \pi(dx) = \pi(\cdot)
\]

and hence \( \nu \ll \pi \). Therefore, on the set \( B \), the measures \( \pi(\cdot) \) and \( R(x, \cdot) \) are not mutually singular (they have a \( \nu \)-component in common), and so, by Theorem 2.14 on p. 78 of [22], the chain \( \{X_n\} \) possesses a stochastically closed subset \( A_1 \subset E \) such that \( \pi(A_1) = 1 \), and the restriction to \( A_1 \) is \( \pi \)-irreducible (this means that \( \pi(B) > 0 \) implies that \( R(x, B) > 0 \) for all \( x \in A_1 \)). Let \( G \) be the potential kernel given by

\[
G(x, \cdot) = \sum_{k=0}^{\infty} p^k(x, \cdot)
\]

for \( x \in A_1 \). We will now show that \( G \) is not a proper kernel. The kernel \( G \) is proper if there exists a sequence of sets \( D_n \)
increasing to \( A_1 \) such that \( G(\cdot, D_n) \) is bounded. But \( \pi(D_n) \)
must be positive for some \( n \), and thus, by (7.4), it must be that \( G(\cdot, D_n) = \pi \) a.e., contradicting the properness.
We now apply Theorem 2.6 on page 73 of [22], which shows that the restriction to $A_1$ itself possesses a stochastically closed subset $A$ with $\pi(A) = \pi(A_1) = 1$ such that the restriction to $A$ is Harris $\pi$-recurrent. Since $\pi$ is clearly invariant for the restriction to $A$, one can apply Theorem 3.8 to complete the proof.

One may interpret condition (7.2) as a kind of smoothness condition on the kernel; the kernel $P$ must not be "too singular." We see that the autoregressive process example of Section 6 fails to possess a well-behaved restriction in the most general possible way, namely by violating the smoothness condition on the kernel.
REFERENCES


The general discrete-event simulation can be viewed, by using the technique of supplementary variables, as a Markov chain living in a general state space. For such chains, we can define in precise terms, the notion of an associated well-posed steady-state simulation problem. We prove that the concept of well-posedness is equivalent to assuming that the Markov chain has regenerative-type structure. These two conditions are, in turn, equivalent to assuming a certain smoothness on the transition probabilities of the chain. We also consider two examples which illustrate how a chain can fail to have regenerative-type structure.
ATE LME