J. Appl. Prob. **11**, 726–741 (1974)
Printed in Israel

QUASI-STATIONARY DISTRIBUTIONS FOR MARKOV CHAINS ON A GENERAL STATE SPACE

RICHARD L. TWEEDIE, *The Australian National University, Canberra*

Abstract

The quasi-stationary behaviour of a Markov chain which is ϕ-irreducible when restricted to a subspace of a general state space is investigated. It is shown that previous work on the case where the subspace is finite or countably infinite can be extended to general chains, and the existence of certain quasi-stationary limits as honest distributions is equivalent to the restricted chain being R-positive with the unique R-invariant measure satisfying a certain finiteness condition.

QUASI-STATIONARY DISTRIBUTIONS; R-THEORY; CONDITIONAL LIMITS; GENERAL MARKOV CHAINS; RATIO LIMITS

1. Introduction

Suppose that $\{X_n\}$ is a discrete-time, temporally homogeneous Markov chain on a general state space $(\mathscr{X}, \mathscr{F})$, and that there exists a set $T \in \mathscr{F}$ such that $\{X_n\}$ restricted to T is ϕ-irreducible for some ϕ (irreducible in the sense of Harris). In this paper we study the ergodic behaviour of $\{X_n\}$ given that it begins in T and conditional on its remaining in T. This problem has been considered by Darroch and Seneta [1] when T is a finite set, and by Seneta and Vere-Jones [6] when T is countable. In the latter paper, the main result hinges on the use of the R-theory for countable state space Markov chains created by Vere-Jones in [8] and [9]. Our extension is based on the generalization of R-theory to Markov chains on a general ϕ-irreducible state space, contained in [7]. As far as possible notation in this paper has been kept compatible with that of both [6] and [7].

2. R-theory and some R-positivity results

For each $n = 1, 2, \cdots$, we write, when $x \in \mathscr{X}$ and $A \in \mathscr{F}$,

$$P^n(x, A) = \Pr\{X_n \in A \mid X_0 = x\}, \tag{2.1}$$

where for each $x \in \mathscr{X}$, $P^n(x, \cdot)$ is a probability measure on $\mathscr{F}$, and for each $A \in \mathscr{F}$, $P^n(\cdot, A)$ is a measurable function on $\mathscr{X}$, and where $\mathscr{F}$ is assumed to be separable (that is, countably generated). We also write, for complex z,

Received in revised form 19 November 1973.

$$G_z(x,A) = \sum_{n=1}^{\infty} P^n(x,A)z^n, \qquad x \in \mathscr{X},\ A \in \mathscr{F};$$

$G_z(x,A)$ always converges for at least $|z| < 1$.

Suppose that T is a set in $\mathscr{F}$ such that the chain $\{X_n\}$ restricted to T is ϕ-irreducible for some ϕ (cf. Orey, [4], p. 4); that is, there exists a σ-finite measure ϕ on $\mathscr{F}$ with $\phi(T) > 0$ such that, if $A \in \mathscr{F}$, $A \subseteq T$, then

$$\phi(A) > 0 \textit{ implies } G_{\frac{1}{2}}(x,A) > 0 \textit{ for all } x \in T.$$

Since we shall be interested in the behaviour of $\{X_n\}$ only whilst it remains in T, *we shall make the assumption that $\mathscr{X}\backslash T$ is absorbing*: that is, $G_{\frac{1}{2}}(x,T) = 0$, $x \notin T$. With this assumption, $\{X_n\}$ restricted to T behaves exactly as $\{X_n\}$ on T. We shall therefore drop the phrase "restricted to T", and use $\{X_n\}$ to denote the chain with transition probabilities $P^n(x,A)$, $x \in T$, $A \in \mathscr{T}$, where

$$\mathscr{T} = \{A \in \mathscr{F},\ A \subseteq T\}.$$

Write, for any $B \in \mathscr{T}$,

$$\bar{B} = \{x \in T : G_{\frac{1}{2}}(x,B) > 0\}.$$

It is shown in [7], Lemma 1.1, that the assumption of ϕ-irreducibility implies the existence of a σ-finite measure M on $\mathscr{T}$ with $M(T) > 0$, satisfying the following condition.

Condition I. (i) *$\{X_n\}$ is M-irreducible,*
(ii) *if $M(B) = 0$, then $M(\bar{B}) = 0$.*

We shall henceforth use M to denote such a measure on $\mathscr{T}$, and unless otherwise specified, such phrases as "almost all" will refer to M-measure. It will be convenient to introduce three pieces of notation: we shall use $\mathscr{T}^+$ to denote the collection of subsets of $\mathscr{T}$ with positive M-measure, and, if $\mathscr{K} = (K(j))$ is a partition of T, we use $\mathscr{T}_{\mathscr{K}}$ to denote those elements of $\mathscr{T}$ which are contained in $K(j)$ for some j, and $\mathscr{T}^+_{\mathscr{K}} = \mathscr{T}_{\mathscr{K}} \cap \mathscr{T}^+$. Finally, we assume that $\{X_n\}$ on $(T,\mathscr{T})$ is *strictly substochastic* (that is, for some set $A \in \mathscr{T}^+$, $P(x,T) < 1$ for every $x \in A$), and that $\{X_n\}$ is aperiodic (cf. Orey, [4], p. 15) on $(T,\mathscr{T})$.

With these assumptions, the following results are all proved in [7] except (d), which is implicit in the proof of Theorem 6 of [7], because of [4], Theorem 7.1(i).

Theorem 1. (a) *There exists a partition $\mathscr{K} = (K(j))$ of T, an M-null set N and a real number $R \geqq 1$ such that:*

(i) *R is the radius of convergence of the power series $G_z(x,A)$ for every $x \in T\backslash N$ and every $A \in \mathscr{T}^+_{\mathscr{K}}$;*

(ii) *either $G_R(x,A) = \infty$ for every such x and A, in which case we shall call $\{X_n\}$ R-recurrent; or $G_R(x,A) < \infty$ for every such x and A, and we shall call $\{X_n\}$ R-transient.*

(b) *When $\{X_n\}$ is R-recurrent, there is a unique (up to a constant multiple) σ-finite measure Q, not identically zero, which satisfies*

$$Q(A) = R\int_T Q(dy)P(y, A), \qquad A \in \mathscr{T}; \tag{2.2}$$

Q is equivalent to M on $\mathscr{T}$. There is also a unique (up to a constant multiple and definition on M-null sets) measurable non-negative function f, positive almost everywhere, which satisfies

$$f(x) = R\int_T P(x, dy)f(y), \text{ almost all } x \in T. \tag{2.3}$$

(c) *When $\{X_n\}$ is R-current, there is a partition $\mathscr{K} = K(j))$* and a null set N_f with $\overline{N_f} \subseteq N_f$, such that for all $x \notin N_f$ and $A \in \mathscr{T}_{\mathscr{X}}$,

$$\pi(x, A) = \lim_{n\to\infty} R^n P^n(x, A) = f(x)Q(A)\Big/\int_T f(y)Q(dy), \tag{2.4}$$

where f and Q are as in (b).

Either $\int_T f(y)Q(dy) = \infty$, and $\pi(x, A) = 0$ for all such x and A, when we shall call $\{X_n\}$ R-null; or $\int_T f(y)Q(dy) < \infty$, and $\pi(x, A) > 0$ for all such x, A, when we shall call $\{X_n\}$ R-positive.

(d) *If $\{X_n\}$ is R-positive, then (2.4) is true for every A such that* $\inf_{y\in A} f(y) > 0$. *If A is such that* $\inf_{y\in A} f(y) > 0$, *and $\{X_n\}$ is R-positive, we can in fact assert that*

$$\left\| R^n\int_T \pi(dy)P^n(y, \cdot) - \left[\int_T \pi(dy)f(y)\right] Q(\cdot)\Big/\int_T f(y)Q(dy)\right\|_A \to 0,$$

for any σ-finite measure $\pi(\cdot)$ (including that which allots unit measure to the point $\{x\}$) which satisfies $\int_T \pi(dy)f(y) < \infty$ and $\pi(N_f) = 0$, where $\|\mu\|_A$ is the total variation of a signed measure μ on A.

In the sequel, the null set N_f of (c) and (d) above will be assumed to contain the null set on which (2.3) fails to hold. Our assumption that $\mathscr{F}$ is separable is needed precisely to ensure the existence of this "global" null set: in the non-separable case the nullset in (c) may depend on the set $A \in \mathscr{F}_{\mathscr{X}}$, and whilst our results in Section 3 can be adjusted to account for this, the notation becomes exceedingly tedious.

Note that, in (c), the criterion of R-positivity for R-recurrent chains is the convergence of $\int_T Q(dy)f(y)$. In [7], Sections 4 and 5, various criteria similar to this for chains not necessarily R-recurrent are given. Another such, which we shall need to use, is the following lemma.

Lemma 1. *Suppose H is a σ-finite measure, not identically zero, on $\mathscr{T}$, satisfying, for some real $r > 0$,*

$$H(A) = r \int_T H(dy)P(y, A), \qquad A \in \mathscr{T}, \tag{2.5}$$

and g is a non-negative measurable function, positive on a set of positive measure, satisfying

$$g(x) \geqq r \int_T P(x, dy)g(y) \tag{2.6}$$

for all $x \in T$; and further that

$$\int_T g(y) H(dy) < \infty. \tag{2.7}$$

Then $r = R$, $\{X_n\}$ is R-positive, and $g = f$ and $H = Q$, the unique R-invariant function and measure for $\{X_n\}$ of Theorem 1(b).

Proof. Since g and H satisfy (2.5), (2.6) and (2.7),

$$\begin{aligned} \infty &> \int_T g(y) H(dy) \\ &\geqq r \int_T \int_T P(y, dw) g(w)\, H(dy) \\ &= \int_T H(dw) g(w). \end{aligned}$$

Hence g satisfies (2.6) with equality for H-almost all (and hence M-almost all, since H satisfies (2.5)) $x \in T$, and so g is r-invariant. The result then follows from [7], Sections 4 and 5.

Finally in this section we give some results which are extensions of Theorem 1(c) and (d) in the R-positive case when $Q(T) < \infty$. From here on *we shall assume that whenever the chain $\{X_n\}$ is R-positive on $(T, \mathscr{T})$, the unique R-invariant measure Q and function f are chosen normalized so that* $\int_T f(w) Q(dw) = 1$.

Lemma 2. *Suppose that $\{X_n\}$ is R-positive and $Q(T) < \infty$. Then for any $A \in \mathscr{T}$, and any probability measure π on $\mathscr{T}$ which is bounded by some multiple of Q,*

$$R^n \int_T \pi(dw) P^n(w, A) \to \left[\int_T \pi(dw) f(w) \right] Q(A). \tag{2.8}$$

Proof. Suppose π is bounded by some multiple κQ of Q. Let $\mathscr{K} = (K(i))$ and N_f be as in Theorem 1(c), and A be an arbitrary set in $\mathscr{T}$. Write

$$A_j = A \cap \left(\bigcup_{k \geqq j} K(k) \right).$$

We have

$$R^n \int_T \pi(dw)P^n(w,A) = R^n \int_T \pi(dw)P^n(w,A \setminus A_j) + R^n \int_T \pi(dw)P^n(w,A_j) \tag{2.9}$$

and from Theorem 1(d), since $A \setminus A_j \uparrow A$ as $j \to \infty$, the first term on the right of (2.9) tends with n and j to $[\int_T \pi(dw)f(w)]Q(A)$, since

$$\int_T \pi(dw)f(w) \leqq \kappa \int_T Q(dw)f(w) = \kappa < \infty,$$

and

$$\pi(N_f) \leqq \kappa Q(N_f) = 0.$$

The second term on the right of (2.9) is bounded, since Q is R-invariant, by

$$R^n \kappa \int_T Q(dw)P^n(w,A_j) = \kappa Q(A_j) \tag{2.10}$$

for all n; since $Q(T) < \infty$, (2.10) tends to zero with j, and so (2.8) holds.

Corollary 1. *For almost all $x \in T$ and all $A \in \mathscr{T}$,*

$$\liminf_{n\to\infty} R^n P^n(x,A) = f(x)Q(A) \tag{2.11}$$

whenever $\{X_n\}$ is R-positive and $Q(T) < \infty$.

Proof. From Theorem 1(c), defining A_j as in the lemma,

$$\begin{aligned}\liminf_{n\to\infty} R^n P^n(x,A) &\geqq \lim_{j\to\infty} \liminf_{n\to\infty} R^n P^n(x,A\setminus A_j) \\ &= f(x)Q(A),\end{aligned} \tag{2.12}$$

for almost all x.

Applying Fatou's lemma to (2.8) shows, on the other hand (with $\pi(\cdot) = Q(\cdot)/Q(T)$),

$$\left[\int_T Q(dw)f(w)\right]Q(A) \geqq \int_T Q(dw) \liminf_{n\to\infty} R^n P^n(w,A)$$

so that (2.12) holds with strict inequality on at most a set of Q-measure zero. Since Q and M are equivalent, the corollary is proved.

Corollary 2. *If $\{X_n\}$ is R-positive and $Q(T) < \infty$, and for almost all $x \in T$,*

$$\lim_{n\to\infty} R^n P^n(x,T) = \pi(x,T) \tag{2.13}$$

exists, then for almost all $x \in T$ and all $A \in \mathscr{T}$,

$$\lim_{n\to\infty} R^n P^n(x,A) = f(x)Q(A).$$

Proof. From the previous corollary $\pi(x,T) = f(x)Q(T)$, and so, as in (2.9), (2.14) $$\lim_{j\to\infty}\lim_{n\to\infty} R^nP^n(x, T_j) = 0$$

for almost all x. But for any $A \in \mathscr{T}$,

$$R^nP^n(x, A) = R^nP^n(x, A\setminus A_j) + R^nP^n(x, A_j)$$

and the first of these terms tends with n and j to $f(x)Q(A)$ for almost all x, whilst the second is bounded by $R^nP^n(x, T_j)$ and hence from (2.14) tends to zero with n and j for almost all x.

3. Quasi-stationarity and the main limit theorem

Suppose that the probability of being absorbed into $\mathscr{X}\setminus T$ is one, from every starting point $x \in T$: the quantities in which we are interested in a study of quasi-stationarity of $\{X_n\}$ on T are then (cf. [6], (1)–(4))

(3.1) $$\alpha(x, A, n) = P^n(x, A)/P^n(x, T), \qquad A \in \mathscr{T},\ x \in T,$$

(3.2) $$\tau(x, A, n) = \left\{\frac{1}{n}\sum_{v=1}^{n}\int_A P^v(x, dy)P^{n-v}(y, T)\right\}\Big/ P^n(x, T), \qquad A \in \mathscr{T}, x \in T,$$

(3.3) $$\sigma(x, A, m, n) = \int_A P^m(x, dy)P^n(y, T)/P^{m+n}(x, T), \qquad A \in \mathscr{T},\ x \in T,$$

(3.4) $$\beta(x, y, n) = P^n(y, T)/P^n(x, T), \qquad x, y \in T.$$

A discussion of the probabilistic meanings of (3.1)–(3.4) in the countable case is given in [6], and we will not repeat it in detail here. We remark only that (3.1) represents the probability distribution of $\{X_n\}$ after n steps conditional on $X_0 = x \in T$ and $X_n \in T$, whilst (3.2)–(3.4) are related quantities conditioned on the same type of event.

If we assume that $G_{\frac{1}{2}}(x, \mathscr{X}\setminus T) > 0$ for $x \in T$, and also that the probability of staying in T forever, starting at $x \in T$, is positive, then the limits in (3.1)–(3.3) are all zero for transient A, and we must define certain more general quantities (cf. [6], (6)–(9)). We shall write

$$\eta(x) = \lim_{n\to\infty} P^n(x, \mathscr{X}\setminus T)$$

for the probability of ultimate absorption in $\mathscr{X}\setminus T$; Condition I and our assumption of strict substochasticity of $\{X_n\}$ on $(T, \mathscr{T})$ imply $\eta(x) > 0$ for $x \in T$. Let us now replace (3.1)–(3.4) with

(3.5) $$a(x, A, n) = \int_A P^n(x, dy)\eta(y)\Big/\int_T P^n(x, dy)\eta(y);$$

(3.6) $$t(x, A, n) = \frac{n^{-1}\ \sum_{v=1}^{n} \int_A P^v(x, dy)\int_T P^{n-v}(y, dw)\eta(w)}{\int_T P^n(x, dw)\eta(w)};$$

(3.7) $$s(x, A, m, n) = \frac{\int_A P^m(x, dy)\ \int_T P^n(y, dw)\eta(w)}{\int_T P^{m+n}(x, dw)\eta(w)};$$

$$b(x, y, n) = \int_T P^n(y, dw)\eta(w) \bigg/ \int_T P^n(x, dw)\eta(w). \tag{3.8}$$

The quantities (3.5)–(3.8) are all related to ergodic properties of $\{X_n\}$ conditional on remaining in T for a time n but being ultimately absorbed into $\mathscr{X} \setminus T$; again, see [6] for a somewhat more detailed explanation in the countable case.

In [6], Theorem 3.1, the convergence of the quantities analogous to (3.1)–(3.4) is discussed, and then in [6], Theorem 3.2, it is stated without proof that essentially the same results hold for the quantities analogous to (3.5)–(3.8), when $\eta(x) \not\equiv 1$. Although, for general T, convergence properties of (3.5)–(3.8) do flow reasonably easily from the results for (3.1)–(3.4) (because of the equivalence of convergence set-wise on all sets in a σ-field and convergence of integrals of bounded functions; see, for example, Gänssler [2]) we shall state and prove the theorem for arbitrary η.

Theorem 2. *Suppose as in the previous section that $\{X_n\}$ on $(T, \mathscr{T})$ is aperiodic, strictly substochastic and M satisfies Condition I. Then the following two sets of conditions are equivalent.*

(A) *There is a null set N such that, for $x \in T \setminus N$, the limits* (3.5)–(3.8) *all exist and the limits* (3.5)–(3.7) *define honest probability measures on $\mathscr{T}$, whilst* (3.8) *is not almost everywhere zero.*

(B) *The chain $\{X_n\}$ on $(T, \mathscr{T})$ is R-positive (as defined in Theorem* 1) *with $R > 1$, and the unique R-invariant measure Q satisfies*

$$\int_T \eta(x)\, Q(dx) < \infty; \tag{3.9}$$

and for x outside an M-null set $N_1 \in \mathscr{T}$,

$$\lim_{n\to\infty} R^n \int_T P^n(x, dy)\eta(y) = \pi(x, T) \tag{3.10}$$

exists.

Either (A) *or* (B) *implies that the probability of remaining in T for n steps or longer conditional on ultimately being absorbed into $\mathscr{X} \setminus T$ tends geometrically to zero as R^{-n} from almost all starting points in T.*

When (A) *holds, the limit distributions* (3.5)–(3.7) *are independent of the initial state $x \in T \setminus N$ and the limits of* (3.5)–(3.8) *are related to the unique R-invariant measure Q and function f for almost all $x \in T$ by*

$$\lim_{n\to\infty} a(x, A, n) = \int_A Q(dw)\eta(w) \bigg/ \int_T Q(dw)\eta(w), \qquad A \in \mathscr{T}; \tag{3.11}$$

$$\begin{aligned} \lim_{n\to\infty} t(x, A, n) &= \lim_{n\to\infty} \lim_{m\to\infty} s(x, A, m, n) \\ &= \int_A Q(dw) f(w), \qquad A \in \mathscr{T}; \end{aligned} \tag{3.12}$$

(3.13) $$\lim_{n\to\infty} b(x,y,n) = f(y)/f(x), \text{ almost all } y \in T.$$

4. Proof of the main theorem

We now prove that (B) implies (A). Notice first that throughout Lemma 2 and its corollaries we could have assumed (3.9) and looked at the sequences $R^n \int_A P^n(x,dy)\eta(y)$ rather than assuming $Q(T) < \infty$ and examining sequences $R^nP^n(x,A)$, and the results would have remained unchanged. Thus R-positivity and the assumptions (3.9) and (3.10) imply, as in the second corollary to Lemma 2, that

(4.1) $$\lim_{n\to\infty} R^n \int_A P^n(x,dw)\eta(w) = f(x)\int_A Q(dw)\eta(w)$$

for almost all $x \in T$ and *all* $A \in \mathscr{T}$.

It follows immediately that (3.11) and (3.13) are both true: a little more work is necessary to prove (3.12). Let $\mathscr{K} = (K(j))$ be a partition of $T\backslash N_f$ such that $\inf_{y\in K(j)} f(y) = \delta_j > 0$, and write

$$T(j) = \bigcup_{k\geqq j} K(k), \qquad H(j) = \bigcup_{k<j} K(k).$$

Let $\zeta \in T\backslash N_f$ satisfy (4.1). If we write

(4.2) $$s(\zeta,A,m,n) = \frac{\int_A R^mP^m(\zeta,dy)\int_T R^nP^n(y,dw)\eta(w)}{\int_T R^{n+m}P^{n+m}(\zeta,dw)\eta(w)},$$

then the denominator of (4.2) tends with n to $f(\zeta)\int_T Q(dw)\eta(w)$, for any m. Write the numerator as

(4.3) $$\int_A R^mP^m(\zeta,dy)\int_{H(j)} R^nP^n(y,dw)\eta(w) + \int_A R^mP^m(\zeta,dy)\int_{T(j)} R^nP^n(y,dw)\eta(w);$$

for fixed m, $R^mP^m(\zeta,\cdot\,)$ satisfies $R^mP^m(\zeta,N_f) = 0$ since $\zeta \notin N_f$, and

$$\int_A R^mP^m(\zeta,dy)f(y) \leqq f(\zeta) < \infty.$$

Since η is a bounded function on T, from Theorem 1(d) the first term in (4.3) tends as $n\to\infty$ to

$$\left[\int_A R^mP^m(\zeta,dy)f(y)\right]\left[\int_{H(j)} Q(dw)\eta(w)\right],$$

which tends $\to\infty$ to

$$\left[\int_A R^mP^m(\zeta,dy)f(y)\right]\left[\int_T Q(dw)\eta(w)\right].$$

But in [7], Theorem 6, it is shown that for all $\zeta \in T \backslash N_f$ and *all* $A \in \mathscr{T}$,

$$\int_A R^m P^m(\zeta, dy) f(y) \to f(\zeta) \int_A Q(dy) f(y)$$

(it is from this that the first statement of Theorem 1(d) above is deduced). Thus the first term in (4.3) tends, for almost all $\zeta \in T$, to

$$f(\zeta)\left[\int_A Q(dy) f(y)\right]\left[\int_T Q(dw)\eta(w)\right] \tag{4.4}$$

as successively n, j and m tend to infinity.

The second term in (4.3), on the other hand, is bounded above by

$$\int_T R^m P^m(\zeta, dy) \int_{T(j)} R^n P^n(y, dw)\eta(w) = \int_{T(j)} R^{n+m} P^{n+m}(\zeta, dw)\eta(w)$$

which tends to zero as first n and then j tend to infinity, from (4.1). Thus (4.4) is the limit of the numerator of (4.2) as n and m tend to infinity.

We can in fact interchange the limits with m and n in this calculation without affecting the limit; for if we fix n, $\int_T R^n P^n(y, dw)\eta(w)$ is a bounded function of y in A, and so for almost all ζ the numerator of (4.2) tends with m to

$$f(\zeta) \int_A Q(dy) \int_T R^n P^n(y, dw)\eta(w), \tag{4.5}$$

and the limit of (4.5) with $n \to \infty$, when divided by the limit of the denominator of (4.2), is then (3.12) from Lemma 2.

Finally, we evaluate the limit of (3.6). We need to show the convergence of

$$\frac{1}{n} \sum_{\nu=1}^{n} \int_A R^\nu P^\nu(x, dy) \int_T P^{n-\nu}(y, dw)\eta(w) R^{n-\nu}. \tag{4.6}$$

Choose N to be a fixed large number. Since $\pi_x(\cdot) = \sum_{\nu=1}^N R^\nu P^\nu(x, \cdot)$ satisfies, for all $x \in T \backslash N_f$, $\int_T \pi_x(dy) f(y) = N f(x) < \infty$, and $\pi_x(N_f) = 0$, it follows from Theorem 1(d) that for $x \in T \backslash N_f$,

$$\begin{aligned} \lim_{n\to\infty} \sum_{\nu=1}^{N} \int_A R^\nu P^\nu(x, dy) \int_T P^{n-\nu}(y, dw)\eta(w) R^{n-\nu} \\ = \sum_{\nu=1}^{N} \int_A R^\nu P^\nu(x, dy) f(y) \int_T Q(dw)\eta(w) \\ \leqq N f(x) \int_T Q(dw)\eta(w). \end{aligned} \tag{4.7}$$

Next, note that as $n \to \infty$

$$(4.8) \qquad \frac{1}{n} \sum_{\nu=1}^{n} \int_T R^{n-\nu} P^{n-\nu}(y, dw)\eta(w) \to f(y) \int_T Q(dw)\eta(w)$$

for all $x \in T \backslash (N_f \cup N_1)$, from (3.10).

By Egorov's Theorem and Condition I we can therefore construct a partition $\mathscr{K} = (K(j))$ of T with the property that $M(K(0)) = 0$ and

$$K(0) \supseteq \overline{(K(0)} \cup N_f \cup N_1 \cup \overline{N_1}),$$

and for $j > 0$, $M(K(j)) > 0$, (3.10) and (4.8) both hold uniformly for $y \in K(j)$, and both f and f^{-1} are bounded functions on $K(j)$.

From these properties it follows that for any fixed N,

$$(4.9) \qquad \frac{1}{n} \sum_{\nu=N+1}^{n} \int_T R^{n-\nu} P^{n-\nu}(y, dw)\eta(w) \to f(y) \int_T Q(dw)\eta(w)$$

uniformly in $y \in K(j)$, $j > 0$.

Let A be any set in $\mathscr{T}^+_{\mathscr{K}}$ and pick $\zeta \in T \backslash N_f$. For arbitrary $\delta > 0$ we find $N = N(\delta, A, \zeta)$ large enough that

$$\| R^n P^n(\zeta, \cdot) - f(\zeta) Q(\cdot) \|_A \leqq \delta$$

for all $n \geqq N$, from Theorem 1(d). Consequently, for $n \geqq N$,

$$(4.10) \qquad \frac{1}{n} \sum_{\nu=N+1}^{n} \int_A R^{\nu} P^{\nu}(\zeta, dy) \int_T R^{n-\nu} P^{n-\nu}(y, dw)\eta(w)$$

differs from

$$(4.11) \qquad f(\zeta) \int_A Q(dy) \left[\frac{1}{n} \sum_{\nu=N+1}^{n} \int_T R^{n-\nu} P^{n-\nu}(y, dw)\eta(w) \right]$$

by at most δS_A, where $S_A < \infty$ is the upper bound on the left-hand side of (4.9) for $y \in A$. Since (4.9) holds uniformly for $y \in A$, as $n \to \infty$ (4.11) tends to

$$(4.12) \qquad f(\zeta) \left[\int_A Q(dy) f(y) \right] \left[\int_T Q(dw)\eta(w) \right];$$

so for $A \in \mathscr{T}^+$, $A \subseteq K(j)$ for some $j > 0$, (4.7) and (4.10)–(4.12) imply that the limit as $n \to \infty$ of (4.6) is (4.12) for almost all $x \in T$. Write $T(j) = \bigcup_{k \geqq j} K(k)$; as before, (4.6) tends to (4.12) for arbitrary $A \in \mathscr{T}$ and $x \in T \backslash K(0)$ if and only if

$$(4.13) \qquad \lim_{j \to \infty} \lim_{n \to \infty} \frac{1}{n} \sum_{\nu=1}^{n} \int_{T(j)} R^{\nu} P^{\nu}(x, dy) \int_T P^{n-\nu}(y, dw)\eta(w) R^{n-\nu} = 0.$$

Since (4.6) for $A = T$ is merely $R^n \int_T P^n(x,dw)\eta(w)$ which tends by assumption to $f(x) \int_T Q(dw)\eta(w)$ for $x \notin N_1 \subseteq K(0)$, (4.13) is true. Putting (4.12) into (3.6) leads to (3.12), as claimed.

We next prove that (A) implies (B). Unlike the countable case, this is in fact the easier direction, mainly due to our assumption in (A) of the existence of those limits, the proof of whose existence from (B) has taken up so much of this section. In the countable case, Seneta and Vere-Jones [6] were able to use standard Abelian arguments to establish many of these limits, but here we have had to resort to the specific R-invariance of Q and f, and the semigroup nature of $\{P^n(\cdot,\cdot)\}$, to do this. On the other hand, we have established, even in the countable case, the ergodic nature of (2.12)–(2.14) for all $A \in \mathscr{T}$, not merely for $A \in \mathscr{T}_{\mathscr{X}}$ for some partition $\mathscr{X}$.

When we are given (A) of Theorem 2, we have in fact a plethora of limits, and we need only some of these, as noted in [6]. Specifically, we choose a particular $\zeta \in T$ which satisfies:

A(i) $\lim_{n\to\infty} a(\zeta, A, n) = \alpha(A)$ exists for each $A \in \mathscr{T}$, and $\alpha(\cdot)$ is a proper probability measure on $\mathscr{T}$;

A(ii) $\beta(y) = \liminf_{n\to\infty} b(\zeta, y, n)$ is a measurable non-negative function on T with $M\{y: \beta(y) > 0\} > 0$, and for all y in $T\setminus N$, where $M(N) = 0$,

$$\beta(y) = \lim_{n\to\infty} b(\zeta, y, n);$$

A(iii) α and β in A(i) and A(ii) satisfy

$$\int_T [\beta(y)/\eta(y)]\alpha(dy) < \infty. \tag{4.14}$$

That ζ can be found to satisfy A(i) and A(ii) is clear from the assumption of the existence of limits for (3.5) and (3.8); A(iii) follows on writing the double limit (3.7) as

$$\frac{\int_A P^m(\zeta, dy)\,[\int_T P^n(y, dw)\eta(w) / \int_T P^n(\zeta, dw)\eta(w)] / \int_T P^m(\zeta, dw)\eta(w)}{\int_T P^m(\zeta, dy)\,[\int_T P^n(y, dw)\eta(w) / \int_T P^n(\zeta, dw)\eta(w)] / \int_T P^m(\zeta, dw)\eta(w)};$$

for A on which $\beta(y)$ is approached uniformly and on which β/η is bounded (from A(ii) and Egorov's Theorem, $T\setminus N'$ can be partitioned into such A, where $M(N') = 0$): the numerator tends with n and then m to the finite limit $\int_A \alpha(dy)\beta(y)/\eta(y)$ and hence if $s(\zeta, \cdot, m, n)$ is to approach a probability measure, the denominator must also remain finite. From Fatou's lemma, this denominator is greater than (4.14), and so A(iii) holds.

The proof now follows that in [6]. We have, firstly,

$$\begin{aligned} &\int_T a(\zeta, dy, n)\left[\int_A P(y, dw)\eta(w)/\eta(y)\right] \\ &\quad = \int_A P^{n+1}(\zeta, dw)\eta(w) \Big/ \int_T P^n(\zeta, dw)\eta(w) \\ &\quad = a(\zeta, A, n+1) \int_T P^{n+1}(\zeta, dw)\eta(w) \Big/ \int_T P^n(\zeta, dw)\eta(w). \end{aligned} \tag{4.15}$$

Write

$$(4.16)\qquad \rho(\zeta, n) = \int_T P^{n+1}(\zeta, dw)\eta(w) \Big/ \int_T P^n(\zeta, dw)\eta(w).$$

Since $\eta(y)$ is the probability of ultimate absorption starting from y,

$$\eta(y) = 1 - P(y, T) + \int_T P(y, dw)\eta(w),$$

so for any $A \in \mathcal{T}$ and all y, $\int_A P(y, dw)\eta(w)/\eta(y) \leqq 1$; by Assumption A($i$), the left-hand side of (4.15) therefore converges, for every $A \in \mathcal{T}$, to

$$(4.17)\qquad \int_T [\alpha(dy)/\eta(y)] \int_A P(y, dw)\eta(w).$$

The right-hand side of (4.15) is $a(\zeta, A, n+1)\rho(\zeta, n)$: the first of these factors converges to $\alpha(A)$, from A(i), and so

$$(4.18)\qquad \rho = \lim_{n\to\infty} \rho(\zeta, n)$$

exists, and from (4.17) satisfies, for all $A \in \mathcal{T}$,

$$(4.19)\qquad \alpha(A)\rho = \int_T [\alpha(dy)/\eta(y)] \int_A P(y, dw)\eta(w).$$

Clearly from (4.19), $\rho > 0$. We can rewrite (4.19) as

$$(4.20)\qquad \rho \int_A \alpha(dw)/\eta(w) = \int_T [\alpha(dy)/\eta(y)]P(y, A), \qquad A \in \mathcal{T};$$

that is, $H(A) = \int_A \alpha(dw)/\eta(w)$, $A \in \mathcal{T}$ is a ρ^{-1}-invariant measure for $\{X_n\}$ (H is σ-finite because α is a probability measure and η is nowhere zero).

Secondly, we have for all $x \in T$

$$(4.21)\qquad \begin{aligned} \int_T P(x, dy)b(\zeta, y, n) &= \int_T P^{n+1}(x, dw)\eta(w) \Big/ \int_T P^r(\zeta, dw)\eta(w) \\ &= b(\zeta, x, n+1)\rho(\zeta, n) \end{aligned}$$

where $\rho(\zeta, n)$ is defined by (4.16). Applying Fatou's Lemma and A(ii) to (4.21) gives, for all $x \in T$,

$$(4.22)\qquad \rho\beta(x) \geqq \int_T P(x, dy)\beta(y).$$

From (4.20), (4.22) and Assumption A(iii), it is clear that $H(A) = \int_A \alpha(dw)/\eta(w)$ and $g = \beta$ satisfy the conditions of Lemma 1. Thus we conclude immediately that:

(i) $\{X_n\}$ is R-positive, with $R = \rho^{-1}$;

(ii) $\beta = f$ almost everywhere and $H = Q$, where f and Q are the unique R-invariant function and measure for $\{X_n\}$.

The assumption that α is a probability measure thus implies that $\int_T Q(dw)\eta(w) = \alpha(T) = 1$. The radius of convergence R must be strictly greater than unity, for a 1-positive chain must be stochastic, and we have assumed that $\{X_n\}$ on $(T, \mathscr{T})$ is strictly substochastic.

Now let $N' = \{\zeta : \text{A(i) fails to hold}\}$; by assumption $M(N') = 0$. For all $x \notin N'$, we have

$$\alpha_x(A) = \lim_{n\to\infty} \frac{R^n \int_A P^n(x, dw)\eta(w)}{R^n \int_T P^n(x, dw)\eta(w)}, \qquad A \in \mathscr{T} \tag{4.23}$$

exists, and α_x is a probability measure on $\mathscr{T}$. The numerator in (4.23) tends to $f(x)\int_A Q(dw)\eta(w)$ for all $A \in \mathscr{T}^+_{\mathscr{K}}$ and $x \notin N_f$, where $\mathscr{K}$ is a partition of T as in Theorem 1, since $\{X_n\}$ is R-positive; thus the denominator in (4.23) must tend to a finite limit $\pi(x, T)$ for all $x \notin N_1 = N_f \bigcup N'$, and (3.10) is true.

Finally, the probability of remaining in T for n steps or more, conditional on final absorption into $\mathscr{X} \setminus T$, is

$$\int_T P^n(x,dw)\eta(w)\,/\eta(x);$$

and when (3.10) holds, this clearly goes to zero geometrically as R^{-n}, when $n \to \infty$, for all $x \notin N_1$.

5. Arbitrary initial distributions

As remarked in [6], Section 4, the results of Theorem 2 do not carry over completely to arbitrary initial distributions. For example, consider the natural analogue of (3.5)

$$\alpha(\pi, A, n) = \int_T \pi(dx)\int_A P^n(x, dw)\eta(w) \Big/ \int_T \pi(dx)\int_T P^n(x, dw)\eta(w) \tag{5.1}$$

where η is as before and π is a probability measure on T. We leave to the reader the proof of the following general analogue of [6], Theorem 4.1: *if $\alpha_\pi(\cdot) = \lim_{n\to\infty}\alpha(\pi, \cdot, n)$ exists and is an honest probability measure on T, then $Q_\pi(A) = \int_A \alpha_\pi(dy)/\eta(y)$ is ρ-invariant for some $\rho \leqq R$; but if Q_ρ is any ρ-invariant measure with $\rho \leqq R$ and $Q_\rho(T) = 1$, then*

$$\alpha_\rho(A) = \int_A Q_\rho(dy)\eta(y) \Big/ \int_T Q_\rho(dy)\eta(y)$$

is obtained as the limit of $\alpha(Q_\rho, A, n)$.

The situation envisaged in this result can occur: thus R-positivity is essentially a condition equivalent to *all* of (3.11)–(3.13) holding. In the countable case this is shown by the semi-infinite random walk considered in [6], Section 6, whilst the behaviour may also be exhibited on a continuous state space: for example, the skeletons of the (continuous-time) conditioned waiting time process of an $M/G/1$ queue considered by Kyprianou in [3] are not R-positive, but have quasi-stationary limits $\alpha_\pi(\cdot)$ as defined above (for $\pi(\cdot) = \delta(y, \cdot)$, $y \in T$) which correspond to $\rho = R$.

Our final result, for which we shall only sketch a proof, is the analogue of [6], Theorem 4.2.

Theorem 3. *Suppose that $\{X_n\}$ on $(T, \mathscr{T})$ is R-positive, and π is an initial probability distribution for X_0 on $\mathscr{T}$. Sufficient conditions to ensure the validity of the analogues of* (3.11)–(3.13) *are either*:

(A′) *the R-invariant measure Q satisfies* (3.9), *and the initial distribution $\pi(A)$ is bounded by some multiple of $\int_A Q(dw)\eta(w)$*; or

(B′) *the R-invariant function f satisfies $\int_T f(y)\pi(dy) < \infty$, and $f(y)/\eta(y)$ is bounded away from zero.*

Proof. If (A′) holds we can appeal to Lemma 2 to show that

$$\lim_{n\to\infty} R^n \int_T \pi(dw) \int_T P^n(w, dy)\eta(y) \tag{5.2}$$

exists and has the value

$$\left[\int_T \pi(dw) f(w)\right]\left[\int_T Q(dy)\eta(y)\right] \underset{<\infty}{\leqq} \kappa \int_T Q(dw) f(w) \int_T Q(dy)\eta(y)$$

by R-positivity and (3.9).

If (B′) holds we can appeal to Theorem 1(d) to show that (5.2) tends to

$$\left[\int_T \pi(dw) f(w)\right]\left[\int_T Q(dy)\eta(y)\right]$$

which is again finite, the first factor by assumption, the second because $\int_T Q(dy) f(y) < \infty$ and $f(y)/\eta(y)$ is bounded away from zero.

The analysis then proceeds as in the preceding section.

We have given this theorem to show that in the uncountable case, an initial starting distribution may prove more tractable when not concentrated at a single point: the reader should note that in the theorem, under either (A′) or (B′), we do not need the assumption that the limits analogous to (3.10) exist. I conjecture that in fact, even in the uncountable case, the limits (3.10) always do exist, and so, from Corollary 2 to Lemma 2, have the desired form: but I have not so far been able to prove this.

In the uncountable case, one might often be willing to assume that the chain starts, not at a single point, but rather with a distribution over some set of positive M-measure. The theorem then ensures that for a large class of initial distributions, the 'correct' quasi-stationary limit is obtained, and is independent of the initial distribution.

In particular this theorem covers the following two cases when the initial distribution is actually concentrated at a point $\{x\}$:

(i) T is countable and $\{X_n\}$ is irreducible in the usual sense. Here $M(\{x\}) > 0$, and so $0 < Q(\{x\}) < \infty$, so that (A′) holds for $\pi(\cdot) = \delta(x, \cdot)$.

(ii) For some $m > 0$, $P^m(x, \cdot) \ll Q(\cdot)$ and the density $p^m(x, y)$ of $P^m(x, \cdot)$ with respect to Q is a bounded function of y. Here the 'initial distribution' $P^m(x, \cdot)$ satisfies (A′), and we may use

$$R^n P^n(x, T) = R^{n-m} \int_T [R^m P^m(x, dy)] P^{n-m}(y, T)$$

to see that $\lim_{n \to \infty} R^n P^n(x, T)$ exists, and so (3.10) need not be assumed.

This second case, and variations of this approach, may well show the way to checking (3.10) in particular cases. However, we conclude with an example to show that the bounded density of this case is by no means necessary to ensure the existence of the limit (3.10).

Example. (This is an adaptation of Example 4 in [5], and details of the working may be found there.)

Let $\{Y_n\}$ be a chain on T such that (B) of Theorem 2 holds for $\{Y_n\}$, and assume that Q_Y, the R_Y-invariant measure for $\{Y_n\}$, allots zero measure to points in T. Define $\{X_n\}$ by

$$\Pr\{X_n \in A \mid X_0 = x\} = \alpha \Pr\{Y_n \in A \mid Y_0 = x\} + \beta\, \delta(x, A), \quad \alpha, \beta > 0, \quad \alpha + \beta = 1.$$

Then it is shown in [5] that $\{X_n\}$ is R_X-positive, with $R_X = R_Y/(\alpha + \beta R_Y)$, that the R_X-invariant measure and function for $\{X_n\}$ are the R_Y-invariant measure and function for $\{Y_n\}$, and that

$$\lim_{n \to \infty} R_X^n \Pr\{X_n \in A \mid X_0 = x\} = \lim_{n \to \infty} R_Y^n \Pr\{Y_n \in A \mid Y_0 = x\}$$

for every $A \in \mathscr{T}$ such that the right-hand limit exists. Thus, since $\{Y_n\}$ satisfies (B) of Theorem 2, so does $\{X_n\}$; and in particular, (3.10) holds for $\{X_n\}$.

However, by construction, for every m

$$\Pr\{X_m = x \mid X_0 = x\} = \beta^m > 0,$$

and $\{X_n\}$ does not satisfy Case (ii) above.

References

[1] DARROCH, J. N. AND SENETA, E. (1965) On quasi-stationary distributions in absorbing discrete finite Markov chains. *J. Appl. Prob.* **2**, 88–100.

[2] GÄNSSLER, P. (1971) Compactness and sequential compactness in spaces of measures. *Z. Wahrscheinlichkeitsth.* **17**, 124–146.

[3] KYPRIANOU, E. K. (1971) On the quasi-stationary distribution of the virtual waiting time in queues with Poisson arrivals. *J. Appl. Prob.* **8**, 494–507.

[4] OREY, S. (1971) *Lecture Notes on Limit Theorems for Markov Chain Transition Probabilities.* van Nostrand, London.

[5] POLLARD, D. AND TWEEDIE, R. L. *R*-theory for Markov chains on a topological space. Submitted for publication.

[6] SENETA, E. AND VERE-JONES, D. (1966) On quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states. *J. Appl. Prob.* **3**, 403–434.

[7] TWEEDIE, R. L. (1974) *R*-theory for Markov chains on a general state space I: solidarity properties and *R*-recurrent chains. *Ann. Probability* **2**. To appear.

[8] VERE-JONES, D. (1962) Geometric ergodicity in denumerable Markov chains. *Quart. J. Math. Oxford* (*2nd ser.*) **13**, 7–28.

[9] VERE-JONES, D. (1967) Ergodic properties of non-negative matrices. *Pacific J. Math.* **22**, 361–386.