Stochastic Processes and their Applications 84 (1999) 177–186

stochastic processes and their applications

www.elsevier.com/locate/spa

A quasi-ergodic theorem for evanescent processes

L.A. Breyer *, G.O. Roberts

Department of Mathematics and Statistics, Lancaster University, Lancaster, LA1 4YF, UK

Received 25 August 1997; accepted 19 February 1999

Abstract

We prove a conditioned version of the ergodic theorem for Markov processes, which we call a quasi-ergodic theorem. We also prove a convergence result for conditioned processes as the conditioning event becomes rarer. © 1999 Published by Elsevier Science B.V. All rights reserved.

MSC: primary 60J; secondary 60F

Keywords: Ergodic theorems; Markov processes; Quasi-stationary distributions

1. Introduction

Consider a Markov process X evolving on its state space E, and which is killed at some a.s. finite random time τ. The first result of this paper states that, under suitable conditions on the process (essentially positive λ-recurrence), the following quasi-ergodic limit theorem holds:

$$\lim_{t\to\infty} \mathbb{E}_x\left[\frac{1}{t}\int_0^t g(X_s)\,\mathrm{d}s \,\middle|\, \tau > t\right] = \int_E g\,\mathrm{d}m, \quad g \in L^1(\mathrm{d}m),\ x \in E$$

for some probability measure m on E.

The measure m is the stationary distribution for a Markov process Y, which can be interpreted as X, conditioned such that a.s. $\tau = \infty$. In terms of Y, the main assumption we shall make on X is that Y be positive Harris recurrent.

One motivation for proving this result comes from Markov chain Monte Carlo techniques (see, for instance, Smith and Roberts, 1993). The basic idea is to simulate a positive recurrent Markov chain in order to estimate properties of its stationary measure by considering suitable ergodic averages along sample paths. In practice, it is frequently the case that the actual distribution which is ultimately estimated is the stationary distribution of a conditioned chain. As an example of this consider the practice

* Corresponding author.

E-mail addresses: l.breyer@lancaster.ac.uk (L.A. Breyer), G.O.roberts@lancaster.ac.uk (G.O. Roberts)

PII: S0304-4149(99)00018-6

of simulating a Markov chain on a computer. The simulation run is terminated and restarted if numerical overflows are achieved. Thus, in effect, successful runs of the Markov chain are conditioned to not achieve these numerical problems.

In this way it is often the case that, when computing the invariant measure π using time averages, one actually computes the measure m above instead. It is therefore of interest to know how far apart m and π really are. Our second main result demonstrates that under mild assumptions, when the conditioning event is sufficiently rare, π is well approximated by m.

Of course in general, it may be, in fact, that m exists while π does not, for instance if X is Brownian motion in more than three dimensions, and τ is the first exit time from a compact set. However in this paper (guided by the Markov chain Monte Carlo motivation) we shall be assuming that all the relative positive recurrence properties hold.

2. Positive λ-recurrence

We describe here the main assumption which ensures the validity of the quasi-ergodic theorem. The facts laid out here shall be used in the proof of the theorem, given in the next section.

Let X be a strong Markov process evolving on a state space E, endowed with some countably generated σ-algebra $\mathscr{E}$. Typically, E is locally compact, and $\mathscr{E}$ denotes the Borel σ-algebra. The notation we use is as follows: X_t denotes the sample path, defined as usual on the canonical space of all right continuous trajectories with left limits. The lifetime of X is the stopping time

$$\tau = \inf\{t > 0\colon X_t \notin E\}$$

and we denote by $\mathbb{P}_x$ the unique law of the process satisfying $\mathbb{P}_x[\lim_{t\downarrow 0} X_t = x] = 1$.

Let f be a positive (i.e. nonnegative) measurable function (all the real-valued functions in this paper will be assumed $\mathscr{E}$-measurable). Given a real number λ, we say that f is *λ-invariant* provided that the equation

$$\mathbb{P}_x(f(X_t), \tau > t) \overset{\text{def}}{=} \int 1_{\{\tau > t\}} f(X_t)\, \mathrm{d}\mathbb{P}_x = \mathrm{e}^{\lambda t} f(x)$$

holds for each $t > 0$. The existence of such a function can often be reduced to a search for a positive solution to the equation $\mathfrak{A} f = \lambda f$, where $\mathfrak{A}$ is a suitable generator for X. Dually, one can look for measures μ which satisfy an equation $\int (\mathfrak{A} h)\, \mathrm{d}\mu = \lambda \int h\, \mathrm{d}\mu$ for all test functions h. This equation is typically satisfied by a *λ-invariant measure*, that is a (σ-finite) measure μ with the property that, for all positive functions h and $t > 0$,

$$\mathbb{P}_\mu(h(X_t), \tau > t) \overset{\text{def}}{=} \int \mu(\mathrm{d}x) \mathbb{P}_x(h(X_t), \tau > t) = \mathrm{e}^{\lambda t} \int h(x) \mu(\mathrm{d}x).$$

When μ is a probability measure, it is often interpreted as a quasi-stationary distribution. This is defined as an initial distribution with the property that X is, when started according to μ, stationary given τ has not occurred:

$$\mathbb{P}_\mu(f(X_t) \mid \tau > t) = \int f \, \mathrm{d}\mu.$$

It was shown by Nair and Pollett (1993) that this last equation is equivalent to λ-invariance (with $\lambda \leqslant 0$), and then $\mathbb{P}_\mu(\tau > t) = \mathrm{e}^{\lambda t}$. Thus under $\mathbb{P}_\mu$, τ is independent of X. Many Markov chain models used in biology seem to settle down to a quasi-stationary distribution after a short time, even though on longer time scales transience (typically associated with extinction of the population) is exhibited.

Every λ-invariant function can be used to construct a new law $\mathbb{Q}$ for X, under which the process is again a strong Markov process, with state space $E_f = \{0 < f < \infty\}$. The probability measure $\mathbb{Q}$ is characterized on path space by the formula

$$f(x)\mathbb{Q}_x(F, \tau > t) = \mathbb{P}_x(F, \mathrm{e}^{-\lambda t} f(X_t), \tau > t),$$

which is valid for all positive $\mathscr{F}_t$ measurable random variables F, $t > 0$ and $x \in E_f$. The λ-invariance implies that $\mathbb{Q}_x[\tau = \infty] = 1$, and X never leaves E_f in a finite time if it begins its trajectory there.

Our main assumption is the following:

Positive (Harris) λ-recurrence

For some $\lambda \leqslant 0$, there exists a λ-invariant function f such that X is, under the probability measure $\mathbb{Q}$, positive Harris recurrent.

We remind the reader that, under the original probability law $\mathbb{P}$, the given process may well be transient. Indeed, if $\lambda < 0$, it can be shown that $\mathbb{P}_x(\tau < \infty) = 1$ for all $x \in E_f$.

When X is positive λ-recurrent, the state space E_f must necessarily be irreducible under $\mathbb{P}$. The stationary distribution m of X under $\mathbb{Q}$ has the property that

$$m(A) > 0 \;\Rightarrow\; \mathbb{Q}_x \left[\int_0^\infty 1_A(X_t) \, \mathrm{d}t = \infty \right] = 1, \quad x \in E_f$$

and so $m(A) > 0$ implies that

$$\mathbb{P}_x \int_0^\infty 1_A(X_t) \, \mathrm{d}t = f(x)\mathbb{Q}_x \int_0^\infty \mathrm{e}^{\lambda t} 1_A(X_t) f(X_t)^{-1} \, \mathrm{d}t > 0$$

for all $x \in E_f$. This means that, under $\mathbb{P}$, m satisfies the definition of an irreducibility measure for X (see Meyn and Tweedie, 1993).

Now define a measure μ on E_f by the prescription $\int g \, \mathrm{d}\mu = \int (g/f) \, \mathrm{d}m$. A simple calculation using the stationarity of m under $\mathbb{Q}$ shows that μ so defined is λ-invariant under $\mathbb{P}$. Conversely, suppose that X has, under $\mathbb{P}$, an irreducibility measure on E. If, for some $\lambda \leqslant 0$, there exists a nontrivial λ-invariant measure μ and a strictly positive λ-invariant function f, then, according to a test of Tweedie (see Tuominen and Tweedie, 1979) it suffices that $0 < \int f \, \mathrm{d}\mu < \infty$ for the process to be positive λ-recurrent. In that case, the measure $m(\mathrm{d}x) = f(x)\mu(\mathrm{d}x)$ is the unique stationary distribution for X under $\mathbb{Q}$.

Tweedie's test allows us to easily identify examples of positive λ-recurrent processes. If X is a Markov chain on a finite and irreducible state space, it is λ-recurrent, for the Perron–Frobenius theorem guarantees the existence of a pair of positive eigenvectors for the generator matrix and its transpose, with common eigenvalue λ. These vectors, which we denote by f and μ, respectively, obviously satisfy $\int f\,\mathrm{d}\mu = \sum_i f_i\mu_i < \infty$.

Similarly, suppose that X is a uniformly elliptic diffusion on a bounded domain with smooth boundary. It is well known that there exists a pair of positive continuous functions φ, φ^* which vanish on the boundary and are, respectively, eigenfunctions for the generator of X and its adjoint. Taking $\mu(\mathrm{d}x) = \varphi^*(x)\,\mathrm{d}x$ and $f(x) = \varphi(x)$, one has $\int f\,\mathrm{d}\mu < \infty$, and again the process turns out positive λ-recurrent.

The asymptotic behaviour of $t \mapsto \mathbb{P}_x(\tau > t)$ is known, provided μ is finite (see Tuominen and Tweedie, 1979). To state it, we shall use the following normalizations: $\mu(E) = 1$, and f satisfies $\langle \mu, f\rangle = 1$, where $\langle \mu, f\rangle = \int f\,\mathrm{d}\mu$. Then,

$$\lim_{t\to\infty} \mathrm{e}^{-\lambda t}\mathbb{P}_x(\tau > t) = \lim_{t\to\infty} f(x)\mathbb{Q}_x(f(X_t)^{-1}) = f(x) \tag{1}$$

for every $x \in E_f$, by Harris recurrence under $\mathbb{Q}$ (see Revuz, 1979; Meyn and Tweedie, 1993).

Jacka and Roberts (1995), see also Breyer (1997), have given an interpretation of this result in terms of a conditioned process as follows: Fix a starting position x and a time interval $[0,t]$. Their result asserts that the laws $\mathbb{P}_x(\cdot\,|\,\tau > T)$ converge weakly to $\mathbb{Q}_x$ as T tends to infinity, on the space of paths with time interval $[0,t]$. Thus if we wait a long time T, then given that X is still alive, its law up to time t is well approximated by $\mathbb{Q}$. Their motivation was to understand so-called quasi-stationary limit theorems, which typically give rise to λ-invariant measures. For example, a positive λ-recurrent process typically exhibits the feature

$$\lim_{t\to\infty} \mathbb{P}_x(f(X_t)\,|\,\tau > t) = \int f\,\mathrm{d}\mu.$$

In this context, the probability measure μ is known as a quasi-stationary distribution for X. The λ-recurrence of X is, however, *not* necessary for the existence of the above limit. Theorems of this type are useful in modelling the persistence of stochastic models (for instance, epidemics and branching processes).

In view of its close connection with quasi-stationarity, we refer to the theorem below as a quasi-ergodic theorem.

3. Quasi-Ergodic theorem

Theorem 1. *Let X be irreducible and positive (Harris) λ-recurrent, with associated λ-invariant function f and measure μ. If $\mu(E_f) < \infty$, then for every bounded measurable function g and every $x \in E_f$,*

$$\lim_{t\to\infty} \mathbb{P}_x\left[\frac{1}{t}\int_0^t g(X_s)\,\mathrm{d}s\,\middle|\,\tau > t\right] = \int g\,\mathrm{d}m, \tag{2}$$

where $m(\mathrm{d}x) = f(x)\mu(\mathrm{d}x)/\langle f, \mu\rangle$ on E_f, and $\langle f, \mu\rangle = \int f\,\mathrm{d}\mu$.

Proof. Assume first that g is bounded and positive. For fixed u, put

$$h_u(x) = \inf\{e^{-\lambda r}\mathbb{P}_x(\tau > r)/f(x)\colon r \geqslant u\}$$

and note that, by (1), $h_u(x)\uparrow 1$, as $u \to \infty$. By definition of the law $\mathbb{Q}$, we also have

$$\frac{e^{-\lambda t}\mathbb{P}_x(\tau>t)}{f(x)}\mathbb{P}_x\left[\frac{1}{t}\int_0^t g(X_s)\,ds\,\middle|\,\tau > t\right] = \frac{1}{t}\int_0^t \mathbb{Q}_x\left[g(X_s)\mathbb{P}_{X_s}(\tau > t-s)\frac{e^{-\lambda(t-s)}}{f(X_s)}\right]ds$$

$$\geqslant \frac{1}{t}\int_0^{t-u} \mathbb{Q}[g(X_s)h_u(X_s)]\,ds.$$

The function gh_u belongs to $L^1(dm)$, for

$$\int g(y)h_u(y)m(dy) \leqslant \int g(y)e^{\lambda u}\mathbb{P}_y(\tau > u)\mu(dy) \leqslant ||g||_\infty$$

and in view of the positive Harris recurrence of X under $\mathbb{Q}$, Fatou's lemma and the above calculations imply that

$$\liminf_{t\to\infty} \mathbb{P}_x\left[\frac{1}{t}\int_0^t g(X_s)\,ds\,\middle|\,\tau > t\right] \geqslant \int g(y)h_u(y)m(dy)$$

$$\uparrow \int g(y)m(dy) \quad \text{as } u \to \infty.$$

The last assertion follows from monotone convergence. Since g is bounded, we can repeat the argument, replacing g by $||g||_\infty - g$, which gives

$$\limsup_{t\to\infty} \frac{1}{t}\mathbb{P}_x\left[\int_0^t g(X_s)\,ds\,|\tau > t\right] \leqslant \int g(y)m(dy).$$

Combining these last two steps gives result (2) when g is bounded, positive, and then for arbitrary bounded g by subtraction. □

The above theorem may be viewed as a generalization of the standard ergodic theorem for positive Harris recurrent Markov processes. Indeed, if $\lambda = 0$, it can be shown that $\tau = \infty$ a.s., and consequently (2) reduces to the well-known result

$$\lim_{t\to\infty} \frac{1}{t}\mathbb{P}_x \int_0^t g(X_s)\,ds = \int g\,dm.$$

A related theorem of ergodic theory in this context states that

$$\mathbb{P}_x\left[\lim_{t\to\infty} \frac{1}{t}\int_0^t g(X_s)\,ds = \langle g, m\rangle\right] = 1. \tag{3}$$

An obvious generalization of this involving the conditioning event $\{\tau > t\}$ is not possible when $\lambda < 0$, for we are discarding eventually every single sample path save for a null set. A natural attempt to give an almost sure interpretation of this result by a branching Markov process also fails as the following example illustrates.

Example. Consider a subcritical continuous-time branching process X on $E = \{1, 2, 3, \ldots\}$ with upward rates $\gamma_i = i\gamma$ and downward rates $\delta_i = \delta i$ ($\gamma < \delta$, $i \geqslant 1$). It is well known (see, for example, Asmussen and Hering, 1983) that X is λ-positive recurrent with $\lambda = \gamma - \delta$, and if τ denotes the extinction time of the branching process,

$$f(x) = \lim_{n\to\infty} \mathbb{P}_x(\tau > t)\mathrm{e}^{-\lambda t} = (1 - \gamma/\delta)^2 x, \quad x \geqslant 1.$$

Construct a branching particle system as follows. Starting in $x \in E$, take a particle X_t and let it evolve according to the law $\mathbb{P}_x$. Let e denote an independent, exponentially distributed random variable with parameter $|\lambda|$. If the particle has not died before time e, replace it with two identical ones, each evolving independently according to the law $\mathbb{P}_{X_e}$. Repeating the previous steps on each of those ad infinitum produces a branching Markov process with rate $|\lambda|$. For each subset A of E, let now $Z_t(A)$ represent the number of particles in A at time t. This induces a random measure $Z_t(\mathrm{d}y)$, in terms of which (see Breyer, 1997)

$$\lim_{t\to\infty} \mathbb{P}_x \left[\frac{1}{t} \int_0^t \langle Z_t, g \rangle \, \mathrm{d}t \right] = \int g \, \mathrm{d}m.$$

However, we also have

$$\mathbb{P}_x \left[\lim_{t\to\infty} \frac{1}{t} \int_0^t \langle Z_t, g \rangle \, \mathrm{d}t = 0 \right] = 1.$$

To see this, let $B_t = (X_t^{(1)}, \ldots, X_t^{(N_t)})$ denote the current state of the branching process, where $N_t = Z_t(E)$, and introduce the Lyapunov function $V(B_t) = \sum_{i=1}^{N_t} X_t^{(i)}$. It is easy to check that $V(B)$ is a nonnegative martingale which converges to 0 almost surely.

4. Convergence of conditional invariant measures

For Markov Chain Monte Carlo applications, we wish to apply Theorem 1 when

$$\tau = \tau_n = \inf\{t > 0\colon X_t \notin E_n\}$$

for some increasing sequence of subsets E_n of E, such that $E_n \uparrow E$.

As we have seen, the result of the limiting operation (2) is a probability measure m_n carried by E_n. Suppose now that X is positive Harris recurrent on the state space E, with invariant distribution π. Clearly $m_n \neq \pi$, since m_n is zero outside E_n, but is it true that $m_n \Rightarrow \pi$ as $n \to \infty$?

The theorem below answers this question in the affirmative. This follows from a straightforward application of Martin boundary theory. For simplicity, we shall make the following assumption on the Markov process X.

Absolute continuity There exists, relative to some excessive measure ξ, a jointly continuous density $p_t(x, y)$ for the transition function, i.e.

$$\mathbb{P}_x(g(X_t), \tau > t) = \int p_t(x, y) g(y) \xi(\mathrm{d}y), \quad g \geqslant 0.$$

This assumption can be dispensed with completely, but the increase in technicality is not worth pursuing here. We refer the reader to Jeulin (1977).

Theorem 2. *Let X be positive Harris recurrent on E with stationary distribution π, and let $E_n \uparrow E$ be an increasing sequence of subsets, with $\tau_n = \inf\{t > 0: X_t \notin E_n\}$.*

Let X^n denote the Markov process with state space E_n obtained by killing X at time τ_n, and suppose that X^n is positive (Harris) λ_n-recurrent for each n, and that a.s. $\tau_n \uparrow \infty$. Let μ_n be the λ_n-invariant measure for X^n, assumed to satisfy $\mu_n(E_n) = 1$. Let f_n be the corresponding λ_n-invariant function, satisfying $\langle f_n, \mu_n\rangle = m_n(E_n) = 1$, where $m_n(\mathrm{d}x) = f_n(x)\mu_n(\mathrm{d}x)$. Then the following statements hold:

(i) *For each $x, y \in E$, $\lim_{n\to\infty} f_n(y)/f_n(x) = 1$.*
(ii) *For any nonzero bounded measurable positive function g, $\lim_{n\to\infty}\langle g, \mu_n\rangle = \langle g, \pi\rangle$.*
(iii) *Let $\mathbb{Q}^n$ denote the law of the (X^n, f_n)-conditioned process, then for each $x \in E$, $t > 0$, $\lim_{n\to\infty} \mathbb{Q}^n_x(\cdot|\mathscr{F}_t) = \mathbb{P}_x(\cdot|\mathscr{F}_t)$.*
(iv) *For any bounded measurable positive function g, $\lim_{n\to\infty}\langle g, m_n\rangle = \langle g, \pi\rangle$.*

Before presenting the proof, we shall describe some of the notation to be used. When $\lambda \leqslant 0$, a λ-invariant function f is *excessive*: $\lim_{t\to 0} P_t f \uparrow = f$, where $P_t(x, \mathrm{d}y) = \mathbb{P}_x(X_t \in \mathrm{d}y, \tau > t)$. Similarly, a λ-invariant measure μ is also *excessive* when $\lambda \leqslant 0$: $\lim_{t\to 0} \mu P_t \uparrow = \mu$.

The following results may be found in Meyer (1968). Let $G(x, y) = \int_0^\infty p_t(x, y)\,\mathrm{d}t$ denote the Green's function of X; given a measure v such that the function $y \mapsto \int v(\mathrm{d}z)G(z, y)$ is continuous into $[0, +\infty]$, there exists a metrizable compactification F of E, unique up to homeomorphism, such that the function

$$y \mapsto K(x, y) \stackrel{\text{def}}{=} G(x, y) \bigg/ \int v(\mathrm{d}z)G(z, y)$$

has a continuous extension to F for each $x \in E$ (this is also denoted $K(x, y)$). If h is an excessive function satisfying $\langle h, v\rangle = 1$, there exists a probability measure θ on F representing it: $h = \int K(\cdot, y)\theta(\mathrm{d}y)$.

If μ is an excessive measure, a similar representation exists. This is shown by exploiting the backwards Markov process $\hat{X}$ with transition function $\hat{P}(x, \mathrm{d}y) = p_t(y, x)\xi(\mathrm{d}y)$ (note that the excessivity of ξ guaranteed that this is a transition function. If μ is excessive for X, it must be absolutely continuous with respect to ξ, and its Radon–Nikodym density $\mathrm{d}\mu/\mathrm{d}\xi$ can be chosen excessive for $\hat{X}$. As above, the function $\mathrm{d}\mu/\mathrm{d}\xi$ then has an integral representation on $\hat{F}$, the Martin compactification of $\hat{X}$.

Proof of Theorem 2. (i) Fix $k > 0$, $x_0 \in E$, and consider for a moment the process X^k. We shall denote all Martin boundary concepts constructed from X^k by the superscript k. For each $n \geqslant k$, the function $\bar{f}_n(y) = f_n(y)/f_n(x_0)$ is excessive for X^k, since

$$\begin{aligned}\mathbb{P}_x(\bar{f}_n(X_t), \tau_k > t) &\leqslant \mathbb{P}_x(f_n(X_t), \tau_n > t)/f_n(x_0)\\ &= \mathrm{e}^{\lambda_n t} f_n(x)/f_n(x_0) \leqslant \bar{f}_n(x)\end{aligned}$$

and Fatou's lemma shows that $\lim_{t\to 0} \mathbb{P}_x(\bar{f}_n(X_t), \tau_k > t) \geqslant \bar{f}_n(x)$. Let θ^k_n be the representing probability on F^k (here the normalizing measure v is the point mass at x_0, assumed to belong to E_k). Since F^k is compact, it carries all weak limit points θ^k of

(θ_n^k) in the Martin topology. In view of the continuity of $K^k(x,\cdot)$, we then have, for some subsequence (n') and every $x \in E_k$,

$$\bar{f}(x) \stackrel{\text{def}}{=} \int K^k(x,y)\theta^k(\mathrm{d}y)$$

$$= \lim_{n'\to\infty} \int K^k(x,y)\theta_{n'}^k(\mathrm{d}y) = \lim_{n'\to\infty} \bar{f}_{n'}(x).$$

The convergence of $\bar{f}_{n'}$ to $\bar{f}$ does not involve the Martin topology, and thus cannot depend on k. The integral representation in terms of θ^k does depend on k however. Since the function $\bar{f}$ is excessive for each X^k, we have $\lim_{t\to 0} \mathbb{P}_x(\bar{f}(X_t), \tau_k > t)\uparrow = \bar{f}(x)$. This being true for each k independently, it follows that the function $\bar{f}$ is excessive for X also: interchanging (as we can always do) increasing limits, we have as required (since $\tau_k \uparrow \infty$)

$$\lim_{t\to 0} \mathbb{P}_x(\bar{f}(X_t)) = \lim_{t\to 0} \lim_{k\to\infty} \mathbb{P}_x(\bar{f}(X_t), \tau_k > t)$$

$$= \lim_{k\to\infty} \lim_{t\to 0} \mathbb{P}_x(\bar{f}(X_t), \tau_k > t) = \bar{f}(x).$$

Now X is positive Harris recurrent, and this is equivalent to having all excessive functions constant (see Getoor, 1980). Since $\bar{f}$ is such a function, and $\bar{f}(x_0) = 1$, it follows that $\bar{f} \equiv 1$, and this is clearly independent of the subsequence (n') chosen to define it.

(ii) The proof of this statement is essentially just that of (i) for $\hat{X}$. For fixed k, the measure μ_n $(n \geqslant k)$ is easily shown to be excessive, so that the function $h_n = \mathrm{d}\mu_n/\mathrm{d}\xi$ can be chosen excessive for $\hat{X}^k$. Moreover, we obviously have $\langle h_n, \xi\rangle = \mu_n(E_n) = 1$, and there exists an integral representation over the Martin compactification based on $\hat{K}(x,y) = G(y,x)/\langle \xi, G(y,\cdot)\rangle$ (note that $\int G(y,x)\xi(\mathrm{d}x) = \mathbb{P}_y(\tau_k) < \infty$). Let γ_n be the representing measure, i.e. $h_n = \int \hat{K}(\cdot,y)\gamma_n(\mathrm{d}y)$. Whenever $\gamma_{n'} \Rightarrow \gamma$ on $\hat{F}$, we also have $h_{n'}(x) \to h(x) \stackrel{\text{def}}{=} \int \hat{K}(x,y)\gamma(\mathrm{d}y)$, and we deduce that h is excessive for $\hat{X}^k$. Equivalently, $\mu(\mathrm{d}x) = h(x)\xi(\mathrm{d}x)$ is excessive for X^k, and since k is arbitrary, μ is excessive for X, thus a multiple of π. Now $\mu(E) = \langle h, \xi\rangle = 1$, since $\langle \xi, \hat{K}(\cdot,y)\rangle = 1$. Hence it follows that $\mu = \pi$, and the limit function h is independent of the subsequence (n'). By Fatou's lemma,

$$\liminf_{n\to\infty} \int h_n \,\mathrm{d}\xi \geqslant \int h \,\mathrm{d}\xi = 1$$

and since $\langle h_n, \xi\rangle = 1$, we have $\int h_n \,\mathrm{d}\xi \to \int h\,\mathrm{d}\xi$. By Scheffé's theorem, h_n converges to h in $L^1(\mathrm{d}\xi)$, which implies that $\mu_n \Rightarrow \pi$.

(iii) Fix $x \in E$, $t > 0$. Since each f_n is λ_n-invariant for X^n, (i) and Fatou's lemma immediately give

$$\lim_{n\to\infty} \mathbb{P}_x\left(\mathrm{e}^{-\lambda_n t}\frac{f_n(X_t)}{f_n(x)}, \tau_n > t\right) \geqslant 1,$$

which implies that the random variables $Z_t^n = \mathrm{e}^{-\lambda t} f_n(X_t)/f_n(x)\cdot 1_{(\tau_n > t)}$ converge strongly in $L^1(d\mathbb{P}_x)$ by Scheffé's theorem. If H is a bounded, $\mathscr{F}_t$ measurable random variable, it follows that

$$\lim_{n\to\infty} \int H \, \mathrm{d}\mathbb{Q}_x^n = \lim_{n\to\infty} \int HZ_t^n \, \mathrm{d}\mathbb{P}_x = \int H \, \mathrm{d}\mathbb{P}_x.$$

(iv) By (i) and (ii), we have $\lim_{n\to\infty}(\mathrm{d}\mu_n/\mathrm{d}\xi)(x)\bar{f}_n(x) = (\mathrm{d}\pi/\mathrm{d}\xi)(x)$; Scheffé's theorem implies that this convergence occurs in $L^1(\mathrm{d}\xi)$. If g is a bounded function, it follows that

$$\lim_{n\to\infty} \langle \mu_n, g\bar{f}_n\rangle = \lim_{n\to\infty} \int g(\mathrm{d}\mu_n/\mathrm{d}\xi)\bar{f}_n \, \mathrm{d}\xi = \int g \, \mathrm{d}\pi.$$

In particular, this holds for $g = 1$, and thus

$$\begin{aligned}\lim_{n\to\infty} \langle m_n, g\rangle &= \lim_{n\to\infty} \langle \mu_n, \bar{f}_n g\rangle / \langle \mu_n, \bar{f}_n\rangle \\ &= \langle \pi, g\rangle / \langle \pi, 1\rangle. \qquad \square\end{aligned}$$

Example. Suppose that E is the open interval $(0,1)$, and denote by X the Brownian motion on E, with reflection at the boundaries. It is easily seen that X is positive Harris recurrent, and therefore it has a unique stationary distribution. Suppose now that we approximate E by the sets $E_n = (1/n, 1-1/n)$. With the notation of Theorem 2, the invariant measure for the process X^n, conditioned on never leaving E_n can be calculated to be

$$m_n(\mathrm{d}x) = \cos\left(\frac{\pi(x-1/2)}{1-2/n}\right)^2 \mathrm{d}x \Big/ \int_0^1 \cos\left(\frac{\pi(y-1/2)}{1-2/n}\right)^2 \mathrm{d}y.$$

If we let $n\to\infty$, the measures m_n converge to

$$m(\mathrm{d}x) = \cos(\pi(x-1/2))^2 \, \mathrm{d}x \Big/ \int_0^1 \cos(\pi(y-1/2))^2 \, \mathrm{d}y,$$

which is certainly different from the stationary distribution of X. Thus the assertion of Theorem 2 fails here, due to the fact that

$$\tau_n \uparrow \inf\{t > 0\colon X_{t-} = 0 \text{ or } X_{t-} = 1\} < \infty.$$

Example. A small modification of the previous example described below also shows that sometimes, the measures m_n can converge to a measure m even though the original process is transient, *without a stationary distribution*: In the context of MCMC simulations on a computer, this means that it is possible to wrongly identify a transient process as recurrent. Thus it is important to take great care when implementing the algorithm. Let $E = (0,1)$ as above, with $E_n = (1/n, 1-1/n)$. Instead of taking X to be reflecting Brownian motion, we construct the process by killing ordinary Brownian motion on first exiting E. Unlike the previous example, we now have $X^n \Rightarrow X$ as $n\to\infty$; however, the process has a finite lifetime, and therefore no invariant measure. As above, $m_n \Rightarrow m$, but now m is simply the stationary distribution of X, conditioned on staying in E forever.

References

Asmussen, S., Hering, H., 1983. Branching Processes. Birkhäuser, Boston.

Breyer, L.A., 1997. Quasistationarity and conditioned Markov processes. Ph.D. Thesis, The University of Queensland.

Getoor, R.K., 1980. Transience and recurrence of Markov processes. Seminar on Probability, vol. XIV, Paris, 1978/1979. Lecture Notes in Mathematics, vol. 784, Springer, Berlin, pp. 397–409.

Jacka, S.D., Roberts, G.O., 1995. Weak convergence of conditioned processes on a countable state space. J. Appl. Probab. 32, 902–916.

Jeulin, T., 1977. Compactification de Martin d'un processus droit. Z. Wahrscheinlichkeitstheorie 42, 229–260.

Meyn, S.P., Tweedie, R.L., 1993. Markov Chains and Stochastic Stability. Springer, Berlin.

Meyer, P.A., 1968. Processus de Markov: la frontière de Martin, Lecture Notes in Mathematics, vol. 77. Springer, New York.

Nair, M.G., Pollett, P.K., 1993. On the relationship between μ-invariant measures and quasi-stationary distributions for continuous-time Markov chains. Adv. Appl. Probab. 25(1), 82–102.

Revuz, D., 1979. A survey of limit theorems for Markov chains and processes on general state spaces. Proceedings of the 42nd session of the International Statistical Institute, vol. 2, Manila, 1979. Bull. Inst. Int. Statist. 48(2), 203–210.

Smith, A.F.M., Roberts, G.O., 1993. Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods. J. Roy. Statist. Soc. Ser. B 55(1), 3–23.

Tuominen, P., Tweedie, R.L., 1979. Exponential decay and ergodicity of general Markov processes and their discrete skeletons. Adv. Appl. Probab. 11, 784–803.