YAGLOM LIMITS VIA COMPACTIFICATIONS

L.A. BREYER

Abstract. Consider a Markov process X with finite lifetime. In this paper, we derive sufficient conditions for the existence of a Yaglom limit, or limiting conditional distribution for X.

1. Introduction

The problem of *Yaglom limits* for Markov processes can be described as follows. Let X be a Markov process with locally compact metric state space E, $X_0 \sim \nu$, and suppose that the lifetime $\zeta = \inf\{s > 0 : X_s \notin E\}$ is a.s. finite. We assume that $X_{s+t} \notin E$, $t \geq 0$, which is a *minimality* requirement on X. Does there exist a probability measure κ on E such that

$$\lim_{t\to\infty} \mathbb{E}_\nu(f(X_t) \,|\, \zeta > t) = \int f d\kappa, \quad f \text{ bounded.} \tag{1}$$

Such a measure, if it exists, is variously known as a *quasistationary distribution* or *limiting conditional distribution*. In this paper, we use the same terminology as in (Kesten, 1995), and refer to κ as a *Yaglom limit*, in honour of A.M. Yaglom, who first showed the existence of such measures for branching processes (Yaglom, 1947).

A first generalization of his results was achieved by Seneta and Vere-Jones (1966), who showed that (1) holds as soon as X is an irreducible Markov chain with finite state space. Note that we can think of (1) as a generalization of the ergodic theory of Markov chains, which corresponds to the "limiting" case when $\zeta = \infty$.

These results were later refined into the theory of λ-recurrence (see Anderson, 1991, for a complete account), where $\lambda \geq 0$ is a real number such that

$$\mathbb{E}_\kappa[f(X_t)] = e^{-\lambda t}\langle \kappa, f\rangle,$$

Date: February, 1998.

1991 *Mathematics Subject Classification.* Primary 60J, Secondary 60F.

Key words and phrases. Markov chain , Quasistationarity, Martin Boundaries.

where $\langle \kappa, f \rangle$ denotes the integral of f with respect to κ. Incidentally, this may be stated without reference to λ as

$$\mathbb{E}_\kappa[f(X_t) \,|\, \zeta > t] = \langle \kappa, f \rangle,$$

which is a common definition of a *quasistationary distribution* κ (see Nair and Pollett, 1993, for the connection). Yaglom limits (1) are always quasistationary distributions in this sense, and conversely (by choosing $\nu = \kappa$). However, the interesting problem is to determine whether (1) holds when ν is a point mass in E.

As mentioned above, this problem has been solved when X is positive λ-recurrent. However, in most applications where quasistationarity plays a role, the relevant process is not λ-recurrent, and the existence question is open. Various special cases have been solved, notably by van Doorn (1991), Kijima (1993), Collet et al. (1995) and Bean et al. (1997). Considerable progress on the general problem was made recently by Kesten (1995). He showed that in discrete time, with $E = \mathbb{N}$, a bounded jump condition implies the validity of (1) independently of the λ-classification.

In this paper, we shall also consider Yaglom limits without λ-recurrence assumptions, but for Markov processes on arbitrary state spaces. We shall see that Kesten's bounded jump type of assumption again enters the picture, but this time by entirely different methods. Moreover, our approach unifies the theory of Yaglom limits for both common diffusions and Markov chains. The methodology of this paper is related to that of Breyer (1998b), to which we sometimes refer. See also Breyer (1997).

We now list three assumptions which we shall make on the Markov process X. The reader will note that all these are satisfied by a wide range of processes often encountered, including diffusions on $\mathbb{R}^d$ and many Markov chains.

Assumption (A1): There exists a Radon measure m on E (finite on compacts) such that

$$\mathbb{E}_x[f(X_t), \zeta > t] = \int p_t(x, y) f(y) m(dy), \quad f \geq 0, \tag{2}$$

and $(t, x, y) \mapsto p_t(x, y) > 0$ is jointly continuous. Moreover, we assume that m is *excessive.*

Note that we implicitly asume some sort of irreducibility of the state space. Recall that an excessive measure m must satisfy $\mathbb{E}_m[f(X_t), \zeta > t] \leq \langle m, f \rangle$ for all

$f \geq 0$, $t \geq 0$. It is well known that, as soon as (2) holds for some Radon measure m, it also holds for some other, excessive measure m'. Thus the last sentence in Assumption (A1) is only there for convenience.

Secondly, we shall require a parabolic Harnack inequality. The validity of (A2) below is well known for common diffusions (Friedman, 1964). When X is a Markov chain, (A2) is also satisfied (Breyer, 1998a). We remind the reader that a function $u : (0,\infty) \times E \to [0,+\infty]$ is called *parabolic* if is satisfies

$$\frac{\partial}{\partial t} u(t,x) = \mathfrak{A} u(t,x) \quad \text{in } (0,\infty) \times E,$$

where $\mathfrak{A}$ is the *local martingale generator* of X (see Breyer, 1998a).

Assumption (A2): The process X satisfies a *parabolic Harnack inequality*: For any compact sets $K \subseteq E$, $\overline{K} \subset (0,\infty) \times E$, let $s > \sup\{t : (t,x) \in \overline{K}\}$. Then there exists a constant $C = C(K, \overline{K})$ such that every function u that is parabolic in $(0,\infty)$ satisfies

$$\sup_{(t,x)\in\overline{K}} u(t,x) \leq C \cdot \inf_{y \in K} u(s,y).$$

Assumption (A3): For each compact set $K \subset E$, there exists another compact set $K' \subset E$ such that

$$\mathbb{P}_x(X_{T_{K^c}} \in K', \zeta > T_{K^c}) = 1, \quad x \in K,$$

where $T_{K^c} = \inf\{t > 0 : X_t \notin K\}$ is the first exit time from K.

Assumption (A3) requires the jumps of X to be bounded (but not uniformly). Armed with (A1)-(A3), we shall derive our results.

2. Entrance laws

Let X be a Markov process on a separable metric space E, with finite lifetime and satisfying Assumption (A1).

For a given probability measure ν such that $X_0 \sim \nu$, we denote by $(\nu_t : t > 0)$ the family of probability measures on E given by

$$\langle \nu_t, f \rangle = \mathbb{E}_\nu(f(X_t) \,|\, \zeta > t), \quad t > 0, f \geq 0. \tag{3}$$

We shall be mainly interested in the family of *entrance laws* $(\eta^s : s > 0)$ defined as

$$\eta_t^s(dy) = \int \nu_s(dx) p_t(x,y) m(dy), \quad t > 0. \tag{4}$$

These entrance laws are said to converge if, for each $t > 0$, the probability measures η_t^s converge weakly, as $s \to \infty$.

Proposition 1. *The Yaglom limit κ exists if and only if the entrance laws η^s converge, as $s \to \infty$, to the entrance law defined by*

$$\kappa_t(dy) = e^{-\Lambda t}\kappa(dy), \quad t > 0.$$

Proof. If the Yaglom limit $\nu_t \Rightarrow \kappa$ exists, then taking g bounded continuous on E we have

$$\begin{aligned}\langle \eta_t^s, g\rangle &= \mathbb{E}_\nu(g(X_{t+s}), \zeta > t+s)/\mathbb{P}_\nu(\zeta > s)\\ &= \langle \nu_{t+s}, g\rangle \mathbb{P}_\nu(\zeta > t+s)/\mathbb{P}_\nu(\zeta > s).\end{aligned}$$

Putting first $g = 1$, we find that

$$L(t) = \lim_{s\to\infty} \mathbb{P}_\nu(\zeta > t+s)/\mathbb{P}_\nu(\zeta > s)$$

exists, and then $L(t) = e^{-\Lambda t}$ for some $\Lambda \geq 0$ by a standard argument. Now letting g range over the bounded continuous functions on E, we get $\eta_t^s \Rightarrow \kappa_t$ for each $t > 0$, as $s \to \infty$. Conversely, suppose that the latter holds for each $t > 0$; choosing $t_0 > 0$ and g bounded and continuous, we find

$$\begin{aligned}\lim_{s\to\infty}\langle \nu_s, g\rangle &= \lim_{s\to\infty}\langle \eta_t^{s-t}, g\rangle/\langle \eta_t^{s-t}, 1\rangle\\ &= \langle \kappa_t, g\rangle/\langle \kappa_t, 1\rangle\\ &= \langle \kappa, g\rangle,\end{aligned}$$

and this establishes the existence of the Yaglom limit. □

Due to the above result, we can (and will) study the convergence of the entrance laws (η^s) rather than that of the probability measures (ν_t) directly. Since these entrance laws are best studied by reversing the direction of time, we recall now some facts and notation we shall need.

Since by (A1), the measure m is excessive, we define a dual transition function $\widehat{P}_t(x, dy) = \widehat{p}_t(x, y)m(dy)$ by setting

$$\widehat{p}_t(x, y) = p_t(y, x), \quad x, y \in E, t > 0. \tag{5}$$

Associated with $(\widehat{P}_t)$ is a Markov process, to be denoted $\widehat{X}$.

Note that Assumptions (A2) and (A3) will often hold for $\widehat{X}$ when they hold for X.

Indeed, consider first (A2). If X is a diffusion with generator L, then $\widehat{X}$ is typically a diffusion with generator $\widehat{L}$ satisfying $\int \widehat{L}f \cdot gdm = \int f \cdot Lgdm$ for all suitable test functions f, g. Conditions which ensure (A2) for $\widehat{L}$ are given in (Friedman, 1964). Supposing instead that X is a minimal Markov chain on a countable state space, it is clear that $\widehat{X}$ is one, too. Thus (A2) holds by (Breyer, 1998a).

Now consider (A3). The process $\widehat{X}$ can be realized as a version of X_{-t}, when $X_0 \sim m$ and $\widehat{X}_0 \sim m$. We can therefore relate the size of jumps of $\widehat{X}$ to the size of the jumps of X. For example, if the jump sizes of X are uniformly bounded, then so are those of $\widehat{X}$. A standard argument can then be used to get (A2) for $(\widehat{P}_t)$, at least for m-almost all $x \in E$.

Let $\overline{E} = (-\infty, 0] \times E$. We shall be interested in the family (k^s) of functions on $\overline{E}$ given by

$$k^s(t,y) = \begin{cases} \int \widehat{p}_{s+t}(y,a)\nu(da)/\mathbb{P}_\nu(\zeta > s) & t \geq -s, y \in E \\ 0 & t < -s, y \in E. \end{cases} \tag{6}$$

Note that it is obvious from (5) and (4) that $k^s(t,\cdot) = d\eta^s_t/dm$ for $t \geq -s$.

Given $(T_0, \widehat{X}_0) \in \overline{E}$, the (backward) spacetime process associated with $\widehat{X}$ is the process $\widetilde{X}_t = (T_0 - t, \widehat{X}_t)$. A function $u : \overline{E} \to [0, \infty]$ is called *parabolic* for $\widehat{X}$ if the process $u(\widetilde{X}_t)$ is a local martingale. We can now state

Lemma 2. *For any $s > 0$, the function k^s is parabolic for $\widehat{X}$ on $(-s, 0] \times E$, and satisfies*

$$\int k^s(0,y)m(dy) = 1. \tag{7}$$

Proof. We prove the stronger statement that k^s is spacetime invariant on $(-s,0]\times E$: take (t,x) in this set, and choose $u < t+s$;

$$\begin{aligned}
\widetilde{\mathbb{E}}_{(t,x)}[k^s(\widetilde{X}_u), \widetilde{\zeta} > u] &= \int \widehat{p}_u(x,y)k^s(t-u,y)m(dy) \\
&= \int m(dy)\widehat{p}_u(x,y) \int \widehat{p}_{t+s-u}(y,a)\nu(da)/\mathbb{P}_\nu(\zeta > s) \\
&= \int \widehat{p}_{t+s}(x,a)\nu(da)/\mathbb{P}_\nu(\zeta > s) \\
&= k^s(t,x).
\end{aligned}$$

It follows that k^s is invariant for $\widetilde{X}$ on $(-s,0]\times E$, and hence that

$$\frac{\partial}{\partial t}k^s = \widehat{\mathfrak{A}}k^s \quad \text{ on } (-s,0]\times E,$$

where $\widehat{\mathfrak{A}}$ is the local martingale generator of $\widehat{X}$. Thus k^s is parabolic (Breyer, 1998a). Finally,

$$\begin{aligned}
\int k^s(0,y)m(dy) &= \int m(dy) \int \widehat{p}_s(y,a)\nu(da)/\mathbb{P}_\nu(\zeta > s) \\
&= \int \nu(da) \int p_s(a,y)m(dy)/\mathbb{P}_\nu(\zeta > s) \\
&= 1,
\end{aligned}$$

and this is the required normalization. □

More generally, it is easy to check that k^s is *excessive* for $\widetilde{X}$ on all of $\overline{E} = (-\infty,0]\times E$. There thus exists an integral representation of k^s on the *Martin compactification* $\widetilde{F}$ of $\widetilde{X}$. This is computed as follows.

Define the *Martin kernel* of $\widetilde{X}$ with normalization m by

$$\text{(8)} \qquad \widetilde{K}(t,x;s,y) = \frac{1_{(-\infty,t]}(s)p_{t-s}(y,x)}{\mathbb{P}_y(\zeta > -s)}, \quad s,t \le 0; x,y \in E$$
$$\left(= \frac{1_{(-\infty,t]}(s)\widehat{p}_{t-s}(x,y)}{\int m(dz)\widehat{p}_{-s}(z,y)} \right)$$

The space $\widetilde{F}$ is the completion of $\overline{E}$ with respect to the metric

$$d(\overline{y},\overline{z}) = \int 1 \wedge \left|\widetilde{K}(t,x;\overline{y}) - \widetilde{K}(t,x;\overline{z})\right| \gamma(t,x)dt \otimes m(dx), \quad \overline{y},\overline{z} \in \overline{E},$$

where γ is any strictly positive $dt \otimes dm$ integrable function. The function $\widetilde{K}(t,x;\cdot)$ has, for each $(t,x) \in \overline{E}$, a continuous extension to $\widetilde{F}$, and $\widetilde{F}$ is compact (Doob, 1984; Bass, 1995).

Proposition 3. *For each $s > 0$, there exists a probability ξ^s on $\widetilde{F}$ such that*

$$k^s(t,x) = \int \widetilde{K}(t,x;\overline{y})\xi^s(d\overline{y}). \tag{9}$$

This measure is concentrated on $\{-s\} \times supp(\nu) \subset \overline{E}$ and has the explicit representation

$$\xi^s(dt,dy) = \epsilon_{-s}(dr)\mathbb{P}_y(\zeta > -s)\nu(dy)/\mathbb{P}_\nu(\zeta > -s), \tag{10}$$

where ϵ_{-s} is the point mass at $\{-s\}$.

Proof. Since k^s is excessive for $\widetilde{X}$ with normalization (7), the standard integral representation on the Martin compactification applies (Meyer, 1968), and this is (9). To get (10), we note that by (6) and (8),

$$\begin{aligned}\int \widetilde{K}(t,x;r,y)\xi^s(dr,dy) &= \int_{\overline{E}} \frac{1_{(-\infty,t]}(r)p_{t-r}(x,y)}{\mathbb{P}_y(\zeta > -r)}\frac{\mathbb{P}_y(\zeta > -s)\nu(dy)}{\mathbb{P}_n u(\zeta > -s)}\epsilon_{-s}(dr)\\ &= 1_{(-\infty,t]}(-s)\cdot\int p_{t+s}(x,y)\nu(dy)/\mathbb{P}_\nu(\zeta > s)\\ &= k^s(t,x).\end{aligned}$$

□

The properties of the functions (k^s) developed above bear a striking resemblance to those of the functions (h_s) studied in (Breyer, 1998b). A similar approach as in that paper can therefore be expected to lead us to our goal. In particular, we can immediately state the following result:

Lemma 4. *Suppose that $\widehat{X}$ satisfies Assumptions (A1)-(A3), then for every set $\{s_n\} \subset (0,\infty)$ with no accumulation point, there exists a subsequence $\{s_{n(j)}\}$ and a (possibly zero) function k, the latter parabolic in $\overline{E}$, such that*

$$\lim_{k\to\infty} k^{s_{n(j)}}(t,x) = k(t,x), \quad (t,x) \in \overline{E}.$$

Proof. Since $\overline{F}$ is compact, the measures ξ^{s_n} are tight, and there exists ξ such that $\xi^{s_n(j)} \Rightarrow \xi$. By weak convergence, we then have

$$\begin{aligned}k(t,x) &:= \int \widetilde{K}(t,x;\overline{y})\xi(d\overline{y})\\ &= \lim_{j\to\infty}\int \widetilde{K}(t,x;\overline{y})\xi^{s_{n(j)}}(d\overline{y})\\ &= \lim_{j\to\infty} k^{s_{n(j)}}(t,x),\end{aligned}$$

for the function $\overline{y} \mapsto \widetilde{K}(t,x;\overline{y})$ is continuous on the closure of $(-\infty,-t)\times E$. The parabolicity of k follows from (Breyer, 1998a, Lemma 9). □

Since each function k is a prospective density for the limiting entrance law κ_t of Lemma 1, we shall need conditions which guarantee that $k>0$ on $\overline{E}$.

Unlike the similar problem treated in (Breyer, 1998b, Lemma 9) for (h_s), it is not the case here that a judicious choice of the initial distribution ν together with (A1)-(A3) suffices for $k>0$. The reason for this difference is that in (Breyer, 1998b), the function h_s was normalized by the measure ν. In the case of k^s, we have no control over the normalizing measure, m. Instead, we shall develop in the next section a combined topological and probabilistic approach to the positivity of k. The uniqueness of k, which we will also need to apply Proposition 1, is considered in the last section.

Lemma 5. *Suppose that $\widehat{X}$ satisfies (A1)-(A3). If the probability measures (ν_t) are tight, then every limit point k of (k^s) described in Lemma 4 is nonzero and satisfies*

$$\int k(0,y)m(dy) = 1. \tag{11}$$

Consequently, there exists a sequence $\{s(j)\}$ such that

$$\lim_{s(j)\to\infty} \mathbb{E}_\nu[f(X_{s(j)})\,|\,\zeta > s(j)] = \int f(y)k(0,y)m(dy), \quad f\in L^1(\kappa). \tag{12}$$

When k is independent of the sequence $\{s(j)\}$, the Yaglom limit exists and $\kappa(dy) = k(0,y)m(dy)$.

Proof. Given $\epsilon>0$, let $K\subset E$ be compact such that $\nu_t(K)\geq 1-\epsilon$ for all t. By (A2) and the bounded convergence theorem,

$$\int_K k(0,y)m(dy) = \lim_{j\to\infty}\int_K k^{t_j}(0,y)m(dy) \geq 1-\epsilon.$$

Letting $K\uparrow E$ gives (11). Now suppose that the limit is independent of the subsequence $\{s(j)\}$, and let $f\in L^1(\kappa)_+$. By Fatou's lemma, it follows that

$$\lim_{s\to\infty}\int k^s(0,z)f(z)m(dz) \geq \int f(z)\kappa(dz),$$

and by Scheffé's Lemma, we deduce that $k^s(0,\cdot)$ converges to $k(0,\cdot)$ in $L^1(dm)$; in particular, this implies (12). □

3. Tightness

In this section we shall make the assumption

Assumption (TA1): There exists $\Lambda \geq 0$ such that

$$\lim_{s\to\infty} \mathbb{P}_\nu(\zeta > t + s \,|\, \zeta > s) = e^{-\Lambda t}.$$

More generally, we could assume that the limit $e^{-\Lambda t}$ in (TA1) is instead of the form $\int e^{-\lambda t}\xi(d\lambda)$ for some probability measure ξ on $\mathbb{R}$. The arguments in this section then apply with trivial changes.

The validity of (TA1) was shown under suitable assumptions in (Breyer, 1998b); see lemmas 9 and 14 of that paper, together with the examples in the last section thereof. We summarize the results found there below.

Proposition 6. *Suppose that the Markov process X satisfies (A1)-(A3) and that the initial measure ν is compactly supported. If, as $s \to \infty$, there exists a time homogeneous Markov process Y such that*

$$(X_r : r \leq t \,|\, \zeta > s) \Rightarrow (Y_r : r \leq t),$$

then (TA1) holds.

It is known that the existence of the conditioned process Y does *not* guarantee the Yaglom limit. We are therefore interested here in conditions which guarantee the tightness of the set of probability measures (ν_t) defined by (3).

According to Prokhorov's criterion, tightness of probability measures always occurs when these are defined on a common compact set. Since all the probabilities ν_t are defined on E, the family (ν_t) is tight on any compactification F of E, and the possible limit points are consequently all probability measures on F, which may or may not charge E. To find out where the probability mass ends up as $t \to \infty$, we require a compactification F with good probabilistic properties.

There will in general exist many "good" compactifications F. Our task here is to show the existence of one such, and to list the properties which make it "good". In applications, any compactification with these properties may then be used.

Proposition 7. *Let (A1) hold. There exists a metric compactification F' of E with the following properties:*

(i) *Every compactly supported function on E which was originally continuous remains continuous on F';*

(ii) *There exists on F' a Strong Markov process (X'_t) extending (X_t): if $T'_E = \inf\{t > 0 : X'_t \in E\}$ then the laws of $(X'_{T'_E+t} : t > 0)$ and $(X_t : t > 0)$ coincide.*

(iii) *The lifetime ζ' of X' has a Laplace transform which is continuous in x:*

$$x \mapsto \int_0^\infty e^{-pt}\mathbb{P}'_x(\zeta' > t)dt \quad \text{is continuous on } F'.$$

The items listed are standard properties of the Ray-Knight compactification procedure, see (Rogers and Williams, 1994). This procedure is normally performed on a process X with infinite lifetime. We outline below the simple changes required in the present setting, where it nearly always happens that $\zeta < \infty$.

Proof. We begin by adding to E a cemetery state ∂, isolated from E, and set $X_{t+\zeta} = \partial$, $t \geq 0$. Now ζ is the first hitting time of ∂, and the process has an infinite lifetime on $E \cup \{\partial\}$. Let F be any Ray-Knight compactification (Rogers and Williams, 1994) of $E \cup \{\partial\}$. Since ∂ was isolated from E, it remains isolated in F. Thus $F' = F\backslash\{\partial\}$ is again compact. By (A1), the resolvent (V_p) of X maps the set of uniformly continuous functions (in the original topology) into itself, and consequently part (i) follows. On F, there exists a unique Markovian resolvent (U_p) with the following properties:

(a) $U_p f = V_p f$ on $E \cup \{\partial\}$, whenever $f = 0$ on $F\backslash(E \cup \{\partial\})$,

(b) $U_p : C_b(F) \to C_b(F)$, where $C_b(F)$ is the set of bounded continuous functions on F.

Associated with (U_p) is a Strong Markov process (Y_t) on F which extends X on $E \cup \{\partial\}$ in the manner of (ii). In particular, ∂ is an absorbing state for Y, and we can construct X' by killing Y the first time it hits ∂. Clearly X' satisfies (i) and (ii). Moreover, since ∂ is isolated, the function $f = 1_{F'}$ is bounded and continuous on F. Applying (a) above gives

$$V_p 1_{F'}(x) = U_p 1_{F'}(x) = \int_0^\infty e^{-pt}\mathbb{P}'_x(\zeta' > t)dt,$$

and this is continuous in x by (b), first on F and hence on F'. □

Armed with the process X', we now investigate the tightness of (ν_t).

Proposition 8. *Let (A1) and (TA1) hold. If μ is any weak limit point of (ν_t) on F', then one of the following holds.*

(i) $\Lambda > 0$ *and*

$$\mu\Big(x \in F' : \mathbb{P}'_x(0 < \zeta' < \infty) = 1\Big) = 1,$$

(ii) $\Lambda = 0$ *and*

$$\mu\Big(x \in F' : \mathbb{P}'_x(\zeta' = \infty) = 1\Big) = 1.$$

In particular, the measures (ν_t) are tight only if $\Lambda > 0$, because $\mathbb{P}_x(\zeta < \infty) = 1$ for all $x \in E$.

Proof. Let $\nu_{t_n} \Rightarrow \mu$ on F'. By Proposition 7, (ii) and (iii), we have for $p > 0$,

$$\begin{aligned}
\int_0^\infty e^{-ps}\mathbb{P}'_\mu(\zeta' > s)ds &= \int \mu(dx) \int_0^\infty e^{-ps}\mathbb{P}'_x(\zeta' > s)ds \\
&= \lim_{n\to\infty} \int \nu_{t_n}(dx) \int_0^\infty e^{-ps}\mathbb{P}'_x(\zeta' > s)ds \\
&= \lim_{n\to\infty} \int \nu_{t_n}(dx) \int_0^\infty e^{-ps}\mathbb{P}_x(\zeta > s)ds \\
&= \lim_{n\to\infty} \int_0^\infty e^{-ps}\mathbb{P}_\nu(\zeta > s + t_n \,|\, \zeta > t_n)ds \\
&= \int_0^\infty e^{-ps}e^{-\Lambda s}ds,
\end{aligned}$$

where we have used the fact that ν_t is concentrated on E, and then (TA1) and the bounded convergence theorem. Now $s \mapsto \mathbb{P}'_x(\zeta' > s)$ is right continuous and decreasing, so by the uniqueness property of Laplace transforms, we must have $\mathbb{P}'_\mu(\zeta' > s) = e^{-\Lambda s}$. Now let $s \to 0$; we see that $\mathbb{P}'_\mu(\zeta' > 0) = 1$, and since $\mathbb{P}'_x(\zeta > 0) \leq 1$, it follows that μ is concentrated on those x for which $\mathbb{P}'_x(\zeta' > 0) = 1$. If we let $s \to \infty$, there are two cases to consider. Either $\Lambda > 0$, in which case $\mathbb{P}'_\mu(\zeta' = \infty) = 0$ by bounded convergence, and μ must be concentrated on those x for which $\mathbb{P}'_x(\zeta' < \infty) = 1$. In the other case, $\Lambda = 0$ and then $\mathbb{P}'_\mu(\zeta' = \infty) = 1$, whence we get statement (ii) above. □

According to the above, there can never exist a nonzero Yaglom limit in the case $\Lambda = 0$, since any possible limit measure ξ is then concentrated on that part of the boundary $F' \backslash E$ from which X' takes forever to die.

It is possible to classify some of the boundary points $x \in F' \backslash E$ as follows:

Definition 9. *Let X' be the process in Proposition 7.*

(i) *A point x is called asymptotically remote (from the cemetery state) if it forms a trap: $\mathbb{P}'_x(X'_t = x \; \forall t) = 1$. This property is denoted (AR).*

(ii) *A point x is called asymptotically proximate if $\mathbb{P}'_x(\zeta' = 0) = \mathbb{P}_x(X'$ hits ∂ immediately$) = 1$. This property is denoted (AP).*

The terminology for (i) is due to Ferrari et al. (1995) and for (ii) it is due to Pakes (1995). The Blumenthal zero one law implies that every state $x \in F$ satisfies $\mathbb{P}'_x(\zeta' = 0) = 0$ or 1. Some of the remaining boundary points may be holding points, where X' waits an exponentially distributed time before jumping. If holding points do not exist on the boundary $F' \backslash E$, then each boundary point is either (AP), (AR), or else the sample path of X' must immediately hit E a.s. (after which it stays there until death); it is not clear how best to classify these other types of points

The value of the (AR)/(AP) classification is that it appears naturally in the conclusions of Proposition 8. Specifically, it is worth mentioning the following corollary:

Corollary 10. *If (A1) and (TA1) hold and every boundary point in $F' \backslash E$ is either (AP) or (AR), then the measures (ν_t) are tight (in the original topology of E) if and only if $\Lambda > 0$.*

Example 11. Let (X_t) be a uniformly elliptic diffusion on $\mathbb{R}^d$, killed upon first exit from the unit disc $E = \{x : \|x\| < 1\}$. Both (A1) and (TA1) hold, the latter being a consequence of the eigenfunction expansion of the semigroup of X. The Euclidean boundary points of E are well known to be regular for E^c, that is $\lim_{x \to \partial E} \mathbb{P}_x(\zeta > t) = 0$ for each $t > 0$. Let $F' = E \cup \{+\infty\}$ be the one-point compactification of E, and put $X'_t = \partial \notin F'$ if $X'_0 = +\infty$, and $X'_t = X_t$ if $X'_0 \in E$. The process X' clearly extends X, and it is easy to see that its semigroup (hence its resolvent) maps $C_b(F')$ into itself. Clearly, every boundary point is (AP) here. By the corollary above, the (ν_t) are tight on E.

Example 12. Let X be one-dimensional Brownian motion, killed upon first hitting zero. The state space is $E = (0, +\infty)$, and (TA1) holds by using the spectral representation of the Laplacian on E (see McKean, 1956). We shall take F' as the Martin boundary $[0, +\infty]$. Here also, the semigroup of X maps $C_b(F')$

into itself, as can be easily checked. Thus there exists a Feller-Dynkin process on F' extending X. The boundary point 0 is (AP), while the boundary point $+\infty$ is (AR). The measures (ν_t) are not tight since $\Lambda = 0$.

Example 13. Let X be a birth and death chain on $\mathbb{Z}$ with constant birth and death parameters, killed upon first hitting 0. The assumption (TA1) was shown in (Jacka and Roberts, 1994). We shall take $E = \{1, 2, 3, \dots\}$. This is the analogue of the previous example. The boundary point $+\infty$ is still (AR), but now the boundary point 0 is not needed. This gives a case where all boundary points are (AR).

Example 14. Let X be a Markov chain on $\{1, 2, 3, \dots\}$ whose behaviour may be described as follows: when started in $x \geq 1$, it may jump up to $x + 1$, or jump catastrophically back to state 1, or disappear from the state space. Such a process was called a pure birth process with catastrophes by Pakes (1995). In that paper, he gave conditions under which the point $+\infty$ is (AP), (AR), or neither. When this last possibility occurs, the point $+\infty$ can be a holding point (where the process waits for an exponentially distributed amount of time before jumping), or else it might split up into several distinct boundary points (this occurs when $\mathbb{P}_x(\zeta > t)$ oscillates as $x \to +\infty$).

4. Yaglom limits

In this section, we end the analysis of (k^s) and prove the existence of Yaglom limits.

Let D be the support of the initial probability measure ν. We define

$$\widetilde{D}_{-s} = d\text{-closure of } (-\infty, -s] \times D, \quad \widetilde{D}_{-\infty} = \cap_{s>0} \widetilde{D}_{-s}.$$

The representing measure ξ^s of k^s (Proposition 3) is concentrated on the closed set $\widetilde{D}_{-s}$. It follows that any weak limit point μ of the family (ξ^s) is concentrated on $\widetilde{D}_{-\infty}$. We are looking for conditions sufficient for having $\widetilde{D}_{-\infty} = \{\overline{z}_0\}$, a singleton. This will not be the case generally, unless D is sufficiently small.

Proposition 15. *Let X satisfy (A1)-(A3), and suppose that $D = supp(\nu)$ is compact in E. If $\overline{z} \in \widetilde{D}_{-\infty}$ and $\widetilde{K}(t, x; \overline{z}) \not\equiv 0$ is minimal parabolic, then there exists a function $\varphi^* > 0$ and a real number Λ such that*

$$\widetilde{K}(t, y; \overline{z}) = e^{-\Lambda t} \varphi^*(y), \quad \widehat{\mathfrak{A}} \varphi^* = -\Lambda \varphi^*. \tag{13}$$

Moreover, the following statements are equivalent:

(i) *For some* $y \in D$, $\underline{\lim}_{s\to-\infty} \widetilde{K}(\overline{x}; s, y) > 0$ *holds for all* $\overline{x}$ *in a set of positive* $dt \otimes dm$ *measure,*

(ii) *There exists a unique point* $\overline{z}_0$ *such that* $\widetilde{D}_{-\infty} = \{\overline{z}_0\}$ *and* $\widetilde{K}(t, x; \overline{z}_0)$ *is a nontrivial minimal parabolic function, hence of the form* (13)*,*

(iii) *For every* $y \in D$, $(t, x) \in \overline{E}$,

$$\lim_{s\to\infty} p_{t+s}(y, x)/\mathbb{P}_y(\zeta > s) = e^{-\Lambda t}\varphi^*(x). \tag{14}$$

Proof. Since the functions $(t, x) \mapsto p_{t+t_0}(x, y)$ are parabolic for all $y \in E$ on $(-t_0, \infty) \times E$, Assumption (A2) implies the existence of some constant $C(r)$ such that, for all $t > 1$,

$$\sup_{x\in D} p_t(x, y) \le C(r) \inf_{z\in D} p_{t+r}(z, y).$$

Using (5), this can be written

$$\widehat{p}_t(y, x) \le C(r)\widehat{p}_{t+r}(y, z), \quad y \in E; x, z \in D; t > 1.$$

This last inequality is identical to Assumption (A4) of (Breyer, 1998b) if we replace the set N there by D. The proof proceeds now identically, mutatis mutandis, to the proofs of Propositions 10 and 15 of that paper. □

The conclusions of Proposition 15 hold when X has a symmetrizing measure (see Breyer, 1998b).

Theorem 16. *Let (A1)-(A3) hold, and* ν *have compact support. If* (ν_t) *is tight and (ii) of Proposition 15 holds, then the Yaglom limit* (1) *exists.*

Proof. By Proposition 3, any limit point k of (k^s) can be written

$$k(t, x) = \int \widetilde{K}(t, x; \overline{y})\xi(d\overline{y}),$$

where ξ is the weak limit (in the Martin topology) of some sequence ξ^{s_n}. Since the latter is concentrated on $\widetilde{D}_{-s_n}$, it follows that $\xi(\widetilde{D}_{-\infty}) = 1$. Since $k > 0$, the measure ξ is concentrated on the parabolic boundary points $\overline{z}_0 \in \widetilde{D}_{-\infty}$ such that $\widetilde{K}(\cdot; \overline{z}_0) > 0$. By uniqueness of $\overline{z}_0$, we see that k is independent of the subsequence s_n, whence by Lemma 5, the Yaglom limit exists. □

In cases when ν charges all of the state space E, all we can deduce from Proposition 15 is that the weak limit point ξ is concentrated on $\widetilde{F}_{-\infty}$. This set usually consists of a large number of points, and by choosing the initial distribution ν appropriately, ξ can often be made to charge any given one, as the next example shows. Compare it with Example 16 of Breyer (1998b).

Example 17. Let $X_t = W_t - \alpha t$, where W is a one-dimensional Brownian motion, and X is killed upon first leaving $(0, \infty)$. The drift is towards zero, i.e. $\alpha \geq 0$. It is well known (Revuz and Yor, 1991) how to transform the Brownian motion into X by a Girsanov transformation. In particular, suppose that (P_t) is the semigroup of Brownian motion B, killed upon leaving $(0, \infty)$. The semigroup (Q_t) of X is then given by the formula

$$Q_t(x, dy) = \frac{1}{e^{-\alpha x}} P_t(x, dy) e^{-\alpha^2 t/2} e^{-\alpha y}, t > 0, x > 0.$$

Stated differently, the spacetime process $(T_0 - t, X_t)$ is the g-transform of $(T_0 - t, B_t)$, with $g(t, x) = e^{\alpha^2 t/2} e^{-\alpha x}$. Thus any excessive function h for $(T_0 - t, X_t)$ can be written $h = k/g$ for some excessive function k of $(T_0 - t, B_t)$. We can use this observation to compute the minimal parabolic Martin boundary of X in terms of that of B. The latter is described in (Doob, 1984, p. 375); see also Example 16 of Breyer (1998b). For any $\tau < 0$, set

$$\mathcal{K}_0^g(t, x; \tau) \propto \begin{cases} \frac{x}{\sqrt{2\pi(t-\tau)^3}} \exp\left(-\frac{(x - \alpha(t-\tau))^2}{2(t-\tau)}\right) & \text{if } t > \tau, \\ 0 & \text{if } t \leq \tau, \end{cases},$$

and $\int \mathcal{K}_0^g(0, x; \tau)\nu(dx) = 1$. Any Martin sequence $(s_n, y_n) \to (\tau, 0)$ converges to the boundary point associated with $\mathcal{K}_0^g(\cdot; \tau)$. For any $\gamma \leq 0$, put $c = (\gamma^2 - \alpha^2)/2$, and

$$\mathcal{K}_1^g(t, x; \gamma) \propto \begin{cases} e^{ct} e^{\alpha x} \sinh(\sqrt{|\alpha^2 + 2C|} \cdot x) & \text{if } c > -\alpha^2/2, \\ e^{ct} x e^{\alpha x} & \text{if } c = -\alpha^2/2. \end{cases}$$

Once again, we normalize by requiring $\int \mathcal{K}_1^g(0, x; \gamma)\nu(dx) = 1$. Any Martin sequence (s_n, y_n) with $\lim_{n\to\infty}(s_n, y_n/s_n) = (-\infty, \gamma)$ converges to the boundary point associated with $\mathcal{K}_1^g(\cdot; \gamma)$.

Using the fact that the measure $m(dx) = e^{2\alpha x} dx$ is symmetrizing, we find immediately that for the time reversal $\widehat{X}$ of X, the part $\widetilde{E}_{-\infty}$ of the parabolic Martin boundary consists of the normalized minimal functions

$$\widetilde{\mathcal{K}}_1^g(t,x;c) \propto \begin{cases} e^{ct}e^{-\alpha x}\sinh(\sqrt{|\alpha^2+2c|}\cdot x) & \text{if } c > -\alpha^2/2 \\ e^{ct}xe^{-\alpha x} & \text{if } c = -\alpha^2/2, \end{cases}$$

where $\int \widetilde{\mathcal{K}}_1^g(0,x;c)dx = 1$. The measures $\mu_c(dx) = \widetilde{\mathcal{K}}_1^g(0,x;c)dx$ are the quasistationary distributions of X. Note that this agrees with results of Martinez and San Martin (1994).

Suppose now that we choose the initial distribution ν in Proposition 3 to be μ_c, which is *not* compactly supported. An easy calculation shows that

$$k^r(t,x) = e^{ct}(d\mu_c/dx)(x), \quad t > -r.$$

The Yaglom limit then becomes

$$\lim_{t\to\infty} \mathbb{E}_{\mu_c}(f(X_t)\,|\,\zeta > t) = \langle \mu_c, f\rangle.$$

Here any Martin sequence (s_n, y_n) with $\lim_{n\to\infty}(s_n, y_n/s_n) = (-\infty, \gamma)$ converges to the minimal boundary point associated with the function $\widetilde{\mathcal{K}}_1^g(\cdot, c)$ where $c = (\gamma^2 - \alpha^2)/2$ as before. To force the Yaglom limit to become $\nu_t \Rightarrow \mu_{-\alpha^2/2}$ (that is, the miminal quasistationary distribution), we need to have Martin sequences (s_n, y_n) such that $\lim_n y_n/s_n = 0$. This is true whenever ν is compactly supported, as described in Proposition 15.

5. Acknowledgement

I thank Phil Pollett and Gareth Roberts for invaluable discussions on the topics of this paper. This research was conducted as part of my Ph.D. work.

References

[1] Anderson, W.J. (1991). *Continuous-Time Markov Chains: An Applications-Oriented Approach.* Springer-Verlag, New York.

[2] Bass, R.F. (1995). *Probabilistic Techniques in Analysis.* Springer-Verlag, New York.

[3] Bean, N.G., Bright, L., Latouche, G., Pearce, C.E.M., Pollett, P.K. and Taylor, P.G. (1997) The quasi-stationary behaviour of quasi-birth-and-death processes. *Ann. Appl. Probab.* **7** no. 1, 134-155.

[4] Breyer, L.A. (1997) *Quasistationarity and Conditioned Markov Processes.* Ph.D. Thesis, The University of Queensland.

[5] Breyer, L.A. (1998a) On a parabolic Harnack inequality for Markov chains. Preprint available at `http://www.statslab.cam.ac.uk/~laird/`

[6] Breyer, L.A. (1998b) Quasistationarity and Martin boundaries: Conditioned processes. Preprint available at `http://www.statslab.cam.ac.uk/~laird/`

[7] Collet, P., Martinez, S. and San Martin, J. (1995). Asymptotic laws for one-dimensional diffusions conditioned to nonabsorption. *Ann. Probab.* **23** (3), 1300-1314.

[8] Doob, J.L. (1984). *Classical Potential Theory and Its Probabilistic Counterpart.* Springer-Verlag, New York.

[9] van Doorn, E.A. (1991). Quasi-stationary distributions and convergence to quasi-stationarity of birth-death processes. *Adv. Appl. Probab.* **23**, 683-700.

[10] Friedman, A. (1964). *Partial Differential Equations of Parabolic Type.* Prentice-Hall, Englewood Cliffs, N.J.

[11] Jacka, S.D. and Roberts, G.O. (1994). Weak convergence of conditioned birth and death processes. *J. Appl. Probab.* **31**, 90-100.

[12] Kesten, H. (1995). A ratio limit theorem for (sub) Markov chains on $\{1, 2, 3, \dots\}$ with bounded jumps. *Adv. Appl. Probab.* **27**, 652-691.

[13] Kijima, M. (1993). Quasi-limiting distributions of Markov chains that are skip-free to the left in continuous-time. *J. Appl. Probab.* **30**, 509-517.

[14] Martinez, S. and San Martin, J. (1994). Quasi-stationary distributions for a Brownian motion with drift and associated limit laws. *J. Appl. Probab.* **31**, 911-920.

[15] McKean, H.P. Jr. (1956). Elementary solutions for certain parabolic partial differential equations. *Trans. Amer. Math. Soc.* **82**, 519-548.

[16] Meyer, P.A. (1968). *Processus de Markov: la frontière de Martin.* Lecture Notes in Mathematics 77. Springer-Verlag, New York.

[17] Nair, N.G. and Pollett, P.K. (1993). On the relationship between μ-invariant measures and quasistationary distributions for continuous-time Markov chains. *Adv. Appl. Probab.* **25**,82-102.

[18] Pakes, A.G. (1995). Quasistationary laws for Markov processes: examples of an always proximate absorbing state. *Adv. Appl. Probab.* **27**, 120-145.

[19] Rogers, L.C.G. and Williams, D. (1994). *Diffusions, Markov processes, and Martingales, Vol. I.* Wiley, New York.

[20] Revuz, D. and Yor, M. (1991). *Continuous Martingales and Brownian Motion.* Springer-Verlag, New York.

[21] Seneta, E. and Vere-Jones, D. (1966). On quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states. *J. Appl. Probab.* **3**, 403-434.

[22] Yaglom, A.M. (1947). Certain limit theorems of the theory of branching stochastic processes (in Russian). *Dokl. Akad. Nauk. SSSR* **56**, 795-798.

Statistical Laboratory, Department of Pure Mathematics and Mathematical Statistics, 16, Mill Lane, Cambridge CB2 1SB

E-mail address: L.Breyer@statslab.cam.ac.uk