QUASISTATIONARITY AND MARTIN BOUNDARIES: CONDITIONED PROCESSES

L.A. BREYER

Abstract. Consider a Markov process X with finite lifetime ζ. In this paper, we derive sufficient conditions which allow the conditioning of X to an infinite lifetime. This is accomplished by showing the weak convergence, as $s \to \infty$, of the laws of $(X_r : r \le t \,|\, \zeta > s)$.

1. Introduction

Markov processes with finite lifetimes arise routinely in Applied Probability. The lifetime is typically a stopping time of the process, marking a transition which is being excluded. As an example, one may be interested in a diffusion before it hits a specified boundary, and so identify the lifetime with this hitting time.

Sometimes, the lifetime is finite but an order of magnitude larger than the time scale of the investigation. This occurs for certain epidemic models (Nåsell, 1995) where extinction of the infected population is certain, but measured in millions of years. The Markov process X measuring the number of infectives quickly settles down to a seemingly stationary regime (this is known as quasistationary behaviour), and the extinction event occurs as a result of a large deviation. On a human time scale, the properties of this process are well described by those of a certain conditioned process, constructed from X by conditioning on the event that extinction is arbitrarily distant.

In this paper, we shall be interested in a generalisation of this idea. Let X be a Markov process with lifetime ζ. We study the feasibility of defining a process Y as the weak limit, on finite time intervals $[0, t]$, of the processes $(X_r : r \le t \,|\, \zeta > s)$

Date: January, 1998.

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

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

as $s \to \infty$. The process Y can be viewed as X, conditioned on the event $\{\zeta = \infty\}$, even though this event has in many interesting cases probability zero.

This problem has been considered previously by various authors, among them Darroch and Seneta (1967), when X is a Markov chain with a finite state space, Jacka and Roberts (1994), Jacka et al. (1997) and Schrijner and van Doorn (1997), when X is a birth-death process, Collet et al. (1995), McKean (1963) and Jacka and Roberts (1997) in the case when X is a one-dimensional diffusion, and Pinsky (1985) in the multidimensional diffusion case under reversibility.

The connection between the above conditioned processes and (quasistationary) limiting conditional distributions are varied. Sometimes, the latter arise as limiting distributions for the former (Pollett, 1988). In a follow-up (Breyer, 1998b) to the present paper, a connection which arises by time reversal will be investigated. Jacka and Roberts (1995) have shown that the existence of a limiting conditional distribution implies the existence of the conditioned process for chains.

Conditioned processes of the type we study here also arise outside the literature on quasistationarity. The best known example to the author concerns the interpretation of the three dimensional Bessel process as a Brownian motion, conditioned on avoiding the origin. See papers by McKean (1963), Williams (1974), and Pitman (1975).

The plan of this paper is as follows: Section 2 introduces notation and the basic Assumption (A1) on X. We also discuss links between the spacetime Martin boundary and the Strong Ratio Limit Property. In Section 7, we prove the weak convergence to the process Y, under Assumptions (A1)-(A4) (Theorem 21). In the other sections, we introduce and discuss the remaining Assumptions, and prove various technical results. Assumptions (A2) and (A3) are introduced in Section 4, and Assumption (A4) in Section 5.

2. Notation, spacetime boundaries and the SRLP

Throughout the remaining sections, we shall fix a Strong Markov process X with right continuous sample paths, locally compact state space E, and lifetime

$$\zeta = \inf\{t > 0 : X_t \notin E\} < \infty.$$

The law of X given $X_0 \sim \nu$ is denoted $\mathbb{P}_\nu$, and $(\mathcal{F}_t)$ is the usual *completed* filtration. We shall also suppose that the semigroup (P_t) of the process admits a reference measure, i.e.

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

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

and $(t,x,y) \mapsto p_t(x,y) > 0$ is jointly continuous.

Note that we have made an implicit assumption of irreducibility in the above.

While X is the process of interest, many calculations will be performed on an auxiliary process, the (backward) *spacetime process* $\overline{X}$. This is the Markov process with state space $\overline{E} = (-\infty, 0] \times E$ which, when started with $\overline{X}_0 = (T_0, X_0)$ say, satisfies $\overline{X}_t = (T_0 - t, X_t)$; Sharpe (1988) makes an in-depth study of $\overline{X}$. Clearly also, the lifetime $\overline{\zeta}$ of $\overline{X}$ coincides with ζ, the lifetime of its second component X.

Let ν be some probability measure on E such that the map

$$(t,y) \mapsto \int \nu(dx) p_t(x,y)$$

is continuous on $[0,\infty) \times E$. The *spacetime Martin kernel with normalization* ν is the function $\overline{K} : \overline{E} \times \overline{E} \to [0, +\infty]$ given by

$$\overline{K}(t,x;s,y) = \frac{1_{(-\infty,t)}(s) p_{t-s}(x,y)}{\int \nu(dz) p_{-s}(z,y)}, \quad (t,x),(s,y) \in \overline{E}. \tag{1}$$

Recall that, given a Markov process X with transition function (P_t), a Borel function $f : E \to [0, +\infty]$ is called *excessive* for X if it satisfies

$$P_t f \leq f \text{ and } \lim_{t \downarrow 0} P_t f = f.$$

It is called *minimal* if, whenever g is excessive and satisfies $g \leq f$, there exists a constant c such that $g = c \cdot f$. For $\overline{y} \in \overline{E}$, the function $\overline{K}(\cdot; \overline{y})$ is minimal excessive for the spacetime process $\overline{X}$ (see Meyer, 1968).

Let now $\eta : \overline{E} \to \mathbb{R}$ be any $dt \otimes dm$ integrable function. We define a metric on $\overline{E}$ by

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

We shall denote by $\overline{F}$ the completion of $\overline{E}$ by d. It is well known (e.g. Doob, 1984), p.197, or Meyer, 1968) that $\overline{F}$ is compact and that $\overline{K}(\overline{x},\cdot)$ has a continuous extension to $\overline{F}$, for each $\overline{x}\in\overline{E}$. The continuity of $\overline{K}(\overline{x},\cdot)$ in the original topology, which comes from the continuity of $p_t(x,y)$, implies that the original and Martin topologies coincide on $\overline{E}$. For each point $\overline{y}$ of the (spacetime) *Martin boundary* $\partial\overline{F}=\overline{F}\backslash\overline{E}$, the function $\overline{K}(\cdot,\overline{y})$ is spacetime excessive; however, it may not always be minimal.

We shall be mainly interested in the part $\overline{F}_{-\infty}$ of the boundary which consists of points $\overline{y}=\lim_{n\to\infty}(s_n,y_n)$ where $y_n\in E$ and $\inf_n s_n=-\infty$. There is a close connection between these points and the *Strong Ratio Limit Property* (SRLP), which states that there exist positive functions φ, φ^* and a real number $\lambda\geq 0$ such that (see for example Anderson, 1991)

$$\lim_{t\to\infty}\frac{p_{t+s}(x,y)}{p_t(x_0,y_0)}=e^{-\lambda t}\,\frac{\varphi(x)\varphi^*(y)}{\varphi(x_0)\varphi^*(y_0)},\quad x,y\in E. \tag{2}$$

Indeed, if we choose ν in (1) as the point mass at x_0, then the SRLP implies that there exists a single point $\overline{z}\in\overline{F}_{-\infty}$ characterized by the two properties

(SRLP1): Any sequence of the form $\overline{y}_n=(s_n,y)$ with $s_n\to-\infty$ converges to $\overline{z}$ in the Martin topology, i.e. $d(\overline{y}_n,\overline{z})\to 0$,

(SRLP2): $\overline{K}(t,x;\overline{z})=e^{-\lambda t}\varphi(x)\varphi^*(y)/\varphi(x_0)\varphi^*(y_0)$.

Note that (SRLP1) implies "half" of (SRLP2):

Proposition 1. *If (SRLP1) holds, then there exist λ and φ such that*

$$\overline{K}(t,x;\overline{z})=e^{-\lambda t}\varphi(x)/\varphi(x_0). \tag{3}$$

Proof. By continuity of $\overline{K}$, we have

$$\overline{K}(t,x;\overline{z})=\lim_{s\to\infty}\frac{p_{t+s}(x,y)}{p_s(x_0,y)},\quad y\in E, t>0.$$

Choosing $t=u+v$ and $t=u$, we find that

$$\overline{K}(u+v,x;\overline{z})=L(u)\overline{K}(v,x;\overline{z}),\text{ where } L(u)=\lim_{s\to\infty}\frac{p_{u+v+s}(x_0,y)}{p_{v+s}(x_0,y)}.$$

It is easy to see that $L(u+u')=L(u)L(u')$ and $L(0)=1$ whence it follows that $L(u)=e^{-\lambda u}$ for some λ. Then since $\overline{K}(0,x_0,\overline{z})=1$ we get (3). □

The "first half" (SRLP1) will later be shown to imply the existence of a suitable conditioned process. The full SRLP generally doesn't follow from (SRLP1). Indeed, notice that (2) is preserved by the transformation $p_t(x, y) \mapsto p_t(y, x)$, and hence by time reversal of X. To get the full SRLP, we would require that (SRLP1) also hold with $\overline{F}$ replaced by the corresponding Martin boundary of the reversed process $\widehat{X}_t = X_{\zeta - t}$.

3. Integral representation

In this section, we take up a study of the class of functions $h_s : \mathbb{R} \times E \to \mathbb{R}$ defined as follows:

$$h_s(t, x) = \begin{cases} \frac{\mathbb{P}_x(\zeta > s+t)}{\mathbb{P}_\nu(\zeta > s)} & x \in E, s > 0, t \geq -s \\ \mathbb{P}_\nu(\zeta > s)^{-1} & x \in E, s > 0, t < -s. \end{cases} \tag{4}$$

We shall only be interested in the values of this function when restricted to $\overline{E}$. Note that

$$\int h_s(0, x)\nu(dx) = 1, \quad s > 0. \tag{5}$$

Moreover, the function h_s is spacetime excessive, as can be easily checked. To describe h_s, we shall use the following concept.

Definition 2. *A subset $N \subset E$ is called a cemetery neighbourhood if*

$$\lim_{t \to \zeta} 1_N(X_t) = 1 \quad a.s.$$

for all starting points $x \in E$.

There always exists at least one cemetery neighbourhood, namely $N = E$. The definition states that the process spends the last segment of its lifetime in N.

Example 3. Suppose X is a Brownian motion killed upon leaving the unit ball $E = \{x : \|x\|^2 < 1\}$. On account of the continuity of sample paths, any annulus $N = \{x : \epsilon < \|x\|^2 < 1\}$ is a cemetery neighbourhood.

Example 4. Let B be a Brownian motion on $E = \{x : \|x\|^2 < 1\}$ as in the previous example, and let $c(x) > 0$ be a bounded function on E. Define a Markov process X by killing B according to the additive functional $A_t = \int_0^t c(B_s)ds$. The semigroup of X is given by the formula

$$\mathbb{E}_x(f(X_t), \zeta > t) = \mathbb{E}_x(f(B_t)e^{-\int_0^t c(B_s)ds}, T_{\partial E} > t), \quad x \in E,$$

here $T_{\partial E} = \inf\{t > 0 : |B_t| = 1\}$. If U is any nonempty open subset of E, the probability that X_t belongs to U at its moment of death is strictly positive. Hence the only cemetery neighbourhood is $N = E$.

Example 5. Let Y be a Markov chain on $E = \{0, 1, 2, 3, \dots\}$ and suppose that Y gets absorbed in 0 in a finite time. Define X as the Markov chain on $E = \{1, 2, 3, \dots\}$ which is constructed by killing Y at the first hitting time of state 0. A cemetery neighbourhood is given by the set of states from which the process Y can directly jump to zero, namely

$$N = \{y > 0 : q_{y0} > 0\};$$

here (q_{ij}) is the q-matrix of Y. The set we defined is clearly the smallest possible cemetery neighbourhood.

Example 6. Let X be an explosive pure birth process on the countable set $E = \{1, 2, 3, \dots\}$. The lifetime of X coincides with the explosion time, that is

$$\zeta = \inf\{t > 0 : |X_t| = \infty\}.$$

A typical cemetery neighbourhood is given by $N = \{n, n + 1, n + 2, \dots\}$. There exists no 'smallest' such set.

We resume the study of h_s. Let $\dagger(X)$ denote the class of all cemetery neighbourhoods of X. We put

$$\overline{N}_{-s} = d\text{-closure of } (-\infty, -s] \times N, \quad \overline{N}_{-\infty} = \cap_{s>o}\overline{N}_{-s}.$$

Lemma 7. *For every $s > 0$, there exists a probability measure μ_s, concentrated on*

$$\mathfrak{N}_{-s}(X) = \cap\{\overline{N}_{-s} : N \in \dagger(X)\},$$

such that

$$h_s(t, x) = \int \overline{K}(t, x; \overline{y})\mu_s(d\overline{y}). \tag{6}$$

Proof. Since the function h_s is excessive for $\overline{X}$ and satisfies (5), the standard representation theory for normalized excessive functions (see Meyer, 1968) asserts the existence of a probability μ_s on $\overline{F}$ such that (6) holds. This measure can always be chosen such that

$$h_s(\overline{x})\overline{\mathbb{P}}^{h_s}_{\overline{x}}(\overline{X}_{\overline{\zeta}-} \in A) = \int_A \overline{K}(\overline{x},\overline{y})\mu_s(d\overline{y}), \quad A \subset \overline{F}, \overline{x} \in \overline{E}. \tag{7}$$

Here $\overline{\mathbb{P}}^{h_s}$ is the law of the h_s-transform of $\overline{X}$, i.e. the Markov process on $\overline{E}$ with transition function $(\overline{P}^{h_s}_t)$ given by

$$\overline{P}^{h_s}_t(\overline{x}, d\overline{y}) = h_s(\overline{x})^{-1}\overline{P}_t(\overline{x}, d\overline{y})h_s(\overline{y}),$$

and where $(\overline{P}_t)$ is the transition function of $\overline{X}$. It remains only to show that $\mu_s(\overline{N}_{-s}) = 1$ for every cemetery neighbourhood $N \in \dagger(X)$. Now for any $\overline{x} = (t, x) \in \overline{E}$,

$$\begin{aligned}\overline{\mathbb{P}}^{h_s}_{\overline{x}}(\overline{\zeta} > s) &= h_s(t,x)^{-1}\mathbb{E}_x[h_s(t-s, X_s), \zeta > s] \\ &= \mathbb{E}_x[\mathbb{P}_{X_s}(\zeta > t), \zeta > s]/\mathbb{P}_x(\zeta > t+s) = 1.\end{aligned}$$

Moreover,

$$\begin{aligned}\lim_{u\to\infty} \overline{\mathbb{P}}^{h_s}_{\overline{x}}(\overline{\zeta} > u) &= \lim_{u\to\infty} \mathbb{E}_x[h_s(t-u, X_u), \zeta > u]/h_s(t,x) \\ &= \lim_{u\to\infty} \mathbb{P}_x(\zeta > u)/\mathbb{P}_x(\zeta > t+s) = 0,\end{aligned}$$

and we deduce that $\overline{\mathbb{P}}^{h_s}_{\overline{x}}(-s \leq \overline{\zeta} < \infty) = 1$. Now a rephrasing of Definition 2 in terms of the spacetime process $\overline{X}$ gives

$$\overline{\mathbb{P}}_{(t,x)}(\overline{X}_{\overline{\zeta}-} \in \overline{N}_{-t}, \overline{\zeta} < \infty) = 1.$$

Hence for $r < -s$,

$$\begin{aligned}\overline{\mathbb{P}}^{h_s}_{\overline{x}}(\overline{X}_{\overline{\zeta}-} \in \overline{N}_{-s}, \overline{\zeta} < \infty) &= \overline{\mathbb{P}}^{h_s}_{\overline{x}}(\overline{X}_{\overline{\zeta}-} \in \overline{N}_{-s}, r < \overline{\zeta} < \infty) \\ &= \overline{\mathbb{E}}_{\overline{x}}[h_s(\overline{X}_r), \overline{X}_{\overline{\zeta}-} \in \overline{N}, r < \overline{\zeta} < \infty]/h_s(\overline{x}) \\ &= \overline{\mathbb{E}}_{\overline{x}}[h_s(\overline{X}_r), \overline{\zeta} > r]/h_s(\overline{x}) = 1.\end{aligned}$$

Thus by (7), taking $A = \overline{N}_{-s}$, we find

$$h_s(\overline{x}) = \int_{\overline{N}_{-s}} \overline{K}(\overline{x},\overline{y})\mu_s(d\overline{y}).$$

Now (5) and Fubini's theorem gives

$$1 = \int h_s(0,x)\nu(dx) = \int_{\overline{N}_{-s}} \Big(\int \overline{K}(0,x;\overline{y})\nu(dx) \Big) \mu_s(d\overline{y})$$
$$\leq \mu_s(\overline{N}_{-s}).$$

This completes the proof, as N was arbitrary. □

Recall that $\overline{F}$ is compact. This implies that the family of measures (μ_s) above has at least one weak limit point, $\mu_{s(k)} \Rightarrow \mu$ say. If $N \in \dagger(X)$, then it follows that

$$1 = \lim_{k\to\infty} \mu_{s(k)}(\overline{N}_{-r}) \leq \mu(\overline{N}_{-r}), \quad r > 0,$$

and μ must therefore be concentrated on the set $\mathfrak{N}_{-\infty}(X)$. Let

$$h(t,x) = \int \overline{K}(t,x;\overline{y})\mu(d\overline{y}). \tag{8}$$

Since $\overline{K}(t,x;\cdot)$ is continuous on the compact set $\overline{N}_{-t}$, it follows by weak convergence and (6) that $h_{s(k)}(t,x) \to h(t,x)$.

The function h is excessive (for $\overline{X}$), but we have as yet no way of checking that $h \not\equiv 0$. Indeed, whereas each function h_s satisfies (5), Fatou's lemma only shows that $\int h(0,x)\nu(dx) \leq 1$.

To see that this is a real possibility, suppose that ν is a quasistationary distribution for X. Thus there exists a real number $\lambda \geq 0$ such that $\mathbb{P}_\nu(\zeta > t) = e^{-\lambda t}$. Then clearly

$$h_s(t,x) = e^{\lambda s}\mathbb{P}_x(\zeta > t+s).$$

If X is λ-transient, this tends to zero as $s \to \infty$, giving $h \equiv 0$.

4. Parabolic functions and further assumptions

To get a sufficient condition for $h \not\equiv 0$, we shall need to discuss parabolic functions.

Definition 8. *A locally bounded Borel function f is said to belong to the domain of the local martingale generator $\mathfrak{A}$ of X if there exists a Borel function $g(x) =: \mathfrak{A}f(x)$ such that the process*

$$M_t^f = f(X_t)1_{(\zeta>t)} - f(X_0) - \int_0^{t\wedge\zeta} \mathfrak{A}f(X_s)ds$$

is, for each $(\Omega, (\mathcal{F}_t), \mathbb{P}_x)$*, a right continuous local martingale up to* ζ *in the following sense: there exists a sequence of stopping times* $T_n \uparrow \zeta$ *such that* $M^f_{t\wedge T_n}$ *is a* $\mathbb{P}_x$ *martingale for each* $x \in E$*.*

Note that the operator $\mathfrak{A}$ is in general multivalued, as $\mathfrak{A}f$ can be arbitrarily modified on a set of potential zero. Two standard examples of generators of this type are as follows:

- If X is a diffusion on a subset of $\mathbb{R}^d$, then by Ito's formula we have $\mathfrak{A} = L$, a second order differential operator. Its domain contains all C^2 (not necessarily bounded) functions.
- If X is a Markov chain on $E = \{1, 2, 3, \dots\}$ then $\mathfrak{A} = Q$, where $Q = (q_{ij})$ is the q-matrix of X. Here the domain of $\mathfrak{A}$ includes all functions $f : E \to \mathbb{R}$ such that $\mathfrak{A}f(i) = \sum_j q_{ij} f(j) < \infty$ for all $i \in E$.

A function $u : (a, b) \times E \to [0, +\infty]$ is called *parabolic* (for X) if

$$\frac{\partial}{\partial t} u(t, x) = \mathfrak{A}u(t, x) \text{ in } (a, b) \times E$$

A standard example of a parabolic function in $(0, \infty) \times E$ which satsifies $u(0, \cdot) = f$ is the function

$$u(t, x) = \mathbb{E}_x[f(X_t), \zeta > t], \quad t \geq 0, x \in E.$$

This immediately shows that the functions h_s of (4) are parabolic in $(-s, \infty) \times E$. One may wonder if the limit functions h of the previous section are therefore parabolic in $(-\infty, \infty) \times E$. This will be the case under the following assumptions:

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 (A2) holds for many processes, including diffusions on regular domains and all minimal Markov chains on a countable state space (for the latter, see Breyer, 1998a). Assumption (A3) is a bounded jump condition; it clearly holds for diffusions since their sample paths are continuous, but (A3) is also satisfied by any Markov chain on countable state space whose q-matrix is such that for fixed x, q_{xy} is nonzero for only a finite number of entries. In that case we can take

$$K' = \{y : q_{xy} \neq 0, x \in K\}.$$

Lemma 9. *Let (A1)-(A3) hold and suppose that the probability* ν *is compactly supported. Then every limit point* h *of the family* (h_s)*, as* $s \to \infty$*, is nonzero, parabolic in* $(-\infty, \infty) \times E$*, and satisfies* $\int h(0,x)\nu(dx) = 1$*. Moreover,* $t \mapsto h(t,x)$ *is decreasing.*

Proof. From (Breyer, 1998a), it follows that every limit function h is parabolic in $(-\infty, \infty) \times E$. The functions $h(t,x)$ are decreasing in t because each function $h_s(t,x)$ is. In particular then, by (A2), there exists a constant C such that

$$\sup\{h_s(0,x) : s > 0, x \in \operatorname{supp}(\nu)\} \leq C.$$

Thus by dominated convergence , we have

$$\int h(0,x)\nu(dx) = \lim_{s\to\infty} \int h_s(0,x)\nu(dx) = 1.$$

A posteriori, this implies that h is nonzero. □

5. Minimal parabolic functions in the set $\mathfrak{N}_{-\infty}(X)$

We will now characterize the limit functions h in terms of eigenfunctions of $\mathfrak{A}$. Before stating the main application (Lemma 14), we need some preliminary technical results.

Assumption (A4): There exists a cemetery neighbourhood N with the following property: for each $r > 0$, there exist $C(r), T(x) > 0$ such that

$$p_t(x,y) \leq C(r)p_{t+r}(x,z), \quad t > T(x), \quad x \in E, \quad y,z \in N.$$

Examples of processes satisfying (A4) will be given below.

The following is an adaptation in our context of a result for diffusions due to Koranyi and Taylor (1985).

Proposition 10. *Let (A4) hold for some $N \in \dagger(X)$, and suppose that $\overline{y} \in \overline{N}_{-\infty}$. If $K(\cdot, \overline{y}) \not\equiv 0$ is parabolic, then*

$$\overline{K}(t, x; \overline{y}) = e^{-\lambda t} g(x),$$

where g is a minimal positive solution to the equation $\mathfrak{A} g = -\lambda g$ in E.

Proof. By the assumption, if $s < t \le 0 < r$, $s < -T(x)$,

$$\begin{aligned} \overline{K}(t-r, x; s, y) &= \frac{p_{t-r-s}(x,y)}{\int \nu(dx) p_{-s}(x,y)} \\ &\le C(r) \cdot \frac{p_{t-s}(x,y)}{\int \nu(dx) p_{-s}(x,y)} \\ &= C(r) \overline{K}(t, x; s, y). \end{aligned}$$

Suppose now that $w(\overline{x}) = \lim_{n\to\infty} \overline{K}(\overline{x}; \overline{y}_n)$ is a minimal parabolic function corresponding to a sequence $\overline{y}_n = (s_n, y_n)$ satisfying $\lim_n s_n = -\infty$ and $y_n \in N$ for all n. For any $r > 0$, the function $w_r(t,x) = w(t-r, x)$ is also parabolic in $(-\infty, 0] \times E$, and the above computation shows that $w_r(t,x) = w(t-r,x) \le C(r) w(t,x)$. By minimality of w (see Section 2), this means there exists a constant $L(r)$ such that $w(t,x) = L(r) w(t-r, x)$. Now L satisfies $L(a+b) = L(a)L(b)$ and $L(0) = 1$. Moreover, L is continuous since $t \mapsto w(t,x)$ is. Hence $L(t) = e^{-\lambda t}$ for some constant $\lambda \in \mathbb{R}$, and then $w(t,x) = e^{\lambda t} w(0,x)$. Finally, since w is parabolic, it satisfies the equation

$$\frac{\partial}{\partial t} w(t,x) = \mathfrak{A} w(t,x) \qquad \text{in } (-\infty, 0] \times E,$$

and hence the function $g(x) = w(0,x)$ is an eigenfunction of $\mathfrak{A}$ with eigenvalue $-\lambda$. For the minimality, put $g = k + l$ where both k and l are positive eigenfunctions with eigenvalue $-\lambda$. The functions $k'(t,x) = e^{-\lambda t} k(x)$ and $l'(t,x) = e^{-\lambda t} l(x)$ are both parabolic, and dominated by w; hence they are constant multiples of w, and multiplying by $e^{\lambda t}$ shows that k and l are both multiples of g. □

Note that in the above, we do not require the full force of (A4), but merely that $p_t(x,y) \le C(r) \cdot p_{t+r}(x,y)$ for all $y \in N$. We now give some examples of processes satisfying Assumption (A4).

Example 11. Let X be a uniformly elliptic diffusion (bounded coefficients) on a bounded open set E with regular boundary ∂E; the transition density $p_t(x,y)$

(with respect to Lebesgue measure) is the fundamental solution of the parabolic operator $\mathfrak{A} - \partial/\partial t$, where

$$\mathfrak{A}f(x) = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(x)\frac{\partial^2}{\partial x^i \partial x^j}f(x) + \sum_{j=1}^{d} b_j(x)\frac{\partial}{\partial x^j}f(x), \quad f \in C^2(E)$$

The regularity of the boundary ensures that $\lim_{x\to\partial E} p_t(x,y) = 0$. The function $(t,x) \mapsto p_t(x,y)$ is of course parabolic in $(0,\infty)\times E$. If we let

$$\widehat{\mathfrak{A}}f(x) = \frac{1}{2}\sum_{i,j=1}^{d} D_{ij}(a_{ij}f)(x) - \sum_{j=1}^{d} D_j(b_j f)(x)$$

be the formal adjoint of $\mathfrak{A}$, then $(t,y) \mapsto p_t(x,y)$ is parabolic for the operator $\widehat{\mathfrak{A}}$. With a suitable choice of coefficients, the parabolic Harnack inequality will hold for $\widehat{\mathfrak{A}}$ with K the closure of D (see Friedman, 1964), and applying this to $(t,y) \mapsto p_t(x,y)$ gives Assumption (A4).

The above example can be modified for diffusions on unbounded domains E, provided a sufficiently small cemetery neighbourhood exists. The following example illustrates the procedure.

Example 12. Let X be Brownian motion on $E = (0,\infty)$, killed upon first hitting zero. Its generator is $(1/2)d^2/dx^2$ on $C_K^2((0,\infty))$. The set $(0,1)$ is a cemetery neighbourhood, and the transition function of X is,

$$p_t(x,y) = \sqrt{2/\pi t}\exp\left(-\frac{x^2+y^2}{2t}\right)\sinh(xy/t), \quad x,y,t > 0.$$

Since

$$\frac{p_t(x,y)}{p_{t+r}(x,y)} = \sqrt{\frac{t+r}{t}}\cdot\frac{\sinh(xy/t)}{\sinh(xy/t+r)},$$

we can take $C(r) = 4$, provided we take

$$T(x,r) = r \vee \inf\left\{t > 0 : \frac{\sinh(x/t)}{\sinh(x/t+r)} \le 2\right\}.$$

It is straightforward to generalize these two examples to the case when X is a uniformly elliptic diffusion on $\mathbb{R}^d$, when the lifetime coincides with the hitting time of some compact set with regular boundary. We give one further example, within the realm of Markov chains.

Example 13. Let X be an irreducible Markov chain on a countable state space, and suppose that a finite cemetery neighbourhood N exists. Assumption (A4) must hold, since if $p_t(x,y)$ is the density of the transition function with respect to counting measure,

$$p_{t+r}(x,y) \geq p_t(x,z)p_r(z,y), \quad x \in E, \quad y,z \in N,$$

and we can take $C(r)^{-1} = \min_{z,y\in N} p_r(z,y) > 0$ by irreducibility. This argument also shows, by taking $y = z$, that the conclusion of Proposition 10 holds as soon as there exists a (not necessarily finite) cemetery neighbourhood N satisfying

$$\sup_{y\in N} |q(y)| < +\infty.$$

In particular, this is always true when the q-matrix is bounded.

We resume our study of the family (h_s).

Lemma 14. *Let Assumptions (A1)-(A4) hold, and suppose that the probability measure ν is compactly supported. If h is a limit point of (h_s), then there exists a probability measure ξ and a family of functions $(g_\lambda : \lambda \geq 0)$ such that $\mathfrak{A} g_\lambda = -\lambda g_\lambda$, $\langle \nu, g_\lambda\rangle = 1$, and*

$$h(t,x) = \int e^{-\lambda t} g_\lambda(x)\xi(d\lambda). \tag{9}$$

Proof. Let h be a limit point of (h_s). It can therefore be represented, via (8), by some probability μ on $\overline{F}$ satisfying

$$\mu\Big(\cap\{\overline{N}_{-\infty} : N \in \dagger(X)\}\Big) = 1. \tag{10}$$

The functions $K(\cdot,\overline{y})$ with $\overline{y}$ in the support of μ may not all be minimal, however it is always possible to represent h by another measure μ' which *is* concentrated on minimal functions $K(\cdot,\overline{z})$, as follows (Meyer, 1968):

$$h(\overline{x})\overline{\mathbb{P}}^h_{\overline{x}}(\overline{X}_{\overline{\zeta}-} \in A) = \int_A \overline{K}(\overline{x};\overline{z})\mu'(d\overline{z}).$$

Since h is parabolic, the same is true for the functions $\overline{K}(\cdot;\overline{z})$ in this representation. We will show that μ' also satisfies (10), which will establish the existence of minimal parabolic functions of the form $\overline{K}(\cdot;\overline{z})$ with $\overline{z} \in \overline{N}_{-\infty}$, and then, via Proposition 10, the representation (9). Now consider the event $\Lambda_N = \{\overline{X}_{\overline{\zeta}-} \in \overline{N}_{-\infty}\}$, where

$N \in \dagger(X)$. We have $\overline{\mathbb{P}}^h_{\overline{x}}(\Lambda_N) = 1$, for if Γ is any event in $\overline{\mathcal{F}}_r$, $r > 0$, then

$$\begin{aligned}\overline{\mathbb{P}}^h_{\overline{x}}(\Gamma \cap \Lambda_N) &= \overline{\mathbb{E}}_{\overline{x}}[h(\overline{X}_r), \Gamma \cap \Lambda_N, \overline{\zeta} > r]/h(\overline{x}) \\ &= \overline{\mathbb{E}}_{\overline{x}}[h(\overline{X}_r), \Gamma, \overline{\zeta} > r]/h(\overline{x}) \\ &= \overline{\mathbb{P}}^h_{\overline{x}}(\Gamma),\end{aligned}$$

by Definition 2. Thus we have

$$1 = \overline{\mathbb{P}}^h_{\overline{x}}(\Lambda_N) = \int \overline{\mathbb{P}}^{\overline{K}(\cdot;\overline{z})}_{\overline{x}}(\Lambda_N)\mu'(d\overline{z}),$$

which gives $\overline{\mathbb{P}}^{\overline{K}(\cdot;\overline{z})}_{\overline{x}}(\Lambda_N) = 1$ for almost every $\overline{z}$ in the support of μ'. Now (7) holds also when h_s is replaced by $\overline{K}(\cdot;\overline{z})$. Taking $A = \overline{N}_{-\infty}$ in this equation, we get that the spacetime excessive function $\overline{K}(\cdot;\overline{z})$ is represented by a measure on $\overline{N}_{-\infty}$. This is of course (by minimality) the point mass at $\overline{z}$, establishing that $\overline{z} \in \overline{N}_{-\infty}$. Repeating this argument for almost every $\overline{z}$ in the support of μ' gives (10) for μ'. The fact that $\lambda \geq 0$ follows because $t \mapsto h(t,x)$ is decreasing, by Lemma 9. The normalization $\langle \nu, g_\lambda \rangle = 1$ is an easy consequence of $\langle \nu, h(0,\cdot)\rangle = 1$ and Fubini's theorem applied to (9). □

It is important to realize that the fact that h is parabolic can be made stronger. Recall from the proof of Lemma 7 that $\mathbb{P}^{h_s}_{\overline{x}}(\overline{\zeta} \geq s) = 1$. Hence for $s > r$, $\overline{\mathbb{E}}_{\overline{x}}[h_s(\overline{X}_r), \overline{\zeta} \leq r] = 0$ and Fatou's lemma gives $\overline{\mathbb{E}}_{\overline{x}}[h(\overline{X}_r), \overline{\zeta} \leq r] = 0$, so that

$$\overline{\mathbb{P}}^h_{\overline{x}}[\overline{\zeta} < \infty] = 0.$$

This last statement is equivalent to the function $h(t,x)$ being *invariant* for $\overline{X}$, i.e. $\overline{P}_t h = h$. Equivalently, almost every function g_λ entering in the representation (9) must satisfy $P_t g_\lambda = e^{-\lambda t} g_\lambda$.

6. When is $\mathfrak{N}_{-\infty}(X)$ a singleton?

In the course of the proof of Lemma 14, it was shown that $\mathfrak{N}_{-\infty}(X)$ always contains a minimal parabolic point $\overline{z}_0$. In this section, we investigate cases when the set is in fact a singleton. In such a case, we can then assert by Lemmas 9 and 14 that

$$\lim_{s\to\infty} \frac{\mathbb{P}_x(\zeta > t+s)}{\mathbb{P}_\nu(\zeta > s)} = \overline{K}(t,x;\overline{z}_0) = e^{-\lambda t} g_\lambda(x). \tag{11}$$

We begin with a general result:

Proposition 15. *Under Assumptions (A1)-(A4), the following three statements are equivalent:*

(i) *For some $y \in N$,*

$$\underline{\lim}_{s\to-\infty}\overline{K}(\overline{x};s,y) > 0$$

holds for all $\overline{x}$ in a set of positive $dt \otimes dm$-measure,

(ii) *The set $\overline{N}_{-\infty}$ contains a single nontrivial, minimal parabolic boundary point $\overline{z}_0$,*

(iii) *For every sequence $\overline{y}_n = (s_n, y_n)$ such that $s_n \to -\infty$ and $y_n \in N$ for all n,*

$$\lim_{n\to\infty}\overline{K}(\overline{x};\overline{y}_n) = \overline{K}(\overline{x},\overline{z}_0) = e^{-\lambda t}g(x).$$

Proof. (i) $\Rightarrow$ (ii): By Assumption (A4), if $\epsilon > 0$ is fixed, we have the inequality

$$\begin{aligned}\overline{K}(t,x;s,y) &= \frac{p_{t-s}(x,y)}{\int\nu(dw)p_{-s}(w,y)}\\ &\le C(\epsilon)^{-2}\frac{p_{t-s+\epsilon}(x,z)}{\int\nu(dw)p_{-s-\epsilon}(w,z)}\\ &= C(\epsilon)^{-2}\overline{K}(t+2\epsilon,x;s+\epsilon,z),\end{aligned}$$

for all $y,z \in N$ and $t-s, -s > T(x)$. Now let $(s_n+\epsilon, z_n) \to \overline{z}$, where $\overline{z}$ is any nonzero minimal parabolic boundary point belonging to $\overline{N}_{-\infty}$. The function $k(\overline{x}) = \underline{\lim}_{s\to-\infty}\overline{K}(\overline{x};s,y)$ satisfies $\overline{P}_r k \le k$ by Fatou's lemma. It is thus bounded below by the excessive function $\widetilde{k} = \lim_{r\to 0}\overline{P}_r k$, from which it differs at most on a set of $dt\otimes dm$-measure zero; the minimality of $\overline{z}$ together with the bound $k(t-2\epsilon,x) \le C(\epsilon)^{-2}\overline{K}(t,x;\overline{z})$ implies that $\widetilde{k}(t-2\epsilon,x)$ is a constant multiple of $\overline{K}(t,x;\overline{z})$. Since $\widetilde{k}$ is necessarily parabolic, we have $\widetilde{k}(t,x) = e^{c\epsilon}k(t-2\epsilon,x)$ for some constant c, whence $\widetilde{k}$ is also a multiple of $\overline{K}(\cdot;\overline{z})$. If $\overline{z}'$ is another minimal parabolic boundary point, the same argument shows that $\overline{K}(\cdot;\overline{z}')$ is a constant multiple of $\widetilde{k}$, and hence of $\overline{K}(\cdot,\overline{z})$. Thus there exists at most one mimimal point in $\overline{N}_{-\infty}$, and hence all sequences converge to it.

(ii) $\Rightarrow$ (i): Conversely, if there are two distinct minimal points $\overline{z},\overline{z}' \in \overline{N}_{-\infty}$, then $\widetilde{k}(\cdot)$ must be proportional to both $\overline{K}(\cdot;\overline{z})$ and $\overline{K}(\cdot;\overline{z}')$, and hence identically zero. $\square$

Here are now some examples of processes.

Example 16. Let X denote the one dimensional Brownian motion on $E = (0, \infty)$, killed upon first hitting zero. We saw earlier that this process satisfies (A1)-(A4). Since its generator is given by $\mathfrak{A} = (1/2)d^2/dx^2$, positive solutions to $\mathfrak{A}g = -\lambda g$ exist only for $\lambda \le 0$. The solutions to $\mathfrak{A}g = 0$ are given by $g(x) = a + bx$, for some positive constants a and b. If $b = 0$, g is a multiple of the excessive function 1, which is not invariant for X since the process has finite lifetime under the h-transformed law $\mathbb{P}^1_x = \mathbb{P}_x$. The Martin compactification of X is well known to be $[0, +\infty]$, and the excessive function 1 corresponds here to the point 0. If $a = 0$, then g is a multiple of the excessive function x. Now this function is invariant, since the corresponding h-transform (with $h(x) = x$) is a three dimensional Bessel process with generator

$$\mathfrak{A}^h f(x) = h(x)^{-1}\mathfrak{A}(hf)(x) = \frac{1}{2}\frac{d^2}{dx^2}f(x) + \frac{1}{x}\frac{d}{dx}f(x),$$

and it is well known that this process is nonexplosive. Thus we see that $\mathfrak{N}_{-\infty}(X)$ consists of only the point $\overline{z}_0$ for which $\overline{K}(t, x; \overline{z}_0) = x/\langle \nu, x\rangle$. In terms of the Martin boundary of X, the excessive function x corresponds to the point $+\infty$. We end this example with some further remarks on the parabolic Martin boundary, taken from Doob (1984, p. 375). For every $\tau < 0$, set

$$\mathcal{K}_0(t, x; \tau) \propto \begin{cases} \frac{x}{\sqrt{2\pi(t-\tau)^3}} \exp\left(-\frac{x^2}{2(t-\tau)}\right) & \text{if } t > \tau, \\ 0 & \text{if } t \le \tau, \end{cases} \quad , \quad \int \mathcal{K}_0(0, x; \tau)\nu(dx) = 1.$$

For each $\gamma \le 0$, set

$$\mathcal{K}_1(t, x; \gamma) \propto \begin{cases} \sinh(-\gamma x) \exp\left(\frac{\gamma^2 t}{2}\right) & \text{if } \gamma < 0, \\ x & \text{if } \gamma = 0, \end{cases} \quad , \quad \int \mathcal{K}_1(0, x; \gamma)\nu(dx) = 1.$$

The Martin sequences are as follows: if $\overline{y}_n \to (\tau, 0)$, then $\lim_n \overline{K}(\overline{x}, \overline{y}_n) = \mathcal{K}_0(\overline{x}; \tau)$; if $s_n \to -\infty$ and $y_n/s_n \to \gamma \le 0$, then $\lim_n \overline{K}(\overline{x}, \overline{y}_n) = \mathcal{K}_1(\overline{x}, \gamma)$, and if either $y_n \to +\infty$ with $y_n/(1 + |s_n|) \to +\infty$ or else $s_n \to 0$ with no restriction on y_n, then $\lim_n \overline{K}(\overline{x}, \overline{y}_n) = 0$. Every positive parabolic function u with $\underline{\lim}_{s\to 0} u(s, x) < +\infty$ then has the Martin representation

$$u(t, x) = \int \mathcal{K}_0(t, x; \tau)\xi_0(d\tau) + \int \mathcal{K}_1(t, x; \gamma)\xi_1(d\gamma).$$

For the function which interests us, namely $h_s(t,x)$, the quickest way to get this representation explicitly is to use the Bachelier-Lévy formula

$$\mathbb{P}_x(\zeta > r) = \int_r^\infty \frac{x}{\sqrt{2\pi u^3}} \exp\bigl(-x^2/2u\bigr) du,$$

and make the change of variable $\tau = r - u$. One then finds that

$$\begin{aligned} \frac{\mathbb{P}_x(\zeta > t+s)}{\mathbb{P}_\nu(\zeta > s)} &= \int_{-\infty}^s \mathcal{K}_0(t,x;\tau)\biggl(\int \nu(dz) \frac{z}{\sqrt{-2\pi\tau^3}} \exp(z^2/2\tau) d\tau\biggr) \\ &= \int \mathcal{K}_0(t,x;\tau) 1_{(-\infty,s]}(\tau) \mu_s(d\tau), \end{aligned}$$

and a representing probability measure μ_s is concentrated on $\overline{E}_{-s}$ as predicted. The set $\overline{E}_{-\infty}$ here consists of the half-line $\gamma \leq 0$, where each point γ is identified with the function $\mathcal{K}_1(\cdot;\gamma)$. A cemetery neighbourhood is given by the set $(0,1) \subset E$. Now the points belonging to $\overline{N}_{-\infty}$ must be arrived at through sequences (s_n, y_n) such that $y_n < 1$ for all n. In view of the characterization of Martin sequences above, every such sequence must give the function $\mathcal{K}_1(t,x;0) = x/\langle \nu, x\rangle$.

In the example above, the fact that $\mathfrak{A}g \leq -\lambda g$ has positive solutions only when $\lambda \leq 0$ is a consequence of the fact that the Brownian motion X has *zero decay parameter* and in fact is *zero-transient*. (for the theory of decay parameters, see Tuominen and Tweedie, 1979). Clearly, when a Markov process X has this property, we can always identify the set $\mathfrak{N}_{-\infty}$, which belongs to the spacetime Martin boundary of X, with a subset of the ordinary Martin boundary of X, as we have done above, where $\mathfrak{N}_{-\infty}$ was identified with the point $+\infty$ of $[0,+\infty]$.

Example 17. Suppose that X is positive λ-recurrent (definition in Tuominen and Tweedie, 1979; see also Anderson, 1991, for chains). In this case, there exists a unique positive solution φ to the equation $\mathfrak{A}g \leq -\lambda g$, $\lambda \geq 0$, and in fact $P_t\varphi = e^{-\Lambda t}\varphi$, where Λ is the decay parameter of X (recall that X is the minimal process with generator $\mathfrak{A}$). Two examples of positive λ-recurrent processes are Markov chains on a finite set $E = \{1,\ldots,n\}$ and uniformly elliptic diffusions on a bounded regular domain (see Breyer, 1997, Chapter 3). Both these examples satisfy (A1)-(A4) with $N = E$, and we have

$$\mathfrak{N}_{-\infty}(X) = \overline{E}_{-\infty} = \{\overline{z}_0\}, \quad \overline{K}(t,x;\overline{z}_0) = e^{-\Lambda t}\varphi(x)/\langle \nu, \varphi\rangle.$$

Indeed, this follows by Proposition 15 (i), since the positive λ-recurrence implies that

$$\lim_{t\to\infty} e^{\Lambda t} p_t(x,y) = \varphi(x),$$

and so we have for $t < 0$:

$$\lim_{s\to-\infty} \overline{K}(t,x;s,y) = \lim_{s\to-\infty} \frac{e^{-\Lambda s} p_{t-s}(x,y)}{\int \nu(dw) e^{-\Lambda s} p_{-s}(w,y)} = e^{\Lambda t}\varphi(x)/\langle \nu, \varphi\rangle > 0.$$

Example 18. Let X be a uniformly elliptic diffusion on an open set $E \subset \mathbb{R}^d$, not necessarily bounded. Assume that the generator is in divergence form,

$$\mathfrak{A} f(x) = \sum_{i,j=1}^{d} \frac{\partial}{\partial x^i}\left(a_{ij}(x) \frac{\partial f}{\partial x^j}\right)(x), \quad f \in C^2(E).$$

Bass and Burdzy (1992) have shown that, if the set E is given locally by the graph of an L^p function with $p > d-1$, then the following parabolic boundary Harnack principle holds: for every $u > 0$, there exists $C(u)$ such that

$$\frac{p_a(y,x)}{p_a(z,x)} \geq C(u) \cdot \frac{p_b(y,v)}{p_b(z,v)}, \quad a,b \geq u, \tag{12}$$

for all $v,x,y,z \in E$. See their paper and references therein for a precise definition of L^p domains, and for other conditions ensuring the validity of (12).

Suppose now that a slightly weaker form of (12) holds, namely for v,x merely belonging to N. Integrating both sides of (12) over $y \in E$ with respect to the measure $\int \nu(dw) p_r(w,\cdot) m(\cdot)$ and inverting gives

$$\frac{p_a(z,x)}{\int \nu(dw) p_{a+r}(w,x)} \leq C(u) \cdot \frac{p_b(z,v)}{\int \nu(dw) p_{b+r}(w,v)}, \quad a,b \geq u; x,v \in N; z \in E.$$

Changing variables according to $t = -r$, $-s' = a + r$, $-s' = b + r$ we find $\overline{K}(t,z;s',x) \leq C(u)\overline{K}(t,z;s,v)$, and this is enough to guarantee that $\overline{N}_{-\infty}$ contains a single minimal boundary point. Indeed, let (s_n, z_n) converge to some minimal $\overline{z} \in \overline{N}_{-\infty}$, and let (s_n', y_n) converge to any parabolic point $\overline{y}$ in $\overline{E}_{-\infty}$ with the correct normalization. The inequality ensures that

$$e^{\lambda' t} g'(z) = \overline{K}(t,z;\overline{y}) \leq C(u) \cdot \overline{K}(t,z;\overline{z}) = e^{\lambda t} g(z),$$

and the minimality of $\overline{z}$ ensures that these two parabolic functions are proportional. Thus there exists a constant c such that $e^{\lambda' t} g'(z) = c \cdot e^{\lambda t} g(z)$. Integrating both sides with respect to $\nu(dz)$ gives $c = 1$, and hence $\overline{y} = \overline{z}$. Since $\overline{y}$ was arbitrary, the set $\overline{N}_{-\infty}$ must consist of a single minimal point.

Example 19. Suppose that X is a Markov chain on a countable state space E, whose transition function is symmetric, i.e. $p_t(x, y) = p_t(y, x)$ for the transition density associated with m in (A1). In dealing with Markov chains, it is usual to work with respect to counting measure, so that m has the representation $m(A) = \sum_{x \in A} m(x)$. Let us denote by $p_{xy}(t)$ the transition density of $P_t(x, dy)$ with respect to counting measure. We then have the formula $p_t(x, y) = p_{xy}(t)/m(y)$, so that the symmetry requirement becomes the formula

$$m(x)p_{xy}(t) = m(y)p_{yx}(t), \quad x, y \in E.$$

Kendall (1959) showed the existence, for each $x, y \in E$, of finite signed measures $\mu(x, y; d\lambda)$ on $[\Lambda, \infty)$ such that

$$p_t(x, y) = \int_{\Lambda}^{\infty} e^{-\lambda t} \mu(x, y; d\lambda). \tag{13}$$

It is also known that the measure $\mu_x(d\lambda) := \mu(x, x; d\lambda)$ is positive for $x \in E$, and that, for each $y \in E$, $\mu(x, y; d\lambda) << \mu_y(d\lambda)$, independently of $x \in E$. Take $N \in \dagger(X)$ and let us now fix once and for all a state $y_0 \in N$. We shall set ν equal to the point mass at y_0. Then the spacetime Martin kernel normalized by ν satisfies

$$\overline{K}(t, y_0; s, y_0) = \frac{\int_{\Lambda}^{\infty} e^{-\lambda(t-s)} \mu_{y_0}(d\lambda)}{\int_{\Lambda}^{\infty} e^{\lambda s} \mu_{y_0}(d\lambda)}, \quad s, t < 0.$$

We claim the following:

Proposition 20. *Let X be an irreducible Markov chain on a countable state space. We suppose that the transition function of X is symmetric, and that there exists a finite cemetery neighbourhood N. If ν is a point mass, then $\mathfrak{N}_{-\infty}(X) = \{\overline{z}_0\}$.*

Proof. Consider the probability measures on $[\Lambda, \infty)$ given by

$$\gamma_s(d\lambda) = \frac{e^{-\lambda s} \mu(d\lambda)}{\int_{\Lambda}^{\infty} e^{-\theta s} \mu(d\theta)}, \quad s > 0.$$

Using the bound $\int_{\Lambda}^{\infty} e^{-\lambda s} \mu(d\lambda) \geq e^{-(\Lambda+\epsilon)s} \mu((\Lambda, \Lambda + \epsilon])$, it follows that

$$\gamma_s([\Lambda + \epsilon, \infty)) \leq \mu([\Lambda, \Lambda + \epsilon))^{-1} e^{(\Lambda+\epsilon)s} \int_{\Lambda+\epsilon}^{\infty} e^{-\lambda s} \mu(d\lambda),$$

and this tends to zero as $s \to \infty$. Consequently, the measures γ_s are tight, and converge weakly to the point mass at Λ. If the normalizing measure is the point mass at y_0, it follows that

$$\lim_{s \to -\infty} \overline{K}(t, y_0; s, y_0) = \lim_{s \to \infty} \int e^{-\lambda t} \gamma_s(d\lambda)$$
$$= e^{-\Lambda t} > 0.$$

Since the set $(-\infty, 0) \times \{y_0\}$ clearly has positive $dt \otimes dm$-measure, Proposition 15 applies. □

By the parabolic Harnack inequality, the proof above can easily be modified to take into account the case of a measure ν with finite (i.e. compact) support. Note also that in the above proof, the only hypothesis we have really used is that the function $t \mapsto p_t(y_0, y_0)$ is the Laplace transform of some finite positive measure. Kijima (1993) showed that such a representation also holds for Markov chains which are skip-free to the left on $E = \{1, 2, 3 \ldots\}$. McKean (1956) proved a representation similar to (13) for one dimensional diffusions (which are always symmetric). Again, the analogue of Proposition 20 goes through.

7. Conditioned processes

Here, we shall use the theory developed in the previous sections to prove the existence of a conditioned process.

Theorem 21. *Let X be a Markov process with finite lifetime ζ such that Assumptions (A1)-(A4) hold, and suppose that ν is a compactly supported probability measure on E. If $\mathfrak{N}_{-\infty}(X) = \{\overline{z}_0\}$, let $\overline{K}(t, x; \overline{z}_0) = e^{-\Lambda t}\varphi(x)$ with $\langle \nu, \varphi \rangle = 1$; then*

$$\lim_{s \to \infty} \mathbb{P}_\nu(d\omega \,|\, \zeta > s) = e^{-\Lambda t}\varphi(X_t(\omega))1_{\{\zeta > t\}}(\omega)\mathbb{P}_\nu(d\omega) \text{ on } \mathcal{F}_t, \quad t > 0.$$

The limiting law is that of a nonexplosive, time homogeneous Markov process, with semigroup $Q_t(x, dy) = e^{-\Lambda t}P_t(x, dy)\varphi(y)/\varphi(x)$ and initial distribution $g(x)\nu(dx)$.

Proof. Let H denote any bounded $\mathcal{F}_t$ measurable random variable, $t \geq 0$. By (11), we have $h_s(t, x) \to e^{-\Lambda t}\varphi(x)$, and by the parabolic Harnack inequality (A2), this convergence occurs boundedly on compact subsets of $\overline{E}$. Now if $T_{K^c} = \inf\{t >$

$0 : X_t \notin K\}$ denotes the first exit time from a compact set $K \subset E$, the bounded convergence theorem gives

$$\begin{aligned}\lim_{s\to\infty} \mathbb{E}_\nu(H, T_{K^c} > t \,|\, \zeta > s) &= \lim_{s\to\infty} \mathbb{E}_\nu\left(H, T_{K^c} > t, \frac{\mathbb{P}_{X_t}(\zeta > s-t)}{\mathbb{P}_\nu(\zeta > s)}\right) \\ &= \mathbb{E}_\nu[H, T_{K^c} > t, e^{-\Lambda t}\varphi(X_t)].\end{aligned}$$

We now dispense with the set $\{T_{K^c} > t\}$ above. Taking $H = 1$ and remembering that φ is Λ-invariant and satisfies $\langle \varphi, \nu\rangle = 1$,

$$\begin{aligned}\lim_{s\to\infty} \mathbb{E}_\nu(T_{K^c} \le t \,|\, \zeta > s) &= 1 - \lim_{s\to\infty} \mathbb{E}_\nu(T_{K^c} \le t \,|\, \zeta > s) \\ &= 1 - \mathbb{E}_\nu[T_{K^c} \le t, e^{-\Lambda t}\varphi(X_t)] \\ &= \int \nu(dx)\Big(\varphi(x) - \mathbb{E}_x[e^{-\Lambda t}\varphi(X_t), T_{K^c} \le t]\Big) \\ &= \mathbb{E}_\nu[e^{-\Lambda t}\varphi(X_t), \zeta > t, T_{K^c} \le t] \\ &= \mathbb{Q}(T_{K^c} \le t).\end{aligned}$$

Here $\mathbb{Q}$ is the law of the Markov process with transition function Q_t and initial distribution $\varphi(x)\nu(dx)$. Now

$$\begin{aligned}&\overline{\lim}_{n\to\infty}\Big|\mathbb{E}_\nu(H \,|\, \zeta > r_n) - \int H d\mathbb{Q}\Big| \\ &\qquad \le \overline{\lim}_{n\to\infty}\left|\mathbb{E}_\nu(H, T_{K^c} > t \,|\, \zeta > s) - \int_{\{T_{K^c} > t\}} H d\mathbb{Q}\right| \\ &\qquad\qquad + \overline{\lim}_{n\to\infty} 2\,\|H\|\,\Big(\mathbb{E}_\nu(T_{K^c} \le t \,|\, \zeta > s) + \mathbb{Q}(T_{K^c} \le t)\Big) \\ &\qquad = 4\,\|H\|\,\mathbb{Q}(T_{K^c} \le t),\end{aligned}$$

and, since $T_{K^c} \uparrow \zeta$ as $K \uparrow E$ and $\zeta = \infty$ a.s. under $\mathbb{Q}$, the right hand side can be made arbitrarily small by choosing K arbitrarily large; the result follows. □

Recall the last paragraph of Section 3. By choosing ν as a quasistationary distribution (which can never be compactly supported, unless the process is Λ-recurrent), we can force $h \equiv 0$. In that case, we also have

$$\lim_{s\to\infty} \mathbb{P}_\nu(d\omega \,|\, \zeta > s) = 0.$$

Consider now the following example, due to Jacka and Roberts (1995), of a process for which the conditioning problem *does not* have a solution, irrespective of how ν is chosen.

Example 22. Let X be the Markov chain on $E = \{1, 2, 3, \dots\}$ with nonconservative q-matrix given by

$$q(x,y) = \begin{bmatrix} -1 & 2^{-2} & 2^{-3} & 2^{-4} & \cdots \\ 2^{-2} & -2^{-2} & & & \\ 2^{-3} & & -2^{-3} & & \\ 2^{-4} & & & -2^{-4} & \\ \vdots & & & & \ddots \end{bmatrix}.$$

When started in state 1, X waits for an exponential time with mean 1 before either jumping to state k with probability 2^{-k+1} or getting killed with probability 1/2. In state $k > 1$, it first waits for an exponential time with mean 2^{k+1} and then jumps back to state 1. Thus the process is irreducible, and the smallest cemetery neighbourhood is given by $N = \{1\}$. Moreover, X is clearly symmetric, so that $\overline{N}_{-\infty}$ consists of at most one point. Nevertheless, there is no conditioned process. Indeed, suppose that $g(x)$ is a positive solution to the equation $Qg = -\lambda g$. We must therefore solve the system

$$-g(1) + \sum_{k=2}^{\infty} 2^{-k} g(k-1) = -\lambda g(1),$$
$$2^{-k} g(1) - 2^{-k} g(k-1) = -\lambda g(k-1), \quad k \geq 2.$$

Clearly the only solution is $g(x) = 0$. Now Proposition 10 guarantees that the minimal parabolic functions associated with points of $\overline{N}_{-\infty}$ are of the form $e^{-\lambda t} g(x)$; thus there are no nonzero minimal parabolic points in $\overline{N}_{-\infty}$ for this process, and in particular the sequence (h_s) cannot converge. The source of this failure lies in the fact that, even though (A1), (A2), and (A4) are satisfied, Assumption (A3) is violated. Thus we have uniqueness, but not existence of a nontrivial limit for (h_s). Here, the process can jump into each and every state from state 1.

8. Acknowledgement

I thank Terry Lyons for suggesting a connection with Martin boundaries to me. I also thank Phil Pollett and Gareth Roberts for many useful discussions. 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. and Burdzy, K. (1992). Lifetimes of conditioned diffusions. *Probab. Theory and Rel. Fields* **91**, 405-444.

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

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

[5] Breyer, L.A. (1998b) *Yaglom limits via compactifications.* Preprint available at `http://www.statslab.cam.ac.uk/ ~laird/`

[6] Darroch, J.N. and Seneta, E. (1967). On quasi-stationary distributions in absorbing continuous-time finite Markov chains. *J. Appl. Probab.* **4**, 192-196.

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

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

[9] Jacka, S.D. and Roberts, G.O. (1995). Weak convergence of conditioned processes on a countable state space. *J. Appl. Probab.* **32**, 902-916.

[10] Jacka, S.D., Pollett, P.K. and Roberts, G.O. (1997). Non-explosivity of limits of conditioned birth and death processes. *J. Appl. Probab.* **34**, 35-45.

[11] Jacka, S.D. and Roberts, G.O. (1997) Conditional one-dimensional diffusions. Preprint.

[12] 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.

[13] Koranyi, A. and Taylor, J.C. (1985). Minimal solutions of the heat equation and uniqueness of the positive Cauchy problem on homogeneous spaces. *Proc. Amer. Math. Soc.* **94** (2), 273-278.

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

[15] McKean, H.P. Jr. (1963). Excursions of a non-singular diffusion. *Z. Wahrscheinlichkeitstheorie* **1**, 230-239.

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

[17] Nåsell, I. (1995). The quasi-stationary distribution of the closed endemic SIS model. *Adv. Appl. Probab.* **28**, 895-932.

[18] Pinsky, R.G. (1985). On the convergence of diffusion processes conditioned to remain in a bounded region for large time to limiting positive recurrent diffusion processes. *Ann. Probab.* **13** (2), 363-378.

[19] Pitman, J.W. (1975). One-dimensional Brownian motion and the three-dimensional Bessel process. *Adv. Appl. Probab.* **7** (3), 511-526.

[20] Pollett, P.K. (1988). Reversibility, invariance and μ-invariance. *Adv. Appl. Probab.* **20**, 600-621.

[21] Schrijner, P. and van Doorn, E. (1997). Weak convergence of conditioned birth-deat processes in discrete time. *J. Appl. Probab.* **34** (1), 46-53.

[22] Sharpe, M. (1988). *General Theory of Markov Processes.* Pure and Applied Mathematics 133. Academic Press, Boston.

[23] Tuominen, P. and Tweedie, R.L. (1979). Exponential decay and ergodicity of general Markov processes and their discrete skeletons. *Adv. Appl. Probab.* **11**, 784-803.

[24] Williams, D. (1974). Path decomposition and continuity of local time for one-dimensional diffusions, I *Proc. London Math. Soc. (3)* **28**, 738-768.

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

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