Chapter 2
Quasi-Stationary Distributions: General Results

In this chapter, we introduce the main concepts in a general context of killed processes. Thus, in Sect. 2.2, we give the definition of quasi-stationary distributions (QSDs). In Theorem 2.2 of Sect. 2.3, we show that starting from a QSD the killing time is exponentially distributed, and in Theorem 2.6 of Sect. 2.4, we show that the killing time and the state of killing are independent random variables. In Theorem 2.11 of Sect. 2.7, we give a theorem of existence of a QSD in a topological setting without any assumption on compactness or spectral properties.

2.1 Notation

We will consider a Markov process $Y = (Y_t : t \geq 0)$ taking values in a state space $\mathcal{X}$ which is endowed with a σ-field $\mathcal{B}(\mathcal{X})$. Let us fix some notation:

- $\mathcal{M}(\mathcal{X})$ is the set of real measurable functions defined on $\mathcal{X}$; $\mathcal{M}_+(\mathcal{X})$ (respectively, $\mathcal{M}_b(\mathcal{X})$) is the set of positive (respectively, bounded) elements in $\mathcal{M}(\mathcal{X})$;
- $\mathcal{P}(\mathcal{X})$ is the set of probability measures on $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$.

When $\mathcal{X}$ has a topological structure, $\mathcal{B}(\mathcal{X})$ denotes its Borel σ-field. In this case, we denote by $\mathcal{C}(\mathcal{X})$ the set of continuous functions and by $\mathcal{C}_b(\mathcal{X})$ its bounded elements.

Let v be a measure on $\mathcal{X}$ and $f \in \mathcal{M}(\mathcal{X})$. If $\int f \, dv$ is well defined then we set $v(f) = \int f \, dv$. This is the case if $f \in L^1(v)$, or when $f \in \mathcal{M}_+(\mathcal{X})$, or when $v \in \mathcal{P}(\mathcal{X})$ and $f \in \mathcal{M}_b(\mathcal{X})$.

Let us fix $\mathcal{X}$ a topological space endowed with its Borel σ-field $\mathcal{B}(\mathcal{X})$. Let Y be a Markov process taking values in $\mathcal{X}$ satisfying the conditions we fix below.

Let Ω be the set of right continuous trajectories on $\mathcal{X}$, indexed by points in $[0, \infty)$. So $\omega = (\omega_s : s \in [0, \infty)) \in \Omega$ means that $\omega_s \in \mathcal{X}$ and $\omega_s = \lim_{h \to 0^+} \omega_{s+h}$ for all $s \in [0, \infty)$. Let $(\Omega, \mathcal{F})$ be a measurable space where $\mathcal{F}$ contains the σ-field generated by all the projections $\mathrm{pr}_s : \Omega \to \mathcal{X}, \omega \to \mathrm{pr}_s(\omega) = \omega_s, s \geq 0$. For $t \geq 0$ let $\Theta_t : \Omega \to \Omega$ be the operator shifting the trajectories by t, so $\Theta_t(\omega) = (\omega_s : s \geq t)$.

P. Collet et al., *Quasi-Stationary Distributions*, Probability and Its Applications,
DOI 10.1007/978-3-642-33131-2_2,

The process $Y = (Y_t : t \geq 0)$ takes values on $\mathcal{X}$ and its trajectories belong to Ω. We denote by $\mathbb{F} = (\mathcal{F}_t : t \geq 0)$ a filtration of σ-fields such that Y is adapted to $\mathbb{F}$, so for all $t \geq 0$ $\mathcal{F}_t \supseteq \sigma(Y_s : s \leq t)$ the σ-field generated by the process up to time t. Let $(\mathbb{P}_x : x \in \mathcal{X})$ be a family of probability measures defined on $(\Omega, \mathcal{F})$. We assume that Y is a Markov process with respect to the family $(\mathbb{P}_x : x \in \mathcal{X})$, that is,

- $\mathbb{P}_x(Y_0 = x) = 1$ for all $x \in \mathcal{X}$;
- For all $A \in \mathcal{F}$ the function $x \to \mathbb{P}_x(A)$ is $\mathcal{B}(\mathcal{X})$-measurable;
- $\mathbb{P}_x(\Theta_t(Y) \in A | \mathcal{F}_t) = \mathbb{P}_{Y_t}(A)$ $\mathbb{P}_x$- a.s., $\forall x \in \mathcal{X}$ and $\forall A \in \mathcal{F}$.

We assume that $\mathcal{X}$ is a Polish space (metric, separable and complete) and that the Markov process Y is standard in the sense of Definition 9.2 in Blumenthal and Getoor (1968). In particular, Y is strong Markov. We denote by

$$T_B = \inf\{t > 0 : Y_t \in B\}, \quad B \in \mathcal{B}(\mathcal{X}),$$

the hitting time of B, where as usual we put $\inf \emptyset = \infty$. Since the Markov process Y is standard, T_B is a stopping time. When $B = \{x_0\}$ is a singleton, we put T_{x_0} instead of $T_{\{x_0\}}$.

2.2 Killed Process and QSD

In this theory, there is a set of forbidden states for the process, denoted by $\partial\mathcal{X}$. We assume that $\partial\mathcal{X} \in \mathcal{B}(\mathcal{X})$ and $\emptyset \neq \partial\mathcal{X} \neq \mathcal{X}$, this last condition is to avoid trivial situations. Its complement $\mathcal{X}^{\mathbf{a}} := \mathcal{X} \setminus \partial\mathcal{X}$ is called the set of allowed states. Let

$$T = T_{\partial\mathcal{X}},$$

be the hitting time of $\partial\mathcal{X}$, it is called the killing time (or extinction time in population models).

All the quantities in this study will only depend on $Y^T = (Y_{t \wedge T} : t \geq 0)$, the process stopped at $\partial\mathcal{X}$, and so there is no loss of generality in assuming $Y_t = Y_T$ for $t \geq T$. Then $Y = Y^T$ and T is an absorption time. We note that except when we look for the region where the process is absorbed, we will be mainly concerned with the trajectory before killing $(Y_t : t < T)$.

In our framework, we will assume that there is sure killing at $\partial\mathcal{X}$,

$$\forall x \in \mathcal{X}^{\mathbf{a}}: \quad \mathbb{P}_x(T < \infty) = 1. \tag{2.1}$$

In particular, this means that there is no explosion. Condition (2.1) is equivalent to

$$\forall \rho \in \mathcal{P}\big(\mathcal{X}^{\mathbf{a}}\big): \quad \mathbb{P}_\rho(T < \infty) = 1,$$

where as usual we put $\mathbb{P}_\rho = \int_{\mathcal{X}^{\mathbf{a}}} \mathbb{P}_x \, d\rho(x)$ for any $\rho \in \mathcal{P}(\mathcal{X}^{\mathbf{a}})$. Relation (2.1) obviously implies

$$\forall x \in \mathcal{X}^{\mathbf{a}}: \quad \mathbb{P}_x(T < \infty) > 0, \tag{2.2}$$

then,

$$\forall \rho \in \mathcal{P}\big(\mathcal{X}^{\mathbf{a}}\big) \; \exists t(\rho) \; \forall t \geq t(\rho): \quad \mathbb{P}_\rho(T > t) < 1. \tag{2.3}$$

Condition (2.2) (and hence the stronger condition (2.1)) implies that there cannot exist stationary distributions supported by $\mathcal{X}^{\mathbf{a}}$ because if ρ is one of them, from (2.3) we should have $\rho(Y_t \in \mathcal{X}^{\mathbf{a}}) < 1$ for $t \geq t(\rho)$, which contradicts the stationarity because $\rho(Y_0 \in \mathcal{X}^{\mathbf{a}}) = 1$.

Let us define the quasi-stationary distributions.

Definition 2.1 $\nu \in \mathcal{P}(\mathcal{X}^{\mathbf{a}})$ is said to be a QSD (for the process killed at $\partial\mathcal{X}$) if

$$\forall B \in \mathcal{B}\big(\mathcal{X}^{\mathbf{a}}\big)\ \forall t \geq 0: \quad \mathbb{P}_\nu(Y_t \in B | T > t) = \nu(B).$$

So, ν is a QSD if

$$\forall B \in \mathcal{B}\big(\mathcal{X}^{\mathbf{a}}\big)\ \forall t \geq 0: \quad \mathbb{P}_\nu(Y_t \in B, T > t) = \nu(B)\mathbb{P}_\nu(T > t). \tag{2.4}$$

Since Y is stopped at time T, we have $\mathbb{P}_\nu(Y_t \in B, T > t) = \mathbb{P}_\nu(Y_t \in B)$ for all $B \in \mathcal{B}(\mathcal{X}^{\mathbf{a}})$, and the condition of QSD takes the form

$$\forall B \in \mathcal{B}\big(\mathcal{X}^{\mathbf{a}}\big)\ \forall t \geq 0: \quad \mathbb{P}_\nu(Y_t \in B) = \nu(B)\mathbb{P}_\nu(T > t). \tag{2.5}$$

From the previous observation, condition (2.2) ensures that a QSD cannot be stationary.

2.3 Exponential Killing

Let ν be a QSD. The first remark is that when starting from ν, the killing time at $\partial\mathcal{X}$ is exponentially distributed (see Ferrari et al. 1995b). Let us state it.

Theorem 2.2 *If ν be a QSD, then*

$$\exists \theta(\nu) \geq 0 \quad \textit{such that} \quad \mathbb{P}_\nu(T > t) = e^{-\theta(\nu)t} \quad \forall t \geq 0,$$

that is, starting from ν, T is exponentially distributed with parameter $\theta(\nu)$.

Proof Since ν is a QSD, from relation (2.4) we get that all $g \in \mathcal{M}_+(\mathcal{X}^{\mathbf{a}})$ and all $g \in \mathcal{M}_b(\mathcal{X}^{\mathbf{a}})$ verify

$$\forall t \geq 0: \quad \mathbb{E}_\nu\big(g(Y_t)\mathbb{1}_{T>t}\big) = \nu(g)\mathbb{P}_\nu(T > t). \tag{2.6}$$

By taking the measurable function $g(x) = \mathbb{P}_x(T > s)$, $x \in \mathcal{X}^{\mathbf{a}}$, and since $\int_{\mathcal{X}^{\mathbf{a}}} g\, d\nu = \mathbb{P}_\nu(T > s)$, we get

$$\forall t \geq 0: \quad \mathbb{E}_\nu\big(\mathbb{P}_{Y_t}(T > s)\mathbb{1}_{T>t}\big) = \mathbb{P}_\nu(T > s)\mathbb{P}_\nu(T > t).$$

From the Markov property, it is verified that $\forall t, s \geq 0$

$$\begin{aligned}\mathbb{P}_\nu(T > t + s) = \mathbb{E}_\nu(\mathbb{1}_{T>t+s}) &= \mathbb{E}_\nu\big(\mathbb{1}_{T>t}\mathbb{E}(\mathbb{1}_{T>t+s}|\mathcal{F}_t)\big)\\ &= \mathbb{E}_\nu\big(\mathbb{1}_{T>t}\mathbb{E}_{Y_t}(\mathbb{1}_{T>s})\big) = \mathbb{P}_\nu(T > s)\mathbb{P}_\nu(T > t).\end{aligned}$$

So, there exists some $\theta \geq 0$ such that $\mathbb{P}_\nu(T > t) = e^{-\theta t}$ $\forall t \geq 0$. Then, the result follows. $\square$

Remark 2.3 For discrete time, the argument in the proof of Theorem 2.2 shows that when starting from a QSD ν, the killing time at $\partial\mathcal{X}$ is geometrically distributed, so $\mathbb{P}_\nu(T > n) = \kappa(\nu)^n$ for all $n \geq 0$, where $\kappa(\nu) = \mathbb{P}_\nu(T > 1)$.

The coefficient $\theta(\nu)$ verifies $\theta(\nu) = -\frac{1}{t}\log \mathbb{P}_\nu(T > t)\ \forall t > 0$. So, when the process starts from ν, $\theta(\nu)$ is its exponential rate of survival, which is stationary in time.

Theorem 2.2 gives $\mathbb{P}_\nu(Y_{0^+} \in \partial\mathcal{X}) = 0$. Then, from (2.3) we get $0 < \mathbb{P}_\nu(T > t) < 1$ for ν a QSD, and so $\theta(\nu) \in (0, \infty)$.

From (2.4) and Theorem 2.2, we get that $\nu \in \mathcal{P}(\mathcal{X}^{\mathbf{a}})$ is a QSD if and only if it verifies

$$\exists \theta \in (0, \infty)\ \forall B \in \mathcal{B}(\mathcal{X}^{\mathbf{a}})\ \forall t \geq 0\colon \quad \mathbb{P}_\nu(Y_t \in B, T > t) = \nu(B)e^{-\theta t}. \tag{2.7}$$

In this case, $\theta = \theta(\nu)$.

When ν is a QSD, we get

$$\forall \theta < \theta(\nu)\colon \quad \mathbb{E}_\nu\left(e^{\theta T}\right) < \infty,$$

and we deduce

$$\forall \theta < \theta(\nu),\ \nu\text{-a.s. in } x\colon \quad \mathbb{E}_x\left(e^{\theta T}\right) < \infty. \tag{2.8}$$

The exponential moment condition can be written in the following way.

Proposition 2.4 *We have the equality*

$$\theta_x^* := \sup\left\{\theta : \mathbb{E}_x\left(e^{\theta T} < \infty\right)\right\} = \liminf_{t\to\infty} -\frac{1}{t}\log \mathbb{P}_x(T > t), \tag{2.9}$$

and a necessary condition for the existence of a QSD is the existence of a positive exponential moment, or equivalently, a positive exponential rate of survival:

$$\exists x \in \mathcal{X}^{\mathbf{a}}\colon \quad \theta_x^* > 0. \tag{2.10}$$

Proof The fact that (2.10) is necessary for the existence of a QSD follows from (2.8) and (2.9). We now prove the equality (2.9). Note that the result will follow once we show the following relation for a nonnegative random variable W:

$$\sup\left\{\theta : \mathbb{E}\left(e^{\theta W}\right) < \infty\right\} = \liminf_{t\to\infty} -\frac{1}{t}\log \mathbb{P}(W > t). \tag{2.11}$$

The inequality $\leq$ in (2.11) follows from Markov's Inequality applied to all nonnegative θ such that $\mathbb{E}(e^{\theta W}) < \infty$,

$$\mathbb{E}\left(e^{\theta W}\right) \geq e^{\theta t}\mathbb{P}(W > t) \quad \forall t \geq 0.$$

Let us prove the other inequality. First, observe that for every distribution function F, Fubini's Theorem gives

$$\begin{aligned}\int_0^\infty \theta e^{\theta t}\big(1 - F(t)\big)\,dt &= \int_0^\infty e^{\theta t}\,dt \int_{t^+}^\infty dF(u) = \int_{0^+}^\infty \left(\int_0^u \theta e^{\theta t}\,dt\right) dF(u) \\ &= \int_0^\infty \left(e^{\theta u} - 1\right) dF(u) = \int_0^\infty e^{\theta u}\,dF(u) - \big(1 - F(0)\big).\end{aligned}$$

Then

$$\int_0^\infty e^{\theta t}\, dF(t) = \big(1 - F(0)\big) + \int_0^\infty \theta e^{\theta t}\big(1 - F(t)\big)\, dt. \tag{2.12}$$

Note that

$$\liminf_{t\to\infty} -\frac{1}{t} \log \mathbb{P}(W > t) = \sup\big\{\theta' : \exists C < \infty\ \forall t \geq 0, \mathbb{P}(W > t) \leq Ce^{-\theta' t}\big\}. \tag{2.13}$$

Now, if $\theta' > 0$ satisfies $\mathbb{P}(W > t) \leq Ce^{-\theta' t}$ $\forall t \geq 0$, from (2.12) we get that any $\theta \in (0, \theta')$ verifies:

$$\begin{aligned}\mathbb{E}\big(e^{\theta W}\big) &= \mathbb{P}(W > 0) + \int_0^\infty \theta e^{\theta t}\mathbb{P}(W > t)\, dt \leq \mathbb{P}(W > 0) + C\theta \int_0^\infty e^{-(\theta'-\theta)t}\, dt \\ &= \mathbb{P}(W > 0) + C\left(\frac{1}{\theta' - \theta}\right) < \infty.\end{aligned} \tag{2.14}$$

The inequality $\geq$ in (2.11) follows from (2.13) and (2.14). □

Remark 2.5 Note that if ν is a QSD then $\mathbb{E}_\nu(e^{\theta(\nu)T}) = \infty$. Then, if $\theta > 0$ satisfies the condition

$$\sup\big(\mathbb{E}_x\big(e^{\theta T}\big) : x \in \mathcal{X}^{\mathbf{a}}\big) < \infty,$$

there cannot exists a QSD ν with $\theta(\nu) = \theta$ because otherwise we should have

$$\infty = \mathbb{E}_\nu\big(e^{\theta(\nu)T}\big) \leq \sup\big(\mathbb{E}_x\big(e^{\theta(\nu)T}\big) : x \in \mathcal{X}^{\mathbf{a}}\big) < \infty.$$

2.4 Independence Between the Exit Time and the Exit State

Let us show that when the process starts from a QSD, the killing time and the state where it is killed are independent random variables (see Martínez 2008).

Theorem 2.6 *Let ν be a QSD (for the process killed at $\partial\mathcal{X}$). Then T and Y_T are $\mathbb{P}_\nu$-independent random variables.*

Proof We already proved that the QSD property implies $\mathbb{P}_\nu(T \in (s, s + ds]) = \theta(\nu)e^{-\theta(\nu)s}\, ds$. Let $E \in \mathcal{B}(\partial\mathcal{X})$ be a measurable set contained in $\partial\mathcal{X}$. The Markov property and the QSD definition $\mathbb{P}_\nu(Y_s \in \bullet | T > s) = \nu(\bullet)$ imply that for all real $a > 0$ and every integer $n \geq 0$,

$$\begin{aligned}\mathbb{P}_\nu\big(T \in \big(na, (n+1)a\big], Y_T \in E\big) &= \mathbb{P}_\nu(T > na, Y_{(n+1)a} \in E) \\ &= \mathbb{P}_\nu(Y_{(n+1)a} \in E | T > na)\mathbb{P}_\nu(T > na) \\ &= \mathbb{P}_\nu(Y_a \in E)e^{-\theta(\nu)na}.\end{aligned}$$

We sum over all $n \geq 0$ to get

$$\mathbb{P}_\nu(Y_T \in E) = \left(\frac{1}{1 - e^{-\theta(\nu)a}}\right)\mathbb{P}_\nu(Y_a \in E).$$

So, the independence relation follows:

$$\begin{aligned}\mathbb{P}_\nu(T \le a, Y_T \in E) = \mathbb{P}_\nu(Y_a \in E) &= \big(1 - e^{-\theta(\nu)a}\big)\mathbb{P}_\nu(Y_T \in E)\\ &= \mathbb{P}_\nu(T \le a)\mathbb{P}_\nu(Y_T \in E).\end{aligned}$$ □

From this relation, we find

$$\frac{d}{da}\mathbb{P}_\nu(Y_a \in E)\bigg|_{a=0} = \theta(\nu)\mathbb{P}_\nu(Y_T \in E). \tag{2.15}$$

Remark 2.7 Let us give the proof of the above property in discrete time. For $n \ge 0$, we have

$$\begin{aligned}\mathbb{P}_\nu(T = n+1, Y_T \in E) &= \mathbb{P}_\nu(T > n, Y_{n+1} \in E)\\ &= \mathbb{P}_\nu(Y_{n+1} \in E | T > n)\mathbb{P}_\nu(T > n)\\ &= \mathbb{P}_\nu(Y_1 \in E)\mathbb{P}_\nu(T > n).\end{aligned}$$

When we sum over all $n \ge 0$ and use $\mathbb{P}_\nu(T > n) = \kappa(\nu)^n$, we obtain

$$\mathbb{P}_\nu(Y_T \in E) = \frac{1}{1-\kappa(\nu)}\mathbb{P}_\nu(Y_1 \in E).$$

Since $\mathbb{P}(T = n+1) = (1-\kappa(\nu))\kappa(\nu)^n$, we find the result

$$\mathbb{P}_\nu(T = n+1, Y_T \in E) = \mathbb{P}_\nu(Y_T \in E)\mathbb{P}_\nu(T = n+1) \quad \forall n \ge 0.$$

Example: The independent case Let Y^k, $k = 1, \dots, K$, be K independent Markov processes taking values in $\mathcal{X}_k$. Assume that each Y_k is killed at time T_k when it hits $\partial\mathcal{X}_k$. Let $\partial\mathcal{X} = \bigcup_{k=1}^K \partial\mathcal{X}_k \times \prod_{l \ne k} \mathcal{X}_l$. Then, the product measure $\nu = \bigotimes_{k=1}^K \nu^k$ is a QSD for the product process $Y = (Y^k : k = 1, \dots, K)$ killed at the region $\partial\mathcal{X}$. Indeed, the killing time is $T = \inf\{T^k : k = 1, \dots, K\}$. Moreover, in this case the exponential rate of survival is $\theta = \sum_{k=1}^K \theta_k$. From the equalities $\mathbb{P}_\nu(T^k \in (t, t+dt], T^l > t \ \forall l \ne k) = \theta_k e^{-\theta t}\, dt$ and $\mathbb{P}_\nu(T \in (t, t+dt]) = \theta K e^{-\theta t}\, dt$, we find $\mathbb{P}_\nu(T = T_k) = \mathbb{P}_\nu(Y_T \in \partial\mathcal{X}_k \times \prod_{l \ne k} \mathcal{X}_l) = \theta_k/\theta$.

2.5 Restricting the Exit State

Let $E \subset \partial\mathcal{X}$ be such that $\mathbb{P}(Y_T \in E) < 1$. We will consider the process that hits the boundary at $\partial\mathcal{X} \setminus E$. For this purpose, let $\Omega^* = \Omega \setminus \{Y_T \in E\} = \{Y_T \notin E\}$ with the restricted σ-field $\mathcal{F}^* = \mathcal{F}|_{\Omega^*}$. We endow this space with the family of conditional probability measures

$$\mathbb{P}_x^*(A) = \mathbb{P}_x(A | Y_T \notin E), \quad x \in \mathcal{X}^* := \mathcal{X} \setminus E, A \in \mathcal{F}^*.$$

Consider the process $Y^* = (Y_t^* : t \ge 0)$, with $Y_t^* := Y_{t\wedge T}$ on Ω^*, so it is the restriction of Y^T to Ω^*. The process Y^* takes values in $\mathcal{X}^*$. We define $\partial\mathcal{X}^* = \partial\mathcal{X} \setminus E$

and the killing time $T^* = \inf\{t \geq 0 : Y_t^* \in \partial\mathcal{X}^*\}$. The set of allowable states is $\mathcal{X}^{\mathbf{a}} = \mathcal{X}^* \setminus \partial\mathcal{X}^*$.

Proposition 2.8 *Y^* is a Markov process with respect to the probability measures $(\mathbb{P}_x^* : x \in \mathcal{X}^*)$. The process $Y_{T^*}^*$ takes values in $\partial\mathcal{X}^*$ with the measure $\mathbb{P}_x^*(Y_{T^*}^* \in C) = \mathbb{P}_x(Y_T \in C | Y_T \notin E)$ for $x \in \mathcal{X}^*$, $C \in \mathcal{B}(\partial\mathcal{X} \setminus E)$.*

Moreover, if ν is a QSD for Y killed at T, the probability measure ν^* defined on $\mathcal{X}^{\mathbf{a}}$ by

$$\nu^*(B) = \frac{1}{\mathbb{P}_\nu(Y_T \notin E)} \int_B \mathbb{P}_x(Y_T \notin E)\nu(dx), \quad B \in \mathcal{B}\big(\mathcal{X}^{\mathbf{a}}\big), \tag{2.16}$$

is a QSD for the process Y^* killed at T^*.

Proof Let us see that Y^* is a Markov process. Since $\mathbb{P}_x^*(Y_t^* = x \ \forall t \geq 0) = 1$ for $x \in \partial\mathcal{X}^*$, we can assume that the starting point satisfies $x \in \mathcal{X}^{\mathbf{a}}$. For $s \geq 0$ let $\mathcal{F}_s^*$ be the σ-field $\mathcal{F}_s$ restricted to $\Omega^* = \{Y_T \notin E\}$. For $0 \leq s < t$ we have

$$\begin{aligned}
\mathbb{E}_x^*\big(g\big(Y_t^*\big)\big|\mathcal{F}_s^*\big) &= \mathbb{E}_x\left(\frac{1}{\mathbb{P}_x(Y_T \notin E)} g(Y_{t\wedge T})\mathbb{1}_{Y_T \notin E}\middle|\mathcal{F}_s^*\right)\mathbb{1}_{Y_T \notin E} \\
&= \mathbb{E}_{Y_{s\wedge T}}\left(g(Y_{(t-s)\wedge T})\mathbb{1}_{Y_T \notin E}\frac{1}{\mathbb{P}_x(Y_T \notin E)}\right)\mathbb{1}_{Y_T \notin E} = \mathbb{E}_{Y_s^*}^*\big(g\big(Y_{t-s}^*\big)\big).
\end{aligned}$$

Then, the Markov property holds.

The only thing left to prove is that the measure ν^* given in (2.16) is a QSD for the process Y^* killed at T^*. We have

$$\begin{aligned}
&\mathbb{P}_{\nu^*}^*\big(Y_t^* \in B, T^* > t\big) \\
&= \int_{\mathcal{X}^{\mathbf{a}}} \mathbb{P}_x^*\big(Y_t^* \in B, T > t\big)\nu^*(dx) \\
&= \int_{\mathcal{X}^{\mathbf{a}}} \frac{1}{\mathbb{P}_x(Y_T \notin E)}\mathbb{P}_x(Y_t \in B, T > t, Y_T \notin E)\frac{\mathbb{P}_x(Y_T \notin E)}{\mathbb{P}_\nu(Y_T \notin E)}\nu(dx) \\
&= \frac{\mathbb{P}_\nu(Y_t \in B, T > t, Y_T \notin E)}{\mathbb{P}_\nu(Y_T \notin E)}.
\end{aligned}$$

Then, from the QSD property of ν and the Markov property satisfied by Y, we find for $B \in \mathcal{B}(\mathcal{X}^{\mathbf{a}})$ that

$$\begin{aligned}
\mathbb{P}_{\nu^*}^*\big(Y_t^* \in B, T^* > t\big) &= \frac{1}{\mathbb{P}_\nu(Y_T \notin E)} \int_B \mathbb{P}_y(Y_T \notin E)\mathbb{P}_\nu(T > t)\nu(dy) \\
&= \frac{\mathbb{P}_\nu(T > t)}{\mathbb{P}_\nu(Y_T \notin E)} \int_B \mathbb{P}_y(Y_T \notin E)\nu(dy).
\end{aligned}$$

We conclude that

$$\mathbb{P}_{\nu^*}^*\big(Y_t^* \in B\big|T^* > t\big) = \frac{\int_B \mathbb{P}_y(Y_T \notin E)\nu(dy)}{\int_{\mathcal{X}^{\mathbf{a}}} \mathbb{P}_z(Y_T \notin E)\nu(dz)} = \nu^*(B).$$

Then ν^* fulfills the QSD requirements. □

2.6 Characterization of QSD by the Semigroup

Let $(P_t : t \geq 0)$ be the semigroup of the process before killing at $\partial\mathcal{X}$ (often this semigroup is only denoted by (P_t)). Then

$$P_t f(x) = \mathbb{E}_x\big(f(Y_t), T > t\big), \quad x \in \mathcal{X}^{\mathbf{a}}, \tag{2.17}$$

and it acts on the set $\mathcal{M}_b(\mathcal{X}^{\mathbf{a}})$ of real measurable bounded functions defined on $\mathcal{X}^{\mathbf{a}}$. The Markov property implies that it verifies the semigroup equation

$$P_t \circ P_s = P_{t+s} = P_s \circ P_t \quad \text{on } \mathcal{M}_b\big(\mathcal{X}^{\mathbf{a}}\big).$$

We denote $\mathbf{1} = \mathbb{1}_{\mathcal{X}^{\mathbf{a}}}$. Then $\mathbb{P}_x(T > t) = P_t\mathbf{1}(x)$ for $x \in \mathcal{X}^{\mathbf{a}}$.

Let $(\mathcal{M}_b(\mathcal{X}^{\mathbf{a}}), \|\cdot\|_\infty)$ be the Banach space, where

$$\|f\|_\infty = \sup\big\{\big|f(x)\big| : x \in \mathcal{X}^{\mathbf{a}}\big\}.$$

Note that P_t is a linear contraction on $(\mathcal{M}_b(\mathcal{X}^{\mathbf{a}}), \|\cdot\|_\infty)$,

$$\|P_t\|_\infty \leq 1 \quad \text{where } \|P_t\|_\infty := \sup\big\{\|P_t f\|_\infty : f \in \mathcal{M}_b\big(\mathcal{X}^{\mathbf{a}}\big), \|f\|_\infty \leq 1\big\}.$$

Notice that for any $g \in \mathcal{M}_b(\mathcal{X})$ and $x \in \mathcal{X}^{\mathbf{a}}$ we have

$$\mathbb{E}_x\big(g(Y_t)\big) = \mathbb{E}_x\big(g(Y_t), T > t\big) + g(Y_T)\mathbb{P}_x(T \leq t).$$

In particular, if $g(y) = 0\ \forall y \in \partial\mathcal{X}$, we get $\mathbb{E}_x(g(Y_t)) = \mathbb{E}_x(g(Y_t), T > t)$ for all $x \in \mathcal{X}^{\mathbf{a}}$.

For any measure v on $\mathcal{X}^{\mathbf{a}}$, the set of quantities $(v(P_t\mathbb{1}_B) : B \in \mathcal{B}(\mathcal{X}^{\mathbf{a}}))$ defines a measure on $\mathcal{X}^{\mathbf{a}}$ which is denoted by $P_t^\dagger v$. So, by linearization this action can be also defined for all finite signed measures. Hence the action of the semigroup on the space of measures on $\mathcal{X}^{\mathbf{a}}$, or the space of signed finite measures on $\mathcal{X}^{\mathbf{a}}$, is defined by

$$\forall f \in \mathcal{M}_b\big(\mathcal{X}^{\mathbf{a}}\big): \quad \big(P_t^\dagger v\big)(f) = v(P_t f).$$

We note that $P_t^\dagger$ is also a semigroup acting in these spaces of measures, that is,

$$P_{t+s}^\dagger = P_t^\dagger \circ P_s^\dagger.$$

In fact,

$$P_t^\dagger\big(\big(P_s^\dagger v\big)f\big) = \big(P_s^\dagger v\big)(P_t f) = v(P_s P_t f) = v(P_{t+s} f) = \big(P_{t+s}^\dagger v\big)f.$$

For all $v \in \mathcal{P}(\mathcal{X})$ we have

$$\forall f \in \mathcal{M}_b\big(\mathcal{X}^{\mathbf{a}}\big)\ \forall t \geq 0: \quad v(P_t f) = \mathbb{E}_v\big(f(Y_t)\mathbb{1}_{T>t}\big).$$

Hence, from (2.6) and (2.7), $\nu \in \mathcal{P}(\mathcal{X})$ is a QSD if and only if it satisfies

$$\exists\theta > 0 \quad \text{such that} \quad \forall f \in \mathcal{M}_b\big(\mathcal{X}^{\mathbf{a}}\big)\ \forall t \geq 0: \quad \nu(P_t f) = e^{-\theta t}\nu(f),$$

or if and only if

$$\exists\theta > 0 \quad \text{such that} \quad \forall t \geq 0: \quad P_t^\dagger \nu = e^{-\theta t}\nu. \tag{2.18}$$

The parameter θ turns out to be equal to $\theta(\nu)$, the exponential rate of survival of ν.

Let us show that there exists some QSD when the eigenmeasure equation (2.18) is satisfied by a finite measure but for a fixed strictly positive time (result shown in Collet et al. 2011). For this purpose, the following observation is useful. Since by definition the process Y is adapted to the filtration $\mathbb{F} = (\mathcal{F}_t : t \geq 0)$, it is progressively measurable (for instance, see Revuz and Yor 1999, Proposition 4.8), that is, for all $t \geq 0$ the map

$$[0,t] \times \Omega \to \mathcal{X}, \qquad (s,\omega) \to Y_s(\omega),$$

is $\mathcal{B}([0,t]) \otimes \mathcal{F}_t$ measurable. Then also the killed process Y^T is progressively measurable with respect to the filtration $(\mathcal{F}_{t\wedge T} : t \geq 0)$ (see Revuz and Yor 1999, Proposition 4.10) and by integrating over Ω we deduce that $P_t \mathbb{1}_B$ is measurable in t.

Lemma 2.9 *Let $\widetilde{\nu} \in \mathcal{P}(\mathcal{X}^{\mathbf{a}})$ and $\beta > 0$ be such that $P_1^\dagger \widetilde{\nu} = \beta \widetilde{\nu}$. Then $\beta < 1$ and there exists a QSD ν whose exponential rate of survival is $\theta := -\log \beta > 0$.*

Proof Let $\mathbf{1} = \mathbb{1}_{\mathcal{X}^{\mathbf{a}}}$. From condition (2.3), there exists some integer $n > 0$ such that $\mathbb{P}_{\widetilde{\nu}}(T > n) < 1$. From $\beta^n = (P_n^\dagger \widetilde{\nu})\mathbf{1} = \mathbb{P}_{\widetilde{\nu}}(T > n) < 1$, we get $\beta^n < 1$, so $\beta < 1$ and $\theta := -\log \beta > 0$. We must show that there exists $\nu \in \mathcal{P}(\mathcal{X}^{\mathbf{a}})$ such that $P_t^\dagger \nu = e^{-\theta t}\nu$ for all $t \geq 0$.

Since $P_t \mathbb{1}_B$ is measurable in t, for all $B \in \mathcal{B}(\mathcal{X}^{\mathbf{a}})$, the following quantity is well-defined:

$$\nu(B) := \int_0^1 e^{\theta s} \widetilde{\nu}(P_s \mathbb{1}_B)\, ds.$$

By linearity and monotonicity, we get that ν is a finite measure on $(\mathcal{X}^{\mathbf{a}}, \mathcal{B}(\mathcal{X}^{\mathbf{a}}))$. By definition of $P_s^\dagger$, we can write

$$\nu(B) := \int_0^1 e^{\theta s} \big(P_s^\dagger \widetilde{\nu}\big)(B)\, ds.$$

Let $t \in (0,1]$. We have

$$\begin{aligned}
P_t^\dagger \nu(B) = \nu(P_t \mathbb{1}_B) &= \int_0^1 e^{\theta s} P_s^\dagger \widetilde{\nu}(P_t \mathbb{1}_B)\, ds = \int_0^1 e^{\theta s} P_{t+s}^\dagger \widetilde{\nu}(B)\, ds \\
&= \int_0^{1-t} e^{\theta s} P_{t+s}^\dagger \widetilde{\nu}(B)\, ds + \int_{1-t}^1 e^{\theta s} P_{t+s}^\dagger \widetilde{\nu}(B)\, ds \\
&= \int_t^1 e^{\theta(u-t)} P_u^\dagger \widetilde{\nu}(B)\, du + \int_1^{1+t} e^{\theta(u-t)} P_u^\dagger \widetilde{\nu}(B)\, du \\
&= e^{-\theta t} \int_t^1 e^{\theta u} P_u^\dagger \widetilde{\nu}(B)\, du + e^{-\theta t} \int_0^t e^{\theta u} e^{\theta} \big(P_u^\dagger P_1^\dagger \widetilde{\nu}\big)(B)\, du \\
&= e^{-\theta t} \nu(B),
\end{aligned}$$

where in the last equality we use $P_1^\dagger \widetilde{\nu} = e^{-\theta}\widetilde{\nu}$. (Note that the integrals $\int_0^{1-t}$ and $\int_t^1$ vanish when $t = 1$.) In particular, we proved $P_1^\dagger \nu = e^{-\theta}\nu$ and so $P_n^\dagger \nu = e^{-n\theta}\nu$.

For $t > 1$, we write $t = n + r$ with $0 \le r < 1$ and $n \in \mathbb{N} = \{1, 2, \ldots\}$. We have

$$P_t^\dagger \nu = P_r^\dagger P_n^\dagger \nu = e^{-n\theta} P_r^\dagger \nu = e^{-n\theta} e^{-r\theta} \nu = e^{-\theta t} \nu.$$

Now we normalize ν. We have shown that it verifies the QSD condition (2.18). $\square$

2.7 Continuity Assumptions

Let us recall one of the theorems of Krein in Oikhberg and Troitsky (2005), we refer to Theorem 4 therein. Let $(B, \le, \|\cdot\|)$ be an ordered normed space and assume that there exists an element $e \in B$ that satisfies $\|e\| = 1$ and $e \ge x$ for all $x \in B$ with $\|x\| \le 1$. Then, every positive operator $R : B \to B$ is such that its adjoint R^* has a positive eigenvector v. The positivity property of $v \in B^*$ means that $v(b) \ge 0$ for all $b \ge 0$. This theorem implies the following result on the existence of QSD.

Proposition 2.10 *Assume that $\mathcal{X}^{\mathbf{a}}$ is a compact Hausdorff space and that P_1 preserves the set of continuous functions: $P_1(\mathcal{C}(\mathcal{X}^{\mathbf{a}})) \subseteq \mathcal{C}(\mathcal{X}^{\mathbf{a}})$. Then there exists a QSD.*

Proof The set $\mathcal{C}(\mathcal{X}^{\mathbf{a}})$ is an ordered normed space, endowed with the usual order between functions and the supremum $\|\cdot\|_\infty$ norm. The function $\mathbf{1}$ verifies the condition $f \le \mathbf{1}$ for all $f \in \mathcal{C}(\mathcal{X}^{\mathbf{a}})$ with $\|f\|_\infty \le 1$. Since $P_1 : \mathcal{C}(\mathcal{X}^{\mathbf{a}}) \to \mathcal{C}(\mathcal{X}^{\mathbf{a}})$ is a positive operator, the above Krein's Theorem states that $P_1^* : \mathcal{C}(\mathcal{X}^{\mathbf{a}})^* \to \mathcal{C}(\mathcal{X}^{\mathbf{a}})^*$ has a positive eigenfunction. The Riesz Representation Theorem implies that this eigenfunction is a positive finite measure v. Hence $P^* v = \beta v$ for some β, and we necessarily have $\beta > 0$. The probability measure $\nu = v(\mathbf{1})^{-1} v$ verifies $P_1^* \nu = \beta \nu$. Since $P_1^* = P_1^\dagger$ on $\mathcal{C}(\mathcal{X}^{\mathbf{a}})^*$, from Lemma 2.9 we get that ν is a QSD. $\square$

Now we state an existence result of a QSD when $\mathcal{X}^{\mathbf{a}}$ is not necessarily compact. It can be found in Collet et al. (2011). We recall that the set of bounded continuous functions $\mathcal{C}_b(\mathcal{X}^{\mathbf{a}})$ becomes a Banach space when equipped with the supremum norm.

Theorem 2.11 *Assume that P_1 maps $\mathcal{C}_b(\mathcal{X}^{\mathbf{a}})$ into itself, $P_1 : \mathcal{C}_b(\mathcal{X}^{\mathbf{a}}) \to \mathcal{C}_b(\mathcal{X}^{\mathbf{a}})$. Also assume that the following hypothesis holds: $\exists \varphi_0 \in \mathcal{C}_b(\mathcal{X}^{\mathbf{a}})$ such that $\varphi_0 \ge 1$ and it verifies*

Hypothesis 2.1 *For any $u \ge 0$, the set $\varphi_0^{-1}([0, u])$ is compact.*

Assume also that there exist three constants $c_1 > \gamma > 0$ and $\beta > 0$ such that

$$P_1 \mathbf{1} \ge c_1 \mathbf{1}$$

and for any $h \in \mathcal{C}_b(\mathcal{X}^{\mathbf{a}})$ with $0 \le h \le \varphi_0$

$$P_1 h \le \gamma \varphi_0 + \beta.$$

Then, there exists a QSD.

Proof In the dual space $\mathcal{C}_b(\mathcal{X}^{\mathbf{a}})^*$, we define for any real $K>0$ the convex set $\mathcal{K}_K$ given by

$$\mathcal{K}_K=\Big\{v\in\mathcal{C}_b\big(\mathcal{X}^{\mathbf{a}}\big)^*: v\geq 0, v(1)=1, \sup_{h\in\mathcal{C}_b(\mathcal{X}^{\mathbf{a}}),0\leq h\leq\varphi_0} v(h)\leq K\Big\}.$$

We observe that for any K large enough the set $\mathcal{K}_K$ is nonempty. It suffices to consider a Dirac measure δ_x at a point $x\in\mathcal{X}^{\mathbf{a}}$ and to take $K\geq\varphi_0(x)$.

Since for any $K\geq 0$, $\mathcal{K}_K$ is an intersection of weak* closed subsets, it is closed in the weak* topology. We now introduce the nonlinear operator R with domain $\mathcal{K}_K$ and defined by

$$Rv=\frac{P_1^*v}{v(P_1\mathbf{1})}.$$

Note that since $P_1\mathbf{1}>c_1\mathbf{1}$, we have $v(P_1\mathbf{1})\geq c_1 v(\mathbf{1})$ and the operator R is well defined on $\mathcal{K}_K$. We obviously have $Rv(\mathbf{1})=1$. We now prove that R maps $\mathcal{K}_K$ into itself.

Consider $h\in\mathcal{C}_b(\mathcal{X}^{\mathbf{a}})$ with $0\leq h\leq\varphi_0$. Since

$$P_1h\leq\gamma\varphi_0+\beta,$$

and obviously

$$0\leq P_1h\leq\gamma\frac{\|P_1h\|}{\gamma},$$

we get

$$0\leq P_1h\leq\gamma\min\big(\varphi_0,\|P_1h\|/\gamma\big)+\beta.$$

Therefore, since the function $h'=\min(\varphi_0,\|P_1h\|/\gamma)$ satisfies $h'\in\mathcal{C}_b(\mathcal{X}^{\mathbf{a}})$ and $0\leq h'\leq\varphi_0$, we conclude that for $v\in\mathcal{K}_K$

$$Rv(h)\leq\frac{\gamma v(h')+\beta}{c_1}.$$

From the bound $v(h')\leq K$, we get

$$Rv(h)\leq\frac{\gamma v(h')+\beta}{c_1}\leq\frac{\gamma}{c_1}K+\frac{\beta}{c_1}\leq K$$

if $K>\beta/(c_1-\gamma)$. Therefore, for any K large enough, the set $\mathcal{K}_K$ is nonempty and mapped into itself by R.

It is easy to show that R is continuous on $\mathcal{K}_K$ in the weak* topology. This follows at once from the continuity of the operator P_1. We can now apply Tychonov's Fixed Point Theorem (see Tychonov 1935) to deduce that R has a fixed point. This implies that there is a point $v\in\mathcal{K}_K$ such that $P_1^*v=v(P_1\mathbf{1})v$. The proof will be finished once we prove that v can be represented by a measure ν on $\mathcal{X}^{\mathbf{a}}$, that is,

$$\forall f\in\mathcal{C}_b\big(\mathcal{X}^{\mathbf{a}}\big):\quad v(f)=\int f\,d\nu.$$

(Since $v(1)=1$, ν will be necessarily be a probability measure.) Let $\varpi:[0,\infty)\to[0,1]$ be a continuous and nonincreasing with $\varpi=1$ on the interval $[0,1]$ and

$\varpi(2) = 0$ (so $\varpi = 0$ on $[2, \infty)$). For any integer $m \geq 1$ let υ_m be the continuous positive linear form defined on $\mathcal{C}(\mathcal{X}^{\mathbf{a}})$ by

$$\upsilon_m(f) = \upsilon\big(\varpi(\varphi_0/m)f\big).$$

This linear form has support in the set $\varphi_0^{-1}([0, 2m])$ in the sense that it vanishes on functions which vanish on this set. Since $\varphi_0^{-1}([0, 2m])$ is compact by Hypothesis 2.1, the form υ_m can be identified with a nonnegative measure ν_m on $\mathcal{X}^{\mathbf{a}}$, namely, for any $f \in \mathcal{C}_b(\mathcal{X}^{\mathbf{a}})$ we have

$$\forall f \in \mathcal{C}_b\big(\mathcal{X}^{\mathbf{a}}\big): \quad \upsilon_m(f) = \int f \, d\nu_m.$$

We now prove that this sequence of measures $(\nu_m : m \geq 1)$ is tight. Let $u > 0$ and define the set

$$\mathcal{K}_u = \varphi_0^{-1}\big([0, u]\big).$$

By Hypothesis 2.1, for any $u > 0$ this is a compact set. We now observe that

$$\mathbb{1}_{\mathcal{K}_u^c} \leq \big(\mathbf{1} - \varpi(2\varphi_0/u)\big).$$

Therefore,

$$\begin{aligned}\nu_m\big(\mathcal{K}_u^c\big) \leq \nu_m\big(\mathbf{1} - \varpi(2\varphi_0/u)\big) &= \upsilon_m\big(\mathbf{1} - \varpi(2\varphi_0/u)\big)\\ &= \upsilon\big(\varpi(\varphi_0/m)\big(\mathbf{1} - \varpi(2\varphi_0/u)\big)\big).\end{aligned}$$

We now use the following facts: the function $\varphi_0\varpi(\varphi_0/m)(\mathbf{1} - \varpi(2\varphi_0/u))$ is in $\mathcal{C}_b(\mathcal{X}^{\mathbf{a}})$ and satisfies

$$\frac{u}{2}\varpi(\varphi_0/m)\big(\mathbf{1} - \varpi(2\varphi_0/u)\big) \leq \varpi(\varphi_0/m)\big(\mathbf{1} - \varpi(2\varphi_0/u)\big)\varphi_0 \leq \varphi_0,$$

and $\upsilon \in \mathscr{K}_K$; then from definition of $\mathscr{K}_K$ we obtain that

$$\upsilon\big(\varpi(\varphi_0/m)\big(\mathbf{1} - \varpi(2\varphi_0/u)\big)\big) \leq \frac{2}{u}\upsilon\big(\varpi(\varphi_0/m)\big(\mathbf{1} - \varpi(2\varphi_0/u)\big)\varphi_0\big) \leq \frac{2K}{u}.$$

In other words, for any $u > 0$ we have for any integer $m \geq 1$

$$\nu_m\big(\mathcal{K}_u^c\big) \leq \frac{2K}{u}.$$

The sequence of measures $(\nu_m : m \geq 1)$ is therefore tight, and we denote by ν an accumulation point which is a nonnegative measure on $\mathcal{X}^{\mathbf{a}}$. We now finish the proof by showing that for any $f \in \mathcal{C}_b(\mathcal{X}^{\mathbf{a}})$ we have $\nu(f) = \upsilon(f)$. For this purpose, we write

$$\upsilon(f) = \upsilon\big(\varpi(\varphi_0/m)f\big) + \upsilon\big(\big(\mathbf{1} - \varpi(\varphi_0/m)\big)f\big).$$

We now use the inequality

$$\varphi_0 \geq \big(\mathbf{1} - \varpi(\varphi_0/m)\big)\varphi_0 \geq m\big(\mathbf{1} - \varpi(\varphi_0/m)\big),$$

and $(\mathbf{1} - \varpi(\varphi_0/m))\varphi_0 \in \mathcal{C}_b(\mathcal{X}^{\mathbf{a}})$ to conclude

$$\big|\upsilon\big(\big(\mathbf{1} - \varpi(\varphi_0/m)\big)f\big)\big| \leq \upsilon\big(\big(\mathbf{1} - \varpi(\varphi_0/m)\big)|f|\big) \leq \|f\|\upsilon\big(\mathbf{1} - \varpi(\varphi_0/m)\big) \leq \frac{K}{m}.$$

Hence, for any $f \in \mathcal{C}_b(\mathcal{X}^{\mathbf{a}})$ we have

$$\left|\upsilon(f) - \nu_m(f)\right| \le \frac{K}{m}.$$

From the tightness, we get $\lim_{m\to\infty} \nu_m(f) = \nu(f)$ for all $f \in \mathcal{C}_b(\mathcal{X}^{\mathbf{a}})$ (see Billingsley 1968), and therefore $\nu(f) = \upsilon(f)$, which completes the proof of the theorem.

□

Springer

http://www.springer.com/978-3-642-33130-5

Quasi-Stationary Distributions
Markov Chains, Diffusions and Dynamical Systems
Collet, P.; Martínez, S.; San Martín, J.
2013, XVI, 280 p., Hardcover
ISBN: 978-3-642-33130-5