J Theor Probab (2016) 29:1458–1484
DOI 10.1007/s10959-015-0616-x

# Risk-Sensitive Control and an Abstract Collatz–Wielandt Formula

**Ari Arapostathis**[1] · **Vivek S. Borkar**[2] · **K. Suresh Kumar**[3]

Received: 17 June 2014 / Revised: 21 April 2015 / Published online: 15 May 2015


**Abstract** The 'value' of infinite horizon risk-sensitive control is the principal eigenvalue of a certain positive operator. For the case of compact domain, Chang has built upon a nonlinear version of the Krein–Rutman theorem to give a 'min–max' characterization of this eigenvalue which may be viewed as a generalization of the classical Collatz–Wielandt formula for the Perron–Frobenius eigenvalue of a nonnegative irreducible matrix. We apply this formula to the Nisio semigroup associated with risk-sensitive control and derive a variational characterization of the optimal risk-sensitive cost. For the linear, i.e., uncontrolled case, this is seen to reduce to the celebrated Donsker–Varadhan formula for principal eigenvalue of a second-order elliptic operator.

Ari Arapostathis: Supported in part by the Office of Naval Research through the Electric Ship Research and Development Consortium.
Vivek S. Borkar: Supported in part by Grant 11IRCCSG014 from IIT Bombay and a J. C. Bose Fellowship from the Department of Science and Technology, Government of India.
K. Suresh Kumar: Supported in part by DST project: SR/S4/MS:751/12.

✉ Ari Arapostathis
ari@ece.utexas.edu

Vivek S. Borkar
borkar.vs@gmail.com

K. Suresh Kumar
suresh@math.iitb.ac.in

1 Department of Electrical and Computer Engineering, The University of Texas at Austin, 1 University Station, Austin, TX 78712, USA

2 Department of Electrical Engineering, Indian Institute of Technology, Powai, Mumbai 400076, India

3 Department of Mathematics, Indian Institute of Technology, Powai, Mumbai 400076, India

**Keywords** Risk-sensitive control · Collatz–Wielandt formula · Nisio semigroup · Variational formulation · Principal eigenvalue · Donsker–Varadhan functional

**Mathematics Subject Classification (2010)** Primary 60J60 · Secondary 60F10 · 93E20

## 1 Introduction

We consider the infinite horizon risk-sensitive control problem for a controlled reflected diffusion in a bounded domain. This seeks to minimize the asymptotic growth rate of the expected 'exponential of integral' cost, which in turn coincides with the principal eigenvalue of a quasi-linear elliptic operator defined as the pointwise envelope of a family of linear elliptic operators parametrized by the 'control' parameter. The Kreĭn–Rutman theorem has been widely applied to study the time-asymptotic behavior of linear parabolic equations [15, Chapter 7]. A recent extension of the Kreĭn–Rutman theorem to positively 1-homogeneous compact (nonlinear) operators and the ensuing variational formulation for the positive eigenpair extends the classical Collatz–Wielandt formula for the Perron–Frobenius eigenvalue of irreducible nonnegative matrices. Using this, we are able to obtain a variational formulation for the positive eigenpair that reduces to the celebrated Donsker–Varadhan characterization thereof in the linear case. In the linear case, the eigenvalue in the positive eigenpair coincides with the principal eigenvalue. This is not in general true for the nonlinear case. Hence we obtain a Collatz–Wielandt formula for the unique positive eigenpair (see the example in Remark 4.2). This establishes interesting connections between theory of risk-sensitive control, nonlinear Kreĭn–Rutman theorem, and Donsker–Varadhan theory.

## 2 Risk-Sensitive Control

Let $Q \subset \mathbb{R}^d$ be an open bounded domain with a $C^3$ boundary $\partial Q$ and $\bar{Q}$ denote its closure. Consider a reflected controlled diffusion $X(\cdot)$ taking values in the bounded domain $\bar{Q}$ satisfying

$$\begin{aligned} \mathrm{d}X(t) &= b(X(t), v(t))\,\mathrm{d}t + \sigma(X(t))\,\mathrm{d}W(t) - \gamma(X(t))\,\mathrm{d}\xi(t),\\ \mathrm{d}\xi(t) &= I\,\{X(t) \in \partial Q\}\,\mathrm{d}\xi(t) \end{aligned} \tag{2.1}$$

for $t \geq 0$, with $X(0) = x$ and $\xi(0) = 0$. Here:

(a) $b : \bar{Q} \times \mathcal{V} \to \mathbb{R}^d$ for a prescribed compact metric control space $\mathcal{V}$ is continuous and Lipschitz in its first argument uniformly with respect to the second,
(b) $\sigma : \bar{Q} \to \mathbb{R}^{d\times d}$ is continuously differentiable, its derivatives are Hölder continuous with exponent $\beta_0 > 0$, and is uniformly non-degenerate in the sense that the minimum eigenvalue of

$$a(x) = [[a_{ij}(x)]]{:=}\sigma(x)\sigma^T(x)$$

is bounded away from zero.

(c) $\gamma : \mathbb{R}^d \to \mathbb{R}^d$ is co-normal, i.e., $\gamma(x) = [\gamma_1(x), \ldots, \gamma_d(x)]^T$, where

$$\gamma_i(x) = \sum_{i=1}^{d} a_{ij}(x) n_j(x), \quad x \in \partial Q,$$

$n(x) = [n_1(x), \ldots, n_d(x)]^T$ is the unit outward normal.

(d) $W(\cdot)$ is a $d$-dimensional standard Wiener process,

(e) $v(\cdot)$ is a $\mathcal{V}$-valued measurable process satisfying the non-anticipativity condition: for $t > s \geq 0$, $W(t) - W(s)$ is independent of $\{v(y), W(y) : y \leq s\}$. A process $v$ satisfying this property is called an 'admissible control.'

Let $r : \bar{Q} \times \mathcal{V} \to \mathbb{R}_+$ be a continuous 'running cost' function which is Lipschitz in its first argument uniformly with respect to the second. We define

$$r_{\max} := \max_{(x,v) \in \bar{Q} \times \mathcal{V}} |r(x, v)|.$$

The infinite horizon risk-sensitive problem aims to minimize the cost

$$\limsup_{T \uparrow \infty} \frac{1}{T} \log E\left[e^{\int_0^T r(X(s), v(s)) \mathrm{d}s}\right], \tag{2.2}$$

i.e., the mean asymptotic growth rate of the exponential of the total cost. See [16] for background and motivation.

We define

$$\begin{aligned} \mathcal{G} f(x) &:= \frac{1}{2} \operatorname{tr}\Big(a(x) \nabla^2 f(x)\Big) + \mathcal{H}\big(x, f(x), \nabla f(x)\big), \text{ where,} \\ \mathcal{H}(x, f, p) &:= \min_{v \in \mathcal{V}} \big[\langle b(x, v), p \rangle + r(x, v) f\big], \end{aligned} \tag{2.3}$$

and

$$C^2_{\gamma,+}(\bar{Q}) := \Big\{ f \in C^2(\bar{Q}) : f \geq 0, \ \nabla f \cdot \gamma = 0 \text{ on } \partial Q \Big\}.$$

The main result of the paper is the following.

**Theorem 2.1** *There exists a unique pair* $(\rho, \varphi) \in \mathbb{R} \times C^2_{\gamma,+}(\bar{Q})$ *satisfying* $\|\varphi\|_{0;\bar{Q}} = 1$ *which solves the p.d.e.*

$$\rho \varphi(x) = \mathcal{G}\varphi(x) \quad \text{in } Q, \quad \langle \nabla \varphi, \gamma \rangle = 0 \quad \text{on } \partial Q,$$

*Moreover,*

$$\rho = \inf_{f \in C^2_{\gamma,+}(\bar{Q}),\, f>0} \sup_{\nu \in \mathcal{P}(\bar{Q})} \int \frac{\mathcal{G} f}{f}\, \mathrm{d}\nu$$
$$= \sup_{f \in C^2_{\gamma,+}(\bar{Q}),\, f>0} \inf_{\nu \in \mathcal{P}(\bar{Q})} \int \frac{\mathcal{G} f}{f}\, \mathrm{d}\nu, \tag{2.4}$$

*where $\mathcal{P}(\bar{Q})$ denotes the space of probability measures on $\bar{Q}$ with the Prohorov topology.*

The first part of the theorem is contained in Lemma 4.5. The second part is proved in Sect. 4.2.

The notation used in the paper is summarized below.

*Notation 2.1* The standard Euclidean norm in $\mathbb{R}^d$ is denoted by $|\cdot|$. The set of nonnegative real numbers is denoted by $\mathbb{R}_+$, and $\mathbb{N}$ stands for the set of natural numbers. The closure, the boundary, and the complement of a set $A \subset \mathbb{R}^d$ are denoted by $\overline{A}$, $\partial A$, and $A^c$, respectively.

We adopt the notation $\partial_t := \frac{\partial}{\partial t}$, and for $i, j \in \mathbb{N}$, $\partial_i := \frac{\partial}{\partial x_i}$ and $\partial_{ij} := \frac{\partial^2}{\partial x_i \partial x_j}$. For a nonnegative multi-index $\alpha = (\alpha_1, \dots, \alpha_d)$, we let $D^\alpha := \partial_1^{\alpha_1} \cdots \partial_d^{\alpha_d}$ and $|\alpha| := \alpha_1 + \cdots + \alpha_d$. For a domain $Q$ in $\mathbb{R}^d$ and $k = 0, 1, 2, \dots$, we denote by $C^k(Q)$ the set of functions $f : Q \to \mathbb{R}$ whose derivatives $D^\alpha f$ for $|\alpha| \le k$ are continuous and bounded. For $k = 0, 1, 2, \dots$, we define

$$[f]_{k;Q} := \max_{|\alpha|=k} \sup_Q |D^\alpha f| \quad \text{and} \quad \|f\|_{k;Q} := \sum_{j=0}^k [f]_{j;Q}.$$

Also for $\delta \in (0, 1)$, we define

$$[g]_{\delta;Q} := \sup_{\substack{x,y \in Q \\ x \neq y}} \frac{|g(x) - g(y)|}{|x-y|^\delta} \quad \text{and} \quad \|f\|_{k+\delta;Q} := \|f\|_{k;Q} + \max_{|\alpha|=k} [D^\alpha f]_{\delta;Q}.$$

For $k = 0, 1, 2, \dots$, and $\delta \in (0, 1)$, we denote by $C^{k+\delta}(Q)$ the space of all real-valued functions $f$ defined on $Q$ such that $\|f\|_{k+\delta;Q} < \infty$. Unless indicated otherwise, we always view $C^{k+\delta}(Q)$ and $C^k(Q)$ as topological spaces under the norms $\|\cdot\|_{k+\delta;Q}$ and $\|\cdot\|_{k;Q}$, respectively. We also write $C^{k+\delta}(\bar{Q})$ and $C^k(\bar{Q})$ if the derivatives up to order $k$ are continuous on $\bar{Q}$. Thus $C^\delta(\bar{Q})$ stands for the Banach space of real-valued functions defined on $\bar{Q}$ that are Hölder continuous with exponent $\delta \in (0, 1)$.

Let $G$ be a domain in $\mathbb{R}_+ \times \mathbb{R}^d$. Recall that $C^{1,2}(G)$ stands for the set of bounded continuous real-valued functions $\varphi(t, x)$ defined on $G$ such that the derivatives $D^\alpha \varphi$, $|\alpha| \le 2$ and $\partial_t \varphi$ are bounded and continuous in $G$. Let $\delta \in (0, 1)$. We define

$$[\varphi]_{\delta/2,\delta;G} := \sup_{\substack{(t,x)\neq(s,y)\\(t,x),\,(s,y)\in G}} \frac{|\varphi(t,x)-\varphi(s,y)|}{|x-y|^\delta+|t-s|^{\delta/2}},$$

$$\|\varphi\|_{\delta/2,\delta;G} := \|\varphi\|_{0;G} + [\varphi]_{\delta/2,\delta;G}.$$

By $C^{\delta/2,\delta}(G)$, we denote the space of functions $\varphi$ such that $\|\varphi\|_{\delta/2,\delta;G} < \infty$. The parabolic Hölder space $C^{1+\delta/2,2+\delta}(G)$ is the set of all real-valued functions defined on $G$ for which

$$\|\varphi\|_{1+\delta/2,2+\delta;G} := \max_{|\alpha|\le 2} \|D^\alpha\varphi\|_{\delta/2,\delta;G} + \|\partial_t\varphi\|_{\delta/2,\delta;G}$$

is finite. It is well known that $C^{1+\delta/2,2+\delta}(G)$ equipped with the norm $\|\varphi\|_{1+\delta/2,2+\delta;G}$ is a Banach space.

For a Banach space $\mathcal{Y}$ of continuous functions on $\bar{Q}$, we denote by $\mathcal{Y}_+$ its positive cone and by $\mathcal{Y}_\gamma$ the subspace of $\mathcal{Y}$ consisting of the functions $f$ satisfying $\nabla f\cdot\gamma = 0$ on $\partial Q$. Also let $\mathcal{Y}^*$ denote the dual of $\mathcal{Y}$ and $\mathcal{Y}^*_+$ the dual cone of $\mathcal{Y}_+$. For example, $\big(C^2_\gamma(\bar{Q})\big)^*_+$ is defined by

$$\big(C^2_\gamma(\bar{Q})\big)^*_+ := \Big\{\Lambda \in \big(C^2_\gamma(\bar{Q})\big)^* : \Lambda(f)\ge 0 \quad \forall f\in C^2_{\gamma,+}(\bar{Q})\Big\}.$$

We define the operator $\mathcal{L}_v$ on $C^2(\bar{Q})$ by

$$\mathcal{L}_v f(\cdot) := \frac{1}{2}\operatorname{tr}\Big(a(\cdot)\nabla^2 f(\cdot)\Big) + \langle b(\cdot,v),\nabla f(\cdot)\rangle, \quad v\in\mathcal{V}, \tag{2.5}$$

where $\nabla^2$ denotes the Hessian.

## 3 The Nisio Semigroup

Associated with the above control problem, define for each $t\ge 0$ the operator $S_t : C(\bar{Q})\to C(\bar{Q})$ by

$$S_t f(x) := \inf_{v(\cdot)} E_x\left[e^{\int_0^t r(X(s),v(s))\mathrm{d}s} f(X(t))\right], \tag{3.1}$$

where the 'inf' is over all admissible controls.

A standard consequence of the dynamic programming principle is that this defines a semigroup, the so-called Nisio semigroup. In fact, the following well-known properties thereof can be proved along the lines of [14, Theorem 1, pp. 298–299]. Let

$$T^u_t f(x) := E_x\left[e^{\int_0^t r(X^u(s),u)\mathrm{d}s} f(X^u(t))\right], \tag{3.2}$$

where $X^u(\cdot)$ is the reflected diffusion in (2.1) for $v(\cdot)\equiv u\in\mathcal{V}$.

**Theorem 3.1** $\{S_t, t \geq 0\}$ *satisfies the following properties:*

(1) *Boundedness:* $\|S_t f\|_{0;\bar{Q}} \leq e^{r_{\max} t}\|f\|_{0;\bar{Q}}$. *Furthermore,* $S_t \mathbf{1} \geq e^{r_{\min} t}\mathbf{1}$, *where* $\mathbf{1}$ *is the constant function* $\equiv 1$, *and* $r_{\min} = \min_{(x,u)} r(x,u)$.
(2) *Semigroup property:* $S_0 = I$ *and* $S_t \circ S_s = S_{t+s}$ *for* $s, t \geq 0$.
(3) *Monotonicity:* $f \geq$ *(resp.,* $>$*)* $g \implies S_t f \geq$ *(resp.,* $>$*)* $S_t g$.
(4) *Lipschitz property:* $\|S_t f - S_t g\|_{0;\bar{Q}} \leq e^{r_{\max} t}\|f - g\|_{0;\bar{Q}}$.
(5) *Strong continuity:* $\|S_t f - S_s f\|_{0;\bar{Q}} \to 0$ *as* $t \to s$.
(6) *Envelope property:* $T_t^u f \geq S_t f$ *for all* $u \in U$, *and* $S_t f \geq S_t' f$ *for any other* $\{S_t'\}$ *satisfying this along with the foregoing properties.*
(7) *Generator: the infinitesimal generator of* $\{S_t\}$ *is* $\mathcal{G}$ *defined in* (2.3).

We can say more by invoking p.d.e. theory. We start with the following theorem that characterizes $S_t$ as the solution of a parabolic p.d.e.

**Theorem 3.2** *For each* $f \in C_\gamma^{2+\delta}(\bar{Q})$, $\delta \in (0, \beta_0)$, *and* $T > 0$, *the quasi-linear parabolic p.d.e.*

$$\frac{\partial}{\partial t}\psi(t,x) = \inf_{v \in \mathcal{V}} \left(\mathcal{L}_v \psi(t,x) + r(x,v)\psi(t,x)\right) \quad \text{in } (0,T] \times Q, \tag{3.3}$$

*with* $\psi(0,x) = f(x)$ *for all* $x \in \bar{Q}$ *and*

$$\langle \nabla\psi(t,x), \gamma(x)\rangle = 0 \qquad \text{for all } (t,x) \in (0,T] \times \partial Q,$$

*has a unique solution in* $C^{1+\delta/2,2+\delta}\left([0,T] \times \bar{Q}\right)$. *The solution* $\psi$ *has the stochastic representation*

$$\psi(t,x) = \inf_{v(\cdot)} E_x\left[e^{\int_0^t r(X(s),v(s))\,\mathrm{d}s} f(X(t))\right] \qquad \forall (t,x) \in [0,T] \times \bar{Q}. \tag{3.4}$$

*Moreover,*

$$\begin{aligned}
\|\psi\|_{1,2;[0,T]\times\bar{Q}} &\leq K_1, \\
\|\nabla^2\psi(s,\cdot)\|_{\delta;Q} &\leq K_2 \quad \text{for all } s \in [0,T],
\end{aligned}$$

*where the constants* $K_1, K_2 > 0$ *depend only on* $T$, $\|a\|_{1+\beta_0;Q}$, *the Lipschitz constants of* $b, r$, *the lower bound on the eigenvalues of* $a$, *the boundary* $\partial Q$ *and* $\|f\|_{2+\delta;Q}$.

*Proof* This follows by [11, Theorem 7.4, p. 491] and [11, Theorem 7.2, pp. 486–487].

**Lemma 3.1** *Let* $\delta \in (0, \beta_0)$. *For each* $t > 0$, *the map* $S_t : C_\gamma^{2+\delta}(\bar{Q}) \to C_\gamma^{2+\delta}(\bar{Q})$ *is compact.*

*Proof* Suppose $f \in C_\gamma^{2+\delta}(\bar{Q})$ for some $\delta \in (0, \beta_0)$. Fix any $T > 0$. Let $g : [0,\infty) \to [0,\infty)$ be a smooth function such that $g(0) = 0$ and $g(s) = 1$ for $s \in [T/2, \infty)$. Define $\tilde{\psi}(t,x) = g(t)\psi(t,x)$, with $\psi$ as in Theorem 3.2. Then $\tilde{\psi}$ satisfies

$$\frac{\partial}{\partial t}\tilde{\psi}(t,x) - \frac{1}{2}\operatorname{tr}\Big(a(x)\nabla^2\tilde{\psi}(t,x)\Big) = \frac{\partial g}{\partial t}(t)\psi(t,x) + g(t)\mathcal{H}\big(x,\psi(t,x),\nabla\psi(t,x)\big) \tag{3.5}$$

in $(0,\infty)\times Q$, $\tilde{\psi}(0,x)=0$ on $\bar{Q}$ and $\langle\nabla\tilde{\psi}(t,x),\gamma(x)\rangle = 0$ for all $(t,x)\in(0,\infty)\times\partial Q$. It is well known that $\partial_i$ is a bounded operator from $C^{1+\delta/2,2+\delta}\big([0,T]\times\bar{Q}\big)$ to $C^{(1+\delta)/2,1+\delta}\big([0,T]\times\bar{Q}\big)$ [10, p. 126]. In particular,

$$\sup_{x\in\bar{Q}}\sup_{s\neq t}\frac{\big|\partial_i\psi(s,x)-\partial_i\psi(t,x)\big|}{|s-t|^{(1+\delta)/2}} < \infty.$$

Since $\mathcal{H}$ is Lipschitz in its arguments and $g$ is smooth, it follows that the r.h.s. of (3.5) is in $C^{\beta/2,\beta}\big([0,T]\times\bar{Q}\big)$ for any $\beta\in(0,1)$. Then it follows by the interior estimates in [11, Theorem 10.1, pp. 351–352] that $\tilde{\psi}\in C^{1+\beta/2,2+\beta}\big([T,T+1]\times\bar{Q}\big)$ for all $\beta\in(0,\beta_0)$. Since $\psi=\tilde{\psi}$ on $[T,T+1]$, it follows that $S_T f\in C^{2+\beta}_{\gamma}(\bar{Q})$ for all $\beta\in(0,\beta_0)$. Since the inclusion $C^{2+\beta}_{\gamma}(\bar{Q})\hookrightarrow C^{2+\delta}_{\gamma}(\bar{Q})$ is compact for $\beta>\delta$, the result follows. □

## 4 An Abstract Collatz–Wielandt Formula

The classical Collatz–Wielandt formula (see [5,17]) characterizes the principal (i.e., the Perron–Frobenius) eigenvalue $\kappa$ of an irreducible nonnegative matrix $Q$ as (see [13, Chapter 8])

$$\begin{aligned}\kappa &= \max_{\{x=(x_1,\dots,x_d)\,:\,x_i\geq 0\}}\ \min_{\{i\,:\,x_i>0\}}\left(\frac{(Qx)_i}{x_i}\right)\\ &= \min_{\{x=(x_1,\dots,x_d)\,:\,x_i>0\}}\ \max_{\{i\,:\,x_i>0\}}\left(\frac{(Qx)_i}{x_i}\right).\end{aligned}$$

An infinite dimensional version of this was recently given by Chang [4] as follows. Let $\mathcal{X}$ be a real Banach space with order cone $P$, i.e., a nontrivial closed subset of $\mathcal{X}$. Define $-P{:=}\{-x : x\in P\}$ and $\dot{P} := P\backslash\{\theta\}$. We assume that the cone $P$ satisfies

(a) $tP\subset P$ for all $t\geq 0$, where $tP=\{tx : x\in P\}$;
(b) $P+P\subset P$;
(c) $P\cap(-P)=\{\theta\}$, where $\theta$ denotes the zero vector of $\mathcal{X}$.

We write $x\preceq y$ if $y-x\in P$, and $x\prec y$ if $x\preceq y$ and $x\neq y$. Define the dual cone

$$P^* := \{x\in\mathcal{X}^* : \langle x^*, x\rangle\geq 0\quad\forall x\in P\}.$$

A map $T:\mathcal{X}\to\mathcal{X}$ is said to be *increasing* if $x\preceq y\implies T(x)\preceq T(y)$, and *strictly increasing* if $x\prec y\implies T(x)\prec T(y)$. If $\operatorname{int}(P)\neq\varnothing$, and $T:\dot{P}\to\operatorname{int}(P)$, then $T$ is called *strongly positive*, and if $x\prec y\implies T(y)-T(x)\in\operatorname{int}(P)$ it is called *strongly increasing*. It is called *positively 1-homogeneous* if $T(tx)=tT(x)$ for all

$t > 0$ and $x \in \mathcal{X}$. Also, a map $T : \mathcal{X} \to \mathcal{X}$ is called *completely continuous* if it is continuous and compact. A generalization of the Kreĭn–Rutman theorem appears in [12]. However the hypotheses in [12, Theorem 2] are not sufficient for uniqueness of an eigenvector in $P$, so the conclusions of that theorem are not correct. The same error has propagated in [4, Theorems 1.4, 4.8, and 4.13]. For a detailed discussion on this, see the forthcoming paper [1]. A corrected version of [12, Theorem 2] is as follows:

**Theorem 4.1** *Let $T : \mathcal{X} \to \mathcal{X}$ be an increasing, positively 1-homogeneous, completely continuous map such that for some $u \in P$ and $M > 0$, $MT(u) \succeq u$. Then there exist $\lambda > 0$ and $\hat{x} \in \dot{P}$ such that $T(\hat{x}) = \lambda\hat{x}$. Moreover, if $T$ is strongly increasing then $\lambda$ is the unique eigenvalue with an eigenvector in $P$.*

The following is proved in [4]:

**Theorem 4.2** *Let $T$ and $\lambda$ be as in the preceding theorem. Define:*

$$P^*(x) := \left\{x^* \in P^* : \langle x^*, x\rangle > 0\right\},$$
$$r_*(T) := \sup_{x \in \dot{P}} \inf_{x^* \in P^*(x)} \frac{\langle x^*, T(x)\rangle}{\langle x^*, x\rangle},$$
$$r^*(T) := \inf_{x \in \dot{P}} \sup_{x^* \in P^*(x)} \frac{\langle x^*, T(x)\rangle}{\langle x^*, x\rangle}.$$

*If $T$ is strongly increasing then $\lambda = r^*(T) = r_*(T)$.*

Uniqueness of the positive eigenvector can be obtained under additional assumptions. In this paper, we are concerned with *superadditive* operators $T$, in other words operators $T$ which satisfy

$$T(x + y) \succeq T(x) + T(y) \qquad \forall x, y \in \mathcal{X}.$$

We have the following simple assertion:

**Corollary 4.1** *Let $T : \mathcal{X} \to \mathcal{X}$ be a superadditive, positively 1-homogeneous, strongly positive, completely continuous map. Then there exists a unique $\hat{x} \in \dot{P}$ with $\|\hat{x}\| = 1$, where $\|\cdot\|$ denotes the norm in $\mathcal{X}$, such that $T(\hat{x}) = \lambda\hat{x}$, with $\lambda > 0$.*

*Proof* It is clear that strong positivity implies that for any $x \in \mathcal{X}$, there exists $M > 0$ such that $MT(x) \succeq x$. By superadditivity, $T(x - y) \preceq T(x) - T(y)$. Hence if $x \succ y$, by strong positivity, we obtain $T(x) - T(y) \in \operatorname{int}(P)$. Therefore every superadditive, strongly positive map is strongly increasing. Existence of a unique eigenvalue with an eigenvector in $P$ then follows by Theorem 4.1. Suppose $\hat{x}$ and $\hat{y}$ are two distinct unit eigenvectors in $P$. Since by strong positivity $\hat{x}$ and $\hat{y}$ are in $\operatorname{int}(P)$, there exists $\alpha > 0$ such that $\hat{x} - \alpha\hat{y} \in \dot{P} \setminus \operatorname{int}(P)$. Since $T$ is strongly increasing, we obtain

$$\lambda(\hat{x} - \alpha\hat{y}) = T(\hat{x}) - T(\alpha\hat{y}) \succeq T(\hat{x} - \alpha\hat{y}) \in \operatorname{int}(P),$$

a contradiction. Uniqueness of a unit eigenvector in $P$ follows. □

An application of Theorem 4.1 and Corollary 4.1 provides us with the following result for strongly continuous semigroups of operators.

**Corollary 4.2** *Let $\mathcal{X}$ be a Banach space with order cone $P$ having non-empty interior. Let $\{S_t,\ t \geq 0\}$ be a strongly continuous semigroup of superadditive, strongly positive, positively 1-homogeneous, completely continuous operators on $\mathcal{X}$. Then there exists a unique $\rho \in \mathbb{R}$ and a unique $\hat{x} \in \operatorname{int}(P)$, with $\|\hat{x}\| = 1$, such that $S_t\hat{x} = e^{\rho t}\hat{x}$ for all $t \geq 0$.*

*Proof* By Theorem 4.1 and Corollary 4.1, there exists a unique $\lambda(t) > 0$ and a unique $x_t \in P$ satisfying $\|x_t\| = 1$, such that $S_t x_t = \lambda(t)x_t$. By the uniqueness of a unit eigenvector in $P$ and the semigroup property, it follows that there exists $\hat{x} \in \mathcal{X}$ such that $x_t = \hat{x}$ for all dyadic rational numbers $t > 0$. On the other hand, from the strong continuity, it follows that if a sequence of dyadic rationals $t_n \geq 0$, $n \geq 1$ converges to some $t > 0$, then $\lambda(t_n)$ is a Cauchy sequence and its limit point $\lambda'$ is an eigenvalue of $S_t$ corresponding to the eigenvector $\hat{x}$ and therefore $\lambda(t) = \lambda'$ and $x_t = \hat{x}$ by the uniqueness thereof. Strong continuity then implies that $\lambda(\cdot)$ is continuous, and by the semigroup property and positive 1-homogeneity, we have $\lambda(t+s) = \lambda(t)\lambda(s)$ for all for $t, s > 0$. It follows that $\lambda(t) = e^{\rho t}$ for some $\rho \in \mathbb{R}$. □

### 4.1 Stability

Concerning the time-asymptotic behavior of $S_t x$, we have the following.

**Theorem 4.3** *Let $\mathcal{X}$, $\{S_t\}$, $\rho$ and $\hat{x}$ be as in Corollary 4.2. Then*

(i) *The set*

$$\mathcal{O}_1 := \left\{ e^{-\rho t} S_t x : x \in P,\ \|x\| \leq 1,\ t \geq 1 \right\}$$

*is relatively compact in $\mathcal{X}$.*

(ii) *There exists $\alpha^*(x) \in \mathbb{R}_+$ such that*

$$\lim_{t\to\infty} \left\| e^{-\rho t} S_t x - \alpha^*(x)\,\hat{x} \right\| \xrightarrow[t\to\infty]{} 0 \quad \forall x \in \dot{P}.$$

(iii) *Suppose that additionally the following properties hold:*

(P1) *For every $M > 0$, there exist $\tau \in (0,1)$ and a positive constant $\zeta_0 = \zeta_0(M)$ such that*

$$\|S_\tau(\hat{x} - z)\| + \|S_\tau z\| \;\geq\; \zeta_0$$

*for all $z \in P$ such that $z \preceq \hat{x}$ and $\|z\| \leq M$.*

(P2) *For every compact set $\mathcal{K} \subset P$, there exists a constant $\zeta_1 = \zeta_1(\mathcal{K})$ such that $x \in \mathcal{K}$ and $x \preceq \alpha\,\hat{x}$ imply $\|x\| \leq \alpha\,\zeta_1$.*

*Then the convergence is exponential: There exists $M_0 > 0$ and $\theta_0 > 0$ such that*

$$\|e^{-\rho t} S_t x - \alpha^*(x)\,\hat{x}\| \;\le\; M_0 e^{-\theta_0 t}\,\|x\| \quad \text{for all } t \ge 0 \ \text{ and all } x \in \dot{P}.$$

*Proof* Without loss of generality, we can assume $\rho = 0$. For $t \ge 0$ and $x \in P$, we define

$$\underline{\alpha}(x) := \sup\,\{a \in \mathbb{R} : x - a\,\hat{x} \in P\}$$
$$\overline{\alpha}(x) := \inf\,\{a \in \mathbb{R} : a\,\hat{x} - x \in P\}.$$

Since $\hat{x} \in \text{int}(P)$, it follows that $\underline{\alpha}(x)$ and $\overline{\alpha}(x)$ are finite and $\overline{\alpha}(x) \ge \underline{\alpha}(x) \ge 0$. Note also that for $x \in \dot{P}$, we have $\overline{\alpha}(x) > 0$ and since $S_t x \in \text{int}(P)$, we have $\underline{\alpha}(S_t x) > 0$ for all $t > 0$. It is also evident from the definition that

$$\underline{\alpha}(\lambda x) = \lambda\,\underline{\alpha}(x) \quad \text{and} \quad \overline{\alpha}(\lambda x) = \lambda\,\overline{\alpha}(x) \quad \text{for all } x \in \dot{P},\ \lambda \in \mathbb{R}_+.$$

By the increasing property and the positive 1-homogeneity of $S_t$, we obtain $S_{t+s}x - \underline{\alpha}(S_s x)\,\hat{x} \in P$ for all $x \in P$ and $t \ge 0$ and this implies that $\underline{\alpha}(S_{t+s}x) \ge \underline{\alpha}(S_s x)$ for all $t \ge 0$ and $x \in P$. It follows that for any $x \in P$, the map $t \mapsto \underline{\alpha}(S_t x)$ is non-decreasing. Similarly, the map $t \mapsto \overline{\alpha}(S_t x)$ is non-increasing.

We next show that the orbit $\mathcal{O}$ of the unit ball in $P$ defined by

$$\mathcal{O} := \left\{S_t x : x \in P,\ \|x\| \le 1,\ t \ge 0\right\}$$

is bounded. Suppose not. Then we can select a sequence $\{x_n\} \subset \dot{P}$ with $\|x_n\| = 1$, and an increasing sequence $\{t_n,\ n \in \mathbb{N}\}$ such that $\|S_{t_n} x_n\| \to \infty$ as $n \to \infty$ and such that $\|S_{t_n} x_n\| \ge \|S_t x_n\|$ for all $t \le t_n$. By the properties of the sequence $\{S_{t_n}\}$, the sequence $\left\{\frac{S_{t_n-2}\,x_n}{\|S_{t_n} x_n\|}\right\}$ is bounded and this implies that $\left\{\frac{S_{t_n-1}\,x_n}{\|S_{t_n} x_n\|}\right\}$ is relatively compact. Let $y \in \mathcal{X}$ be any limit point of $\frac{S_{t_n-1}\,x_n}{\|S_{t_n} x_n\|}$ as $n \to \infty$. By continuity of $S_1$, it follows that $\|S_{t_n}\,x_n\| \le k_1 \|S_{t_n-1}\,x_n\|$ for some $k_1 > 0$. This implies that $\|y\| \ge k_1^{-1}$. Therefore $y \in \dot{P}$ which in turn implies that $\underline{\alpha}(S_1 y) > 0$. It is straightforward to show that the map $x \mapsto \underline{\alpha}(x)$ is continuous. Therefore, we have

$$\underline{\alpha}\left(\frac{S_{t_n} x_n}{\|S_{t_n} x_n\|}\right) = \underline{\alpha}\left(S_1\left(\frac{S_{t_n-1} x_n}{\|S_{t_n} x_n\|}\right)\right) \xrightarrow[n\to\infty]{} \underline{\alpha}(S_1 y). \tag{4.1}$$

On the other hand, it holds that

$$\underline{\alpha}(S_{t_n} x_n) = \|S_{t_n} x_n\|\;\underline{\alpha}\left(\frac{S_{t_n} x_n}{\|S_{t_n} x_n\|}\right). \tag{4.2}$$

Since $\hat{x} \in \text{int}(P)$, the constant $\kappa_1$ defined by

$$\kappa_1 := \sup_{x \in P,\ \|x\|=1} \overline{\alpha}(x) \tag{4.3}$$

is finite. Since $\underline{\alpha}(S_1 y) > 0$ and $\|S_{t_n} x_n\|$ diverges, (4.1)–(4.2) imply that $\underline{\alpha}(S_{t_n} x_n)$ diverges which is impossible since

$$\underline{\alpha}(S_{t_n} x_n) \le \overline{\alpha}(S_{t_n} x_n) \le \overline{\alpha}(x_n) \le \kappa_1.$$

Since $\mathcal{O}$ is bounded in $\mathcal{X}$, there exists a constant $k_0$ such that

$$\|S_t x\| \le k_0 \|x\| \qquad \forall t \in [0, 1], \quad \forall x \in P. \tag{4.4}$$

That the set $\mathcal{O}_1$ is relatively compact for each $x \in \mathcal{X}$ now easily follows. Indeed, since $\mathcal{O}(x)$ is bounded, by the semigroup property, we obtain

$$\mathcal{O}_1 = \bigl\{S_1(S_{t-1}x) : x \in P, \ \|x\| = 1, \ t \ge 1\bigr\} \subset S_1\bigl(\mathcal{O}\bigr),$$

and the claim follows since by hypothesis $S_1$ is a compact map.

For all $t \ge s \ge 0$, we have

$$S_t\bigl(S_s x - \underline{\alpha}(S_s x)\,\hat{x}\bigr) \preceq S_{t+s}x - \underline{\alpha}(S_s x)\,\hat{x}, \tag{4.5}$$

$$S_t\bigl(\overline{\alpha}(S_s x)\,\hat{x} - S_s x\bigr) \preceq \overline{\alpha}(S_s x)\,\hat{x} - S_{t+s}x. \tag{4.6}$$

Let $s = t_n$ in (4.5) and take limits along some converging sequence $S_{t_n}x \to \bar{x}$ as $n \to \infty$, for some $\bar{x} \in P$, to obtain

$$\underline{\alpha}^*(x)\hat{x} + S_t\bigl(\bar{x} - \underline{\alpha}^*(x)\hat{x}\bigr) \preceq S_t\bar{x}, \tag{4.7}$$

where $\underline{\alpha}^*(x) := \lim_{t\uparrow\infty} \underline{\alpha}(S_t x)$. Since $\bar{x}$ is an $\omega$-limit point of $S_t x$, it follows that $\underline{\alpha}(S_t\bar{x}) = \underline{\alpha}^*(x)$ for all $t \ge 0$. Therefore $S_t\bar{x} - \underline{\alpha}^*(x)\hat{x} \notin \mathrm{int}(P)$ for all $t \ge 0$, which implies by (4.7) and the strong positivity of $S_t$ that $\bar{x} - \underline{\alpha}^*(x)\hat{x} = 0$. A similar argument shows that $\bar{x} = \overline{\alpha}^*(x)\hat{x}$, where $\overline{\alpha}^*(x) := \lim_{t\uparrow\infty} \overline{\alpha}(S_t x)$. We let $\alpha^* := \overline{\alpha}^* = \underline{\alpha}^*$.

It remains to prove that convergence is exponential. Since the orbit $\mathcal{O}$ is bounded and $\hat{x} \in \mathrm{int}(P)$, it follows that the set $\{\overline{\alpha}(S_t x) : t \ge 0, \ x \in P, \ \|x\| \le 1\}$ is bounded. Therefore since the orbit $\mathcal{O}_1$ is also relatively compact, it follows that the set

$$\mathcal{K}_1 := \bigl\{S_k x - \underline{\alpha}(S_k x)\hat{x}, \ \overline{\alpha}(S_k x)\hat{x} - S_k x : k \ge 1, \ x \in P, \ \|x\| \le 1\bigr\}$$

is a relatively compact subset of $P$. Define

$$\eta(S_k x) := \overline{\alpha}(S_k x) - \underline{\alpha}(S_k x), \quad k = 1, 2, \ldots$$

By property (P2), since

$$S_k x - \underline{\alpha}(S_k x)\hat{x} \preceq \eta(S_k x)\,\hat{x},$$

$$\overline{\alpha}(S_k x)\hat{x} - S_k x \preceq \eta(S_k x)\,\hat{x},$$

it follows that for some $\zeta_1 = \zeta_1(\mathcal{K}_1)$, we have

$$\max\left\{\|S_k x - \underline{\alpha}(S_k x)\hat{x}\|,\ \|\overline{\alpha}(S_k x)\hat{x} - S_k x\|\right\} \le \zeta_1\, \eta(S_k x) \tag{4.8}$$

for all $k \ge 1$ and $x \in P$ with $\|x\| \le 1$. Define

$$\underline{\mathcal{Z}}_k(x) := \frac{S_k x - \underline{\alpha}(S_k x)\hat{x}}{\eta(S_k x)}, \qquad \overline{\mathcal{Z}}_k(x) := \frac{\overline{\alpha}(S_k x)\hat{x} - S_k x}{\eta(S_k x)},$$

provided $\eta(S_k x) \neq 0$, which is equivalent to $S_k x \neq \hat{x}$. By (4.8) the set

$$\tilde{\mathcal{K}}_1 := \left\{\underline{\mathcal{Z}}_k(x),\ \overline{\mathcal{Z}}_k(x) : k \ge 1,\ x \in \dot{P} \setminus \{\hat{x}\},\ \|x\| \le 1\right\}$$

lies in the ball of radius $\zeta_1$ centered at the origin of $\mathcal{X}$. Therefore, since $\overline{\mathcal{Z}}_k(x) = \hat{x} - \underline{\mathcal{Z}}_k(x)$, by property (P1), there exists $\zeta_0 = \zeta_0(\zeta_1) > 0$ and $\tau \in (0, 1)$ such that

$$\|S_\tau \underline{\mathcal{Z}}_k(x)\| + \|S_\tau \overline{\mathcal{Z}}_k(x)\| \ge \zeta_0 \qquad \forall k = 1, 2, \ldots, \quad \forall x \in \dot{P} \setminus \{\hat{x}\},\ \|x\| \le 1 \tag{4.9}$$

Let

$$A_k(x) := \sup\left\{\alpha \in \mathbb{R} : \{S_1 \underline{\mathcal{Z}}_k(x) - \alpha\,\hat{x}\} \cup \{S_1 \overline{\mathcal{Z}}_k(x) - \alpha\,\hat{x}\} \subset P\right\}.$$

We claim that

$$\zeta_2 := \inf\left\{A_k(x) : k \ge 1,\ x \in \dot{P} \setminus \{\hat{x}\},\ \|x\| \le 1\right\} > 0. \tag{4.10}$$

Indeed, if the claim is not true then by (4.9) and the definition of $A_k$, there exists a sequence $z_k$ taking values in

$$\left\{\underline{\mathcal{Z}}_k(x), \overline{\mathcal{Z}}_k(x) : x \in \dot{P} \setminus \{\hat{x}\},\ \|x\| \le 1\right\}$$

for each $k = 1, 2, \ldots$, such that $\|S_\tau z_k\| \ge \zeta_0/2$ and such that $\underline{\alpha}(S_1 z_k) \to 0$ as $k \to \infty$. However, since $\tilde{\mathcal{K}}_1$ is bounded, it follows that $S_\tau\big(\tilde{\mathcal{K}}_1\big)$ is a relatively compact subset of $\mathrm{int}(P)$. Therefore the limit set of $S_\tau z_k$ is non-empty and any limit point $y \in P$ of $S_\tau z_k$ satisfies $\|y\| \ge \zeta_0/2$. Since $\underline{\alpha}(S_1 z_k) = \underline{\alpha}(S_{1-\tau} S_\tau z_k)$ and $z \mapsto \underline{\alpha}(S_{1-\tau} z)$ is continuous on $P$, any such limit point $y$ satisfies $\underline{\alpha}(S_{1-\tau} y) = 0$ which contradicts the strong positivity hypothesis.

Equation (4.10) implies that

$$\underline{\alpha}\big(S_1\big(\overline{\alpha}(S_k x)\,\hat{x} - S_k x\big)\big) + \underline{\alpha}\big(S_1\big(S_k x - \underline{\alpha}(S_k x)\,\hat{x}\big)\big) \ge \zeta_2\big(\overline{\alpha}(S_k x) - \underline{\alpha}(S_k x)\big) \tag{4.11}$$

for all $x \in \dot{P} \setminus \{\hat{x}\}$ with $\|x\| \le 1$, and by 1-homogeneity, for all $x \in \dot{P} \setminus \{\hat{x}\}$.

By (4.5)–(4.6), we have

$$S_1\big(S_k x - \underline{\alpha}(S_k x)\,\hat{x}\big) \preceq S_{k+1}x - \underline{\alpha}(S_k x)\,\hat{x},$$
$$S_1\big(\overline{\alpha}(S_k x)\,\hat{x} - S_k x\big) \preceq \overline{\alpha}(S_k x)\,\hat{x} - S_{k+1}x. \tag{4.12}$$

In turn, (4.12) implies that

$$\underline{\alpha}(S_{k+1}x) \geq \underline{\alpha}(S_k x) + \underline{\alpha}\big(S_1\big(S_k x - \underline{\alpha}(S_k x)\,\hat{x}\big)\big),$$
$$\overline{\alpha}(S_{k+1}x) \leq \overline{\alpha}(S_k x) - \underline{\alpha}\big(S_1\big(\overline{\alpha}(S_k x)\,\hat{x} - S_k x\big)\big). \tag{4.13}$$

By (4.11) and (4.13), we obtain that

$$\eta(S_k x) - \eta(S_{k+1}x) \geq \zeta_2\,\eta(S_k x),$$

which we write as

$$\eta(S_{k+1}x) \leq (1-\zeta_2)\,\eta(S_k x), \qquad k = 1, 2, \ldots \tag{4.14}$$

We add the inequalities

$$\|S_k x - \alpha^*(x)\,\hat{x}\| \leq \|S_k x - \underline{\alpha}(S_k x)\,\hat{x}\| + \alpha^*(x) - \underline{\alpha}(S_k x),$$
$$\|\alpha^*(x)\,\hat{x} - S_k x\| \leq \|\overline{\alpha}(S_k x)\,\hat{x} - S_k x\| + \overline{\alpha}(S_k x) - \alpha^*(x)$$

and use (4.8) and (4.14) to obtain

$$\begin{aligned} 2\,\|S_k x - \alpha^*(x)\,\hat{x}\| &\leq 2\zeta_1\,\eta(S_k x) + \eta(S_k x) \\ &\leq (2\zeta_1+1)\eta(S_k x) \\ &\leq (2\zeta_1+1)(1-\zeta_2)^{k-1}\,\eta(S_1 x), \qquad k = 1, 2, \ldots \end{aligned} \tag{4.15}$$

We have

$$\begin{aligned} \eta(S_1 x) &= \overline{\alpha}(S_1 x) - \underline{\alpha}(S_1 x) \\ &\leq \overline{\alpha}(S_1 x) \\ &\leq \kappa_1\,\|S_1 x\| \\ &\leq \kappa_1\,k_0\,\|x\|, \end{aligned} \tag{4.16}$$

where $k_0$ is the continuity constant in (4.4), and $\kappa_1$ is defined in (4.3). Let $\lfloor t \rfloor$ denote the integral part of a number $t \in \mathbb{R}_+$. We define

$$M_0 := \frac{\kappa_1\,k_0^2\,(2\zeta_1+1)}{2} \qquad \text{and} \qquad \theta_0 := -\log(1-\zeta_2),$$

and combine (4.15)–(4.16) to obtain

$$\begin{aligned}\|S_t x - \alpha^*(x)\,\hat{x}\| &\leq \frac{M_0}{k_0}(1-\zeta_2)^{\lfloor t\rfloor -1}\big\|S_{t-\lfloor t\rfloor}x\big\| \\ &\leq M_0 e^{-\theta_0 t}\,\|x\|.\end{aligned}$$

The proof is complete. □

*Remark 4.1* Recall that the cone $P$ is called *normal* if there exists a constant $K$ such that $\|x\| \leq K\|y\|$ whenever $0 \preceq x \preceq y$. It might appear that property (P2) in Theorem 4.3 is weaker than normality of the cone. However it turns out that (P2) together with the fact that $\hat{x}$ is an interior point imply that $P$ is normal. This is shown in Lemma 4.1 below.

Also $\tau$ in (P1) in Theorem 4.3 can be any positive constant and need not be restricted to lie in (0, 1). The proof of geometric convergence follows in the same manner, by using the iterates $S_{k(\tau+1)}$ instead of $S_k$.

**Lemma 4.1** *Consider the following properties:*

(P2′) *There exists a constant* $\zeta_1' > 0$ *such that* $x \in P$ *and* $x \preceq \hat{x}$ *imply* $\|x\| \leq \zeta_1'$.
(P2″) $P$ *is normal.*

*Then* (P2) $\Longleftrightarrow$ (P2′) $\Longleftrightarrow$ (P2″)

*Proof* If (P2′) does not hold, then there exists $\{x_n\} \subset P$ with $x_n \preceq \hat{x}$ and $\|x_n\| \nearrow \infty$. Hence $\{\|x_n\|^{-2}x_n\}$ is precompact, and since $\|x_n\|^{-2}x_n \preceq \|x_n\|^{-2}\hat{x}$, this implies by (P2) that $\|x_n\|^{-2}\|x_n\| \leq \|x_n\|^{-2}\zeta_1$ for some $\zeta_1 > 0$. This contradicts $\|x_n\| \nearrow \infty$ and so (P2) cannot hold. Therefore (P2) $\Longrightarrow$ (P2′). The other direction is obvious.

Since $\hat{x} \in \operatorname{int}(P)$, there exists $\varepsilon > 0$ such that $\|y\| \leq \varepsilon$ implies that $y \preceq \hat{x}$. Suppose $0 \preceq x \preceq y$. By scaling, we have

$$0 \preceq \frac{\varepsilon}{\|y\|}x \preceq \frac{\varepsilon}{\|y\|}y \preceq \hat{x}. \tag{4.17}$$

Then (P2′) and (4.17) imply that $\frac{\varepsilon}{\|y\|}\|x\| \leq \zeta_0$ or that $\|x\| \leq \frac{\zeta_0}{\varepsilon}\|y\|$. Therefore (P2′) is equivalent to normality of the cone $P$. □

It is also the case that (P1)–(P2) are weaker than *uniform strong positivity* property which is defined as

(H1) There exists $\tau > 0$ and $\xi > 0$ such that $S_\tau x \succeq \xi\|x\|\,\hat{x}$ for all $x \in P$,

or in a seemingly weaker form as

(H1′) For any compact subset $\mathcal{K} \subset P$, there exists $\tau = \tau(\mathcal{K}) > 0$ and $\xi = \xi(\mathcal{K}) > 0$ such that $S_\tau x \succeq \xi\|x\|\,\hat{x}$ for all $x \in \mathcal{K}$.

We first show that (H1) and (H1′) are equivalent.

**Lemma 4.2** (H1) $\Longleftrightarrow$ (H1′).

*Proof* Obviously (H1) $\Longrightarrow$ (H1′).

To prove the converse, suppose (H1) does not hold. Then there exists a sequence $\{x_n\} \subset P$ with $\|x_n\| = 1$ and a sequence $\tau_n \nearrow \infty$ such that $\underline{\alpha}(S_{\tau_n} x_n) \searrow 0$. Hence $\underline{\alpha}(S_{\tau_n} x_n)x_n \searrow 0$, so that the set $\{\underline{\alpha}(S_{\tau_n} x_n)x_n\}$ is precompact. Therefore by (H1′), there exists $\tau > 0$ and $\xi > 0$ such that $S_\tau(\underline{\alpha}(S_{\tau_n} x_n)x_n) \succeq \xi \underline{\alpha}(S_{\tau_n} x_n)\hat{x}$ which is equivalent (by 1-homogeneity) to $S_\tau x_n \succeq \xi \hat{x}$. But $S_\tau x_n \succeq \xi \hat{x}$ implies that $\underline{\alpha}(S_\tau x_n) \geq \xi$. Since $\underline{\alpha}(S_\tau x_n) \leq \underline{\alpha}(S_{\tau_n} x_n)$ whenever $\tau_n \geq \tau$, we obtain a contradiction with the property $\underline{\alpha}(S_{\tau_n} x_n) \searrow 0$. Therefore (H1′) cannot hold and the proof is complete. □

We need the following lemma.

**Lemma 4.3** *Provided* $\operatorname{int}(P) \neq \varnothing$, *then for every* $x \in \dot{P}$, *there exists* $C_0 = C_0(x) > 0$ *such that* $y \succeq x$ *implies* $\|y\| \geq C_0$.

*Proof* Fix any $x_0 \in \operatorname{int}(P)$. If the assertion in the lemma is not true, there exists $\{y_n\} \subset P$ with $\|y_n\| \searrow 0$ such that $y_n \succeq x$. Then since $x_0 \in \operatorname{int}(P)$ there exists a sequence $\varepsilon_n \searrow 0$, such that $\varepsilon_n x_0 \succeq y_n$. But this implies $\varepsilon_n x_0 \succeq x$ and taking limits as $n \to \infty$, we have $0 \succeq x$ which contradicts $x \in \dot{P}$. □

We next show that uniform strong positivity implies (P1)–(P2).

**Lemma 4.4** (H1) $\Longrightarrow$ (P1)–(P2).

*Proof* By (H1), we have

$$\begin{aligned} S_\tau(\hat{x} - z) + S_\tau z \;&\succeq\; \xi \|\hat{x} - z\|\, \hat{x} + \xi \|z\|\, \hat{x} \\ &\succeq\; \xi \|\hat{x}\|\, \hat{x}. \end{aligned} \tag{4.18}$$

By (4.18) and Lemma 4.3, we have

$$\begin{aligned} \|S_\tau(\hat{x} - z)\| + \|S_\tau z\| \;&\geq\; \|S_\tau(\hat{x} - z) + S_\tau z\| \\ &\geq\; C_0\, \xi \|\hat{x}\|. \end{aligned} \tag{4.19}$$

It is clear that (4.19) is stronger than (P1), since it holds for any $z \preceq \hat{x}$.

Next we show that (H1) $\Longrightarrow$ (P2). By Lemma 4.2, it is enough to show that (H1′) $\Longrightarrow$ (P2). By the increasing property, $x \preceq \alpha\hat{x}$ implies $S_\tau x \preceq \alpha\hat{x}$, which combined with (H1′) implies that $\xi \|x\| \hat{x} \preceq \alpha\hat{x}$, which in turn implies $\|x\| \leq \xi^{-1}\alpha$. □

### 4.2 The Positive Eigenpair of the Nisio Semigroup

We now return to the Nisio semigroup in (3.1).

**Lemma 4.5** *There exists a unique pair* $(\rho, \varphi) \in \mathbb{R} \times C^2_{\gamma,+}(\bar{Q})$ *satisfying* $\|\varphi\|_{0;\bar{Q}} = 1$ *such that*

$$S_t\varphi = e^{\rho t}\varphi, \quad t \geq 0.$$

*The pair* $(\rho, \varphi)$ *is a solution to the p.d.e.*

$$\rho\,\varphi(x) = \mathcal{G}\varphi(x) = \inf_{v\in\mathcal{V}} \big(\mathcal{L}_v\varphi(x) + r(x,v)\varphi(x)\big) \quad in\ Q, \quad \langle\nabla\varphi, \gamma\rangle = 0 \quad on\ \partial Q, \tag{4.20}$$

*where* (4.20) *specifies* $\rho$ *uniquely in* $\mathbb{R}$ *and* $\varphi$, *with* $\|\varphi\|_{0;\bar{Q}} = 1$, *uniquely in* $C^2_{\gamma,+}(\bar{Q})$.

*Proof* It is clear that $S_t$ is superadditive. If $f \in C^2_{\gamma,+}(\bar{Q})$, then (3.4) implies that the solution $\psi$ of (3.3) is nonnegative. Moreover by the strong maximum principle [9, Theorem 3, p. 38] and the Hopf boundary lemma [9, Theorem 14, p. 49], it follows that $\psi(t,\cdot) > 0$ for all $t > 0$. Hence the strong positivity hypothesis in Corollary 4.2 is satisfied. Since also the compactness hypothesis holds by Lemma 3.1, the first statement follows by Corollary 4.2. That (4.20) holds follows from (7) of Theorem 3.1 (see also [3, pp. 73–75]). Uniqueness follows from the following argument. Suppose $\hat{\rho} \in \mathbb{R}$ and $\hat{\varphi} \in C^2_{\gamma,+}(\bar{Q})$ solve

$$\hat{\rho}\,\hat{\varphi}(x) \;=\; \inf_{v\in\mathcal{V}} \big(\mathcal{L}_v\hat{\varphi}(x) + r(x,v)\hat{\varphi}(x)\big).$$

Then by direct substitution, we have

$$\begin{aligned}\frac{\partial}{\partial t}\big(e^{\hat{\rho}t}\,\hat{\varphi}(x)\big) &= \hat{\rho}\,e^{\hat{\rho}t}\,\hat{\varphi}(x)\\ &= \inf_{v\in\mathcal{V}} \big[\mathcal{L}_v\big(e^{\hat{\rho}t}\hat{\varphi}(x)\big) + r(x,v)\big(e^{\hat{\rho}t}\hat{\varphi}(x)\big)\big].\end{aligned}$$

Therefore, $S_t\hat{\varphi} = e^{\hat{\rho}t}\,\hat{\varphi}$, and by the uniqueness assertion in Corollary 4.2, we have $\hat{\rho} = \rho$ and $\hat{\varphi} = C\varphi$ for some positive constant $C$. □

*Remark 4.2* Consider the operator $R_t : C^{2+\delta}_{\gamma}(\bar{Q}) \to C^{2+\delta}_{\gamma}(\bar{Q})$ defined by $R_t f = -S_t(-f)$. Then by same arguments as in the proof of Lemma 4.5 using Corollary 4.2, there exists a unique $\beta \in \mathbb{R}$ and $\psi > 0$ in $C^{2+\delta}_{\gamma}(\bar{Q})$ such that

$$R_t\psi \;=\; e^{\beta t}\psi.$$

Hence the pair $(e^{\beta t}, -\psi)$ is an eigenvalue–function pair of $S_t$. Now the same arguments as in the proof of Lemma 4.5 lead to the conclusion that $(\beta, \psi)$ is the unique positive solution pair of

$$\beta\,\psi(x) = \sup_{v\in\mathcal{V}} \big(\mathcal{L}_v\psi(x) + r(x,v)\psi(x)\big) \quad \text{in } Q, \quad \langle\nabla\psi, \gamma\rangle = 0 \quad \text{on } \partial Q,$$

Hence $(\beta, -\psi)$ is the unique solution pair of (4.20) satisfying $-\psi < 0$. Moreover it is easy to see that $\rho \le \beta$ and that $\beta$ is the principal eigenvalue of both operators $R_t$, $S_t$. This leads to the conclusion that the risk-sensitive control problem where the controller tries to maximize the risk-sensitive cost (2.2) leads to the value $\beta$ which is the principal eigenvalue.

*Remark 4.3* The p.d.e. in (4.20) is the Hamilton–Jacobi–Bellman equation for the risk-sensitive control problem [2].

**Lemma 4.6** *Let $\mathcal{M}(\bar{Q})$ denote the space of finite Borel measures on $\bar{Q}$. Then*

$$\big(C^2_\gamma(\bar{Q})\big)^*_+ = \mathcal{M}(\bar{Q}).$$

*Proof* Let $\Lambda \in \big(C^2_\gamma(\bar{Q})\big)^*_+$. Then for $f \in C^2_\gamma(\bar{Q})$ by positivity of $\Lambda$ we have

$$\begin{aligned}
\big|\Lambda(f)\big| &= \big|\Lambda(f + \|f\|_{0;\bar{Q}} \cdot \mathbf{1}) - \Lambda(\|f\|_{0;\bar{Q}} \cdot \mathbf{1})\big| \\
&\le \max\big\{\Lambda(f + \|f\|_{0;\bar{Q}} \cdot \mathbf{1}), \Lambda(\|f\|_{0;\bar{Q}} \cdot \mathbf{1})\big\} \\
&\le \Lambda(2\|f\|_{0;\bar{Q}} \cdot \mathbf{1}) \\
&= 2\|f\|_{0;\bar{Q}}\Lambda(\mathbf{1}).
\end{aligned}$$

It follows that $\Lambda$ is a bounded linear functional on the linear subspace $C^2_\gamma(\bar{Q})$ of $C(\bar{Q})$. By the Hahn–Banach theorem, $\Lambda$ can be extended to some $\psi \in \big(C(\bar{Q})\big)^*$. Clearly $\psi$ is a positive linear functional. By the Riesz representation theorem, there exists $\mu \in \mathcal{M}(\bar{Q})$ such that $\psi(f) = \int_{\bar{Q}} f\,d\mu$ for all $f \in C(\bar{Q})$. Therefore $\Lambda(f) = \int_{\bar{Q}} f\,d\mu$ for all $f \in C^2_\gamma(\bar{Q})$. This shows that $\big(C^2_\gamma(\bar{Q})\big)^*_+ \subset \mathcal{M}(\bar{Q})$. It is clear that $\mathcal{M}(\bar{Q}) \subset \big(C^2_\gamma(\bar{Q})\big)^*_+$, so equality follows. □

**Lemma 4.7** *Let $\delta \in (0, \beta_0)$. Then for any $f \in C^{2+\delta}_{\gamma,+}(\bar{Q})$, we have*

$$\limsup_{t\downarrow 0} \inf_{\substack{\mu\in\mathcal{M}(\bar{Q})\\ \int f\,\mathrm{d}\mu=1}} \int_{\bar{Q}} \frac{S_t f(x) - f(x)}{t}\,\mu(\mathrm{d}x) = \inf_{\substack{\mu\in\mathcal{M}(\bar{Q})\\ \int f\,\mathrm{d}\mu=1}} \int_{\bar{Q}} \mathcal{G} f(x)\,\mu(\mathrm{d}x)$$

*and*

$$\liminf_{t\downarrow 0} \sup_{\substack{\mu\in\mathcal{M}(\bar{Q})\\ \int f\,\mathrm{d}\mu=1}} \int_{\bar{Q}} \frac{S_t f(x) - f(x)}{t}\,\mu(\mathrm{d}x) = \sup_{\substack{\mu\in\mathcal{M}(\bar{Q})\\ \int f\,\mathrm{d}\mu=1}} \int_{\bar{Q}} \mathcal{G} f(x)\,\mu(\mathrm{d}x).$$

*Proof* Note that

$$\lim_{t\downarrow 0} \frac{S_t f(x) - f(x)}{t} = \mathcal{G} f(x), \quad x \in \bar{Q}.$$

Hence using the dominated convergence theorem,[1] we obtain, for all $\mu \in \mathcal{M}(\bar{Q})$ satisfying $\int f d\mu = 1$,

$$\lim_{t\downarrow 0} \int_{\bar{Q}} \frac{S_t f(x) - f(x)}{t} \mu(\mathrm{d}x) = \int_{\bar{Q}} \mathcal{G} f(x)\, \mu(\mathrm{d}x).$$

Therefore

$$\begin{aligned}\limsup_{t\downarrow 0} \inf_{\substack{\tilde{\mu}\in\mathcal{M}(\bar{Q})\\ \int f\,\mathrm{d}\tilde{\mu}=1}} \int_{\bar{Q}} \frac{S_t f(x) - f(x)}{t} \tilde{\mu}(\mathrm{d}x) &\le \lim_{t\downarrow 0} \int_{\bar{Q}} \frac{S_t f(x) - f(x)}{t}\, \mu(\mathrm{d}x)\\ &= \int_{\bar{Q}} \mathcal{G} f(x)\, \mu(\mathrm{d}x)\end{aligned}$$

for all $\mu \in \mathcal{M}(\bar{Q})$ satisfying $\int f d\mu = 1$. Hence

$$\limsup_{t\downarrow 0} \inf_{\substack{\mu\in\mathcal{M}(\bar{Q})\\ \int f\,\mathrm{d}\mu=1}} \int_{\bar{Q}} \frac{S_t f(x) - f(x)}{t} \mu(\mathrm{d}x) \le \inf_{\substack{\mu\in\mathcal{M}(\bar{Q})\\ \int f\,\mathrm{d}\mu=1}} \int_{\bar{Q}} \mathcal{G} f(x)\, \mu(\mathrm{d}x). \tag{4.21}$$

Since for each $t > 0$, the map $\mu \mapsto \int_{\bar{Q}} \frac{S_t f(x) - f(x)}{t} \mu(\mathrm{d}x)$ from $\mathcal{M}(\bar{Q}) \to \mathbb{R}$ is continuous, there exists a $\mu_t \in \mathcal{M}(\bar{Q})$ satisfying $\int f \mathrm{d}\mu_t = 1$ such that

$$\inf_{\substack{\mu\in\mathcal{M}(\bar{Q})\\ \int f\,\mathrm{d}\mu=1}} \int_{\bar{Q}} \frac{S_t f(x) - f(x)}{t}\, \mu(\mathrm{d}x) = \int_{\bar{Q}} \frac{S_t f(x) - f(x)}{t}\, \mu_t(\mathrm{d}x).$$

Clearly $\{\mu_t\}$ is tight. Let $\hat{\mu}$ be a limit point of $\mu_t$ as $t \to 0$. Suppose $\mu_{t_n} \to \hat{\mu}$ in $\mathcal{M}(\bar{Q})$ as $t_n \downarrow 0$. Then $\int f d\hat{\mu} = 1$. Note that for $f \in C^{2+\delta}_{\gamma,+}(\bar{Q})$,

$$\frac{S_t f(x) - f(x)}{t} = \frac{1}{t} \int_0^t \partial_s u^f(s, x)\, \mathrm{d}s, \tag{4.22}$$

with $u^f(t, \cdot) := S_t f(\cdot)$. By the Hölder continuity of $\partial_s u^f$ on $[0,\ 1] \times \bar{Q}$, there exists $k_1 > 0$ such that

$$|\partial_s u^f(s, x) - \partial_s u^f(s, y)| < k_1 |x - y|^\delta \qquad \forall x, y \in \bar{Q},\ s \in [0,\ 1]. \tag{4.23}$$

[1] Note that

$$\begin{aligned}\left|\frac{S_t f(x) - f(x)}{t}\right| &\le \inf_{v(\cdot)} \frac{1}{t} E_x\left[\int_0^t e^{\int_0^s r(X_z, v_z)\mathrm{d}z} |\mathcal{L}_{v_s} f(X_s) + r(X_s, v_s) f(X_s)|\, \mathrm{d}s\right]\\ &\le K e^{r_{\max}}, \quad 0 \le t \le 1,\end{aligned}$$

for some constant $K > 0$.

Therefore by (4.22) and (4.23), $x \mapsto \frac{S_t f(x) - f(x)}{t}$ is Hölder equicontinuous over $t \in (0,\ 1]$, and the convergence

$$\lim_{t \downarrow 0} \frac{S_t f(x) - f(x)}{t} = \mathcal{G} f(x)$$

is uniform in $\bar{Q}$. Hence from

$$\int_{\bar{Q}} \frac{S_{t_n} f(x) - f(x)}{t_n} \mu_{t_n}(\mathrm{d}x) = \int_{\bar{Q}} \left( \frac{S_{t_n} f(x) - f(x)}{t_n} - \mathcal{G} f(x) \right) \mu_{t_n}(\mathrm{d}x) + \int_{\bar{Q}} \mathcal{G} f(x)\, \mu_{t_n}(\mathrm{d}x),$$

it follows that

$$\lim_{n \to \infty} \int_{\bar{Q}} \left( \frac{S_{t_n} f(x) - f(x)}{t_n} \right) \mu_{t_n}(\mathrm{d}x) = \int_{\bar{Q}} \mathcal{G} f(x)\, \hat{\mu}(\mathrm{d}x) \geq \inf_{\substack{\mu \in \mathcal{M}(\bar{Q}) \\ \int f\, \mathrm{d}\mu = 1}} \int_{\bar{Q}} \mathcal{G} f(x)\, \mu(\mathrm{d}x).$$

Hence

$$\limsup_{t \downarrow 0} \int_{\bar{Q}} \left( \frac{S_t f(x) - f(x)}{t} \right) \mu_t(\mathrm{d}x) \ \geq \inf_{\substack{\mu \in \mathcal{M}(\bar{Q}) \\ \int f\, \mathrm{d}\mu = 1}} \int_{\bar{Q}} \mathcal{G} f(x)\, \mu(\mathrm{d}x). \quad (4.24)$$

From (4.21) and (4.24), the result follows. The proof of the second limit follows by a symmetric argument. □

We next prove the main result.

*Proof of Theorem 2.1* Let $\delta \in (0, \beta_0)$. Since $\rho\, \varphi = \mathcal{G}\varphi$ by Lemma 4.5, we obtain

$$\rho = \inf_{\substack{\mu \in \mathcal{M}(\bar{Q}) \\ \int \varphi\, \mathrm{d}\mu = 1}} \int \mathcal{G}\varphi\, \mathrm{d}\mu \leq \sup_{f \in C^{2+\delta}_{\gamma,+}(\bar{Q})} \inf_{\substack{\mu \in \mathcal{M}(\bar{Q}) \\ \int f\, \mathrm{d}\mu = 1}} \int \mathcal{G} f\, \mathrm{d}\mu.$$

To show the reverse inequality, we use Theorem 4.2 and Lemma 4.6. We have

$$e^{\rho t} = \sup_{g \in C^{2+\delta}_{\gamma,+}(\bar{Q})} \inf_{\substack{\mu \in \mathcal{M}(\bar{Q}) \\ \int g\, \mathrm{d}\mu = 1}} \int S_t g\, \mathrm{d}\mu.$$

Therefore, using Lemma 4.7, we obtain

$$\begin{aligned} \rho &= \lim_{t\downarrow 0} \sup_{g\in C^{2+\delta}_{\gamma,+}(\bar{Q})} \inf_{\substack{\mu\in\mathcal{M}(\bar{Q})\\ \int g\,\mathrm{d}\mu=1}} \int \frac{S_t g - g}{t}\,\mathrm{d}\mu \\ &\geq \limsup_{t\downarrow 0} \inf_{\substack{\mu\in\mathcal{M}(\bar{Q})\\ \int f\,\mathrm{d}\mu=1}} \int \frac{S_t f - f}{t}\,\mathrm{d}\mu \\ &= \inf_{\substack{\mu\in\mathcal{M}(\bar{Q})\\ \int f\,\mathrm{d}\mu=1}} \int \mathcal{G} f\,\mathrm{d}\mu \end{aligned}$$

for all $f \in C^{2+\delta}_{\gamma,+}(\bar{Q})$. Therefore,

$$\rho \;\geq\; \sup_{f\in C^{2+\delta}_{\gamma,+}(\bar{Q})} \inf_{\substack{\mu\in\mathcal{M}(\bar{Q})\\ \int f\,\mathrm{d}\mu=1}} \int \mathcal{G} f\,\mathrm{d}\mu.$$

Using a symmetric argument to establish the first equality in (4.25) below, we obtain

$$\begin{aligned} \rho &= \inf_{f\in C^{2+\delta}_{\gamma,+}(\bar{Q})} \sup_{\substack{\mu\in\mathcal{M}(\bar{Q})\\ \int f\,\mathrm{d}\mu=1}} \int \mathcal{G} f\,\mathrm{d}\mu \\ &= \sup_{f\in C^{2+\delta}_{\gamma,+}(\bar{Q})} \inf_{\substack{\mu\in\mathcal{M}(\bar{Q})\\ \int f\,\mathrm{d}\mu=1}} \int \mathcal{G} f\,\mathrm{d}\mu \end{aligned} \tag{4.25}$$

for all $\delta \in (0, \beta_0)$. Note that the outer 'inf' and 'sup' in (4.25) are realized at the function $\varphi$ in Lemma 4.5. Therefore, since $\varphi > 0$, equation (4.25) remains valid if we restrict the outer 'inf' and 'sup' on $f > 0$. Hence using the probability measure $\mathrm{d}\nu = f\,\mathrm{d}\mu$, we can write (4.25) as

$$\begin{aligned} \rho &= \inf_{f\in C^{2+\delta}_{\gamma,+}(\bar{Q}),\, f>0} \sup_{\nu\in\mathcal{P}(\bar{Q})} \int \frac{\mathcal{G} f}{f}\,\mathrm{d}\nu \\ &= \sup_{f\in C^{2+\delta}_{\gamma,+}(\bar{Q}),\, f>0} \inf_{\nu\in\mathcal{P}(\bar{Q})} \int \frac{\mathcal{G} f}{f}\,\mathrm{d}\nu. \end{aligned}$$

Therefore

$$\inf_{f\in C^{2}_{\gamma,+}(\bar{Q}),\, f>0} \sup_{\nu\in\mathcal{P}(\bar{Q})} \int \frac{\mathcal{G} f}{f}\,\mathrm{d}\nu \;\leq\; \rho \;\leq\; \sup_{f\in C^{2}_{\gamma,+}(\bar{Q}),\, f>0} \inf_{\nu\in\mathcal{P}(\bar{Q})} \int \frac{\mathcal{G} f}{f}\,\mathrm{d}\nu. \tag{4.26}$$

Suppose that the inequality on the r.h.s. of (4.26) is strict. Then for some $\hat{f} \in C^2_{\gamma,+}(\bar{Q})$, we have

$$\inf_{\nu \in \mathcal{P}(\bar{Q})} \int \frac{\mathcal{G}\hat{f}}{\hat{f}} \, \mathrm{d}\nu \;>\; \rho.$$

Since $\mathcal{G} : C^2_{\gamma,+}(\bar{Q}) \to C^0(\bar{Q})$ is continuous and since $C^{2+\delta}_{\gamma,+}(\bar{Q})$ is dense in $C^2_{\gamma,+}(\bar{Q})$ in the $\|\cdot\|_{2;\bar{Q}}$ norm, there exists $g \in C^{2+\delta}_{\gamma,+}(\bar{Q})$, $g > 0$, such that $\min_{\bar{Q}} \frac{\mathcal{G}g}{g} > \rho$. However this contradicts Theorem 4.2 which means that the first equality in (2.4) must hold. The proof of the second equality in (2.4) is similar. The last assertion of the theorem follows via the change of measure $f \, \mathrm{d}\mu = \mathrm{d}\nu$. □

*Remark 4.4* As pointed out in the proof of Theorem 2.1, the outer 'inf,' respectively, 'sup' in (2.4) and (4.25) are in fact 'min,' 'max' attained by $\varphi$.

Concerning the stability of the semigroup, we have the following lemma.

**Lemma 4.8** *There exist $M > 0$ and $\theta > 0$ such that for any $f \in C^2_{\gamma,+}(\bar{Q})$, we have*

$$\left\| e^{-\rho t} S_t f - \alpha^*(f)\varphi \right\|_{0;\bar{Q}} \;\le\; M e^{-\theta t} \|f\|_{0;\bar{Q}} \qquad \forall t \ge 1,$$

*for some $\alpha^*(f) \in \mathbb{R}_+$.*

*Proof* Without loss of generality, we assume $\varrho = 0$. We first verify that property (P1) of Theorem 4.3 holds. Let $\tau = 1/2$. We claim that there exists a constant $c_0 > 0$ such that

$$\left(E^v_x[f(X_\tau)]\right)^2 \;\le\; c_0 \, E^{v'}_x[f(X_\tau)] \qquad \forall f \in C(\bar{Q}),\; 0 \le f \le \varphi, \tag{4.27}$$

and for all Markov controls $v$, $v'$ and $x \in \bar{Q}$. The proof of (4.27) is as follows. To distinguish between processes, let $Y$, $Z$ denote the processes corresponding to the controls $v$, $v'$, respectively. Then using Girsanov's theorem, it follows that if we define

$$\begin{aligned} F(\tau) := & \int_0^\tau \sigma^{-1}(Y_t)\left[b(Y_t, v_t) - b(Y_t, v'_t)\right] \mathrm{d}W_t \\ & -\frac{1}{2}\int_0^\tau \|\sigma^{-1}(Y_t)\left[b(Y_t, v_t) - b(Y_t, v'_t)\right]\|^2 \mathrm{d}t, \end{aligned}$$

then

$$\begin{aligned} E_x[f(Y_\tau)] &= E_x\left[e^{F(\tau)} f(Z_\tau)\right] \\ &\le \left(E_x\left[f^2(Z_\tau)\right]\right)^{1/2}\left(E_x\left[e^{2F(\tau)}\right]\right)^{1/2} \\ &\le \left(E_x\left[f^2(Z_\tau)\right]\right)^{1/2}\left(E_x\left[e^{\int_0^\tau \|\sigma^{-1}(Y_t)[b(Y_t,v_t)-b(Y_t,v'_t)]\|^2 \mathrm{d}t}\right)^{1/2} \\ &\le c_1\left(E_x\left[f^2(Z_\tau)\right]\right)^{1/2} \\ &\le c_1 \|\varphi\|^{1/2}_{0;Q}\left(E_x\left[f(Z_\tau)\right]\right)^{1/2} \end{aligned}$$

where $c_1 > 0$ is a constant which only depends on the bounds of $\sigma^{-1}$ and $b$. This proves (4.27). For $f \in C(\bar{Q})$ satisfying $0 \le f \le \varphi$ and for any fixed $v$, we have

$$\begin{aligned} S_\tau(\varphi - f)(x) \;&\ge\; e^{r_{\min}} E_x^{v_1}\bigl[\varphi(X_\tau) - f(X_\tau)\bigr] \\ &\ge\; e^{r_{\min}} c_0^{-1}\bigl(E_x^{v}\bigl[\varphi(X_\tau) - f(X_\tau)\bigr]\bigr)^2 \end{aligned} \tag{4.28}$$

and

$$\begin{aligned} S_\tau(f)(x) \;&\ge\; e^{r_{\min}} E_x^{v_2}\bigl[f(X_\tau)\bigr] \\ &\ge\; e^{r_{\min}} c_0^{-1}\bigl(E_x^{v}\bigl[f(X_\tau)\bigr]\bigr)^2, \end{aligned} \tag{4.29}$$

where $v_1$, $v_2$ are the corresponding minimizers. Note that[2]

$$\bigl(E_x^v\bigl[\varphi(X_\tau) - f(X_\tau)\bigr]\bigr)^2 + \bigl(E_x^v\bigl[f(X_\tau)\bigr]\bigr)^2 \;\ge\; \frac{1}{2}\bigl(E_x^v\bigl[\varphi(X_\tau)\bigr]\bigr)^2 \;\ge\; \frac{1}{2}\bigl(\min \varphi\bigr)^2. \tag{4.30}$$

Adding (4.28) and (4.29) and using (4.30), it follows that

$$\|S_\tau(\varphi - f)\| + \|S_\tau f\| \;\ge\; \frac{e^{r_{\min}}}{2c_0}\bigl(\min \varphi\bigr)^2,$$

which establishes property (P1). On the other hand, property (P2) of Theorem 4.3 is trivially satisfied under the $\|\cdot\|_{0;\bar{Q}}$ norm. Hence the result follows by Theorem 4.3 (iii). □

### 4.3 The Donsker–Varadhan Functional

Let $U = \{u\}$, i.e., a singleton, and $v(\cdot) \equiv v{:=}\delta_u$, thus reducing the problem to an uncontrolled one. Thus $\mathcal{G} = \mathcal{L}_v + r(x, v)$ is a linear operator. By [6, Lemma 2, pp. 781–782], the first equality in (2.4) equals the Donsker–Varadhan functional

$$\sup_{\nu \in \mathcal{P}(\bar{Q})} \left(\int_{\bar{Q}} r(x, v)\, \nu(\mathrm{d}x) - I(\nu)\right),$$

where

$$I(\nu) := -\inf_{f \in C^2_{\gamma,+}(\bar{Q}),\, f > 0} \int \frac{\mathcal{L}_v f}{f}\, \mathrm{d}\nu.$$

[2] The first part of the inequality below follows from the fact that $(a - x)^2 + x^2$, $0 \le x \le a$ attains it minimum at $x = \frac{a}{2}$.

More generally, if $r(x, v)$ does not depend on $v$, say $r(x, v) = r(x)$ and $\mathcal{A}$ is defined by

$$\mathcal{A} f(x) := \frac{1}{2} \operatorname{tr}\Big(a(x)\nabla^2 f(x)\Big) + \min_{v\in\mathcal{V}} \big[\langle b(x, v), \nabla f(x)\rangle\big],$$

then

$$\rho = \sup_{\nu\in\mathcal{P}(\bar{Q})} \left(\int_{\bar{Q}} r(x)\,\nu(\mathrm{d}x) - I(\nu)\right),$$
$$I(\nu) = -\inf_{f\in C^2_{\gamma,+}(\bar{Q}),\ f>0} \int \frac{\mathcal{A} f}{f}\,\mathrm{d}\nu.$$

This also takes the form

$$\rho = \sup_{x\in\bar{Q}} \Big(r(x) - \tilde{I}(x)\Big),$$
$$\tilde{I}(x) := -\inf_{f\in C^2_{\gamma,+}(\bar{Q}),\ f>0} \frac{\mathcal{A} f(x)}{f(x)}.$$

Our results thus provide a counterpart of the Donsker–Varadhan functional for the nonlinear case arising from control.

It is also interesting to consider the substitution $f = e^{\psi}$. Then we obtain

$$\begin{aligned}
\rho &= \inf_{\psi\in C^2_{\gamma}(\bar{Q})} \sup_{\nu\in\mathcal{P}(\bar{Q})} \int \inf_{v\in\mathcal{V}} \sup_{w\in\mathbb{R}^d} \Big(r(\cdot, v) - \frac{1}{2}\|w\|^2 + \mathcal{L}_v\psi + \langle\nabla\psi, \sigma w\rangle\Big)\,\mathrm{d}\nu \\
&= \sup_{\psi\in C^2_{\gamma}(\bar{Q})} \inf_{\nu\in\mathcal{P}(\bar{Q})} \int \inf_{v\in\mathcal{V}} \sup_{w\in\mathbb{R}^d} \Big(r(\cdot, v) - \frac{1}{2}\|w\|^2 + \mathcal{L}_v\psi + \langle\nabla\psi, \sigma w\rangle\Big)\,\mathrm{d}\nu \\
&= \inf_{\psi\in C^2_{\gamma}(\bar{Q})} \sup_{\nu\in\mathcal{P}(\bar{Q})} \int \sup_{v\in\mathcal{V}} \inf_{w\in\mathbb{R}^d} \Big(r(\cdot, v) - \frac{1}{2}\|w\|^2 + \mathcal{L}_v\psi + \langle\nabla\psi, \sigma w\rangle\Big)\,\mathrm{d}\nu \\
&= \sup_{\psi\in C^2_{\gamma}(\bar{Q})} \inf_{\nu\in\mathcal{P}(\bar{Q})} \int \sup_{v\in\mathcal{V}} \inf_{w\in\mathbb{R}^d} \Big(r(\cdot, v) - \frac{1}{2}\|w\|^2 + \mathcal{L}_v\psi + \langle\nabla\psi, \sigma w\rangle\Big)\,\mathrm{d}\nu,
\end{aligned}$$

where the last two expressions follow from the standard Ky Fan min–max theorem [7]. This is the standard logarithmic transformation to convert the Hamilton–Jacobi–Bellman equation for risk-sensitive control to the Hamilton–Jacobi–Isaacs equation for an associated zero-sum ergodic stochastic differential game [8], given by

$$\inf_{v\in\mathcal{V}} \sup_{w\in\mathbb{R}^d} \Big(r(\cdot, v) - \frac{1}{2}\|w\|^2 + \mathcal{L}_v\psi + \langle\nabla\psi, \sigma w\rangle\Big) = \rho \tag{4.31}$$

in $Q$, with $\langle\nabla\psi, \gamma\rangle = 0$ on $\partial Q$. The expressions above bear the same relationship with (4.31) as what Lemma 4.5 and Remark 4.3 spell out for (4.20).

## 5 Risk-Sensitive Control with Periodic Coefficients

In this section, we consider risk-sensitive control with periodic coefficients. Consider a controlled diffusion $X(\cdot)$ taking values in $\mathbb{R}^d$ satisfying

$$\mathrm{d}X(t) = b(X(t), v(t))\,\mathrm{d}t + \sigma(X(t))\,\mathrm{d}W(t) \tag{5.1}$$

for $t \geq 0$, with $X(0) = x$.

We assume that

(1) The functions $b(x, v)$, $\sigma(x)$ and the running cost $r(x, v)$ are periodic in $x_i$, $i = 1, 2, \ldots, d$. Without loss of generality, we assume that the period equals 1.
(2) $b : \mathbb{R}^d \times \mathcal{V} \to \mathbb{R}^d$ is continuous and Lipschitz in its first argument uniformly with respect to the second,
(3) $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ is continuously differentiable, its derivatives are Hölder continuous with exponent $\beta_0 > 0$, and is non-degenerate,
(4) $r : \mathbb{R}^d \times \mathcal{V} \to \mathbb{R}$ is continuous and Lipschitz in its first argument uniformly with respect to the second. We let $r_{\max} := \max_{(x,v)\in\bar{Q}\times\mathcal{V}} |r(x, v)|$.

Admissible controls are defined as in (e).

We consider here as well the infinite horizon risk-sensitive problem which aims to minimize the cost in (2.2) under the controlled process governed by (5.1). Recall the notation defined in Sect. 2 and note that $C^0(\mathbb{R}^d)$ is the space of all continuous and bounded real-valued functions on $\mathbb{R}^d$. We define the semigroups of operators $\{S_t,\ t \geq 0\}$ and $\{T_t^u,\ t \geq 0\}$ acting on $C^0(\mathbb{R}^d)$ as in (3.1)–(3.2) relative to the controlled process governed by (5.1). Also the operators $\mathcal{L}_v : C^2(\mathbb{R}^d) \to C^0(\mathbb{R}^d)$ are as defined in (2.5).

Let $C_p(\mathbb{R}^d)$ denote the set of all $C^0(\mathbb{R}^d)$ functions with period 1, and in general, if $\mathcal{X}$ is a subset of $C^0(\mathbb{R}^d)$, we let $\mathcal{X}_p(\mathbb{R}^d){:=}\mathcal{X} \cap C_p(\mathbb{R}^d)$.

We start with the following theorem which is analogous to Theorem 3.1.

**Theorem 5.1** $\{S_t,\ t \geq 0\}$ *acting on* $C^0(\mathbb{R}^d)$ *satisfies the following properties:*

(1) *Boundedness:* $\|S_t f\|_{0;\mathbb{R}^d} \leq e^{r_{\max} t}\|f\|_{0;\mathbb{R}^d}$. *Furthermore,* $S_t\mathbf{1} \geq e^{r_{\min} t}\mathbf{1}$, *where* $\mathbf{1}$ *is the constant function* $\equiv 1$.
(2) *Semigroup property:* $S_0 = I$, $S_t \circ S_s = S_{t+s}$ *for* $s, t \geq 0$.
(3) *Monotonicity:* $f \geq$ *(resp.,* $>$*)* $g \implies S_t f \geq$ *(resp.,* $>$*)* $S_t g$.
(4) *Lipschitz property:* $\|S_t f - S_t g\|_{0;\mathbb{R}^d} \leq e^{r_{\max} t}\|f - g\|_{0;\mathbb{R}^d}$.
(5) *Strong continuity:* $\|S_t f - S_s f\|_{0;\mathbb{R}^d} \to 0$ *as* $t \to s$.
(6) *Envelope property:* $T_t^u f \geq S_t f$ *for all* $u \in U$ *and* $S_t f \geq S_t' f$ *for any other* $\{S_t'\}$ *satisfying this along with the foregoing properties.*
(7) *Generator: the infinitesimal generator of* $\{S_t\}$ *is given by* (2.3).
(8) *For* $f \in C_p(\mathbb{R}^d)$, $S_t f \in C_p(\mathbb{R}^d)$, $t \geq 0$.

*Proof* Properties (1)–(4) and (6) follow by standard arguments from (3.1) and the bound on $r$. That $S_t : C^0(\mathbb{R}^d) \to C^0(\mathbb{R}^d)$ is well known. See Remark 5.1 below. Property (8) follows from (3.1) and the periodicity of the data. □

**Theorem 5.2** *For $f \in C_p^{2+\delta}(\mathbb{R}^d)$, $\delta \in (0, \beta_0)$, the p.d.e.*

$$\frac{\partial}{\partial t} u(t,x) = \inf_{v \in \mathcal{V}} \big(\mathcal{L}_v u(t,x) + r(x,v)u(t,x)\big) \quad \text{in } \mathbb{R}_+ \times \mathbb{R}^d, \tag{5.2}$$

*with $u(0,x) = f(x)$ $\forall x \in \mathbb{R}^d$ has a unique solution in $C_p^{1+\delta/2,2+\delta}\big([0,T] \times \mathbb{R}^d\big)$, $T > 0$. The solution $\psi$ has the stochastic representation*

$$u(t,x) = \inf_{v(\cdot)} E_x\left[e^{\int_0^t r(X(s),v(s))\,\mathrm{d}s} f(X(t))\right] \qquad \forall (t,x) \in [0,\infty) \times \mathbb{R}^d. \tag{5.3}$$

*Moreover, for some $K_T > 0$ depending on $T$, $\delta$, $\|f\|_{2+\delta;\mathbb{R}^d}$ and the bounds on the data, we have*

$$\|u\|_{1+\delta/2,2+\delta;[0,T]\times B_R} \le K_T.$$

*Proof* Without loss of generality, we assume that $f$ is nonnegative. Consider the p.d.e.

$$\frac{\partial}{\partial t} u^R(t,x) = \inf_v \big(\mathcal{L}_v u^R(t,x) + r(x,v)u^R(t,x)\big) \quad \text{in } \mathbb{R}_+ \times B_R,$$

with $u^R = 0$ on $\mathbb{R}_+ \times \partial B_R$ and with $u^R(0,x) = f(x)g(R^{-1}x)$ for all $x \in B_R$, where $g$ is a smooth nonnegative, radially non-decreasing function which equals 1 on $\bar{B}_{\frac{1}{2}}$ and 0 on $B^c_{\frac{3}{4}}$. From [11, Theorem 6.1, pp. 452–453], the p.d.e. (5.2) has a unique solution $u^R$ in $C^{1+\delta/2,2+\delta}\big([0,T] \times \bar{B}_R\big)$, $T > 0$. This solution has the stochastic representation

$$u^R(t,x) = \inf_{v(\cdot)} E_x\left[e^{\int_0^{t\wedge\tau_R} r(X(s),v(s))\,\mathrm{d}s} f(X(t\wedge\tau_R))g(R^{-1}X(t\wedge\tau_R))\right]$$

for all $(t,x) \in [0,\infty) \times \mathbb{R}^d$, where $\tau_R$ denotes the first exit time from the ball $B_R$. Clearly then $R \mapsto u^R$ is non-decreasing. By [11, Theorem 5.2, p. 320], for each $T > 0$, there exists a constant $K_T$ such that

$$\|u^R\|_{1+\delta/2,2+\delta;[0,T]\times B_R} \le K_T.$$

Therefore $u^R$ converges to a function $u \in C^{1+\delta/2,2+\delta}\big([0,T] \times \bar{\mathbb{R}}^d\big)$, as $R \to \infty$, which satisfies (5.2)–(5.3). The periodicity of $u(t,x)$ in $x$ follows by (5.3) and the periodicity of the coefficients. □

*Remark 5.1* The regularity of the initial condition $f$ is only needed to obtain continuous second derivatives at $t = 0$. It is well known that for each $f \in C^0(\mathbb{R}^d)$, (5.2) has a solution in $C\big([0,T] \times \mathbb{R}^d\big) \cap C^{1+\delta/2,2+\delta}_{\mathrm{loc}}\big((0,T) \times \mathbb{R}^d\big)$, for $T > 0$.

**Theorem 5.3** *There exists a unique $\rho \in \mathbb{R}$ and a $\varphi > 0$ in $C_p^2(\mathbb{R}^d)$ unique up to a scalar multiple such that*

$$S_t\varphi = e^{\rho t}\varphi, \quad t > 0.$$

*Proof* Using Theorem 5.2, one can show as in the proof of Lemma 3.1 that $S_t : C_p^2(\mathbb{R}^d) \to C_p^2(\mathbb{R}^d)$ is compact for each $t \geq 0$. Now with $\mathcal{X} = C_p^2(\mathbb{R}^d)$ and $P = \{f \in C_p^2(\mathbb{R}^d) : f \geq 0\}$ and $T = S_t$ for some $t \geq 0$, the conditions of Theorems 4.1 and 4.2 are easily verified using Theorem 5.1. Repeating the same argument as in the proof of Corollary 4.2 completes the proof. □

**Lemma 5.1** *The pair* $(\rho, \varphi)$ *given in Theorem 5.3 is a solution to the p.d.e.*

$$\rho\,\varphi(x) = \inf_v \big(\mathcal{L}_v\varphi(x) + r(x, v)\varphi(x)\big), \tag{5.4}$$

*where (5.4) specifies* $\rho$ *uniquely in* $\mathbb{R}$ *and* $\varphi$ *uniquely in* $C_p^2(\mathbb{R}^d)$ *up to a scalar multiple. Moreover,* $\inf_{\mathbb{R}^d} \varphi > 0$.

*Proof* The proof is directly analogous to that of Lemma 4.5. □

**Lemma 5.2** $(C_p^2(\mathbb{R}^d))^* \simeq \mathcal{M}(Q)$, *with* $Q = [0, 1)^d$.

*Proof* Let $\pi$ denote the projection of $\mathbb{R}^d$ to $[0, 1)^d$. Set

$$\mathcal{D} = \{f \circ \pi \in C(Q) : f \in C_p(\mathbb{R}^d)\}.$$

Then $\mathcal{D}$ is a linear subspace of $C^0(Q)$.

For $\Lambda \in (C_p(\mathbb{R}^d))^*$, define the linear map $\tilde{\Lambda} : \mathcal{D} \to \mathbb{R}$ by

$$\tilde{\Lambda}(f \circ \pi) = \Lambda(f).$$

Then

$$|\tilde{\Lambda}(f \circ \pi)| \;\leq\; \|\Lambda\|\|f\|_{0;\mathbb{R}^d} \;\leq\; \|\Lambda\|\|f \circ \pi\|_{0;Q}.$$

i.e., $\tilde{\Lambda} \in \mathcal{D}^*$. Using the Hahn–Banach theorem, there exists a continuous linear extension $\Lambda' : C^0(Q) \to \mathbb{R}$ of $\tilde{\Lambda}$ such that $\|\Lambda'\| = \|\tilde{\Lambda}\|$.

Since $\big(C^0(Q)\big)^* = \mathcal{M}(Q)$, the set of all finite signed Radon measures, we have $(C_p(\mathbb{R}^d))^* \subseteq \mathcal{M}(Q)$. The reverse inequality follows easily. Hence $(C_p(\mathbb{R}^d))^* = \mathcal{M}(Q)$. Now the analogous argument in Lemma 4.6 can be used to complete the proof. □

Now by closely mimicking the proofs of Lemma 4.7 and Theorem 2.1, we have

**Theorem 5.4** $\rho$ *satisfies*

$$\begin{aligned}\rho &= \inf_{f \in C_+^2(Q)\cap\mathcal{D}} \; \sup_{\mu \in \mathcal{M}(Q)\,:\, \int f\,\mathrm{d}\mu = 1} \int \mathcal{G} f \,\mathrm{d}\mu \\ &= \sup_{f \in C_+^2(Q)\cap\mathcal{D}} \; \inf_{\mu \in \mathcal{M}(Q)\,:\, \int f\,\mathrm{d}\mu = 1} \int \mathcal{G} f \,\mathrm{d}\mu,\end{aligned}$$

*where* $\mathcal{G}$ *given in Theorem 5.1.*

The stability of the semigroup also follows as in Lemma 4.8. It is well known that (5.1) has a transition probability density $p(t, x, y)$ which is bounded away from zero, uniformly over all Markov controls $v$, for $t = 1$ and $x$, $y$ in a compact set. It is straightforward to show that this implies property (P1). Therefore exponential convergence follows by Theorem 4.3 (iii).

## References

1. Arapostathis, A.: A Correction to a Nonlinear Version of the Kreĭn–Rutman Theorem by Rajesh Mahadevan (preprint)
2. Biswas, A.: An eigenvalue approach to the risk sensitive control problem in near monotone case. Syst. Control Lett. **60**, 181–184 (2011)
3. Borkar, V.S.: Optimal Control of Diffusion Processes. Pitman Research Notes in Mathematics Series, vol. 203. Longman Scientific & Technical, Harlow (1989)
4. Chang, K.C.: A nonlinear Krein Rutman theorem. J. Syst. Sci. Complex. **22**, 542–554 (2009)
5. Collatz, L.: Einschließungssatz für die charakteristischen Zahlen von Matrizen. Math. Z. **48**, 221–226 (1942)
6. Donsker, M.D., Varadhan, S.R.S.: On a variational formula for the principal eigenvalue for operators with maximum principle. Proc. Nat. Acad. Sci. USA **72**, 780–783 (1975)
7. Fan, Ky: Fixed-point and minimax theorems in locally convex topological linear spaces. Proc. Nat. Acad. Sci. USA **38**, 121–126 (1952)
8. Fleming, W.H., McEneaney, W.M.: Risk-sensitive control on an infinite time horizon. SIAM J. Control Optim. **33**, 1881–1915 (1995)
9. Friedman, A.: Partial Differential Equations of Parabolic Type. Prentice-Hall, Englewood Cliffs (1964)
10. Krylov, N.V.: Lectures on Elliptic and Parabolic Equations in Hölder Spaces. Graduate Studies in Mathematics, vol. 12. American Mathematical Society, Providence (1996)
11. Ladyženskaja, O. A., Solonnikov, V. A., Ural'ceva, N. N.: Linear and Quasilinear Equations of Parabolic Type, Translated from the Russian by S. Smith. Translations of Mathematical Monographs, vol. 23. American Mathematical Society, Providence (1968)
12. Mahadevan, R.: A note on a non-linear Krein–Rutman theorem. Nonlinear Anal. **67**, 3084–3090 (2007)
13. Meyer, C.: Matrix Analysis and Applied Linear Algebra. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (2000)
14. Nisio, M.: On stochastic optimal controls and envelope of Markovian semigroups. In: Proceedings of the International Symposium on Stochastic Differential Equations (Res. Inst. Math. Sci., Kyoto Univ., Kyoto, 1976), pp. 297–325. New York, Wiley (1978)
15. Smith, H.L.: Monotone Dynamical Systems. Mathematical Surveys and Monographs, vol. 41. American Mathematical Society, Providence (1995)
16. Whittle, P.: Risk-Sensitive Optimal Control, Wiley-Interscience Series in Systems and Optimization. Wiley, Chichester (1990)
17. Wielandt, H.: Unzerlegbare, nicht negative Matrizen. Math. Z. **52**, 642–648 (1950)