Appeared in STOCHASTIC ANALYSIS AND APPLICATIONS, Vol. 21, pp.763–799, 2003

Lyapunov Functions and Stationary Distributions of Stochastic Evolution Equations

Gottlieb Leha and Gunter Ritter

Fakultät für Mathematik und Informatik, Universität Passau
Passau, Germany

We investigate stationary distributions of diffusion processes on Hilbert spaces by means of Lyapunov functions: existence, uniqueness, and attractivity. The emphasis is on solutions to stochastic differential equations with *nonlinear* diffusion coefficients and *unbounded* and *nonlinear* drift. The methods are applicable to stochastic partial differential equations such as multidimensional stochastic reaction-diffusion equations. The method provides also information on the supports of the stationary distributions.

Key words: Stationary distributions of stochastic evolution equations; Lyapunov functions; stochastic differential equations; multidimensional stochastic reaction-diffusion equations; forward iequality

AMS Subject classification: 60J60

1 Introduction

This communication deals with the approach via *Lyapunov functions* to existence and uniqueness of stationary distributions of time-homogeneous diffusion processes in Hilbert spaces and to their trend to stationary distributions. We consider (mild) solutions X to stochastic initial-value problems of Itô's type

$$(1)\quad \begin{aligned} &dX(t) = (-AX(t) + f(X(t)))dt + \sigma(X(t))dW(t),\\ &X(0) = x \in \mathbf{E}. \end{aligned}$$

Here, the state space $\mathbf{E}$ of X is measurably imbedded in a Hilbert space $\mathbf{H}$ with inner product $(\cdot,\cdot)$ and $(W(t))_{t\geq 0}$ is a (possibly cylindrical) Wiener process on another Hilbert space $\mathbf{K}$ with a (continuous) covariance operator Q on $\mathbf{K}$ and defined on a probability space $(\Omega, \mathcal{A}, P)$. The linear drift part $A : \mathcal{D}(A) \to \mathbf{H}$ is a self-adjoint and uniformly positive, linear operator in $\mathbf{H}$. With $\mathcal{L}(\mathbf{K},\mathbf{H})$ denoting the space of all continuous, linear operators from $\mathbf{K}$ to $\mathbf{H}$, $f : \mathbf{E} \to \mathbf{H}$ and $\sigma : \mathbf{E} \to \mathcal{L}(\mathbf{K},\mathbf{H})$ are measurable mappings, nonlinear in general. Equations of the form (1) arise when stochastic parabolic partial differential equations

$$(2)\quad \frac{\partial u}{\partial t}(t,\xi) = \sum_{i,j=1}^{d} \frac{\partial}{\partial \xi_i}\left(a_{ij}(\xi)\frac{\partial u}{\partial \xi_j}(t,\xi)\right) + F(u(t,\xi)) + G(u(t,\xi))\eta(t,\xi)$$

are treated in a Hilbert-space context, cf. Section 6.1 where we also explain the notation. The solution X to Eq. (1) gives rise to a semi-group $(\mathcal{P}_t)_{t\geq 0}$ of transition kernels on $\mathbf{E}$ in the usual way. Often, this semigroup can be extended to a semigroup on a larger set $\mathbf{D} \supseteq \mathbf{E}$, e.g. via a generalized solution; cf. Section 3.2.2. We denote the extended semigroup again by $(\mathcal{P}_t)_t$.

Our chief results on stationary distributions are, on the one hand, conditions for the Feller property, existence, uniqueness, and trend in the abstract Hilbert-space setting couched in terms of the coefficients of the equation alone; see Proposition 5.3 and Corollary 5.5, where we use quadratic Lyapunov functions of the form $(y, A^{\alpha}y)$, $\alpha \in \mathbf{R}$. No prior information on moments of the diffusion is needed. In the linear case, these conditions are sharp in the sense that they reduce to Zabczyk's [28] well-known criterion if this Lyapunov function is used with exponent $\alpha = -1$, cf. Section 5.6. On the other hand, these conditions may be applied to stochastic reaction-diffusion equations with bounded, nonconstant diffusion coefficients and essentially decreasing drift functions of unlimited growth on a bounded, open subset $\mathcal{O} \subseteq \mathbf{R}^d$, cf. Theorem 6.5. We also gain information on the supports of the stationary distributions. In important cases, our conditions ensure that the stationary distribution is supported by the domain of definition of $A^{\frac{1+\alpha}{2}}$ for suitable $\alpha \in \mathbf{R}$, cf. Corollary 5.5. Unfortunately, our approach via Young's inequality does not allow the above Lyapunov function with $\alpha = -1$ in the presence of a *nonlinear* drift part f so that, in this case, we have to confine ourselves to a nuclear diffusion term.

The outlines of the paper are as follows. In order to verify the tightness properties necessary for having stationary distributions we introduce in Section 2 a completely regular Radon space $\mathbf{X}$ which contains $\mathbf{D}$ as a measurable, dense subset. Since the semi-group $(\mathcal{P}_t)_t$ is defined on $\mathbf{D}$, only, the usual notion of a stationary distribution γ on $\mathbf{X}$ does not make sense unless γ is supported by $\mathbf{D}$. We, therefore, first weaken the notion of a stationary distribution, cf. 2.4. Well-known properties assuring existence, uniqueness, and trend are a combination of the Feller property and tightness, attractivity in the mean, and attractivity, respectively. We provide sufficient conditions for a semigroup $(\mathcal{P}_t)_t$ of transition kernels on a dense subset $\mathbf{D}$ of a completely regular Radon space $\mathbf{X}$ to possess weakened versions of these properties, cf. Proposition 2.14. These sufficient conditions, stated in 2.11 and 2.13, are formulated in terms of Lyapunov functions for the underlying semigroup $(\mathcal{P}_t)_t$ and a family $(\bar{\mathcal{P}}_t)_t$ of kernels on the squared state space $\mathbf{D} \times \mathbf{D}$ with marginal distributions equal to $\mathcal{P}_t$, cf. 2.12. In the case of the Feller property and the attractivities, the Lyapunov functions are powers of a metric. Using the weakened versions we prove existence, uniqueness, and trend first in an abstract setting, Theorem 2.15.

In Section 3, we recall the concept of π–solution thus laying the ground for our main tool, the *Forward Inequality* 4.2. This inequality serves to apply Itô's calculus to irregular Lyapunov functions in the infinite dimensional situation. It is a generalization of Kolmogorov's forward equation and extends also the so-called *energy inequality* which is related to the Lyapunov function "norm square."

In Section 4, we reformulate the conditions of Section 2 in terms of "superharmonic functions" of the process X, cf. 4.4 and 4.5, and of superharmonic functions of the linked process 4.6. As in our earlier paper [17], we obtain again a unified view of Feller property and attractivity, cf. 4.7 and 4.8. Section 5 considers generally unbounded quadratic forms on $\mathbf{H}$ of the type $(x, A^\alpha x)$ for some real number α as Lyapunov functions. Here, we use the completion $\mathbf{X}$ of $\mathbf{D}$ equipped with the uniform structure induced by some continuous operator on $\mathbf{H}$.

We conclude our paper by applying the general results of Section 5 to stochastic partial differential equations. In the cases specified in Theorem 6.5, we obtain sufficient conditions for existence of and trend to stationary distributions in the usual sense. Our treatment of the drift term is based on Young's inequality, Lemma 6.2, cf. also [18], Lemma 6.3; it enables us to deal with measurable, *nonlinear* drift functions f of *unlimited* growth. To our knowledge, there are only few communications constructing stationary distributions for equations with highly nonlinear drift terms f not everywhere defined on $\mathbf{H}$ and nonconstant diffusions σ such as stochastic reaction-diffusion equations.

The use of Lyapunov functions for analyzing stochastic differential equations was initiated by Khasminskii [13] in finite dimensions; it was transferred to infinite dimensions, among others, by Ichikawa [12], Maslowski [20], and, more recently, by Leha and Ritter [17] and Chow and Khasminskii [3]. The last-mentioned authors couch their theory in a framework more general than ours, including the stochastic Navier-Stokes equation, but use the norm square as a Lyapunov function. For this reason, they need nuclear covariance of the noise for having stationary distributions also in the linear case.

Other recent papers on the subject matter are Da Prato and Pardoux [4], Gątarek and Gołdys [10], and Cerrai [2]. Da Prato and Pardoux deal with stochastic partial differential equations on an interval in the real line. They use a decomposition technique which enables them to deal with polynomially bounded drift coefficients and a cylindrical diffusion. Gątarek and Gołdys use certain conditions on the solution X and related processes in order to show existence and uniqueness of stationary distributions for fairly general continuous functions f and σ on $\mathbf{E}$. Their result applies in particular to the case of a stochastic parabolic partial differential equation (2) with $\mathbf{H} = \mathbf{L}^2([0,1])$, $\mathbf{E} = \mathcal{C}([0,1])$, and continuous functions F and G, F lying between two nonincreasing functions and $G > 0$ bounded away from zero. Cerrai treats systems of stochastic partial differential equations with polynomially bounded drift coefficients and constant correlation operators of the diffusion commuting with A; the diffusion may be cylindrical in the one-dimensional case.

2 Existence and uniqueness of stationary distributions of semigroups on completely regular Radon spaces

Let $\mathbf{X} = (\mathbf{X}, \mathcal{D})$ be a uniform space, $\mathcal{D}$ being a family of semi-metrics on $\mathbf{X}$. For details on uniform and completely regular spaces we refer the reader to Gillman and Jerison [11], Chapters 3 and 15. We denote the spaces of all continuous, all bounded continuous, and all bounded, uniformly continuous, real-valued functions on $\mathbf{X}$ by $\mathcal{C}(\mathbf{X})$, $\mathcal{C}_b(\mathbf{X})$, and $\mathcal{C}_{ub}(\mathbf{X})$, respectively. The norm $\|g\|_\infty$ of a bounded, real-valued function g on $\mathbf{X}$ is the least upper bound $\sup_{u\in\mathbf{X}} \mid g(u) \mid$. For any $d \in \mathcal{D}$, the space of bounded, uniformly d–continuous functions on $\mathbf{X}$ is denoted by $\mathcal{C}^d_{ub}(\mathbf{X})$. Let $\mathbf{D}$ be a Borel measurable, dense subset of $\mathbf{X}$ and let $(\mathcal{P}_t)_{t\geq 0}$ be a semigroup of transition kernels on $\mathbf{D}$.

2.1 Prohorov's theorem on completely regular Radon spaces We want to resort to Prohorov's theorem on $\mathbf{X}$ for proving existence of stationary distributions of $(\mathcal{P}_t)_t$; therefore, we consider "stationary" distributions not only on $\mathbf{D}$ but also on the enlarged space $\mathbf{X}$.

(a) The completely regular Hausdorff space $\mathbf{X}$ is endowed with its Borel σ–field $\mathcal{B}$. A finite *Radon measure* μ on $\mathbf{X}$ is a finite Borel measure which is inner regular with respect to the compact sets of $\mathbf{X}$, i.e., satisfies $\mu(A) = \sup\{\mu(K)/K \subseteq A, K \text{ compact}\} < \infty$ for all $A \in \mathcal{B}$ (cf. Schwartz [23], p. 13). We denote the set of all Radon probability measures on $\mathbf{X}$ by $\mathrm{Prob}(\mathbf{X})$. The space $\mathbf{X}$ is called a **Radon space** (cf. Schwartz, loc. cit.) if every finite Borel measure μ on $\mathbf{X}$ is a Radon measure. Details on Radon spaces can be found in Gardner [8], Gardner and Pfeffer [9], and Dellacherie and Meyer [7], Chapter III. Since $\mathbf{X}$ is completely regular, the topology of *narrow convergence* on the vector space of all finite, signed Radon measures on $\mathbf{X}$ is the coarsest topology for which the mappings $\mu \to \mu(g), g \in \mathcal{C}_b(\mathbf{X})$, are continuous (cf. Schwartz, loc. cit., pp. 249 and 371, Dellacherie and Meyer, loc. cit., p. 71).

(b) A family $M \subseteq \mathrm{Prob}(\mathbf{X})$ is **(uniformly) tight** if, for every $\varepsilon > 0$, there exists a compact set $K \subseteq \mathbf{X}$ such that $\mu(\mathbf{X} \setminus K) < \varepsilon$ for all $\mu \in M$. We will use the generalization of Prohorov's classical compactness criterion to completely regular spaces based on ideas of Topsøe [24], bottom of p. 98. It states that any uniformly tight family of Radon probability measures on a completely regular space $\mathbf{X}$ is narrowly relatively compact (cf. Topsøe [25], Cor. 1 (i), p. 203, Topsøe [26], Theorem 9.1 (ii), Schwartz [23], p. 379, and Dellacherie and Meyer [7], p. 72, Theorem 59).

(c) If K is a transition kernel, μ a probability measure, and g a function then the probability measure μK and the function Kg are defined in the usual way.

2.2 Feller property Let $\mathcal{F}$ be a system of bounded, continuous, real-valued functions on $\mathbf{X}$.

(a) The system $\mathcal{F}$ **separates** $\mathrm{Prob}(\mathbf{X})$ if, for any two different measures $\mu, \nu \in \mathrm{Prob}(\mathbf{X})$, there exists a function $f \in \mathcal{F}$ such that $\int f d\mu \neq \int f d\nu$.

(b) A semigroup $(\mathcal{P}_t)_{t\geq 0}$ of kernels on $\mathbf{D}$ operating on the space $\mathcal{C}_b(\mathbf{D})$ is usually said to possess the Feller property. The following is a modification of this notion. We say that the semigroup $(\mathcal{P}_t)_{t\geq 0}$ possesses the **$\mathcal{F}$-Feller property** if $\mathcal{F}$ is a separating subspace of $\mathcal{C}_b(\mathbf{X})$ such that $\mathcal{P}_t\mathcal{F} \subseteq \mathcal{F}_{|\mathbf{D}}$ for all $t > 0$. Then, for all $g \in \mathcal{F}$, there exists a uniquely defined continuous extension of $\mathcal{P}_t g$ to all of $\mathbf{X}$ which we denote by $\mathcal{T}_t g$ $(\in \mathcal{F})$.

If $(\mathcal{P}_t)_{t\geq 0}$ possesses the $\mathcal{F}$-Feller property then $(\mathcal{T}_t)_t$ is a semigroup of continuous, linear operators on $\mathcal{F}$. Indeed, $\mathcal{T}_t$ defines a continuous, linear operator $\mathcal{F} \to \mathcal{F}$, $t \geq 0$. Moreover, for $g \in \mathcal{F}$, $s, t \geq 0$, and $x \in \mathbf{D}$, we have

$$\begin{aligned}\mathcal{T}_t(\mathcal{T}_s g)(x) &= P_t(\mathcal{T}_s g)(x) = \int_{\mathbf{D}} P_t(x, dy)\mathcal{T}_s g(y) = \int_{\mathbf{D}} \mathcal{P}_t(x, dy)\mathcal{P}_s g(y) \\ &= \mathcal{P}_{s+t} g(x) = \mathcal{T}_{s+t} g(x).\end{aligned}$$

Since both functions $\mathcal{T}_t\mathcal{T}_s g$ and $\mathcal{T}_{s+t} g$ are continuous we have equality for $x \in \mathbf{X}$ and the claim follows.

2.3 Example The following example shows that the operator $\mathcal{T}_t$ on $\mathcal{F}$ need not be representable by integration with respect to a kernel on $\mathbf{X}$, not even in the deterministic linear case. Let

$$\mathbf{X} = c_0 = \{(x_k) \in \mathbf{R}^{\mathbf{N}} / \lim_k x_k = 0\}$$

be endowed with the product uniform structure, let

$$\mathbf{D} = \mathbf{R}^{(\mathbf{N})} = \{(x_k) \in \mathbf{R}^{\mathbf{N}} / x_k = 0 \text{ for all but finitely many indices } k\},$$

and let $A : \mathbf{D} \to \ell^2(\mathbf{N})$ be defined by $Ay = (y_k \ln k)_k$. Note that $\mathbf{D}$ is dense both in $\mathbf{X}$ and in $\mathbf{R}^{\mathbf{N}}$. Define $(S_t x)_k = x_k k^t$, $k \in \mathbf{N}$, $t \geq 0$, $x \in \mathbf{R}^{\mathbf{N}}$. Then $X(t) = S_t x$ solves the equation $dX(t) = AX_t dt$ with initial value $X(0) = x \in \mathbf{D}$. The transition kernel of X is $\mathcal{P}_t(x, \cdot) = \delta_{S_t x}$. As a measurable subset of the Polish space $\mathbf{R}^{\mathbf{N}}$, $\mathbf{X}$ is a Radon space; moreover, $(\mathcal{P}_t)_t$ is a semigroup of transition kernels on $\mathbf{D}$ possessing the $\mathcal{F}$–Feller property with respect to $\mathcal{F} = \mathcal{C}_{ub}(\mathbf{X})$. Indeed, the mapping $x \to S_t x$ is uniformly continuous on $\mathbf{D}$ and $\mathcal{P}_t g = g \circ S_t$. Now choose $a \in c_0$ such that $S_t a \notin c_0$. Since $\mathcal{T}_t g(a) = \tilde{g}(S_t a)$, where $\tilde{g}$ is the unique continuous extension of g to $\mathbf{R}^{\mathbf{N}}$, and since $\mathcal{F}$ is separating, the linear form $g \to \mathcal{T}_t g(a)$, $g \in \mathcal{F}$, cannot be represented by integration on $\mathbf{X}$.

2.4 Stationary distributions A probability distribution γ on $\mathbf{D}$ is called **stationary** with respect to $(\mathcal{P}_t)_t$ if $\int \mathcal{P}_t g d\gamma = \int g d\gamma$ for all bounded, measurable functions g on $\mathbf{D}$ and all t. The following is a generalization of this notion.

Let $(\mathcal{P}_t)_t$ possess the $\mathcal{F}$-Feller property and let $(\mathcal{T}_t)_t$ be the associated semigroup of operators on $\mathcal{F}$ as defined in 2.2(b). We call $\gamma \in \text{Prob}(\mathbf{X})$ **$\mathcal{F}$-stationary** with respect to the semigroup $(\mathcal{P}_t)_t$ if, for all $t > 0$ and all $g \in \mathcal{F}$,

$$\int \mathcal{T}_t g \, d\gamma = \int g \, d\gamma.$$

$\mathcal{F}$–stationarity of a probability measure on $\mathbf{D}$ amounts to its stationarity with respect to $(\mathcal{P}_t)_t$. $\mathcal{F}$–stationarity of a distribution γ on $\mathbf{X}$ says that, for all $t > 0$, the continuous, linear form $g \to \gamma(\mathcal{T}_t g)$, $g \in \mathcal{F}$, is represented by integration with respect to γ.

2.5 Notation Let $\mathcal{M}_t$ be the kernel on $\mathbf{D}$ defined by

$$\mathcal{M}_t(x, B) := \frac{1}{t} \int_0^t \mathcal{P}_s(x, B) ds, \qquad x \in \mathbf{D}, B \in \mathcal{B}(\mathbf{D}).$$

If $(\mathcal{P}_t)_t$ possesses the $\mathcal{F}$-Feller property then we also define the operators $\mathcal{N}_t : \mathcal{F} \to \mathcal{C}_b(\mathbf{X})$:

$$\mathcal{N}_t g(u) := \frac{1}{t} \int_0^t \mathcal{T}_s g(u) ds, \qquad u \in \mathbf{X}, g \in \mathcal{F}.$$

2.6 Existence principle Let $(\mathcal{P}_t)_t$ possess the $\mathcal{F}$-Feller property. If there exists $\mu \in \text{Prob}(\mathbf{D})$ such that the family $(\mu \mathcal{M}_t)_{t \geq 1}$ is narrowly relatively sequentially compact on $\mathbf{X}$ then $(\mathcal{P}_t)_t$ has an $\mathcal{F}$–stationary distribution; in fact, each narrow cluster point γ of $(\mu \mathcal{M}_t)_t$ as $t \to \infty$ is $\mathcal{F}$–stationary.

Proof. Let $\gamma \in \text{Prob}(\mathbf{X})$ be a cluster point of $(\mu \mathcal{M}_t)_{t \geq 1}$ as $t \to \infty$, $\gamma = \lim_{n \to \infty} \mu \mathcal{M}_{t_n}$ for some sequence $(t_n)_n \to \infty$ of real numbers ≥ 1. Let $t > 0$ and $g \in \mathcal{F}$. A standard computation shows

$$\lim_n \int_{\mathbf{D}} (\mu \mathcal{M}_{t_n})(dx) \mathcal{P}_t g(x) = \lim_n \int_{\mathbf{D}} (\mu \mathcal{M}_{t_n})(dx) g(x).$$

Since $g \in \mathcal{C}_b(\mathbf{X})$, the right side equals $\int_{\mathbf{X}} \gamma(du) g(u)$ and, since $\mathcal{T}_t g \in \mathcal{C}_b(\mathbf{X})$, the left side equals $\int_{\mathbf{X}} \gamma(du) \mathcal{T}_t g(u)$. This is $\mathcal{F}$–stationarity of γ. $\square$

2.7 Attractivity and trend (a) The semigroup $(\mathcal{P}_t)_t$ is said to exhibit **trend** to the distribution γ on $\mathbf{X}$ if

(3) $\quad \mu\mathcal{P}_t \to \gamma \quad$ narrowly as $t \to \infty$

for all probability distributions μ on $\mathbf{D}$. Note that (3) holds if and only if $\mathcal{P}_t(x, \cdot) \to \gamma$ narrowly as $t \to \infty$ for all $x \in \mathbf{D}$.

A related notion is attractivity which we now recall.

(b) A subset $\mathcal{A} \subseteq \mathcal{C}_b(\mathbf{X})$ **determines convergence** if $\mathcal{A}$ generates the narrow topology on $\mathrm{Prob}(\mathbf{X})$, i.e., if the narrow topology is the coarsest topology on $\mathrm{Prob}(\mathbf{X})$ for which all mappings $\mu \to \mu(g)$, $g \in \mathcal{A}$, are continuous. Part of the *portmanteau theorem* in the context of a completely regular Radon space $\mathbf{X}$ shows: the space $\mathcal{C}_{ub}(\mathbf{X})$ of all bounded, uniformly continuous, real-valued functions is convergence determining. This follows, e.g., from Topsøe [26], Theorem 8.1, and the fact that finite Radon measures are τ–smooth. In the case of a *metric* space $\mathbf{X}$ this theorem is also true; cf. Billingsley [1], p. 12. Obviously, any convergence-determining system is separating.

(c) Let $\mathcal{A}$ be a subspace of $\mathcal{C}_b(\mathbf{X})$. We call the semigroup $(\mathcal{P}_t)_t$ **$\mathcal{A}$-attractive** if

(i) $\mathcal{A}$ determines convergence;

(ii) $\mathcal{P}_t g(y) - \mathcal{P}_t g(x) \to 0$ as $t \to \infty$ for all $g \in \mathcal{A}$ and all $x, y \in \mathbf{D}$;

(iii) the family $(\mathcal{P}_t g)_{t \geq 1}$ is uniformly equicontinuous on $\mathbf{D}$ for all $g \in \mathcal{A}$.

2.8 Trend principle Let $(\mathcal{P}_t)_t$ possess the $\mathcal{A}$-Feller property. If $(\mathcal{P}_t)_t$ is $\mathcal{A}$-attractive and possesses an $\mathcal{A}$–stationary distribution γ then we have trend to γ.

Proof. Let $\mu \in \mathrm{Prob}(\mathbf{D})$ and let $g \in \mathcal{A}$. By 2.7(ii),(iii) and denseness of $\mathbf{D} \subseteq \mathbf{X}$, we have $\mathcal{T}_t g(v) - \mathcal{T}_t g(u) \to 0$ as $t \to \infty$ for all $g \in \mathcal{A}$ and all $u, v \in \mathbf{X}$. Hence,

$$\lim_{t\to\infty} \int g\, d(\mu\mathcal{P}_t) - \int g d\gamma = \lim_{t\to\infty} \left\{ \int \mathcal{T}_t g\, d\mu - \int \mathcal{T}_t g\, d\gamma \right\}$$
$$= \int_{\mathbf{X}\times\mathbf{X}} \gamma \otimes \mu(du, dv) \lim_{t\to\infty} (\mathcal{T}_t g(v) - \mathcal{T}_t g(u)) = 0$$

by Lebesgue's theorem on bounded convergence. Since $\mathcal{A}$ is convergence determining the claim follows. □

2.9 Attractivity in the mean (a) Let $\mathcal{AM}$ be a subspace of $\mathcal{C}_b(\mathbf{X})$. We call the semigroup $(\mathcal{P}_t)_t$ **$\mathcal{AM}$-attractive in the mean** if

(i) $\mathcal{AM}$ separates $\mathrm{Prob}(\mathbf{X})$;

(ii) $\mathcal{M}_t g(y) - \mathcal{M}_t g(x) \to 0$ as $t \to \infty$ for all $g \in \mathcal{AM}$ and all $x, y \in \mathbf{D}$;

(iii) the family $(\mathcal{M}_t g)_{t \geq 1}$ is uniformly equicontinuous on $\mathbf{D}$ for all $g \in \mathcal{AM}$.

Note that $\mathcal{A}$-attractivity implies $\mathcal{A}$-attractivity in the mean. The latter is sufficient for uniqueness:

2.10 Uniqueness Principle Let $(\mathcal{P}_t)_t$ possess the $\mathcal{AM}$-Feller property. If $(\mathcal{P}_t)_t$ is $\mathcal{AM}$-attractive in the mean then there exists at most one $\mathcal{AM}$–stationary distribution.

Proof. By 2.9(ii),(iii) and denseness of $\mathbf{D}$ in $\mathbf{X}$, we have $\mathcal{N}_t g(v) - \mathcal{N}_t g(u) \to 0$ as $t \to \infty$ for all $g \in \mathcal{AM}$ and all $u, v \in \mathbf{X}$. Let μ and γ be two $\mathcal{AM}$–stationary distributions. The proof is now carried out with the aid of separation of $\mathcal{AM}$ as in 2.8 with $\mathcal{N}_t$ instead of $\mathcal{T}_t$. □

Our general objective is to give Lyapunov-type conditions for existence and uniqueness of stationary distributions and for trend to equilibrium. In view of the three principles formulated above it is sufficient to concentrate on relative compactness of $(\mu\mathcal{M}_t)_t$, the Feller property, attractivity, and attractivity in the mean.

2.11 The Condition (T) The following is a general condition of Lyapunov's type ensuring tightness of the family $(\mu\mathcal{M}_t)_t$ of distributions.

(T) There exists a probability measure μ on $\mathbf{D}$ and a lower bounded, measurable function $\eta : \mathbf{D} \to \mathbf{R}$ such that

(i) $\{\eta \leq r\} \subseteq \mathbf{D}$ is relatively compact in $\mathbf{X}$ for all real numbers $r > 0$;

(ii) $\int_{\mathbf{D}} \eta(y)(\mu\mathcal{M}_t)(dy) \leq c$ for all $t \geq 1$ and some $c \in \mathbf{R}$.

In many cases, it is possible to choose Dirac measures for the measure μ; in Section 4.4, we will indicate how to find functions η with the desired properties in the case of diffusion processes on a Hilbert space $\mathbf{H}$.

Interestingly, Feller property, attractivity, and attractivity in the mean allow a unified approach by means of Lyapunov functions on the product space $\mathbf{D} \times \mathbf{D}$ equipped with the product structures, i.e., the product topology and the corresponding Borel structure. We need a few preliminaries.

2.12 Notation In the sequel, $\bar{\mathcal{P}}_t$ will denote any kernel on $\mathbf{D} \times \mathbf{D}$ such that

(i) $\bar{\mathcal{P}}_t((x, y), B \times \mathbf{D}) = \mathcal{P}_t(x, B)$ and $\bar{\mathcal{P}}_t((x, y), \mathbf{D} \times B) = \mathcal{P}_t(y, B)$ for all measurable subsets $B \subseteq \mathbf{D}$ and all $x, y \in \mathbf{D}$.

Examples are the tensor product $\bar{\mathcal{P}}_t((x, y), \cdot) = \mathcal{P}_t(x, \cdot) \otimes \mathcal{P}_t(y, \cdot)$ and the linked semigroup which we will use extensively in the Hilbert-space context, cf. 4.6. We also write $\bar{\mathcal{M}}_t((x, y), \cdot) = \frac{1}{t}\int_0^t \bar{\mathcal{P}}_s((x, y), \cdot)ds$ for $t > 0$ and $x, y \in \mathbf{D}$.

2.13 The Conditions (F), (A), and (AM)

(F) There exists a metric $d \in \mathcal{D}$ and some $p > 0$ such that

(i) $\bar{\mathcal{P}}_t d^p(x, y) \leq c(t) d^p(x, y)$ for some function $c : \mathbf{R}_+ \to \mathbf{R}_+$, all $t \geq 0$, and all $x, y \in \mathbf{D}$,

(ii) $\mathcal{C}^d_{ub}(\mathbf{X})$ separates $\mathrm{Prob}(\mathbf{X})$.

(A) There exists a metric $d \in \mathcal{D}$ and some $p > 0$ such that

(i) the Condition (F)(i) is satisfied with some bounded function $c : \mathbf{R}_+ \to \mathbf{R}_+$ vanishing at infinity,

(ii) $\mathcal{C}^d_{ub}(\mathbf{X})$ is convergence determining.

(AM) There exists a metric $d \in \mathcal{D}$ and some $p > 0$ such that

(i) the Condition (F)(i) is satisfied with $\bar{\mathcal{M}}_t$ instead of $\bar{\mathcal{P}}_t$ and some bounded function $c : \mathbf{R}_+ \to \mathbf{R}_+$ vanishing at infinity,

(ii) $\mathcal{C}^d_{ub}(\mathbf{X})$ separates $\mathrm{Prob}(\mathbf{X})$.

If the metric d appearing in these conditions generates the $\mathcal{D}$-Borel structure of $\mathbf{X}$ then the Parts (ii) of the foregoing conditions are satisfied, cf. 2.7(b).

We recall that the **lower semi-continuous regularization** $\hat{\eta} : \mathbf{X} \to \mathbf{R}$ of a lower bounded function $\eta : \mathbf{D} \to \mathbf{R}$ is the largest lower semi-continuous function on $\mathbf{X}$ dominated by η on $\mathbf{D}$.

2.14 Proposition Let $\mathbf{X}$ be a completely regular Radon space.

(a) If Condition (T) (cf. 2.11) is satisfied with $\mu \in \text{Prob}(\mathbf{D})$ then the family $(\mu\mathcal{M}_t)_{t\geq 1}$ is sequentially narrowly relatively compact. Let γ be a narrow cluster point of $(\mu\mathcal{M}_t)_{t\geq 1}$ as $t \to \infty$. We have $\gamma(\{\hat{\eta} < \infty\}) = 1$, where $\hat{\eta}$ is the lower semi-continuous regularization of the function η appearing in Condition (T).

(b) Condition (F) implies the $\mathcal{C}^d_{ub}(\mathbf{X})$-Feller property of $(\mathcal{P}_t)_t$ with the metric d appearing in (F).

(c) Condition (A) implies Conditions (F), (AM), and the $\mathcal{C}^d_{ub}(\mathbf{X})$-attractivity of $(\mathcal{P}_t)_t$ with the metric d appearing in (A).

(d) If Condition (AM) is satisfied and if $(\mathcal{P}_t)_t$ has the $\mathcal{C}^d_{ub}(\mathbf{X})$-Feller property with the metric d appearing in (AM) then the semigroup $(\mathcal{P}_t)_t$ is $\mathcal{C}^d_{ub}(\mathbf{X})$-attractive in the mean.

Proof. (a) Choose a lower bound c_T of η and let $r > 0$. Putting $\eta_0 = \eta - c_T$ and $K_r := \{\eta_0 \leq r\}$ it suffices to use Markov's inequality and to estimate for $t \geq 1$

$$\mu\mathcal{M}_t(\mathbf{X} \setminus \bar{K}_r) \leq \mu\mathcal{M}_t(\mathbf{X} \setminus K_r) \leq \frac{1}{r}\int_{\mathbf{X}} \eta_0(u)(\mu\mathcal{M}_t)(du) \leq \frac{c - c_T}{r}.$$

The first claim follows from Prohorov's theorem, cf. 2.1(b). For the second claim note that, by the portmanteau theorem, $\int \hat{\eta} d\gamma \leq \liminf \int \hat{\eta} d(\mu\mathcal{M}_t) \leq \liminf \int \eta d(\mu\mathcal{M}_t) \leq c$.

(b) Let d and p be as in Condition (F) and let $\mathcal{F} = \mathcal{C}^d_{ub}(\mathbf{X})$. Since $\mathcal{F}$ is separating by (F)(ii), it remains to show that $\mathcal{P}_t\mathcal{F} \subseteq \mathcal{F}_{|\mathbf{D}}$. To this end, let $g \in \mathcal{F}$, let $\varepsilon > 0$, and choose $\delta > 0$ such that $| g(v) - g(u) | < \varepsilon$ whenever $d^p(u,v) < \delta$. Using 2.12(i) and F(i), we may estimate for $x, y \in \mathbf{D}$

$$\begin{aligned} | \mathcal{P}_t g(y) - \mathcal{P}_t g(x) | &\leq \int_{\mathbf{X}\times\mathbf{X}} \bar{\mathcal{P}}_t((x,y), d(u,v)) \mid g(v) - g(u) \mid = \int_{\{d^p<\delta\}} + \int_{\{d^p\geq\delta\}} \\ &\leq \varepsilon + \frac{2\|g\|_\infty}{\delta}\bar{\mathcal{P}}_t d^p(x,y) \leq \varepsilon + \frac{2\|g\|_\infty}{\delta} c(t) d^p(x,y). \end{aligned} \tag{4}$$

Hence, $\mathcal{P}_t g$ is uniformly d-continuous on $\mathbf{D}$ and may, therefore, be extended to a uniformly d-continuous function on $\mathbf{X}$, i.e., $\mathcal{P}_t g \in \mathcal{F}_{|\mathbf{D}}$.

(c) Hypothesis A(i) together with (4) implies 2.7(ii), (iii) and 2.7(i) is just A(ii).

The proof of Part (d) is analogous to that of (c) after we replace $\mathcal{P}_t$ ($\bar{\mathcal{P}}_t$) with $\mathcal{M}_t$ ($\bar{\mathcal{M}}_t$). □

Summarizing Proposition 2.14 and the Principles 2.6, 2.8, and 2.10, we obtain the following main result of this section.

2.15 Theorem Let $\mathbf{X}$ be a completely regular Radon space.

(a) Assume the Condition (T) (cf. 2.11) and let (F) (cf. 2.13) be satisfied with some metric $d \in \mathcal{D}$. Then $(\mathcal{P}_t)_t$ has a $\mathcal{C}^d_{ub}(\mathbf{X})$–stationary distribution γ satisfying $\gamma(\{\hat{\eta} < \infty\}) = 1$, where $\hat{\eta}$ is the lower semi-continuous regularization on $\mathbf{X}$ of the function η appearing in Condition (T). In particular, if $\hat{\eta} = \infty$ on $\mathbf{X} \setminus \mathbf{D}$ then γ is a stationary distribution with respect to $(\mathcal{P}_t)_t$.

(b) Suppose that Condition (A) (cf. 2.13) is satisfied with some metric $d \in \mathcal{D}$ and that $(\mathcal{P}_t)_t$ possesses a $\mathcal{C}^d_{ub}(\mathbf{X})$–stationary distribution γ. Then we have trend to γ; i.e., we have $\mu\mathcal{P}_t \to \gamma$ narrowly as $t \to \infty$ for all probability distributions μ on $\mathbf{D}$. In particular, there exists exactly one $\mathcal{C}^d_{ub}(\mathbf{X})$–stationary distribution.

(c) If Condition (AM) (cf. 2.13) is satisfied with some metric $d \in \mathcal{D}$ and if $(\mathcal{P}_t)_t$ has the $\mathcal{C}^d_{ub}(\mathbf{X})$-Feller property then there exists at most one $\mathcal{C}^d_{ub}(\mathbf{X})$–stationary distribution.

Condition (T) is also necessary for the existence of a stationary distribution carried by $\mathbf{D}$: Given such a stationary distribution γ one can construct, due to inner regularity of γ, a function η as required in Condition (T) by choosing a suitable sequence of compact subsets of the Radon space $\mathbf{X}$.

3 Itô diffusions on Hilbert spaces

3.1 Notation From now on, $\mathbf{H}$ denotes a real Hilbert space and $\mathbf{E}$ a measurable subset of $\mathbf{H}$. Often $\mathbf{E}$ carries the structure of a separable Banach space with a norm finer than that of the Hilbert space. Then measurability of $\mathbf{E}$ is guaranteed by a theorem of Lusin's, cf. Schwartz [23], p. 101, Theorem 5. We denote the inner product and the norm of $\mathbf{H}$ by $(\cdot,\cdot)$ and $\|\cdot\|$, respectively, and the domain of any linear operator B in $\mathbf{H}$ by $\mathcal{D}(B)$. The symbol tr stands for the trace of a linear operator and λ denotes Lebesgue measure on $\mathbf{R}_+$. We consider a formal stochastic differential equation of the form (1).

3.2 Concepts of solutions We first recall some concepts of solutions to the differential equation (1). We suppose that a solution $X = (X(t))_{0\leq t<\infty}$ with initial point $x \in \mathbf{E}$ enjoys the following basic properties:

(i) X is adapted to the filtration $(\mathcal{A}_t)_{t\geq 0}$ on $(\Omega, \mathcal{A}, P)$ induced by the Wiener process W;

(ii) $X(t,\omega) \in \mathbf{E}$ for $\lambda \otimes P$-a.a. $(t,\omega) \in \mathbf{R}_+ \times \Omega$.

Concerning the treatment of the singular drift coefficients A and f and the diffusion coefficient σ there exist several approaches. Among them are mild solutions, cf. Da Prato and Zabczyk [5], pp. 152, 182, generalized solutions, cf. Da Prato and Zabczyk [6], p.81, and π-solutions [18] which we next recall. We abbreviate

$$\mathcal{A}(y) = \sigma(y)Q\sigma^*(y), \qquad y \in \mathbf{E}.$$

3.2.1 Mild solutions Let $(S(t))_{t\geq 0}$ be the C_0-semigroup generated by $-A$. A process X is called a *mild* solution of (1) if it satisfies (i), (ii) and if the following hold:

(iii) P-a.s., we have

$$\int_0^t \|S(t-s)f(X(s))\|ds < \infty, \quad t \geq 0;$$

(iv) $\sigma(X)$ is predictable and, P-a.s.,

$$\int_0^t tr S(t-s)\mathcal{A}(X(s))S(t-s)ds < \infty, \quad t \geq 0;$$

(v) the *variation-of-constants* formula

$$X(t) = S(t)x + \int_0^t S(t-s)f(X(s))ds + \int_0^t S(t-s)\sigma(X(s))dW(s)$$

is satisfied for all $t \geq 0$.

3.2.2 Generalized solutions A *generalized* solution to Eq. (1) is an $\mathbf{H}$–valued process X such that there exists a sequence $(X_n)_n$ of mild solutions to Eq. (1) converging in norm to X locally uniformly on $[0,\infty[$. Of course, any mild solution is a generalized solution.

3.2.3 π-solutions In the Forward Inequality 4.2 we will need yet another concept of solution, the π-solution introduced in [18]. This concept is useful for dealing with Itô's calculus in the presence of an unbounded operator A. Let $(\pi_k)_{k\in\mathbf{N}}$ be a sequence of symmetric Hilbert-Schmidt operators on $\mathbf{H}$ such that

- the sequence $(\pi_k)_k$ converges to the identity operator I in the strong operator topology,

- the operators $A\pi_k$ are bounded and self-adjoint and extend the operators $\pi_k A$.

It follows that the sequence of vectors $A\pi_k y$ converges to Ay for $y \in \mathcal{D}(A)$ as $k \to \infty$. Examples for π_k are the operators $kR_k(-A)$, where $R_k(-A) = (kI + A)^{-1}$ is the k-resolvent operator of $-A$, and the operators $1_{[0,k]}(A)$. We abbreviate for $y \in \mathbf{E}$

(5) $\quad b_k(y) = -A\pi_k y + \pi_k f(y),$

(6) $\quad \mathcal{A}_k(y) = \pi_k \sigma(y) Q \sigma^*(y) \pi_k.$

The operator $\mathcal{A}_k(y)$ is nuclear. A process $X = (X(t))_{0 \le t < \infty}$ is a π-*solution* to the stochastic differential equation (1) if it satisfies (i), (ii) and if the following hold:

(vi) P-a.s., we have for all $k \in \mathbf{N}$

$$\int_0^t \|\pi_k f(X(s))\| ds < \infty, \quad t \ge 0;$$

(vii) the $\mathcal{L}(\mathbf{K}, \mathbf{H})$-valued mappings $t \to \pi_k \sigma(X(t))$, $k \in \mathbf{N}$, are locally Itô integrable on $[0, \infty[$ with respect to W; i.e., $\pi_k \sigma(X)$ is predictable and, P-a.s.,

$$\int_0^t tr \mathcal{A}_k(X(s)) ds < \infty, \quad t \ge 0;$$

(viii) for all $k \in \mathbf{N}$ and $t \ge 0$, the equality

$$\pi_k X(t) = \pi_k x + \int_0^t \{-A\pi_k X(s) + \pi_k f(X(s))\} ds + \int_0^t \pi_k \sigma(X(s)) dW(s)$$

is satisfied.

In Leha et al. [18], Proposition 3.6, the following connection between mild and π-solutions is proved.

3.3 Proposition Let X be a P-a.s. locally Bochner integrable mild solution to Eq. (1) satisfying P-a.s.

(7) $\quad \int_0^t \|f(X(s))\| ds < \infty,$

(8) $\quad \int_0^t tr \mathcal{A}(X(s)) ds < \infty$

for all $t \ge 0$. Then X is also a π-solution to (1).

There are several results on existence and uniqueness of mild solutions to stochastic differential equations of the form (1), cf. Da Prato and Zabczyk [5], Theorem 7.10, Da Prato and Zabczyk [6], and Peszat [21], and the literature cited there. These are mainly based on dissipativity of the drift term and a Banach-space topology on $\mathbf{E}$.

4 Lyapunov conditions for Itô diffusions

Let $\mathbf{E} \subseteq \mathbf{H}$, A, f, σ, and π_k be as defined in the introduction and in Section 3.2.3. For each $x \in \mathbf{E}$, let $X = X^x = (X^x(t))_{t \ge 0}$ be an $\mathbf{H}$-continuous π-solution 3.2.3 to (1) starting from x with infinite lifetime, e.g. a mild solution 3.2.1 satisfying the assumptions of Proposition 3.3.

4.1 Notation (a) We denote the space of all twice continuously (Fréchet-) differentiable real-valued functions on $\mathbf{H}$ by $\mathcal{C}^2(\mathbf{H})$. For any $U \in \mathcal{C}^2(\mathbf{H})$, U' and U'' stand for the first and second Fréchet derivatives of U. If $U \in \mathcal{C}^2(\mathbf{H})$, $y \in \mathbf{H}$, and if $B \in \mathcal{L}(\mathbf{H})$ then $\operatorname{tr} U''(y)(B\cdot,\cdot)$ is the trace of the bilinear form $(u,v) \to U''(y)(Bu,v)$, $u, v \in \mathbf{H}$, if it exists. $\mathcal{C}^2_u(\mathbf{H})$ denotes the space of functions $U \in \mathcal{C}^2(\mathbf{H})$ such that U'' is uniformly continuous on bounded sets.

(b) For $U \in C^2(\mathbf{H})$, $k \in \mathbf{N}$, and $y \in \mathbf{E}$, we put

(9) $$\mathcal{L}^{(k)}U(y) = U'(\pi_k y)b_k(y) + \frac{1}{2} tr U''(\pi_k y)(\mathcal{A}_k(y)\cdot,\cdot),$$

where b_k and $\mathcal{A}_k$ are defined in (5) and (6), respectively. In order to control the growth of the diffusion process X in terms of a Lyapunov function U we have repeatedly used some kind of *forward inequality*; cf. Leha and Ritter [15], [16], [17], Leha, Ritter and Wakolbinger [18], Lemma 4.2, and Leha, Maslowski, and Ritter [14], Section 2. It is applicable to nonquadratic and even irregular Lyapunov functions. The following version of the forward inequality is adapted from [14], Section 2, for our present needs.

4.2 The Forward Inequality Let $V_n \in \mathcal{C}^2_u(\mathbf{H})$, $V_n \geq 0$, be a sequence of (Lyapunov) functions such that $V := \underline{\lim}_n V_n$ is lower semi-continuous (with respect to the norm topology on $\mathbf{H}$) and let $x \in \mathbf{E}$. Suppose that

(i) $V(x) < \infty$;

(ii) for all n, there exists a P-a.s. locally-integrable random function $\varphi_n : \mathbf{R}_+ \to [0,\infty]$ such that, $\lambda \otimes P$-a.s.,
$$\sup_k V_n'(\pi_k X)(-A\pi_k X) \leq \varphi_n;$$

(iii) there exists a constant $C \geq 0$ such that, for all n and all $y \in \mathbf{E}$,
$$V_n(y) \leq C(1 + V(y)) \quad \text{and} \quad \overline{\lim}_k \mathcal{L}^{(k)} V_n(y) \leq C(1 + V_n(y)).$$

Then, we have
$$EV(X(t)) \leq V(x) + E\int_0^t \overline{\lim}_n \overline{\lim}_k \mathcal{L}^{(k)} V_n(X(s))ds, \quad t \geq 0.$$

Proof. Invoking [14], Theorem 2.7, it is sufficient to provide a P-a.s. locally-integrable random function $\tilde{\varphi}_n : \mathbf{R}_+ \to [0,\infty]$ such that, $\lambda \otimes P$-a.s., we have

(10) $$\sup_k \mathcal{L}^{(k)} V_n(X) \leq \tilde{\varphi}_n.$$

Indeed, from (9), we infer
$$\begin{aligned}\mathcal{L}^{(k)}V_n(X) \quad \leq \quad & V_n'(\pi_k X)(-A\pi_k X) + \|V_n'(\pi_k X)\|_{\mathbf{H}'}\|\pi_k\|\|f(X)\| \\ & + \frac{1}{2}\|V_n''(\pi_k X)\|^2\|\pi_k\|^2 tr\, \sigma(X)Q\sigma^*(X)\end{aligned}$$

and an estimate of the form (10) follows from (ii) and the properties of V_n and those of X, π_k, f, σ, and Q stipulated in 3.2.3. □

4.3 Explanation We now specify the semigroup $(\mathcal{P}_t)_{t\geq 0}$ appearing in Section 2 as the semigroup associated with a mild solution to Eq. (1) on $\mathbf{D} = \mathbf{E}$ or the semigroup associated with a generalized solution 3.2.2 on a measurable subset $\mathbf{D} \subseteq \mathbf{H}$ containing $\mathbf{E}$. The advantage of using generalized solutions and the extended semigroup is that $\mathcal{F}$–stationary distributions are stationary more often. In Da Prato and Zabczyk [6], Section 5.5.3, there appears a theorem guaranteeing an extension to all of $\mathbf{H}$. We endow $\mathbf{D}$ with a uniform structure assuming that

- the measurable structures on $\mathbf{D}$ induced by this uniformity and by $\mathbf{H}$ are equal and
- the completion $\mathbf{X}$ of $\mathbf{D}$ is a Radon space.

An example is given in Remark 5.1(a).

4.4 The Condition (T_0) We next deal with a condition sufficient for (T), see 2.11, in the present situation; cf. Leha and Ritter [17]. In the special case of a constant sequence (V_n), cf. also Ichikawa [12], Proposition 4.1 and Corollary 4.3.

(T_0) There exists a sequence $(V_n)_{n\in\mathbf{N}}$ of nonnegative functions $V_n \in \mathcal{C}_u^2(\mathbf{H})$ and an initial point $x \in \mathbf{E}$ for X such that

(i) $V := \underline{\lim}_n V_n$ is lower semi-continuous and $V(x) < \infty$;

(ii) for all n, there exists a P-a.s. locally Lebesgue-integrable random function $\varphi_n : \mathbf{R}_+ \to [0,\infty]$ such that, $\lambda \otimes P$-a.s.,

$$\sup_k V_n'(\pi_k X)(-A\pi_k X) \leq \varphi_n;$$

(iii) there exists a constant $C \geq 0$ such that, for all n and all $y \in \mathbf{E}$, we have

$$V_n(y) \leq C(1+V(y)) \quad \text{and} \quad \overline{\lim}_k \mathcal{L}^{(k)} V_n(y) \leq C(1+V_n(y)).$$

(iv) $\overline{\lim}_n \overline{\lim}_k \mathcal{L}^{(k)} V_n$ is bounded above on $\mathbf{E}$;

(v) with

$$\eta(y) := \begin{cases} -\overline{\lim}_n \overline{\lim}_k \mathcal{L}^{(k)} V_n(y), & y \in \mathbf{E} \\ \infty, & y \in \mathbf{D} \setminus \mathbf{E}, \end{cases}$$

the subset $\{\eta \leq r\} \subseteq \mathbf{E}$ is relatively compact in $\mathbf{X}$ for all $r > 0$.

4.5 Proposition Condition (T_0) implies Condition (T). More precisely:
If $x \in \mathbf{E}$, η, and c are as in (T_0) then (T) is satisfied with $\mu = \delta_x$, the Dirac measure at the point x.

Proof. By definition, η is bounded below and, since $\mathbf{E}$ is measurable in $\mathbf{D}$, it is also measurable. Condition (T)(i) is just (T_0)(v). Positivity of V and the Forward Inequality 4.2 first imply

$$0 \leq EV(X(t)) \leq V(x) + E\int_0^t \overline{\lim}_n \overline{\lim}_k \mathcal{L}^{(k)} V_n(X(s))ds, \quad t \geq 0.$$

Since $\mu\mathcal{M}_t$ is supported by $\mathbf{E}$ we have, for $t \geq 1$,

$$\begin{aligned} \int_{\mathbf{D}} \eta(y)\mu\mathcal{M}_t(dy) &= -\int_{\mathbf{E}} \overline{\lim}_n \overline{\lim}_k \mathcal{L}^{(k)} V_n(y)\mathcal{M}_t(x,dy) \\ &= -\frac{1}{t}E\int_0^t \overline{\lim}_n \overline{\lim}_k \mathcal{L}^{(k)} V_n(X(s))ds \\ &\leq V(x); \end{aligned}$$

together with (T_0)(i) this implies (T)(ii). □

In order to formulate Lyapunov conditions for (F) and (A) in terms of the coefficients of X we follow [17] using the linked diffusion. Recall that the superscript x denotes the starting point of the process X^x.

4.6 The linked diffusion $\bar{X}$ The *linked diffusion* $\bar{X} = \bar{X}^{x,y}$ defined by

$$\bar{X}^{x,y}(t) := (X^x(t), X^y(t)), \quad x,y \in \mathbf{E}, t \geq 0,$$

is an Itô diffusion in $\mathbf{E} \times \mathbf{E}$ satisfying the initial value problem

$$\begin{aligned} d\bar{X}(t) &= (-\bar{A}\bar{X}(t) + \bar{f}(\bar{X}(t)))dt + \bar{\sigma}(\bar{X}(t))dW(t), \quad t \geq 0, \\ \bar{X}(0) &= (x,y) \in \mathbf{E} \times \mathbf{E}. \end{aligned}$$

Here, $\bar{A}$ is the self-adjoint operator in $\mathbf{H}\times\mathbf{H}$ defined by $\bar{A}(u,v)=(Au,Av)$, $u,v\in\mathcal{D}(A)$, $\bar{f}:\mathbf{E}\times\mathbf{E}\to\mathbf{H}\times\mathbf{H}$ is defined by $\bar{f}(u,v)=(f(u),f(v))$, and $\bar{\sigma}:\mathbf{E}\times\mathbf{E}\to\mathcal{L}(\mathbf{K},\mathbf{H}\times\mathbf{H})$ by $\bar{\sigma}(u,v)(k)=(\sigma(u)k,\sigma(v)k)$. If X is a generalized solution with initial values in $\mathbf{D}$, see Section 4.3, then $\bar{X}$ is defined for initial values in $\mathbf{D}\times\mathbf{D}$ and the associated *linked semigroup* $\bar{\mathcal{P}}_t$ satisfies the requirements of Section 2.12. The sequence π_k defined in 3.2.3 gives rise to the approximation of identity associated with $\bar{A}$, $\bar{\pi}_k:\mathbf{H}\times\mathbf{H}\to\mathbf{H}\times\mathbf{H}$, $\bar{\pi}_k(u,v)=(\pi_k u,\pi_k v)$. Let $U\in\mathcal{C}^2(\mathbf{H})$ and $k\in\mathbf{N}$. The counterpart of the operator $\mathcal{L}^{(k)}$ defined in Eq. (9) in the present situation of the linked diffusion, applied to the function $\bar{U}\in\mathcal{C}^2(\mathbf{H}\times\mathbf{H})$, $\bar{U}(u,v):=U(v-u)$, is the operator $\mathcal{L}_d^{(k)}$ defined by

$$\begin{aligned}
&\mathcal{L}_d^{(k)}U(x,y)\\
=\;&\bar{U}'(\bar{\pi}_k(x,y))\left[-\bar{A}\bar{\pi}_k(x,y)+\bar{\pi}_k\bar{f}(x,y)\right]+tr\,\frac{1}{2}\bar{U}''(\bar{\pi}_k(x,y))(\bar{\pi}_k\bar{\sigma}(x,y)Q\bar{\sigma}^*(x,y)\bar{\pi}_k\cdot,\cdot)\\
=\;&U'(\pi_k(y-x))\left[-A\pi_k(y-x)+\pi_k(f(y)-f(x))\right]+\frac{1}{2}tr\,U''(\pi_k(y-x))(\bar{\mathcal{A}}_k(x,y)\cdot,\cdot),
\end{aligned}$$

$x,y\in\mathbf{E}$, where $\bar{\mathcal{A}}_k(x,y)=\pi_k(\sigma(y)-\sigma(x))Q(\sigma^*(y)-\sigma^*(x))\pi_k$. Of course, $\bar{\mathcal{A}}_k=0$ if σ is constant.

We next formulate sufficient conditions for (F) and (A) (cf. 2.13) in terms of drift and diffusion of the process, more precisely of the linked differential operators $\mathcal{L}_d^{(k)}$.

4.7 The Conditions (F_0) and (A_0)

(F_0) There exists a metric $d\in\mathcal{D}$, a function $U\in\mathcal{C}^2(\mathbf{H})$, and a real number $p>0$ such that $U(y-x)=d^p(x,y)$ for all $x,y\in\mathbf{D}$ with the following properties:

(i) there exists a P-a.s. locally integrable random function $\varphi:[0,\infty[\to[0,\infty]$ such that, for all $x,y\in\mathbf{E}$, we have $\lambda\otimes P$-a.s.

$$\sup_k U'\left(\pi_k(X^y-X^x)\right)\left(-A\pi_k(X^y-X^x)\right)\le\varphi;$$

(ii) there exists a constant $c\in\mathbf{R}$ such that

$$\overline{\lim}_k\mathcal{L}_d^{(k)}\,U(u,v)\le cU(v-u),\quad u,v\in\mathbf{E};$$

(iii) $\mathcal{C}_{ub}^d(\mathbf{X})$ separates $\mathrm{Prob}(\mathbf{X})$.

(A_0) Condition (F_0) with $c<0$ in (ii) and convergence determination instead of separation in (iii).

4.8 Proposition The Conditions (F_0) and (A_0) imply Conditions (F) and (A), respectively, with the same metric d and $c(t)=e^{ct}$, $t\ge0$.

Proof. Let $x,y\in\mathbf{E}$, $t>0$. The Forward Inequality 4.2 admits a straightforward generalization to the case where V_n and V depend also on time t. An application of this generalized version to the process $\bar{X}=\bar{X}^{x,y}$ and the functions $\bar{V}_n(t,x,y)=\bar{V}(t,x,y):=e^{-ct}U(y-x)$ instead of X and V_n yields

$$E\bar{V}(t,\bar{X}(t))\le\bar{V}(0,x,y)+E\int_0^t\overline{\lim}_k\tilde{\mathcal{L}}^{(k)}\bar{V}(s,\bar{X}(s))ds$$

with $\tilde{\mathcal{L}}^{(k)}\bar{V}(s,x,y)=e^{-cs}\left(-cU(y-x)+\mathcal{L}_d^{(k)}U(x,y)\right)$ and the constant c appearing in (F_0)(ii). Indeed, 4.2(ii) and 4.2(iii) are just Conditions (F_0)(i) and (F_0)(ii), respectively; cf. also 4.6. Now, again by Condition (F_0)(ii), $\overline{\lim}_k\tilde{\mathcal{L}}^{(k)}\bar{V}(s,\bar{X}(s))\le0$ and, hence,

(11) $e^{-ct}\bar{\mathcal{P}}_t\bar{U}(x,y)\le\bar{U}(x,y),\quad x,y\in\mathbf{E}.$

It follows from the definition of a generalized solution and from Fatou's lemma that this estimate is even true for $x,y\in\mathbf{D}$; it implies (F)(i) with $c(t)=e^{ct}$. Condition (F)(ii) is just (F_0)(iii). If $c<0$ then (11) plainly implies (A)(i). This proves the claim. □

Theorem 2.15 and Propositions 4.5 and 4.8 combine to show the following general result for Itô diffusions.

4.9 Theorem (a) If Conditions (T_0) and (F_0) (cf. 4.4 and 4.7) are satisfied then $(\mathcal{P}_t)_t$ has the $\mathcal{C}_{ub}^d(\mathbf{X})$-Feller property and possesses a $\mathcal{C}_{ub}^d(\mathbf{X})$–stationary distribution γ on $\mathbf{X}$ with the metric d appearing in (F_0). We have $\gamma(\{\hat{\eta} < \infty\}) = 1$, where $\hat{\eta}$ is the lower semi-continuous regularization of the function η appearing in (T_0). If $\hat{\eta} = \infty$ on $\mathbf{X} \setminus \mathbf{D}$ then γ is a stationary distribution with respect to $(\mathcal{P}_t)_t$.

(b) Suppose that Condition (A_0) (cf. 4.7) is satisfied and let d be the metric appearing in (F_0). If $(\mathcal{P}_t)_t$ possesses a $\mathcal{C}_{ub}^d(\mathbf{X})$–stationary distribution γ then there is trend to γ; i.e., $\mu\mathcal{P}_t \to \gamma$ narrowly on $\mathbf{X}$ as $t \to \infty$ for all probability distributions μ on $\mathbf{D}$. In particular, γ is unique.

5 Stationary distributions of Itô diffusions – existence, uniqueness, and trend

Let the set-up be as in Sections 3 and 4.3 and recall the definition of $\mathcal{C}_{ub}^d(\mathbf{X})$ in Section 2. In this and the following section, the Hilbert-Schmidt operators π_k are chosen as spectral theoretic functions of A, $\pi_k = h_k(A)$, $h_k \colon \mathbf{R}_+ \to \mathbf{R}_+$ such that $h_k \uparrow 1$ as $k \to \infty$ and $rh_k(r)$ is bounded for $r \geq 0$. We denote

$$R_n^{(\alpha)} = R_n(-A^\alpha) = (nI + A^\alpha)^{-1}$$

the resolvent operator of $-A^\alpha$ and

$$A_n^{(\alpha)} = A^\alpha n R_n^{(\alpha)} = R_{1/n}(-A^{-\alpha})$$

the Yosida approximation of A^α, $\alpha \in \mathbf{R}$. We also define the norm $\|By\|$ of a positive, self-adjoint operator B at a point $y \notin \mathcal{D}(B)$ as ∞. Note also that, given another positive, self-adjoint operator T, the trace $\mathrm{tr} BT := \sum_k \|\sqrt{B}\sqrt{T}c_k\|^2$ (c_k an orthonormal basis of $\mathbf{H}$) is always defined as an element in $[0, \infty]$.

We now illustrate how a uniform structure on $\mathbf{D}$ and Lyapunov functions V_n and U adapted to X and satisfying the requirements stated in Conditions (T_0), (F_0), and (A_0) may be chosen. In view of (F_0), the uniform structure is chosen in accordance with U. A crucial point is the proper choice of V_n and U related to the parameters of the given differential equation. We restrict matters to positive quadratic forms V_n and U and assume from here on $0 \in \mathbf{E}$.

5.1 Construction of X, U, and V_n (a) Any positive-definite operator $S \in \mathcal{L}(\mathbf{H})$ induces the quadratic form $U(y) = (y, Sy)$, $y \in \mathbf{H}$, and the S–metric d on $\mathbf{D}$ defined by $d(y,z) = \|z-y\|_S = \sqrt{(z-y, S(z-y))}$. Let $\mathbf{X}$ be the completion of $\mathbf{D}$ with respect to d. We call the extension of d to $\mathbf{X}$ the S–metric on $\mathbf{X}$. This metric being separable, $\mathbf{X}$ is a completely regular Radon space, cf. [17], Section 2.5. The Borel structures of $\mathbf{X}$ and $(\mathbf{H}, \|\cdot\|)$ coincide on $\mathbf{D}$ (cf. again [17], Section 2.5), i.e., the assumptions stated in 4.3 are satisfied. Note that $\sqrt{S}$ induces an isometry $\Phi_S : \mathbf{X} \to \Phi_S \mathbf{X} \subseteq \mathbf{H}$ and we have $\mathcal{C}_{ub}(\mathbf{X}) = \mathcal{C}_{ub}^d(\mathbf{X})$. Moreover, any uniformly d-continuous, bounded function on $\mathbf{D}$ has a unique extension to a function in $\mathcal{C}_{ub}^d(\mathbf{X})$.

Conversely, let $\mathbf{X}$ be a subset of some Hilbert space $\mathbf{L}$ with inner product $(\cdot,\cdot)_\mathbf{L}$ and assume that the linear structures on $\mathbf{D}$ induced by $\mathbf{H}$ and $\mathbf{L}$ are equal. Put $\mathbf{D}_0 := \langle \mathbf{D} \rangle$, the linear hull of $\mathbf{D}$ in $\mathbf{H}$ endowed with the inner product of $\mathbf{H}$ and let $J : \mathbf{D}_0 \to \mathbf{L}$ be the natural embedding. Suppose that J is continuous and let $J^* : \mathbf{L} \to \mathbf{H}$ be its adjoint operator. Choose a positive-definite, linear operator $S \in \mathcal{L}(\mathbf{H})$ such that $S_{|\mathbf{D}_0} = J^*J$. For $y, z \in \mathbf{D}$, we have $\|z-y\|_S = \|z-y\|_\mathbf{L}$ and we are in the situation of the previous paragraph. If $\mathbf{D}$ is total in $\mathbf{H}$ then, of course, S is uniquely determined.

(b) We choose the functions V_n required in (T_0) as quadratic forms $V_n(y) = \frac{1}{2}(y, B_n y)$, $y \in \mathbf{H}$, where $(B_n)_n$ is an increasing sequence of positive, self-adjoint operators in $\mathcal{L}(\mathbf{H})$.

5.2 Remarks (a) Let us compute the operators $\mathcal{L}^{(k)}$ and $\mathcal{L}_d^{(k)}$ appearing in (T_0), (F_0), and (A_0) in the special case where the functions U in (F_0) and (A_0) and V_n in (T_0) are quadratic forms $W(y) = \frac{1}{2}(y, By)$ with positive, self-adjoint operators $B \in \mathcal{L}(\mathbf{H})$. For $x, y \in \mathbf{E}$, we have $W'(\pi_k y)A\pi_k y =$

$(B\pi_k y, A\pi_k y)$,

$$
\begin{aligned}
(12)\quad \mathcal{L}^{(k)}W(y) &= (B\pi_k y, -A\pi_k y + \pi_k f(y)) + \frac{1}{2} tr\, B\pi_k \sigma(y) Q\sigma^*(y)\pi_k, \quad \text{and} \\
(13)\quad \mathcal{L}_d^{(k)}W(y,z) &= (B\pi_k(z-y), -A\pi_k(z-y) + \pi_k(f(z)-f(y))) \\
&\quad + \frac{1}{2} tr\, B\pi_k(\sigma(z)-\sigma(y))Q(\sigma^*(z)-\sigma^*(y))\pi_k.
\end{aligned}
$$

If σ is constant then the last summand on the right side of (13) vanishes.

(b) Parts of the Conditions (F_0), (A_0), and (T_0) are automatically satisfied in the situation of Section 5.1: From the choice of $\mathbf{X}$ and U it follows that Parts (iii) in (F_0) and (A_0) are satisfied. Moreover, the definition of V_n implies Condition (T_0)(i) with $x = 0$, (T_0)(ii) with $\varphi_n = 0$, and also the first inequality in (T_0)(iii). According to (T_0), (F_0), and (A_0), we need to take superior limits in $\mathcal{L}_d^{(k)}U$ and $\mathcal{L}^{(k)}V_n$ with respect to k and n. In passing to the superior limits of the terms on the right sides of (12) and (13) we have to exclude the situation $-\infty + \infty$. This is the reason for the subdivisions appearing in 5.3(i) (for $\overline{\lim}_k$) and 5.4(a) (for $\overline{\lim}_n$).

5.3 Proposition Let $S \in \mathcal{L}(\mathbf{H})$ be positive definite, let $\mathbf{X}$ be constructed with the aid of the S-metric as described in 5.1(a), let $\alpha \in \mathbf{R}$, and let $\theta : \mathbf{D} \to \mathbf{R}$ be a lower bounded, measurable function.

(a) Suppose that

(i) there exists a constant $C \geq 0$ such that, for all $n \in \mathbf{N}$, we have

$$
\begin{cases}
-\|\sqrt{A_n^{(\alpha)}A}y\|^2 + (A_n^{(\alpha)}y, f(y)) + \frac{1}{2} tr\, A_n^{(\alpha)}\sigma(y)Q\sigma^*(y) \leq C(1 + (A_n^{(\alpha)}y, y)), \\
\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad y \in \mathbf{E} \cap \mathcal{D}(A^{\frac{1-\alpha^-}{2}}), \\
tr A_n^{(\alpha)}\sigma(y)Q\sigma^*(y) < \infty, \qquad\qquad\qquad y \in \mathbf{E} \setminus \mathcal{D}(A^{\frac{1-\alpha^-}{2}});
\end{cases}
$$

(ii) $\overline{\lim}_n \left(-\|\sqrt{A_n^{(\alpha)}A}y\|^2 + (A_n^{(\alpha)}y, f(y)) + \frac{1}{2} tr\, A_n^{(\alpha)}\sigma(y)Q\sigma^*(y) \right) \leq -\theta(y)$ for all $y \in \mathbf{E} \cap \mathcal{D}(A^{\frac{1-\alpha^-}{2}})$;

(iii) the set $\{\theta \leq r\} \subseteq \mathbf{D}$ is relatively compact in $\mathbf{X}$ for all $r > 0$;

(iv) for all $x, y \in \mathbf{E}$, there exists a P-a.s. locally–integrable random function $\varphi : [0,\infty[\to [0,\infty]$ such that, $\lambda \otimes P$-a.s., we have

$$\sup_k (S\pi_k(X^y - X^x), -A\pi_k(X^y - X^x)) \leq \varphi;$$

(v) there exists a constant $c \in \mathbf{R}$ such that, for all $x, y \in \mathbf{E}$, we have

$$
\begin{aligned}
& \overline{\lim}_k (S\pi_k(y-x), -A\pi_k(y-x)) + (S(y-x), f(y)-f(x)) \\
& + \frac{1}{2} trS(\sigma(y)-\sigma(x))Q(\sigma^*(y)-\sigma^*(x)) \\
\leq\ & c(S(y-x), y-x).
\end{aligned}
$$

Then $(\mathcal{P}_t)_t$ has the $\mathcal{C}_{ub}(\mathbf{X})$-Feller property and possesses a $\mathcal{C}_{ub}(\mathbf{X})$–stationary distribution γ on $\mathbf{X}$. We have $\gamma(\{\hat{\theta} < \infty\}) = 1$, where $\hat{\theta}$ is the lower semi-continuous regularization of θ on $\mathbf{X}$. If $\hat{\theta} = \infty$ on $\mathbf{X} \setminus \mathbf{D}$ then γ is a stationary distribution with respect to $(\mathcal{P}_t)_t$.

(b) Suppose that (v) is satisfied with $c < 0$. If there exists a $\mathcal{C}_{ub}(\mathbf{X})$–stationary distribution γ then there is trend to γ.

Proof. In view of applying Theorem 4.9 we first verify (T_0) with $V_n(y) = \frac{1}{2}(A_n^{(\alpha)}y, y)$, $y \in \mathbf{H}$, and $x = 0$. Conditions (T_0)(i),(ii) and the first part of (T_0)(iii) follow from positivity of A as noted in 5.2(b). In order to prove the second part of (T_0)(iii), note first that we have for all $y \in \mathbf{H}$

$$(14)\quad \overline{\lim}_k(-\|\sqrt{A_n^{(\alpha)}}A\pi_k y\|^2) = \overline{\lim}_k \int_0^\infty \frac{-n\, t^{1+\alpha}}{t^\alpha + n} h_k^2(t)\mu_y(dt) = -\int_0^\infty \frac{n\, t^{1+\alpha}}{t^\alpha + n}\mu_y(dt) = -\|\sqrt{A_n^{(\alpha)}}Ay\|^2,$$

where μ_y denotes the spectral measure associated with the operator A and y. (By definition, the last line is $-\infty$ if $y \notin \mathcal{D}\left(\sqrt{A_n^{(\alpha)}}A\right)$. Since there exists a number $\beta(\alpha, n, \varepsilon) > 0$ such that

$$\beta(\alpha, n, \varepsilon)\; t^{1-\alpha^-} \le \frac{t^{1+\alpha}}{t^\alpha + n} \le t^{1-\alpha^-}, \quad t \ge \varepsilon > 0,$$

it follows from uniform positivity of A that the domain of definition $\mathcal{D}(\sqrt{A_n^{(\alpha)}}A)$ equals $\mathcal{D}(A^{\frac{1-\alpha^-}{2}})$ for all n. Hence,

$$(15)\quad \overline{\lim}_k \mathcal{L}^{(k)}V_n(y) = -\infty, \quad y \in \mathbf{E} \setminus \mathcal{D}(A^{\frac{1-\alpha^-}{2}}),$$

for all n by (i) and (14). If $y \in \mathbf{E} \cap \mathcal{D}(A^{\frac{1-\alpha^-}{2}})$ then the second part of (T_0)(iii) follows again from (i) and (14). Thus, we have the second half of (T_0)(iii). For $y \in \mathbf{E} \setminus \mathcal{D}(A^{\frac{1-\alpha^-}{2}})$, (T_0)(iv) follows from (15) and for $y \in \mathbf{E} \cap \mathcal{D}(A^{\frac{1-\alpha^-}{2}})$ this is (ii).

We show next (T_0)(v). If $y \in \mathbf{E} \cap \mathcal{D}(A^{\frac{1-\alpha^-}{2}})$ then, according to (ii),

$$\eta(y) = -\overline{\lim}_n\, \overline{\lim}_k \mathcal{L}^{(k)}V_n(y) \ge \theta(y).$$

If $y \in \mathbf{E} \setminus \mathcal{D}(A^{\frac{1-\alpha^-}{2}})$ then, by (15),

$$\eta(y) = -\overline{\lim}_n \overline{\lim}_k \mathcal{L}^k V_n(y) = \infty.$$

Hence, $\eta \ge \theta$ and $\{y \in \mathbf{D}/\eta(y) \le r\} \subseteq \{y \in \mathbf{D}/\theta(y) \le r\}$; therefore, (iii) implies Condition (T_0)(v).

The Condition (F_0) with the S–metric d on $\mathbf{X}$ and $p = 2$ follows from (iv) and (v) and from 2.7(b). Now, Theorem 4.9(a) guarantees the existence of a $\mathcal{C}_{ub}(\mathbf{X})$–stationary distribution γ on $\mathbf{X}$ satisfying $\gamma(\{\hat{\eta} < \infty\}) = 1$.

Finally, $\eta \ge \theta$ implies $\hat{\eta} \ge \hat{\theta}$ and $\{\hat{\eta} < \infty\} \subseteq \{\hat{\theta} < \infty\}$. This proves Part (a).

Part (b) is a direct consequence of Theorem 4.9(b). □

5.4 Remarks (a) If $\overline{\lim}_n \left((A_n^{(\alpha)}y, f(y)) + \frac{1}{2} tr\, A_n^{(\alpha)}\sigma(y)Q\sigma^*(y)\right) < \infty$ for all $y \in \mathcal{D}(A^{\frac{1-\alpha^-}{2}})$ then the left hand side of 5.3(ii) is $-\infty$ for $y \notin \mathcal{D}(A^{\frac{1+\alpha}{2}})$ and it is sufficient to require 5.3(ii) for $y \in \mathbf{E} \cap \mathcal{D}(A^{\frac{1+\alpha}{2}})$ instead of $\mathbf{E} \cap \mathcal{D}(A^{\frac{1-\alpha^-}{2}})$.

(b) If S is chosen such that SA is positive then Condition 5.3(iv) is satisfied with $\varphi = 0$ since the range of π_k is contained in $\mathcal{D}(A)$ by the choice of h_k at the beginning of this section.

5.5 Corollary Let $S \in \mathcal{L}(\mathbf{H})$ be positive definite and let $\alpha \in \mathbf{R}$.

(a) Suppose that

(i) 5.3(i) and 5.3(v) are satisfied;

(ii) $$\overline{\lim}_n \left((A_n^{(\alpha)}y, f(y)) + \tfrac{1}{2} tr\, A_n^{(\alpha)}\sigma(y)Q\sigma^*(y)\right) \begin{cases} \le c_T + c_1\|A^{\frac{1+\alpha}{2}}y\|^2, & y \in \mathbf{E} \cap \mathcal{D}(A^{\frac{1+\alpha}{2}}), \\ < \infty, & y \in \mathbf{E} \cap \mathcal{D}(A^{\frac{1-\alpha^-}{2}}), \end{cases}$$
for some $c_T \in \mathbf{R}$, $c_1 < 1$;

(iii) $\mathcal{D}(A^{-\frac{1+\alpha}{2}}) = \mathbf{H}$ and the operator $T := \sqrt{S}A^{-\frac{1+\alpha}{2}}$ is compact on $\mathbf{H}$;

(iv) $(Sy, Ay) \ge 0, \quad y \in \mathcal{D}(A)$.

Let $\mathbf{X}$ be constructed with the aid of the S-metric as described in 5.1(a). Then $(\mathcal{P}_t)_t$ has the $\mathcal{C}_{ub}(\mathbf{X})$-Feller property and possesses a $\mathcal{C}_{ub}(\mathbf{X})$–stationary distribution γ with respect to $(\mathcal{P}_t)_t$ on $\mathbf{X}$. Moreover, γ is supported by $\mathcal{D}(A^{\frac{1+\alpha}{2}})$. If $\mathcal{D}(A^{\frac{1+\alpha}{2}}) \subseteq \mathbf{D}$ then γ is a stationary distribution.

(b) If, in addition, 5.3(v) is satisfied with $c < 0$ then there is trend to γ.

Proof. Let us first verify the Conditions of Proposition 5.3 with

$$\theta(y) = -c_T + (1-c_1)\ \|A^{\frac{1+\alpha}{2}} y\|^2, \quad y \in \mathbf{D}.$$

Condition 5.3(iv) with $\varphi = 0$ follows from (iv) as noted in 5.4(b). For $y \in \mathbf{E} \cap \mathcal{D}(A^{\frac{1+\alpha}{2}})$, the left side of 5.3(ii) may be estimated

$$\begin{aligned} & -\|A^{\frac{1+\alpha}{2}} y\|^2 + \overline{\lim}_n \left((A_n^{(\alpha)} y, f(y)) + \frac{1}{2} tr\ A_n^{(\alpha)} \sigma(y) Q \sigma^*(y) \right) \\ \leq\ & -\|A^{\frac{1+\alpha}{2}} y\|^2 + c_T + c_1 \|A^{\frac{1+\alpha}{2}} y\|^2 \\ =\ & c_T - (1-c_1)\|A^{\frac{1+\alpha}{2}} y\|^2 \\ =\ & -\theta(y) \end{aligned}$$

by (ii). Together with 5.4(a), this proves 5.3(ii).

In view of 5.3(iii), we have for $r \geq -c_T$

$$\{\theta \leq r\} = \mathbf{D} \cap \{y \in \mathcal{D}(A^{\frac{1+\alpha}{2}}) / \|A^{\frac{1+\alpha}{2}} y\|^2 \leq \frac{r + c_T}{1 - c_1}\} \subseteq \mathbf{D} \cap A^{-\frac{1+\alpha}{2}} B =: D,$$

where B is the ball of radius $\sqrt{(r + c_T)/(1 - c_1)}$ in $\mathbf{H}$. By isometry and (iii), D is a totally bounded subset of $\mathbf{D}$ and, hence, relatively compact in $\mathbf{X}$; this is 5.3(iii).

It remains to show that the set $\{\hat{\theta} < \infty\}$ is contained in $\mathcal{D}(A^{\frac{1+\alpha}{2}})$. Since S and A are injective so is T. Moreover, the range of T is dense in $\mathbf{H}$ since S is positive definite and $A^{\frac{1+\alpha}{2}}$ is densely defined. Hence, T^* and TT^* are both injective. Now, we have $\left((TT^*)^{-1/2}\right)^* (TT^*)^{-1/2} = (T^{-1})^*(T^{-1})$; hence

(16) $\mathcal{D}\left((TT^*)^{-1/2}\right) = \mathcal{D}\left(T^{-1}\right)$ and

(17) $\|(TT^*)^{-1/2} y\| = \|T^{-1} y\|, \quad y \in \mathcal{D}(T^{-1}),$

cf., e.g., Weidmann [27], Satz 5.40. Let us show

(18) $\|A^{\frac{1+\alpha}{2}} y\| = \|(TT^*)^{-1/2} \sqrt{S} y\|, \quad y \in \mathbf{H},$

where the right side is infinite if $\sqrt{S} y \notin \mathcal{D}((TT^*)^{-1/2})$. Indeed, for $y \in \mathcal{D}(A^{\frac{1+\alpha}{2}})$ we have $TA^{\frac{1+\alpha}{2}} y = \sqrt{S} y$, i.e., $\sqrt{S} y \in \mathcal{R}(T) = \mathcal{D}(T^{-1})$, and (16) and (17) imply

$$\|(TT^*)^{-1/2} \sqrt{S} y\| = \|T^{-1} \sqrt{S} y\| = \|A^{\frac{1+\alpha}{2}} y\|.$$

For $y \notin \mathcal{D}(A^{\frac{1+\alpha}{2}})$, we have $\sqrt{S} y \notin \mathcal{R}(T) = \mathcal{D}(T^{-1}) = \mathcal{D}((TT^*)^{-1/2})$ by (16). Hence, both sides of (18) are infinite in this case.

By positivity of TT^*, the function $z \mapsto \|(TT^*)^{-1/2} z\|$ is lower semi-continuous on $\mathbf{H}$. Therefore, by isometry, the function $x \mapsto \|(TT^*)^{-1/2} \Phi_S x\|$ is lower semi-continuous on $\mathbf{X}$ and (18) implies $\hat{\theta}(x) \geq -c_T + (1 - c_1)\|(TT^*)^{-1/2} \Phi_S x\|^2$ for all $x \in \mathbf{X}$. Consequently, if $x \in \mathbf{X}$ is such that $\hat{\theta}(x) < \infty$ then

$$\Phi_S x \in \mathcal{D}\left((TT^*)^{-1/2}\right) = \mathcal{D}\left(T^{-1}\right) = \mathcal{R}(T),$$

i.e., $\Phi_S x = \sqrt{S} y$ for some $y \in \mathcal{R}(A^{-\frac{1+\alpha}{2}})$. Therefore, $x = y \in \mathcal{D}(A^{\frac{1+\alpha}{2}})$. The Claims (a) and (b) now follow from Proposition 5.3. □

5.6 The linear case Zabczyk [28], see also Da Prato and Zabczyk [5], Theorem 11.7, gave a complete answer to the question of existence of stationary distributions in the linear case, i.e., if $f = 0$ and $\sigma =$ identity. He considered a not necessarily positive operator A, generator of a C_0-semigroup $S(t)$, and showed that a stationary distribution of (1) exists if and only if $tr \int_0^\infty S(t)QS(t)ds < \infty$. This is equivalent to the condition

(19) $\sup_n tr\, R_{\frac{1}{n}}(-A)\, Q < \infty.$

Let us show that, in the linear case, Corollary 5.5 is sharp in the present set-up, i.e., if A is uniformly positive, self-adjoint, and not necessarily bounded and if Q is not necessarily nuclear, cf. also [17], 2.25. To this end, let $\mathbf{E} = \mathbf{D} = \mathbf{H}$, $\alpha = -1$, and let S be any compact, self-adjoint, positive-definite operator on $\mathbf{H}$ commuting with A such that SA is continuous and positive. Such an operator always exists since $A \geq 0$. Note that $A_n^{(-1)} = R_{1/n}(-A)$. The Condition 5.5(ii) is equivalent to (19) and implies 5.3(i). Condition 5.5(iii) is just compactness of S.

Plainly, all conditions of Corollary 5.5(a) are satisfied and we have $\mathcal{D}(A^{\frac{1+\alpha}{2}}) = \mathbf{H} = \mathbf{E}$. It follows that $(\mathcal{P}_t)_t$ has a stationary distribution γ on $\mathbf{H}$. Moreover, if the spectrum of A is bounded away from 0 then 5.5(b) shows that there is trend to equilibrium.

6 Stationary distributions of nonlinear stochastic partial differential equations

In this section we show that the foregoing material is applicable to stochastic partial differential equations, in particular, to reaction-diffusion processes.

6.1 Notation Let us consider a formal stochastic partial differential equation

(20) $$\frac{\partial u}{\partial t}(t,\xi) = \sum_{i,j=1}^{d} \frac{\partial}{\partial \xi_i}\left(a_{ij}(\xi)\frac{\partial}{\partial \xi_j}u(t,\xi)\right) u(t,\xi) + F(u(t,\xi)) + G(u(t,\xi))\eta(t,\xi),$$

$(t,\xi) \in \mathbf{R}_+ \times \mathcal{O}$. Here $\mathcal{O} \subset \mathbf{R}^d$ is a bounded domain with a smooth boundary, η stands for a Gaussian noise correlated in space and white in time, $F : \mathbf{R} \to \mathbf{R}$ is measurable and locally bounded, $G : \mathbf{R} \to \mathbf{R}$ is Lipschitz-continuous, the second-order differential operator on the right side is uniformly-elliptic, and $a_{ij} \in \mathcal{C}^\infty(\bar{\mathcal{O}})$, $i,j = 1,2,\ldots d$, $a_{ij} = a_{ji}$. We consider Equation (20) with initial condition

(21) $u(0,\cdot) = u_0 \in \mathbf{L}^2(\mathcal{O})$

and boundary conditions of *Dirichlet* type, i.e.,

(22) $u(t,\xi) = 0, \quad (t,\xi) \in \mathbf{R}_+ \times \partial\mathcal{O}.$

The formal system (20), (21), and (22) can be given a precise meaning in the sense of Equation (1). To this end, we assume that η gives rise to a Wiener process $W(t)$ with covariance operator Q on a Hilbert space $\mathbf{K}$ continuously embedded in $\mathbf{L}^\infty(\mathcal{O})$ by means of a continuous, linear mapping $J : \mathbf{K} \to L^\infty(\mathcal{O})$; moreover, we specify $\mathbf{H} = \mathbf{L}^2(\mathcal{O})$, $\mathbf{E} = \mathcal{C}(\bar{\mathcal{O}})$, and

(23) $A =$ the closed extension of $-\mathcal{L}$ to $\mathcal{D}(A) = H_0^1(\mathcal{O}) \cap H^2(\mathcal{O})$,

$$f(y)(\xi) := F(y(\xi)), \quad \xi \in \mathcal{O},\ y \in \mathbf{E},$$
$$[\sigma(y)h](\xi) := G(y(\xi))Jh(\xi), \quad \xi \in \mathcal{O},\ y \in \mathbf{E},\ h \in \mathbf{K}.$$

It follows from (23) that $A : \mathcal{D}(A) \to \mathbf{H}$ is a self-adjoint and uniformly positive operator. Since F is locally bounded, f maps $\mathbf{E}$ into $\mathbf{H}$ and we have

(24) $\|f(y)\| \leq M\|y\|_\infty, \qquad y \in \mathbf{E}.$

Moreover, since G is Lipschitzian,

$$\|\sigma(y)h\|_{L^2}^2 = \int_{\mathcal{O}} |G(y(\xi))Jh(\xi)|^2 d\xi \leq C^2(\|y\|^2+1)\|J\|^2\|h\|_{\mathbf{K}}^2, \qquad y \in \mathbf{E},\ h \in \mathbf{K},$$

and hence

$$\begin{aligned} tr\mathcal{A}(y) &= tr\sigma(y)Q\sigma^*(y) \leq \|\sigma(y)Q\sigma^*(y)\| \leq \|Q\|\|\sigma(y)\|^2 \\ &\leq C^2\|J\|^2\|Q\|(\|y\|^2+1), \qquad y \in \mathbf{E}. \end{aligned} \tag{25}$$

Let $(X(t))_{t\geq 0}$ be a mild solution to Equation (1). If $t \to X(t)$ is P-a.s. locally bounded in $\mathcal{C}(\bar{\mathcal{O}})$ then (24) and (25) imply that both Assumptions (7) and (8) of Proposition 3.3 are satisfied. This proposition now shows that X is also a π-solution. Concerning the existence of such mild solutions we refer the reader to the literature, e.g., Da Prato and Zabczyk [5], Theorem 7.10 and Example 7.11.

Our aim is to derive from Corollary 5.5 a theorem on stationary distributions of mild solutions for fairly general functions F and G. In order to control the nonlinear drift part we first state two lemmas which are adapted from Leha et al. [18], Lemma 6.3 and Remark 6.5. The first one is more general than we actually need here.

6.2 Lemma (dissipativity) Let $(M, \mathcal{B}, \mu)$ be a measure space, let K be a submarkovian kernel on M such that $\mu K \leq \mu$, and let $H : \mathbf{R} \to \mathbf{R}$ be a function vanishing at zero.

(a) If H is strictly increasing then, for all bounded, measurable functions y on M, we have

$$\int_M \mu(d\xi)\, |H(y(\xi))|\, K|y|(\xi) \leq \int_M \mu(d\xi) H(y(\xi))y(\xi).$$

(b) If μ is finite then the same is true if H is (not necessarily strictly) increasing.

Proof. (a) Let $\Phi(u) = \int_0^u H(r)dr$ and $\Psi(v) = \int_0^v H^{-1}(s)ds$, where H^{-1} is the inverse function with respect to composition. Using Young's inequality and the properties of K we first estimate

$$\begin{aligned} & \int_M \mu(d\xi) \left(\int_M K(\xi, d\eta)|y(\eta)| \right) |H(y(\xi))| \\ \leq & \int_M \mu(d\xi) \int_M K(\xi, d\eta)[\Phi(|y(\eta)|) + \Psi(|H(y(\xi))|)] \\ \leq & \int_M \mu(d\eta)\Phi(|y(\eta)|) + \int_M \mu(d\xi)\Psi(|H(y(\xi))|) \\ = & \int_M \mu(d\xi) \left[\Phi(|y(\xi)|) + \Psi(|H(y(\xi))|)\right]. \end{aligned} \tag{26}$$

It is well known that $\Phi(a) + \Psi(H(a)) = aH(a)$ for $a \geq 0$ (cf. the equality case in Young's inequality). Hence, (26) equals $\int_M \mu(d\xi)|y(\xi)||H(y(\xi))|$ which, by positivity of $uH(u)$, yields Claim (a).

(b) If H is only increasing then $H + \varepsilon\, id$ is strictly increasing and we obtain the claim from (a) by letting $\varepsilon \to 0$ since μ is finite here. □

6.3 Lemma Let $H : \mathbf{R} \to \mathbf{R}$ be an increasing function vanishing at zero, let $n \in \mathbf{N}$, and assume that $nR_n^{(\alpha)}$ induces a kernel K on $\mathcal{O}$ which satisfies the assumptions of Lemma 6.2 with $M = \mathcal{O}$ and Lebesgue measure μ on $\mathcal{O}$. Then, for all bounded, measurable functions $y : \mathcal{O} \to \mathbf{R}$, we have

$$(A_n^{(\alpha)}y, H \circ y)_{\mathbf{L}^2} \geq 0.$$

Proof. The claim follows from the identity $A_n^{(\alpha)} = n(I - nR_n^{(\alpha)})$ and Lemma 6.2(b).

□

Our next lemma deals with the diffusion term.

6.4 Lemma (a) If $G : \mathbf{R} \to \mathbf{R}$ is Lipschitz continuous with constant $L > 0$ then we have for all $n \in \mathbf{N}$ and all $y \in \mathbf{L}^2(\mathcal{O})$

$$(A_n^{(\alpha)} G \circ y, G \circ y)^{1/2} \leq L(A_n^{(\alpha)} y, y)^{1/2} + |G(0)| \|A^{\alpha/2}\mathbf{1}\|.$$

(b) If $g, h \in \mathbf{L}^\infty(\mathcal{O})$ then

$$(A_n^{(\alpha)}(gh), gh)^{1/2} \leq \|g\|_\infty (A_n^{(\alpha)} h, h)^{1/2} + \|h\|_\infty (A_n^{(\alpha)} g, g)^{1/2}.$$

Proof. (a) The function $T := (G - G(0))/L$ is a normal contraction in the sense of Ma and Röckner [19]. Hence, the identity $A_n^{(\alpha)} = n(I - nR_n^{(\alpha)})$ and Formula (4.10) in Ma and Röckner [19], Chapter I, imply

$$(A_n^{(\alpha)} T \circ y, T \circ y)^{1/2} = (n(I - nR_n^{(\alpha)}) T \circ y, T \circ y)^{1/2} \leq (n(I - nR_n^{(\alpha)}) y, y)^{1/2} = (A_n^{(\alpha)} y, y)^{1/2}.$$

Positivity of $A_n^{(\alpha)}$ now permits to estimate

$$\begin{aligned} (A_n^{(\alpha)} G \circ y, G \circ y)^{1/2} &= (A_n^{(\alpha)}(LT \circ y + G(0)), LT \circ y + G(0))^{1/2} \\ &\leq L(A_n^{(\alpha)} T \circ y, T \circ y)^{1/2} + (A_n^{(\alpha)} G(0), G(0))^{1/2} \\ &\leq L(A_n^{(\alpha)} y, y)^{1/2} + |G(0)| \|A^{\alpha/2}\mathbf{1}\|. \end{aligned}$$

(b) Let $u := gh$, let $u_1 = \|g\|_\infty h$, and let $u_2 = \|h\|_\infty g$. Then, for all $x, y \in \mathcal{O}$,

$$|u(x) - u(y)| \leq |u_1(x) - u_1(y)| + |u_2(x) - u_2(y)| \quad \text{and} \quad |u(x)| \leq |u_1(x)| + |u_2(x)|.$$

Observing again $A_n^{(\alpha)} = n(I - nR_n^{(\alpha)})$, we obtain the claim from Ma and Röckner [19], Chapter I, Formula (4.13). □

A function s on $\mathcal{O}$ is called A-excessive if $s \geq 0$, $s \in \mathcal{D}(A)$, and $As \geq 0$.

6.5 Theorem Let notation be as specified in Section 6.1, let s be a bounded, strictly positive, A-excessive function, and let $\alpha \geq 0$. Suppose that

(i) F is of the form

$$F(r) = b \cdot r - H(r),$$

where b is a constant and $H : \mathbf{R} \to \mathbf{R}$ is an increasing function vanishing at the origin;

(ii) G is bounded and Lipschitzian with Lipschitz constant L;

(iii) (α) for eventually all n, the operators $nR_n^{(\alpha)}$ induce submarkovian kernels K_n on $\mathcal{O}$ such that $\mu K_n \leq \mu$, where μ is Lebesgue measure on $\mathcal{O}$, cf. Remark 6.6(a);

(β) $tr A^\alpha \tilde{J} Q \tilde{J}^* < \infty$, where $\tilde{J}$ is the embedding $\mathbf{K} \to \mathbf{L}^\infty(\mathcal{O})$ considered as a mapping $\tilde{J} : \mathbf{K} \to L^2(\mathcal{O})$;

(γ) there exists an orthonormal basis (v_l) of $\mathbf{K}$ such that
$C := \sum_l \|J\sqrt{Q} v_l\|_\infty^2 < \infty$;

(δ) $b + L^2 C < \lambda_0$, where λ_0 is the greatest lower bound of the spectrum of A.

Suppose that Eq. (1) possesses a P-a.s. locally bounded mild solution for all initial values in $\mathcal{C}(\bar{\mathcal{O}})$. Let $(\mathcal{P}_t)_t$ be the transition semigroup on $\mathbf{D}$, $\mathcal{C}(\bar{\mathcal{O}}) \subseteq \mathbf{D} \subseteq \mathbf{L}^2(\mathcal{O})$, associated with some generalized solution, and let $\mathbf{X}$ be the metric space $\mathbf{L}^2(\mathcal{O}, s)$. Then we have:

(a) $(\mathcal{P}_t)_t$ has the $\mathcal{C}_{ub}(\mathbf{X})$-Feller property and X has a $\mathcal{C}_{ub}(\mathbf{X})$–stationary distribution γ in $\mathbf{L}^2(\mathcal{O})$; γ is supported by $\mathcal{D}(A^{\frac{1+\alpha}{2}})$. Moreover, there is trend to γ with respect to the narrow topology on $\mathrm{Prob}(\mathbf{X})$.

(b) If $\mathcal{D}(A^{\frac{1+\alpha}{2}}) \subseteq \mathbf{D}$ then γ is a stationary distribution of $(\mathcal{P}_t)_t$ in the usual sense.

(c) If $\alpha \geq [d/2]$ then γ is supported by $\mathcal{C}(\bar{\mathcal{O}})$.

Proof. Let $\mathbf{E} = \mathcal{C}(\bar{\mathcal{O}})$ and let S be the operator multiplication by s on $\mathbf{H} = \mathbf{L}^2(\mathcal{O})$. Plainly, the space $\mathbf{X}$ is the metric space associated with $\mathbf{D}$ and S specified in 5.1(a). Let us verify the assumptions of Corollary 5.5.

In order to verify 5.3(i) we let $n \in \mathbf{N}$ and $y \in \mathbf{E}$ and estimate first, using Lemma 6.3 and (i),

$$(27) \quad (A_n^{(\alpha)}y, f(y)) = (A_n^{(\alpha)}y, by) - (A_n^{(\alpha)}y, H \circ y) \leq b(A_n^{(\alpha)}y, y).$$

Next, we have by Lemma 6.4(b) and (a) for $y \in \mathbf{H}$ and $h \in \mathbf{K}$

$$\begin{aligned}
& (A_n^{(\alpha)}\sigma(y)h, \sigma(y)h)^{1/2} \\
= \; & (A_n^{(\alpha)}(G \circ y\, Jh), G \circ y\, Jh)^{1/2} \\
\leq \; & (A_n^{(\alpha)}G \circ y, G \circ y)^{1/2}\|Jh\|_\infty + (A_n^{(\alpha)}Jh, Jh)^{1/2}\|G \circ y\|_\infty \\
\leq \; & \left(L(A_n^{(\alpha)}y, y)^{1/2} + |G(0)|\|A^{\alpha/2}\mathbf{1}\|\right)\|Jh\|_\infty + (A_n^{(\alpha)}Jh, Jh)^{1/2}\|G \circ y\|_\infty.
\end{aligned}$$

Applying this estimate with $h = \sqrt{Q}v_l$, where (v_l) is as in (iii)(β), we have

$$\begin{aligned}
& \frac{1}{2} tr\, A_n^{(\alpha)}\sigma(y)Q\sigma^*(y) \\
= \; & \frac{1}{2} tr\, \sqrt{Q}\sigma^*(y)A_n^{(\alpha)}\sigma(y)\sqrt{Q} \\
= \; & \frac{1}{2}\sum_l \left(A_n^{(\alpha)}\sigma(y)\sqrt{Q}v_l, \sigma(y)\sqrt{Q}v_l\right) \\
(28) \qquad \leq \; & \left(L(A_n^{(\alpha)}y, y)^{1/2} + |G(0)|\,\|A^{\alpha/2}\mathbf{1}\|\right)^2 \sum_l \|J\sqrt{Q}v_l\|_\infty^2 \\
& + \|G \circ y\|_\infty^2 \sum_l (A_n^{(\alpha)}J\sqrt{Q}v_l, J\sqrt{Q}v_l) \\
\leq \; & 2\left(L^2(A_n^{(\alpha)}y, y) + G(0)^2\|A^{\alpha/2}\mathbf{1}\|\right)\sum_l \|J\sqrt{Q}v_l\|_\infty^2 + \|G \circ y\|_\infty^2 tr A_n^{(\alpha)}\tilde{J}Q\tilde{J}^* \\
= \; & c_T + 2CL^2(A_n^{(\alpha)}y, y)
\end{aligned}$$

with $c_T = 2\,C\,G(0)^2\|A^{\alpha/2}\mathbf{1}\|^2 + \|G\|_\infty^2 \sup_n tr A_n^{(\alpha)}\tilde{J}Q\tilde{J}^* < \infty$ by (iii)(α),(β). Estimates (27) and (28) combine to show

$$(29) \quad (A_n^{(\alpha)}y, f(y)) + \frac{1}{2} tr\, A_n^{(\alpha)}\sigma(y)Q\sigma^*(y) \leq c_T + (b + CL^2)(A_n^{(\alpha)}y, y)$$

for $y \in \mathbf{E}$. Hence, 5.3(i) follows. Moreover, (29) implies for $y \in \mathbf{E}$

$$\begin{aligned}
& \overline{\lim}_n \left[(A_n^{(\alpha)}y, f(y)) + \frac{1}{2} tr\, A_n^{(\alpha)}\sigma(y)Q\sigma^*(y)\right] \\
\leq \; & c_T + (b + CL^2)\lim_n (A_n^{(\alpha)}y, y) \\
= \; & c_T + (b + CL^2)\|A^{\alpha/2}y\|^2.
\end{aligned}$$

Now, $\|A^{\alpha/2}y\|^2 \leq \frac{1}{\lambda_0}\|A^{(1+\alpha)/2}y\|^2$ for all $y \in \mathbf{H}$; thus, we have 5.5(ii) with $c_1 = (b + CL^2)/\lambda_0 < 1$.

Since $\mathcal{L}$ is uniformly elliptic and $\mathcal{O}$ is bounded and smooth, A^{-1} is compact, i.e., 5.5(iii) is satisfied, cf. Renardy and Rogers [22], Lemma 8.20. Since s is A-excessive we also have 5.5(iv).

Finally,

$$(S(y - x), f(y) - f(x)) = b\|\sqrt{s}(y - x)\|^2 - (s(y - x), H \circ y - H \circ x) \leq b\|\sqrt{s}(y - x)\|^2$$

since H is increasing; moreover, we have

$$\begin{aligned}
&\frac{1}{2}\, tr S(\sigma(y)-\sigma(x))Q(\sigma^*(y)-\sigma^*(x)) = \frac{1}{2}\, tr\sqrt{Q}(\sigma(y)-\sigma(x)^*S(\sigma(y)-\sigma(x))\sqrt{Q}\\
= \ & \frac{1}{2}\sum_l \|\sqrt{s}(\sigma(y)-\sigma(x))\sqrt{Q}v_l\|^2 = \frac{1}{2}\sum_l \int_{\mathcal{O}} s(\xi)(G(y(\xi))-G(x(\xi)))^2(J\sqrt{Q}v_l(\xi))^2 d\xi\\
\leq \ & \frac{L^2}{2}\|\sqrt{s}(y-x)\|^2\|\sum_l (J\sqrt{Q}v_l)^2\|_\infty \leq \frac{L^2}{2}C\|\sqrt{s}(y-x)\|^2.
\end{aligned}$$

Hence, 5.3(v) is satisfied with $c := b + \frac{L^2}{2}C - \lambda_0 \quad (< 0)$. This concludes the proof of Part (a) and proves also Part (b).

(c) If $1+\alpha \geq [d/2]+1$ then $\mathcal{D}\left(A^{\frac{1+\alpha}{2}}\right) \subseteq H_{1+\alpha}(\mathcal{O}) \subseteq \mathcal{C}(\bar{\mathcal{O}})$ by Sobolev's lemma. Hence, Part (c) follows. □

6.6 Remarks (a) Assumption (iii)(α) of the previous theorem is satisfied, e.g., in the following three cases

- $\alpha = 0$,
- $\alpha = 1$,
- $0 \leq \alpha \leq 1$ and $A = -\Delta$.

Information on the last two cases can be found, e.g., in Ma and Röckner [19], Chapter II. In particular, the property $\mu K_n \leq \mu$ follows from symmetry of R_n.

(b) In Section 5.6, we pointed out that the quadratic Lyapunov function $y \to (y, A^{-1}y)$ yields a necessary and sufficient condition for the existence of a stationary distribution in the linear case. Unfortunately, Conditions 5.3(i),(ii) and 5.5(ii) are often not satisfied with (the Yosida approximation $(y, A_n^{(-1)}y)$ of) this function and it cannot be used together with the present methods to obtain stationary distributions for equations with nonlinear drift coefficients. The reason is that $nR_n^{(-1)}$ cannot be represented by a kernel and, therefore, Lemma 6.3 fails to be applicable with $\alpha = -1$. In fact,

$$(A^{-1}y, H\circ y) \text{ is not } \geq 0 \text{ in general.}$$

As an example, let $\mathcal{O} =]0,1[$, let $H(r) = r^3$, and let $A = -d^2/dx^2$ so that $A^{-1}y(\xi) = \int_0^1 g(\xi,\eta)y(\eta)d\eta$ with $g(\xi,\eta) = \xi(1-\eta)$, if $\xi \leq \eta$, and $g(\xi,\eta) = (1-\xi)\eta$, otherwise. Define $\sigma_r = 2\delta_{\frac{1}{2}-r} + 3\delta_{\frac{1}{2}} - 4\delta_{\frac{1}{2}+r}$, $0 < r < \frac{1}{2}$, and let $y_{r,k} = k\, 1_{[-\frac{1}{2k},\frac{1}{2k}]} * \sigma_r$. We have $y_{r,k} \to \sigma_r$ and $y_{r,k}^3 \to \tau_r := 8\delta_{\frac{1}{2}-r} + 27\delta_{\frac{1}{2}} - 64\delta_{\frac{1}{2}+r}$ weakly as $k \to \infty$. Hence

$$\begin{aligned}
&\lim_{k\to\infty}(A^{-1}y_{r,k}, y_{r,k}^3)\\
= \ & \lim_{k\to\infty}\int_0^1\int_0^1 g(\xi,\eta)y_{r,k}(\xi)y_{r,k}^3(\eta)d\xi d\eta\\
= \ & \int_0^1\int_0^1 g(\xi,\eta)\sigma_r(d\xi)\tau_r(d\eta)\\
= \ & \int_{\xi\leq\eta}\xi(1-\eta)\sigma_r(d\xi)\tau_r(d\eta) + \int_{\xi>\eta}(1-\xi)\eta\sigma_r(d\xi)\tau_r(d\eta)\\
= \ & 272(\frac{1}{4}-r^2) - 160(\frac{1}{2}-r)^2 - 234\cdot\frac{1}{2}\cdot(\frac{1}{2}-r) + 81(\frac{1}{2})^2\\
\to \ & -41/4 \quad \text{as } r \to 0.
\end{aligned}$$

Therefore, there exists $r > 0$ and $k \in \mathbf{N}$ such that $(A^{-1}y_{r,k}, y_{r,k}^3) < 0$. It is now plain that Conditions 5.3(i),(ii) and 5.5(ii) are not satisfied in this situation.

Acknowledgment. We thank Professor Bohdan Maslowski for making a number of valuable suggestions on the first draft of this paper.

References

[1] Billingsley, P., Convergence of Probability Measures. Wiley, New York-Chichester-Brisbane-Toronto, 1968.

[2] Cerrai, S., Ergodicity for Stochastic Reaction-Diffusion Systems with Polynomial Coefficients. Stochastics and Stochastics Reports 67, 1999, 17–51.

[3] Chow, P.-L., Khasminskii, R.Z., Stationary solutions of nonlinear stochastic evolution equations. Stochastic Analysis and Applications 15, 1997, 671–699.

[4] Da Prato, G., Pardoux, E., Invariant Measures for White Noise Driven Stochastic Partial Differential Equations. Stochastic Analysis and Applications 13, 1995, 295–305.

[5] Da Prato, G., Zabczyk, J., Stochstic Equations in Infinite Dimensions. Cambridge University Press, 1992

[6] Da Prato, G., Zabczyk, J., Ergodicity for Infinite Dimensional Systems. Cambridge University Press, 1996.

[7] Dellacherie, C., Meyer, P.A., Probabilities and Potential. North Holland, Amsterdam-New York-Oxford, 1978.

[8] Gardner, R.J., The regularity of Borel measures. Springer, Berlin, Lecture Notes in Math. 945, 1982, 42–100.

[9] Gardner, R.J., Pfeffer, W.F., Conditions that imply a space is Radon. Springer, Berlin, Lecture Notes in Math. 1089, 1984, 11–22.

[10] Gątarek, D., Gołdys, B., On invariant measures for diffusions on Banach spaces. Potential Analysis 7, 1997, 539–553.

[11] Gillman, L., Jerison, M., Rings of Continuous Functions. Springer, New York-Heidelberg-Berlin, 1976.

[12] Ichikawa, A., Semilinear stochastic evolution equations: boundedness, stability and invariant measures. Stochastics 12, 1984, 1–39.

[13] Khasminskii, R.Z., Stochastic Stability of Differential Equations. Sijthoff and Noordhoff, Alphen aan den Rijn-Rockville, 1980.

[14] Leha, G., Maslowski, B., Ritter, G., Stability of solutions to semilinear stochastic evolution equations. Stochastic Analysis and Applications 17, 1999, 1009–1051.

[15] Leha, G., Ritter, G., On diffusion processes and their semigroups in Hilbert spaces with an application to interacting stochastic systems. Ann. Prob. 12, 1984, 1077–1112.

[16] Leha, G., Ritter, G., On solutions to stochastic differential equations with discontinuous drift in Hilbert space. Math. Ann. 270, 1985, 109–123.

[17] Leha, G., Ritter, G., Lyapunov-type conditions for stationary distributions of diffusion processes on Hilbert spaces. Stochastics and Stochastic Reports 48, 1994, 195–225.

[18] Leha, G., Ritter, G., Wakolbinger, A., An improved Lyapunov-function approach to the behavior of diffusion processes in Hilbert spaces. Stochastic Analysis and Applications 15, 1997, 59–89.

[19] Ma, Z.-M., Röckner, M., Dirichlet Forms. Springer, Berlin–Heidelberg, 1992.

[20] Maslowski, B., Uniqueness and stability of invariant measures for stochastic differential equations in Hilbert spaces. Stochastics and Stochastics Reports 28, 1989, 85–114.

[21] Peszat, S., Existence and uniqueness of the solution for stochastic equations on Banach spaces. Stochastics and Stochastics Reports 55, 1995, 167–193.

[22] Renardy, M., Rogers, R.C., An Introduction to Partial Differential Equations. Springer, New York, 1993.

[23] Schwartz, L., Radon Measures on Arbitrary Topological Spaces and Cylindrical Measures. Oxford University Press, 1973.

[24] Topsøe, F., A criterion for weak convergence of measures with an application to measures on $D[0,1]$. Math. Scand., 25, 97–104, 1969.

[25] Topsøe, F., Compactness in spaces of measures. Studia Math. 36, 1970, 195–212.

[26] Topsøe, F., Topology and Measures. Springer, Berlin, Lecture Notes in Math. 133, 1970.

[27] Weidmann, J., Lineare Operatoren in Hilberträumen. Teubner, Stuttgart, 1976.

[28] Zabczyk, J., Structural properties and limit behaviour of linear stochastic systems in Hilbert spaces. Math. Control Theory, Banach Center Publications, Vol. 14, Polish Scientific Publishers, 1985.