Technische Universität Ilmenau
Institut für Mathematik

Preprint No. M 06/08

Measure Attractors and Markov Attractors

Crauel, Hans

2006

Impressum:
Hrsg.: Leiter des Instituts für Mathematik
Weimarer Straße 25
98693 Ilmenau
Tel.: +49 3677 69 3621
Fax: +49 3677 69 3270
http://www.tu-ilmenau.de/ifm/

ISSN xxxx-xxxx

ilmedia

Measure Attractors and Markov Attractors

Hans Crauel
Institut für Mathematik
Technische Universität Ilmenau
Weimarer Straße 25
98693 Ilmenau

hans.crauel@tu-ilmenau.de

The actions induced by a random dynamical system on spaces of probability measures on the state space are investigated, and generalisations of the notion of an attractor are discussed and compared. For the particular case of a random dynamical system generated by a stochastic differential equation the notion of an attractor for the associated Markov semigroup had previously been discussed in several instances in the literature. It is re-discovered here as a special case of a more general notion of an attractor in the space of measures.

2000 Mathematics Subject Classification Primary 37B25 Secondary 37A50 37H99 37L30 37L40 37L55 60G57 60H10 60H15 93E03

1 Introduction

The notion of an attractor is one of the basic concepts of the theory of dynamical systems. Whilst for deterministic systems this notion has been of importance since several decades, it is only during the last couple of years that attractors for stochastic systems have been investigated. The crucial obstacle came from the fact that the classical approach, using the Markov property of individual solutions to define the Markov semigroup and an associated generator, could not deal with the joint motions of two or more points. This began to change after the introduction of stochastic flows, going back to Kunita and to Elworthy, and the introduction of the notion of random dynamical systems. All approaches to random attractors use the theory of random dynamical systems (RDS), starting with Crauel and Flandoli [13], Schmalfuß [20], and Schenk-Hoppé [19].

There has been, however, another notion of an attractor for a stochastic differential equation (SDE), going back to Morimoto [18]. It uses the classical notions of deterministic systems for the induced Markov semigroup on the space of probability measures. This approach, which works for SDE as well as for products of independent identically distributed (i.i.d.) maps, has been pursued by Schmalfuß [21], who showed that this concept gives results for a certain class of SDE on a separable Banach space. Provided the associated RDS has a random attractor, one may define a compact set of deterministic probability measures on the state space of the system. This set can be shown to be an attractor for the action of the Markov semigroup on a certain class of sets of probability measures. This attractor and its relations to other concepts of measure attractors are discussed in the present paper.

An approach initiated by Capiński and Cutland [5] and pursued by Cutland and Keisler [14] investigates, roughly speaking, measure attractors for 3D Navier-Stokes equations, where no RDS – not even unique solutions for the SDE – are available.

What attractors for RDS are concerned, we will distinguish between two notions, which are the *global set attractor* and the *global point attractor*. We mention that further classes of attractors for RDS have been investigated, see, e. g., Ashwin and Ochs [2]. A global point attractor is required to attract every individual point, whereas the global set attractor is assumed to attract every compact (or every bounded) set. Set attractors sometimes are also addressed to as *uniform* attractors. In between these two extremes of global attractors there is a large variety of further attractors for classes of sets in between finite and compact sets. Although this will become important also here, when attraction of the class of Markov measures is considered, we will not go into detail in this respect and refer to Crauel [10] instead.

For deterministic systems it is not a very severe restriction to consider global attractors, since every local attractor is a global attractor for the restriction of the system to a suitable open subset of the state space. For random systems it is more difficult to restrict to local attractors, since there localisation will have to be random in general. In fact, typically for random systems there is no bounded deterministic set which is invariant for the system. See Crauel et al. [12] for the related discussion of a concept for a Morse theory for RDS.

We start by considering the action of an RDS on the space of probability measures of the state space, which gives a new RDS on the space of measures. For this RDS the common notions of random point and set attractors are available. We obtain what one might have expected already: The (random) set of all probability measures supported by a random set attractor is the set attractor for the action of the RDS on measures. This may be considered not particularly exciting. Furthermore, this setup does not allow access to what we really are interested in, which are *invariant measures*. Invariant measures for an RDS are elements of the set of *random probability measures* on the state space.

Therefore we consider the action induced by an RDS on the space of random probability measures. We do not obtain an RDS acting on the space of random measures, but rather a dynamical system, given by the skew product flow induced by the RDS. We again obtain what might have been expected: The set of all random probability measures supported by the set attractor for the RDS is the set attractor for the action of the skew product flow on random measures.
The concepts of measure attractors sketched above look very similar. Their difference is like the difference between the tangent bundle and the set of all vector fields of a manifold, where differentiability is replaced by measurability here.
Being interested in invariant measures, we abandon the first possibility of a measure attractor, and concentrate on the notion given by the action of the RDS on random measures. It is known that the set of all invariant measures is a compact invariant subset of the measure attractor. However, in general the set of invariant measures itself is not attracting.

On first view the notion of a measure attractor as introduced by Morimoto and by Schmalfuß is yet another concept. It is restricted to systems induced by stochastic differential equations (or by products of independent identically distributed (i.i.d.) maps), using the Markov semigroup induced by this type of systems. We will show that it suffices to assume a *white noise* RDS to define a Markov semigroup. For white noise systems the distinction between a point and the set attractor is crucial, since a point attractor always supports all invariant Markov measures, but not necessarily all invariant measures, which are supported by the set attractor only. See Crauel [10] for details.

We proceed by investigating the relations between measure attractors for RDS and the Schmalfuß attractor in more detail. It will become immediate that the Schmalfuß attractor is nothing but the "expectation of the Markov measure attractor". So everything said about the Markov measure attractor applies to the Schmalfuß attractor. One consequence, applying to general white noise RDS on a Polish space already, is that the (convex and closed) set of all invariant probability measures for the Markov semigroup always is a subset of the Schmalfuß attractor. But, in general, the set of invariant measures for the Markov semigroup is a proper subset of the Schmalfuß attractor, and it is *not* an attractor, not even a local one. In fact, it can be shown that in general there are measures arbitrarily close to the set of invariant Markov measures, and which are not attracted by the Schmalfuß attractor under the action of the Markov semigroup. These measures can even be chosen to be elements of the Schmalfuß attractor.

The work of Capiński and Cutland [5] and of Cutland and Keisler [14] on measure attractors for 3D Navier-Stokes equations deals with a situation where no RDS – not even unique solutions for the SDE – have to exist. The present contribution investigates a different situation, simpler in a certain sense, insofar existence of an RDS is always assumed, and the actions induced on the spaces of deterministic and random measures on the state space are investigated.

2 Basic Notions for Random Dynamical Systems

We will give here only the basic notions as needed for the purposes of the present paper. See Arnold [1] for a more comprehensive account of the fundamentals of random dynamical systems.

Suppose that $(\Omega, \mathscr{F}, P)$ is a probability space, and

$$\begin{aligned} \vartheta : \mathbb{R} \times \Omega &\rightarrow \Omega \\ (t, \omega) &\mapsto \vartheta_t \omega \end{aligned}$$

is a measurable map, such that $\vartheta_t : \Omega \rightarrow \Omega$ preserves P, and such that $\vartheta_{t+s} = \vartheta_t \circ \vartheta_s$ for all $s, t \in \mathbb{R}$. Thus (ϑ_t) is a classical measurable dynamical system on $(\Omega, \mathscr{F}, P)$.

2.1 Definition Given $(\Omega, \mathscr{F}, P)$ and (ϑ_t) as above, X a Polish space, and T either $\mathbb{R}$, $\mathbb{R}^+$, $\mathbb{Z}$, or $\mathbb{N}$. A measurable map

$$\begin{aligned} \varphi : T \times X \times \Omega &\rightarrow X \\ (t, x, \omega) &\mapsto \varphi(t, \omega)x \end{aligned}$$

is a *random dynamical system (RDS)*, if

(i) $\varphi(t, \omega) : X \rightarrow X$ is continuous P-almost surely for every $t \in T$

(ii) $\varphi(t+s, \omega) = \varphi(t, \vartheta_s\omega) \circ \varphi(s, \omega)$ for all $s, t \in T$, and $\varphi(0, \omega) = \mathrm{id}$, for P-almost all ω

Note that we do not have to assume continuity in t here.
An RDS is said to be two-sided if T is two-sided. For a two-sided RDS φ the maps $\varphi(t, \omega)$ are invertible, and $\varphi(t, \omega)^{-1} = \varphi(-t, \vartheta_t\omega)$ a. s.

An RDS φ induces a (semi-) flow $\Theta_t : X \times \Omega \rightarrow X \times \Omega$ on the product space, given by $(x, \omega) \mapsto \big(\varphi(t, \omega)x, \vartheta_t\omega\big)$ for $t \in T$. The (semi-) flow $(\Theta_t)_{t \in T}$ is a skew product (semi-) flow.

Every RDS defines a canonical (two-parameter) *filtration* on $(\Omega, \mathscr{F}, P)$ by

$$\mathscr{F}_s^t = \sigma\big\{\omega \mapsto \varphi(\tau_1, \vartheta_{s+\tau_2}\omega)x : \tau_1, \tau_2 \geq 0,\ \tau_1 + \tau_2 \leq t - s,\ x \in X\big\} \tag{1}$$

for $s, t \in [-\infty, \infty]$, $s \leq t$.

2.2 Lemma *The family of σ-algebras $\mathscr{F}_s^t$, $s \leq t$, satisfies*

(i) *for all $s, t, u \in \mathbb{R}$, $s \leq t$,*

$$\vartheta_u^{-1}\mathscr{F}_s^t = \mathscr{F}_{s+u}^{t+u};$$

(ii) *for all $s, t, u, v \in \mathbb{R}$ with $u \leq s \leq t \leq v$*

$$\mathscr{F}_s^t \subset \mathscr{F}_u^v.$$

The proof is straightforward.

The rather technical definition of $\mathscr{F}_s^t$ is needed for Lemma 2.2 (ii). For a two-sided time RDS (1) simplifies to

$$\mathscr{F}_s^t = \sigma\big\{\omega \mapsto \varphi(\tau, \vartheta_s\omega)x : 0 \le \tau \le t-s,\ x \in X\big\}.$$

We further put

$$\begin{aligned} \mathscr{F}^- &= \sigma\big\{\omega \mapsto \varphi(s, \vartheta_{-t}\omega)x : x \in X,\ 0 \le s \le t\big\}, \\ \mathscr{F}^+ &= \sigma\big\{\omega \mapsto \varphi(t, \vartheta_s\omega)x : x \in X,\ 0 \le s, t\big\}, \end{aligned} \tag{2}$$

addressing these as the *past* and the *future* of φ, respectively. Clearly $\mathscr{F}^- = \mathscr{F}_{-\infty}^0 = \sigma\{\mathscr{F}_s^0 : s \le 0\}$, compare (1). Note that $\mathscr{F}_s^t = \vartheta_t^{-1}\mathscr{F}^- \cap \vartheta_s^{-1}\mathscr{F}^+$, compare Arnold [1] Definition 2.3.4 and Remark 2.3.5.

A *closed random set* is a set-valued map $C : \Omega \to 2^X$, the set of all subsets of X, such that $C(\omega)$ is closed P-a. s., and such that, for every open $U \subset X$, we have $\{\omega : C(\omega) \cap U \neq \emptyset\} \in \mathscr{F}$. For several equivalent conditions see, e. g., Castaing and Valadier [6] or Crauel [11]. An *open random set* is a set-valued map $U : \Omega \to 2^X$ such that $\omega \mapsto U^c(\omega)$ defines a closed random set, where B^c denotes the complement of a set $B \subset X$. If C is a closed random set, a random variable $c : \Omega \to X$ is said to be a *selection of* C if $c(\omega) \subset C(\omega)$ for P-almost all $\omega \in \Omega$. Given a sub-σ-algebra $\mathscr{G} \subset \mathscr{F}$, a $\mathscr{G}$*-selection* is a $\mathscr{G}$-measurable selection.

If φ is an RDS, then a closed random set $D : \Omega \to 2^X$ is said to be *invariant* or *strictly invariant*, respectively, with respect to φ, if $\varphi(t, \omega)D(\omega) \subset D(\vartheta_t\omega)$ or $\varphi(t, \omega)D(\omega) = D(\vartheta_t\omega)$ P-a. s., respectively, for every $t \ge 0$.

The set of all Borel probability measures on X is denoted by $Pr(X)$. It is understood to be equipped with the topology of weak convergence, which makes it a Polish space. Measurability of random variables taking values in $Pr(X)$ is understood with respect to the Borel σ-algebra induced by this topology. A measure $\rho \in Pr(X)$ is said to be a *Dirac measure* if $\rho(\{x\}) = 1$ for some $x \in X$; it is denoted by δ_x then.
A *random probability measure* on a Polish space X is a measurable map $\mu : \Omega \to Pr(X)$, $\omega \mapsto \mu_\omega$, where two such maps are identified if they coincide for P-almost all $\omega \in \Omega$. The space of all random probability measures is denoted by $Pr_\Omega(X)$. We will henceforth be concerned with probability measures only, so "random measure" will always stand for "random probability measure" in the following.

A random measure μ induces a measure on the product space by first assigning to a Borel set $B \subset X$ and a measurable set $F \in \mathscr{F}$ the number $\int_\Omega \mu_\omega(F)\, dP(\omega)$, and then extending it to a measure on the product σ-algebra of $X \times \Omega$. The random variable $\omega \mapsto \mu_\omega$ and the measure on the product space are in one-to-one correspondence, so the latter will be denoted by μ again. In particular, $Pr_\Omega(X)$ can be considered as the subspace of $Pr(X \times \Omega)$ consisting of all measures on the product space with marginal P on the second component.

Suppose that $f : X \times \Omega \to \mathbb{R}$ is a function, such that $x \mapsto f(x, \omega)$ is continuous for P-almost all $\omega \in \Omega$, $\omega \mapsto f(x, \omega)$ is measurable for all $x \in X$, and $\omega \mapsto \sup\{|f(x, \omega)| : x \in X\}$ is integrable with respect to P. Then f is said to be a *random continuous function*. For a random continuous function f and a random measure μ, the integral $\mu(f) = \int_\Omega \int_X f(x, \omega)\, d\mu_\omega(x)\, dP(\omega)$ is well defined. The smallest topology on the space of random measures such that $\mu \mapsto \mu(f)$ is continuous for every random continuous function f is said to be the *narrow topology* or the *topology of weak convergence* on the space of random measures. See Crauel [11] for a more detailed study of these topics.

Given an RDS φ, a random measure μ is said to be an *invariant measure* for φ if $\Theta_t \mu = \mu$ for all $t \in T$, or, equivalently, if

$$\varphi(t, \omega)\mu_\omega = \mu_{\vartheta_t \omega}$$

P-almost surely for every $t \in T$; here $\varphi\mu$ denotes the image measure of a measure μ under a measurable map φ, $\varphi\mu(B) = \mu\big(\varphi^{-1}(B)\big)$.
A random measure μ is said to be *supported by a closed random set* C, if $\mu_\omega(C(\omega)) = 1$ a. s.
A random measure μ is said to be a *Markov measure* (with respect to φ), if $\omega \mapsto \mu_\omega$ is measurable with respect to $\mathscr{F}^-$ (the past of φ, see (2)). Markov measures are of particular interest for several reasons. They provide the link between the notion of an invariant measure in the theory of random dynamical systems and the notion of an invariant measure in the classical approach, which uses the Markov semigroup induced by a stochastic differential equation (SDE). The essential relation in this respect is given in Proposition 4.2 below. Furthermore, several dynamical properties of an RDS — related to Lyapunov exponents, stability, hyperbolicity, attractors — can be classified in terms of the Markov property of invariant measures. For further characterisations of random measures in the context of RDS see Arnold [1] Chapter 1 Sections 4–8, particularly Section 7 for Markov measures, and Crauel [11] and references therein.

2.3 Definition Suppose that φ is an RDS on a Polish space X, and let d be a complete metric on X. A closed random set $\omega \mapsto A(\omega)$ is said to be a *global point* or a *global set attractor*, respectively, if

(i) $A(\omega)$ is compact P-a. s.

(ii) A is strictly invariant

(iii) A attracts every finite $B \subset X$ for the point attractor, or every compact $B \subset X$ for the set attractor, respectively, in the sense that

$$\lim_{t\to\infty} d\big(\varphi(t, \vartheta_{-t}\omega)B, A(\omega)\big) = 0 \qquad P\text{-almost surely.} \tag{3}$$

Here $d(B, A) = \sup_{b \in B} d(b, A)$ for $A, B \subset X$ arbitrary (which defines, on compact sets, the *Hausdorff semi-metric*); we have $d(B, A) = 0$ if, and only if, $B \subset \bar{A}$, $\bar{A}$ denoting the closure of A in X).

2.4 Remark (i) Clearly for A to be a point attractor it suffices to have

$$\lim_{t\to\infty} d\big(\varphi(t,\vartheta_{-t}\omega)x, A(\omega)\big) = 0 \qquad P\text{-a.s.}$$

for every $x \in X$.

(ii) The definition of attractors for an RDS φ on a Polish space X makes use of a choice of a metric d on X, entering into (3). However, (3) is easily seen to be independent of the choice of a metric, using compactness of A. Thus, *existence* of a point or a set attractor, respectively, is a property of an RDS which is independent of the choice of a metric on X.

(iii) Often a set attractor is supposed to attract all *bounded* sets instead of just all compact ones. For this version of the definition the choice of the metric does affect the class of sets which are supposed to be attracted. Therefore there may be choices of the metric for which a compact random set $\omega \mapsto A(\omega)$ is a global set attractor for all bounded sets, and other choices where the same A does not attract all sets bounded in the other metric. The most obvious case where the second situation may occur would be to consider a bounded metric, i. e., a metric such that the whole state space X is bounded. Compare also Remark 2.6 (ii) below.

Recall the notion of an Ω-limit set: For $B \subset X$ put

$$\Omega_B(\omega) = \bigcap_{t\geq 0} \overline{\bigcup_{\tau\geq t} \varphi(\tau,\vartheta_{-\tau}\omega)B}$$

For arbitrary $B \subset X$ the Ω-limit set Ω_B is invariant (Crauel [9], Lemma 5.1; in fact, this is true for Ω-limit sets of random sets).

The following theorem on the existence of attractors has appeared in several variations in the literature during the last years. We refrain from trying to give a comprehensive list of references.

2.5 Theorem *Suppose that φ is an RDS on a Polish space X, and let d be a complete metric. Then there exists a global set or global point attractor, respectively, if, and only if, there exists a globally attracting compact random set $\omega \mapsto K(\omega)$, i. e., for every $B \subset X$ compact or finite, respectively,*

$$\lim_{t\to\infty} d\big(\varphi(t,\vartheta_{-t}\omega)B, K(\omega)\big) = 0 \qquad P\text{-almost surely.}$$

In this case the Ω-limit set Ω_B is strictly invariant for every B compact or finite, respectively. A global set or point attractor, respectively, is given by

$$A(\omega) = \overline{\bigcup_{B} \Omega_B(\omega)} \qquad P\text{-a.s.} \tag{4}$$

with B ranging over the compact or finite sets, respectively. Furthermore,

(i) *both the global set and the global point attractor given by* (4) *are* $\mathscr{F}^-$*-measurable.*

(ii) *The global set attractor is unique almost surely.*

(iii) *The global set attractor satisfies* $A(\omega) = \Omega_B(\omega)$ *for every compact* B *for which* $P\{A(\omega) \subset B\} > 0$*, and there always exist compact deterministic* $K \subset X$ *with* $P\{A(\omega) \subset K\} > 0$*, for which therefore* $A = \Omega_K(\omega)$ *almost surely.*

For a proof (of a more general version) see Crauel [10].

2.6 Remark (i) Point attractors are not unique in general, but there always exists a unique minimal one, which is given by (4).

(ii) The global set attractor is unique almost surely. Furthermore, it is uniquely determined by attracting compact (deterministic) sets already (see Crauel [9] Section 5). This means, in particular, that any attractor for a larger class of sets automatically coincides with the attractor for compact deterministic sets $\omega \mapsto A(\omega)$ given by Theorem 2.5. Thus, in particular, a compact strictly invariant random set which attracts all bounded deterministic sets with respect to some choice of the metric must coincide with A already. Furthermore, also a compact invariant set attracting all compact or all bounded *random* sets cannot give anything new, but it is equal to A a priori.
It is an *additional* feature of an attractor for compact deterministic sets to attract deterministic or random sets which are bounded with respect to some choice of a metric. The attractor remains the same. It has not to be enlarged in order to attract these (substantially bigger) classes of sets.
For a discussion of attractors for more general classes of sets instead of just finite, compact or bounded ones see Crauel [10].

2.7 Remark Clearly, for the definition of a point attractor it suffices to consider attraction or Ω-limit sets, resp., of single points only instead of finite sets. From this observation one obtains immediately that also the Ω-limit set of an X-valued random variable, which takes only finitely many values P-a. s., is a subset of the point attractor. In fact, if x is such a random variable then $\Omega_x(\omega) \subset \Omega_B(\omega)$ P-a. s., where $B \subset X$ is any (finite) set comprising the values x takes. On the other hand, viewing a single point as a constant random variable, we may conclude that the minimal point attractor is given by $\bigcup_x \Omega_x(\omega)$, where the union is taken over all finite-valued random variables x.

2.8 Definition An RDS φ is said to be *white noise* if the past $\mathscr{F}^-$ of φ is independent of the future $\mathscr{F}^+$ (see (2) for the definition).

The most prominent example for white noise RDS are RDS induced by stochastic differential equations, see Arnold [1], Section 2.3, and RDS given by products of independent (identically distributed) maps.
It should be noted that deterministic systems are automatically white noise systems with respect to the present notion, since the past and the future of deterministic systems are both trivial.

The relation between point and set attractors and invariant (Markov-) measures for φ is given by the next theorem.

2.9 Theorem (i) *Suppose that φ is an* arbitrary *RDS on a Polish space which has a global set attractor (which only has to be assumed to attract every compact deterministic set here). Then every invariant measure for φ is supported by A, i. e., $\mu_\omega(A(\omega)) = 1$ almost surely for every invariant μ.*

(ii) *Suppose that φ is a* white noise *RDS on a Polish space which has a global point attractor. Then the minimal global point attractor supports every invariant Markov measure.*

See Crauel [9] Corollary 4.4 for the proof of (i) and [10] Theorem 4.3 for the proof of (ii).

3 Measure Attractors

Starting from an RDS on a Polish space X we discuss two different notions of measure attractors. The first one is concerned with the action induced by the RDS on the space $Pr(X)$ of deterministic probability measures on X. This defines an RDS on $Pr(X)$, which is a Polish space again. Secondly, we consider the action induced by the associated skew product flow on the space $Pr_\Omega(X)$ of random probability measures on X. This is a deterministic dynamical system on a rather abstract state space.
We will encounter several (semi-) metrics on different spaces, which will often be denoted by the same letter d, specified in the corresponding context.

3.1 Measure Attractors on $Pr(X)$

Suppose that φ is an RDS on a Polish space X. Then φ induces an RDS on the Polish space $Pr(X)$, denoted by Φ, by

$$\begin{aligned} \Phi : T \times Pr(X) \times \Omega &\rightarrow Pr(X) \\ (t, \rho, \omega) &\mapsto \varphi(t, \omega)\rho, \end{aligned}$$

where $\varphi(t, \omega)\rho$ is the image measure of the deterministic measure ρ under the random map $\varphi(t, \omega)$.

Using the fact that the action induced by the continuous $\varphi(t, \omega) : X \rightarrow X$ on $Pr(X)$ is continuous as well, it is straightforward to verify the conditions of Definition 2.1 for Φ on $Pr(X)$.

According to Definition 2.3, a compact random set $\omega \mapsto \mathcal{A}(\omega)$ in $Pr(X)$ is a random set attractor or a random point attractor, respectively, for Φ on $Pr(X)$ if it is strictly invariant with respect to Φ, and if it attracts every finite or compact subset $\mathcal{B} \subset Pr(X)$, respectively, in the sense of (3). The Hausdorff semi-metric d invoked in (3) is induced by (any choice of) a complete metric on $Pr(X)$ metrising the topology of weak convergence.

We will make use of the following lemmas.

3.1 Lemma *Suppose that $\psi : Y \to Z$ is a continuous surjective map from a compact metrisable Y to a metrisable Z. Then the map induced by ψ on $Pr(Y)$, denoted again by $\psi : Pr(Y) \to Pr(Z)$ and given by $\psi\gamma(B) = \gamma(\psi^{-1}B)$ for $\gamma \in Pr(Y)$, B a Borel set in Z, is continuous and surjective. In particular, $\psi(Pr(Y)) = Pr(Z)$.*

PROOF Continuity of ψ is immediate. In order to obtain surjectivity, first note that every Dirac measure δ_z for $z \in Z$ is the image of an arbitrary Borel probability measure supported by the closed set $\psi^{-1}(z) \subset Y$, e. g., of a Dirac measure in some $y \in Y$ with $\psi(y) = z$. The set of all probability measures on Z being the closure of the set of all convex combinations of Dirac measures, the assertion follows in view of the fact that ψ is affine (and continuous) on $Pr(Y)$. □

3.2 Lemma *Suppose that $\omega \mapsto C(\omega)$ is a closed random set in a Polish space X. Then $\omega \mapsto Pr\big(C(\omega)\big)$, given by*

$$Pr(C(\omega)) = \big\{\rho \in Pr(X) : \rho\big(C(\omega)\big) = 1\big\},$$

is a closed random set in $Pr(X)$.

PROOF We use the fact that, for a Polish space Y and a dense set $D \subset Y$, one has

$$Pr(Y) = \mathrm{cl}\,\Big(\mathrm{conv}\,\{\delta_p : p \in D\}\Big),$$

where "conv" means the convex hull, and "cl" the closure (in the topology of weak convergence). Denoting $\mathscr{D} = \mathrm{cl}\,\mathrm{conv}\,\{\delta_p : p \in D\}$, this follows by observing first that for every $y \in Y$ one has $\delta_y \in \mathrm{cl}\,\{\delta_p : p \in D\}$, and hence $\delta_y \in \mathscr{D}$. Now the closure of a convex set is convex again, which implies, in particular, that $\mathscr{D}$ is convex. Therefore $\{\delta_y : y \in Y\} \subset \mathscr{D}$ gives $\mathrm{conv}\,\{\delta_y : y \in Y\} \subset \mathscr{D}$, and thus $\mathrm{cl}\,\mathrm{conv}\,\{\delta_y : y \in Y\} \subset \mathscr{D}$. The assertion follows in view of $Pr(Y) = \mathrm{cl}\,\mathrm{conv}\,\{\delta_y : y \in Y\}$.
In order to prove measurability of $\omega \mapsto Pr\big(C(\omega)\big)$ we will show that $\omega \mapsto d\big(\rho, Pr(C(\omega))\big)$ is measurable for every $\rho \in Pr(X)$, where d is any metric metrising the topology of weak convergence on $Pr(X)$. The selection theorem (Castaing and Valadier [6] Theorem III.9) gives existence of a countable family $\{\omega \mapsto a_n(\omega) : n \in \mathbb{N}\}$ of measurable selections of A which is dense in A. Therefore $Pr\big(C(\omega)\big) = \mathrm{cl}\,\mathrm{conv}\,\{\delta_{a_n(\omega)} : n \in \mathbb{N}\}$. Since $\omega \mapsto d(\rho, \sum_{k=1}^{n} r_k \delta_{a_k(\omega)})$ is measurable for every non-negative real numbers r_k summing up to one one gets measurability of $\omega \mapsto d(\rho, \mathrm{conv}\,\{\delta_{a_n(\omega)} : n \in \mathbb{N}\})$ by noting that it is determined by taking convex combinations with rational coefficients only. Now the assertion follows, since $d(\rho, D) = d(\rho, \mathrm{cl}\,D)$ for any set $D \subset Pr(X)$. □

3.3 Lemma *Suppose that φ is an RDS on a Polish space X with a set attractor $\omega \mapsto A(\omega)$. Then for every compact set $\Gamma \subset Pr(X)$ of probability measures on X and for*

P-almost every $\omega \in \Omega$ the following holds. For every $\varepsilon > 0$ and $\delta > 0$ there exists $T = T(\Gamma, \varepsilon, \delta, \omega)$, such that

$$\varphi(t, \vartheta_{-t}\omega)\gamma\big(U_\delta\big(A(\omega)\big)\big) \geq 1 - \varepsilon$$

for all $t \geq T$ and $\gamma \in \Gamma$. Here $U_\delta(B) = \{x \in X : d(x, B) < \delta\}$ denotes the δ-neighbourhood of $B \subset X$.

Proof By the Prohorov theorem (see, e. g., Ethier and Kurtz [15] Theorem 2.2, pp. 104–105), for every $\varepsilon > 0$ there exists $K \subset X$ compact with $\gamma(K) \geq 1 - \varepsilon$ for every $\gamma \in \Gamma$. Since A is a set attractor, for P-a. e. ω there exists $T = T(K, \delta, \omega)$ such that $d\big(\varphi(t, \vartheta_{-t}\omega)K, A(\omega)\big) \leq \delta$, i. e.,

$$\varphi(t, \vartheta_{-t}\omega)K \subset U_\delta\big(A(\omega)\big). \tag{5}$$

Consequently, for every $\gamma \in \Gamma$,

$$\begin{aligned} \varphi(t, \vartheta_{-t}\omega)\gamma\big[U_\delta(A(\omega))\big] &\geq \varphi(t, \vartheta_{-t}\omega)\gamma\big[\varphi(t, \vartheta_{-t}\omega)K\big] \\ &= \gamma\big[\varphi(t, \vartheta_{-t}\omega)^{-1}\big(\varphi(t, \vartheta_{-t}\omega)K\big)\big] \\ &\geq \gamma(K) \;\geq\; 1 - \varepsilon \end{aligned}$$

invoking (5) in the first inequality. □

3.4 Theorem *Suppose that φ is an RDS on a Polish space X.*

(i) *If A is a set attractor for φ, then $\omega \mapsto Pr\big(A(\omega)\big)$, given by*

$$Pr\big(A(\omega)\big) = \{\rho \in Pr(X) : \rho(A(\omega)) = 1\},$$

is a (and thus the*) set attractor for the induced RDS Φ on $Pr(X)$.*

(ii) *For every point attractor $A_{\rm pt}$ for φ also $\omega \mapsto Pr\big(A_{\rm pt}(\omega)\big)$, given by*

$$Pr\big(A_{\rm pt}(\omega)\big) = \{\rho \in Pr(X) : \rho(A_{\rm pt}(\omega)) = 1\},$$

is a point attractor for the induced RDS Φ on $Pr(X)$.

(iii) *Denoting the minimal point attractor for Φ on $Pr(X)$ by $\omega \mapsto \mathcal{A}_{\rm pt}(\omega)$, and denoting for a moment by $\mathrm{Dirac}(B) = \{\delta_x : x \in B\}$ the set of Dirac measures supported by $B \subset X$, it holds that*

$$\mathrm{Dirac}\big(A_{\rm pt}^{\min}(\omega)\big) \subset \mathcal{A}_{\rm pt}(\omega) \subset Pr\big(A_{\rm pt}^{\min}(\omega)\big) \qquad P\text{-a. s.},$$

where $A_{\rm pt}^{\min}$ denotes the minimal point attractor for φ. In particular, the set of Dirac measures supported by $A_{\rm pt}^{\min}$ always is a subset of the minimal point attractor for Φ.

Proof By Lemma 3.2 both $\omega \mapsto Pr\big(A(\omega)\big)$ and $\omega \mapsto Pr\big(A_{\mathrm{pt}}(\omega)\big)$ are closed random sets in $Pr(X)$. Furthermore, the set of probability measures supported by a compact set in the state space being compact in the space of probability measures we get that both $Pr(A)$ and $Pr(A_{\mathrm{pt}})$ are compact a. s.
Strict Φ-invariance of $\omega \mapsto Pr\big(A(\omega)\big)$ and of $\omega \mapsto Pr\big(A_{\mathrm{pt}}(\omega)\big)$ follows from strict φ-invariance of A and A_{pt}, respectively, using Lemma 3.1 to conclude that the Φ-invariance is strict as well.

In order to conclude attraction of $Pr(A)$ in (i), let d be any metric on $Pr(X)$ metrising the topology of weak convergence. Then Lemma 3.3 implies that for any compact $\Gamma \subset Pr(X)$ and for any $\varepsilon > 0$ there exists $T = T(\varepsilon, \omega)$, such that

$$d\big(\varphi(t, \vartheta_{-t}\omega)\Gamma, Pr(A(\omega))\big) < \varepsilon \qquad \text{for every } t \geq T,$$

where now, by misuse of notation, d denotes the Hausdorff semi-metric on compact subsets of $Pr(X)$ induced by the metric d on $Pr(X)$.

Turning to (ii), we note that A_{pt} being a point attractor for φ implies

$$\big\{x \in X : \lim_{t\to\infty} d\big(\varphi(t, \vartheta_{-t}\omega)x, A_{\mathrm{pt}}(\omega)\big) = 0\big\} = X \qquad P\text{-a. s.}$$

Therefore for any $\rho \in Pr(X)$

$$\rho\Big(\big\{x \in X : \lim_{t\to\infty} d\big(\varphi(t, \vartheta_{-t}\omega)x, A_{\mathrm{pt}}(\omega)\big) = 0\big\}\Big) = 1 \qquad P\text{-a. s. for every } \varepsilon > 0,$$

which implies, in particular,

$$\lim_{t\to\infty} \rho\big\{x : \varphi(t, \vartheta_{-t}\omega)x \in U_\varepsilon\big(A_{\mathrm{pt}}(\omega)\big)\big\} = 1 \qquad P\text{-a. s.},$$

where $U_\varepsilon(B) = \{x \in X : d(x, B) < \varepsilon\}$ denotes the ε-neighbourhood of $B \subset X$. This means that

$$\lim_{t\to\infty} \varphi(t, \vartheta_{-t}\omega)\rho\big(U_\varepsilon(A_{\mathrm{pt}}(\omega))\big) = 1,$$

which gives

$$\lim_{t\to\infty} d\big(\varphi(t, \vartheta_{-t}\omega)\rho, Pr\big(A_{\mathrm{pt}}(\omega)\big)\big) = 0 \qquad P\text{-a. s.},$$

which holds regardless of the choice of the metric d on $Pr(X)$. This proves that $Pr(A_{\mathrm{pt}})$ is a point attractor for Φ on $Pr(X)$.

Concerning (iii) we first note that there always exists a minimal point attractor for Φ on $Pr(X)$, compare Remark 2.6 (i). The second inclusion is then immediate from (ii), which says that $Pr\big(A_{\mathrm{pt}}^{\min}(\omega)\big)$ is a point attractor for Φ. What the first inclusion is concerned, note that for every $x \in A_{\mathrm{pt}}^{\min}(\omega)$ there exists a sequence of times t_n, tending to infinity for $n \to \infty$, and a point $x_0 \in X$, say, such that $\varphi(t_n, \vartheta_{-t_n}\omega)x_0$ converges to x. Consequently, $\varphi(t_n, \vartheta_{-t_n}\omega)\delta_{x_0}$ converges to δ_x, proving the assertion. $\square$

3.5 Remark It is not clear whether $Pr(A_{\mathrm{pt}}^{\min})$ always is the minimal point attractor for Φ, or whether it may occur that the minimal point attractor is smaller. This is open already for deterministic systems.

Theorem 3.4 indicates that an introduction of measure attractors for general RDS as random attractors for the induced RDS on the space $Pr(X)$ of Borel probability measures might not yield a very useful notion. Since both the global set attractor and the minimal point attractor are uniquely determined (compare Remarks 2.4 (iii) and 2.6 (i)), there is no chance to find more interesting set or point attractors on $Pr(X)$ by this approach, though. We therefore turn to another approach to a notion of measure attractors for RDS.

3.2 Measure Attractors on $Pr_\Omega(X)$

We consider the action of an RDS φ on the set of random probability measures $Pr_\Omega(X)$. The first thing to note is that φ does not define an RDS on $Pr_\Omega(X)$, since $(t, \gamma, \omega) \mapsto \varphi(t, \omega)\gamma_\omega$ does not satisfy Definition 2.1. However, we may refer to the skew product flow $(\Theta_t)_{t\in T}$, given by

$$\begin{aligned} \Theta : T \times X \times \Omega &\to X \times \Omega \\ (t, x, \omega) &\mapsto \big(\varphi(t, \omega)x, \vartheta_t\omega\big). \end{aligned}$$

As before, we may consider the action induced by Θ_t on the space $Pr(X \times \Omega)$ of measures on the product space, which we will denote by Θ_t again. On first view this may appear not very meaningful, since, for every $t \in T$, the skew product Θ_t is just measurable and not continuous. However, Θ_t leaves $Pr_\Omega(X)$ invariant, and $\Theta_t : Pr_\Omega(X) \to Pr_\Omega(X)$ is continuous with respect to the narrow topology for every t, see Crauel [11] Lemma 6.7.

There appears to be a very natural notion of an attractor now. Simply define a measure attractor as a *deterministic* attractor for the continuous semigroup Θ_t, $t \geq 0$, on the topological space $Pr_\Omega(X)$. There is one obstacle. While $Pr(X)$ is a Polish space again, in general $Pr_\Omega(X)$ is not. In fact, the narrow topology on $Pr_\Omega(X)$ is metrisable if and only if it is separable, and this holds if and only $\mathscr{F}$ is countably generated (mod P), see Crauel [11] Chapter 4, in particular Theorem 4.16 and Corollaries 4.24 and 4.31. Thus $Pr_\Omega(X)$ is Polish if, and only if, the σ-algebra $\mathscr{F}$ is countably generated (mod P). We therefore have to extend the definition of an attractor for a deterministic flow on a non-metrisable space. This is done in a straightforward fashion. It should be mentioned that for a non-metrisable space one cannot define set attractors for bounded sets, since there is no notion of a bounded set (compare Remark 2.4 (iii)).

3.6 Definition A compact, strictly invariant set $\mathcal{A} \subset Pr_\Omega(X)$ is a *point measure attractor* or *set measure attractor*, respectively, for the skew product flow Θ induced by an RDS φ, if for every $\Gamma \subset Pr_\Omega(X)$ finite or compact, respectively, and for every neighbourhood $\mathcal{U}$ of $\mathcal{A}$ in $Pr_\Omega(X)$ there exists a time $t_0 = t_0(\Gamma, \mathcal{U})$ such that $\Theta_t\Gamma \subset \mathcal{U}$ for all $t \geq t_0$.

Note that a set measure attractor attracts itself by assumption. Observing that the narrow topology has the Hausdorff property it is immediate that a set measure attractor is unique by its very definition. This is the same situation as encountered for deterministic systems, or for attractors for RDS which are assumed to attract classes of random sets including the random attractor.

We will make use of the following lemma, the proof of which is straightforward (see, e. g., Arnold [1] Lemma 1.4.4, or Crauel [11] Lemma 6.16).

3.7 Lemma *If Θ_t, $t \in T$, is the skew product flow induced by an RDS φ, then the disintegration of $\Theta_t \nu$ for $\nu \in Pr_\Omega(X)$, $t \in T$, is given by*

$$(\Theta_t \nu)_\omega = \varphi(t, \vartheta_{-t}\omega)\nu_{\vartheta_{-t}\omega} \qquad P\text{-a. s.}$$

3.8 Proposition *Suppose that φ is an RDS on a Polish space X with a set attractor $\omega \mapsto A(\omega)$. Then for every compact set $\Gamma \subset Pr_\Omega(X)$ of random probability measures, and for every $\varepsilon > 0$ and $\delta > 0$, there exists $T = T(\Gamma, \varepsilon, \delta, \omega)$, such that for every $\gamma \in \Gamma$*

$$\varphi(t, \vartheta_{-t}\omega)\gamma_{\vartheta_{-t}\omega}\big[U_\delta(A(\omega))\big] \geq 1 - \varepsilon \qquad P\text{-a. s. for every } t \geq T.$$

Proof Compactness of Γ in the narrow topology implies tightness by Crauel [11] Theorem 4.29. This in turn implies that, for every $\varepsilon > 0$, there exists a compact deterministic $K \subset X$ such that $\gamma_\omega(K) \geq 1 - \varepsilon$ a. s. for every $\gamma \in \Gamma$ by Crauel [11] Proposition 4.3. Since A is a set attractor, for P-a. e. ω there exists $T = T(K, \delta, \omega)$ such that, for all $t \geq T$, $d\big(\varphi(t, \vartheta_{-t}\omega)K, A(\omega)\big) \leq \delta$, i. e.,

$$\varphi(t, \vartheta_{-t}\omega)K \subset U_\delta(A(\omega)). \tag{6}$$

Consequently, for every $\gamma \in \Gamma$,

$$\begin{aligned} \varphi(t, \vartheta_{-t}\omega)\gamma_{\vartheta_{-t}\omega}\big[U_\delta(A(\omega))\big] &\geq \varphi(t, \vartheta_{-t}\omega)\gamma_{\vartheta_{-t}\omega}\big[\varphi(t, \vartheta_{-t}\omega)K\big] \\ &= \gamma_{\vartheta_{-t}\omega}\big[\varphi(t, \vartheta_{-t}\omega)^{-1}\big(\varphi(t, \vartheta_{-t}\omega)K\big)\big] \\ &\geq \gamma_{\vartheta_{-t}\omega}(K) \;\geq\; 1 - \varepsilon \qquad P\text{-a. s.,} \end{aligned}$$

invoking (6) in the first inequality. □

In order to make use of this result we have to include an elementary lemma on the Prohorov metric on the space of Borel probability measures. For a metric space (X, d), the Prohorov metric on $Pr(X)$ associated with d on X is given by

$$d_P(\rho, \eta) = \inf\{r \geq 0 : \rho(B) \leq \eta(U_r(B)) + r \text{ for every Borel set } B \subset X\}$$

for $\rho, \eta \in Pr(X)$.

3.9 Lemma *For the Prohorov metric d_P on the space of Borel probability measures over a metric space (X, d) the following holds.*

(i) *For any* $\rho_1, \rho_2 \in Pr(X)$, $\eta_1, \eta_2 \in Pr(X)$, *and* $p \in [0,1]$, *for the convex combinations* $\rho = p\rho_1 + (1-p)\rho_2$ *and* $\eta = p\eta_1 + (1-p)\eta_2$ *one has*

$$d_P(\rho, \eta) \le \max\big\{d_P(\rho_1, \eta_1), d_P(\rho_2, \eta_2)\big\}. \tag{7}$$

(ii) *For any* $\eta_1, \eta_2 \in Pr(X)$, $p \in [0,1]$, *the convex combination* $\eta = p\eta_1 + (1-p)\eta_2$ *satisfies*

$$d_P(\eta, \eta_1) \le 1 - p.$$

PROOF Concerning (i), for any $r \ge \max\big\{d_P(\rho_1, \eta_1), d_P(\rho_2, \eta_2)\big\}$

$$\begin{aligned} \rho(B) &= p\rho_1(B) + (1-p)\rho_2(B) \\ &\le p\big(\eta_1(U_r(B)) + r\big) + (1-p)\big(\eta_2(U_r(B)) + r\big) \\ &= \eta(U_r(B)) + r \end{aligned}$$

for every Borel set $B \subset X$, which implies (7).

What (ii) is concerned, we get

$$\eta(B) = p\eta_1(B) + (1-p)\eta_2(B) \le \eta_1(U_r(B)) + 1 - p$$

for any $r \ge 0$ and any Borel set $B \subset X$. Consequently, for $r \ge 1-p$ we have $\eta(B) \le \eta_1(U_r(B)) + r$ for any Borel set $B \subset X$, which implies $d_P(\eta, \eta_1) \le 1 - p$. □

The next result aims at a quantitative relation of attraction on the state space with respect to some metric with attraction in the space of probability measures with respect to the Prohorov metric associated with the metric on the state space.

Suppose that $\omega \mapsto A(\omega)$ is a compact random set with $U_\delta(A(\omega))$ denoting the δ-neighbourhood of $A(\omega)$ for some $\delta > 0$. In general $\omega \mapsto U_\delta(A(\omega))$ may not be an open random set, i. e., it need not be measurable with respect to the original σ-algebra $\mathscr{F}$, but only with respect to its universal completion (see Crauel [11] Chapter 2, in particular Remark 2.11 (ii)). In the following result we will use arguments which implicitly rely on measurability of $\omega \mapsto \gamma_\omega(U_\delta(A(\omega)))$. We will assume in such cases that random variables with values in separable metric spaces, whose construction only yields measurability with respect to the universal completion, are tacitly replaced by $\mathscr{F}$-measurable versions.

3.10 Proposition *Suppose that* $\omega \mapsto A(\omega)$ *is a compact random set in a Polish space* X, *and let* d *be a metric on* X *inducing the topology. Suppose further that* $\gamma \in Pr_\Omega(X)$. *Then, for every* $\delta > 0$, *there exists* $\alpha \in Pr_\Omega(X)$ *with* $\alpha_\omega(A(\omega)) = 1$ P-a. s. *such that*

$$d_P(\gamma_\omega, \alpha_\omega) \le 2\delta + \gamma_\omega(U_\delta(A(\omega))^c) \qquad P\text{-a. s.}, \tag{8}$$

where d_P *is the Prohorov metric associated with* d.

Proof First suppose that $\gamma_\omega(U_\delta(A(\omega))) = 1$ P-a. s. Since $\gamma_\omega \in \mathrm{cl\ conv}\,\{\delta_x : x \in \overline{U_\delta}(A(\omega))\}$ we may choose random variables $\omega \mapsto n(\omega) \in \mathbb{N}$, $x_1(\omega), \dots, x_n(\omega) \in \overline{U_\delta}(A(\omega))$, and $p_1(\omega), \dots, p_n(\omega) \in [0, 1]$ with $\sum p_j = 1$, such that $d_P\big(\gamma_\omega, \sum_{j=1}^{n(\omega)} p_j(\omega)\delta_{x_j(\omega)}\big) < \delta$. Here we made use of the fact that the support of γ is a closed random set (see Crauel [11] Remark 3.8) in order to apply the selection theorem (Castaing and Valadier [6] Theorem III.9). For every j, $1 \le j \le n(\omega)$, choose a random variable y_j such that $y_j(\omega) \in A(\omega)$ and $d(x_j, y_j) \le \delta$ a. s. Putting $\alpha_\omega = \sum_j p_j(\omega)\delta_{y_j(\omega)}$ defines $\alpha \in Pr_\Omega(X)$ which, furthermore, satisfies

$$d_P(\gamma_\omega, \alpha_\omega) \le d_P(\gamma_\omega, \textstyle\sum_j p_j(\omega)\delta_{x_j(\omega)}) + d_P(\textstyle\sum_j p_j(\omega)\delta_{x_j(\omega)}, \alpha_\omega) \le 2\delta, \tag{9}$$

invoking Lemma 3.9 (i) together with the fact that $d_P(\delta_u, \delta_v) \le d(u, v)$. This proves the assertion for the particular case $\gamma_\omega(U_\delta(A(\omega))) = 1$ a. s.

Turning to the general case with $\gamma_\omega(U_\delta(A(\omega)))$ not necessarily equal to one almost surely, define a random measure $\tilde\gamma$ by

$$\tilde\gamma_\omega(B) = \begin{cases} \dfrac{\gamma_\omega\big(U_\delta(A(\omega)) \cap B\big)}{\gamma_\omega(U_\delta(A(\omega)))} & \text{for } \gamma_\omega(U_\delta(A(\omega))) > 0 \\ \gamma_\omega(B) & \text{for } \gamma_\omega(U_\delta(A(\omega))) = 0. \end{cases}$$

Then γ can be written as the convex combination

$$\gamma = p\tilde\gamma + (1 - p)\tilde\gamma',$$

where the random variable p is given by $p(\omega) = \gamma_\omega(U_\delta(A(\omega)))$, and where $\tilde\gamma'_\omega$ is uniquely determined for $p(\omega) \ne 1$, and defined to be a fixed element of $Pr(X)$, say, for ω with $p(\omega) = 1$. Choose some random variable x_0 with $x_0(\omega) \in A(\omega)$ a. s., and put $p_0(\omega) = 1$ for $\gamma_\omega(U_\delta(A(\omega))) = 0$, $p_0(\omega) = 0$ else. Redefining for a moment $\tilde\gamma$ by putting $\tilde\gamma^0_\omega = \delta_{x_0(\omega)}$ for ω with $\gamma_\omega(U_\delta(A(\omega))) = 0$ gives a random measure $\tilde\gamma^0$ with $\tilde\gamma^0_\omega(U_\delta(A(\omega))) = 1$. Defining α associated with this modified $\tilde\gamma^0$ exactly as in the first step we get $d_P(\tilde\gamma_\omega, \alpha_\omega) \le 1 \le 2\delta + 1$ for ω with $\gamma_\omega(U_\delta(A(\omega))) = 0$ from the fact that the Prohorov metric is bounded by one. For ω with $\gamma_\omega(U_\delta(A(\omega))) \ne 0$ we may infer from Lemma 3.9 (ii) together with (9) that

$$d_P(\gamma_\omega, \alpha_\omega) \le d_P(\gamma_\omega, \tilde\gamma_\omega) + d_P(\tilde\gamma_\omega, \alpha_\omega) \le (1 - \gamma_\omega(U_\delta(A\omega))) + 2\delta,$$

which proves (8). □

We note in passing that the estimate in (8) can be made slightly sharper, insofar 2δ may be replaced by $c\delta$ with $c > 1$ arbitrary.

We are now in a position to prove the main result of the present subsection.

3.11 Theorem *Suppose that φ is an RDS on a Polish space X with a set attractor $\omega \mapsto A(\omega)$. Then the set*

$$\mathcal{A} = \{\nu \in Pr_\Omega(X) : \nu_\omega\big(A(\omega)\big) = 1 \ P\text{-a. s.}\}$$

is a (and thus the*) set measure attractor in the sense of Definition* 3.6.

One might suggestively write $Pr_\Omega(A)$ instead of $\mathcal{A}$, but for notational brevity we will stick to $\mathcal{A}$.

Proof The set of all random measures supported by a compact random set is compact in the narrow topology by Crauel [11] Proposition 4.3, so $\mathcal{A}$ is compact. To get strict Θ_t-invariance of $\mathcal{A}$ first note that by Lemma 3.7, for $\nu \in \mathcal{A}$,

$$\begin{aligned}(\Theta_t\nu)_\omega\big(A(\omega)\big) &= \varphi(t,\vartheta_{-t}\omega)\nu_{\vartheta_{-t}\omega}\big(A(\omega)\big)\\ &= \nu_{\vartheta_{-t}\omega}\big(\varphi(t,\vartheta_{-t}\omega)^{-1}(A(\omega))\big)\\ &\geq \nu_{\vartheta_{-t}\omega}\big(A(\vartheta_{-t}\omega)\big) = 1 \quad P\text{-a. s.},\end{aligned}$$

invoking $\varphi(t,\vartheta_{-t}\omega)A(\vartheta_{-t}\omega) = A(\omega)$ to conclude $A(\vartheta_{-t}\omega) \subset \varphi(t,\vartheta_{-t}\omega)^{-1}(A(\omega))$. This implies $\Theta_t\mathcal{A} \subset \mathcal{A}$.
To see that also the other inclusion holds, fix $\nu \in \mathcal{A}$. The sets $\{\rho \in Pr(X) : \rho\big(A(\vartheta_{-t}\omega)\big) = 1\}$ and $\{\rho \in Pr(X) : \varphi(t,\vartheta_{-t}\omega)\rho = \nu_\omega\}$ define a random compact and a random closed set, respectively, in $Pr(X)$. Surjectivity of $\varphi(t,\vartheta_{-t}\omega) : A(\vartheta_{-t}\omega) \to A(\omega)$ together with $\nu_\omega\big(A(\omega)\big) = 1$ a. s. allows to invoke Lemma 3.1 in order to conclude that the intersection is nonempty, so that

$$\omega \mapsto \big\{\rho \in Pr(X) : \rho\big(A(\vartheta_{-t}\omega)\big) = 1 \text{ and } \varphi(t,\vartheta_{-t}\omega)\rho = \nu_\omega\big\}$$

is an a. s. nonempty compact random set. We may therefore use the selection theorem (Castaing and Valadier [6] Theorem III.9) to conclude the existence of a $Pr(X)$-valued random variable $\omega \to \bar{\eta}_\omega$ with $\bar{\eta}_\omega\big(A(\vartheta_{-t}\omega)\big) = 1$ and $\varphi(t,\vartheta_{-t}\omega)\bar{\eta}_\omega = \nu_\omega$ a. s. Defining η by $\eta_\omega = \bar{\eta}_{\vartheta_t\omega}$ gives $\eta \in \mathcal{A}$ with $\Theta_t\eta = \nu$.
Finally, in order to prove attraction of $\mathcal{A}$ suppose that $\Gamma \subset Pr(X)$ is compact, and that $\mathcal{U}$ is a neighbourhood of $\mathcal{A}$ in $Pr_\Omega(X)$, both in the narrow topology of $Pr_\Omega(X)$. We have to show that $\Theta_t\Gamma \subset \mathcal{U}$ for t sufficiently large. In order to achieve this we are going to use the fact that the metric on $Pr_\Omega(X)$ given by

$$(\gamma,\eta) \mapsto \int_\Omega d_P(\gamma_\omega,\eta_\omega)\,dP(\omega), \tag{10}$$

where d_P denotes the Prohorov metric on $Pr(X)$, induces a topology which is stronger than the narrow topology, see Crauel [11] Chapter 5, in particular Proposition 5.4 and Lemma 5.2. Consequently, for every neighbourhood $\mathcal{U}$ of $\mathcal{A}$ with respect to the narrow topology there exists $\delta_0 > 0$ such that the δ-neighbourhood of $\mathcal{A}$ with respect to the metric given by (10) is contained in $\mathcal{U}$ for every $\delta < \delta_0$.
Fix a neighbourhood $\mathcal{U}$ of $\mathcal{A}$ in the narrow topology, and let $\delta_0 > 0$ be such that $U_{\delta_0}(\mathcal{A}) \subset \mathcal{U}$, where $U_\delta(\mathcal{A})$ denotes the δ-neighbourhood of $\mathcal{A}$ in the metric given by (10). Fix δ with $0 < \delta < \delta_0/4$. Proposition 3.8 implies existence of a $T = T(\Gamma,\delta,\omega)$ such that for every $\gamma \in \Gamma$

$$\varphi(t,\vartheta_{-t}\omega)\gamma_{\vartheta_{-t}\omega}\big[U_\delta(A(\omega))\big] \geq 1-\delta \qquad P\text{-a. s. for every } t \geq T(\Gamma,\delta,\omega). \tag{11}$$

Choose $T_0 \in [0, \infty)$ so large that $P\{T(\Gamma, \delta, \omega) > T_0\} < \delta$. Fix $\gamma \in \Gamma$. Proposition 3.10, applied for $\Theta_t\gamma$, yields existence of an $\alpha \in \mathcal{A}$ such that

$$d_P(\varphi(t, \vartheta_{-t}\omega)\gamma_{\vartheta_{-t}\omega}, \alpha_\omega) \leq 2\delta + \varphi(t, \vartheta_{-t}\omega)\gamma_{\vartheta_{-t}\omega}(U_\delta(A(\omega))^c) \qquad P\text{-a. s.}$$

For $t \geq T_0$ this implies

$$\begin{aligned}
&\int_\Omega d_P(\Theta_t\gamma, \alpha)\, dP \\
&= \int_\Omega d_P(\varphi(t, \vartheta_{-t}\omega)\gamma_{\vartheta_{-t}\omega}, \alpha_\omega)\, dP(\omega) \\
&\leq \int\limits_{\{T(\omega) \leq T_0\}} \Big(2\delta + \varphi(t, \vartheta_{-t}\omega)\gamma_{\vartheta_{-t}\omega}(U_\delta(A(\omega))^c)\Big)\, dP(\omega) + P\{T(\omega) > T_0\} \\
&\leq 4\delta,
\end{aligned}$$

using (11) to conclude that $\varphi(t, \vartheta_{-t}\omega)\gamma_{\vartheta_{-t}\omega}(U_\delta(A(\omega))^c) \leq \delta$ for $t \geq T(\omega)$, where we dropped the dependence of $T(\omega)$ on δ and on Γ notationally. This implies

$$\inf_{\alpha \in \mathcal{A}} \int_\Omega d_P(\Theta_t\gamma, \alpha)\, dP < 4\delta$$

or, equivalently, $\Theta_t\gamma \in U_{4\delta}(\mathcal{A})$ for every $t \geq T_0$ and for every $\gamma \in \Gamma$, and therefore $\Theta_t\Gamma \subset U_{4\delta}(\mathcal{A})$ for $t \geq T_0$. In view of $U_{4\delta}(\mathcal{A}) \subset \mathcal{U}$ this proves the desired attraction property of $\mathcal{A}$. □

3.12 Remark The proof of Theorem 3.11 establishes, roughly speaking, attraction with respect to a stronger topology, which in addition happens to be metrisable. One might wonder whether the notion of an attractor should rather be defined with respect to this stronger topology on the space of random measures. However, in this topology $\mathcal{A}$ fails to be compact in general.

3.13 Remark Since a set attractor always is a point attractor, the assumptions of Theorem 3.11 imply existence of both point attractors $\omega \mapsto A_{\text{pt}}(\omega)$ for φ and point measure attractors for Θ. However, trying to carry over the idea of Theorem 3.11 to point attractors fails. The set $\{\nu \in Pr_\Omega(X) : \nu(A_{\text{pt}}) = 1\}$ is a compact, strictly invariant set of random measures, but in general it need not be a point measure attractor for Θ in the sense of Definition 3.6. See Example 5.7 below. Since for deterministic systems the two concepts of measure attractors introduced above coincide, Theorem 3.4 (ii) applies, whence the above set always gives a point measure attractor. This shows that this is a phenomenon which cannot occur for deterministic systems, but which is specific for stochastic systems.

We emphasize once again that for general RDS the concepts of measure attractors on $Pr(X)$ and on $Pr_\Omega(X)$ as developed in Sections 3.1 and 3.2, respectively, are not the same. However, they coincide for deterministic systems.

Markov measure attractors for general RDS

We now turn to Markov measures as a particular class of random measures in $Pr_\Omega(X)$, defined in reference to an RDS φ. Denoting by

$$\mathcal{M} = \{\nu \in Pr_\Omega(X) : \omega \mapsto \nu_\omega \text{ is } \mathscr{F}^-\text{-measurable}\},$$

where $\mathscr{F}^-$ is the past of an RDS φ, the set of Markov measures with respect to φ, we have $\Theta_t\mathcal{M} \subset \mathcal{M}$ for $t \geq 0$. In general the inclusion is strict for $t > 0$. For instance, identifying deterministic measures $\rho \in Pr(X)$ with constant elements of $Pr_\Omega(X)$ we get $Pr(X) \subset Pr_\Omega(X)$. Clearly, $Pr(X) \subset \mathcal{M}$, but $Pr(X) \not\subset \Theta_t\mathcal{M}$ for $t > 0$ in general.

3.14 Definition A set $\mathcal{C} \subset Pr_\Omega(X)$ is said to be a (*set* or *point*, respectively) *Markov measure attractor* if it is compact and strictly (Θ_t)-invariant, and if for every $\Gamma \subset \mathcal{M}$ compact or finite, respectively, and for every neighbourhood $\mathcal{U}$ of $\mathcal{C}$ in $Pr_\Omega(X)$ there exists a time $t_0 = t_0(\Gamma, \mathcal{U})$ such that $\Theta_t\Gamma \subset \mathcal{U}$ for all $t \geq t_0$.

In general a Markov attractor need not be unique. In fact, in the situation of Theorem 3.11 the general measure attractor $\mathcal{A}$ is also a Markov measure attractor. However, in general the minimal Markov measure attractor is smaller.

3.15 Theorem *Suppose that φ is an RDS on a Polish space X with a set attractor $\omega \mapsto A(\omega)$. Then there exists a minimal set Markov measure attractor $\mathcal{A}^{\mathcal{M}} \subset Pr_\Omega(X)$, and it is given by*

$$\mathcal{A}^{\mathcal{M}} = \bigcap_{t\geq 0} \Theta_t(\mathcal{A} \cap \mathcal{M}), \tag{12}$$

where $\mathcal{A}$ is the global set measure attractor given by Theorem 3.11. *In particular, $\mathcal{A}^{\mathcal{M}} \subset \mathcal{M}$, i. e., the minimal Markov attractor consists itself of Markov measures.*

Proof The set $\mathcal{A} \cap \mathcal{M}$ is compact in the narrow topology, since $\mathcal{M}$ is closed in $Pr_\Omega(X)$ by Crauel [11] Theorem 4.21. Thus $\bigcap_{t\geq 0}\Theta_t(\mathcal{A}\cap\mathcal{M})$ is compact and nonempty, invoking continuity of Θ_t.
In order to establish strict invariance of $\bigcap_{t\geq 0}\Theta_t(\mathcal{A}\cap\mathcal{M})$, first note that $\Theta_t\bigcap_{\tau\geq 0}\Theta_\tau(\mathcal{A}\cap\mathcal{M}) \subset \bigcap_{\tau\geq 0}\Theta_\tau(\mathcal{A}\cap\mathcal{M})$ for every $t \geq 0$, proving invariance. With regard to strict invariance suppose that $\gamma \in \bigcap_{\tau\geq 0}\Theta_\tau(\mathcal{A}\cap\mathcal{M})$. We have to show that there exists $\zeta \in \mathcal{A}\cap\mathcal{M}$ with $\Theta\zeta = \gamma$. Choose a sequence (τ_n) with $\tau_n \to \infty$ for $n \to \infty$. Then there exist $\alpha_n \in \mathcal{A}\cap\mathcal{M}$ with $\Theta_{\tau_n}\alpha_n = \gamma$. Now consider the sequence (ζ_n), given by $\zeta_n = \Theta_{\tau_n - t}\alpha_n$ for n with $\tau_n \geq t$, and note that $\zeta_n \in \mathcal{A}\cap\mathcal{M}$. Since $\mathcal{A}\cap\mathcal{M}$ is sequentially compact by Crauel [11] Proposition 4.28 there is a subsequence, denoted by (ζ_n) again, converging to some $\zeta \in \mathcal{A}\cap\mathcal{M}$. Continuity of Θ_t implies $\Theta_t\zeta = \lim_{n\to\infty}\Theta_t\zeta_n = \lim_{n\to\infty}\Theta_{t\tau_n}\alpha_n = \gamma$.
To show that $\bigcap_{t\geq 0}\Theta_t(\mathcal{A}\cap\mathcal{M})$ is attracting, suppose that Γ is a compact set of Markov measures. Then Γ is attracted by $\mathcal{A}$ by virtue of Theorem 3.11. Consequently, Γ is attracted by its own Ω-limit set (with respect to (Θ_t)), given by $\Omega(\Gamma) = \bigcap_{t\geq 0}\overline{\bigcup_{\tau\geq t}\Theta_\tau\Gamma}$, and $\Omega(\Gamma) \subset \mathcal{A}$. On the other hand, $\Gamma \subset \mathcal{M}$ implies $\Omega(\Gamma) \subset \mathcal{M}$ in view of the fact that $\mathcal{M}$ is both closed and Θ_t-invariant for $t \geq 0$. Therefore, $\mathcal{A}\cap\mathcal{M}$ attracts Γ, and

$\Omega(\Gamma) \subset \mathcal{A}\cap\mathcal{M}$. Since Γ is attracted by the compact set $\mathcal{A}\cap\mathcal{M}$, $\Omega(\Gamma)$ is strictly invariant. Consequently, $\Omega\big(\Omega(\Gamma)\big) = \Omega(\Gamma)$, hence $\Omega(\Gamma) \subset \mathcal{A}\cap\mathcal{M}$ implies $\Omega(\Gamma) \subset \Omega(\mathcal{A}\cap\mathcal{M})$. Since $\Theta_t(\mathcal{A}\cap\mathcal{M}) \subset \Theta_t\mathcal{A}\cap\Theta_t\mathcal{M} = \mathcal{A}\cap\Theta_t\mathcal{M} \subset \mathcal{A}\cap\mathcal{M}$ for every $t \geq 0$, $\mathcal{A}\cap\mathcal{M}$ is (in general not strictly) invariant, which implies $\Omega(\mathcal{A}\cap\mathcal{M}) = \bigcap_{t\geq 0}\Theta_t(\mathcal{A}\cap\mathcal{M})$. This holds for every compact $\Gamma \subset \mathcal{M}$, proving that the right hand side of (12) is, indeed, a Markov set attractor.
Finally, in order to show that the set Markov measure attractor given by (12) is the minimal set Markov measure attractor, note that $\Theta_t(\mathcal{A}\cap\mathcal{M}) \subset \mathcal{M}$ for every $t \geq 0$. Therefore also $\bigcap_{t\geq 0}\Theta_t(\mathcal{A}\cap\mathcal{M}) \subset \mathcal{M}$, hence $\bigcap_{t\geq 0}\Theta_t(\mathcal{A}\cap\mathcal{M})$ is itself a compact set of Markov measures. Thus any other Markov measure attractor has to attract $\mathcal{A}^{\mathcal{M}}$. But $\mathcal{A}^{\mathcal{M}}$ is invariant, whence any other Markov measure attractor must contain $\mathcal{A}^{\mathcal{M}}$. $\square$

As noted in the proof of Theorem 3.15, Θ_t-invariance of the set $\mathcal{M}$ of Markov measures for $t \geq 0$ implies that its Ω-limit set equals $\bigcap_{t\geq 0}\Theta_t\mathcal{M}$. Furthermore, one immediately verifies $\Omega(\Gamma) \subset \Omega(\mathcal{M})$ for every compact set Γ of Markov measures. One might be tempted to conjecture that the Ω-limit set of the set of all Markov measures is a Markov measure attractor already. However, this fails to be true in general, as the case of a two-sided deterministic system shows. There the set of Markov measures is $Pr(X)$ itself, and it is strictly invariant.

3.16 Corollary *Under the assumptions of Theorem* 3.15, *the global set Markov measure attractor* $\mathcal{A}^{\mathcal{M}}$*, given by* (12), *satisfies*

$$\mathcal{A}^{\mathcal{M}} \subset \mathcal{A}\cap\Big(\bigcap_{t\geq 0}\Theta_t\mathcal{M}\Big).$$

If the RDS φ has two-sided time then equality holds, i. e.,

$$\mathcal{A}^{\mathcal{M}} = \mathcal{A}\cap\Big(\bigcap_{t\geq 0}\Theta_t\mathcal{M}\Big). \tag{13}$$

Proof For every $t \geq 0$

$$\Theta_t(\mathcal{A}\cap\mathcal{M}) \subset \Theta_t\mathcal{A}\cap\Theta_t\mathcal{M} = \mathcal{A}\cap\Theta_t\mathcal{M},$$

where we used that $\mathcal{A}$ is invariant with respect to Θ_t by Theorem 3.11. This implies, invoking (12), $\mathcal{A}^{\mathcal{M}} = \bigcap_{t\geq 0}\Theta_t(\mathcal{A}\cap\mathcal{M}) \subset \mathcal{A}\cap\big(\bigcap_{t\geq 0}\Theta_t\mathcal{M}\big)$. If φ is two-sided then also (Θ_t) is two-sided, and so Θ_t is invertible for every t. Consquently, then $\Theta_t(\mathcal{A}\cap\mathcal{M}) = \Theta_t\mathcal{A}\cap\Theta_t\mathcal{M} = \mathcal{A}\cap\Theta_t\mathcal{M}$ for every t, and (13) follows by taking the intersection over $t \geq 0$. $\square$

The following example shows that in general $\mathcal{A}^{\mathcal{M}} \neq \mathcal{A}\cap\bigcap_{t\geq 0}\Theta_t\mathcal{M}$ for the minimal set Markov measure attractor $\mathcal{A}^{\mathcal{M}}$. In fact, the example shows that $\mathcal{A}\cap\bigcap_{t\geq 0}\Theta_t\mathcal{M}$ need not be strictly invariant.

3.17 Example Suppose that η is a non-degenerate $\{0,1\}$-valued random variable such that $\left(\eta \circ \vartheta_n\right)_{n\in\mathbb{N}}$ is i.i.d. Define a discrete time RDS φ on $X = \mathbb{N}$ by its time one map

$$\varphi(\omega)x = \begin{cases} x-1 & \text{for } x \geq 2 \\ \eta(\omega) & \text{for } x \in \{0,1\}, \end{cases}$$

The (set and point) attractor of φ is (the deterministic set) $A = \{0,1\}$. Clearly $\delta_1 \in \mathcal{A}$, where $\mathcal{A}$ is the set measure attractor of φ given by Theorem 3.11. Furthermore, $\delta_1 \in \Theta_n\mathcal{M}$ for every $n \geq 0$, since $\delta_1 = \varphi(n, \vartheta_{-n}\omega)\delta_{n+1}$, hence $\delta_1 \in \mathcal{A} \cap \bigcap_{n\geq 0} \Theta_n\mathcal{M}$. However, $\delta_1 \notin \mathcal{A}^{\mathcal{M}}$, since there is no $\nu \in \mathcal{M}$ with $\Theta\nu = \delta_1$.

Though we cannot characterize point measure attractors as precisely as Theorem 3.11 does with respect to set measure attractors, it is straightforward to carry the arguments over for Markov point attractors.

3.18 Corollary *Under the assumptions of Theorem* 3.15, *let* $\mathcal{A}_{\mathrm{pt}}$ *be some point measure attractor (not necessarily the minimal one). Then the set*

$$\bigcap_{t\geq 0} \Theta_t(\mathcal{A}_{\mathrm{pt}} \cap \mathcal{M}) \tag{14}$$

of random probability measures is a point Markov measure attractor.

Proof The arguments of the proof of Theorem 3.15 go through, replacing in the proof of attractivity of (14) the compact set Γ by a finite set of Markov measures. □

3.19 Remark (i) As stated in Theorem 2.9 (ii), in case of a white noise system every *invariant* Markov measure is supported by the minimal point attractor of the RDS. The analogue statement for the Markov set measure attractor instead of the set of invariant Markov measures does not hold. In general $\mathcal{A}^{\mathcal{M}} \not\subset Pr(A_{\mathrm{pt}})$, which can be seen already for deterministic systems. A deterministic system is trivially white noise, and every Borel probability measure on the state space is Markov, whence $\mathcal{A}^{\mathcal{M}} = \mathcal{A}$, and $\mathcal{A} = Pr(A_{\mathrm{set}})$ by Theorem 3.4 (i) or Theorem 3.11.

(ii) However, denoting the minimal point Markov measure attractor by $\mathcal{A}_{\mathrm{pt}}^{\mathcal{M}}$, one should get $\mathcal{A}_{\mathrm{pt}}^{\mathcal{M}} \subset Pr_\Omega(A_{\mathrm{pt}})$, where $Pr_\Omega(A_{\mathrm{pt}}) = \{\nu \in Pr_\Omega(X) : \nu_\omega\big(A(\omega)\big) = 1\ P\text{-a.s.}\}$. Let us briefly sketch an argument: Approximate a Markov measure by a convex combination of a finite set of Dirac measures. Then use attraction of A_{pt} for finite valued random variables (compare Remark 2.7) with respect to φ to conclude that the Markov measure is attracted by $Pr_\Omega(A_{\mathrm{pt}})$. In order not to deviate too far from the main theme we refrain from entering the details of the argument.

(iii) But note that as soon as (ii) has been established one may then use the idea of Corollary 3.18, restricting it to attraction of Markov measures, in order to conclude

$$\mathcal{A}_{\mathrm{pt}}^{\mathcal{M}} \subset \bigcap_{t\geq 0} \Theta_t\big(Pr_\Omega(A_{\mathrm{pt}}) \cap \mathcal{M}\big).$$

It is not clear whether one always has equality here. This question is closely related to the one from Remark 3.5, and an affirmative answer here would settle the latter by considering a deterministic system as an RDS.

Until now the discussion of Markov properties referred to general abstract RDS. Here the Markov property reduces, roughly speaking, to measurability with respect to the past. It is only for white noise RDS that the classical Markov theory can be invoked.

4 Markov Measures for white noise RDS

We now turn to the particular case of white noise RDS. Associated with a white noise RDS there is a Markov semigroup, describing the behaviour of distributions of individual trajectories, or one point motions. It might be worth mentioning, though, that also multiple point motions give rise to Markov semigroups on a corresponding product of the state space. See Baxendale [3] for a systematic study.

4.1 Proposition *Suppose that φ is a white noise RDS. Then, with $(\mathscr{F}_s^t)$, $s \le t$, denoting the filtration associated with φ (compare* (1)*), the σ-algebras $\mathscr{F}_u^v$ and $\mathscr{F}_s^t$ are independent for any u, v, s, t with $-\infty \le u \le v \le s \le t \le \infty$.*

Proof For $-\infty \le u \le v \le s \le t \le \infty$ choose $\tau \in T$ with $v \le \tau \le s$. Then $\vartheta_{-\tau}^{-1}\mathscr{F}_u^v = \mathscr{F}_{u-\tau}^{v-\tau} \subset \mathscr{F}^-$ and $\vartheta_{-\tau}^{-1}\mathscr{F}_s^t = \mathscr{F}_{s-\tau}^{t-\tau} \subset \mathscr{F}^+$, invoking Lemma 2.2 (i). The RDS φ being white, $\vartheta_{-\tau}^{-1}\mathscr{F}_u^v$ and $\vartheta_{-\tau}^{-1}\mathscr{F}_s^t$ are independent, so by invariance of P with respect to ϑ_τ also $\mathscr{F}_u^v$ and $\mathscr{F}_s^t$ are independent. $\square$

As a consequence of Proposition 4.1, the filtration induced by a white noise RDS gives a *filtered dynamical system* $(\Omega, \mathscr{F}, P, \mathscr{F}_s^t, \{\vartheta_t\}_{t\in\mathbb{R}})$ in the sense of Schmalfuß [21].

Independence of the 'increments of φ', as given by Proposition 4.1, allows to define Markov semigroups along standard lines. We refrain from going into details, and only state the basic relations. See Arnold [1] Sections 2.1.3 and 2.3.9 for a more detailed exposition.

Suppose that φ is a white noise RDS. Then

$$P_t f(x) = Ef\big(\varphi(t,\omega)x\big)$$

for $t \ge 0$ defines a semigroup on the space of bounded measurable as well as on the space of bounded continuous functions $f : X \to \mathbb{R}$, i. e., $P_{t+s} = P_t \circ P_s$ for $s, t \ge 0$. Furthermore, P_t can be defined on Borel measures ρ on X by

$$P_t\rho(B) = E\Big(\int_X 1_B(\varphi(t,\omega)x)\,d\rho(x)\Big)$$

for $t \ge 0$, which can be rewritten as $P_t\rho = E\big(\varphi(t,\cdot)\rho\big)$. One has $\int_X P_t f\,d\rho = \int_X f\,d(P_t\rho)$ for every bounded measurable $f : X \to \mathbb{R}$ and every finite Borel measure ρ on X.

The semigroup $(P_t)_{t\geq 0}$ defined by a white noise RDS φ is addressed to as the *Markov semigroup induced by* φ. It is denoted by the same symbol $(P_t)_{t\geq 0}$ both on functions and on measures.
A Borel probability measure ρ on X with $P_t\rho = \rho$ for all $t \geq 0$ is said to be an *invariant measure for the Markov semigroup* P_t, or briefly P_t*-invariant*.

The relation between invariant measures for the RDS and invariant measures for the Markov semigroup is well known. For a proof of the following Proposition see, e. g., Arnold [1] Theorems 1.7.2 and 2.3.45.

4.2 Proposition *Suppose that φ is a white noise RDS, and denote by (P_t) its associated Markov semigroup. Then*

(i) *for any $\rho \in Pr(X)$ with $P_t\rho = \rho$ for all $t \geq 0$*

$$\lim \varphi(t, \vartheta_{-t}\omega)\rho = \mu_\omega \tag{15}$$

exists, and the limiting random measure $\omega \mapsto \mu_\omega$ is an invariant Markov measure for φ.

(ii) *For any invariant Markov measure μ for φ, the Borel probability measure ρ on X, given by $\rho(B) = \int_\Omega \mu_\omega(B)\, dP(\omega)$, satisfies $P_t\rho = \rho$ for all $t \geq 0$.*

The limit in (15) exists in the sense that, for every bounded measurable function $f : X \to \mathbb{R}$, the real valued stochastic process

$$(t, \omega) \mapsto \int_X f(x)\, d(\varphi(t, \vartheta_{-t}\omega)\rho)(x) = \int_X f(\varphi(t, \vartheta_{-t}\omega))\, d\rho(x)$$

is a bounded martingale, and therefore converges P-almost surely. This implies, in particular, P-almost sure convergence of the measure-valued random variables $\omega \mapsto \varphi(t, \vartheta_{-t}\omega)\rho$ to μ_ω in the topology of weak convergence on $Pr(X)$. But almost sure convergence in the topology of weak convergence implies convergence in the narrow topology of $Pr_\Omega(X)$ (see Crauel [11] Chapter 5).

By Proposition 4.2, invariant Markov measures for φ and invariant measures for the Markov semigroup (P_t) may be identified. The set of all invariant Markov measures is a convex and closed subset of the set of all invariant measures (closed in the narrow topology on random measures as well as in the topology of weak convergence on $Pr(X)$).

5 Measure Attractors and the Schmalfuß Attractor; Invariant Markov Measures

Suppose that φ is an RDS on a Polish space X with an attractor A. For $t \geq 0$ define

$$\mathcal{D}_t = \text{cl conv } \{E(\delta_{\varphi(t,\vartheta_{-t}\omega)b(\vartheta_{-t}\omega)}) : b \text{ an } \mathscr{F}^-\text{-selection of } A\},$$

where the convex hull is taken in $Pr(X)$, and the closure is taken with respect to the topology of weak convergence in $Pr(X)$. Put

$$\mathfrak{A} = \bigcap_{t \geq 0} \mathcal{D}_t, \tag{16}$$

which is a deterministic subset of $Pr(X)$. In case of a white noise RDS φ the Markov semigroup induced by φ will be denoted by $(P_t)_{t\geq 0}$. The following result is proved in Schmalfuß [21].

5.1 Theorem *Suppose that φ is a white noise RDS on a separable Banach space H with norm $|\cdot|$, and A is a (and thus* the*) global set attractor for φ. Suppose that A is, in addition,* tempered, *i. e., $\lim_{t\to\infty}\big[\log^+\big(\sup\{|a| : a \in A(\vartheta_t\omega)\}\big)\big]/t = 0$, and that A attracts in addition to compact sets also every tempered random set. Suppose furthermore that H_1 is a Banach space with norm $|\cdot|_1$, which is compactly embedded into H. Assume the following integrability conditions.*

(i) *For every $t > 0$ and for every $\rho \in Pr(H)$ with $\int_H |x|^2\, d\rho(x) < \infty$ it holds that*

$$E\left(\int |\varphi(t,\omega)x|^2\, d\rho(x) < \infty\right).$$

(ii) *For P-almost all $\omega \in \Omega$ and for all $r > 0$ and for all $t > 0$ it holds that*

$$\sup\{|\varphi(t,\omega)x|_1 : |x| \leq r\} < \infty.$$

(iii) *There exists a real valued random variable $\omega \mapsto r(\omega)$ with $Er^2 < \infty$ such that $A(\omega) \subset B(r(\omega))$ a. s., where $B(r)$ denotes the ball of radius r in H.*

Then the set $\mathfrak{A}$ as given by (16) *is compact, satisfies $P_t\mathfrak{A} = \mathfrak{A}$ for every $t \geq 0$, and it attracts in the sense that, for every $r > 0$,*

$$\lim_{t\to\infty} d(P_t B_r, \mathfrak{A}) = 0$$

where B_r is the set of all $\rho \in Pr(H)$ with $\int_H |x|^2\, d\rho(x) < r$, and d denotes a metric metrising the topology of weak convergence on $Pr(H)$.

We will refer to the set $\mathfrak{A}$ defined in (16) as the *Schmalfuß attractor* associated with a white noise RDS φ with a global set attractor A.

We will now investigate the relation between the Schmalfuß attractor and general measure attractors. It will turn out that essentially the Schmalfuß attractor is the expectation of the minimal set Markov measure attractor.

For a set $M \subset Pr_\Omega(X)$ we write $\mathcal{Y}(M)$ for the set of Young measures in M. Young measures are those random measures on X which are Dirac measures supported by an X-valued random variable, or, more precisely,

$$\mathcal{Y}(M) = \{\nu \in M : \nu_\omega = \delta_{n(\omega)} \text{ a.s. for some random variable } n\}.$$

Clearly, $\mathcal{Y}(M) = \emptyset$ is not excluded.

5.2 Theorem *Suppose that φ is an RDS on a Polish space with a set attractor. Let $\mathcal{A}^{\mathcal{M}}$ be the minimal set Markov measure attractor given by Theorem* 3.15, *and define $\mathfrak{A}$ by* (16). *Then*

$$E(\mathcal{A}^{\mathcal{M}}) = \mathfrak{A},$$

where E denotes expectation.

Proof Since $\mathcal{A}$ is the set of all random measures supported by A, and $\mathcal{M}$ is the set of all random measures which are measurable with respect to $\mathscr{F}^-$, the set of Young measures in $\mathcal{A} \cap \mathcal{M}$ is given by

$$\mathcal{Y}(\mathcal{A} \cap \mathcal{M}) = \{\delta_b : b \text{ an } \mathscr{F}^-\text{-selection of } A\}.$$

Writing $\mathcal{Y}$ instead of $\mathcal{Y}(\mathcal{A} \cap \mathcal{M})$ for brevity, Lemma 3.7 implies

$$\Theta_t \mathcal{Y} = \{\Theta_t \delta_b : b \text{ an } \mathscr{F}^-\text{-selection of } A\},$$

where the disintegration of $\Theta_t \delta_b$ is given by $(\Theta_t \delta_b)_\omega = \delta_{\varphi(t,\vartheta_{-t}\omega)b(\vartheta_{-t}\omega)}$ almost surely (see Lemma 3.7). Now we use that Θ_t acts 'linearly' on $Pr_\Omega(X)$ – which can be stated more precisely for the present purposes by saying that Θ_t commutes with convex combinations – to conclude that

$$\Theta_t \mathrm{conv}\, \mathcal{Y} = \mathrm{conv}\, \{\Theta_t \delta_b : b \text{ an } \mathscr{F}^-\text{-selection of } A\}.$$

Since $\mathrm{conv}\, \mathcal{Y} \subset \mathcal{A}$, and $\mathcal{A}$ is compact, continuity of Θ_t on $Pr_\Omega(X)$ implies

$$\Theta_t(\mathrm{cl\ conv}\, \mathcal{Y}) = \mathrm{cl\ conv}\, \{\Theta_t \delta_b : b \text{ an } \mathscr{F}^-\text{-selection of } A\}.$$

The set of all convex combinations of Young measures is a dense subset of $\mathcal{A} \cap \mathcal{M}$, hence $\mathrm{cl\ conv}\, \mathcal{Y} = \mathcal{A} \cap \mathcal{M}$. Therefore we have

$$\Theta_t(\mathcal{A} \cap \mathcal{M}) = \mathrm{cl\ conv}\, \{\Theta_t \delta_b : b \text{ an } \mathscr{F}^-\text{-selection of } A\}.$$

Now we use that the expectation $E : Pr_\Omega(X) \to Pr(X)$ is continuous with respect to the narrow topologies (see Crauel [11] Corollary 4.27, and Remark 3.23), together with $f(\bar{M}) = \overline{f(M)}$ for f continuous and $\bar{M}$ compact, to conclude that

$$\begin{aligned}
&E\big(\Theta_t(\mathcal{A} \cap \mathcal{M})\big) \\
&\quad = E\Big(\mathrm{cl\ conv}\, \{\Theta_t \delta_b : b \text{ an } \mathscr{F}^-\text{-selection of } A\}\Big) \\
&\quad = \mathrm{cl\ conv}\, \{E(\Theta_t \delta_b) : b \text{ an } \mathscr{F}^-\text{-selection of } A\} \\
&\quad = \mathrm{cl\ conv}\, \{E\big(\varphi(t, \vartheta_{-t}\omega) b(\vartheta_{-t}\omega)\big) : b \text{ an } \mathscr{F}^-\text{-selection of } A\} \\
&\quad = \mathcal{D}_t,
\end{aligned}$$

using the notation from (16). Again invoking continuity of the expectation from $Pr_\Omega(X)$ to $Pr(X)$ together with the fact that $\big(\Theta_t(\mathcal{A} \cap \mathcal{M})\big)_{t\geq 0}$ forms a decreasing family of compact sets in $Pr_\Omega(X)$, we get

$$E\Big(\bigcap_{t\geq 0} \Theta_t(\mathcal{A} \cap \mathcal{M})\Big) = \bigcap_{t\geq 0} E\big(\Theta_t(\mathcal{A} \cap \mathcal{M})\big) = \bigcap_{t\geq 0} \mathcal{D}_t = \mathfrak{A}. \qquad \square$$

We note that in general $\mathfrak{A} \subsetneq E\mathcal{A}$, where $\mathcal{A}$, the global set measure attractor given by Theorem 3.11, is the set of all random probability measures supported by the set attractor of φ. In fact, in Example 3.17, $\delta_1 \in \mathcal{A}$ is a deterministic measure with $E\delta_1 = \delta_1 \in E\mathcal{A}$, whereas $\delta_1 \notin \mathcal{A}^{\mathcal{M}}$, and therefore also $\delta_1 \notin E\mathcal{A}^{\mathcal{M}}$. Here we used that for a random measure $\nu \in Pr_\Omega(X)$ one has that $E\nu$ is a Dirac measure (if and) only if ν_ω is almost surely constant and, in addition, this constant measure is a Dirac measure.

Note also that Theorem 5.2 is not restricted to white noise RDS, but that it holds for general RDS. As a first conclusion we obtain that for white noise RDS the assertion of Theorem 5.1 holds true under much more general conditions.

5.3 Corollary *If φ is a white noise RDS on a Polish space with a global set attractor, then the Schmalfuß attractor given by* (11) *exists, and it is a (and therefore* the*) set attractor for the Markov semigroup (P_t).*

Proof Immediate from Theorem 5.2 in view of the fact that $P_t \circ E = E \circ \Theta_t$ on $\mathcal{M} \subset Pr_\Omega(X)$. $\square$

5.4 Remark Corollary 5.3 implies that every compact set of Borel probability measures on the state space is attracted by the Schmalfuß attractor under the action of the Markov semigroup. In the case of an RDS on a separable Hilbert space H, however, Corollary 5.3 does not assert attraction of $B_r = \{\rho \in Pr(H) : \int_H |x|^2 \, d\rho(x) < r\}$. It seems that one would need more assumptions in the direction of those made by Schmalfuß [21] in order to derive Theorem 5.1. It is not clear which of the assumptions made there is needed in order to obtain that B_r, $r > 0$, is attracted by the Schmalfuß attractor.

5.5 Corollary *For a white noise RDS φ on a Polish space with a global set attractor suppose that, in addition, the induced Markov semigroup has a set attractor. Then this set attractor equals $E(\mathcal{A}^{\mathcal{M}})$, where $\mathcal{A}^{\mathcal{M}}$ is the minimal set Markov attractor given by* Theorem 3.15, *which in turn equals the Schmalfuß attractor.*

Proof Immediate from uniqueness of the set attractor for the Markov semigroup together with Theorem 5.2. $\square$

Next suppose that ρ is an invariant measure for the Markov semigroup P_t. Theorem 5.1 allows to infer $\rho \in \mathfrak{A}$ as soon as one knows that $\int |x|^2 \, d\rho(x) < \infty$. As another consequence of Theorem 5.2 we obtain that every P_t-invariant measure ρ is an element of the Schmalfuß attractor, also in case that $\int |x|^2 \, d\rho(x)$ is not finite.

The next question would be as to whether the closed convex set of all invariant Markov measures coincides with the Schmalfuß attractor. The answer is negative. This can be demonstrated by a deterministic example already.

5.6 Example Let $S^1 \subset \mathbb{R}^2$ denote the unit circle, identified with $\mathbb{R}/2\pi\mathbb{Z}$, equipped with the Riemannian structure inherited from the standard inner product on $\mathbb{R}$. Consider the deterministic flow $\varphi(t) : S^1 \to S^1$, given by $\varphi(t)x = x + t \bmod 1$. Every deterministic flow is a trivial white noise system, all σ-algebras being trivial. The global set attractor (which is also the global point attractor) is the whole state space S^1. The unique invariant (Markov) measure is the Lebesgue measure. The set

$$\{E(\delta_{\varphi(t,\vartheta_{-t}\omega)b(\vartheta_{-t}\omega)}) : b \text{ an } \mathscr{F}^{-}\text{-selection of } A\}$$

here simply is $\{\delta_{\varphi(t)b} : b \in S^1\} = \{\delta_x : x \in S^1\}$. The closed convex hull of this set is the set of all Borel probability measures on S^1, independent of t. The Schmalfuß attractor is thus the whole $Pr(S^1)$. The set of invariant measures consists, however, of just one element, the Lebesgue measure.
The example can easily be modified to obtain a system where S^1 becomes (part of) a nontrivial attractor. Take, e. g., the system in $\mathbb{R}^2$ given by $\dot{x} = (h(|x|)\mathrm{Id} + C)x$, where Id is the identity matrix, C is the 2×2-matrix $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, and

$$h(r) = \begin{cases} \sin \pi r & \text{for } 0 \le r \le \frac{3}{2} \\ -1 & \text{for } \frac{3}{2} \le r \end{cases}$$

Then $h(1) = 0$, and the system restricted to S^1 is just the one discussed above. The origin is an unstable equilibrium now, whereas S^1 is a local set attractor. The global point attractor becomes $S^1 \cup \{0\}$, and the global set attractor becomes $D^1 = \{x \in \mathbb{R}^2 : |x| \le 1\}$. There are two ergodic invariant (Markov) measures now, namely the Lebesgue measure on S^1 and δ_0. The Schmalfuß attractor is the set of all convex combinations of $Pr(S^1)$ and δ_0.

However, if one takes the still rather degenerate system given by $\varphi(t,\omega)x = x + W(t,\omega) \bmod 1$ on S^1, where W is a Brownian motion (which is induced by the elliptic SDE $dx = dW(t)$ on S^1), one obtains a non-degenerate white noise system. This system again has the Lebesgue measure as the unique invariant – in fact, fixed – measure. The global attractor is S^1 again. We only sketch the argument: Now $E(\delta_{\varphi(t,\vartheta_{-t}\omega)b})$ converges to the Lebesgue measure on S^1 for every $b \in S^1$. The convergence is uniform in $b \in S^1$, so the convex combination over all $b \in S^1$ is close to the Lebesgue measure for t large. Therefore the Schmalfuß attractor has one element only, which is the Lebesgue measure on S^1.

There is a simple case, however, in which the Schmalfuß attractor automatically coincides with the set of all Markov measures. This is the case where the set attractor $\omega \mapsto A_{\mathrm{set}}(\omega)$ consists of just one point, i. e., $A_{\mathrm{set}}(\omega) = \{a(\omega)\}$ almost surely. Now all invariant measures are supported by the attractor, see Theorem 2.9. So in this case there is

only one invariant measure for φ, which is the random Dirac measure $\omega \mapsto \delta_{a(\omega)}$. Since the attractor, and hence a, is $\mathscr{F}^-$-measurable by Theorem 2.5, this invariant measure is a Markov measure. By Proposition 4.2, the deterministic measure $\rho = E\delta_a$, which is nothing but the distribution of the random variable a, is an invariant measure for Markov semigroup.
On the other hand, the Schmalfuß attractor is obtained by taking first the closed convex hull of the set of all $E\delta_{\varphi(t,\vartheta_{-t})b(\vartheta_{-t}\omega)}$, where b ranges over all $\mathscr{F}^-$-measurable selections of A_{set}, and then taking the intersection of all these sets for nonnegative t. But here there is only one selection of $\omega \mapsto A_{\text{set}}(\omega) = \{a(\omega)\}$, which is $\omega \mapsto a(\omega)$. Since with A_{set} a forteriori also a is invariant, i. e., $\varphi(t,\omega)a(\omega) = a(\vartheta_t\omega)$ a. s. for all $t \in T$. Shifting this identity for t fixed by ϑ_t gives $\varphi(t,\vartheta_{-t}\omega)a(\vartheta_{-t}\omega) = a(\omega)$ a. s., whence

$$E\delta_{\varphi(t,\vartheta_{-t})b(\vartheta_{-t}\omega)} = E\delta_{a(\omega)} \equiv \rho,$$

independently of t. Thus in the present situation the Schmalfuß attractor is nothing but the one point set $\{\rho\}$. This observation reproves Theorem 5.3 of Schmalfuß [21] without having to make use of Theorem 3.4 of [21].

We give another example of a white noise RDS whose set measure attractor does not coincide with its minimal point measure attractor, and which has an invariant measure which is not supported by the minimal point attractor.

5.7 Example The stochastic differential equation

$$d\xi = \sin(\xi) \circ dW_1(t) + \cos(\xi) \circ dW_2(t) \tag{17}$$

defines an RDS on S^1, the generator of which is $\Delta/2$, so the solutions of (17) are Brownian motions on S^1. The invariant Markov measure is a random Dirac measure in a point $\omega \mapsto a(\omega)$, say, and $\varphi(t,\vartheta_{-t}\omega)x$ converges to $a(\omega)$ almost surely, for every fixed $x \in S^1$. Thus, the random one point set made of $a(\omega)$ is a point attractor – in fact, the minimal point attractor – for the RDS induced by (17). Denoting $A_{\text{pt}}(\omega) = \{a(\omega)\}$, clearly $\big\{\nu \in Pr_\Omega(S^1) : \nu(A_{\text{pt}}) = 1\big\}$ contains only one element, which is the Young measure whose disintegration is given by $\delta_{a(\omega)}$. However, there exists another invariant measure for φ, which is not a Markov measure, and which thus does not coincide with the invariant Markov measure. This follows since φ does not leave a Borel measure ρ on S^1 fixed (in the sense that $\varphi(t,\omega)\rho = \rho$ almost surely for some $t > 0$). Existence of another invariant measure is then a (not completely immediate) consequence of Theorem 4.2 and Corollary 4.4 of Baxendale [4], which can also be inferred from Crauel [7] Theorem 8.5; see also Section 8.6.1 of [7], and Crauel [8] Corollary 5.4 and Remark 5.6. This invariant measure cannot be an element of $\{\nu \in Pr_\Omega(S^1) : \nu(A_{\text{pt}}) = 1\}$, the set of all probability measures supported by the minimal point attractor of φ.

References

[1] L. Arnold, *Random Dynamical Systems*, Springer, New York 1998

[2] P. Ashwin and G. Ochs, Convergence to local random attractors, *Dynamical Systems* 18 (2003) 139–158

[3] P. Baxendale, Brownian motions in the diffeomorphism group I, *Compositio Mathematica* 53 (1984) 19–50

[4] P. Baxendale, Lyapunov exponents and relative entropy for a stochastic flow of diffeomorphisms, *Probab. Theory Rel. Fields* 81 (1989) 522–544

[5] M. Capiński and N. J. Cutland, Measure attractors for stochastic Navier-Stokes equations, *Electronic Journal of Probability* 3, paper no. 8 (1998) 1–15

[6] C. Castaing and M. Valadier, *Convex Analysis and Measurable Multifunctions*, Lecture Notes in Mathematics 580, Springer, Berlin 1977

[7] H. Crauel, Extremal exponents of random dynamical systems do not vanish, *J. Dyn. Differ. Equations* 2 (1990) 245–291

[8] H. Crauel, Non-Markovian invariant measures are hyperbolic, *Stochastic Processes Appl.* 45 (1993) 13–28

[9] H. Crauel, Global random attractors are uniquely determined by attracting deterministic compact sets, *Ann. Mat. Pura Appl., IV. Ser.*, Vol. CLXXVI (1999) 57–72

[10] H. Crauel, Random point attractors versus random set attractors, *J. London Math. Soc., II. Ser.*, 63 (2001) 413–427

[11] H. Crauel, *Random Probability Measures on Polish Spaces*, Series *Stochastics Monographs*, Volume 11, Taylor & Francis, London and New York 2002

[12] H. Crauel, Luu Hoang Duc, and S. Siegmund, Towards a Morse Theory for Random Dynamical Systems, *Stochastics and Dynamics* 4 (2004) 277–296

[13] H. Crauel and F. Flandoli, Attractors for random dynamical systems, *Probab. Theory Relat. Fields* 100 (1994) 365–393

[14] N. J. Cutland and H. J. Keisler, Global attractors for 3-dimensional stochastic Navier-Stokes equations, *Journal of Dynamics and Differential Equations* 16 (2004) 205–266

[15] S. N. Ethier and T. G. Kurtz, *Markov Processes, Characterization and Convergence*, Wiley, Chichester 1989

[16] H. Keller and G. Ochs, Numerical approximation of random attractors, pp. 93–115 in *Stochastic Dynamics*, H. Crauel and V. M. Gundlach (eds), Springer, New York 1999

[17] H. Keller and B. Schmalfuß, Attractors for stochastic differential equations with nontrivial noise, *Bul. Acad. Stiinte Repub. Mold., Mat.* 26 (1998) 43–54

[18] H. Morimoto, Attractors of probability measures for semilinear stochastic evolution equations, *Stochastic Anal. Appl.* 10 (1992) 205–212

[19] K. R. Schenk-Hoppé, Random Attractors – General Properties, Existence and Applications to Stochastic Bifurcation Theory, *Discrete Contin. Dyn. Syst.* 4 (1998) 99–130

[20] B. Schmalfuß, Stochastische Attraktoren des stochastischen Lorenz-Systems, *Z. Angew. Math. Mech.* 74, No. 6 (1994) T 627–T 628

[21] B. Schmalfuß, Measure attractors and random attractors for stochastic partial differential equations, *Stochastic Anal. Appl.* 17 (1999) 1075–1101