Imperial College London
Department of Mathematics

Synchronisation in Random Dynamical Systems

Julian Magnus Ike Newman

A thesis presented
for the degree of Doctor of Philosophy
at Imperial College London.

Declaration

I certify that the research documented in this thesis is entirely my own research, and that all ideas in this thesis which either originate from other people or are the fruit of my discussions with other people have been explicitly acknowledged as such.

Copyright

Abstract

In this thesis, we develop a deeper and much more extensive theory of synchronisation of trajectories of random dynamical systems (RDS) than currently exists. In particular, focusing on random dynamical systems with memoryless noise, we achieve two main goals: Firstly, we demonstrate that the notion of "statistical equilibria" is purely a property of the measurable dynamics of a RDS on a standard Borel space; and yet, within such statistical equilibria is "encoded" the phenomenon of noise-induced synchronisation (which may then be observed in *any* compatible metric on the phase space). Secondly, we provide new, widely applicable criteria for synchronisation in RDS, considerably improving upon some of the existing criteria for synchronisation.

Acknowledgements

I would like to thank my supervisors, Dr Martin Rasmussen and Prof Jeroen Lamb, for all their support and guidance throughout my PhD. In addition, I would like to thank the following people for very helpful and interesting discussions: Dr Doan Thai Son, Prof Thomas Kaijser, Prof Dan Crisan, Prof Gerhard Keller, Prof Gary Froyland, Prof Franco Flandoli, Prof Hans Crauel, Dr Thomas Cass, Prof Ken Palmer, Prof Ale Jan Homburg, Prof Michael Scheutzow, Prof Anthony Quas, Prof Maxim Arnold, Prof Sanjeeva Balasuriya, Dr Anna Maria Cherubini, and Maximilian Engel. I would also like to thank Prof Nick Bingham for the several occasions when he has sacrificed time to help give me general guidance and advice.

I would like to thank Ai Guan for her immense help with the typesetting of this document.

Finally, I would like to thank my family and my church for all their encouragement, support and prayers.

The research documented in this thesis has been supported by an EPSRC Doctoral Training Account.

"Before the mountains were brought forth,
or ever thou hadst formed the earth and the world,
even from everlasting to everlasting, thou art God."
Psalm 90:2

A notice concerning other concurrent research

While carrying out the research on circle RDS documented in Section 5.2 of this thesis, the author was unaware that other research was being carried out on the same topic (seemingly at more-or-less the same time), the results of which were documented in a preprint in December 2014.[1] Although working specifically in discrete time, this preprint contains several remarkable results, including conditions for stable synchronisation in invertible RDS on the circle that are weaker than our conditions presented in Theorem 5.19 of this thesis: specifically, if there are no deterministic fixed points then stable synchronisation is equivalent to contractibility.[2] Moreover, remarkably, in this case we are guaranteed to have *exponential-rate* stable synchronisation. Results are also obtained for RDS on a bounded interval (again involving exponential-rate stability), with partial overlap with Theorem 3.18 of this thesis.

[1] Dominique Malicet. "Random walks on Homeo(S^1)". arXiv:1412.8618v1 [math.DS]. 2014. (See, in particular, Theorems A and E.)

[2] Here, contractibility is formulated in terms of the possibility for two points to come arbitrarily close together; but as we prove in Proposition 4.68 this thesis, it is sufficient just to show that two distinct points are able to come strictly closer than their initial separation.

Contents

Notations and terminology

We write $\mathbb{N}$ for the strictly positive integers, and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ for the nonnegative integers.

For a finite set P, $|P|$ denotes the number of elements of P.

A *partial order* $\leq$ on a set X is a binary relation that is transitive and has the property that for all $x, y \in X$,

$$x = y \quad \Leftrightarrow \quad (x \leq y \text{ and } y \leq x).$$

A *linear order* (or *total order*) $\leq$ on a set X is a partial order on X with the additional property that for all $x, y \in X$, either $x \leq y$ or $y \leq x$.

Given a smooth map $f{:}M \to N$ between smooth manifolds M and N, we write $\mathrm{d}f{:}TM \to TN$ for the derivative of f, and $(\mathrm{d}f)_x{:}T_xM \to T_{f(x)}N$ for the restriction of $\mathrm{d}f$ as mapping between T_xM and $T_{f(x)}N$.

Given a topological space X, $\mathcal{B}(X)$ denotes the Borel σ-algebra of X. Given a measurable space (X, Σ), we interchangeably use the phrases "probability measure on X" and "probability measure on (X, Σ)"; given a topological space X, a "probability measure on X" always means a *Borel* probability measure on X (i.e. a probability measure on $(X, \mathcal{B}(X))$).

Given a metric space (X, d), a point $x \in X$, and a number $\delta > 0$, we write $B_\delta(x) := \{y \in X : d(x, y) < \delta\}$ and $\bar{B}_\delta(x) := \{y \in X : d(x, y) \leq \delta\}$; this notation obviously makes implicit reference to the underlying metric space (X, d) from which the point x is taken.

We always use the term "neighbourhood" to refer to an *open* neighbourhood. (Nonetheless, we will sometimes use the phrase "open neighbourhood" to emphasise this.)

Given a measure space $(\Omega, \mathcal{F}, m)$, a measurable space (X, Σ), and a measurable map $f{:}\Omega \to X$, we write f_*m for the image measure of m under f (that is, $f_*m(A) := m(f^{-1}(A))$ for all $A \in \Sigma$). Given an m-integrable function $g{:}\Omega \to \mathbb{R}$, we sometimes write $m(g)$ to mean $\int_\Omega g\, dm$. (So $m(E) = m(\mathbb{1}_E)$ for any $E \in \mathcal{F}$.)

Given a collection $\mathcal{C}$ of subsets of some set Ω, we write $\sigma(\mathcal{C})$ for the σ-algebra on Ω generated by $\mathcal{C}$. Note that this notation makes implicit reference to the underlying set Ω; still, when using this notation, it will always be clear what the underlying set is. Given a family $(\mathcal{F}_\alpha)_{\alpha \in I}$ of σ-algebras $\mathcal{F}_\alpha$ on Ω, we write $\sigma(\mathcal{F}_\alpha : \alpha \in I)$ for the smallest σ-algebra on Ω containing the σ-algebra $\mathcal{F}_\alpha$ for every α. Given a family $((X_\alpha, \Sigma_\alpha))_{\alpha \in I}$ of measurable spaces $(X_\alpha, \Sigma_\alpha)$ and a family $(f_\alpha)_{\alpha \in I}$ of functions $f_\alpha{:}\Omega \to X_\alpha$, we write $\sigma(f_\alpha : \alpha \in I)$ for the smallest σ-algebra on Ω with respect to which the function f_α is measurable for every $\alpha \in I$.

We write $\mathbb{E}_{(\rho)}[\cdot]$ to denote *expectation* with respect to a probability measure ρ; in cases where the underlying probability measure happens to be denoted $\mathbb{P}$, we will sometimes omit the subscript $_{(\mathbb{P})}$ and just write $\mathbb{E}[\cdot]$. Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a $\mathbb{P}$-integrable function $g : \Omega \to \mathbb{R}$ and a sub-σ-algebra $\mathcal{G}$ of $\mathcal{F}$, we write $\mathbb{E}[g|\mathcal{G}] : \Omega \to \mathbb{R}$ to denote an (arbitrary) version of the conditional expectation given $\mathcal{G}$ of g with respect to $\mathbb{P}$. Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a set $E \in \mathcal{F}$, and a sub-σ-algebra $\mathcal{G}$ of $\mathcal{F}$, we write $\mathbb{P}(E|\mathcal{G}) : \Omega \to [0,1]$ to denote an (arbitrary) version of the conditional probability given $\mathcal{G}$ of E according to $\mathbb{P}$; that is to say, $\mathbb{P}(E|\mathcal{G})$ is a $\mathcal{G}$-measurable function satisfying

$$\mathbb{P}(E \cap G) = \int_G \mathbb{P}(E|\mathcal{G})(\omega)\, \mathbb{P}(d\omega) \quad \forall\, G \in \mathcal{G}.$$

Note that $\mathbb{P}(E|\mathcal{G}) \overset{\mathbb{P}\text{-a.s.}}{=} \mathbb{E}[\mathbb{1}_E|\mathcal{G}]$.

For any statement P, we define

$$\chi_P := \begin{cases} 1 & \text{if } P \text{ is true} \\ 0 & \text{if } P \text{ is false.} \end{cases}$$

Given a measurable space (X, Σ) and a point $x \in X$, we write δ_x for the *Dirac mass* at x, namely the probability measure on (X, Σ) defined by $\delta_x(A) = \mathbb{1}_A(x)$ for all $A \in \Sigma$. Obviously, if the singletons in X are Σ-measurable, then Dirac masses at distinct points are distinct measures.

As will be introduced in the body of the thesis, when considering stochastic processes and random dynamical systems, the symbol $\mathbb{T}$ is used to refer to $\mathbb{Z}$ in discrete time and $\mathbb{R}$ in continuous time; and $\mathbb{T}^+$ will denote $\mathbb{T} \cap [0, \infty)$. Given a topological space X, a $\mathbb{T}^+$-indexed family $(x_t)_{t \in \mathbb{T}^+}$ of elements of X, a point $x \in X$, and a set $S \subset \mathbb{T}$ that is unbounded above, we say that *$x_t \to x$ as t tends to ∞ in S* if for every neighbourhood U of x there exists $T \in \mathbb{T}^+$ such that for all $t \in S$ with $t \geq T$, $x_t \in U$.

Chapter 1. Introduction

In this introduction, we will first give a summary of our main results and an overview of the thesis; we will then give a general introduction to the topic of synchronisation in random dynamical systems (RDS) and an overview of some existing results (including how they relate to our results); finally, we will present a (non-rigorous) exposition of one of the key results in the theory of RDS that is tied to much of the work in this thesis, namely the correspondence between two-sided-time and one-sided-time invariant measures. (Still, all of the theory of RDS that is needed for this thesis *will* be introduced rigorously in subsequent chapters.)

1.1 Summary of main results

In this thesis, we study synchronisation of trajectories of memoryless time-homogeneous random dynamical systems. The vast majority of studies on synchronisation in RDS have either been particular case studies, or slightly more general results giving sufficient conditions for synchronisation, in either RDS taking a particular form or RDS satisfying some special property. By contrast, the aim of this thesis is to develop much more general results on synchronisation in RDS. In particular, we will provide:

(A) deep insights concerning asymptotic statistics and synchronising behaviour in measurable RDS;

(B) new, broadly applicable criteria for synchronisation (in topological RDS) that are weaker and easier to verify than criteria given in previous literature.

We will consider several different modes of synchronisation of trajectories; and in contrast to virtually all previously existing results, each of our new tests for synchronisation provides conditions that are either *necessary and sufficient* for the mode of synchronisation in question, or at least *become* necessary and sufficient when a mild additional assumption is made.

We now give a summary of our main results. Here, φ is a memoryless-noise RDS on a phase space X, and ρ is a stationary probability measure for the associated Markov transition probabilities of the one-point motion. A "right-continuous RDS" is a RDS that is jointly continuous in space and right-continuous in time.

Our main general results concerning measurable RDS are:

- Theorem 3.6: Taking X to be a standard measurable space, we show that there is a probability measure Q_ρ on the space of probability measures on X with the property that under any metrisation of the measurable structure of X, the limiting distribution of the measure-valued Markov process obtained by letting ρ evolve under the flow of φ is precisely Q_ρ. (We also prove some important further properties of Q_ρ.) Q_ρ is called the *statistical equilibrium* associated to ρ. In previous literature, it is only when working with a topology in which φ is (spatially) continuous that a limiting distribution Q_ρ has been obtained; by contrast, we remove all continuity requirements, and show that the limiting distribution Q_ρ is a measurable invariant.[1] Foundational to the proof of Theorem 3.6 is Theorem 3.49, where we prove the one-to-one correspondence between stationary probability measures and Markov invariant measures without any continuity requirements.

- Corollary 3.9: Taking X to be a standard measurable space and taking ρ to be ergodic, we show that there is a (deterministic) number $n \in \mathbb{N} \cup \{\infty\}$ such that under any metrisation of the measurable structure of X, when one observes how all the trajectories evolve simultaneously under the same realisation of the noise, one finds that either:

  (a) [the case that $n < \infty$] after a long time, the trajectories of a very large proportion (according to ρ) of the initial conditions have separated out (in equal proportions) into n "clusters" of very small diameter; or

  (b) [the case that $n = \infty$] there is no significant synchronisation phenomenon in the asymptotic dynamics.

  (This is formalised rigorously, using notions based on convergence in probability.) We refer to n as the *ρ-clustering number* of φ. In the case that $n = 1$ (meaning that a very large proportion of the trajectories of φ concentrate into a very small region), we say that φ is *statistically synchronising with respect to ρ*. Once again, results akin to Corollary 3.9 have been obtained when restricting to a topology in which φ is continuous (see [LeJ87]); but we remove all continuity requirements and show that the ρ-clustering number of φ is a measurable invariant.

Our main general tests for synchronisation in right-continuous RDS are:

- Theorem 6.1: Taking X to be a compact metric space and taking φ to be right-continuous, we have almost sure synchronisation of the trajectories of any given pair of initial conditions, together with almost sure local asymptotic stability of any given initial condition, if and only if the following statements hold:

  (i) there is a unique (deterministic) minimal set $K \subset X$;

  (ii) for any two distinct initial conditions in K, there is a positive probability that the subsequent trajectories will at some time be closer together than their initial separation;

  (iii) with positive probability there exist locally asymptotically stable initial conditions in K.

  (For RDS on a manifold, condition (iii) is typically verified by showing that the maximal Lyapunov exponent associated to some ergodic probability measure supported by K is negative.)

[1] This fact is discussed in the introduction of [New15b] and also mentioned in the author's open problem in [GGTQ15]; but the author has not published a full statement and proof before now.

- Theorem 6.6: We essentially answer the question, "When do we have $n = 1$ in Propositions 2 and 3 of [LeJ87]?" More precisely, taking X to be a Borel subset of a Polish space, taking φ to be right-continuous, and taking ρ to be ergodic, we provide necessary and sufficient conditions for the phenomenon that there is almost surely a ρ-full-measure open set of initial conditions whose trajectories are asymptotically stable and synchronise with each other. Our conditions for this phenomenon involve (i) the support of ρ admitting locally asymptotically stable trajectories, and (ii) a condition concerning the ability of pairs of trajectories to simultaneously come close to at least one "typical" point within the support of ρ.

Our main results regarding synchronisation in monotone RDS on linearly ordered spaces[2] are:

- Theorem 3.13: Taking X to be a standard measurable space, equipped with a Borel linear order that is preserved by φ, and taking ρ to be ergodic, we show that φ is always statistically synchronising with respect to ρ.
- Theorem 3.18: Taking X to be a Borel subset of $\bar{\mathbb{R}}$ and taking φ to be monotone, we show that if ρ is ergodic then there is an "attracting random fixed point" whose law is ρ. (As a consequence, we obtain easily verifiable necessary and sufficient conditions for a stationary measure ρ to be ergodic.)
- Proposition 4.59: We take X to be a Borel subset of $\bar{\mathbb{R}}$ and take φ to be monotone, and we suppose that there exists an ergodic probability measure ρ such that the only ρ-full-measure interval in X is the whole of X. Under this assumption, we show that all trajectories synchronise almost surely if and only if there exist (with positive, or equivalently, with full probability) locally asymptotically stable trajectories in X.

Our main result regarding synchronisation in orientation-preserving invertible RDS on the circle is:

- Theorem 5.19: Taking X to be the circle $\mathbb{S}^1$, and taking φ to be a right-continuous RDS with $\varphi(t,\omega)$ being an orientation-preserving homeomorphism for all t and ω, we show that if

  (i) there are no deterministic fixed points; and
  (ii) for any ordered pair of distinct initial conditions in $\mathbb{S}^1$, there is a positive probability that the anticlockwise separation of their subsequent trajectories will, at some point in time, be less than their original anticlockwise separation;

  then almost sure synchronisation of the trajectories of any given pair of initial conditions occurs, together with almost sure local asymptotic stability of any given initial condition. These sufficient conditions are also necessary in the case that there are no deterministic non-empty open proper subsets of $\mathbb{S}^1$ that are almost surely forward-invariant under φ. (If such an open invariant set *does* exist, it may be possible to reduce the question of synchronisation to the question of whether synchronisation occurs on this set; see [New15c, Proposition 2.16].)

[2] The topic of synchronisation in monotone RDS on partially ordered spaces is an important one; for a deep study on this topic, see [FGS15]. Nonetheless, we do not consider this topic here.

Intuitively, Theorem 5.19 guarantees that for any orientation-preserving invertible RDS on $\mathbb{S}^1$ with "enough noisiness", synchronisation is guaranteed to occur. Our main application of Theorem 5.19 is:

- Theorem 5.25: We consider the RDS φ on $\mathbb{S}^1$ formed by projecting onto $\mathbb{S}^1$ the solutions of the SDE
$$dX_t = b(X_t)dt + \sigma dW_t$$
on $\mathbb{R}$, where $b{:}\mathbb{R} \to \mathbb{R}$ is a 1-periodic Lipschitz function and (W_t) is a Wiener process. We show that if $\sigma \neq 0$ and 1 is the least period of b, then φ is guaranteed to exhibit the "stable synchronisation" phenomenon described in Theorem 5.19 (and in Theorem 6.1).

(A simple example is discussed where it is found that additive noise destroys a saddle-node bifurcation.)

Structure of the thesis

In the remainder of Chapter 1, we will first present an introduction to the topic of synchronisation in RDS, and an overview of the existing results and how they relate to our results. We will then give an exposition of the well-known correspondence between one-sided-time and two-sided-time invariant measures (of which the correspondence between Markov invariant measures and stationary measures is a particular case), including an explanation of the new contribution to this topic made by Theorem 3.49. This exposition is worth providing, because (i) the correspondence between one-sided-time and two-sided-time invariant measures is one of the most fundamental results in the theory of random dynamical systems, and (ii) this correspondence (or rather, the particular case of it for Markov invariant measures) forms the basis of our results on statistical equilibria, and consequently also of our test for ρ-almost stable synchronisation (Theorem 6.6).

In Chapter 2, we introduce random dynamical systems formally. (We specifically consider memoryless[3] RDS.) We provide some basic examples, and extensively develop the foundational material that will be needed later on in the thesis. There does not currently exist a general in-depth exposition of RDS with the memoryless property. Therefore, while all the non-trivial concepts and results presented in Chapter 2 have already appeared in some form in previous literature (except perhaps some of the results in Section 2.9, as well as some of the peripheral results in Section 2.7), several of these concepts and results have not been formulated as rigorously and in as much generality as we will do here.

In Chapter 3, we develop the theory of statistical equilibria, clustering numbers, and statistical synchronisation, and also apply this to monotone RDS. The concept of statistical equilibria has already been established within the setting of continuous RDS on a pre-defined topological space. (Likewise, therefore, the notions of "clustering numbers" and "statistical synchronisation" have also been studied within this framework, although not under these names. Nonetheless, a precise mathematical description of the

[3] Our formalism of the "memorylessness" property is similar to that in [FGS14].

"clustering" phenomenon, as provided by Corollary 3.9, does not appear to have been provided before now.) Our key new contribution is to show that these concepts are not topology-specific but concern the dynamical behaviour of measurable RDS. Also, while fairly general results regarding synchronisation in monotone RDS already exist, we provide the most general existing results for the presence of attractive random fixed points in monotone RDS on linearly ordered spaces.

In Chapter 4, we introduce some further notions of synchronisation and local stability: We say that a RDS is *synchronising* (without further qualification) to mean the trajectories of any given pair of initial conditions synchronise almost surely. Given a probability measure ρ on the phase space, we say that a RDS is *ρ-almost everywhere synchronising* to mean that there is a ρ-full set A such that the trajectories of any given pair of initial conditions in A synchronise almost surely. We also introduce the notions of Lyapunov and asymptotic stability. (Our definition of asymptotic stability is different from the conventional definition, but is "very nearly" equivalent[4] and much easier to work with.) Using the notion of asymptotic stability, we define what it means for a RDS to be *stable with respect to ρ* (where ρ is an ergodic measure of the Markov transition probabilities), and we define what it means for a RDS to be *pointwise-stably synchronising*, (*uniformly*) *stably synchronising*, and *ρ-almost stably synchronising*. (For a RDS on $\mathbb{R}^d$ or a more general manifold, stability with respect to ρ is typically verified by showing that the maximal Lyapunov exponent associated to ρ is negative.) The notions of stable synchronisation and pointwise-stable synchronisation are important "improvements" on the more general notion of a synchronising RDS, since (i) they overcome potential problems related to instability of trajectories, and (ii) they appear to be the more "mathematically natural" notions to consider (as suggested by Theorem 5.19 and Theorem 6.1—there do not appear to exist similarly simple characterisations of when a RDS is merely "synchronising"). Likewise, the notion of ρ-almost stable synchronisation is an important improvement on the notion of ρ-almost everywhere synchronisation. We also present the most general existing result on forward-time synchronisation in monotone RDS on subsets of $\bar{\mathbb{R}}$. Finally, we briefly consider "synchronisation at a deterministic rate", and explain that although noise can create synchronisation, there will never be an almost sure upper bound on how long one has to wait for such synchronisation to occur. As in Chapter 2, most of the non-trivial results in this chapter are already understood conceptually; however, our set of definitions for the different modes of forward-time synchronisation is new, and most of the results here have not been formulated in the level of generality that we do. (The result on synchronisation in monotone RDS, namely Proposition 4.59, is also new, although it is conceptually only a slight extension of already understood facts.)

In Chapter 5, we carry out an in-depth study of stable synchronisation in orientation-preserving RDS on the circle. We first provide a geometrical characterisation of stable synchronisation (Theorem 5.6) in terms of "crack points" (a notion adapted from [Kai93]). We then provide our main test for stable synchronisation, which is essentially the "generalised form" of results in [DKN07, Section 5.1]. As an application, we give the

[4]Indeed, by [New15b, Theorem A11, Remarks A9 and A13], it is precisely equivalent in the case of a fixed point of a continuous (deterministic) dynamical system on a locally compact metric space.

first complete description of synchronisation in Wiener-driven additive-noise SDE on the circle with Lipschitz drift.

In Chapter 6, we present our test for stable synchronisation on compact spaces, and our test for ρ-almost stable synchronisation on general Polish spaces (or Borel subsets thereof). Once again, these tests are new, and provide considerable improvement on existing tests for almost sure forward-time synchronisation.

In Appendix A, we present various concepts and facts from measure theory and probability theory that will be used throughout the thesis. (In particular, we provide an exposition of the narrow topology.) In Appendix B, we present fundamental results concerning the "topology of uniform convergence on compact sets". In Appendix C, we introduce some foundational ergodic theory for both dynamical systems and Markov transition probabilities. In none of these appendices do we intend to provide full expositions of the subjects in question; the main aim is simply to present some of the key facts that will be needed in the thesis.

1.2 An overview of synchronisation in random dynamical systems

Noise-induced synchronisation

Let us motivate the whole study of synchronisation in random dynamical systems with an examplary scenario.

Suppose we have an array of identical non-interacting self-oscillators,[5] where the time-evolution of each oscillator is governed by the differential equation

$$\dot{x} \;=\; b(x) \tag{1.1}$$

for some vector field b on $\mathbb{R}^d$ (where $\mathbb{R}^d$ represents the space of possible "states" of one oscillator). Suppose these oscillators start at different states from each other. Assuming that the oscillators are not purely dissipative, since there is no interaction between the oscillators, there is obviously no reason for the oscillators to ever synchronise with each other.

But now suppose we subject all the oscillators simultaneously to some external forcing (which acts equally on all the oscillators); for example, we can suppose that the time-evolution of each oscillator is now governed by the equation

$$x(t) - x(0) \;=\; \int_0^t b(x(s))\,ds + F(t) \tag{1.2}$$

[5]The term "self-oscillator" refers generally to any oscillatory physical system for which, even in the absence of external driving forces, the total energy does not have to be a monotonically decreasing function of time. Oscillators whose energy is constrained to decrease over time will normally settle towards an equilibrium state, but self-oscillators need not do so.

for some function $F\colon[0,\infty) \to \mathbb{R}^d$ with $F(0) = \mathbf{0}$. A natural question to ask is whether $F(t)$ can be chosen in such a manner that the oscillators will eventually synchronise with each other. To be more precise: can $F(t)$ can be chosen in such a manner that the difference in state between any two of these oscillators tends to 0 as time tends to ∞? Not surprisingly, in many cases the answer to this question is *yes*.

But now suppose that the external forcing is not deterministic, but *random*. Indeed, let us suppose that this external forcing is a completely memoryless random process, and that its statistical properties do not change over time. (For example, the external forcing could consist of a sequence of "i.i.d. random kicks", where the time-separations between consecutive kicks are i.i.d. exponentially distributed random variables.) We now regard the time-evolution of the oscillators as being governed by an equation of the form

$$x(t) - x(0) \; = \; \int_0^t b(x(s))\,ds \; + \; F_\omega(t) \tag{1.3}$$

where the function F_ω depends on a sample point ω drawn randomly from some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. (The "random kicks" example is precisely the situation that the stochastic process $(F_\omega(t))_{t\geq 0}$ is a compound Poisson process.) The question now becomes: *is it possible that with strictly positive probability, or even with full probability, the processes will synchronise with each other?* Remarkably, the answer is often *yes*.

The phenomenon that processes can be caused to synchronise with each other due to being subjected to the same external random forcing is known as *noise-induced synchronisation*. In mathematical models, this phenomenon will typically appear as the phenomenon that as time tends to infinity, the difference in state between any two of the processes converges to 0, either in probability or under a positive-measure (often full-measure) set of sample paths of the external random forcing.

Noise-induced synchronisation was first described in the 1980s (see [Pik84], which considers synchronisation of non-interacting self-oscillators subjected to a sequence of random kicks,[6] or [Ant84], which considers synchronisation of cyclic phenomena). And since then, there have been numerous case studies of the phenomenon (see e.g. [Tor+01] and the references therein), as well as some general rigorous theoretical results, which will be described later.

Random dynamical systems

Just as equation (1.1) generates a *flow* on $\mathbb{R}^d$, and equation (1.2) generates a *non-autonomous flow* on $\mathbb{R}^d$, so equation (1.3) generates a *random dynamical system* (RDS) on $\mathbb{R}^d$. A random dynamical system is a dynamical system that is not deterministic but influenced by a random process (which we refer to as the "noise").

The natural mathematical framework within which to study the phenomenon of noise-induced synchronisation is precisely the framework of random dynamical systems. In this framework, the question of noise-induced synchronisation becomes the question of when

[6] For later work on the same topic, see e.g. [Nak+05].

different trajectories of a RDS converge towards each other (under the same realisation of the noise).

Memoryless RDS and Markov processes

When the term "random dynamical system" is used, it is normally assumed that the RDS in question is "time-homogeneous" in the sense that it can be defined without reference to any kind of "external clock". For a deterministic dynamical system, "time-homogeneity" would mean that the dynamical system is an *autonomous dynamical system*; for a random dynamical system, "time-homogeneity" means that the following two statements hold:

(a) the precise rule specifying how the behaviour of the system is determined by the behaviour of the noise does not change over time: in other words, writing $\varphi(t,\omega)x$ to denote the position at time t of the trajectory whose position at time 0 is x when the realised behaviour of the noise is given by ω, and writing $\theta^s\omega$ to denote the time-shifted version of ω forward through time s, we have that $\varphi(t,\theta^s\omega)x$ is the position at time $s+t$ of the trajectory whose position at time s is x when the realised behaviour of the noise is given by ω;

(b) the probability distribution for the precise behaviour of the noise is invariant under any time-shift.

(A noise process satisfying property (b) is said to be (*strictly*) *stationary.* A RDS satisfying property (a) is sometimes said to be an "autonomous RDS". When (a) and (b) are both satisfied, we will continue to use the "$\varphi(t,\omega)x$" notation, with noise realisations ω being taken from some probability space $(\Omega,\mathcal{F},\mathbb{P})$.)

When considering RDS satisfying the above two properties, it is also natural (for mathematical purposes) to treat the noise as having no specified "starting time", i.e. as being a *two-sided-time* random process. Hence we can consider trajectories starting at any time on the two-sided timeline. (Under a noise realisation ω, the trajectory starting at position x at time τ is given by $(\varphi(t,\theta^\tau\omega)x)_{t\geq 0}$. If $\tau<0$, then $\theta^\tau\omega$ denotes the time-shifted version of ω *backward* through time $|\tau|$.)

Of course, properties (a) and (b) above can be formulated rigorously; an in-depth study of random dynamical systems based on these two properties can be found in [Arn98]. But for now, let us illustrate these two properties with an example: suppose we have a stochastic differential equation of the form

$$dX_t \;=\; b(t,X_t)\,dt + \sigma(t,X_t)\,dL_t \tag{1.4}$$

for some semimartingale $(L_t)_{t\in\mathbb{R}}$,[7] and suppose that this equation generates a well-defined random dynamical system (where a "noise realisation ω" corresponds to a sample path of the stochastic process $(L_t)_{t\in\mathbb{R}}$). Here, we regard the "noise process" as being represented

[7] We regard $(L_t)_{t\in\mathbb{R}}$ as being a semimartingale if for every $\tau\in\mathbb{R}$, the forward-time stochastic process $(L_{\tau+t})_{t\geq 0}$ is a semimartingale (with respect to its natural filtration).

by the *increments* of the stochastic process (L_t).[8] If the functions $b(t,x)$ and $\sigma(t,x)$ do not depend on t but only depend on x, then property (a) is satisfied; if the stochastic process (L_t) has strictly stationary increments, then property (b) is satisfied.

In most studies of random dynamical systems, the systems involved are not only "time-homogeneous" but also "memoryless"; specifically, this means the following:

(c) the probability distribution for how the noise behaves from a given time t onwards is statistically independent of how the noise behaves up until time t;

(d) how the system behaves over any given time interval is not affected by how the noise behaves *outside* of this time interval.

Once again, illustrating this with equation (1.4): if the stochastic process (L_t) has independent increments, then property (c) is satisfied; in every case, property (d) is satisfied.

Throughout this thesis, we study RDS satisfying properties (a)–(d) above. For any such RDS, we have associated *Markov transition probabilities*; that is to say, we can define a Markov transition function P by

$$P(t,x,A) \;:=\; \mathbb{P}(\omega : \varphi(t,\omega)x \in A),$$

and given any initial condition in the phase space, the subsequent trajectory is a Markov process for which the associated transition probabilities are given by P. (This is proved rigorously in Section 2.5; there, and throughout this thesis, the notation "$\varphi_x^t(A)$" is preferred over the slightly more conventional "$P(t,x,A)$" notation.)

As we will see, much of the study of random dynamical systems revolves around the stationary distributions of the associated Markov transition probabilities. (See Appendix C for an introduction to Markov transition probabilities and their stationary probability distributions.) Conversely, in many situations where one encounters a homogeneous Markov process, this process can naturally be seen as a trajectory of some RDS.

From now on, when we use the term "random dynamical system", we will always mean a random dynamical system satisfying properties (a)–(d) above except where stated otherwise; and we will also assume for the rest of Section 1.2 that every random dynamical system has the property of continuous dependence on initial conditions.

Investigating synchronisation in RDS

When one wishes to investigate mathematically the occurrence or otherwise of noise-induced synchronisation, typically one of the key concepts involved is that of "Lyapunov exponents" (which, heuristically, are a measure of the "infinitesimal-scale repulsivity" of

[8] In other words, the noise process is not a classical stochastic process, but is a kind of "generalised stochastic process", where the behaviour of the noise over a time interval $[t_1,t_2]$ is identified with $(L_t - L_{t_1})_{t_1 \leq t \leq t_2}$.

trajectories). Specifically, the *maximal Lyapunov exponent*[9] associated to a trajectory $(\varphi(t,\omega)x)_{t\geq 0}$ of a spatially differentiable RDS φ on a Riemannian manifold (or a suitable subset thereof) is defined as the limit

$$\lambda(\omega,x) \;:=\; \lim_{t\to\infty} \tfrac{1}{t} \log \|(\mathrm{d}\varphi(t,\omega))_x\|$$

if this limit exists. (When the phase space is a compact manifold, this is independent of the choice of Riemannian structure.[10]) If this limit is negative, it typically follows that the trajectories (evolving under the noise realisation ω) which start sufficiently close to x exhibit some degree of mutual synchronisation.[11] Given an ergodic probability measure ρ for the Markov transition probabilities associated to the RDS φ, under fairly weak conditions on the spatial derivatives of φ, $\lambda(\omega,x)$ exists and is constant across $(\mathbb{P}\otimes\rho)$-almost all (ω,x);[12] let us denote this constant by λ_ρ.

The negativity of λ_ρ will often imply that ρ-almost every initial condition is $\mathbb{P}$-almost surely locally asymptotically stable. But in and of itself, this does not say anything about whether two trajectories starting at distant initial conditions will synchronise. From a practical point of view: negativity of Lyapunov exponents will not automatically guarantee that an array of processes evolving according to φ under the influence of the same noise realisation will synchronise.

A typical "test" for synchronising behaviour in random dynamical systems takes roughly the following form:

(A) If the range of possible behaviours that the system can undergo on a finite time-scale is "sufficiently broad", and if the system exhibits some local-scale synchronising behaviour (as suggested by negativity of Lyapunov exponents), then the RDS will exhibit some "large-scale" synchronising behaviour (in the sense that many initially distant trajectories will approach each other in the long run).

The basic principle behind this is that, given enough opportunity (i.e. after a sufficiently long time), the trajectories of distant initial conditions will eventually be brought within

[9] The word "maximal" is used because, in more than one dimension, there is typically a "spectrum" of Lyapunov exponents representing the growth rates of the different possible tangent vectors at the initial condition x.

[10] This is an immediate consequence of the fact that the norms induced by two different Riemannian metrics on a compact smooth manifold X are Lipschitz equivalent on T_xX uniformly across all $x \in X$; see [MO10].

[11] Rigorous results to this effect include [LeJ87, Lemme 3] (where we warn that the characters $\|^2$ are missing from the end of the denominator in the formula for $\delta_2(T)$), [Car85, Proposition 2.2.3], [MS99, Theorem 3.1, Remark (iii)], [Rue79, Theorem 5.1] and [Arn98, Theorem 7.5.15]. Nonetheless, a useful task for future research would be to provide (especially in continuous time) a more general result than those given in these references.

[12] See e.g. the start of Section 2 of [LeJ87] for discrete time; or for continuous time, see [Arn98, Theorem 4.2.6] (which deals with the entire Lyapunov spectrum), restricted to one-sided time and applied to the measure $\mu = \mathbb{P}\otimes\rho$. As in [New15b, Remark 2.2.12], it is not hard to show that $\lambda_\rho < 0$ if and only if there exists $t > 0$ such that the "average finite-time Lyapunov exponent" $\int \log \|(\mathrm{d}\varphi(t,\omega))_x\| \, \mathbb{P}\otimes\rho(d(\omega,x))$ is negative.

some very small region, where synchronisation then occurs. The "sufficiently broad behaviour" condition typically consists of two parts:

- a "transitivity" condition—i.e. a condition to the effect that individual trajectories, or small-diameter clusters of trajectories, can be transferred from anywhere in the phase space (or some relevant subset thereof) to anywhere else in the phase space (or some relevant subset thereof);
- a "contractibility" condition—i.e. a condition to the effect that distant trajectories can be brought (to some sufficient degree) closer together.

Now as it happens, there is a second important category of tests for synchronising behaviour, where the requirement of local-scale synchronisation is replaced by some "structure-preserving" property of the RDS; in other words, tests of this type take roughly the following form:

(B) If the phase space of the RDS has some special structure that is respected by the RDS, and (where necessary) if the range of possible behaviours that the system can undergo on a finite time-scale is "sufficiently broad", then the RDS will exhibit some "large-scale" synchronising behaviour.

We now give a brief overview of some results from each category, as well as a further result that does not really come under either category:

Tests of "category (A)"

[Bax91] considers random dynamical systems on compact connected smooth manifolds generated by stochastic differential equations of the form

$$dX_t \;=\; b(X_t)\,dt + \sum_{i=1}^{k} \sigma_i(X_t) \circ dW_t^i \tag{1.5}$$

where b and $\sigma_1, \ldots, \sigma_k$ are smooth vector fields and $(W_t^1)_{t\in\mathbb{R}}, \ldots, (W_t^k)_{t\in\mathbb{R}}$ are independent Wiener processes. It is assumed that the set of vector fields $\{b, \sigma_1, \ldots, \sigma_k\}$ satisfies certain "non-degeneracy" conditions (playing the role of the "transitivity" part of the "sufficiently broad behaviour" requirement). As in [BS88], these conditions imply that there exists a unique stationary probability measure ρ for the Markov transition probabilities of the SDE (1.5), and moreover ρ is equivalent to the Riemannian measure (under any Riemannian metric on the manifold). One of the results proved in [Bax91] is that if $\lambda_\rho < 0$ and the trajectories of any two distinct initial conditions are capable of being brought closer together than their initial separation, then the RDS is "synchronising" in the sense that for any two initial conditions in the phase space, with full probability the distance between their subsequent trajectories will tend to 0 as time tends to ∞. (See [Bax91, Theorem 4.10(i)].)

In Theorem 6.1 of this thesis, we provide general criteria for synchronisation on a compact phase space. [Bax91, Theorem 4.10(i)] is a particular case of this more general result; moreover, as a consequence of this result, the conditions on the vector fields in [Bax91, Theorem 4.10(i)] can be replaced with the weaker condition that there is a unique stationary probability measure for the Markov transition probabilities.[13]

[FGS14] considers general RDS on separable complete metric spaces, and finds sufficient conditions for certain notions of synchronisation to hold, that are based on convergence in probability (rather than almost sure convergence). As an application, very broad classes of ordinary differential equations on $\mathbb{R}^n$ are shown to exhibit synchronisation when perturbed by n-dimensional additive Gaussian white noise. (See also Example 6.7 of this thesis.)

In [Hom13], discrete-time diffeomorphic RDS on a compact manifold are considered. Theorems 1.1[14] and 1.2 of [Hom13] provide sufficient conditions for synchronisation to occur (on either the whole manifold or a suitable open subset thereof). Once again, Theorems 1.1 and 1.2 of [Hom13] are particular cases of Theorem 6.1 of this thesis (although Proposition 4.70 of this thesis is needed in order to derive [Hom13, Theorem 1.1] as a special case of Theorem 6.1 of this thesis). Nonetheless, the basic idea of the proof of [Hom13, Theorem 1.1] can be generalised well beyond the context of diffeomorphic RDS on a compact manifold. Specifically, the basic idea of the proof is that, given any set S of initial conditions, if the subsequent trajectories are able to simultaneously reach an arbitrarily small neighbourhood of some point p, and if the trajectory starting at p is itself able to reach an open region U such that it is possible for all trajectories starting in U to synchronise with each other, then it is possible for all the trajectories starting in S to eventually synchronise with each other. It is precisely by combining this idea with [LeJ87, Proposition 2] (which concerns the "statistical equilibrium" associated to an ergodic probability measure of the Markov transition probabilities) that Theorem 6.6 of this thesis has been obtained.

Tests of "category (B)"

In [CF98], it is shown that for a monotone (i.e. order-preserving) continuous RDS on $\mathbb{R}$ (or a subinterval thereof) whose Markov transition probabilities admit a unique stationary probability measure, if there exists a "strictly invariant compact absorbing random set" that is determined by the past of the noise, then the RDS admits a *globally attracting random fixed point* (in the "pullback" sense); this implies, in particular, that the distance between the trajectories of any two given initial conditions converges in probability to 0 as time tends to ∞. In Theorem 3.18 of this thesis, we give a similar result, in which the unique ergodicity and "absorbing set" conditions are not needed, but rather for each stationary probability measure ρ, we conclude that there is a random fixed point that is attracting within the support of ρ. [CF98] demonstrates, as its main application, that adding Gaussian white noise to the right-hand side of the differential equation

$$\dot{x} = \alpha x - x^3 \tag{1.6}$$

[13] In particular, consideration of the "lifted" vector fields onto the unit sphere bundle is not needed.

[14] In Theorem 1.1, it seems that the required additional assumption that m is the only stationary probability measure is missing.

causes the pitchfork bifurcation exhibited by (1.6) to be destroyed, creating instead the scenario that a globally attracting random fixed point persists across all values of α;[15] see also Example 3.21 of this thesis, where the same result is obtained as a consequence of Theorem 3.18.

Synchronisation in monotone RDS on partially ordered phase spaces has been considered in [CS04] (with an application being found in [CCK06, Proposition 5.6]) and further developed in [FGS15] (which is not specific to memoryless RDS, but also considers non-memoryless time-homogeneous RDS).

The first major result to the effect that "sufficiently noisy" invertible RDS on the circle are synchronising (and in fact, perhaps the first rigorous theoretical study concerning the phenomenon of noise-induced synchronisation) is due to Antonov in 1984 ([Ant84]). [Ant84] considers discrete-time RDS on the circle $\mathbb{S}^1$ generated by a finite family of orientation-preserving homeomorphisms $\{f_1,\ldots,f_k\}$, where at each stage one of these maps is selected at random (independently of all previous stages) according to some fixed probability distribution. It is shown that if the whole circle is a minimal invariant closed set under both the original RDS φ and its time-reversal, then either

(i) the maps $f_1,\ldots,f_k$ are simultaneously conjugate to rotations (in other words, after a continuous coordinate change on $\mathbb{S}^1$, the RDS φ just consists of random rotations); or

(ii) there exists an orientation-preserving homeomorphism $g\colon\mathbb{S}^1\to\mathbb{S}^1$ such that

- $g^n = \mathrm{id}_{\mathbb{S}^1}$ for some $n\in\mathbb{N}$ (so all orbits are n-periodic);
- g commutes with f_i for all $1\le i\le k$; so letting S_g denote the set of orbits of g, we can define the maps $\hat{f}_1,\ldots,\hat{f}_k\colon S_g\to S_g$ as the projections of $f_1,\ldots,f_k$ (respectively) onto S_g;
- equipping S_g with the obvious topology (making it a topological circle), the RDS $\hat{\varphi}$ on S_g generated by the maps $\{\hat{f}_1,\ldots,\hat{f}_k\}$ is synchronising (in that the trajectories of two given initial conditions will almost surely mutually converge).

Of course, the case that $n=1$ (i.e. g is the identity function) is the case that the original RDS φ is synchronising. Note that both scenario (i) and the "$n\ge 2$" case of scenario (ii) are "atypical" situations, and so the "typical" scenario for iterated function systems that

[15]Nonetheless (as discussed in [Cal+13]), after the addition of noise, synchronisation of trajectories continues to occur *faster than some deterministic rate* when $\alpha<0$, while for $\alpha>0$ there is no almost sure upper bound on how long one has to wait in order to observe synchronisation of two trajectories. (Indeed, wherever synchronisation does not occur in the absence of noise, one can never expect the addition of noise to create an almost sure deterministic rate of synchronisation.) Viewing this from another perspective: as pointed out to the author by Maxim Arnold, for $\alpha<0$ we have synchronisation of all trajectories under *every* sample path of the noise, while for $\alpha>0$ there is a non-empty Wiener-null set on which some trajectories will never synchronise. This is a highly relevant observation, because the space of sample paths is naturally equipped with a *topology* in which trajectories of the RDS depend continuously on the sample path. (See Remark 2.7.) We will discuss this further in Section 4.7 and in Example 6.7.

are minimal in both forward time and reverse time is global-scale synchronisation.

In Section 5.1 of [DKN07], discrete-time invertible orientation-preserving RDS on the circle are considered.[16] It is shown ([DKN07, Proposition 5.2]) that if the whole circle is a minimal invariant closed set and any arc is able to contract to an arbitrarily small length under the action of the RDS, then there is a pullback-attracting random fixed point which attracts almost the whole circle. Consequently, by reversing time, it is obtained ([DKN07, Proposition 5.3, Remarque 5.4]) that if the whole circle is a minimal invariant closed set under the time-reversal of the RDS[17] and any arc is able to contract to an arbitrarily small length under the action of the RDS, then the RDS is synchronising (in that there is a global random repeller whose law is atomless). This generalises a previous result ([KN04, Theorem 1]) on synchronisation in iterated function systems on the circle.

In Section 5.2 of this thesis, we improve this result by showing that, given each of the two conditions in [DKN07, Proposition 5.3], the other condition can be replaced with a weaker condition. Moreover, our version of the contractibility of arcs condition is simply that every arc is able to contract to a length less than its original length; as in Proposition 5.18 of this thesis, this implies that every arc is able to contract to an arbitrarily small length. We formulate our results in such a way as to cover both discrete and continuous time.

[Kai93] considers the RDS generated by iterations of an orientation-preserving analytic diffeomorphism f on the circle subject to a sequence of independent (but not necessarily identically distributed)[18] random perturbations, and finds conditions under which the RDS is synchronising. (Specifically, these conditions are that f has an irrational rotation number and has no subperiodicity, together with an additional condition on the probability distributions of the random perturbations.)

There are also several results to the effect that a "generic" order- or orientation-preserving RDS on a one-dimensional phase space exhibits negative Lyapunov exponents, when Lyapunov exponents exist. See, for example: [LeJ87, Proposition 1(b)] with $d = 1$ (which applies not only to invertible RDS on $\mathbb{S}^1$ but also to strictly monotone RDS on $\mathbb{R}$); [CF98, Remark 3.7] (which concerns SDE on $\mathbb{R}$); [Cra02a, Corollary 4.4] (which concerns continuous RDS on $\mathbb{S}^1$ in continuous time); and [Kai93, Theorem 2.1(c)].

Vanishing maximal Lyapunov exponents

We have mentioned that when investigating noise-induced synchronisation, Lyapunov exponents are often considered. Now in general, a maximal Lyapunov exponent of exactly 0 indicates nothing about local-scale attractivity or repulsivity of trajectories.[19]

[16] Some rather restrictive additional assumptions are made, but these assumptions are not needed for the proofs of the synchronisation results.

[17] Due to the additional assumptions made in [DKN07, Section 5.1], no distinction is made between minimality under the forward action of the RDS and minimality under the time-reversal of the RDS

[18] Of the four properties (a)–(d) given further above, this RDS satisfies (a), (c) and (d), but not necessarily (b).

[19] There is no reason to expect that in general, maximal Lyapunov exponents being exactly 0 is a "degenerate" situation; see e.g. [BBD14].

Nonetheless, there is one synchronisation result that is worth mentioning, where synchronising behaviour can be deduced from a *non-positive* (i.e. negative *or zero*) maximal Lyapunov exponent. This result is provided by [Bax91, Corollary 5.12] (together with [Bax91, Theorem 4.10] mentioned further above): Once again, random dynamical systems on a compact connected smooth manifold M generated by a SDE of the form (1.5) are considered. A stronger "non-degeneracy" condition on the vector fields is assumed than those required for [Bax91, Theorem 4.10].[20] Let ρ be the unique stationary probability measure of the Markov transition probabilities associated to equation (1.5). In [Bax91, Corollary 5.12] it is shown that if $\lambda_\rho = 0$ and the trajectories of any two distinct initial conditions are capable of being brought closer together than their initial separation, then the RDS is statistically synchronising with respect to ρ (in the sense of Chapter 3 of this thesis); heuristically, this implies that after a sufficiently long time, with extremely high probability the trajectories of all but an extremely small proportion (according to the Riemannian measure) of the initial conditions in the phase space will lie within some region of extremely small diameter. (The same conclusion, and indeed a much stronger conclusion, holds when λ_ρ is strictly negative rather than 0, as given by [Bax91, Theorem 4.10].)

1.3 On the one-to-one correspondence $\mathcal{I}_\varphi \leftrightarrow \mathcal{I}_\varphi^+$

Foundational to the proof of Theorem 3.6 is the well-known one-to-one correspondence between Markov invariant measures and stationary measures, which is itself a special case of the more general one-to-one correspondence between two-sided-time invariant measures and one-sided-time invariant measures. We will first explain the former special case (as studied in this thesis), and then explain the general case.

This section will assume basic familiarity with random dynamical systems.

Stationary measures and Markov invariant measures

Let φ be a RDS on a phase space X over a measure-preserving dynamical system $(\Omega, \mathcal{F}, \mathbb{P}, (\theta^t)_{t\in\mathbb{T}})$. We write $\pi_\Omega : (\omega, x) \mapsto \omega$ and $\pi_X : (\omega, x) \mapsto x$ for the projections from $\Omega \times X$ onto Ω and X respectively.

An *invariant measure* of φ is a probability measure μ on $\Omega \times X$ with $\pi_{\Omega *}\mu = \mathbb{P}$ that is invariant under the dynamical system $(\Theta^t)_{t\in\mathbb{T}^+}$ given by

$$\Theta^t(\omega, x) = (\theta^t\omega, \varphi(t,\omega)x).$$

[20]Specifically: Let $T_{\neq 0}M$ be the set of non-zero tangent vectors on M; and for any vector field σ on M, let $\hat{\sigma}$ be the vector field on the manifold $T_{\neq 0}M$ given in local coordinates by $\hat{\sigma}(x,v) = (x, v, f(x), f'(x)v)$ where $\sigma(x) = (x, f(x))$. (In other words, $\hat{\sigma}$ is the restriction to $T_{\neq 0}M$ of $\kappa \circ \mathrm{d}\sigma$, where κ denotes the canonical flip on TTM and $\mathrm{d}\sigma : TM \to TTM$ is the derivative of σ; see also [Ver14] and [Arn98, Theorem 2.3.42(ii)].) For [Bax91, Corollary 5.12], it is assumed that the vector fields $\{\hat{\sigma}_1, \ldots, \hat{\sigma}_k\}$ satisfy Hörmander's condition—that is to say, the union of the images of the vector fields contained in the Lie algebra generated by $\{\hat{\sigma}_1, \ldots, \hat{\sigma}_k\}$ is equal to the whole of $T(T_{\neq 0}M)$.

Provided the measurable structure of X is that of a *standard measurable space* (i.e. it can be regarded as a Borel subset of a Polish space, equipped with the Borel σ-algebra), any probability measure μ on $\Omega \times X$ with $\pi_{\Omega*}\mu = \mathbb{P}$ admits a *disintegration*—that is, a random probability measure μ_ω such that μ can be expressed as

$$\mu(A) = \int_\Omega \mu_\omega(A_\omega)\,\mathbb{P}(d\omega)$$

where A_ω denotes the ω-section of a measurable set $A \subset \Omega \times X$. Moreover, such a disintegration μ_ω is *unique* up to $\mathbb{P}$-almost everywhere equality. With this, one can show that μ is an invariant measure if and only if

$$\varphi(t,\omega)_*\mu_\omega = \mu_{\theta^t\omega} \ \mathbb{P}\text{-a.s.}$$

for each $t \in \mathbb{T}^+$.

Now suppose we place a two-sided filtration $(\mathcal{F}_s^{s+t})_{s\in\mathbb{T},t\in\mathbb{T}^+}$ on the underlying probability space, such that the dynamical system (θ^t) becomes a *filtered dynamical system* and φ is adapted to this same filtration (see Section 2.2 for precise details).[21] With this, we can say that φ is a *memoryless RDS* if the σ-algebras $\mathcal{F}_{-\infty}^0$ and $\mathcal{F}_0^\infty$ (defined in the natural way) are independent σ-algebras under $\mathbb{P}$. (This then implies that for all $t \in \mathbb{T}$, $\mathcal{F}_{-\infty}^t$ and $\mathcal{F}_t^\infty$ are independent.)

Assume φ is memoryless. We say that an invariant measure μ is a *Markov invariant measure* if μ admits a disintegration μ_ω that depends $\mathcal{F}_{-\infty}^0$-measurably on ω. Now we will say that a probability meausre ρ on X is *stationary* if

$$\rho = \int_\Omega \varphi(t,\omega)_*\rho(\cdot)\,\mathbb{P}(d\omega).$$

This is the same as saying that ρ is stationary under the Markov transition probabilities associated to φ.

Now then, it is well-known that for a spatially continuous memoryless RDS φ on a Polish space X, the map

$$\mu \mapsto \int_\Omega \mu_\omega(\cdot)\,\mathbb{P}(d\omega) = \pi_{X*}\mu$$

serves as a bijection from the set $\mathcal{I}_M$ of Markov invariant measures to the set $\mathcal{S}$ of stationary probability measures. Moreover, if we restrict $\mathcal{F}$ to being the smallest σ-algebra containing every member of the two-sided filtration $(\mathcal{F}_s^{s+t})$, then the same map serves as a bijection between the set of ergodic Markov invariant measures and the set of ergodic stationary probability measures. (This can be obtained as a consequence of the more general one-to-one correspondence between invariant measures and forward-time invariant measures, which we will soon describe; see also Section 1.7 of [Arn98].)

As in [KS12, Theorem 4.2.9], the inverse of the above bijective map is constructed as

[21] If X is a separable metric space and φ is spatially continuous, then we can just take the "natural filtration" $\mathcal{F}_s^{s+t} := \sigma(\omega \mapsto \varphi(\theta^u\omega, v)x : x \in X, s \le u \le u+v \le t)$, as is done in some expositions of the topic.

follows: given $\rho \in \mathcal{S}$, if we take an arbitrary unbounded increasing sequence (t_n) in $\mathbb{T}^+$, the limit

$$\mu_\omega := \lim_{n\to\infty} \varphi(t_n, \theta^{-t_n}\omega)_*\rho \tag{1.7}$$

exists in the narrow topology for $\mathbb{P}$-almost all $\omega \in \Omega$; moreover, this limiting random measure μ_ω stays the same (up to almost sure equality) when the sequence (t_n) is changed. This map $\rho \mapsto \mu$ (where μ_ω is a disintegration of μ) serves as the inverse of the map $\mu \mapsto \pi_{X*}\mu$. As in [LeJ87, Lemme 1(b)], the measure-valued stochastic process $(\varphi(t,\cdot)_*\rho)_{t\in\mathbb{T}^+}$ is convergent in distribution, with the limiting distribution Q_ρ (called the "statistical equilibrium" associated to ρ) being precisely equal to the law of the random measure μ_ω.[22]

Now let us outline the proof presented in [KS12] that the above construction for the inverse map is well-defined. Fix $\rho \in \mathcal{S}$. Let (t_n) be an unbounded increasing sequence in $\mathbb{T}^+$, and let $\mu_{n,\omega} := \varphi(t_n, \theta^{-t_n}\omega)_*\rho$ for each n and ω. It is not hard to show that for any bounded continuous $g: X \to \mathbb{R}$, the stochastic process $\mu_{n,\omega}(g)$ is a martingale, and therefore converges almost surely. Now it has been proved in [BPR06] that for *any* sequence of random probability measures $\mu_{n,\omega}$ on a Polish space X, if for every bounded continuous $g: X \to \mathbb{R}$ the stochastic process $\mu_{n,\omega}(g)$ is almost surely convergent, then the measure-valued stochastic process $\mu_{n,\omega}$ is itself almost surely convergent in the narrow topology. Hence the limit (1.7) exists in the narrow topology almost surely. It is clear that if we took a different sequence (t'_n) then we would obtain the same limit (almost everywhere), since the two sequences (t_n) and (t'_n) can be expressed as subsequences of one "larger" increasing sequence (s_n) on which the above construction can be applied. Thus the construction of the measure μ is well-defined. Moreover, this construction does not make any use of continuity properties of φ. Of course, once the measure μ has been obtained, the next stage in the proof of the one-to-one correspondence between $\mathcal{I}_M$ and $\mathcal{S}$ is to show that the constructed measure μ is an invariant measure. This is a straightforward task, assuming the spatial continuity of φ.

Now it is worth mentioning that the result in [BPR06] is much easier to prove in the particular case that X is compact. (Indeed, it can be obtained as an immediate consequence of Corollary A.18.) As a consequence of the result holding when X is compact, one can easily deduce that the result holds whenever X is a Borel subset of a Polish space and the sequence of random probability measures $\mu_{n,\omega}$ has the property that

$$\int_\Omega \mu_{n,\omega}(\cdot)\,\mathbb{P}(d\omega) = \int_\Omega \mu_{m,\omega}(\cdot)\,\mathbb{P}(d\omega)$$

for all $m, n \in \mathbb{N}$. (This consequence is obtained simply by embedding X into the Hilbert cube $[0,1]^\mathbb{N}$ and regarding $\mu_{n,\omega}$ as a probability measure on $[0,1]^\mathbb{N}$ with $\mu_{n,\omega}(X) = 1$. Letting μ_ω be the limiting random measure on $[0,1]^\mathbb{N}$, the dominated convergence theorem gives that $\int_\Omega \mu_\omega(\cdot)\,\mathbb{P}(d\omega) = \int_\Omega \mu_{n,\omega}(\cdot)\,\mathbb{P}(d\omega)$ for any n, and so $\mu_\omega(X) = 1$ almost surely.)

In the case of the above construction for the inverse of the map $\mu \mapsto \pi_{X*}\mu$, we have

[22] In some references, such as in [LeJ87] itself, the term "statistical equilibrium" is used to refer to the Markov invariant measure μ rather than to Q_ρ.

that

$$\int_\Omega \mu_{n,\omega}(\cdot)\,\mathbb{P}(d\omega) \;=\; \rho$$

for all n. Consequently, [KS12, Theorem 4.2.9] is not specific to the case that X is Polish, but holds whenever X is a Borel subset of a Polish space.[23]

Our new approach

The notion of a Markov invariant measure is not a topological notion but purely a measurable notion. Consequently, for any stationary probability measure $\rho \in \mathcal{S}$, the associated statistical equilibrium Q_ρ remains the same under any change of metric that preserves both the measurable structure of X and the spatial continuity of φ. Of course, most changes of metric that merely preserve the *measurable* structure of X do *not* preserve the spatial continuity of φ.

However, if we can remove the requirement that φ is spatially continuous[24] and still obtain that the measure μ as constructed in (1.7) is an invariant measure, then the whole picture changes. The one-to-one correspondence between $\mathcal{I}_M$ and $\mathcal{S}$ no longer relies on any continuity assumptions. As a consequence, the statistical equilibrium Q_ρ becomes meaningful without reference to a topology, i.e. it becomes a measurable invariant. Now as in [Bax91], the statistical equilibrium Q_ρ encodes the statistical asymptotic behaviour of the n-point motions of φ. Hence, key properties of the asymptotic n-point dynamics also become measurable invariants—not least, synchronisation of trajectories. (The precise sense in which synchronisation is preserved under measurable changes of metric is described in Chapter 3.)

As in Chapter 3, the requirement of spatial continuity to prove the invariance of μ can indeed be removed. This can be achieved as follows: Since μ_ω is $\mathcal{F}_{-\infty}^0$-measurable, we may assume without loss of generality that $\mathcal{F}$ is the smallest σ-algebra containing every member of the two-sided filtration $(\mathcal{F}_s^{s+t})$. For each $t \in \mathbb{T}^+$, letting $\mu_{t,\omega} := \varphi(t, \theta^{-t}\omega)_*\rho$ for all ω, we may regard the random measure $\mu_{t,\omega}$ as a disintegration of a probability measure μ^t on the measurable space $(\Omega \times X, \mathcal{F}_{-t}^\infty \otimes \mathcal{B}(X))$. One can show that the measure μ as constructed in (1.7) agrees with μ^t on $\mathcal{F}_{-t}^\infty \otimes \mathcal{B}(X)$ for all t. One can also show that for each t, μ^t is invariant under the dynamical system $(\Theta^s)_{s\in\mathbb{T}^+}$ acting on $\mathcal{F}_{-t}^\infty \otimes \mathcal{B}(X)$. Hence, by the uniqueness of extensions of premeasures to measures, μ is invariant under the dynamical system $(\Theta^s)_{s\in\mathbb{T}^+}$ acting on $\mathcal{F} \otimes \mathcal{B}(X)$. So μ is an invariant measure. (Moreover, again using uniqueness of extensions, it is not hard to show that if ρ is ergodic then μ is ergodic.)

The above approach makes no reference to the continuity or otherwise of φ. In other words, the one-to-one correspondence between $\mathcal{I}_M$ and $\mathcal{S}$ described above holds for any measurable memoryless RDS on a Borel subset of a Polish space.

---

[23] I am grateful to Gerhard Keller and Hans Crauel for the discussions that led to these observations.

[24] We warn, however, that if we remove the spatial continuity of φ, then we can no longer necessarily just work with the filtration $\mathcal{F}_s^{s+t} := \sigma(\omega \mapsto \varphi(\theta^u\omega, v)x : x \in X, s \le u \le u+v \le t)$, as φ may no longer be adapted to this filtration. In practice, this will rarely if ever be a problem: a natural choice for the underlying filtration $(\mathcal{F}_s^{s+t})$ will usually come from the structure of the noise itself, independently of the RDS that is defined over the noise.

Now the martingale convergence arguments used in [KS12, Theorem 4.2.9], when combined with the consideration of a random measure on X as being a disintegration of a measure on $\Omega \times X$, actually yield a kind of "extension theorem". Specifically: one can prove (Theorem 3.35) that for any one-parameter filtered probability space $(\Omega, \mathcal{F}, (\mathcal{G}_t)_{t\in\mathbb{T}^+}, \mathbb{P})$ and any standard measurable space (X, Σ), given a consistent family $\{\mu^t\}_{t\in\mathbb{T}^+}$ of probability measures μ^t on $\mathcal{G}_t \otimes \Sigma$ whose Ω-projection coincides with $\mathbb{P}|_{\mathcal{G}_t}$, there exists a unique probability measure μ on $\sigma(\mathcal{G}_t : t \in \mathbb{T}^+) \otimes \Sigma$ whose restriction to $\mathcal{G}_t \otimes \Sigma$ agrees with μ^t for all t. In Chapter 3 of this thesis, the way in which we present the theory of Markov invariant measures is to *first* develop this extension theorem as a general result, and *then* employ it in the proof of the one-to-one correspondence between $\mathcal{I}_M$ and $\mathcal{S}$ (Theorem 3.49). (So the martingale convergence theorem does not *directly* appear within our proof of the correspondence.)

It is also worth adding that in continuous time, by Lévy's upward theorem together with the fact that the narrow topology is determined by a countable family of bounded continuous functions, one can obtain that for any unbounded countable $S \subset [0, \infty)$, in the narrow topology we have the convergence

$$\varphi(t, \theta^{-t}\omega)_*\rho \to \mu_\omega \quad \text{as } t \text{ tends to } \infty \text{ in } S \tag{1.8}$$

$\mathbb{P}$-almost surely. (See also Theorem 3.33.) Again, this statement does not rely on any continuity properties of φ. Nonetheless, if we are working with a topology in which the map $t \mapsto \varphi(t, \theta^{-t}\omega)x$ is left-continuous for every ω and x (or right-continuous for every ω and x),[25] then it follows that

$$\varphi(t, \theta^{-t}\omega)_*\rho \to \mu_\omega \quad \text{as } t \to \infty \tag{1.9}$$

$\mathbb{P}$-almost surely, where t is not restricted to a countable set but ranges throughout $[0, \infty)$. (In general, (1.9) does *not* follow from the fact that (1.8) holds almost surely for each unbounded countable S,[26] since there are uncountably many unbounded countable subsets of $[0, \infty)$.)

For fuller details, see Sections 3.3–3.5 of this thesis (and in particular, Theorem 3.49).

[25]More generally, the requirement is that for every bounded continuous function $g: X \to \mathbb{R}$, the stochastic process $\varphi(t, \theta^{-t}\omega)_*\rho(g)$ is a separable stochastic process.

[26]This is perhaps most easily demonstrated by the (rather pathological) example of a RDS describing "random kicks that immediately undo themselves": Let Ω be the set of surjective increasing right-continuous functions $\omega: \mathbb{R} \to \mathbb{Z}$ with $\omega(0) = 0$ (equipped with the σ-algebra $\mathcal{F}$ generated by the projections $\omega \mapsto \omega(t)$); let $(\theta^t)_{t\in\mathbb{R}}$ be the shift system $\theta^t\omega(s) = \omega(s+t) - \omega(t)$; and let $\mathbb{P}$ be the probability measure on Ω such that the stochastic processes $(\omega(t))_{t\geq 0}$ and $(-\omega(-t))_{t\geq 0}$ are independent Poisson processes with the same parameter λ. Let $\mathcal{F}_s^t := \bigcap_{\delta>0} \tilde{\mathcal{F}}_{s-\delta}^t$ where $\tilde{\mathcal{F}}_u^v := \sigma(\omega \mapsto \omega(r) - \omega(u) : u \leq r \leq v)$. [Alternatively, one can just take the natural filtration of the RDS φ that we will introduce.] Let $X = \{-1, 1\}$. For any $t \geq 0$, $\omega \in \Omega$ and $x \in X$, let $\varphi(t, \omega)x = x$ if the map $\tau \mapsto \theta^\tau\omega(t)$ is left-continuous at 0 and let $\varphi(t, \omega)x = -x$ otherwise. Then we can take $\rho = \mu_\omega = \delta_1$ for all ω; in this case, (1.8) will hold almost surely for any given unbounded countable S, but (1.9) will not hold for any $\omega \in \Omega$.

One-sided-time invariant measures

Our new approach described above can easily be extended to the more general one-to-one correspondence between two-sided-time and one-sided-time invariant measures, again yielding a statement that requires no continuity assumptions.

We still equip the noise with a two-sided filtration $(\mathcal{F}_s^{s+t})$ as described above, and assume moreover that $\mathcal{F}$ is the smallest σ-algebra containing all members of this filtration. We still assume that the RDS φ is adapted to this filtration, but we do *not* need to assume that φ is memoryless. We write $\mathcal{I}_\varphi$ for the set of invariant measures of φ. Again, X is a Borel subset of a Polish space.

A *one-sided-time invariant measure* is a probability measure μ^+ on the measurable space $(\Omega \times X, \mathcal{F}_0^\infty \otimes \mathcal{B}(X))$ such that $\pi_{\Omega *}\mu^+ = \mathbb{P}|_{\mathcal{F}_0^\infty}$ and μ^+ is invariant under $(\Theta^t)_{t\in\mathbb{T}^+}$. Let us write $\mathcal{I}_\varphi^+$ for the set of one-sided-time invariant measures.

By much the same arguments as in Section 3.5 of this thesis, one can show that without any continuity requirements, the map

$$\mu \;\mapsto\; \mu|_{\mathcal{F}_0^\infty \otimes \mathcal{B}(X)}$$

serves as a bijection from $\mathcal{I}_\varphi$ to $\mathcal{I}_\varphi^+$, and the inverse map is constructed as follows: given any $\mu^+ \in \mathcal{I}_\varphi^+$, taking μ_ω^+ to be a disintegration of μ^+ (over the probability space $(\Omega, \mathcal{F}_0^\infty, \mathbb{P}|_{\mathcal{F}_0^\infty})$), for any unbounded countable $S \subset \mathbb{T}^+$ we have that the limit

$$\mu_\omega \;:=\; \lim_{\substack{t\to\infty \\ \text{in } S}} \varphi(t, \theta^{-t}\omega)_* \mu_{\theta^{-t}\omega}^+ \tag{1.10}$$

exists in the narrow topology for $\mathbb{P}$-almost all $\omega \in \Omega$; moreover, this limiting random measure μ_ω stays the same (up to almost sure equality) when the set S is changed. This map $\mu^+ \mapsto \mu$ (where μ_ω is a disintegration of μ) serves as the inverse of the above map $\mu \mapsto \mu|_{\mathcal{F}_0^\infty \otimes \mathcal{B}(X)}$.

Now in continuous time, even if we are working in a topology in which the map $(t,x) \mapsto \varphi(t, \theta^{-t}\omega)x$ is jointly continuous, it is *not* the case that for every version of the disintegration μ_ω^+ of μ the limit

$$\lim_{t\to\infty} \varphi(t, \theta^{-t}\omega)_* \mu_{\theta^{-t}\omega}^+$$

exists almost surely: indeed, if we have one version in which the limit does exist almost surely, it will generally be possible to modify this version on a null set in such a manner that the limit no longer almost surely exists. The natural question is then whether there exist some "reasonable" conditions under which there is guaranteed to exist at least one version of the disintegration such that this limit does exist almost surely. This question remains open.

Chapter 2. Foundations

We will develop the foundational theory of random dynamical systems that will be needed for results later on in the thesis.

2.1 Standard measurable spaces

Before introducing random dynamical systems, we first introduce the kind of space on which they will always be assumed to act throughout this thesis.

Recall that a topology or a topological space is said to be *Polish* if it is separable and completely metrisable. Note that the completion of any separable metric space is a Polish space.

Definition 2.1. A measurable space (X, Σ) is said to be *standard* if there exists a Polish topology on X whose Borel σ-algebra is Σ.

The term *standard Borel space* is also often used for a measurable space that is standard. The *Borel isomorphism theorem* (e.g. [Sri98, Theorem 3.3.13]) states that for a measurable space (X, Σ), the following are equivalent:

(i) (X, Σ) is standard;

(ii) (X, Σ) is measurably isomorphic to either a finite discrete space, a countable discrete space, or $([0,1], \mathcal{B}([0,1]))$.

(Obviously it follows, in particular, that (X, Σ) is standard if and only if there exists a compact metrisable topology on X whose Borel σ-algebra is Σ.)

For a full proof of the above theorem, see e.g. [Sri98, Theorem 3.3.13] or [KL14, Theorem A.17]. We now present a sketch of the proof:

Sketch-proof of the Borel isomorphism theorem. The statement is clear if X is finite or countable. Suppose X is uncountable. First observe that via binary expansions, $[0,1)$ is measurably isomorphic to a cocountable subset of the set $C := \{0,1\}^{\mathbb{N}}$; and therefore $[0,1]$ is measurably isomorphic to C (since any uncountable measurable space in which all singletons are measurable is clearly isomorphic to any of its cocountable subsets). Moreover, since $\mathbb{N}$ and $\mathbb{N} \times \mathbb{N}$ are isomorphic as sets, it follows that C is measurably isomorphic to $[0,1]^{\mathbb{N}}$. Now it is known that for two measurable spaces to be isomorphic, it is sufficient that each can be measurably embedded as a measurable subset of the other; so, to prove the desired result, we show that **(I)** $[0,1]^{\mathbb{N}}$ contains a copy of X as a measurable subset, and **(II)** X contains a copy of C as a measurable subset. **(I)** Fix a complete metrisation d of the topology of X in which $\operatorname{diam} X \leq 1$. Given a countable dense subset $\{x_n\}_{n \in \mathbb{N}}$ of X, the map $x \mapsto (d(x, x_n))_{n \in \mathbb{N}}$ serves as a topological embedding of X into $[0,1]^{\mathbb{N}}$. If we let U_n be the union of all open subsets V of $[0,1]^{\mathbb{N}}$ satisfying

$\mathrm{diam}(X \cap V) \leq \frac{1}{n}$ under the metric d, one can use the completeness of d to show that $X = \bar{X} \cap \bigcap_{n=1}^{\infty} U_n$ (where $\bar{X}$ is the closure of X in $[0,1]^{\mathbb{N}}$). **(II)** If X is a perfect space,[1] then (as in the "Cantor middle thirds" construction) one can obtain a copy of C as the intersection of a nested sequence of unions of closed balls (under a complete metrisation of the topology of X). If X is not perfect, then letting $N \subset X$ be the set of points admitting a neighbourhood that is at most countable, one can show (using the separability of X) that N is itself countable, from which it follows that $X \setminus N$ is a perfect space. □

It can also be shown ([Sri98, Proposition 3.3.7]) that for any separable metric space (X,d), the following are equivalent:

- $\mathcal{B}(X)$ is standard;
- X is homeomorphic to a Borel subset of a Polish space;[2]
- (X,d) is isometric to a Borel subset of a separable complete metric space;
- X is a Borel subset of the d-completion of X;
- for every metric d' on X that is topologically equivalent to d, X is a Borel subset of the d'-completion of X.

Remark 2.2. It is known that, assuming the axiom of choice, *every* metrisable topology whose Borel σ-algebra is standard is separable ([Sri98, Remark 3.3.8, Theorem 4.3.8]). Now there are several results in this thesis (particularly in Chapter 3) where, working with a standard measurable space (X,Σ), we prove that in every separable metrisable topology on X with $\mathcal{B}(X) = \Sigma$ some particular phenomenon occurs. Since every metrisable topology whose Borel σ-algebra is standard is separable, we can in fact say that in *every* metrisable topology on X with $\mathcal{B}(X) = \Sigma$, the desired phenomenon occurs. (Nonetheless, we do still choose to keep the word "separable" in the statements of these results.)

Recall that a sequence of probability measures (μ_n) on a measurable space (X,Σ) is said to *converge strongly* to a probability measure μ if for every $A \in \Sigma$, $\mu_n(A) \to \mu(A)$ as $n \to \infty$.

Lemma 2.3. *Let (X,Σ) be a standard measurable space, let (μ_n) be a sequence of probability measures on X, and let μ be a probability measure on X. Then μ_n converges strongly to μ as $n \to \infty$ if and only if for every Polish topology on X generating Σ, μ_n converges weakly to μ.*

Proof. It is clear that strong convergence always implies weak convergence. Conversely, suppose that for every Polish topology on X generating Σ, μ_n converges weakly to μ. Fix any $A \in \Sigma$. By [KL14, Theorem A.11], there exists a Polish topology on X generating Σ in which A is both open and closed; since μ_n converges weakly to μ in this topology, we have that $\mu_n(A) \to \mu(A)$. □

[1] A topological space X is said to be a *perfect space* if every point in X is an accumulation point of X.

[2] The term "Lusin space" is sometimes used to mean a topological space that is a Borel subset of a Polish space; however, we avoid the term here, since it can have other meanings also.

2.2 The formal setup

We introduce the notion of a "memoryless random dynamical system". This consists of two components: a *noise space* (equipped with a "memoryless" filtration); and an *action* of this noise space ("filtered" with respect to the same filtration) on some *phase space*.

The noise space

Let $\mathbb{T}$ be either $\mathbb{Z}$ or $\mathbb{R}$, and let $\mathbb{T}^+ := \mathbb{T} \cap [0, \infty)$. (The set $\mathbb{T}$ represents the "time set" of the noise, which we regard as being two-sided.) Let $\bar{\mathbb{T}} := \mathbb{T} \cup \{-\infty, \infty\}$, and let $\bar{\mathbb{T}}^+ := \mathbb{T}^+ \cup \{\infty\}$.

Let $(\Omega, \mathcal{F})$ be a measurable space. We refer to an element $\omega \in \Omega$ as a *sample point* or *noise realisation*. Let $(\mathcal{F}_s^{s+t})_{s \in \mathbb{T}, t \in \mathbb{T}^+}$ be a family of sub-σ-algebras of $\mathcal{F}$ such that

(i) $\mathcal{F}_{t_1}^{t_2} \subset \mathcal{F}_{t_0}^{t_3}$ for all $t_0 \leq t_1 \leq t_2 \leq t_3$ in $\mathbb{T}$;

(ii) $\sigma(\mathcal{F}_s^{s+t} : s \in \mathbb{T}, t \in \mathbb{T}^+) = \mathcal{F}$.

We refer to $(\mathcal{F}_s^{s+t})_{s \in \mathbb{T}, t \in \mathbb{T}^+}$ as an *exhaustive two-sided filtration of* $\mathcal{F}$. Now let $(\theta^t)_{t \in \mathbb{T}}$ be a family of functions $\theta^t : \Omega \to \Omega$ such that

(i) $\theta^0 = \mathrm{id}_\Omega$, and $\theta^{s+t} = \theta^t \circ \theta^s$ for all $s, t \in \mathbb{T}$;

(ii) $\theta^\tau \mathcal{F}_s^t = \mathcal{F}_{s-\tau}^{t-\tau}$ for all $s, t, \tau \in \mathbb{T}$ with $s \leq t$.

We refer to $(\Omega, \mathcal{F}, (\mathcal{F}_s^{s+t})_{s \in \mathbb{T}, t \in \mathbb{T}^+}, (\theta^t)_{t \in \mathbb{T}})$ as an *exhaustively filtered dynamical system.* We sometimes refer to the group of maps $(\theta^t)_{t \in \mathbb{T}}$ as the *time-shift* system. We will use the following notations:

$$\begin{aligned}
\mathcal{F}_s^\infty &:= \sigma(\mathcal{F}_s^{s+t} : t \in \mathbb{T}^+) \quad \text{for any } s \in \mathbb{T} \\
\mathcal{F}_\infty^\infty &:= \bigcap_{s \in \mathbb{T}} \mathcal{F}_s^\infty \\
\mathcal{F}_{-\infty}^t &:= \sigma(\mathcal{F}_{t-s}^t : s \in \mathbb{T}^+) \quad \text{for any } t \in \mathbb{T} \\
\mathcal{F}_{-\infty}^{-\infty} &:= \bigcap_{t \in \mathbb{T}} \mathcal{F}_{-\infty}^t
\end{aligned}$$

(The σ-algebras $\mathcal{F}_{-\infty}^{-\infty}$ and $\mathcal{F}_\infty^\infty$ are referred to as the *tail σ-algebras.*) It will also be useful to have the convention that $\mathcal{F}_{-\infty}^\infty := \mathcal{F}$.

It is easy to show that (given any $s, t, \tau \in \mathbb{T}$) the following hold:

$$\begin{aligned}
\theta^\tau \mathcal{F} &= \mathcal{F} \\
\theta^\tau \mathcal{F}_s^\infty &= \mathcal{F}_{s-\tau}^\infty \\
\theta^\tau \mathcal{F}_\infty^\infty &= \mathcal{F}_\infty^\infty \\
\theta^\tau \mathcal{F}_{-\infty}^t &= \mathcal{F}_{-\infty}^{t-\tau} \\
\theta^\tau \mathcal{F}_{-\infty}^{-\infty} &= \mathcal{F}_{-\infty}^{-\infty}
\end{aligned}$$

So in particular, θ^t serves as a measurable transformation of the measurable space $(\Omega, \mathcal{F})$ for all $t \in \mathbb{T}$. Let $\mathbb{P}$ be a probability measure on $(\Omega, \mathcal{F})$ with the following properties:

(i) $\theta^t_*\mathbb{P} = \mathbb{P}$ for all $t \in \mathbb{T}$ (i.e. $\mathbb{P}$ is an invariant measure of the time-shift system (θ^t));

(ii) for each $t \in \mathbb{T}$, $\mathcal{F}^t_{-\infty}$ and $\mathcal{F}^\infty_t$ are independent σ-algebras under $\mathbb{P}$.

To verify condition (i), it is sufficient just to consider $t \in \mathbb{T}^+$. Property (i) implies that for any measurable space (X, Σ) and any $(\mathcal{F}, \Sigma)$-measurable function $q : \Omega \to X$, the stochastic process $(q \circ \theta^t)_{t\in\mathbb{T}}$ is strictly stationary. Note that given property (i), a sufficient condition for property (ii) is that *there exists* $t \in \mathbb{T}$ such that $\mathcal{F}^t_{-\infty}$ and $\mathcal{F}^\infty_t$ are independent σ-algebras under $\mathbb{P}$.

Property (ii) implies in particular that for any $t_0 \le t_1 \le t_2 \le t_3$ in $\bar{\mathbb{T}}$, $\mathcal{F}^{t_1}_{t_0}$ and $\mathcal{F}^{t_3}_{t_2}$ are independent under $\mathbb{P}$. This statement is obvious, except perhaps in the case that either $t_0 = t_1 = t_2 = -\infty$ or $t_1 = t_2 = t_3 = \infty$. Indeed, for this case, the statement reduces to being a statement of *Kolmogorov's 0-1 law*, which asserts the following:

Lemma 2.4. *The tail σ-algebras $\mathcal{F}^{-\infty}_{-\infty}$ and $\mathcal{F}^\infty_\infty$ consist of only $\mathbb{P}$-null sets and $\mathbb{P}$-full sets.*

Proof. For any $t \in \mathbb{T}$, $\mathcal{F}^\infty_\infty$ is clearly independent of $\mathcal{F}^t_{-\infty}$. Hence, by Corollary A.9, $\mathcal{F}^\infty_\infty$ independent of the whole of $\mathcal{F}$. In particular, $\mathcal{F}^\infty_\infty$ is independent of itself, i.e. $\mathcal{F}^\infty_\infty$ consists of only $\mathbb{P}$-null sets and $\mathbb{P}$-full sets. One can argue similarly for $\mathcal{F}^{-\infty}_{-\infty}$. □

We will refer to $(\Omega, \mathcal{F}, (\mathcal{F}^{s+t}_s)_{s\in\mathbb{T}, t\in\mathbb{T}^+}, (\theta^t)_{t\in\mathbb{T}}, \mathbb{P})$ as a (*memoryless*, *stationary*) *noise space*.[3]

Basic examples and results

Let us mention a couple of important typical examples of what the noise space $(\Omega, \mathcal{F}, (\mathcal{F}^{s+t}_s), (\theta^t), \mathbb{P})$ could be.

Example 2.5. Let $\mathbb{T} = \mathbb{Z}$. Given a probability space $(I, \mathcal{I}, \nu)$, define

$$\begin{aligned} \Omega &:= I^{\mathbb{Z}} \\ \mathcal{F} &:= \mathcal{I}^{\otimes\mathbb{Z}} \\ \mathbb{P} &:= \nu^{\otimes\mathbb{Z}} \\ \theta^n((\alpha_r)_{r\in\mathbb{Z}}) &:= (\alpha_{n+r})_{r\in\mathbb{Z}} \text{ for each } n \in \mathbb{Z} \\ \mathcal{F}^n_m &:= \sigma(\, (\alpha_r)_{r\in\mathbb{Z}} \mapsto \alpha_k \,:\, m+1 \le k \le n \,) \text{ for each } m, n \in \mathbb{Z} \text{ with } m \le n. \end{aligned}$$

(In particular, $\mathcal{F}^m_m$ is the trivial σ-algebra $\{\varnothing, \Omega\}$ for each $m \in \mathbb{Z}$.)

It is not hard to show that $(\Omega, \mathcal{F}, (\mathcal{F}^{m+r}_m)_{m\in\mathbb{Z}, r\in\mathbb{N}_0}, (\theta^n)_{n\in\mathbb{Z}}, \mathbb{P})$ is a noise space (in accordance with our formalism). Note that if I is equipped with a second-countable topology generating $\mathcal{I}$, then the product topology on Ω generates $\mathcal{F}$; in this case, if ν has full support in I then $\mathbb{P}$ has full support in Ω.

[3]In general, a "(stationary) noise space" may be defined as an exhaustively filtered dynamical system together with a probability measure $\mathbb{P}$ satisfying property (i) (but not necessarily property (ii)). However, throughout this thesis, we will always require a "noise space" to be *memoryless*, meaning that $\mathbb{P}$ also satisfies property (ii).

Example 2.6 (Gaussian white noise)**.** As in Sections A.2 and A.3 of [Arn98], a d-dimensional Gaussian white noise process may be described within our framework as follows: Let $\mathbb{T} = \mathbb{R}$. Let $\Omega := \{\omega \in C(\mathbb{R},\mathbb{R}^d) : \omega(0) = \mathbf{0}\}$. Let $\mathcal{F}$ be the smallest σ-algebra on Ω with respect to which the projection $W_r : \omega \mapsto \omega(r)$ is measurable for every $r \in \mathbb{R}$. For each $s,t \in \mathbb{R}$ with $s \leq t$, let $\mathcal{F}_s^t$ be the smallest σ-algebra on Ω with respect to which $W_r - W_s$ is measurable for every $r \in [s,t]$. Let $\mathbb{P}$ be the *Wiener measure* on $(\Omega,\mathcal{F})$—that is, $\mathbb{P}$ is the unique probability measure under which the stochastic processes $(W_t)_{t\geq 0}$ and $(W_{-t})_{t\geq 0}$ are independent d-dimensional Wiener processes. For each $s,t \in \mathbb{R}$, set $\theta^t\omega(s) := \omega(s+t) - \omega(t)$. Once again, it is not hard to show that $(\Omega,\mathcal{F},(\mathcal{F}_s^{s+t})_{s\in\mathbb{R},\,t\geq 0},(\theta^t)_{t\in\mathbb{R}},\mathbb{P})$ is a noise space (in accordance with our formalism).

Remark 2.7. As in Lemma B.9, on the set $C(\mathbb{R},\mathbb{R}^d)$, the evaluation σ-algebra $\sigma(\omega \mapsto \omega(t) : t \in \mathbb{R})$ is precisely the Borel σ-algebra of the topology of uniform convergence on compact sets. So in Example 2.6, equipping Ω with the topology inherited from $C(\mathbb{R},\mathbb{R}^d)$, we have that $\mathcal{F}$ is the Borel σ-algebra of Ω. As in [Fre13, Proposition 477F], the Wiener measure $\mathbb{P}$ has full support. It is easy to show (using Lemma B.6) that the map $(t,\omega) \mapsto \theta^t\omega$ is jointly continuous.

We will now show that "memoryless stationary noise is always ergodic":

Lemma 2.8 (cf. [New15c, Lemma 5.1])**.** *For any $t \in \mathbb{T} \setminus \{0\}$, θ^t is an ergodic transformation of $(\Omega,\mathcal{F},\mathbb{P})$.*

Proof. Since the inverse of an invertible ergodic transformation is ergodic, it is sufficient just to consider positive t. Fix $t \in \mathbb{T}^+ \setminus \{0\}$. Let $E \in \mathcal{F}$ be a set with $\theta^{-t}(E) = E$, and let $h : \Omega \to [0,1]$ be a version of $\mathbb{P}(E|\mathcal{F}_0^\infty)$. By Lemma A.14, for every $n \in \mathbb{Z}$, $h \circ \theta^{nt}$ is a version of $\mathbb{P}(E|\mathcal{F}_{nt}^\infty)$; and by Lemma 2.4, the constant map $\omega \mapsto \mathbb{P}(E)$ is a version of $\mathbb{P}(E|\mathcal{F}_\infty^\infty)$. Therefore, by Lévy's downward theorem ([Wil91, Theorem 14.4]), $h \circ \theta^{nt}(\omega) \to \mathbb{P}(E)$ as $n \to \infty$ for $\mathbb{P}$-almost all $\omega \in \Omega$. But since θ^t is itself $\mathbb{P}$-preserving, it follows that for each $n \in \mathbb{Z}$, $h \circ \theta^{nt}(\omega) = \mathbb{P}(E)$ for $\mathbb{P}$-almost all $\omega \in \Omega$. In other words, the constant map $\omega \mapsto \mathbb{P}(E)$ is a version of $\mathbb{P}(E|\mathcal{F}_{nt}^\infty)$ for each n, i.e. E is independent of $\mathcal{F}_{nt}^\infty$ for each n. It follows by Corollary A.9 that E is independent of $\mathcal{F}$. In particular, E is independent of itself, i.e. $\mathbb{P}(E) \in \{0,1\}$. □

Action of the noise

Let (X,Σ) be a standard measurable space. We write $\mathcal{M}_{(X,\Sigma)}$ for the set of probability measures on (X,Σ), and we equip $\mathcal{M}_{(X,\Sigma)}$ with its "evaluation σ-algebra" $\mathfrak{K}_{(X,\Sigma)}$, namely the smallest σ-algebra with respect to which the map $\rho \mapsto \rho(A)$ is measurable for all $A \in \Sigma$. We say that a probability measure ρ on X is *atomless* if $\rho(\{x\}) = 0$ for all $x \in X$. We define the projections $\pi_\Omega : \Omega \times X \to \Omega$ and $\pi_X : \Omega \times X \to X$ by $\pi_\Omega(\omega,x) = \omega$ and $\pi_X(\omega,x) = x$. We write $\Delta_X := \{(x,x) : x \in X\} \subset X \times X$, and for any $A \subset X$, we write $\Delta_A := \{(x,x) : x \in A\}$. It is not hard to show that for any $A \in \Sigma$, $\Delta_A \in \Sigma \otimes \Sigma$.

Let $\varphi = (\varphi(t,\omega))_{t\in\mathbb{T}^+,\,\omega\in\Omega}$ be a $(\mathbb{T}^+ \times \Omega)$-indexed family of functions $\varphi(t,\omega) : X \to X$ such that:

(i) $\varphi(0,\omega) = \mathrm{id}_X$ for all $\omega \in \Omega$;

(ii) $\varphi(s+t,\omega) = \varphi(t,\theta^s\omega)\circ\varphi(s,\omega)$ for all $s,t\in\mathbb{T}^+$ and $\omega\in\Omega$;

(iii) for each $t\in\mathbb{T}^+$, the map $(\omega,x)\mapsto\varphi(t,\omega)x$ from $\Omega\times X$ to X is $(\mathcal{F}_0^t\otimes\Sigma,\Sigma)$-measurable.

Properties (i) and (ii) are summarised by saying that φ is a (*forward*) *cocycle* over $(\theta^t)_{t\in\mathbb{T}}$. Property (iii) is summarised by saying that φ is *adapted to* (or *filtered with respect to*) the filtration $(\mathcal{F}_s^{s+t})_{s\in\mathbb{T},t\in\mathbb{T}^+}$. Note that for each $s\in\mathbb{T}$ and $t\in\mathbb{T}^+$, the map $(\omega,y)\mapsto\varphi(t,\theta^s\omega)y$ from $\Omega\times X$ to X is $(\mathcal{F}_s^{s+t}\otimes\Sigma,\Sigma)$-measurable.

We refer to φ as a (*filtered*) *random dynamical system* on the phase space (X,Σ) over the noise space $(\Omega,\mathcal{F},(\mathcal{F}_s^{s+t}),(\theta^t),\mathbb{P})$. We will sometimes refer to the family of time-shifts $(\theta^t)_{t\in\mathbb{T}}$ as the *base system* of the RDS.

In the case that the map $(t,\omega,x)\mapsto\varphi(t,\omega)x$ is $(\mathcal{B}(\mathbb{T}^+)\otimes\mathcal{F}\otimes\Sigma,\Sigma)$-measurable, we will say that *φ is (forward-)measurable*.[4] In the case that the map $(s,t,\omega,x)\mapsto\varphi(t,\theta^s\omega)x$ is $(\mathcal{B}(\mathbb{T})\otimes\mathcal{B}(\mathbb{T}^+)\otimes\mathcal{F}\otimes\Sigma,\Sigma)$-measurable, we will say that *φ is two-way measurable*. (Obviously if $\mathbb{T}=\mathbb{Z}$ then φ is automatically two-way measurable.) Note that if φ is measurable and the map $(t,\omega)\mapsto\theta^t\omega$ is jointly measurable, then φ is two-way measurable.

A path in X taking the form $(\varphi(t,\omega)x)_{t\in\mathbb{T}^+}$ for some $\omega\in\Omega$ and $x\in X$ will sometimes be called a *(forward) trajectory of φ*. A path in X taking the form $(\varphi(t,\theta^{-t}\omega)x)_{t\in\mathbb{T}^+}$ for some $\omega\in\Omega$ and $x\in X$ will sometimes be called a *pullback trajectory of φ*.

Note that in the "deterministic case" where Ω is a singleton $\{\omega\}$, $(\varphi(t,\omega))_{t\in\mathbb{T}^+}$ is an autonomous dynamical system on (X,Σ). (See Section C.2.)

Physical interpretation of the formalism

The formalism that we have just presented (involving both the noise space and its "action" on the state space) is intended to be a precise mathematical way of representing the notion of a non-deterministic dynamical system whose non-determinism is specifically due to the moment-by-moment influence of the behaviour of some time-homogeneous random noise process which dictates the evolution of the the system in a time-homogeneous manner. (Our formalism also incorporates the additional notion that the noise process is statistically memoryless.)

The term "time-homogeneous" is not inherently a mathematically rigorous term, but can be understood physically as meaning "making no reference to any kind of external clock". Let us first illustrate the concept in terms of *deterministic* systems.

A "general deterministic dynamical system" (that is not necessarily time-homogeneous) can be represented mathematically as a two-parameter family $(f_s^t)_{s,t\in T,t\geq s}$ of functions $f_s^t\colon X\to X$, where T represents some "set of times". The physical interpretation is that

[4] Here, we are using the word in a stricter sense than in Chapter 1, where the term "measurable RDS" was used (in contrast to the term "topological RDS") simply to mean an RDS acting on a standard measurable space with no pre-defined topological structure.

if a process governed by this dynamical system is at state $x \in X$ at time s, then it will be at state $f_s^t(x)$ at time t. Of course, for this to make physical sense, we require the consistency relations that (I) f_s^s is the identity for all $s \in T$, and (II) $f_s^u = f_t^u \circ f_s^t$ for all $s, t, u \in T$ with $s \leq t \leq u$.

By contrast, a deterministic dynamical system that is time-homogeneous can be represented mathematically in a slightly more succinct manner, namely as a *one-parameter* family $(f^t)_{t \in \mathbb{T}^+}$ of functions $f^t : X \to X$, with the physical interpretation being as follows: if a process governed by this dynamical system is at state x at time s, then (regardless of what the time s is) it will be at state $f^t(x)$ at time $s + t$. For this to make physical sense, we require the consistency relations that (I) f^0 is the identity, and (II) $f^{s+t} = f^t \circ f^s$ for all $s, t \in \mathbb{T}^+$. These relations form the definition (in the purely set-theoretic setting) of an "autonomous dynamical system on X".

Let us now provide a heuristic interpreration for our formalism of a "*random* dynamical system".

We imagine that we have some "time-homogeneous" noise process, which (for mathematical purposes) we regard as being "eternal", i.e. having no start and no end. (Indeed, it is not surprising that we should view a "time-homogeneous" noise process in this way, since there should be no "clock" to specify when the noise process starts or ends.)

Suppose we fix an arbitrary time to be our "reference time $t = 0$", and imagine that we have some mechanism for "plotting" precisely how the noise behaves over time. (Here, we imagine that the plot is able to display how the noise behaves over the *whole* timeline, both the future $\{t \geq 0\}$ and the past $\{t \leq 0\}$.) Since the noise is random, there are (uncountably) many possibilities for how the plot will turn out. The set of all physical possibilities for how the plot will turn out is denoted by Ω; since the noise is "time-homogeneous", the set Ω does not depend on which time was chosen as our reference time.

Now suppose the plot turns out to be ω (which is some element of the set Ω); and suppose that someone else observing the same noise process chooses their reference time to be τ later than our chosen reference time (where $\tau \in \mathbb{T}$). Then the plot that this person will obtain (assuming the same plotting mechanism as ours) is denoted by $\theta^\tau \omega$. Of course, for this to make physical sense we require the consistency relations that (I) $\theta^0 \omega = \omega$ for all $\omega \in \Omega$, and (II) $\theta^{s+t} \omega = \theta^t \theta^s \omega$ for all $s, t \in \mathbb{T}$ and $\omega \in \Omega$.

Now we assume that for each $s, t \in \mathbb{T}$ with $s \leq t$, there is some natural σ-algebra $\mathcal{F}_s^t$ on Ω, representing all the information concerning how the noise behaves over the time interval $\mathbb{T} \cap [s, t]$ according to our plot; time-homogeneity will then give that $\theta^\tau \mathcal{F}_s^t = \mathcal{F}_{s-\tau}^{t-\tau}$ for all $\tau \in \mathbb{T}$. Naturally, the σ-algebra $\mathcal{F}$ is taken to be the smallest σ-algebra containing $\mathcal{F}_s^t$ for all $s, t \in \mathbb{T}$ with $s \leq t$.

Since the noise is random, we assume that we have a probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ representing the probability distribution for how our plot will turn out; again, time-homogeneity will mean that the probability measure $\mathbb{P}$ does not depend on which time was chosen as our reference time. Of course, for consistency, this demands that $\mathbb{P}$ is θ^t-invariant for all $t \in \mathbb{T}$.

Now imagine that we have some process that is affected by the noise process, in the following manner: if the process is at state $x \in X$ at our reference time 0, and if our plot of the noise is given by ω, then the process will be at state $\varphi(t,\omega)x$ at time $t \in \mathbb{T}^+$, where $\varphi(t,\omega)$ is some function from X to X. (Of course, for consistency, we must have that $\varphi(0,\omega)$ is the identity for all ω.) Let us also assume that the precise manner in which the behaviour of the noise dictates the evolution of the process is itself time-homogeneous. This implies that the function $\varphi(t,\omega)$ does not depend on our chosen reference time. Hence in particular, if the process is at state $y \in X$ at time $s \in \mathbb{T}$ (relative to our chosen reference time), and if our plot of the noise is given by ω, then (for any $t \in \mathbb{T}^+$) the process will be at state $\varphi(t,\theta^s\omega)y$ at time $s+t$. Hence, for consistency, we require that $\varphi(s+t,\omega) = \varphi(t,\theta^s\omega) \circ \varphi(s,\omega)$ for all $s,t \in \mathbb{T}^+$ and $\omega \in \Omega$.

Not surprisingly, we also wish to assume that the behaviour of our process during a given time interval is not affected by the behaviour of the noise *outside* of that same time interval. This is represented mathematically by the assumption that the map $(\omega,x) \mapsto \varphi(t,\theta^s\omega)x$ is $(\mathcal{F}_s^{s+t} \otimes \Sigma)$-measurable for each $s \in \mathbb{T}$ and $t \in \mathbb{T}^+$. (For this, it is sufficient just to consider $s = 0$.)

Finally, we also assume that the noise is statistically "memoryless", which is represented by the assumption that $\mathcal{F}_0^\infty$ and $\mathcal{F}_{-\infty}^0$ are independent under $\mathbb{P}$.

2.3 Examples of random dynamical systems

"Standard form" of a discrete-time RDS

Let $(I,\mathcal{I},\nu)$ be a probability space, and let $(\Omega,\mathcal{F},(\mathcal{F}_m^{m+r}),(\theta^n),\mathbb{P})$ be as is Example 2.5. Let $(f_\alpha)_{\alpha\in I}$ be a family of functions $f_\alpha : X \to X$ such that the map $(\alpha,x) \mapsto f_\alpha(x)$ is measurable. Then we can define the family $\varphi = (\varphi(n,\omega))_{n\in\mathbb{N}_0,\,\omega\in\Omega}$ of functions $\varphi(n,\omega) : X \to X$ by

$$\varphi(n,(\alpha_r)_{r\in\mathbb{Z}}) \;=\; f_{\alpha_n} \circ \ldots \circ f_{\alpha_1}.$$

It is not hard to show that φ is a RDS (in accordance with our above formalism). We refer to φ as *the RDS generated by the random map* $(I,\mathcal{I},\nu,(f_\alpha)_{\alpha\in I})$.

Stochastic differential equations

Just as the prototypical continuous-time deterministic dynamical systems are those generated by differential equations, so likewise the prototypical continuous-time *random* dynamical systems are those generated by *stochastic differential equations* (SDE). If X is a Euclidean space $\mathbb{R}^d$ or a more general smooth manifold,[5] such equations will often

[5]Processes whose state space is *infinite-dimensional* (e.g. heat distribution in a room) are often described by *stochastic partial differential equations*. (See e.g. [KS12], which considers RDS generated by stochastically perturbed Navier-Stokes equations.)

take the form

$$dx_t \;=\; b(x_t)dt + \sum_{i=1}^{k} \sigma_i(x_t) \circ dW_t^i \tag{2.1}$$

where $b, \sigma_1, \ldots, \sigma_k$ are vector fields on X, and $(W_t^1), \ldots, (W_t^k)$ are independent one-dimensional Wiener processes (although more general Lévy processes can certainly be considered). We refer to b as the *drift vector field* (or *drift coefficient*), and we refer to $\sigma_1, \ldots, \sigma_k$ as the *diffusion vector fields* (or *diffusion coefficients*). The circle $\circ$ indicates that this equation is to be interpreted as a Stratonovich integral equation. (If X is a manifold, this is done locally, via charts.) In the case that $k = 1$, an alternative sample-pathwise interpretation (making no reference to the underlying probability measure) exists, namely, to regard (2.1) as the result of a kind of "linear superposition" of the drift vector field b over the non-autonomous flow on X whose time-t mappings are the time-W_t mappings of the flow generated by the diffusion vector field σ; this interpretation can be formulated rigorously, and the equivalence (modulo null sets) of the two interpretations is provided by the *Doss-Sussmann* theorem (e.g. [Sus78]).

If $X = \mathbb{R}^d$, one can alternatively work with equations of the form

$$dx_t \;=\; b(x_t)dt + \sum_{i=1}^{k} \sigma_i(x_t) dW_t^i \tag{2.2}$$

where the lack of the circle indicates that the equation is to be interpreted as an Itō integral equation. The equation is said to be an *additive noise SDE* if the diffusion coefficients $\sigma_1, \ldots, \sigma_k$ are all constant. In this case, there is no difference between the Itō and the Stratonovich formulation: both reduce to the natural sample-pathwise interpretation, namely as a *Volterra integral equation*

$$x_t \;=\; x_0 + \int_0^t b(x_s)\, ds + \sum_{i=1}^{k} \sigma_i W_t^i. \tag{2.3}$$

Now (2.3) can be re-expressed as a (classical) differential equation

$$\dot{y} \;=\; b\left(y + \sum_{i=1}^{k} \sigma_i W_t^i \right) \tag{2.4}$$

where $y(t) = x_t - \sum_{i=1}^{k} \sigma_i W_t^i$. If b is locally Lipschitz then solutions are unique (over any time interval), and exist for as long as they do not blow up in magnitude to ∞. (See [Bur83, Theorems 3.1.3 and 3.3.1].) If, in addition, there exists $\lambda \in \mathbb{R}$ such that

$$(b(y) - b(x)) \cdot (y - x) \;\leq\; \lambda |y - x|^2 \quad \forall\, x, y \in \mathbb{R}^d \tag{2.5}$$

then (among other useful properties) solutions never blow up in forward time. A function b with this property is said to be *one-sided Lipschitz*, and such a value $\lambda \in \mathbb{R}$ is called a *one-sided Lipschitz constant of* b. If b is C^1, then it is easy to show that (2.5) is equivalent to

$$h \cdot Db(x)h \;\leq\; \lambda \quad \forall\, x, h \in \mathbb{R}^d \text{ with } |h| = 1. \tag{2.6}$$

The one-sided Lipschitz property provides an upper bound on how quickly different trajectories can separate under the same realisation of the noise: if $(x_t^1)_{t\geq 0}$ and $(x_t^2)_{t\geq 0}$

are forward-time solutions of (2.3) under the same sample paths of the Wiener processes (W_t^i), and if b is locally Lipschitz and satisfies (2.5), then Grönwall's inequality (applied to $|x^1 - x^2|^2$) gives that

$$|x^1_{s+t} - x^2_{s+t}| \;\leq\; e^{\lambda t}|x^1_s - x^2_s| \quad \forall\, s,t \geq 0. \tag{2.7}$$

A detailed study of synchronisation in additive noise SDE with locally Lipschitz and one-sided Lipshitz drift can be found in [FGS14]. A stochastic differential equation whose diffusion coefficients are not all constant is sometimes called a *multiplicative noise SDE.*

The question of exactly when and how a SDE generates a RDS is quite a complicated issue. (See [Arn98, Section 2.3] for some details.) Therefore, in this thesis, when we consider examples involving SDE, we will deal with them in a relatively informal manner—*except* in the particularly "simple" case of SDE on the circle with rigidly rotating noise, which we shall present in detail below.

Gaussian-white-noise perturbation of a vector field on the circle

We identify the circle $\mathbb{S}^1$ with $\mathbb{R}/\mathbb{Z}$ in the obvious manner, with $\pi\colon\mathbb{R}\to\mathbb{S}^1$ denoting the associated projection mapping. Given a continuous function $h\colon[0,\infty)\to\mathbb{S}^1$, it is not hard to show that there exists a continuous function $\hat{h}\colon[0,\infty)\to\mathbb{R}$ such that $\pi\circ\hat{h} = h$, and that this function is unique up to addition by a constant integer; we refer to such a function $\hat{h}$ as a *lift of* h.

Given a Lipschitz 1-periodic function $b\colon\mathbb{R}\to\mathbb{R}$ and a value $\sigma\in\mathbb{R}$, we may formally define *the RDS on* $\mathbb{S}^1$ *generated by the SDE* $d\phi_t = b(\phi_t)dt + \sigma dW_t$ to be the RDS described by the following result:

Proposition 2.9. *Let $(\Omega,\mathcal{F},(\mathcal{F}_s^{s+t}),(\theta^t),\mathbb{P})$ be as in Example 2.6 with $d = 1$, and equip Ω with the topology of uniform convergence on compact sets. Let $b\colon\mathbb{R}\to\mathbb{R}$ be a Lipschitz 1-periodic function, and fix any $\sigma\in\mathbb{R}$. There exists a RDS φ on $\mathbb{S}^1$ over $(\Omega,\mathcal{F},(\mathcal{F}_s^{s+t}),(\theta^t),\mathbb{P})$ with the following properties:*

(i) the map $(t,\omega,x)\mapsto\varphi(t,\omega)x$ from $[0,\infty)\times\Omega\times\mathbb{S}^1$ to $\mathbb{S}^1$ is continuous;

(ii) given any $y\in\mathbb{R}$ and $\omega\in\Omega$, letting $r\colon[0,\infty)\to\mathbb{R}$ denote the unique lift of the function $t\mapsto\varphi(t,\omega)\pi(y)$ satisfying $r(0)=y$, r has the property that for all $T>0$, the only function $u\colon[0,T]\to\mathbb{R}$ satisfying $b\circ u\in\mathcal{L}^1([0,T])$ and

$$u(t) \;=\; y + \int_0^t b(u(s))\,ds + \sigma\omega(t) \quad \forall\, t\in[0,T]$$

is precisely the function $u := r|_{[0,T]}$.

(Note that if $\sigma = 0$ then $\varphi(t,\omega)$ is independent of ω, and is simply the solution flow for the ODE $\dot{\phi} = b(\phi)$ on $\mathbb{S}^1$.)

Proof of Proposition 2.9. We will first show that the SDE $dX_t = b(X_t)dt + \sigma dW_t$ on $\mathbb{R}$ naturally generates an RDS on $\mathbb{R}$ with the desired continuity properties. We will then show that the solution of this SDE depends 1-periodically on its initial condition, enabling us to project the RDS onto $\mathbb{S}^1$.

Since b is globally Lipschitz, by standard results regarding existence and uniqueness of solutions of integral equations (see e.g. Section 3.2 of [Bur83]), we have that for every $y \in \mathbb{R}$ and $\omega \in \Omega$ there exists a continuous function $r_{y,\omega}\colon[0,\infty) \to \mathbb{R}$ with the property that for all $T > 0$, the only function $u\colon[0,T] \to \mathbb{R}$ satisying $b \circ u \in L^1([0,T])$ and

$$u(t) \;=\; y + \int_0^t b(u(s))\,ds + \sigma\omega(t) \quad \forall\, t \in [0,T]$$

is the function $u := r_{y,\omega}|_{[0,T]}$. We now show that the map $(t,\omega,y) \mapsto r_{y,\omega}(t)$ is continuous. Fix any convergent sequence (t_n) in $[0,\infty)$ with limit t, any sequence (ω_n) in Ω which converges uniformly on compact sets to a sample point ω, any convergent sequence (y_n) in $\mathbb{R}$ with limit y. Fix any $\varepsilon > 0$. Let $\delta > 0$ be such that for all $s \in [0,\infty)$ with $|s-t| < \delta$, $|r_{y,\omega}(s) - r_{y,\omega}(t)| < \frac{\varepsilon}{2}$. Let $L > 0$ be a Lipschitz constant of b, and let $N \in \mathbb{N}$ be sufficiently large that for all $n \geq N$, the following statements hold:

(i) $|t_n - t| < \delta$;

(ii) $|y_n - y| + \max_{s\in[0,t+\delta]} |\omega_n(s) - \omega(s)| < \frac{\varepsilon}{2e^{L(t+\delta)}}$.

By [Bur83, Theorem 3.4.1], for all $n \geq N$ and $s \in [0, t+\delta]$ we have that

$$|r_{y_n,\omega_n}(s) - r_{y,\omega}(s)| \;<\; \left(\frac{\varepsilon}{2e^{L(t+\delta)}}\right) e^{Ls} \;\leq\; \frac{\varepsilon}{2}.$$

Therefore, for all $n \geq N$,

$$\begin{aligned} |r_{y_n,\omega_n}(t_n) - r_{y,\omega}(t)| \;&\leq\; |r_{y_n,\omega_n}(t_n) - r_{y,\omega}(t_n)| + |r_{y,\omega}(t_n) - r_{y,\omega}(t)| \\ &<\; \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &=\; \varepsilon \end{aligned}$$

as required.

Note in particular that for each t, since the map $(y, \omega|_{[0,t]}) \mapsto r_{y,\omega}(t)$ from $\mathbb{R} \times C_0([0,t],\mathbb{R})$ to $\mathbb{R}$ is continuous (where $C_0([0,t],\mathbb{R})$ is the set of all $f \in C([0,t],\mathbb{R})$ with $f(0) = 0$, equipped with the topology of uniform convergence), it follows by Lemma B.10 that the map $(y,\omega) \mapsto r_{y,\omega}(t)$ is $(\mathcal{B}(\mathbb{R}) \otimes \mathcal{F}_0^t, \mathcal{B}(\mathbb{R}))$-measurable.

Now obviously $r_{y,\omega}(0) = y$ for all y and ω. Moreover, fixing any $\tau \geq 0$, $\omega \in \Omega$ and $y \in \mathbb{R}$, setting $u(t) := r_{y,\omega}(\tau + t)$ for all $t \geq 0$, we have that

$$\begin{aligned} u(t) \;&=\; y + \int_0^{\tau+t} b(r_{y,\omega}(s))\,ds + \sigma\omega(\tau+t) \\ &=\; y + \int_0^{\tau} b(r_{y,\omega}(s))\,ds + \int_0^t b(r_{y,\omega}(\tau+s))\,ds + \sigma\omega(\tau+t) \\ &=\; y + \int_0^{\tau} b(r_{y,\omega}(s))\,ds + \int_0^t b(r_{y,\omega}(\tau+s))\,ds + \sigma\omega(\tau) + \sigma\theta^{\tau}\omega(t) \\ &=\; r_{y,\omega}(\tau) + \int_0^t b(u(s))\,ds + \sigma\theta^{\tau}\omega(t). \end{aligned}$$

So, for all $t \geq 0$, we have that $r_{y,\omega}(\tau + t) = r_{r_{y,\omega}(\tau),\theta^\tau\omega}(t)$.

So then, defining the map $\hat{\varphi}(t,\omega)\colon y \mapsto r_{y,\omega}(t)$ for all $t \geq 0$ and $\omega \in \Omega$, we have shown that $(\hat{\varphi}(t,\omega))_{t\geq 0,\,\omega\in\Omega}$ is a RDS on $\mathbb{R}$ over the noise space $(\Omega, \mathcal{F}, (\mathcal{F}_s^{s+t}), (\theta^t), \mathbb{P})$, with the additional property that the map $(t,\omega,y) \mapsto \hat{\varphi}(t,\omega)y$ is continuous.

Now fixing any $y \in \mathbb{R}$, $\omega \in \Omega$ and $k \in \mathbb{Z}$, setting $u(t) := r_{y,\omega}(t) + k$ for all $t \geq 0$, we have that

$$\begin{aligned} u(t) &= y + \int_0^t b(r_{y,\omega}(s))\,ds + \sigma\omega(t) + k \\ &= y + k + \int_0^t b(u(s) - k)\,ds + \sigma\omega(t) \\ &= y + k + \int_0^t b(u(s))\,ds + \sigma\omega(t). \end{aligned}$$

So, for all $t \geq 0$, we have that $r_{y,\omega}(t) + k = r_{y+k,\omega}(t)$, and therefore $\pi(r_{y,\omega}(t)) = \pi(r_{y+k,\omega}(t))$.

So then, for each $t \geq 0$ and $\omega \in \Omega$ we can define the map $\varphi(t,\omega)\colon \mathbb{S}^1 \to \mathbb{S}^1$ by

$$\varphi(t,\omega)\pi(y) = \pi(\hat{\varphi}(t,\omega)y) \quad \forall\, y \in \mathbb{R}.$$

It is easy to show that $(\varphi(t,\omega))_{t\geq 0,\,\omega\in\Omega}$ is itself an RDS, with the desired properties. □

Time-discretisation

Let $(\Omega, \mathcal{F}, (\mathcal{F}_s^{s+t})_{s\in\mathbb{R},\,t\geq 0}, (\theta^t)_{t\in\mathbb{R}}, \mathbb{P})$ be a noise space, and let φ be a RDS over this noise space. Fix any $\tau > 0$. Then over the noise space $(\Omega, \mathcal{F}, (\mathcal{F}_{m\tau}^{(m+r)\tau})_{m\in\mathbb{Z},r\in\mathbb{N}_0}, (\theta^{n\tau})_{n\in\mathbb{Z}}, \mathbb{P})$ we may define the RDS $\mathring{\varphi}_\tau$ by $\mathring{\varphi}_\tau(n,\omega) = \varphi(n\tau,\omega)$ for all $n \in \mathbb{N}_0$ and $\omega \in \Omega$. We refer to $\mathring{\varphi}_\tau$ as the *time-τ discretisation of φ*.

2.4 Other formalisms of random dynamical systems

(See also Chapter 1 of [Arn98].)

Throughout the rest of this thesis (except this section), when we refer to a "random dynamical system", we will specifically mean a "memoryless random dynamical system" as defined in accordance with our formalism in Section 2.2. However, it is worth mentioning some alternative (mostly, more general) notions of a RDS.

(For convenience we will still, throughout this section, regard the state space of any kind of "random dynamical system" as having the measurable structure of a *standard measurable space*; although some references do not specifically include this in the definition, it is crucial for the most basic tools of the theory of RDS to be applied.)

RDS without a filtration

Central to our formalism is the underlying two-sided filtration $(\mathcal{F}_s^{s+t})_{s\in\mathbb{T},\,t\in\mathbb{T}^+}$ on the sample space Ω. However, in general, the term "random dynamical system" does not necessarily imply the presence or relevance of any pre-defined filtration on Ω; a "random dynamical system" merely consists of:

- a group $\theta := (\theta^t)_{t\in\mathbb{T}}$ of measure-preserving transformations of a probability space $(\Omega, \mathcal{F}, \mathbb{P})$;[6]
- a standard measurable space (X, Σ), and a family $\varphi = (\varphi(t,\omega))_{t\in\mathbb{T}^+, \omega\in\Omega}$ of functions $\varphi(t,\omega)\colon X \to X$, satisfying the "forward cocycle" property (i.e. properties (i) and (ii) in our formalism) and

  (iii)' for each $t \in \mathbb{T}^+$, the map $(\omega, x) \mapsto \varphi(t,\omega)x$ is $(\mathcal{F}\otimes\Sigma, \Sigma)$-measurable.

By identifying (θ, φ) with the "product system" $\Theta^t\colon(\omega, x) \mapsto (\theta^t\omega, \varphi(t,\omega)x)$, we have an equivalent, slightly more succint formulation: A "random dynamical system" is a skew-product dynamical system on $\Omega \times X$ (with X standard) whose base system is invertible and is equipped with an invariant probability measure.[7]

The above definition of a "general" random dynamical system still captures the notion of a time-homogeneous noise process determining the evolution of some system in a time-homogeneous manner; but it does not incorporate a notion of "how the noise behaves over a given finite time interval".

Filtered RDS that are not necessarily memoryless

One can also consider a "random dynamical system" as defined in accordance with our formalism, *except* without having to satisfy that $\mathcal{F}_{-\infty}^t$ and $\mathcal{F}_t^\infty$ are independent under $\mathbb{P}$ for each t.

Examples of non-memoryless filtered RDS include: (a) RDS generated by SDE (with time-independent vector fields) that are driven by processes with strictly stationary but not independent incremements; and (b) RDS generated by random differential equations (RDE) of the form

$$\dot{x}(t) \;=\; f(\,\omega(t)\,,\,x(t))$$

where the sample points ω are functions from $\mathbb{R}$ to some suitable space W, $\mathbb{P}$ is invariant under the shift $\theta^t\omega(s) = \omega(s+t)$, and $f(y,\cdot)$ is a vector field on X for each $y \in W$.[8]

"Semifiltered" RDS

The term "filtered RDS" has been used elsewhere to describe what we will call a "semifiltered" RDS, where the system may be able to "remember" the past behaviour of

[6]Sometimes one also assumes that $\mathbb{P}$ is an ergodic probability measure of the family of transformations (θ^t), in which case the RDS is sometimes referred to as an *ergodic RDS*.

[7]A *skew-product dynamical system* on a product space $(\Omega \times X, \mathcal{F}\otimes\Sigma)$ is an autonomous dynamical system $(\Theta^t)_{t\in\mathbb{T}^+}$ on $(\Omega \times X, \mathcal{F}\otimes\Sigma)$ such that the Ω-component of $\Theta^t(\omega, x)$ does not depend on x. The *base system* of a skew-product dynamical system (Θ^t) on $\Omega \times X$ is the dynamical system on Ω obtained by projecting the dynamics of Θ^t onto Ω. See also Section 2.6.

[8]A more general time-homogeneous RDE takes the form $\dot{x}(t) = f(\,\theta^t\omega\,,\,x(t))$ where $(\theta^t)_{t\in\mathbb{R}}$ is a measure-preserving group of transformations of $(\Omega, \mathcal{F}, \mathbb{P})$, and $f(\omega,\cdot)$ is a vector field on X for each ω.

the noise, but cannot predict the future behaviour of the noise. In its simplest formulation, a semifiltered RDS consists of:

- a measurable space $(\Omega, \mathcal{F})$, and a $\mathbb{T}$-indexed family $(\mathcal{F}_t)_{t\in\mathbb{T}}$ of sub-σ-algebras of $\mathcal{F}$ such that $\mathcal{F}_s \subset \mathcal{F}_t$ for $s \le t$ and $\mathcal{F} = \sigma(\mathcal{F}_t : t \in \mathbb{T})$;
- a group $(\theta^t)_{t\in\mathbb{T}}$ of functions $\theta^t : \Omega \to \Omega$ such that $\theta^\tau \mathcal{F}_t = \mathcal{F}_{t-\tau}$ for all $t, \tau \in \mathbb{T}$;
- a (θ^t)-invariant probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$;
- a standard measurable space (X, Σ), and a RDS φ on X over $(\Omega, \mathcal{F}, \mathbb{P}, (\theta^t)_{t\in\mathbb{T}})$ such that the map $(\omega, x) \mapsto \varphi(t, \omega)x$ is $(\mathcal{F}_t \otimes \Sigma, \Sigma)$-measurable for each $t \in \mathbb{T}^+$.

Note that any filtered RDS can be regarded as a semifiltered RDS, by setting $\mathcal{F}_t := \mathcal{F}^t_{-\infty}$; and semifiltered RDS can be regarded as a special case of filtered RDS, by setting $\mathcal{F}^t_s := \mathcal{F}_t$. For examples of semifiltered RDS, see e.g. [CSS05], or the random differential equations in [IL02] constructed to aid the study of Wiener-driven SDE.

RDS with non-invertible base

For mathematical reasons, we always consider RDS whose base system is a *group* $(\theta^t)_{t\in\mathbb{T}}$ of $\mathbb{P}$-preserving transformations. However, one can certainly also consider RDS whose base system is a $\mathbb{P}$-preserving dynamical system $(\theta^t)_{t\in\mathbb{T}^+}$ that is not necessarily invertible. (Indeed, this plays a key role in the proof of the correpondence between one-sided-time and two-sided-time invariant measures described in Section 1.3.)

In this case (following [New15a, Section 7]), if we wish the RDS to be a "filtered RDS", we equip Ω with a *one-sided* filtration $(\mathcal{F}_t)_{t\in\mathbb{T}^+}$ of sub-σ-algebras of $\mathcal{F}$ such that for all $s, t \in \mathbb{T}^+$, θ^t is $(\mathcal{F}_{s+t}, \mathcal{F}_s)$-measurable; the RDS φ is filtered with respect to this filtration if for each $t \in \mathbb{T}^+$, the map $(\omega, x) \mapsto \varphi(t, \omega)x$ is $(\mathcal{F}_t \otimes \Sigma, \Sigma)$-measurable. The noise space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{T}^+}, (\theta^t)_{t\in\mathbb{T}^+}, \mathbb{P})$ is *memoryless* if $\mathbb{P}(E \cap \theta^{-t}(F)) = \mathbb{P}(E)\mathbb{P}(F)$ for all $t \in \mathbb{T}^+$, $E \in \mathcal{F}_t$ and $F \in \sigma(\mathcal{F}_s : s \in \mathbb{T}^+)$. Note that a "memoryless noise space" in the sense of Section 2.2 can be regarded as a memoryless noise space in this sense, by setting $\mathcal{F}_t := \mathcal{F}^t_0$.

Local RDS

Now it is, of course, entirely possible to have a stochastic differential equation whose forward-time solutions blow up in finite time. This naturally motivates the study of *local RDS*. Specifically (roughly following [FGS14]), we can define a "local RDS on X" to be a RDS φ on an "extended phase space" $X \cup \{\partial\}$ (equipped with the obvious σ-algebra), such that $\varphi(t, \omega)\partial = \partial$ for all t and ω.

Bundle RDS

Another generalisation of the notion of a RDS is the notion of a *bundle RDS*, where the set of possible states of the system evolves over time in accordance with the noise. We present two possible definitions of a bundle RDS over an invertible measure-preserving dynamical system $(\Omega, \mathcal{F}, \mathbb{P}, (\theta^t)_{t\in\mathbb{T}})$.[9]

[9]The author has not seen explicitly the first of the two definitions elsewhere; however, it is useful for motivating the definition of cohomology of random dynamical systems.

• In the first case: We have a measurable space $(Y, \mathcal{Y})$ and a surjective $(\mathcal{Y}, \mathcal{F})$-measurable function $\pi: Y \to \Omega$ such that for each $\omega \in \Omega$, the set $Y_\omega := \pi^{-1}(\Omega)$ equipped with the induced σ-algebra $\mathcal{Y}_\omega$ from $\mathcal{Y}$ is a standard measurable space. And we have a family $\varphi = (\varphi(t,\omega))_{t\in\mathbb{T}^+, \omega\in\Omega}$ of functions $\varphi(t,\omega): Y_\omega \to Y_{\theta^t\omega}$ such that

(i) $\varphi(0,\omega) = \mathrm{id}_{Y_\omega}$ for all $\omega \in \Omega$;

(ii) $\varphi(s+t,\omega) = \varphi(t,\theta^s\omega) \circ \varphi(s,\omega)$ for all $s,t \in \mathbb{T}^+$ and $\omega \in \Omega$;

(iii) for each $t \in \mathbb{T}^+$, the map $y \mapsto \varphi(t,\pi(y))y$ is $(\mathcal{Y}, \mathcal{Y})$-measurable.

It follows that $\varphi(t,\omega)$ is $(\mathcal{Y}_\omega, \mathcal{Y}_{\theta^t\omega})$-measurable for all t and ω.

• In the second case (cf. [Arn98, Definition 1.9.1]): We have a standard measurable space $(Z, \mathcal{Z})$ and a set $Y \in \mathcal{F} \otimes \mathcal{Z}$, such that for each $\omega \in \Omega$, the ω-section $Y_\omega := \{x \in Z : (\omega, x) \in Y\}$ of Y is nonempty. And we have a family $\varphi = (\varphi(t,\omega))_{t\in\mathbb{T}^+, \omega\in\Omega}$ of functions $\varphi(t,\omega): Y_\omega \to Y_{\theta^t\omega}$ satisfying (i) and (ii) above, as well as

(iii)' for each $t \in \mathbb{T}^+$, the map $(\omega, x) \mapsto \varphi(t,\omega)x$ is $(\mathcal{Y}, \mathcal{Z})$-measurable,

where $\mathcal{Y}$ denotes the set of all $(\mathcal{F} \otimes \mathcal{Z})$-measurable subsets of Y. Note that a RDS φ on a phase space X may be regarded as the "trivial case" of a bundle RDS, simply by taking $Z = X$ and $Y = \Omega \times X$.

Now a bundle RDS φ according to the second definition may be regarded as having the structure of a bundle RDS under the first definition, by the identification $\varphi(t,\omega): (\omega, x) \mapsto (\theta^t\omega, \varphi(t,\omega)x)$. However, a bundle RDS under the second definition has additional structure which a bundle RDS under the first definition does not have: namely, under the second definition, one can consider *intersections* of the form $\bigcap_{\omega\in E} Y_\omega$ where $E \subset \Omega$. (For bundle RDS according to the first definition, if E has more than one element then this intersection is always empty.)

An important motivation for the notion of a bundle RDS (under the second definition) is the following: Given a deterministic dynamical system $(f^t)_{t\in\mathbb{T}^+}$ on (X, Σ) with a forward-invariant set $Y \in \Sigma$, we can restrict (f^t) to Y, to obtain a dynamical system on Y. However, the RDS-analogue of the notion of an "invariant set" is a "random invariant set", which is an ω-dependent subset of X. The restriction of a RDS φ on X to a "random forward-invariant set" $(Y_\omega)_{\omega\in\Omega}$ will be a bundle RDS.

(Now we warn that the term "bundle RDS" is sometimes also used to refer to a RDS on a fixed phase space, whose mappings $\varphi(t,\omega)$ take the form of a skew-product map. [Arn98] uses the term "bundle RDS" in both senses.)

Backward cocycles

Now as we will see, pullback trajectories often play a very significant role in the theory of RDS. Pullback trajectories of a RDS φ can simply be regarded as trajectories of the system $\psi := (\varphi(t,\theta^{-t}\omega))_{t\in\mathbb{T}^+, \omega\in\Omega}$, which forms a "backward cocycle" over the group

$\bar{\theta} := (\theta^{-t})_{t \in \mathbb{T}}$. In general, a *backward cocycle* on X over a group of maps $(\bar{\theta}^t)_{t \in \mathbb{T}}$ on Ω is a family $\psi = (\psi(t,\omega))_{t \in \mathbb{T}^+, \omega \in \Omega}$ of functions $\psi(t,\omega): X \to X$ such that

(i) $\psi(0,\omega) = \mathrm{id}_X$ for all $\omega \in \Omega$;

(ii) $\psi(s+t,\omega) = \psi(s,\omega) \circ \psi(t, \bar{\theta}^s \omega)$ for all $s,t \in \mathbb{T}^+$ and $\omega \in \Omega$.

Given a backward cocycle $(\bar{\theta}, \psi)$, one can obtain a forward cocycle (θ, φ) by $\theta^t := \bar{\theta}^{-t}$ and $\varphi(t,\omega) := \psi(t, \bar{\theta}^{-t}\omega)$. This procedure actually inverts the above procedure for obtaining a backward cocycle from a forward cocycle. Thus forward cocycles and backward cocycles are in one-to-one correspondence.

Additional structure on the phase space

So far, we have not discussed RDS that respect some additional structure on the state space beyond measurable structure (e.g. topological structure, vector space structure, differentiable structure, Riemannian structure, partial or linear ordering). Later in this chapter, we will consider "monotone" RDS on ordered spaces, and "(right)-continuous" RDS on metrisable topological spaces.

In the case that X is equipped with the structure of a measurable vector space, a RDS φ on X is said to be *linear* if $\varphi(t,\omega)$ is a linear map for all t and ω. In the case that X is equipped with the structure of a C^k-smooth manifold, a RDS φ is said to be a *C^k-smooth RDS* if for every partial differential operator ∂_α of order less than or equal to k, $\partial_\alpha \varphi(t,\omega)$ exists for all t and ω and the map $(t,x) \mapsto (\partial_\alpha \varphi(t,\omega))(x)$ is jointly continuous for all ω.

Central to the study of linear RDS and smooth RDS is the *multiplicative ergodic theorem*, which provides the existence of "exponential separation rates" called *Lyapunov exponents*. See Part II (especially, Chapters 3 and 4) of [Arn98] for a detailed exposition of the multiplicative ergodic theorem and its major corollaries.

Isomorphism and cohomology of RDS

Of course, different generalisations of our notion of a random dynamical system give rise to different notions of *isomorphism* of random dynamical systems. We now mention some possible notions of "isomorphism". (These notions do not make reference to any filtrations on the underlying probability space.)

- Two random dynamical systems

  $$(\Omega_1, \mathcal{F}_1, \mathbb{P}_1, (\theta_1^t)_{t \in \mathbb{T}}, X_1, \Sigma_1, \varphi_1) \quad \text{and} \quad (\Omega_2, \mathcal{F}_2, \mathbb{P}_2, (\theta_2^t)_{t \in \mathbb{T}}, X_2, \Sigma_2, \varphi_2)$$

  are *isomorphic as RDS* if there exists an $(\mathcal{F}_1, \mathcal{F}_2)$-measurably invertible function $g: \Omega_1 \to \Omega_2$ and a (Σ_1, Σ_2)-measurably invertible function $h: X_1 \to X_2$ such that

  (i) $g_* \mathbb{P}_1 = \mathbb{P}_2$;

  (ii) $\theta_1^t = g^{-1} \circ \theta_2^t \circ g$ for all $t \in \mathbb{T}$;

  (iii) $\varphi_1(t,\omega) = h^{-1} \circ \varphi_2(t, g(\omega)) \circ h$ for all $t \in \mathbb{T}^+$ and $\omega \in \Omega_1$.

In this case, we refer to (g, h) as a *RDS isomorphism from* φ_1 *to* φ_2.

- Two random dynamical systems

$$(\Omega_1, \mathcal{F}_1, \mathbb{P}_1, (\theta_1^t)_{t\in\mathbb{T}}, X_1, \Sigma_1, \varphi_1) \quad \text{and} \quad (\Omega_2, \mathcal{F}_2, \mathbb{P}_2, (\theta_2^t)_{t\in\mathbb{T}}, X_2, \Sigma_2, \varphi_2)$$

are *isomorphic as bundle RDS* if there exists an $(\mathcal{F}_1 \otimes \Sigma_1, \mathcal{F}_2 \otimes \Sigma_2)$-measurably invertible function $H: \Omega_1 \times X_1 \to \Omega_2 \times X_2$ such that

(i) the Ω_2-component of $H(\omega, x)$ does not depend on x, but only on ω; so we can define the functions $g_H: \Omega_1 \to \Omega_2$ and $h_H: \Omega_1 \times X_1 \to X_2$ such that $H(\omega, x) = (g_H(\omega), h_H(\omega, x))$;

(ii) $g_{H*}\mathbb{P}_1 = \mathbb{P}_2$;

(iii) $(\theta_1^t\omega\,,\, \varphi_1(t,\omega)x) \;=\; H^{-1}(\,\theta_2^t(g_H(\omega))\,,\, \varphi_2(t, g_H(\omega))h_H(\omega, x)\,)$ for all $t \in \mathbb{T}^+$, $\omega \in \Omega_1$ and $x \in X_1$.

In this case, we refer to H as a *bundle RDS isomorphism from* φ_1 *to* φ_2. (Note that this terminology is making reference to the first of our two definitions of a bundle RDS.)

- Two RDS over the same base system $\boldsymbol{\theta}$ are called *cohomologous* if they are "isomorphic as bundle RDS over $\boldsymbol{\theta}$". To be precise: Two random dynamical systems $(X_1, \Sigma_1, \varphi_1)$ and $(X_2, \Sigma_2, \varphi_2)$ defined over a measure-preserving group $(\Omega, \mathcal{F}, \mathbb{P}, (\theta^t)_{t\in\mathbb{T}})$ are said to be *cohomologous* if there exists a bundle RDS isomorphism H from φ_1 to φ_2 such that g_H is the identity function on Ω. In this case, we refer to h_H as a *random measurable conjugacy* or a *cohomology* from φ_1 to φ_2. It is easy to show that a function $h: \Omega \times X_1 \to X_2$ is a random measurable conjugacy from φ_1 to φ_2 if and only if the following hold:

(i) for each $\omega \in \Omega$, $h(\omega, \cdot)$ serves as a bijection between X_1 and X_2; so we can define $h^{-1}: \Omega \times X_2 \to X_1$ such that for every ω, $h^{-1}(\omega, \cdot)$ is the inverse function of $h(\omega, \cdot)$;

(ii) h is $(\mathcal{F} \otimes \Sigma_1, \Sigma_2)$-measurable, and h^{-1} is $(\mathcal{F} \otimes \Sigma_2, \Sigma_1)$-measurable;

(iii) $\varphi_1(t, \omega) \;=\; h^{-1}(\theta^t\omega, \cdot) \circ \varphi_2(t, \omega) \circ h(\omega, \cdot)$ for all $t \in \mathbb{T}^+$ and $\omega \in \Omega$.

- Alternatively (for practical purposes), one may choose to regard two RDS over $(\Omega, \mathcal{F}, \mathbb{P}, (\theta^t)_{t\in\mathbb{T}})$ as being "cohomologous" if there is a strictly $(\theta^t)_{t\in\mathbb{T}}$-invariant $\mathbb{P}$-full set $\Omega' \in \mathcal{F}$ such that the restrictions of the two RDS to Ω' (defined in the obvious manner) are cohomologous under the above definition.

Cohomology can be used to define a "bifurcation" in a parametrised family of actions of the noise space $(\Omega, \mathcal{F}, \mathbb{P}, (\theta^t)_{t\in\mathbb{T}})$ on a topological space X. (Of course, in this context we would specifically consider cohomologies that "respect topological structure" in some suitable sense.)

Remark 2.10. Cohomology can also be characterised in terms of "pullback operators". Let $h: \Omega \times X_1 \to X_2$ be a function satisfying properties (i) and (ii) above. Let $\mathcal{L}^0(\Omega, \mathcal{F}; X_1)$ and $\mathcal{L}^0(\Omega, \mathcal{F}; X_2)$ denote respectively the set of measurable functions from

Ω to X_1 and the set of measurable functions from Ω to X_2. Define the function $\mathcal{H}: \mathcal{L}^0(\Omega, \mathcal{F}; X_1) \to \mathcal{L}^0(\Omega, \mathcal{F}; X_2)$ by $(\mathcal{H}a)(\omega) = h(\omega, a(\omega))$; it is clear that $\mathcal{H}$ is bijective, with inverse $\mathcal{H}^{-1}$ given by $(\mathcal{H}^{-1}b)(\omega) = h^{-1}(\omega, b(\omega))$. For each $t \in \mathbb{T}^+$, define the "pullback operator" $\mathcal{P}^t_{\varphi_1}: \mathcal{L}^0(\Omega, \mathcal{F}; X_1) \to \mathcal{L}^0(\Omega, \mathcal{F}; X_1)$ by $(\mathcal{P}^t_{\varphi_1} a)(\omega) = \varphi_1(t, \theta^{-t}\omega)a(\theta^{-t}\omega)$; and define $\mathcal{P}^t_{\varphi_2}: \mathcal{L}^0(\Omega, \mathcal{F}; X_2) \to \mathcal{L}^0(\Omega, \mathcal{F}; X_2)$ similarly. It is easy to show that h is a random measurable conjugacy if and only if

$$\mathcal{P}^t_{\varphi_1} = \mathcal{H}^{-1} \circ \mathcal{P}^t_{\varphi_2} \circ \mathcal{H} \tag{2.8}$$

for all $t \in \mathbb{T}^+$. For each t, to verify (2.8), it is sufficient just to show that (2.8) holds on a subset $\mathcal{A} \subset \mathcal{L}^0(\Omega, \mathcal{F}; X_1)$ such that $\{a(\omega) : a \in \mathcal{A}\} = X_1$ for all $\omega \in \Omega$. (Pullback operators will be discussed further in Section 2.7.)

Non-homogeneous RDS

Just as, in the deterministic context, a major area of study is the theory of non-autonomous dynamical systems, so also a growing area of research within the setting of systems affected by noise is the theory of RDS that are not time-homogeneous. The inhomogeneity is often due to some external deterministic forcing that is not constant over time, leading to the study of *non-autonomous RDS*, where φ has two time-indeces. (See e.g. [Che+15], or Section 3 of [FZ15].) One can also consider systems where the noise itself is not statistically stationary. (See e.g. [Kai93] or [SH02].[10])

2.5 Markovian dynamics

Throughout the rest of this thesis, $\mathbb{T}$, Ω, $\mathcal{F}$, $\mathcal{F}^{s+t}_s$, θ^t, $\mathbb{P}$, X, Σ and φ will be as in Section 2.2 (although, for much of the thesis, we will consider additional structure on the phase space X).

For each $x \in X$ and $t \in \mathbb{T}^+$, define the probability measure φ^t_x on X by

$$\begin{aligned} \varphi^t_x(A) &:= \mathbb{P}(\omega \in \Omega : \varphi(t,\omega)x \in A) \\ &= \mathbb{P}(\omega \in \Omega : \varphi(t,\theta^s\omega)x \in A) \quad \text{(for any } s \in \mathbb{T}\text{)}. \end{aligned}$$

In heuristic terms, for each t, $(\varphi^t_x)_{x \in X}$ represents the transition probabilities associated with the time-t mapping of the RDS φ.

We will now see that, due to the noise being memoryless, "the statistical dynamics of φ are Markovian".

Lemma 2.11. *Fix an arbitrary probability measure ρ on X. Over the probability space $(\Omega \times X, \mathcal{F} \otimes \Sigma, \mathbb{P} \otimes \rho)$, define the X-valued stochastic process $(M_t)_{t \in \mathbb{T}^+}$ by $M_t(\omega, x) = \varphi(t,\omega)x$. Then (M_t) is a homogeneous Markov process with respect to the filtration $(\mathcal{F}^t_0 \otimes \Sigma)_{t \in \mathbb{T}^+}$, with transition probabilities $(\varphi^t_x)_{x \in X,\, t \in \mathbb{T}^+}$.*

[10] Most of the content of [SH02] really concerns "RDS" *without* a probability measure on the base; but this is so that the theory can be applied to systems influenced (in a time-homogeneous manner) by non-stationary noise.

Proof. We start by verifying that the family of probability measures $(\varphi_x^t)_{x\in X,\,t\in\mathbb{T}^+}$ does indeed satisfy the Chapman-Kolmogorov relations: for any $x \in X$, $s, t \in \mathbb{T}^+$ and $A \in \Sigma$, we have

$$\begin{aligned}
\varphi_x^{s+t}(A) &= \int_\Omega \mathbb{1}_A(\varphi(t,\theta^s\omega)\varphi(s,\omega)x)\,\mathbb{P}(d\omega) \\
&= \int_{\Omega\times\Omega} \mathbb{1}_A(\varphi(t,\theta^s\tilde{\omega})\varphi(s,\omega)x)\,\mathbb{P}\otimes\mathbb{P}(d(\omega,\tilde{\omega})) \\
&\qquad \text{(by Lemma A.10, since } \mathcal{F}_0^s \text{ and } \mathcal{F}_s^{s+t} \text{ are independent)} \\
&= \int_\Omega\int_\Omega \mathbb{1}_A(\varphi(t,\theta^s\tilde{\omega})\varphi(s,\omega)x)\,\mathbb{P}(d\tilde{\omega})\,\mathbb{P}(d\omega) \\
&= \int_\Omega \varphi_{\varphi(s,\omega)x}^t(A)\,\mathbb{P}(d\omega) \\
&= \int_X \varphi_y^t(A)\,\varphi_x^s(dy) \\
&\qquad \text{(since } \varphi_x^s \text{ is precisely the image measure of } \mathbb{P} \text{ under } \omega\mapsto\varphi(s,\omega)x\text{)}
\end{aligned}$$

as required. Now we know that M_t is $(\mathcal{F}_0^t\otimes\Sigma,\Sigma)$-measurable for each $t\in\mathbb{T}^+$. Given any $s,t\in\mathbb{T}^+$ and $A\in\Sigma$, we have that for $(\mathbb{P}\otimes\rho)$-almost all (ω,x),

$$\begin{aligned}
&\mathbb{P}\otimes\rho\,(M_{s+t}\in A\,|\,\mathcal{F}_0^s\otimes\Sigma)(\omega,x) \\
&\quad= \mathbb{E}_{(\mathbb{P}\otimes\rho)}[\,(\tilde{\omega},\tilde{x})\mapsto\mathbb{1}_A(M_{s+t}(\tilde{\omega},\tilde{x}))\,|\,\mathcal{F}_0^s\otimes\Sigma\,](\omega,x) \\
&\quad= \mathbb{E}_{(\mathbb{P}\otimes\rho)}[\,(\tilde{\omega},\tilde{x})\mapsto\mathbb{1}_A(\varphi(t,\theta^s\tilde{\omega})M_s(\tilde{\omega},\tilde{x}))\,|\,\mathcal{F}_0^s\otimes\Sigma\,](\omega,x) \\
&\quad= \mathbb{E}_{(\mathbb{P}\otimes\rho)}[\,(\tilde{\omega},\tilde{x})\mapsto\mathbb{1}_A(\varphi(t,\theta^s\tilde{\omega})M_s(\omega,x))\,] \\
&\qquad \text{(by Corollary A.11, since } \mathcal{F}_0^s\otimes\Sigma \text{ and } \mathcal{F}_s^{s+t}\otimes\{\varnothing,X\} \text{ are } (\mathbb{P}\otimes\rho)\text{-independent)} \\
&\quad= \mathbb{E}_{(\mathbb{P})}[\,\tilde{\omega}\mapsto\mathbb{1}_A(\varphi(t,\theta^s\tilde{\omega})M_s(\omega,x))\,] \\
&\quad= \varphi_{M_s(\omega,x)}^t(A)
\end{aligned}$$

as required. □

Hence in particular:

Corollary 2.12. *Fix any $x_0\in X$. Over $(\Omega,\mathcal{F},\mathbb{P})$, define the X-valued stochastic process $(M_t)_{t\in\mathbb{T}^+}$ by $M_t(\omega)=\varphi(t,\omega)x_0$. Then $(M_t)_{t\in\mathbb{T}^+}$ is a homogeneous Markov process with respect to $(\mathcal{F}_0^t)_{t\in\mathbb{T}^+}$, with transition probabilities $(\varphi_x^t)_{x\in X,\,t\in\mathbb{T}^+}$*

It is easy to prove Corollary 2.12 simply by going through the same proof as for Lemma 2.11; alternatively, one can derive Corollary 2.12 as a special case of Lemma 2.11 using Lemma A.14, with $Y(\omega):=(\omega,x_0)$.

Now for any $t\in\mathbb{T}^+$ and any probability measure ρ on X, define the probability measure $\varphi^{t*}\rho$ on X by

$$\varphi^{t*}\rho(A) := \int_X \varphi_x^t(A)\,\rho(dx).$$

Using Fubini's theorem, we have

$$\begin{aligned}
\varphi^{t*}\rho(A) &\overset{\text{def}}{=} \int_X \mathbb{P}(\omega:\varphi(t,\omega)x\in A)\,\rho(dx) \\
&= \mathbb{P}\otimes\rho(\,(\omega,x)\,:\,\varphi(t,\omega)x\in A\,) \\
&= \int_\Omega \varphi(t,\omega)_*\rho(A)\,\mathbb{P}(d\omega).
\end{aligned}$$

Obviously then, we also have that for any $s \in \mathbb{T}$, $\varphi^{t*}\rho$ is the image measure of $\mathbb{P} \otimes \rho$ under $(\omega, x) \mapsto \varphi(t, \theta^s\omega)x$ and

$$\varphi^{t*}\rho(A) = \int_\Omega \varphi(t, \theta^s\omega)_*\rho(A)\,\mathbb{P}(d\omega)$$

for all $A \in \Sigma$.

Since the family of probability measures $(\varphi^t_x)_{x\in X,\, t\in\mathbb{T}^+}$ satisfies the Chapman-Kolmogorov relations, it is not hard to show that $\varphi^{s+t*}\rho = \varphi^{t*}(\varphi^{s*}\rho)$ for all $s, t \in \mathbb{T}^+$ and $\rho \in \mathcal{M}_{(X,\Sigma)}$.

Note that (by definition) a probability measure ρ on X is stationary under the Markov transition probabilities $(\varphi^t_x)_{x\in X,\, t\in\mathbb{T}^+}$ if and only if $\varphi^{t*}\rho = \rho$ for all $t \in \mathbb{T}^+$.

Remark 2.13. (I) Let ρ be a probability measure on X, and let $(M_t)_{t\in\mathbb{T}^+}$ be as in Lemma 2.11. Then for each $t \in \mathbb{T}^+$, the law $M_{t*}(\mathbb{P} \otimes \rho)$ of the random variable M_t is precisely $\varphi^{t*}\rho$. (II) Fix any $x_0 \in X$, and let $(M_t)_{t\in\mathbb{T}^+}$ be as in Corollary 2.12. Then for each $t \in \mathbb{T}^+$, the law $(M_t)_*\mathbb{P}$ of the random variable M_t is precisely $\varphi^t_{x_0}$.

Remark 2.14. In the deterministic case that Ω is a singleton $\{\omega\}$, writing $f^t := \varphi(t, \omega)$, we have that $\varphi^t_x = \delta_{f^t(x)}$ for all x and t, and therefore $\varphi^{t*}\rho = f^t_*\rho$ for any probability measure ρ on X and any t; so then, a probability measure ρ is stationary under the Markov transition probabilities $(\varphi^t_x)_{x\in X,\, t\in\mathbb{T}^+}$ if and only if ρ is invariant under the dynamical system $(f^t)_{t\in\mathbb{T}^+}$.

Lemma 2.15. *Suppose φ is measurable, and let ρ be a probability measure on X that is ergodic with respect to the Markov transition probabilities $(\varphi^t_x)_{x\in X,\, t\in\mathbb{T}^+}$. Then for each $A \in \Sigma$, for $(\mathbb{P} \otimes \rho)$-almost all $(\omega, x) \in \Omega \times X$, we have*

$$\frac{1}{n}\sum_{i=0}^{n-1} \mathbb{1}_A(\varphi(i, \omega)x) \to \rho(A) \quad \text{as } n \to \infty \qquad \text{if } \mathbb{T} = \mathbb{Z},$$
$$\frac{1}{T}\int_0^T \mathbb{1}_A(\varphi(t, \omega)x)\,dt \to \rho(A) \quad \text{as } T \to \infty \qquad \text{if } \mathbb{T} = \mathbb{R}.$$

Proof. Follows immediately from Lemma 2.11 and the ergodic theorem for Markov processes.[11] (See Sections C.5 and C.6.) □

Definition 2.16. We say that a probability measure ρ on X is *incompressible* (under φ) if $\mathbb{P}$-almost every $\omega \in \Omega$ has the property that for all $t \in \mathbb{T}^+$, $\varphi(t, \omega)_*\rho = \rho$. We say that ρ is *crudely incompressible* (under φ) if for each $t \in \mathbb{T}^+$, $\mathbb{P}(\omega : \varphi(t, \omega)_*\rho = \rho) = 1$.

Obviously if ρ is crudely incompressible then ρ is stationary under the Markov transition probabilities $(\varphi^t_x)_{x\in X,\, t\in\mathbb{T}^+}$.

[11] If, in addition to the hypotheses of Lemma 2.15, the map $(t, \omega) \mapsto \theta^t\omega$ is jointly measurable, then the conclusion of Lemma 2.15 can be obtained by an alternative means: namely, it follows from Lemma 2.21(ii), together with Birkhoff's ergodic theorem for the dynamical system $(\Theta^t)_{t\in\mathbb{T}^+}$ applied to the function $\mathbb{1}_{\Omega\times A}$.

Definition 2.17. We say that a set $A \subset X$ is *invariant* (under φ) if $\mathbb{P}$-almost every $\omega \in \Omega$ has the property that for all $t \in \mathbb{T}^+$, $\varphi(t,\omega)A \subset A$. We say that $A \subset X$ is *crudely invariant* (under φ) if for each $t \in \mathbb{T}^+$, there is a $\mathbb{P}$-full set $\Omega_t \subset \Omega$ such that for all $\omega \in \Omega_t$, $\varphi(t,\omega)A \subset A$. We say that $A \in \Sigma$ is *very crudely invariant* (under φ) if $\varphi_x^t(A) = 1$ for all $x \in A$ and $t \in \mathbb{T}^+$ (i.e. if A is forward-invariant according to the Markov transition probabilities $(\varphi_x^t)_{x\in X,\, t\in\mathbb{T}^+}$).

Obviously a finite or countable set is crudely invariant if and only if it is very crudely invariant.

Definition 2.18. We will say that a point $p \in X$ is a *deterministic fixed point* if $\{p\}$ is invariant; and we will say that $p \in X$ is a *crude deterministic fixed point* if $\{p\}$ is crudely invariant.

2.6 Skew-product dynamics

One important tool for analysing the behaviour of a RDS is its representation as a "skew-product" flow on the product space $\Omega \times X$; while fully encoding the dynamics of φ, this is a *deterministic* dynamical system (thus giving a crucial role to deterministic ergodic theory for the study of the dynamics of a RDS).

For any $\tau \in \mathbb{T}$ and $t \in \mathbb{T}^+$, define the map

$$\begin{aligned} \Theta_\tau^t : \Omega \times X \ &\to\ \Omega \times X \\ \Theta_\tau^t(\omega, x) \ &=\ (\theta^t\omega, \varphi(t, \theta^\tau\omega)x). \end{aligned}$$

Lemma 2.19. *For each $\tau \in \mathbb{T}$, the family $(\Theta_\tau^t)_{t\in\mathbb{T}^+}$ satisfies the (autonomous) flow equations*

$$\begin{aligned} \Theta_\tau^0 \ &=\ \mathrm{id}_{\Omega\times X} \\ \Theta_\tau^{s+t} \ &=\ \Theta_\tau^t \circ \Theta_\tau^s \quad \textit{for all } s, t \in \mathbb{T}^+. \end{aligned}$$

Proof. We have

$$\Theta_\tau^0(\omega, x) \ =\ (\theta^0\omega\,,\, \varphi(0, \theta^\tau\omega)x) \ =\ (\omega, x)$$

and

$$\begin{aligned} \Theta_\tau^t\Theta_\tau^s(\omega, x) \ &=\ \Theta_\tau^t(\,\theta^s\omega\,,\, \varphi(s, \theta^\tau\omega)x\,) \\ &=\ (\,\theta^t\theta^s\omega\,,\, \varphi(t, \theta^\tau\theta^s\omega)\varphi(s, \theta^\tau\omega)x\,) \\ &=\ (\,\theta^{s+t}\omega\,,\, \varphi(t, \theta^s\theta^\tau\omega)\varphi(s, \theta^\tau\omega)x\,) \\ &=\ (\,\theta^{s+t}\omega\,,\, \varphi(s+t, \theta^\tau\omega)x\,) \\ &=\ \Theta_\tau^{s+t}(\omega, x) \end{aligned}$$

as required. □

For each $\tau \in \mathbb{T}$, we will refer to $(\Theta_\tau^t)_{t\in\mathbb{T}^+}$ as the *skew-product dynamical system* (*associated to* φ) *started at time* τ.

Now just as $(\theta^t)_{t\in\mathbb{T}}$ represents time-shifts on Ω, so we can (trivially) define a "time-shift system" $(\tilde{\theta}^t)_{t\in\mathbb{T}}$ on the product space $\Omega \times X$ by $\tilde{\theta}^t(\omega, x) := (\theta^t\omega, x)$.

Lemma 2.20. *For any $\tau_1, \tau_2 \in \mathbb{T}$, $(\Theta^t_{\tau_1})_{t\in\mathbb{T}^+}$ and $(\Theta^t_{\tau_2})_{t\in\mathbb{T}^+}$ are conjugated by*

$$\Theta^t_{\tau_2} \;=\; \tilde{\theta}^{\tau_1-\tau_2} \circ \Theta^t_{\tau_1} \circ \tilde{\theta}^{\tau_2-\tau_1} \quad \forall\, t \in \mathbb{T}^+.$$

Proof. We have

$$\begin{aligned} \tilde{\theta}^{\tau_1-\tau_2} \circ \Theta^t_{\tau_1} \circ \tilde{\theta}^{\tau_2-\tau_1}(\omega,x) &= \tilde{\theta}^{\tau_1-\tau_2} \circ \Theta^t_{\tau_1}(\,\theta^{\tau_2-\tau_1}\omega\,,\,x\,) \\ &= \tilde{\theta}^{\tau_1-\tau_2}(\,\theta^{t+\tau_2-\tau_1}\omega\,,\,\varphi(t,\theta^{\tau_2}\omega)x\,) \\ &= (\,\theta^t\omega\,,\,\varphi(t,\theta^{\tau_2}\omega)x\,) \\ &= \Theta^t_{\tau_2}(\omega,x) \end{aligned}$$

as required. □

So then, since the dynamical systems $\{(\Theta^t_\tau)_{t\in\mathbb{T}^+} : \tau \in \mathbb{T}\}$ are all conjugate to each other via time-shifts, it will suffice for all purposes *just* to consider the skew-product dynamical system $(\Theta^t_0)_{t\in\mathbb{T}^+}$ started at 0. We will therefore drop the subscript 0 and just write

$$\Theta^t(\omega,x) \;:=\; (\theta^t\omega, \varphi(t,\omega)x).$$

For any $t \in \mathbb{T}^+$ and $A \subset \Omega \times X$, we write $\Theta^{-t}(A) := (\Theta^t)^{-1}(A)$.

Now it is clear that for each $t \in \mathbb{T}^+$, Θ^t is $(\mathcal{F}\otimes\Sigma, \mathcal{F}\otimes\Sigma)$-measurable; in other words, we can regard $(\Theta^t)_{t\in\mathbb{T}^+}$ as a dynamical system on the measurable space $(\Omega\times X, \mathcal{F}\otimes\Sigma)$. But moreover: for any $r,t \in \mathbb{T}^+$, since θ^t is $(\mathcal{F}^\infty_{-r}, \mathcal{F}^\infty_{-(r+t)})$-measurable and $(\omega,x) \mapsto \varphi(t,\omega)x$ is $(\mathcal{F}^\infty_0\otimes\Sigma, \Sigma)$-measurable, it follows that Θ^t is $(\mathcal{F}^\infty_{-r}\otimes\Sigma, \mathcal{F}^\infty_{-(r+t)}\otimes\Sigma)$-measurable; obviously, this implies in particular that Θ^t is $(\mathcal{F}^\infty_{-r}\otimes\Sigma, \mathcal{F}^\infty_{-r}\otimes\Sigma)$-measurable. So in summary: *for any $r \in \bar{\mathbb{T}}^+$, $(\Theta^t)_{t\in\mathbb{T}^+}$ can be regarded as a dynamical system on the measurable space $(\Omega\times X, \mathcal{F}^\infty_{-r}\otimes\Sigma)$.*

The following lemma serves as an important "link" between the Markovian properties of φ and the dynamics of the skew-product system (Θ^t):

Lemma 2.21. *For any probability measure ρ on X,*

(i) $(\Omega\times X,\, \mathcal{F}^\infty_0\otimes\Sigma\,,\, \mathbb{P}|_{\mathcal{F}^\infty_0}\otimes\rho\,,\, (\Theta^t)_{t\in\mathbb{T}^+})$ *is a measure-preserving dynamical system if and only if ρ is stationary under the Markov transition probabilities $(\varphi^t_x)_{x\in X,\,t\in\mathbb{T}^+}$;*

(ii) $(\Omega\times X,\, \mathcal{F}^\infty_0\otimes\Sigma\,,\, \mathbb{P}|_{\mathcal{F}^\infty_0}\otimes\rho\,,\, (\Theta^t)_{t\in\mathbb{T}^+})$ *is an ergodic measure-preserving dynamical system if and only if ρ is ergodic with respect to the Markov transition probabilities $(\varphi^t_x)_{x\in X,\,t\in\mathbb{T}^+}$.*

Proof. Apply [New15a, Theorem 143][12] with $\mathcal{F}_t := \mathcal{F}^t_0$. □

Remark 2.22. In Lemma 2.21, it is important that we restrict the underlying probability space to the σ-algebra $\mathcal{F}^\infty_0$. (In fact, we will see in Section 3.4 that for any probability measure ρ on X, for any $r \in \bar{\mathbb{T}}^+ \setminus \{0\}$, $(\Omega\times X,\, \mathcal{F}^\infty_{-r}\otimes\Sigma\,,\, \mathbb{P}|_{\mathcal{F}^\infty_{-r}}\otimes\rho\,,\, (\Theta^t)_{t\in\mathbb{T}^+})$ is a measure-preserving dynamical system if and only if ρ is crudely incompressible.)

[12] This is, in turn, based on [Kif86, Lemma I.2.3 and Theorem I.2.1].

2.7 Pullback operations and random fixed points

The most basic object in the geometric study of autonomous dynamical systems is the notion of a "fixed point". If we wish to generalise this notion to *random* dynamical systems, one possible way to do this is to regard a *deterministic fixed point* (see Definition 2.18) as being the RDS-analogue of a fixed point of a deterministic dynamical system. However, this is a very restrictive notion. An alternative, much broader notion is that of a *random fixed point*: although the RDS φ is defined as acting on the state space X, there is a natural way to regard φ as acting on the space $L^0(\mathbb{P}; X)$ of X-valued random variables identified up to $\mathbb{P}$-almost sure equality, via the *pullback* construction; a "random fixed point" is a fixed point of this action.

In this section, we will introduce some theory of pullback operations and random fixed points. Several of the results presented here are not actually needed explicitly later on, but help build an intuition for the mathematical "role" of random fixed points.

There are various possible ways of motivating the "pullback" construction and the definition of a random fixed point, of which we now present one:

Motivation

Suppose we have a transitive[13] group action of $\mathbb{T}$ on a set $\mathbf{T}$, denoted by $(t, \mathbf{s}) \mapsto \mathbf{s} + t$ (where $t \in \mathbb{T}$ and $\mathbf{s} \in \mathbf{T}$). Heuristically, $\mathbf{T}$ represents a "timeline without a defined origin".

An element $(x_\mathbf{t})_{\mathbf{t} \in \mathbf{T}}$ of $X^\mathbf{T}$ will be referred to as an *(X-valued) $\mathbf{T}$-path*. A $\mathbf{T}$-path $(x_\mathbf{t})$ will be said to be *stationary* if there exists $c \in X$ such that $x_\mathbf{t} = c$ for all $\mathbf{t} \in \mathbf{T}$.

Let $(f^t)_{t \in \mathbb{T}^+}$ be an autonomous dynamical system on (X, Σ). A *$\mathbf{T}$-solution of* (f^t) is a $\mathbf{T}$-path $(x_\mathbf{t})$ such that for all $\mathbf{s} \in \mathbf{T}$ and $t \in \mathbb{T}^+$, $f^t(x_\mathbf{s}) = x_{\mathbf{s}+t}$. It is clear that a stationary $\mathbf{T}$-path $(p)_{t \in \mathbf{T}}$ is a $\mathbf{T}$-solution of (f^t) if and only if p is a fixed point of (f^t). Let us now consider the random case.

Suppose we have a probability space $(\mathbf{\Omega}, \mathbf{F}, \mathbf{P})$ and a $\mathbf{T}$-indexed family $\boldsymbol{\pi} = (\pi_\mathbf{t})_{\mathbf{t} \in \mathbf{T}}$ of $(\mathbf{F}, \mathcal{F})$-measurable functions $\pi_\mathbf{t} : \mathbf{\Omega} \to \Omega$ such that

(i) $\pi_{\mathbf{s}+t} = \theta^t \circ \pi_\mathbf{s}$ for all $\mathbf{s} \in \mathbf{T}$ and $t \in \mathbb{T}$;

(ii) $\pi_{\mathbf{t}*}\mathbf{P} = \mathbb{P}$ for all $\mathbf{t} \in \mathbf{T}$.

[Heuristically: $(\mathbf{\Omega}, \mathbf{F}, \mathbf{P})$ is a probability space that incorporates the behaviour of some noise process over the timeline $\mathbf{T}$, and (following the heuristic description in Section 2.2) $\pi_\mathbf{t}$ constructs a "plot" of the noise with respect to $\mathbf{t}$ as the reference time.]

An *(X-valued) $(\mathbf{\Omega}, \mathbf{T})$-process* is a $\mathbf{T}$-indexed family $(Y_\mathbf{t})_{\mathbf{t} \in \mathbf{T}}$ of $(\mathbf{F}, \Sigma)$-measurable functions $Y_\mathbf{t} : \mathbf{\Omega} \to X$. We will say that an $(\mathbf{\Omega}, \mathbf{T})$-process $(Y_\mathbf{t})$ is *$\boldsymbol{\pi}$-stationary* if there

[13] A group action of a group G on a set S is said to be *transitive* if the whole of S is a single orbit of the action, i.e. if for all $x, y \in S$ there exists $g \in G$ such that $gx = y$.

exists an $(\mathcal{F},\Sigma)$-measurable function $a\colon\Omega\to X$ such that for each $\mathbf{t}\in\mathbf{T}$, $Y_{\mathbf{t}} \overset{\mathbf{P}\text{-a.s.}}{=} a\circ\pi_{\mathbf{t}}$.

We will say that an $(\mathbf{\Omega},\mathbf{T})$-process $(Y_{\mathbf{t}})$ is an $(\mathbf{\Omega},\mathbf{T})$-*solution of* φ if there is a $\mathbf{P}$-full set $\mathbf{\Omega}'\in\mathbf{F}$ such that for all $\mathbf{s}\in\mathbf{T}$, $t\in\mathbb{T}^+$ and $\boldsymbol{\omega}\in\mathbf{\Omega}'$, $\varphi(t,\pi_{\mathbf{s}}(\boldsymbol{\omega}))Y_{\mathbf{s}}(\boldsymbol{\omega}) = Y_{\mathbf{s}+t}(\boldsymbol{\omega})$.

Lemma 2.23. *Suppose we have an* $(\mathbf{\Omega},\mathbf{T})$*-solution* $(Y_{\mathbf{t}})$ *of* φ*, a time* $\mathbf{s}\in\mathbf{T}$ *and a measurable function* $a\colon\Omega\to X$ *such that*

$$Y_{\mathbf{s}} \overset{\mathbf{P}\text{-}a.s.}{=} a\circ\pi_{\mathbf{s}}\,.$$

Then for any $t\in\mathbb{T}^+$*, we will have that*

$$Y_{\mathbf{s+t}} \overset{\mathbf{P}\text{-}a.s.}{=} (\mathcal{P}_\varphi^t a)\circ\pi_{\mathbf{s+t}}$$

where $\mathcal{P}_\varphi^t a\colon\Omega\to X$ *is given by* $\mathcal{P}_\varphi^t a(\omega) \coloneqq \varphi(t,\theta^{-t}\omega)a(\theta^{-t}\omega)$.

Proof. Fix $t\in\mathbb{T}^+$; let $\mathbf{\Omega}'\in\mathbf{F}$ be a $\mathbf{P}$-full set such that $\varphi(t,\pi_{\mathbf{s}}(\boldsymbol{\omega}))Y_{\mathbf{s}}(\boldsymbol{\omega}) = Y_{\mathbf{s}+t}(\boldsymbol{\omega})$ for all $\boldsymbol{\omega}\in\mathbf{\Omega}'$; and let $\mathbf{\Omega}''\in\mathbf{F}$ be a $\mathbf{P}$-full set such that $Y_{\mathbf{s}}(\boldsymbol{\omega}) = a(\pi_{\mathbf{s}}(\boldsymbol{\omega}))$ for all $\boldsymbol{\omega}\in\mathbf{\Omega}''$. Then for all $\boldsymbol{\omega}\in\mathbf{\Omega}'\cap\mathbf{\Omega}''$,

$$(\mathcal{P}_\varphi^t a)\circ\pi_{\mathbf{s+t}}(\boldsymbol{\omega}) = (\mathcal{P}_\varphi^t a)\circ\theta^t\circ\pi_{\mathbf{s}}(\boldsymbol{\omega}) = \varphi(t,\pi_{\mathbf{s}}(\boldsymbol{\omega}))a(\pi_{\mathbf{s}}(\boldsymbol{\omega})) = \varphi(t,\pi_{\mathbf{s}}(\boldsymbol{\omega}))Y_{\mathbf{s}}(\boldsymbol{\omega}) = Y_{\mathbf{s}+t}(\boldsymbol{\omega})$$

as required. □

Pullback operators

Let $\mathcal{L}^0(\Omega,\mathcal{F};X)$ be the set of all $(\mathcal{F},\Sigma)$-measurable functions $a\colon\Omega\to X$. We say that two functions $a,b\in\mathcal{L}^0(\Omega,\mathcal{F};X)$ are *equivalent* if $a(\omega)=b(\omega)$ for $\mathbb{P}$-almost all $\omega\in\Omega$. With this notion of equivalence, let $L^0(\mathbb{P};X)$ denote the set of equivalence classes of $\mathcal{L}^0(\Omega,\mathcal{F};X)$. For any $a\in\mathcal{L}^0(\Omega,\mathcal{F};X)$, we write $[a]\in L^0(\mathbb{P};X)$ for the equivalence class represented by a.

Now for any $t\in\mathbb{T}^+$, define the "pullback operator" $\mathcal{P}_\varphi^t\colon\mathcal{L}^0(\Omega,\mathcal{F};X)\to\mathcal{L}^0(\Omega,\mathcal{F};X)$ by $\mathcal{P}_\varphi^t a(\omega)\coloneqq\varphi(t,\theta^{-t}\omega)a(\theta^{-t}\omega)$. We show that $(\mathcal{P}_\varphi^t)_{t\in\mathbb{T}^+}$ forms a "dynamical system" on $\mathcal{L}^0(\Omega,\mathcal{F};X)$.

Lemma 2.24. $\mathcal{P}_\varphi^0$ *is the identity function on* $\mathcal{L}^0(\Omega,\mathcal{F};X)$*, and for any* $s,t\in\mathbb{T}^+$ *we have that* $\mathcal{P}_\varphi^{s+t}=\mathcal{P}_\varphi^t\circ\mathcal{P}_\varphi^s$.

Proof. Since $\theta^0=\mathrm{id}_\Omega$ and $\varphi(0,\omega)=\mathrm{id}_X$ for all ω, it follows that $\mathcal{P}_\varphi^0 a = a$ for all $a\in\mathcal{L}^0(\Omega,\mathcal{F};X)$. Now for any $s,t\in\mathbb{T}^+$, any $a\in\mathcal{L}^0(\Omega,\mathcal{F};X)$ and any $\omega\in\Omega$, we have

$$\begin{aligned}
((\mathcal{P}_\varphi^t\circ\mathcal{P}_\varphi^s)(a))(\omega) &= \varphi(t,\theta^{-t}\omega)(\mathcal{P}_\varphi^s a)(\theta^{-t}\omega)\\
&= \varphi(t,\theta^{-t}\omega)\varphi(s,\theta^{-s}\theta^{-t}\omega)a(\theta^{-s}\theta^{-t}\omega)\\
&= \varphi(t,\theta^{-t}\omega)\varphi(s,\theta^{-(s+t)}\omega)a(\theta^{-(s+t)}\omega)\\
&= \varphi(s+t,\theta^{-(s+t)}\omega)a(\theta^{-(s+t)}\omega)\\
&= \mathcal{P}_\varphi^{s+t}a(\omega)
\end{aligned}$$

as required. □

Lemma 2.25. *Suppose we have two functions $a, b \in \mathcal{L}^0(\Omega, \mathcal{F}; X)$ that are equivalent. Then for any $t \in \mathbb{T}^+$, $\mathcal{P}^t_\varphi a$ and $\mathcal{P}^t_\varphi b$ are equivalent.*

Proof. Fix $t \in \mathbb{T}^+$. Let $\Omega' \in \mathcal{F}$ be a $\mathbb{P}$-full set such that $a(\omega) = b(\omega)$ for all $\omega \in \Omega$. Then for all $\omega \in \theta^t(\Omega)$, we have that $a(\theta^{-t}\omega) = b(\theta^{-t}\omega)$ and so $\mathcal{P}^t_\varphi a(\omega) = \mathcal{P}^t_\varphi b(\omega)$. □

In view of Lemma 2.25, for each $t \in \mathbb{T}^+$ we can define the function $P^t_\varphi : L^0(\mathbb{P}; X) \to L^0(\mathbb{P}; X)$ by $P^t_\varphi[a] = [\mathcal{P}^t_\varphi a]$. Obviously, we have that P^0_φ is the identity function on $L^0(\mathbb{P}; X)$ and $P^{s+t}_\varphi = P^t_\varphi \circ P^s_\varphi$ for all $s, t \in \mathbb{T}^+$.

Remark 2.26. Suppose we have a separable metrisable topology on X generating Σ, in which the map $\varphi(t, \omega)$ is continuous for all t and ω. Suppose we have a sequence (a_n) in $\mathcal{L}^0(\Omega, \mathcal{F}; X)$ converging $\mathbb{P}$-almost surely to $a \in \mathcal{L}^0(\Omega, \mathcal{F}; X)$; then for any $t \in \mathbb{T}^+$, $a_n \circ \theta^{-t}$ converges $\mathbb{P}$-almost surely to $a \circ \theta^{-t}$, and therefore (by continuity of $\varphi(t, \omega)$) $\mathcal{P}^t_\varphi a_n$ converges $\mathbb{P}$-almost surely to $\mathcal{P}^t_\varphi a$. It follows[14] that if we have a sequence (a_n) in $\mathcal{L}^0(\Omega, \mathcal{F}; X)$ converging in probability to $a \in \mathcal{L}^0(\Omega, \mathcal{F}; X)$, then for any $t \in \mathbb{T}^+$, $\mathcal{P}^t_\varphi a_n$ converges in probability to $\mathcal{P}^t_\varphi a$. (So for each $t \in \mathbb{T}^+$, P^t_φ is continuous in the topology of convergence in probability.)

We also mention that there is a strong link between the dynamics of $(P^t_\varphi)_{t \in \mathbb{T}^+}$ and the dynamics of the skew-product system $(\Theta^t)_{t \in \mathbb{T}^+}$, as exemplified by Lemma 3.36.

Random fixed points

Definition 2.27. A *random fixed point* or *equilibrium* of φ is a measurable function $q : \Omega \to X$ such that $P^t_\varphi[q] = [q]$ for all $t \in \mathbb{T}^+$, i.e. such that

$$\mathbb{P}(\,\omega : \varphi(t, \theta^{-t}\omega) q(\theta^{-t}\omega) = q(\omega)\,) = 1$$

for all $t \in \mathbb{T}^+$.

(Obviously, in the case that $\mathbb{T} = \mathbb{Z}$, it is sufficient just to check that $P^1_\varphi[q] = [q]$.)

Let $q : \Omega \to X$ be a measurable function. For any $t \in \mathbb{T}^+$ and $s \in \mathbb{T}$, the following statements are clearly equivalent:

- $P^t_\varphi[q] = [q]$;
- $\varphi(t, \theta^s\omega) q(\theta^s\omega) = q(\theta^{s+t}\omega)$ for $\mathbb{P}$-almost all $\omega \in \Omega$;
- $\varphi(t, \omega) q(\omega) = q(\theta^t\omega)$ for $\mathbb{P}$-almost all $\omega \in \Omega$;
- for $\mathbb{P}$-almost all $\omega \in \Omega$, $\Theta^t(\omega, q(\omega)) \in \mathrm{graph}(q)$.

(Intuitively, the last of these says that the graph of q is "almost invariant" under Θ^t. The penultimate of these is often easiest to work with practically.)

[14]Since almost sure convergence implies convergence in probability, and convergence in probability implies the existence of an almost surely convergent subsequence, one obtains the following characterisation of convergence in probability: a sequence of random variables (a_n) converges in probability to a random variable a if and only if every subsequence of (a_n) admits a further subsequence that converges almost surely to a. (Cf. [Din13, Proposition 12.2].)

Proposition 2.28. *Let $q_1, q_2 \colon \Omega \to X$ be random fixed points of φ with $[q_1] \neq [q_2]$. Then $\mathbb{P}(\omega : q_1(\omega) = q_2(\omega)) = 0$.*

Proof. Let $E := \{\omega : q_1(\omega) = q_2(\omega)\}$. Given any $t \in \mathbb{T}^+$, letting $\Omega_t \subset \Omega$ be a $\mathbb{P}$-full set such that $\varphi(t,\omega)q_1(\omega) = q_1(\theta^t\omega)$ and $\varphi(t,\omega)q_2(\omega) = q_2(\theta^t\omega)$ for all $\omega \in \Omega_t$, we have that $\theta^t(E \cap \Omega_t) \subset E$. Hence, by Lemma 2.8, $\mathbb{P}(E) \in \{0,1\}$; and since we assume that $[q_1] \neq [q_2]$, it follows that $\mathbb{P}(E) = 0$. □

Proposition 2.29 (Relation between random fixed points and stationary solutions)**.** *Let $\mathbf{T}$, $\boldsymbol{\Omega}$, $\mathbf{F}$, $\mathbf{P}$ and $(\pi_{\mathbf{t}})$ be as in the "Motivation" part of this section, and let $q \colon \Omega \to X$ be a measurable function. The following are equivalent:*

(i) q is a random fixed point of φ;

(ii) there is an $(\boldsymbol{\Omega}, \mathbf{T})$-solution $(Y_{\mathbf{t}})$ of φ such that for each $\mathbf{t} \in \mathbf{T}$, $Y_{\mathbf{t}} \overset{\mathbf{P}\text{-a.s.}}{=} q \circ \pi_{\mathbf{t}}$.

Proof. (ii)⇒(i): Suppose $(Y_{\mathbf{t}})$ is an $(\boldsymbol{\Omega}, \mathbf{T})$-solution with $Y_{\mathbf{t}} \overset{\mathbf{P}\text{-a.s.}}{=} q \circ \pi_{\mathbf{t}}$ for each $\mathbf{t} \in \mathbf{T}$. Then for any $t \in \mathbb{T}^+$, fixing any $\mathbf{s} \in \mathbf{T}$, we have that for $\mathbf{P}$-almost all $\boldsymbol{\omega} \in \boldsymbol{\Omega}$,

$$\varphi(t, \pi_{\mathbf{s}}(\boldsymbol{\omega}))q(\pi_{\mathbf{s}}(\boldsymbol{\omega})) \;=\; \varphi(t, \pi_{\mathbf{s}}(\boldsymbol{\omega}))Y_{\mathbf{s}}(\boldsymbol{\omega}) \;=\; Y_{\mathbf{s}+t}(\boldsymbol{\omega}) \;=\; q(\pi_{\mathbf{s}+t}(\boldsymbol{\omega})) \;=\; q(\theta^t \pi_{\mathbf{s}}(\boldsymbol{\omega})).$$

Since $\pi_{\mathbf{s}*}\mathbf{P} = \mathbb{P}$, it follows that

$$\mathbb{P}(\omega : \varphi(t,\omega)q(\omega) = q(\theta^t\omega)) \;=\; 1.$$

Since t was arbitrary, q is a random fixed point.

(i)⇒(ii): Suppose q is a random fixed point. Let $\tilde{\Omega} \in \mathcal{F}$ be a $\mathbb{P}$-full set such that for all $m \in \mathbb{Z}$, $n \in \mathbb{N}$ and $\omega \in \tilde{\Omega}$, $\varphi(n, \theta^m\omega)q(\theta^m\omega) = q(\theta^{m+n}\omega)$. Fix an arbitrary $\mathbf{s} \in \mathbf{T}$. For each $t \in \mathbb{T}$, let $Y_{\mathbf{s}+t} \colon \boldsymbol{\Omega} \to X$ be a measurable function with

$$Y_{\mathbf{s}+t}(\boldsymbol{\omega}) \;:=\; \varphi(t - \lfloor t \rfloor, \pi_{\mathbf{s}+\lfloor t \rfloor}(\boldsymbol{\omega}))q(\pi_{\mathbf{s}+\lfloor t \rfloor}\boldsymbol{\omega})$$

for all $\boldsymbol{\omega} \in \pi_{\mathbf{s}}^{-1}(\tilde{\Omega})$. We will show that $(Y_{\mathbf{t}})_{\mathbf{t} \in \mathbf{T}}$ is an $(\boldsymbol{\Omega}, \mathbf{T})$-solution with $Y_{\mathbf{t}} \overset{\mathbf{P}\text{-a.s.}}{=} q \circ \pi_{\mathbf{t}}$ for each $\mathbf{t} \in \mathbf{T}$. For each $t \in \mathbb{T}$, let

$$\Omega_t \;:=\; \{\, \omega \in \Omega : \varphi(t - \lfloor t \rfloor, \theta^{\lfloor t \rfloor}\omega)q(\theta^{\lfloor t \rfloor}\omega) \;=\; q(\theta^t\omega) \,\}.$$

Since q is a random fixed point, Ω_t is a $\mathbb{P}$-full set. For each $t \in \mathbb{T}$, for each $\boldsymbol{\omega} \in \pi_{\mathbf{s}}^{-1}(\tilde{\Omega} \cap \Omega_t)$, writing $\omega := \pi_{\mathbf{s}}(\boldsymbol{\omega})$, we have that

$$Y_{\mathbf{s}+t}(\boldsymbol{\omega}) \;=\; \varphi(t - \lfloor t \rfloor, \theta^{\lfloor t \rfloor}\omega)q(\theta^{\lfloor t \rfloor}\omega) \;=\; q(\theta^t\omega) \;=\; q(\pi_{\mathbf{s}+t}(\boldsymbol{\omega})).$$

So it remains to show that $(Y_{\mathbf{t}})$ is an $(\boldsymbol{\Omega}, \mathbf{T})$-solution of φ. For each $\omega \in \pi_{\mathbf{s}}^{-1}(\tilde{\Omega})$, $s \in \mathbb{T}$ and $t \in \mathbb{T}^+$, writing $\omega := \pi_{\mathbf{s}}(\boldsymbol{\omega})$, we have that

$$\begin{aligned}
\varphi(t, \pi_{\mathbf{s}+s}(\boldsymbol{\omega}))Y_{\mathbf{s}+s}(\boldsymbol{\omega}) &= \varphi(t, \theta^s\omega)\varphi(s - \lfloor s \rfloor, \theta^{\lfloor s \rfloor}\omega)q(\theta^{\lfloor s \rfloor}\omega) \\
&= \varphi(s + t - \lfloor s \rfloor, \theta^{\lfloor s \rfloor}\omega)q(\theta^{\lfloor s \rfloor}\omega) \\
&= \varphi(s + t - \lfloor s + t \rfloor, \theta^{\lfloor s+t \rfloor}\omega)\varphi(\lfloor s + t \rfloor - \lfloor s \rfloor, \theta^{\lfloor s \rfloor}\omega)q(\theta^{\lfloor s \rfloor}\omega) \\
&= \varphi(s + t - \lfloor s + t \rfloor, \theta^{\lfloor s+t \rfloor}\omega)q(\theta^{\lfloor s+t \rfloor}\omega) \\
&= Y_{\mathbf{s}+s+t}(\boldsymbol{\omega}).
\end{aligned}$$

So we are done. □

Now it is clear that for an autonomous dynamical system $(f^t)_{t\in\mathbb{T}^+}$ on a metric space (Y,d) with f^t being continuous for all t, if there exists an initial condition $x_0 \in Y$ such that $f^t(x_0)$ converges as $t \to \infty$, then the limit is itself a fixed point of (f^t). We now consider a couple of ways that this generalises to the random case, first looking at "pullback convergence" and then looking at "forward-time convergence".

Proposition 2.30. *Suppose we have a separable metrisable topology on X generating Σ, such that the map $\varphi(t,\omega)$ is continuous for all t and ω. And suppose we have functions $a, q \in \mathcal{L}^0(\Omega,\mathcal{F};X)$ such that $\mathcal{P}^t_\varphi a$ converges in probability to q as $t \to \infty$. Then q is a random fixed point of φ.*

Proof. As we have established, $(P^t_\varphi)_{t\in\mathbb{T}^+}$ is an autonomous dynamical system on $L^0(\mathbb{P};X)$ with P^t_φ being continuous (with respect to convergence in probability) for all t. Hence $[q] = \lim_{t\to\infty} P^t_\varphi[a]$ is a fixed point of $(P^t_\varphi)_{t\in\mathbb{T}^+}$. □

Proposition 2.31. *Suppose we have a separable metrisable topology on X generating Σ, such that the map $\varphi(t,\omega)$ is continuous for all t and ω.*

(A) If φ is two-way measurable, then $\mathbb{P}$-almost every $\omega \in \Omega$ has the following property: for any $x \in X$, if $\varphi(t,\omega)x$ converges as $t \to \infty$ to a point $p \in X$, then p is a crude deterministic fixed point of φ.

(B) Suppose we have functions $a, q \in \mathcal{L}^0(\Omega,\mathcal{F};X)$ and a probability measure $\mathbb{P}'$ on $(\Omega,\mathcal{F})$ that is absolutely continuous with respect to $\mathbb{P}$, such that the stochastic process $(\varphi(t,\cdot)a(\cdot))_{t\in\mathbb{T}^+}$ defined over $(\Omega,\mathcal{F},\mathbb{P}')$ converges in probability to q. Then for $\mathbb{P}'$-almost every $\omega \in \Omega$, $q(\omega)$ is a crude deterministic fixed point of φ.

Remark 2.32. Proposition 2.31 suggests that in random dynamical systems, one should expect to see convergence of pullback trajectories more often than convergence of forward trajectories, since forward trajectories can only converge where there are (crude) *deterministic* fixed points. (As we will often see, convergence of pullback trajectories can easily occur without the presence of crude deterministic fixed points.)

Proof of Proposition 2.31. (A) Suppose φ is two-way measurable. Let $\mathcal{U}$ be a countable base for the topology of X. For each $U \in \mathcal{U}$ and $n \in \mathbb{N}_0$, let

$$A_{U,n} := \{ (\tau,\omega) \in \mathbb{T}^+ \times \Omega : U \cap \varphi(\tau,\theta^{n\tau}\omega)U \neq \varnothing \}.$$

Note that $A_{U,n}$ is $(\mathcal{B}(\mathbb{T}^+)\otimes\mathcal{F})$-measurable, since, letting S be a countable subset of U that is dense in U, we have that

$$A_{U,n} = \bigcup_{x\in S} \{ (\tau,\omega) \in \mathbb{T}^+ \times \Omega : \varphi(\tau,\theta^{n\tau}\omega)x \in U \}.$$

Now for each $U \in \mathcal{U}$, let

$$T_U := \{\tau \in \mathbb{T}^+ : \mathbb{P}(\omega : (\tau,\omega) \in A_{U,0}) < 1\} = \{\tau \in \mathbb{T}^+ : \mathbb{P}(\omega : (\tau,\omega) \in A_{U,0}) < 1\} \quad \text{(for any } n\text{)}$$

and let

$$N^{(U)} := (T_U \times \Omega) \cap \bigcup_{n=0}^{\infty} \bigcap_{m=n}^{\infty} A_{U,m}.$$

In other words, $N^{(U)}$ is the set of all $(\tau,\omega) \in T_U \times \Omega$ with the property that for all m sufficiently large, $U \cap \varphi(\tau,\theta^{m\tau}\omega)U \neq \varnothing$. It is clear that for any $\tau \in \mathbb{T}^+$, the τ-section of

$N^{(U)}$ (that is, the set $\{\omega \in \Omega : (\tau, \omega) \in N^{(U)}\}$) is a $\mathbb{P}$-null set.

Now let $N := \bigcup_{U \in \mathcal{U}} N^{(U)}$. Let λ be the counting measure on $\mathbb{N}_0$ if $\mathbb{T}^+ = \mathbb{N}_0$, and let λ be the Lebesgue measure on $[0, \infty)$ if $\mathbb{T} = [0, \infty)$. For every $\tau \in \mathbb{T}^+$, the τ-section of N is $\mathbb{P}$-null; so by Fubini's theorem, there is a $\mathbb{P}$-full set $\Omega' \in \mathcal{F}$ such that for every $\omega \in \Omega'$, the ω-section of N is a λ-null set.

Now fix any $\omega \in \Omega'$ and $x \in X$, and suppose that $\varphi(t, \omega)x$ converges as $t \to \infty$ to a point $p \in X$. Fixing a metrisation d of the topology of X, let $(U_r)_{r \in \mathbb{N}}$ be a sequence in $\mathcal{U}$ such that $p \in U_r$ for all r and $\mathrm{diam}(U_r) \to 0$ as $r \to \infty$. For any $\tau \in \mathbb{T}^+$ and $r \in \mathbb{N}$, we clearly have that $(\tau, \omega) \in \bigcup_{n=0}^{\infty} \bigcap_{m=n}^{\infty} A_{U_r, m}$. So, letting N_ω be the ω-section of N, we have that for all $\tau \in \mathbb{T}^+ \setminus N_\omega$ and $r \in \mathbb{N}$, τ does not belong to T_{U_r}. Hence, for each $\tau \in \mathbb{T}^+ \setminus N_\omega$, we have that

$$\mathbb{P}(\tilde{\omega} \in \Omega : \text{for all } r \in \mathbb{N},\, U \cap \varphi(\tau, \tilde{\omega})U \neq \varnothing) \;=\; 1.$$

It clearly follows that

$$\mathbb{P}(\tilde{\omega} \in \Omega : \varphi(\tau, \tilde{\omega})p = p) \;=\; 1.$$

So the Dirac mass δ_p is a fixed point of the map $\rho \mapsto \varphi^{\tau *}\rho$ for λ-almost all $\tau \in \mathbb{T}^+$. But since $\varphi^{s+t*}\rho = \varphi^{t*}\varphi^{s*}\rho$ for all $s, t \in \mathbb{T}^+$ and $\rho \in \mathcal{M}_{(X,\Sigma)}$, it follows that δ_p is a fixed point of the map $\rho \mapsto \varphi^{\tau *}\rho$ for *all* $\tau \in \mathbb{T}^+$. So p is a crude deterministic fixed point of φ.

(B) Suppose for a contradiction that the desired statement is false. Let p be a point in the support of $q_*\mathbb{P}'$ that is not a crude deterministic fixed point of φ, and let $\tau \in \mathbb{T}^+$ be such that $\varphi_p^\tau(\{p\}) < 1$. Fixing a metrisation d of the topology of X, let $\delta > 0$ be such that

$$\mathbb{P}(\omega \in \Omega : B_\delta(p) \cap \varphi(\tau, \omega)B_\delta(p) \neq \varnothing) \;<\; 1$$

and let

$$E \;:=\; \{\omega \in \Omega : B_\delta(p) \cap \varphi(\tau, \omega)B_\delta(p) \neq \varnothing\}.$$

Let $c := q_*\mathbb{P}'(B_{\frac{\delta}{2}}(p)) > 0$. Let $\eta > 0$ be such that for all $A \in \mathcal{F}$ with $\mathbb{P}(A) < \eta$, $\mathbb{P}'(A)$ is less than $\frac{c}{2}$.[15] It is clear that

$$\mathbb{P}\left(\bigcap_{n=0}^{\infty} \theta^{-n\tau}(E)\right) \;=\; 0,$$

so let $m \in \mathbb{N}$ be such that

$$\mathbb{P}\left(\bigcap_{n=0}^{m-1} \theta^{-n\tau}(E)\right) \;<\; \eta.$$

For each $t \in \mathbb{T}^+$, let

$$B_t \;:=\; \{\, \omega \in \Omega : d(\varphi(t, \omega)a(\omega), q(\omega)) \;<\; \tfrac{\delta}{2} \,\}.$$

Let $T \in \mathbb{T}^+$ be such that for all $t \geq T$, $\mathbb{P}'(B_t) \geq 1 - \frac{c}{2(m+1)}$. Let

$$F \;:=\; \bigcap_{n=0}^{m-1} \theta^{-(T+n\tau)}(E) \;=\; \theta^{-T}\left(\bigcap_{n=0}^{m-1} \theta^{-n\tau}(E)\right).$$

[15] The absolute continuity of $\mathbb{P}'$ with respect to $\mathbb{P}$ implies that this is possible; see [Doo94, Section IX.4].

Since θ^T is $\mathbb{P}$-invariant, $\mathbb{P}(F) < \eta$ and so $\mathbb{P}'(F) < \frac{c}{2}$. Note that

$$q^{-1}(B_{\frac{\delta}{2}}(p)) \cap \bigcap_{n=0}^{m} B_{T+n\tau} \subset F.$$

Hence

$$\begin{aligned}
\mathbb{P}'(F) &\geq \mathbb{P}'\left(q^{-1}(B_{\frac{\delta}{2}}(p)) \cap \bigcap_{n=0}^{m} B_{T+n\tau}\right) \\
&\geq -(m+1) + \mathbb{P}'\left(q^{-1}(B_{\frac{\delta}{2}}(p))\right) + \sum_{n=0}^{m} \mathbb{P}'(B_{T+n\tau}) \\
&\geq -(m+1) + c + (m+1-\tfrac{c}{2}) \\
&= \tfrac{c}{2},
\end{aligned}$$

contradicting the fact that $\mathbb{P}'(F) < \frac{c}{2}$. □

Remark 2.33. A further obvious fact about fixed points of dynamical systems is the following: if an autonomous dynamical system (f^t) on a metric space (Y,d) admits a globally attracting fixed point p, then δ_p is the only invariant measure of (f^t) (and so in particular, p is the only fixed point of (f^t)). This also generalises to the random case:[16] A *random invariant measure* of φ is a measurable function $\mu\colon\Omega \to \mathcal{M}_{(X,\Sigma)}$ such that

$$\mathbb{P}(\,\omega : \varphi(t,\omega)_*\mu(\omega) = \mu(\theta^t\omega)\,) = 1$$

for all $t \in \mathbb{T}^+$.[17] Given a separable metric d on X generating Σ, a random fixed point q of φ is said to be *globally weakly attracting* if for every bounded set $B \subset X$, the stochastic process $\omega \mapsto \sup_{x\in B} d(\varphi(t,\omega)x, q(\theta^t\omega))$ converges in probability to 0 as $t \to \infty$. It is easy to show that if q is a globally weakly attracting random fixed point then the only random invariant measure (up to $\mathbb{P}$-almost sure equality) is $\omega \mapsto \delta_{q(\omega)}$.

Examples and basic facts

Example 2.34. Let $X = [0,1]$ (with Σ being the Borel σ-algebra). Let $I = \{0,1\}$ (equipped with the discrete σ-algebra $\mathcal{I}$), let ν be any probability measure on I, and define the functions $f_0, f_1\colon[0,1] \to [0,1]$ by

$$\begin{aligned}
f_0(x) &= \tfrac{1}{2}x \\
f_1(x) &= \tfrac{1}{2}(x+1).
\end{aligned}$$

As in Section 2.3, let φ be the RDS on $[0,1]$ generated by the random map $(I,\mathcal{I},\nu,(f_i)_{i\in I})$. Then it is easy to check that the random variable $q\colon\Omega \to [0,1]$ given by

$$q((i_r)_{r\in\mathbb{Z}}) = 0\cdot i_0 i_{-1} i_{-2} i_{-3} \ldots$$

[16] I am grateful to Martin Rasmussen and Doan Thai Son for showing me that a globally pullback attracting random fixed point must be the only random fixed point. Remark 2.33 is a slight generalisation of this fact.

[17] Invariant measures of random dynamical systems will be discussed in much greater detail in Chapter 3.

(where the right-hand side is to be interpreted in binary) is a random fixed point of φ. If $\nu = \delta_0$ then q is obviously a modification of the constant function $\omega \mapsto 0$; for $\mathbb{P}$-almost every $\omega \in \Omega$, all forward trajectories and all pullback trajectories starting in $[0,1)$ converge to 0, and therefore one can show that the only other random fixed point of φ (up to modification) is the constant function $\omega \mapsto 1$. Similarly, if $\nu = \delta_1$ then q is a modification of the constant function $\omega \mapsto 1$, and the only other random fixed point (up to modification) is the constant function $\omega \mapsto 0$. If $\nu = \lambda\delta_0 + (1-\lambda)\delta_1$ for some $\lambda \in (0,1)$, then q is almost surely in the open interval $(0,1)$; for $\mathbb{P}$-almost every $\omega \in \Omega$, all pullback trajectories starting in $(0,1)$ converge to $q(\omega)$, and therefore one can show that the only other random fixed points of φ (up to modification) are $\omega \mapsto 0$ and $\omega \mapsto 1$.

Example 2.35. Let X, Σ, I and $\mathcal{I}$ be as in Example 2.34, let ν be any probability measure on I, and define the functions $f_0, f_1 \colon [0,1] \to [0,1]$ by

$$f_0(x) = \min(2x,1) = \begin{cases} 2x & x \in [0,\frac{1}{2}] \\ 1 & x \in [\frac{1}{2},1] \end{cases}$$
$$f_1(x) = \max(0,2x-1) = \begin{cases} 0 & x \in [0,\frac{1}{2}] \\ 2x-1 & x \in [\frac{1}{2},1]. \end{cases}$$

As in Section 2.3, let φ be the RDS on $[0,1]$ generated by the random map $(I,\mathcal{I},\nu,(f_i)_{i\in I})$. Then it is easy to check that the random variable $q \colon \Omega \to [0,1]$ given by

$$q((i_r)_{r\in\mathbb{Z}}) = 0 \cdot i_1 i_2 i_3 i_4 \ldots$$

(where the right-hand side is to be interpreted in binary) is a random fixed point of φ. If $\nu = \delta_0$ then q is obviously a modification of the constant function $\omega \mapsto 0$; for $\mathbb{P}$-almost every $\omega \in \Omega$, all forward trajectories and all pullback trajectories starting in $(0,1]$ converge to 1, and therefore one can show that the only other random fixed point of φ (up to modification) is the constant function $\omega \mapsto 1$. Similarly, if $\nu = \delta_1$ then q is a modification of the constant function $\omega \mapsto 1$, and the only other random fixed point (up to modification) is $\omega \mapsto 0$. If $\nu = \lambda\delta_0 + (1-\lambda)\delta_1$ for some $\lambda \in (0,1)$, then q is almost surely in the open interval $(0,1)$; for $\mathbb{P}$-almost every $\omega \in \Omega$, all forward trajectories starting in $[0,1] \setminus \{q(\omega)\}$ converge to either 0 or 1, and therefore one can show that the only other random fixed points of φ (up to modification) are $\omega \mapsto 0$ and $\omega \mapsto 1$.

Remark 2.36. In Example 2.34, q is monotone with respect to the lexicographical order on $I^{-\mathbb{N}_0}$ (and continuous with respect to the product topology on Ω). Similarly, in Example 2.35, q is monotone with respect to the lexicographical order on $I^{\mathbb{N}}$ (and again continuous). Obviously, in both of these examples, if $\nu = \frac{1}{2}(\delta_0 + \delta_1)$ then the law of q is the Lebesgue measure. For a simple example of a situation where (once again) $X = [0,1]$, $I = \{0,1\}$, $\nu = \frac{1}{2}(\delta_0 + \delta_1)$, and φ is a continuous monotone RDS admitting an $\mathcal{F}^0_{-\infty}$-measurable random fixed point q whose law is the Lebesgue measure, *but* where q is severely *non*-monotone with respect to the lexicographical order on $I^{-\mathbb{N}_0}$ (and severely discontinuous), see Example 3.19 [with X extended to $[0,1]$] and Remark 3.20.

Example 2.37 (adapted from [CKS04, Lemma 4.1][18]). Let $(\Omega, \mathcal{F}, (\mathcal{F}_s^{s+t})_{s\in\mathbb{R},\, t\geq 0}, (\theta^t)_{t\in\mathbb{R}}, \mathbb{P})$

[18] I am grateful to Thomas Cass for pointing out to me that the strong solution of the Orstein-Uhlenbeck equation, which is typically expressed using integration against Brownian motion, can be re-expressed (through integration by parts) just using integration against $e^{-\alpha t}$.

be as in Example 2.6, with $d = 1$. Let $X = \mathbb{R}$. We consider the equation

$$dx_t \; = \; \alpha x_t \, dt \, + \, d\omega(t) \tag{2.9}$$

where $\alpha \in \mathbb{R}$. (For $\alpha < 0$, a solution of this equation is called an *Ornstein-Uhlenbeck process.*) Setting $y_t := x_t - \omega(t)$, we have that (2.9) takes the form of a linear differential equation $\dot{y} = \alpha y + \alpha\omega(t)$. The solution is given by

$$y_t \; = \; e^{\alpha t}\left(y_0 + \alpha \int_0^t e^{-\alpha s}\omega(s)\, ds\right).$$

Hence, the RDS φ generated by (2.9) is given by

$$\varphi(t,\omega)x \; = \; \omega(t) + e^{\alpha t}\left(x + \alpha \int_0^t e^{-\alpha s}\omega(s)\, ds\right)$$

for all $\omega \in \Omega$, $t \geq 0$ and $x \in \mathbb{R}$. Using integration by parts ([Apo74, Theorem 7.6]), this can be expressed slightly more succinctly as

$$\varphi(t,\omega)x \; = \; e^{\alpha t}\left(x + \int_0^t e^{-\alpha s}\, d\omega(s)\right)$$

where the integral on the right-hand side is a Riemann-Stieltjes integral. We now consider random fixed points.

• For $\alpha < 0$: Let Ω_- be the set of sample points $\omega \in \Omega$ for which the map $s \mapsto e^{-\alpha s}\omega(s)$ is integrable on $(-\infty, 0]$. Obviously $\theta_t(\Omega_-) = \Omega_-$ for all $t \in \mathbb{R}$. It is also not hard to show that $\mathbb{P}(\Omega_-) = 1$: one can show (using the strong law of large numbers) that for $\mathbb{P}$-almost every $\omega \in \Omega$, $\frac{\omega(-t)}{t} \to 0$ as $t \to \infty$; for any such sample point ω, we have that $|e^{-\alpha s}\omega(s)| \leq e^{-\frac{1}{2}\alpha s}|\frac{\omega(s)}{s}|$ for $s < 0$ with $|s|$ sufficiently large, and therefore $\omega \in \Omega_-$. Now define the function $q\colon \Omega \to \mathbb{R}$ by

$$q(\omega) \; = \; \begin{cases} \alpha \int_{-\infty}^0 e^{-\alpha s}\omega(s)\, ds & \omega \in \Omega_- \\ 0 & \text{otherwise.} \end{cases}$$

Integration by parts yiels that for $\mathbb{P}$-almost every $\omega \in \Omega_-$, $q(\omega)$ can be re-expressed as the improper Riemann-Stieltjes integral $\int_{-\infty}^0 e^{-\alpha s}\, d\omega(s)$. We show that q is a random fixed point as follows: for all $\omega \in \Omega_-$ and $t \geq 0$,

$$\begin{aligned}
\varphi(t,\omega)q(\omega) \; &= \; \omega(t) \, + \, e^{\alpha t}\left(\alpha \int_{-\infty}^0 e^{-\alpha s}\omega(s)\, ds \, + \, \alpha \int_0^t e^{-\alpha s}\omega(s)\, ds\right) \\
&= \; \omega(t) \, + \, \alpha \int_{-\infty}^t e^{-\alpha(s-t)}\omega(s)\, ds \\
&= \; \omega(t) \, + \, \alpha \int_{-\infty}^0 e^{-\alpha s}\omega(s+t)\, ds \\
&= \; \omega(t) \, + \, \alpha \int_{-\infty}^0 e^{-\alpha s}(\theta^t\omega(s) + \omega(t))\, ds \\
&= \; \omega(t) \, + \, \alpha \int_{-\infty}^0 e^{-\alpha s}\theta^t\omega(s)\, ds \, + \, \alpha\omega(t) \int_{-\infty}^0 e^{-\alpha s}\, dt \\
&= \; \omega(t) \, + \, q(\theta^t\omega) \, - \, \omega(t) \\
&= \; q(\theta^t\omega).
\end{aligned}$$

For any ω for which the improper integral $\int_{-\infty}^{0} e^{-\alpha s}\, d\omega(s)$ exists, it is easy to see that every pullback trajectory converges to this integral (which, as a function of ω, is a modification of q); so q is "globally attracting", and therefore one can show that it is the only random fixed point (up to almost-everywhere equality).

• For $\alpha = 0$: Obviously we just have that $\varphi(t,\omega)x = x + \omega(t)$ for all $t \geq 0$, $\omega \in \Omega$ and $x \in \mathbb{R}$. We will show that φ has no random fixed points. Let $q\colon\Omega \to \mathbb{R}$ be any measurable function, and let $A \in \mathcal{B}(\mathbb{R})$ be a bounded set with $q_*\mathbb{P}(A) > 0$. Since A is bounded, one can show (e.g. using the central limit theorem applied to the sequence $(\omega(n) - \omega(n-1))_{n\in\mathbb{N}}$) that $\mathbb{P}(\omega : \varphi(n,\omega)q(\omega) \in A) \to 0$ as $n \to \infty$ in the integers. So, since θ^1 is $\mathbb{P}$-preserving, it cannot be the case that $q(\theta^1\omega) = \varphi(1,\omega)q(\omega)$ for $\mathbb{P}$-almost all ω.

• For $\alpha > 0$: Let Ω_+ be the set of sample points $\omega \in \Omega$ for which the map $s \mapsto e^{-\alpha s}\omega(s)$ is integrable on $[0,\infty)$. Again, $\theta^t(\Omega_+) = \Omega_+$ for all $t \in \mathbb{R}$, and $\mathbb{P}(\Omega_+) = 1$. Define the function

$$q(\omega) \;=\; \begin{cases} -\alpha \int_0^\infty e^{-\alpha s}\omega(s)\, ds & \omega \in \Omega_+ \\ 0 & \text{otherwise.} \end{cases}$$

Integration by parts yiels that for $\mathbb{P}$-almost every $\omega \in \Omega_+$, $q(\omega)$ can be re-expressed as the improper Riemann-Stieltjes integral $-\int_0^\infty e^{-\alpha s}\, d\omega(s)$. We show that q is a random fixed point as follows: for all $\omega \in \Omega_+$ and $t \geq 0$,

$$\begin{aligned}
\varphi(t,\omega)q(\omega) &= \omega(t) + e^{\alpha t}\left(-\alpha \int_0^\infty e^{-\alpha s}\omega(s)\, ds \;+\; \alpha \int_0^t e^{-\alpha s}\omega(s)\, ds\right) \\
&= \omega(t) - \alpha \int_t^\infty e^{-\alpha(s-t)}\omega(s)\, ds \\
&= \omega(t) - \alpha \int_0^\infty e^{-\alpha s}\omega(s+t)\, ds \\
&= \omega(t) - \alpha \int_0^\infty e^{-\alpha s}(\theta^t\omega(s) + \omega(t))\, ds \\
&= \omega(t) - \alpha \int_0^\infty e^{-\alpha s}\theta^t\omega(s)\, ds \;-\; \alpha\omega(t) \int_0^\infty e^{-\alpha s}\, dt \\
&= \omega(t) + q(\theta^t\omega) - \omega(t) \\
&= q(\theta^t\omega).
\end{aligned}$$

By considering the *time-reversal of* φ (see Definition 2.54), one obtains—as in the case that $\alpha < 0$—that q is the only random fixed point (up to almost-everywhere equality).

Remark 2.38. Suppose we have a random fixed point $q\colon\Omega \to X$ that is $(\mathcal{F}_{-\infty}^{r}, \Sigma)$-measurable for some $r \in \mathbb{T}^+$. Then q has a modification $\tilde{q}$ that is $(\mathcal{F}_{-\infty}^{0}, \Sigma)$-measurable: observe that $q \circ \theta^{-r}$ is $(\mathcal{F}_{-\infty}^{0}, \Sigma)$-measurable, and so we can take $\tilde{q}(\omega) := \varphi(r, \theta^{-r}\omega)q(\theta^{-r}\omega)$. (By Remark 3.50, if q is an $\mathcal{F}_{-\infty}^{-r}$-measurable random fixed point for some $r \in \mathbb{T}^+ \setminus \{0\}$, then q is a modification of a constant function $\omega \mapsto p$ where p is a crude deterministic fixed point.)

Lemma 2.39. *Let $q\colon\Omega \to X$ be a random fixed point that is $(\mathcal{F}_{-\infty}^{r}, \Sigma)$-measurable for some $r \in \mathbb{T}$. Then $q_*\mathbb{P}$ is ergodic with respect to the Markov kernel $(\varphi_x^t)_{x\in X}$ for every $t \in \mathbb{T}^+ \setminus \{0\}$ (and is therefore ergodic with respect to the family of Markov transition probabilities $(\varphi_x^t)_{x\in X,\, t\in\mathbb{T}^+}$).*

Proof of Lemma 2.39.[19] Due to Remark 2.38, we can assume without loss of generality that q is $(\mathcal{F}^0_{-\infty}, \Sigma)$-measurable. We first show that $q_*\mathbb{P}$ is stationary under the Markov transition probabilities (φ^t_x). Note that $q \circ \theta^{-t}$ is $\mathcal{F}^{-t}_{-\infty}$-measurable for any t. For each $t \in \mathbb{T}^+$ and $A \in \Sigma$, we have

$$\begin{aligned}
\varphi^{t*}(q_*\mathbb{P})(A) &= \int_\Omega \varphi(t,\theta^{-t}\omega)_* q_*\mathbb{P}(A)\,\mathbb{P}(d\omega) \\
&= \int_\Omega \varphi(t,\theta^{-t}\omega)_* q_*\theta^{-t}_*\mathbb{P}(A)\,\mathbb{P}(d\omega) \\
&= \int_\Omega \mathbb{P}(\tilde\omega : \varphi(t,\theta^{-t}\omega)q(\theta^{-t}\tilde\omega) \in A)\,\mathbb{P}(d\omega) \\
&= \mathbb{P}(\omega : \varphi(t,\theta^{-t}\omega)q(\theta^{-t}\omega) \in A) \\
&\qquad \text{(by Lemma A.10, since } \mathcal{F}^{-t}_{-\infty} \text{ and } \mathcal{F}^0_{-t} \text{ are independent)} \\
&= \mathbb{P}(\omega : q(\omega) \in A) \\
&\qquad \text{(since } q \text{ is a random fixed point)} \\
&= q_*\mathbb{P}(A)
\end{aligned}$$

as required. Now to prove ergodicity: fix $t \in \mathbb{T}^+ \setminus \{0\}$, and let $A \in \Sigma$ be such that for $(q_*\mathbb{P})$-almost every $x \in A$, $\varphi^t_x(A) = 1$; we need to show that $q_*\mathbb{P}(A) \in \{0,1\}$. Let $E := q^{-1}(A)$, and let

$$\tilde E_t \;:=\; \{\omega \in \Omega \,:\, \varphi(t,\theta^{-t}\omega)q(\theta^{-t}\omega) \in A\}.$$

Obviously, since q is a random fixed point, $\mathbb{P}(E \triangle \tilde E_t) = 0$. Note that $\theta^t(E) \in \mathcal{F}^{-t}_{-\infty}$. So then,

$$\begin{aligned}
\mathbb{P}(\,E \cap \theta^t(E)\,) &= \int_{\theta^t(E)} \mathbb{P}(E|\mathcal{F}^{-t}_{-\infty})(\omega)\,\mathbb{P}(d\omega) \\
&= \int_{\theta^t(E)} \mathbb{P}(\tilde E_t|\mathcal{F}^{-t}_{-\infty})(\omega)\,\mathbb{P}(d\omega) \\
&= \int_{\theta^t(E)} \mathbb{P}(\tilde\omega : \varphi(t,\theta^{-t}\tilde\omega)q(\theta^{-t}\omega) \in A)\,\mathbb{P}(d\omega) \\
&\qquad \text{(by Corollary A.11, since } \mathcal{F}^{-t}_{-\infty} \text{ and } \mathcal{F}^0_{-t} \text{ are independent)} \\
&= \int_{\theta^t(E)} \varphi^t_{q(\theta^{-t}\omega)}(A)\,\mathbb{P}(d\omega) \\
&= \int_A \varphi^t_x(A)\; q_*\mathbb{P}(dx) \\
&= \int_A 1\; q_*\mathbb{P}(dx) \\
&= q_*\mathbb{P}(A) \\
&= \mathbb{P}(E).
\end{aligned}$$

Hence $\mathbb{P}(E \setminus \theta^t(E)) = 0$. It follows by Lemma 2.8 that $\mathbb{P}(E) \in \{0,1\}$. So $q_*\mathbb{P}(A) \in \{0,1\}$. □

Remark 2.40. It turns out that for any $\mathcal{F}^r_{-\infty}$-measurable random fixed point $q\colon\Omega \to X$ (with r finite), under any separable metric on X whose Borel σ-algebra is Σ, the dynamics

[19] Lemma 2.39 can in fact be obtained as a corollary of Theorem 3.49. However, we present here a much more direct proof.

of φ will be "contracting on average" within the support of $q_*\mathbb{P}$. The precise sense in which this is the case will be expounded in Chapter 3.

2.8 Monotone RDS

In this thesis, we consider monotone RDS only on *linearly* ordered spaces. Monotone RDS on partially ordered spaces have been studied in e.g. [Chu02] and [FGS15].

We first give some preliminaries on linearly ordered spaces.

Given two linearly (or partially) ordered spaces $(Y, \preceq_Y)$ and $(Z, \preceq_Z)$, we say that a function $f\colon Y \to Z$ is $(\preceq_Y, \preceq_Z)$-*monotone* if for all $x, y \in Y$,

$$x \preceq_Y y \;\;\Rightarrow\;\; f(x) \preceq_Z f(y).$$

Given a linearly (or partially) ordered space $(Y, \preceq_Y)$, a set $A \subset Y$ is said to be *convex* if for every $x, z \in A$ and $y \in Y$ with $x \preceq y \preceq z$, we have that $y \in A$. Note that for a function $f\colon Y \to Z$ between linearly (or partially) ordered spaces $(Y, \preceq_Y)$ and $(Z, \preceq_Z)$, if f is $(\preceq_Y, \preceq_Z)$-monotone then for every convex $A \subset Z$, $f^{-1}(A)$ is convex in Y.

Given a linearly ordered space $(Y, \preceq_Y)$, we will say that a set $A \subset Y$ is *downward-inclusive* if for every $x \in Y$ and $y \in A$ with $x \preceq y$, we have that $x \in A$; and we will say that a set $A \subset Y$ is *upward-inclusive* if for every $x \in A$ and $y \in Y$ with $x \preceq y$, we have that $y \in A$. It is easy to prove the following (very intuitive) statements:

- every downward-inclusive set and every upward-inclusive set is convex;
- for any downward-inclusive $A \subset Y$, the complement $Y \setminus A$ is upward-inclusive; and for any upward-inclusive $A \subset Y$, the complement $Y \setminus A$ is downward-inclusive;
- the intersection of two convex sets is convex;
- given two convex sets $A, B \subset Y$, if $B \not\subset A$ then $A \setminus B$ is convex;
- given a convex set $A \subset Y$ and a set $B \subset Y$ that is either downward-inclusive or upward-inclusive, $A \setminus B$ is convex;
- for any $A \subset Y$, the smallest downward-inclusive set containing A exists and is given by
$$A^- \;:=\; \bigcup_{x \in A} \{y \in X : y \preceq x\}\,;$$
and the smallest upward-inclusive set containing A exists and is given by
$$A^+ \;:=\; \bigcup_{x \in A} \{y \in X : x \preceq y\}\,;$$
- for any $A \subset Y$, the smallest convex set containing A exists and is precisely $A^- \cap A^+$.

We will soon introduce the notion of a "monotone RDS" on a linearly ordered phase space. Since the study of RDS fundamentally relies on the measurable structures involved (namely, the filtration $(\mathcal{F}_s^{s+t})$ on the sample space and the σ-algebra Σ on the phase space), we specifically consider linear orders that "respect" the measurable structure of the phase space. This may be formalised as follows:

Definition 2.41. A *Borel linear order on* (X,Σ) is a linear order $\leq$ on the set X with the property that $\{(x,y) \in X \times X : x \leq y\} \in \Sigma \otimes \Sigma$.

As an obvious example, the usual linear order $\leq$ on $\bar{\mathbb{R}}$ is a Borel linear order (assuming that $\bar{\mathbb{R}}$ is equipped with its usual Borel σ-algebra). A very different example is the lexicographical order on $[0,1] \times [0,1]$ (equipped with the usual Borel σ-algebra), that is

$$(x_1,x_2) \leq_{lex} (y_1,y_2) \quad \overset{\text{def}}{\iff} \quad (x_1 < y_1 \text{ or } (x_1 = y_1 \text{ and } x_2 \leq y_2)).$$

Standing Assumption. *For the rest of Section 2.8, we work with a fixed Borel linear order* $\leq$ *on* (X,Σ).

Note that, given any measurable set $A \in \Sigma$, the restriction of $\leq$ to A is itself a Borel linear order on A (equipped with the induced σ-algebra from Σ). Also note that the map

$$(\rho,x) \mapsto \rho(y \in X : y \leq x) = \int_X \mathbb{1}_{\{(u,v):u\leq v\}}(y,x)\,\rho(dy)$$

from $\mathcal{M} \times X$ to $[0,1]$ is measurable.

Definition 2.42. We say that the RDS φ is *monotone* (*with respect to* $\leq$) if for all $t \in \mathbb{T}^+$ and $\omega \in \Omega$, $\varphi(t,\omega)\colon X \to X$ is $(\leq,\leq)$-monotone.

Remark 2.43. Note that if $\mathbb{T} = \mathbb{R}$ and X is a subset of $\mathbb{R}$, equipped with the usual order $\leq$, then (by the intermediate value theorem) any RDS on X with continuous trajectories must be monotone.

Definition 2.44. Suppose φ is monotone with respect to $\leq$. We say that a point $p \in X$ is ($\leq$-)*subinvariant* if $\mathbb{P}$-almost every $\omega \in \Omega$ has the property that for all $t \in \mathbb{T}^+$, $\varphi(t,\omega)p \leq p$; we say that p is *crudely* ($\leq$-)*subinvariant* if for each $t \in \mathbb{T}^+$, $\mathbb{P}(\omega : \varphi(t,\omega)p \leq p) = 1$. We say that a point $p \in X$ is ($\leq$-)*superinvariant* if $\mathbb{P}$-almost every $\omega \in \Omega$ has the property that for all $t \in \mathbb{T}^+$, $p \leq \varphi(t,\omega)p$; we say that p is *crudely* ($\leq$-)*superinvariant* if for each $t \in \mathbb{T}^+$, $\mathbb{P}(\omega : p \leq \varphi(t,\omega)p) = 1$.

Obviously if $\mathbb{T} = \mathbb{Z}$ then the crudely subinvariant (resp. crudely superinvariant) points are subinvariant (resp. superinvariant).

Remark 2.45. Let $A \subset X$ be a crudely invariant set. If $\max A$ exists then $\max A$ is crudely subinvariant; and if $\min A$ exists then $\min A$ is crudely superinvariant.

Lemma 2.46. *Let* X *be a Borel-measurable subset of* $\bar{\mathbb{R}}$*, equipped with the induced topology from* $\bar{\mathbb{R}}$*, and with* Σ *being the Borel* σ*-algebra (which coincides with the induced* σ*-algebra from* $\mathcal{B}(\bar{\mathbb{R}})$*). Take* $\leq$ *to be the usual order* $\leq$ *on* X*, and suppose that* φ *is monotone. Suppose moreover that the map* $t \mapsto \varphi(t,\omega)x$ *is right-continuous for all* ω *and* x*. Then any crudely subinvariant point is subinvariant, and any crudely superinvariant point is superinvariant.*

Proof. Let D be a countable dense subset of $\mathbb{T}^+$. Let $p \in X$ be a crudely subinvariant point. Fix any ω with the property that for all $t \in D$, $\varphi(t,\omega)p \leq p$. Since the map $t \mapsto \varphi(t,\omega)p$ is right-continuous, we have that $\varphi(t,\omega)p \leq p$ for all $t \in \mathbb{T}^+$. So p is subinvariant. Likewise any crudely superinvariant point is superinvariant. □

As an immediate corollary, we have:

Corollary 2.47. *Assume the hypotheses of Lemma 2.46, and let $A \subset X$ be a crudely invariant set. If* $\max A$ *exists then* $\max A$ *is subinvariant, and if* $\min A$ *exists then* $\min A$ *is superinvariant. Hence in particular, if any of the following statements hold:*

(a) A is downward-inclusive and $\max A$ *exists;*

(b) A is upward-inclusive and $\min A$ *exists;*

(c) A is convex and both $\max A$ *and* $\min A$ *exist;*

then A is invariant.

We now go on to consider stationary probability measures for monotone RDS. We first introduce the "convex core" of a probability measure.[20]

Lemma 2.48. *For any probability measure ρ on X, the set of all convex ρ-full subsets of X has a least element X_ρ (with respect to inclusion), and X_ρ is Σ-measurable.*[21]

Proof. For each $x \in X$, let $I_x^- := \{y \in X : y \leq x\}$ and let $I_x^+ := \{y \in X : x \leq y\}$. Let $J^- := \{x \in X : \rho(I_x^-) = 0\}$, and let $J^+ := \{x \in X : \rho(I_x^+) = 0\}$. Note that J^- and J^+ are Σ-measurable. We will show that

(a) $\rho(J^-) = \rho(J^+) = 0$;

(b) the set $X_\rho := X \setminus (J^- \cup J^+)$ is convex;

(c) every ρ-full convex set contains X_ρ.

The proof will then be complete. (a) For any $x \in J^-$, it is clear that $I_x^- \subset J^-$. With this, and Fubini's theorem, we have

$$\begin{aligned}
\rho(J^-)^2 &= \rho \otimes \rho(J^- \times J^-) \\
&\leq \rho \otimes \rho(\,(x,y) \in J^- \times J^- : x \leq y\,) + \rho \otimes \rho(\,(x,y) \in J^- \times J^- : y \leq x\,) \\
&= 2\rho \otimes \rho(\,(x,y) \in J^- \times J^- : x \leq y\,) \\
&= 2\int_{J^-}\int_{J^-} \chi_{x \leq y}\, \rho(dx)\, \rho(dy) \\
&= 2\int_{J^-} \rho(I_y^-)\, \rho(dy) \\
&= 0.
\end{aligned}$$

[20] The term "convex core" has been used in convex geometry for finite measures on Euclidean space, in [CM01].

[21] It is actually the case that *every* convex set is Σ-measurable; see [MO15b]. However, we will not need this fact, except in the case that (X,Σ) is a measurable subspace of $(\bar{\mathbb{R}}, \mathcal{B}(\bar{\mathbb{R}}))$ with $\preceq := \leq$, in which case the fact is clear.

So $\rho(J^-) = 0$. Similarly, we have that $\rho(J^+) = 0$. (b) Since the intersection of two convex sets is convex, it is sufficient to show that $X \setminus J^-$ and $X \setminus J^+$ are convex. But it is clear that J^- is downward-inclusive and J^+ is upward-inclusive, so the result is immediate. (c) Let $C \subset X$ be a convex set that does not contain X_ρ, and let x be a point in $X_\rho \setminus C$. Since C is convex, we either have that $x \leq y$ for all $y \in C$, or that $y \leq x$ for all $y \in C$. In the former case, I_x^- and C are disjoint, and so since $\rho(I_x^-) > 0$, it follows that C is not a ρ-full set. Likewise in the latter case, C is not a ρ-full set. □

Definition 2.49. For any probability measure ρ on X, we refer to the smallest convex ρ-full subset of X as the *convex core of* ρ (*with respect to* $\leq$), and denote it by X_ρ.

Remark 2.50. Let X be a Borel-measurable subset of $\bar{\mathbb{R}}$ (with Σ being the induced σ-algebra from $\mathcal{B}(\bar{\mathbb{R}})$), and take $\leq$ to be the usual order $\leq$ on X. Let ρ be any probability measure on X. Let $a := \inf \operatorname{supp} \rho$ and $b := \sup \operatorname{supp} \rho$ (where $\operatorname{supp} \rho$ is taken with respect to the usual topology). Then we have

$$X_\rho = \begin{cases} (a,b) \cap X & \text{if } \rho(\{a,b\} \cap X) = 0 \\ [\mathrm{a,b}) \cap X & \text{if } a \in X \text{ and } \rho(\{a\}) > 0, \text{ but } \rho(\{b\} \cap X) = 0 \\ (\mathrm{a,b}] \cap X & \text{if } b \in X \text{ and } \rho(\{b\}) > 0, \text{ but } \rho(\{a\} \cap X) = 0 \\ [\mathrm{a,b}] \cap X & \text{if } a,b \in X \text{ and } \rho(\{a\}), \rho(\{b\}) > 0. \end{cases}$$

Lemma 2.51. *For any probability measure ρ on X, for any ρ-null convex set $A \subset X$, we have that either A and X_ρ are disjoint or $A \subset X_\rho$.*

Proof. Suppose we have a probability measure ρ on X and a ρ-null convex set $A \subset X$, such that A and X_ρ are not disjoint. Since A is ρ-null, $X_\rho \setminus A$ is ρ-full; and therefore, since $X_\rho \setminus A$ is a proper subset of X_ρ, it follows that $X_\rho \setminus A$ is not convex. But since X_ρ and A are themselves convex, it follows that $A \subset X_\rho$. □

Lemma 2.52. *Suppose φ is monotone. Let ρ be a stationary probability measure of the Markov transition probabilities $(\varphi_x^t)_{x \in X,\, t \in \mathbb{T}^+}$. Then X_ρ is crudely invariant.*

Proof. Fix any $t \in \mathbb{T}^+$. We have that

$$1 = \rho(X_\rho) = \int_\Omega \rho(\varphi(t,\omega)^{-1}(X_\rho))\, \mathbb{P}(d\omega)$$

and therefore $\rho(\varphi(t,\omega)^{-1}(X_\rho)) = 1$ for $\mathbb{P}$-almost all $\omega \in \Omega$. But since φ is monotone, $\varphi(t,\omega)^{-1}(X_\rho)$ is convex for all ω, and therefore $X_\rho \subset \varphi(t,\omega)^{-1}(X_\rho)$ for $\mathbb{P}$-almost all $\omega \in \Omega$, as required. □

2.9 Invertible RDS

(Several of the results in this section are adapted from Section 4 of [New15c].)

We will say that the RDS φ is *invertible* if $\varphi(t,\omega)$ is bijective for all $t \in \mathbb{T}^+$ and $\omega \in \Omega$, with the map $(\omega, x) \mapsto \varphi(t,\omega)^{-1}(x)$ being $(\mathcal{F}_0^t \otimes \Sigma, \Sigma)$-measurable for all $t \in \mathbb{T}^+$.

We will now show that if φ is invertible, then one can obtain a random dynamical system $\bar{\varphi}$ simply by "running φ in backward time".

Proposition 2.53. *Suppose φ is invertible. Then the family $\bar{\varphi} = (\bar{\varphi}(t,\omega))_{t\in\mathbb{T}^+,\omega\in\Omega}$ of functions $\bar{\varphi}(t,\omega)\colon X \to X$ given by $\bar{\varphi}(t,\omega)x := \varphi(t,\theta^{-t}\omega)^{-1}(x)$ is a random dynamical system over the noise space $(\Omega,\mathcal{F},(\mathcal{F}^{-s}_{-(s+t)})_{s\in\mathbb{T},\,t\in\mathbb{T}^+},(\theta^{-t})_{t\in\mathbb{T}},\mathbb{P})$.*

Proof. Firstly, it is clear that $(\Omega,\mathcal{F},(\mathcal{F}^{-s}_{-(s+t)})_{s\in\mathbb{T},\,t\in\mathbb{T}^+},(\theta^{-t})_{t\in\mathbb{T}},\mathbb{P})$ is indeed a memoryless stationary noise space (as defined according to our formalism in Section 2.2). Obviously, $\bar{\varphi}(0,\omega) = \mathrm{id}_X$ for all $\omega \in \Omega$. For any $s,t \in \mathbb{T}^+$ and $\omega \in \Omega$, we have

$$\begin{aligned} \bar{\varphi}(s+t,\omega) &= \varphi(s+t,\theta^{-(s+t)}\omega)^{-1} \\ &= \big(\varphi(s,\theta^{-s}\omega) \circ \varphi(t,\theta^{-(s+t)}\omega)\big)^{-1} \\ &= \varphi(t,\theta^{-t}\theta^{-s}\omega)^{-1} \circ \varphi(s,\theta^{-s}\omega)^{-1} \\ &= \bar{\varphi}(t,\theta^{-s}\omega) \circ \bar{\varphi}(s,\omega)^{-1}. \end{aligned}$$

Finally, for any $t \in \mathbb{T}^+$, since the map $(\omega,x) \mapsto \varphi(t,\omega)^{-1}(x)$ is $(\mathcal{F}^t_0 \otimes \Sigma, \Sigma)$-measurable, we have that the map $(\omega,x) \mapsto \varphi(t,\theta^{-t}\omega)^{-1}(x)$ is $(\mathcal{F}^0_{-t} \otimes \Sigma, \Sigma)$-measurable. So we are done. □

Definition 2.54. If φ is invertible, then we refer to $\bar{\varphi}$ in Proposition 2.53 as the *inverse* of φ, or the *time-reversed version* of φ.

Remark 2.55. Note that if φ is invertible, then $\bar{\varphi}$ is invertible as an RDS over $(\Omega,\mathcal{F},(\mathcal{F}^{-s}_{-(s+t)})_{s\in\mathbb{T},\,t\in\mathbb{T}^+},(\theta^{-t})_{t\in\mathbb{T}},\mathbb{P})$, with $\bar{\bar{\varphi}} = \varphi$.

Remark 2.56. If φ is invertible, then we can "extend its domain of definition" by allowing negative times as well as positive times: specifically, set $\varphi(-t,\omega) := \bar{\varphi}(t,\omega)$ for all $t \in \mathbb{T}^+$. One can easily check that the "two-sided cocycle equation"

$$\varphi(s+t,\omega) = \varphi(t,\theta^s\omega) \circ \varphi(s,\omega) \quad \forall\, s,t \in \mathbb{T}$$

is satisfied for all ω. We should warn, however, that a stationary probability measure ρ of the Markov transition probabilities $(\varphi^t_x)_{x\in X,\,t\in\mathbb{T}^+}$ generally does *not* satisfy the equation $\rho = \int_\Omega \varphi(t,\omega)_*\rho(\cdot)\,\mathbb{P}(d\omega)$ for negative $t \in \mathbb{T}$.[22]

Example 2.57. For any Lipschitz 1-periodic $b\colon\mathbb{R} \to \mathbb{R}$ and any $\sigma \in \mathbb{R}$, the RDS on $\mathbb{S}^1$ generated by the SDE $d\phi_t = b(\phi_t)dt + \sigma dW_t$ (as defined in Section 2.3) is invertible. (For the proof, see Proposition 2.87.) The inverse of this RDS can be regarded as "the RDS on $\mathbb{S}^1$ generated by the SDE $d\phi_t = -b(\phi_t)dt + \sigma dW_{-t}$".

Now if φ is invertible, then for any $x \in X$ and $t \in \mathbb{T}^+$ we may define the probability measure $\bar{\varphi}^t_x$ on X by

$$\bar{\varphi}^t_x(A) := \mathbb{P}(\omega \in \Omega : x \in \varphi(t,\omega)A).$$

Note that $(\bar{\varphi}^t_x)_{x\in X,\,t\in\mathbb{T}^+}$ is precisely the family of Markov transition probabilities associated with the time-reversed version $\bar{\varphi}$ of φ, and that a probability measure ρ on X is stationary

[22] One interesting exception is the case that X is finite. Here, one can show that if φ is invertible then every stationary probability measure is crudely incompressible, and thus for each $t \in \mathbb{T}^+$ we have that $\varphi(t,\omega)_*\rho = \varphi(-t,\omega)_*\rho = \rho$ almost surely.

under the Markov transition probabilities $(\bar{\varphi}_x^t)_{x\in X,\,t\in\mathbb{T}^+}$ if and only if for all $t \in \mathbb{T}^+$ and $A \in \Sigma$,

$$\rho(A) \;=\; \int_\Omega \rho(\varphi(t,\omega)A)\,\mathbb{P}(d\omega). \tag{2.10}$$

As in Remark 2.56, the set of stationary probability measures under $(\varphi_x^t)_{x\in X,\,t\in\mathbb{T}^+}$ generally does not coincide with the set of stationary probability measures under $(\bar{\varphi}_x^t)_{x\in X,\,t\in\mathbb{T}^+}$.

Now observe that if φ is invertible, then a finite set $P \subset X$ is crudely invariant under φ if and only if P is crudely invariant under $\bar{\varphi}$.

Lemma 2.58. *Suppose φ is invertible, and let ρ be a probability measure that is ergodic with respect to* either *the Markov transition probabilities $(\varphi_x^t)_{x\in X,\,t\in\mathbb{T}^+}$* or *the Markov transition probabilities $(\bar{\varphi}_x^t)_{x\in X,\,t\in\mathbb{T}^+}$. Then either ρ is atomless, or $\rho = \frac{1}{|P|}\sum_{x\in P}\delta_x$ for some finite crudely invariant set $P \subset X$.*

Proof. Suppose ρ is not atomless. Let $m := \max_{x\in X}\rho(\{x\})$, and let $P := \{x : \rho(\{x\}) = m\}$. For any $t \in \mathbb{T}^+$ and $\omega \in \Omega$, since $\varphi(t,\omega)$ is bijective, we have that either

(i) $\varphi(t,\omega)P = \varphi(t,\omega)^{-1}(P) = P$; or

(ii) $\varphi(t,\omega)P \neq P$, and $\rho(\varphi(t,\omega)P)$ and $\rho(\varphi(t,\omega)^{-1}(P))$ are both less than P.

So then, since ρ is stationary under either (φ_x^t) or $(\bar{\varphi}_x^t)$, we have that for each $t \in \mathbb{T}^+$, scenario (i) occurs for $\mathbb{P}$-almost all $\omega \in \Omega$ (i.e. P is crudely invariant). Hence in particular, $\varphi_x^t(P) = \bar{\varphi}_x^t(P) = 1$ for all $x \in P$ and $t \in \mathbb{T}^+$. Since ρ is ergodic under either (φ_x^t) or $(\bar{\varphi}_x^t)$, it follows that $\rho(P) = 1$. □

Lemma 2.59. *Suppose φ is invertible and $q\colon\Omega \to \mathcal{F}$ is a random fixed point of φ. Then q is also a random fixed point of $\bar{\varphi}$ (regarded as a RDS over the noise space $(\Omega, \mathcal{F}, (\mathcal{F}_{-(s+t)}^{-s})_{s\in\mathbb{T},\,t\in\mathbb{T}^+}, (\theta^{-t})_{t\in\mathbb{T}}, \mathbb{P})$).*

Proof. Fix $t \in \mathbb{T}^+$. For $\mathbb{P}$-almost all $\omega \in \Omega$, we have that

$$\varphi(t,\theta^{-t}\omega)q(\theta^{-t}\omega) \;=\; q(\omega)$$

and therefore

$$q(\theta^{-t}\omega) \;=\; \bar{\varphi}(t,\omega)q(\omega)$$

as required. □

So then, applying Lemma 2.39 to $\bar{\varphi}$, we obtain the following:

Corollary 2.60. *Suppose φ is invertible, and let $q\colon\Omega \to X$ be a random fixed point of φ that is $(\mathcal{F}_r^\infty, \Sigma)$-measurable for some $r \in \mathbb{T}$. Then $q_*\mathbb{P}$ is ergodic with respect to the Markov kernel $(\bar{\varphi}_x^t)_{x\in X}$ for each $t \in \mathbb{T}^+ \setminus \{0\}$ (and is therefore ergodic with respect to the family of Markov transition probabilites $(\bar{\varphi}_x^t)_{x\in X,\,t\in\mathbb{T}^+}$).*

We finish this section on invertible RDS with the following:

Lemma 2.61. *Suppose φ is invertible, and let $q\colon\Omega \to X$ be a random fixed point of φ that is either $(\mathcal{F}_{-\infty}^r, \Sigma)$-measurable or $(\mathcal{F}_r^\infty, \Sigma)$-measurable for some $r \in \mathbb{T}$. Then $q_*\mathbb{P}$ is either atomless or a Dirac mass on a crude deterministic fixed point.*

Proof. It suffices just to consider the case that q is $\mathcal{F}_{-\infty}^{r}$-measurable, since we can then just take the inverse of φ to give the case that q is $\mathcal{F}_{r}^{\infty}$-measurable for some $r \in \mathbb{T}$. Due to Remark 2.38, we can assume without loss of generality that q is $\mathcal{F}_{-\infty}^{0}$-measurable.

Suppose that $q_*\mathbb{P}$ is not atomless. Then by Lemmas 2.39 and 2.58, there is a finite crudely invariant set $P \subset \mathbb{S}^1$ such that $q_*\mathbb{P} = \frac{1}{|P|}\sum_{x\in P}\delta_x$. Fix an arbitrary $x \in P$. Let $E := q^{-1}(\{x\})$, and for each $n \in \mathbb{N}$ let

$$\tilde{E}_n \;:=\; \{\omega \in \Omega : \varphi(n,\theta^{-n}\omega)q(\theta^{-n}\omega) = x\}.$$

Obviously, since q is a random fixed point, $\mathbb{P}(E \vartriangle \tilde{E}_n) = 0$ for each n. For each $n \in \mathbb{N}$, for any $F \in \mathcal{F}_{-n}^{\infty}$, we have that

$$\begin{aligned}
\mathbb{P}(E \cap F) &= \int_F \mathbb{P}(E|\mathcal{F}_{-n}^{\infty})(\omega)\,\mathbb{P}(d\omega) \\
&= \int_F \mathbb{P}(\tilde{E}_n|\mathcal{F}_{-n}^{\infty})(\omega)\,\mathbb{P}(d\omega) \\
&= \int_F \mathbb{P}(\tilde{\omega} : \varphi(n,\theta^{-n}\omega)q(\theta^{-n}\tilde{\omega}) = x)\,\mathbb{P}(d\omega) \\
&\qquad \text{(by Corollary A.11, since } \mathcal{F}_{-\infty}^{-n} \text{ and } \mathcal{F}_{-n}^{\infty} \text{ are independent)} \\
&= \int_F \mathbb{P}(\tilde{\omega} : \varphi(n,\theta^{-n}\omega)q(\tilde{\omega}) = x)\,\mathbb{P}(d\omega) \\
&= \int_F q_*\mathbb{P}(\{\bar{\varphi}(n,\omega)x\})\,\mathbb{P}(d\omega) \\
&= \int_F \tfrac{1}{|P|}\,\mathbb{P}(d\omega) \\
&\qquad \text{(since } \bar{\varphi}(n,\omega)x \in P \text{ for } \mathbb{P}\text{-almost all } \omega) \\
&= \tfrac{1}{|P|}\mathbb{P}(F) \\
&= \mathbb{P}(E)\mathbb{P}(F).
\end{aligned}$$

So E is independent of $\mathcal{F}_{-n}^{\infty}$ for each $n \in \mathbb{N}$, and therefore E is independent of $\mathcal{F}$. In particular, E is independent of itself, and so $\mathbb{P}(E) = 1$. Hence $|P| = 1$, i.e. $q_*\mathbb{P}$ is a Dirac mass at a crude deterministic fixed point. □

2.10 Two "derived" RDS

The n-point motion

Since we assume that (X, Σ) is standard, the measurable space $(X^n, \Sigma^{\otimes n})$ is also standard for any $n \in \mathbb{N}$ (as it is the Borel space associated to the n-fold product of any compact metrisable topology on X generating Σ). We define the family $\varphi^{\times n} = (\varphi^{\times n}(t,\omega))_{t\in\mathbb{T}^+,\omega\in\Omega}$ of functions $\varphi^{\times n}(t,\omega) \colon X^n \to X^n$ by

$$\varphi^{\times n}(t,\omega)(x_1,\ldots,x_n) \;=\; (\,\varphi(t,\omega)x_1\,,\,\ldots\,,\,\varphi(t,\omega)x_n\,).$$

It is clear that $\varphi^{\times n}$ is a RDS on X^n (over the same noise space over which the RDS φ is defined). We refer to $\varphi^{\times n}$ as the *n-point motion* of φ. We denote the associated Markov transition probabilities by $(\varphi^t_{\mathbf{x}})_{\mathbf{x}\in X^n,\, t\in\mathbb{T}^+}$. Given any probability measure $\boldsymbol{\rho}$ on X^n, we define the probability measure $\varphi^{t*}_{(n)}\boldsymbol{\rho}$ on X^n by

$$\varphi^{t*}_{(n)}\boldsymbol{\rho}(A) \;:=\; \int_{X^n} \varphi^t_{\mathbf{x}}(A)\, \boldsymbol{\rho}(d\mathbf{x})$$

for all $A \in \Sigma^{\otimes n}$. Note, as before, that this can be re-expressed as

$$\varphi^{t*}_{(n)}\boldsymbol{\rho}(A) \;=\; \int_{\Omega} \varphi^{\times n}(t,\omega)_*\boldsymbol{\rho}(A)\, \mathbb{P}(d\omega)$$

for all $A \in \Sigma^{\otimes n}$.

Proposition 2.62 (cf. [Kun90, Theorem 4.3.2])**.** *For any integer $n \geq 2$, for any probability measure ρ on X, $\rho^{\otimes n}$ is stationary under the Markov transition probabilities $(\varphi^t_{\mathbf{x}})_{\mathbf{x}\in X^n,\, t\in\mathbb{T}^+}$ if and only if ρ is crudely incompressible under φ.*

Proof. If ρ is crudely incompressible then for any $t \in \mathbb{T}^+$ and $E_1, \ldots, E_n \in \Sigma$,

$$\begin{aligned}
\varphi^{t*}_{(n)}(\rho^{\otimes n})(E_1 \times \ldots \times E_n) &= \int_{\Omega} \varphi^{\times n}(t,\omega)_*(\rho^{\otimes n})(E_1 \times \ldots \times E_n)\, \mathbb{P}(d\omega) \\
&= \int_{\Omega} \prod_{i=1}^{n} \varphi(t,\omega)_*\rho(E_i)\, \mathbb{P}(d\omega) \\
&= \int_{\Omega} \prod_{i=1}^{n} \rho(E_i)\, \mathbb{P}(d\omega) \\
&= \rho^{\otimes n}(E_1 \times \ldots \times E_n).
\end{aligned}$$

So $\rho^{\otimes n}$ is stationary under $(\varphi^t_{\mathbf{x}})_{\mathbf{x}\in X^n,\, t\in\mathbb{T}^+}$. Now, conversely, suppose that $\rho^{\otimes n}$ is stationary under $(\varphi^t_{\mathbf{x}})_{\mathbf{x}\in X^n,\, t\in\mathbb{T}^+}$. Fix $t \in \mathbb{T}^+$. For any $A \in \Sigma$, we have that

$$\begin{aligned}
&\mathbb{E}_{(\mathbb{P})}[\omega \mapsto \varphi(t,\omega)_*\rho(A)] \\
&\quad= \mathbb{E}_{(\mathbb{P})}[\omega \mapsto \varphi^{\times n}(t,\omega)_*(\rho^{\otimes n})(A \times X^{n-1})] \\
&\quad= \rho^{\otimes n}(A \times X^{n-1}) \\
&\quad= \rho(A)
\end{aligned}$$

and (writing $\mathrm{Var}_{(\mathbb{P})}[\cdot]$ for the variance of a random variable)

$$\begin{aligned}
&\mathrm{Var}_{(\mathbb{P})}[\omega \mapsto \varphi(t,\omega)_*\rho(A)] \\
&\quad= \mathbb{E}_{(\mathbb{P})}[\omega \mapsto (\varphi(t,\omega)_*\rho(A))^2] - \mathbb{E}_{(\mathbb{P})}[\omega \mapsto \varphi(t,\omega)_*\rho(A)]^2 \\
&\quad= \mathbb{E}_{(\mathbb{P})}[\omega \mapsto \varphi^{\times n}(t,\omega)_*(\rho^{\otimes n})(A^2 \times X^{n-2})] - \rho(A)^2 \\
&\quad= \rho^{\otimes n}(A^2 \times X^{n-2}) - \rho(A)^2 \\
&\quad= 0.
\end{aligned}$$

Hence $\varphi(t,\omega)_*\rho(A) = \rho(A)$ for $\mathbb{P}$-almost all $\omega \in \Omega$. Since (X, Σ) is standard, there exists a countable π-system $\mathcal{C}$ generating Σ (by Remark A.1). $\mathbb{P}$-almost every $\omega \in \Omega$ has the property that for all $A \in \mathcal{C}$, $\varphi(t,\omega)_*\rho(A) = \rho(A)$. Therefore, by Corollary A.6, $\varphi(t,\omega)_*\rho = \rho$ for $\mathbb{P}$-almost all $\omega \in \Omega$. Hence ρ is crudely incompressible. □

The image-measure RDS

Since we assume that (X,Σ) is standard, the measurable space $(\mathcal{M}_{(X,\Sigma)},\mathfrak{K}_{(X,\Sigma)})$ is also standard (as it is the Borel space associated to the narrow topology corresponding to any Polish topology on X generating Σ). Now writing $\varphi(t,\omega)_*\colon\mathcal{M}_{(X,\Sigma)}\to\mathcal{M}_{(X,\Sigma)}$ for the function sending a probability measure ρ on X to the image measure $\varphi(t,\omega)_*\rho$, let $\varphi_* := (\varphi(t,\omega)_*)_{t\in\mathbb{T}^+,\,\omega\in\Omega}$.

It is not hard to show that φ_* is a RDS on $\mathcal{M}_{(X,\Sigma)}$ (over the same noise space over which the RDS φ is defined). The "non-trivial" part is to show that the map $(\omega,\rho)\mapsto\varphi(t,\omega)_*\rho$ is $(\mathcal{F}_0^t\otimes\mathfrak{K}_{(X,\Sigma)},\mathfrak{K}_{(X,\Sigma)})$-measurable, which we justify as follows: for any $A\in\Sigma$, we have

$$\varphi(t,\omega)_*\rho(A) \;=\; \int_X \mathbb{1}_A(\varphi(t,\omega)x)\,\rho(dx),$$

so by Lemma A.12, the map $(\omega,\rho)\mapsto\varphi(t,\omega)_*\rho(A)$ is $(\mathcal{F}_0^t\otimes\mathfrak{K}_{(X,\Sigma)},\mathcal{B}([0,1]))$-measurable, as required.

We refer to φ_* as the *image-measure RDS associated to* φ. We denote the associated Markov transition probabilities by $(\varphi_\rho^t)_{\rho\in\mathcal{M}_{(X,\Sigma)},\,t\in\mathbb{T}^+}$. Given any probability measure Q on $\mathcal{M}_{(X,\Sigma)}$, we define the probability measure $\varphi_*^{t*}Q$ on $\mathcal{M}_{(X,\Sigma)}$ by

$$\varphi_*^{t*}Q(A) \;:=\; \int_{\mathcal{M}_{(X,\Sigma)}} \varphi_\rho^t(A)\,Q(d\rho)$$

for all $A\in\mathfrak{K}_{(X,\Sigma)}$. Once again, this can be re-expressed as

$$\varphi_*^{t*}Q(A) \;=\; \int_\Omega \varphi(t,\omega)_{**}Q(A)\,\mathbb{P}(d\omega)$$

for all $A\in\mathfrak{K}_{(X,\Sigma)}$. (Here, $\varphi(t,\omega)_{**}Q(A) = Q(\{\rho\in\mathcal{M}_{(X,\Sigma)} : \varphi(t,\omega)_*\rho\in A\})$.)

Remark 2.63. One useful intuitive way of visualising a trajectory $(\varphi(t,\omega)_*\rho)_{t\geq0}$ of the image-measure RDS φ_* is as follows: Imagine we endow the phase space X with some *initial distribution of mass* ρ; we then run the RDS according to some noise realisation ω, and see how the distribution of mass evolves over time. At time t, the distribution of mass is given by $\varphi(t,\omega)_*\rho$.

2.11 RDS in a topological setting

Fix a separable metrisable topology on X generating Σ.

We will define certain continuity properties for a RDS, and from then on, we will study properties of "right-continuous" RDS.

(We work with the convention that for *any* subset E of $\mathbb{R}$ and any function $f\colon E\to X$, we say that f is right-continuous at $t\in E$ if the restriction of f to $E\cap[t,\infty)$ is continuous at t, and we likewise say that f is left-continuous at t if the restriction of f to $E\cap(-\infty,t]$ is continuous at t.)

Definition 2.64. We will say that φ is *spatially continuous* if the map $\varphi(t,\omega)\colon X \to X$ is a continuous function for all $t \in \mathbb{T}^+$ and $\omega \in \Omega$.

Definition 2.65. We will say that φ is *continuous* if the map $(t,x) \mapsto \varphi(t,\omega)x$ is continuous for all $\omega \in \Omega$.

Definition 2.66. We will say that φ is *right-continuous* if for any decreasing[23] sequence (t_n) in $\mathbb{T}^+$ converging to a value t and any sequence (x_n) in X converging to a point x, $\varphi(t_n,\omega)x_n \to \varphi(t,\omega)x$ as $n \to \infty$ for all $\omega \in \Omega$.

Obviously, right-continuity of φ implies in particular that (a) φ is spatially continuous; and (b) for each x and ω the trajectory $t \mapsto \varphi(t,\omega)x$ is right-continuous. Note also that if φ is right-continuous then φ is measurable.[24]

Remark 2.67. By Lemma A.20, if φ is right-continuous then for any decreasing sequence (t_n) in $\mathbb{T}^+$ converging to a value t and any sequence (x_n) in X converging to a point x, $\varphi_{x_n}^{t_n}$ converges in the narrow topology to φ_x^t. (In particular, the Markov transition probabilities $(\varphi_x^t)_{x \in X,\, t \in \mathbb{T}^+}$ are Feller-continuous.)

Definition 2.68. We will say that φ is *càdlàg* if φ is right-continuous and for each $t \in \mathbb{T}^+ \setminus \{0\}$ and $\omega \in \Omega$ there exists a continuous function $\varphi_-(t,\omega)\colon X \to X$ such that for any strictly increasing sequence (t_n) in $\mathbb{T}^+$ converging to t and any sequence (x_n) in X converging to a point x, $\varphi(t_n,\omega)x_n \to \varphi_-(t,\omega)x$ as $n \to \infty$.

Definition 2.69. We will say that φ *has left-continuous pullback trajectories* if for every $x \in X$ and $\omega \in \Omega$ the map $t \mapsto \varphi(t,\theta^{-t}\omega)x$ is left-continuous.

Note that if $\mathbb{T} = \mathbb{Z}$ then continuity, spatial continuity, right-continuity and càdlàg are all equivalent, and φ necessarily has left-continuous pullback trajectories.

Definition 2.70. We will say that φ is an *open-mapping RDS* if φ is right-continuous and for every $t \in \mathbb{T}^+$, $\omega \in \Omega$ and open $U \subset X$, $\varphi(t,\omega)U$ is open.

Definition 2.71. Suppose φ is right-continuous. Moreover, let Y be a separable metrisable topological space (with $\mathcal{B}(Y)$ standard), and let φ' be a right-continuous RDS on Y over $(\Omega,\mathcal{F},(\mathcal{F}_s^{s+t}),(\theta^t),\mathbb{P})$. A function $h\colon X \to Y$ is called a *deterministic semiconjugacy from φ to φ'* if h is continuous and surjective and for all $t \in \mathbb{T}^+$, $\omega \in \Omega$ and $x \in X$, $h(\varphi(t,\omega)x) = \varphi'(t,\omega)h(x)$.

Standing Assumption. *For the rest of Section 2.11, we assume that φ is right-continuous.*

Lemma 2.72. *For any $x \in X$ and open $U \subset X$, the set*

$$E_{x,U} \;:=\; \{\, \omega \in \Omega \,:\, \exists\, t \in \mathbb{T}^+ \ \text{s.t.}\ \varphi(t,\omega)x \in U \,\}$$

is $\mathcal{F}_0^\infty$-measurable. Given any dense $D \subset \mathbb{T}^+$, $\mathbb{P}(E_{x,U}) > 0$ if and only if there exists $t \in D$ such that $\varphi_x^t(U) > 0$.

[23] Here, a "decreasing sequence" need *not* be strictly decreasing.
[24] See e.g. [New15a, Lemma 16(B)].

If $\mathbb{P}(E_{x,U}) > 0$, then we say that *U is accessible from x* (under φ). Note that for any open $U \subset X$, the set of points in X from which U is accessible is itself open.

Proof of Lemma 2.72. Let $\tilde{D}$ be any countable dense subset of $\mathbb{T}^+$. For each $\omega \in \Omega$, since the map $t \mapsto \varphi(t,\omega)x$ is right-continuous, it is clear that

$$\exists\, t \in \mathbb{T}^+ \text{ s.t. } \varphi(t,\omega)x \in U \iff \exists\, t \in D \text{ s.t. } \varphi(t,\omega)x \in U.$$

In other words,

$$E_{x,U} = \bigcup_{t \in \tilde{D}} \{\, \omega \in \Omega : \varphi(t,\omega)x \in U \,\}, \tag{2.11}$$

and so $E_{x,U} \in \mathcal{F}_0^\infty$. Now given any dense $D \subset \mathbb{T}^+$, we can take the countable dense set $\tilde{D}$ to be a subset of D. Obviously if there exists $t \in D$ such that $\varphi_x^t(U) > 0$ then $\mathbb{P}(E_{x,U}) > 0$; and conversely, if $\mathbb{P}(E_{x,U}) > 0$ then by equation (2.11) there exists $t \in \tilde{D} \subset D$ such that $\varphi_x^t(U) > 0$. □

Remark 2.73. We have defined "accessibility" in terms of there being a *positive probability* of reaching a given set from a given point. Nonetheless, it is useful to note the following: Suppose that (as is usually the case in practice for continuous RDS) there exists a separable metrisable topology on Ω whose Borel σ-algebra coincides with $\mathcal{F}$, such that $\mathbb{P}$ has full support and the map $\omega \mapsto \varphi(t,\omega)x$ is continuous for each t and x. Then, to show that an open set U is accessible from a point x, it is sufficient just to find one sample point $\omega \in \Omega$ and a time $t \in \mathbb{T}^+$ such that $\varphi(t,\omega)x \in U$. Note in particular that, to show that a given point $p \in X$ is *not* a deterministic fixed point (i.e. that $X \setminus \{p\}$ is accessible from p), it is sufficient to find one sample point ω and a time t such that $\varphi(t,\omega)p \neq p$.

Lemma 2.74. *For any closed $G \subset X$ and $t \in \mathbb{T}^+$, the set $\{\, \omega \in \Omega : \varphi(t,\omega)G \subset G \,\}$ is $\mathcal{F}_0^t$-measurable. This set is a $\mathbb{P}$-full set if and only if $\varphi_x^t(G) = 1$ for all $x \in G$.*

Proof. Let $S \subset G$ be a countable set that is dense in G. For each $\omega \in \Omega$, since $\varphi(t,\omega)$ is continuous, we have that $\varphi(t,\omega)G \subset G$ if and only if $\varphi(t,\omega)S \subset G$; in other words

$$\{\, \omega \in \Omega : \varphi(t,\omega)G \subset G \,\} = \bigcap_{x \in S} \{\, \omega \in \Omega : \varphi(t,\omega)x \in G \,\} \in \mathcal{F}_0^t \tag{2.12}$$

as required. As in the proof of Lemma 2.72, the rest is clear by equation (2.12). □

Recall that a set $A \subset X$ is said to be *invariant* (under φ) if $\mathbb{P}$-almost every $\omega \in \Omega$ has the property that for all $t \in \mathbb{T}^+$, $\varphi(t,\omega)A \subset A$; and a set $A \in \Sigma$ is said to be *very crudely invariant* (under φ) if for each $x \in A$ and $t \in \mathbb{T}^+$, $\varphi_x^t(A) = 1$. (In other words, A is said to be very crudely invariant under φ if A is forward-invariant according to the Markov transition probabilities $(\varphi_x^t)_{x \in X,\, t \in \mathbb{T}^+}$.)

Lemma 2.75. *For any closed $G \subset X$, the set $\{\, \omega \in \Omega : \varphi(t,\omega)G \subset G \ \forall t \in \mathbb{T}^+ \,\}$ is $\mathcal{F}_0^\infty$-measurable. This set is a $\mathbb{P}$-full set if and only if $\varphi_x^t(G) = 1$ for all $x \in G$ and $t \in \mathbb{T}^+$. (In other words: a closed set is invariant if and only if it is very crudely invariant).*

Obviously, as a special case of this, any *crude* deterministic fixed point is in fact a deterministic fixed point.

Proof of Lemma 2.75. For each $\omega \in \Omega$ and $x \in G$, since the map $t \mapsto \varphi(t,\omega)x$ is right-continuous, we have that

$$\varphi(t,\omega)x \in G \ \forall\, t \in \mathbb{T}^+ \quad \Longleftrightarrow \quad \varphi(t,\omega)x \in G \ \forall\, t \in D.$$

In other words,

$$\{\, \omega \in \Omega \,:\, \varphi(t,\omega)G \subset G \ \forall t \in \mathbb{T}^+ \,\} \;=\; \bigcap_{t \in D} \{\, \omega \in \Omega \,:\, \varphi(t,\omega)G \subset G \,\}, \tag{2.13}$$

which is $\mathcal{F}_0^\infty$-measurable due to Lemma 2.74. The rest is clear by Lemma 2.74 and equation (2.13). □

Note in particular that for any probability measure ρ that is stationary under the Markov transition probabilities $(\varphi_x^t)_{x \in X,\, t \in \mathbb{T}^+}$, $\operatorname{supp}\rho$ is invariant. As in Section C.4, for any non-empty compact invariant $G \subset X$ there exists at least one ergodic probability measure ρ with $\rho(G) = 1$.

We will now see that for an ergodic probability measure ρ, almost all trajectories starting in $\operatorname{supp}\rho$ are almost surely dense in $\operatorname{supp}\rho$.

Lemma 2.76. *Let ρ be a probability measure that is ergodic with respect to the Markov transition probabilities $(\varphi_x^t)_{x \in X,\, t \in \mathbb{T}^+}$. Then $(\mathbb{P} \otimes \rho)$-almost every $(\omega, x) \in \Omega \times X$ has the property that for all $T \in \mathbb{T}^+$, $\overline{\{\varphi(t,\omega)x : t \geq T\}} = \operatorname{supp}\rho$.*

Lemma 2.76 can be obtained as an immediately corollary of Lemma 2.15 (using the fact that there is a countable base for the topology of X). However, we give the following more elementary proof:

Proof. For any open $U \subset X$, let

$$\begin{aligned} A_U \;&:=\; \{\, (\omega,x) \in \Omega \times X \,:\, \exists\, t \in \mathbb{T}^+ \text{ s.t. } \varphi(t,\omega)x \in U \,\} \\ &=\; \{\, (\omega,x) \in \Omega \times X \,:\, \exists\, t \in \mathbb{T}^+ \text{ s.t. } \Theta^t(\omega,x) \in \Omega \times U \,\}. \end{aligned}$$

It is clear that for any $(\omega,x) \in \Omega \times X$ and $\tau \in \mathbb{T}^+$, if $\Theta^\tau(\omega,x) \in A_U$ then $(\omega,x) \in A_U$. Moreover, A_U is $(\mathcal{F}_0^\infty \otimes \Sigma)$-measurable, since (due to the right-contintuity of φ) it can be expressed as

$$A_U \;=\; \bigcup_{t \in D} \{\, (\omega,x) \in \Omega \times X \,:\, \varphi(t,\omega)x \in U \,\}$$

where D may be any countable dense subset of $\mathbb{T}^+$. Consequently, by Lemma 2.21(ii), $\mathbb{P} \otimes \rho(A_U)$ is equal to either 0 or 1. Note that $\Omega \times U \subset A_U$, so if $\rho(U) > 0$ then $\mathbb{P} \otimes \rho(\Omega \times U) = \rho(U) > 0$ and therefore $\mathbb{P} \otimes \rho(A_U) = 1$.

Let $\mathcal{U}$ be a countable base for the topology of X, and let $\mathcal{V} := \{U \in \mathcal{U} : U \cap \operatorname{supp}\rho \neq \varnothing\}$. Observe that the set $A^{(0)} \subset \Omega \times X$ of points (ω,x) whose trajectory $\{\varphi(t,\omega)x\}_{t \in \mathbb{T}^+}$ densely covers $\operatorname{supp}\rho$ is given by

$$\begin{aligned} A^{(0)} \;&:=\; \left\{ (\omega,x) \in \Omega \times X \,:\, \operatorname{supp}\rho \subset \overline{\{\varphi(t,\omega)x : t \in \mathbb{T}^+\}} \right\} \\ &=\; \left\{ (\omega,x) \in \Omega \times X \,:\, \operatorname{supp}\rho \subset \overline{\{\, \pi_X(\Theta^t(\omega,x)) \,:\, t \in \mathbb{T}^+ \,\}} \right\} \\ &=\; \bigcap_{U \in \mathcal{V}} A_U. \end{aligned}$$

Consequently, for any $T \in \mathbb{T}^+$, the set $A^{(T)} \subset \Omega \times X$ of points (ω, x) whose trajectory subsequent to time T densely covers $\operatorname{supp}\rho$ is given by

$$\begin{aligned} A^{(T)} &:= \left\{ (\omega, x) \in \Omega \times X \,:\, \operatorname{supp}\rho \subset \overline{\{\varphi(t,\omega)x : t \geq T\}} \right\} \\ &= \left\{ (\omega, x) \in \Omega \times X \,:\, \operatorname{supp}\rho \subset \overline{\{\, \pi_X(\Theta^{T+t}(\omega, x)) \,:\, t \in \mathbb{T}^+ \,\}} \right\} \\ &= \Theta^{-T}\left(\bigcap_{U \in \mathcal{V}} A_U \right). \end{aligned}$$

So then, the set $A \subset \Omega \times X$ of points (ω, x) with the property that for all $T \in \mathbb{T}^+$, $\operatorname{supp}\rho \subset \overline{\{\varphi(t,\omega)x : t \geq T\}}$ is given by

$$A = \bigcap_{T \in \mathbb{N}} \Theta^{-T}\left(\bigcap_{U \in \mathcal{V}} A_U \right).$$

Now for any $U \in \mathcal{V}$, $\rho(U) > 0$ and therefore $\mathbb{P} \otimes \rho(A_U) = 1$. Hence $\mathbb{P} \otimes \rho(A) = 1$.

In other words, $(\mathbb{P} \otimes \rho)$-almost every $(\omega, x) \in \Omega \times X$ has the property that for all $T \in \mathbb{T}^+$, $\operatorname{supp}\rho \subset \overline{\{\varphi(t,\omega)x : t \geq T\}}$. But also, since $\operatorname{supp}\rho$ is an *invariant* closed ρ-full measure set, it is clear that $(\mathbb{P} \otimes \rho)$-almost every $(\omega, x) \in \Omega \times X$ has the property that $\overline{\{\varphi(t,\omega)x : t \in \mathbb{T}^+\}} \subset \operatorname{supp}\rho$. So we are done. □

Now for any $x \in X$, let $G_x \subset X$ be the smallest invariant set containing x; as in Section C.4, this can be written explicitly as

$$G_x = \overline{\bigcup_{t \in \mathbb{T}^+} \operatorname{supp} \varphi_x^t}.$$

Note in particular that for any open $U \subset X$, $U \cap G_x$ is non-empty if and only if U is accessible from x. In other words, G_x is precisely the set of points $y \in X$ such that every neighbourhood of y is accessible from x. Obviously (by definition), for any $x \in X$, for any $y \in G_x$, we have that $G_y \subset G_x$.

Lemma 2.77. *Fix a metrisation d of the topology of X. The map $(x, y) \mapsto d(x, G_y)$ from $X \times X$ to $[0, \infty)$ is upper semicontinuous.*

Proof. Let (x_n) and (y_n) be convergent sequences in X, with limits x and y respectively, such that the sequence $r_n := d(x_n, G_{y_n})$ converges to a value c as $n \to \infty$. And suppose for a contradiction that $c > d(x, G_y) =: r$. Then on the one hand, $B_{\frac{1}{2}(r+c)}(x)$ is accessible from y; but on the other hand, since for every n we have that $B_{r_n}(x_n)$ is not accessible from y_n, we therefore have that for all n sufficiently large, $B_{\frac{1}{2}(r+c)}(x)$ is not accessible from y_n. So the set of points from which $B_{\frac{1}{2}(r+c)}(x)$ is accessible includes y but excludes y_n for sufficiently large n; this contradicts the fact that the set of points from which an open set is accessible is itself open. □

Now we will say that a set $G \subset X$ is *minimal* (with respect to φ) if the following equivalent statements hold:

(i) G is closed and invariant, and the only closed invariant proper subset of G is $\varnothing$;

(ii) G is a non-empty closed invariant set, and for all $x \in G$, $G_x = G$;

(iii) G is a non-empty closed invariant set, and for all $x \in G$ and open $U \subset X$ with $U \cap G \neq \varnothing$, U is accessible from x.

(So G is minimal with respect to φ if and ony if G is minimal according to the Markov transition probabilities $(\varphi_x^t)_{x \in X,\, t \in \mathbb{T}^+}$.) If the whole phase space X is minimal (i.e. if the only closed invariant sets are X and $\varnothing$), we say that φ *has minimal dynamics* (*on* X). Obviously, if φ has minimal dynamics then every stationary probability measure of the Markov transition probabilities $(\varphi_x^t)_{x \in X,\, t \in \mathbb{T}^+}$ has full support.

Remark 2.78. As in Appendix C, every non-empty compact invariant set contains a minimal set. Hence in particular, if X is compact then there exists at least one minimal set. Also note that for any non-empty compact invariant set C, if C contains *only one* minimal set K, then *every* non-empty closed invariant subset of C must contain K.

The following fairly intuitive lemma (which is not really specific to RDS but can be generalised to homogeneous Markov processes with sufficient continuity properties) will play a crucial role in the proofs of some of our results:

Lemma 2.79. *Suppose $K \subset X$ is a compact set possessing no non-empty closed invariant subsets. Then given any $x \in X$, for $\mathbb{P}$-almost every $\omega \in \Omega$ there exist arbitrarily large times $t \in \mathbb{T}^+ \cap \mathbb{Q}$ such that $\varphi(t,\omega)x \notin K$.*

The proof of Lemma 2.79 is essentially the same as the proof of [BS88, Proposition 4.1].

We will use the following general fact:

Lemma 2.80. *Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{T}^+}, \mathbb{P})$ be a filtered probability space, and let $(M_t)_{t \in \mathbb{T}^+}$ be an X-valued homogeneous Markov process with respect to $(\mathcal{F}_t)_{t \in \mathbb{T}^+}$, with transition probabilities $(\mu_x^t)_{x \in X,\, t \in \mathbb{T}^+}$. Fix $s \in \mathbb{T}^+$, let D be a countable subset of $\mathbb{T}^+$, and let $T : \Omega \to D$ be an $\mathcal{F}_s$-measurable function. Then for any $A \in \mathcal{B}(X)$,*

$$\mathbb{P}(M_{s+T} \in A | \mathcal{F}_s) \overset{\mathbb{P}\text{-a.s.}}{=} \mu_{M_s}^T(A).$$

Proof of Lemma 2.80. First observe that $\omega \mapsto \mu_{M_s(\omega)}^{T(\omega)}(A)$ is indeed $\mathcal{F}_s$-measurable: for any $I \in \mathcal{B}([0,1])$, we have

$$\left\{\omega : \mu_{M_s(\omega)}^{T(\omega)}(A) \in I\right\} = \bigcup_{t \in D} \Big(\{\omega : T(\omega) = t\} \cap \{\omega : \mu_{M_s(\omega)}^t(A) \in I\} \Big) \in \mathcal{F}_s.$$

Now for any $E \in \mathcal{F}_s$, we have

$$\begin{aligned}
\int_E \mathbb{1}_A(M_{s+T(\omega)}(\omega))\, \mathbb{P}(d\omega) &= \sum_{t \in D} \int_{E \cap T^{-1}(\{t\})} \mathbb{1}_A(M_{s+t}(\omega))\, \mathbb{P}(d\omega) \\
&= \sum_{t \in D} \int_{E \cap T^{-1}(\{t\})} \mu_{M_s(\omega)}^t(A)\, \mathbb{P}(d\omega) \\
&= \int_E \mu_{M_s(\omega)}^{T(\omega)}(A)\, \mathbb{P}(d\omega)
\end{aligned}$$

as required. □

Proof of Lemma 2.79. Let $D := \mathbb{T}^+ \cap \mathbb{Q}$. Fix $x \in X$ and let $M_t(\omega) := \varphi(t,\omega)x$ for all t and ω. By Corollary 2.12, (M_t) is a Markov process with transition probabilities $(\varphi_x^t)_{x\in X,\,t\in\mathbb{T}^+}$. For each $y \in K$, $G_y \notin K$ (since K admits no non-empty closed invariant subsets), and so $X \setminus K$ is accessible from y; hence, as in the proof of Lemma 2.72, there exists $\tau \in D$ such that $\varphi_y^\tau(K) < 1$.[25] So, defining the function $l\colon K \to [0,1]$ by

$$l(y) \;:=\; \inf_{t\in D} \varphi_y^t(K),$$

we have that l is strictly less than 1 on the whole of K. Note that for each $t \in D$, the map $y \mapsto \varphi_y^t(K)$ is upper semicontinuous; therefore, l is upper semicontinuous. So then, since K is compact, l has a maximum value c', which is strictly less than 1. Fix a value $c \in (c',1)$. Obviously, for all $y \in K$ there exists $\tau \in D$ such that $\varphi_y^\tau(K) \le c$; but moreover, one can easily construct a *measurable* function $\tau : K \to D$ such that $\varphi_y^{\tau(y)}(K) \le c$ for all $y \in K$.[26] We extend τ to the whole of X by setting $\tau(y) = 0$ for all $y \in X \setminus K$.

Now to obtain the desired result, it is sufficient just to show that for each $N \in \mathbb{N}$, for $\mathbb{P}$-almost every $\omega \in \Omega$ there exists $t \in D$ with $t \ge N$ such that $\varphi(t,\omega)x \notin K$. Fix any $N \in \mathbb{N}$, and define an increasing sequence $(T_n)_{n\in\mathbb{N}_0}$ of functions $T_n : \Omega \to D$ by

$$\begin{aligned} T_0(\omega) &= N \\ T_n(\omega) &= T_{n-1}(\omega) + \tau(M_{T_{n-1}(\omega)}(\omega)) \quad (n \ge 1) \end{aligned}$$

for all $\omega \in \Omega$. Note that for each ω, the sequence $T_n(\omega)$ is strictly increasing in n until the first point n^* at which $M_{T_{n^*}(\omega)}(\omega) \notin K$, beyond which the sequence remains constant.

For each $n \in \mathbb{N}_0$, let $E_n := \{\,\omega \in \Omega : M_{T_n(\omega)}(\omega) \in K\,\}$. Obviously if $\bigcap_{n=0}^\infty E_n$ is $\mathbb{P}$-null set, then in particular we have that $\mathbb{P}$-almost every $\omega \in \Omega$ there exists $t \ge N$ such that $\varphi(t,\omega)x \notin K$, as is required. So we will show that $\bigcap_{n=0}^\infty E_n$ is $\mathbb{P}$-null set; to do this, we will prove by induction that for each $n \in \mathbb{N}_0$, $\mathbb{P}(E_n) \le c^n$.

The $n = 0$ case is trivial. Now fix any $m \in \mathbb{N}_0$ such that $\mathbb{P}(E_m) \le c^m$. First observe that for each $s \in D$, the set $E_m \cap T_m^{-1}(\{s\}) \in \mathcal{F}_s$-measurable. One way to see this is as follows: provided $m \ge 1$ and $s > N$, we can express $E_m \cap T_m^{-1}(\{s\})$ as

$$\left(\bigcup_{\substack{t_0,\dots,t_m\in D \\ N=t_0<\dots<t_m=s}} \{\,\omega\in\Omega : \tau(M_{t_r}(\omega)) = t_{r+1}-t_r \text{ for all } 0\le r\le m-1\,\} \right) \cap \{\omega\in\Omega : M_s(\omega)\in K\};$$

otherwise we have

$$E_m \cap T_m^{-1}(\{s\}) \;=\; \begin{cases} \varnothing & s < N \\ \varnothing & m \ge 1 \text{ and } s = N \\ \varnothing & m = 0 \text{ and } s > N \\ M_s^{-1}(K) & m = 0 \text{ and } s = N. \end{cases}$$

[25] Using the compactness of K, one can show that τ can be taken from a *bounded* interval $[0,T]$ (where T is independent of y). Consequently (as in [BS88, Proposition 4.1]) in addition to proving Lemma 2.79, one can make a statistical statement about the length of time taken to escape from K; however, we will not need this for our purposes.

[26] e.g. if $(s_n)_{n\in\mathbb{N}}$ is an enumeration of D, set $\tau(y) := s_{N(y)}$ where $N(y) := \min\{n \in \mathbb{N} : \varphi_y^{s_n}(K) < 1\}$.

So then,

$$\begin{aligned}\mathbb{P}(E_{m+1}) &= \int_{E_m} \mathbb{1}_{E_{m+1}}(\omega)\,\mathbb{P}(d\omega) \quad (\text{since } E_{m+1} \subset E_m) \\ &= \int_{E_m} \mathbb{1}_K(M_{T_{m+1}(\omega)}(\omega))\,\mathbb{P}(d\omega) \\ &= \int_{E_m} \mathbb{1}_K(M_{T_m(\omega)+\tau(M_{T_m(\omega)}(\omega))}(\omega))\,\mathbb{P}(d\omega) \\ &= \sum_{s\in D} \int_{E_m \cap T_m^{-1}(\{s\})} \mathbb{1}_K(M_{s+\tau(M_s(\omega))}(\omega))\,\mathbb{P}(d\omega) \\ &= \sum_{s\in D} \int_{E_m \cap T_m^{-1}(\{s\})} \varphi_{M_s(\omega)}^{\tau(M_s(\omega))}(K)\,\mathbb{P}(d\omega) \\ &\qquad (\text{by Lemma 2.80 with } T := \tau(M_s)) \\ &\le \sum_{s\in D} \int_{E_m \cap T_m^{-1}(\{s\})} c\,\mathbb{P}(d\omega) \\ &= c\,\mathbb{P}(E_m) \\ &\le c^{m+1}.\end{aligned}$$

So we are done. □

We have the following important corollary:

Corollary 2.81. *Let $C \subset X$ be a compact invariant set, and suppose that C contains only one minimal set K. Let $U \subset X$ be an open set with $U \cap K \neq \emptyset$. Then for each $x \in C$, for $\mathbb{P}$-almost every $\omega \in \Omega$ there exist (arbitrarily large) times $t \in \mathbb{T}^+ \cap \mathbb{Q}$ such that $\varphi(t,\omega)x \in U$.*

(Note that one particular case of this is the case that C is itself minimal.)

Proof. By Remark 2.78, $C \setminus U$ cannot possess any non-empty closed invariant subsets. Hence Lemma 2.79 combined with the invariance of C gives the result. □

Now since we assume that the RDS φ is right-continuous, it is easy to check that the n-point motion $\varphi^{\times n}$ is a right-continuous RDS on X^n (equipped with the product topology). For any $(x,y) \in X \times X$ we will write $G_{(x,y)} \subset X \times X$ to denote the smallest closed invariant set under $\varphi^{\times 2}$ containing (x,y).

Let us denote the standard projections from $X \times X$ to X by $\pi_1 \colon (x,y) \mapsto x$ and $\pi_2 \colon (x,y) \mapsto y$.

Lemma 2.82. *For any $x, y \in X$, $\overline{\pi_1(G_{(x,y)})} = G_x$ and $\overline{\pi_2(G_{(x,y)})} = G_y$.*

Hence in particular, if $G_{(x,y)}$ is compact then $\pi_1(G_{(x,y)}) = G_x$ and $\pi_2(G_{(x,y)}) = G_y$.

Proof of Lemma 2.82. Let $A := \pi_1(G_{(x,y)})$; so we need to show that $\bar{A} = G_x$. We first show that $\bar{A}$ is invariant; for this, it is sufficient to show that for every $u \in A$, $X \setminus \bar{A}$ is not accessible from u. Fix $u \in A$, and let $v \in X$ be such that $(u,v) \in G_{(x,y)}$. Obviously (by definition) the sets $(X \setminus A) \times X$ and $G_{(x,y)}$ are mutually disjoint; and so, since $G_{(x,y)}$ is invariant, $(X \setminus \bar{A}) \times X$ is not accessible from (u,v). Hence $X \setminus \bar{A}$ is not accessible from u.

It remains to show that $\bar{A}$ admits no closed invariant proper subsets containing x. Let C be a closed proper subset of $\bar{A}$ containing x; we will show that C is not invariant. Since C is closed and $\bar{A} \not\subset C$, it follows that $A \not\subset C$, and thus the sets $(X \setminus C) \times X$ and $G_{(x,y)}$ have non-empty intersection. Therefore $(X \setminus C) \times X$ is accessible from (x,y), and so $X \setminus C$ is accessible from x. Thus C is not invariant.

Hence we have shown that $\overline{\pi_1(G_{(x,y)})} = G_x$. Likewise, $\overline{\pi_2(G_{(x,y)})} = G_y$. □

2.12 Invertible RDS in a topological setting

As in Section 2.11, fix a separable metrisable topology on X generating Σ.

Lemma 2.83. *If $\varphi(t,\omega)\colon X \to X$ is a homeomorphism for all t and ω, then φ is invertible (in the sense of Section 2.9).*

Proof. Fix $t \in \mathbb{T}^+$; we need to show that the map $(\omega, x) \mapsto \varphi(t,\omega)^{-1}(x)$ is $(\mathcal{F}_0^t \otimes \Sigma, \Sigma)$-measurable. Since X is separable and the map $x \mapsto \varphi(t,\omega)^{-1}(x)$ is continuous for each ω, it is sufficient[27] to show that the map $\omega \mapsto \varphi(t,\omega)^{-1}(x)$ is $(\mathcal{F}_0^t, \Sigma)$-measurable for each x. Fix $x \in X$ and a closed set $G \subset X$. Let $S \subset G$ be a countable set that is dense in G. Note that for every $\omega \in \Omega$, $\varphi(t,\omega)G$ is closed and $\varphi(t,\omega)S$ is dense in $\varphi(t,\omega)G$; hence we have that $x \in \varphi(t,\omega)G$ if and only if every neighbourhood of x intersects $\varphi(t,\omega)S$. In other words, fixing a metrisation of the topology of X, we have

$$\{\omega \in \Omega : \varphi(t,\omega)^{-1}(x) \in G\} = \bigcap_{n=1}^{\infty} \bigcup_{y \in S} \{\omega \in \Omega : \varphi(t,\omega)y \in B_{\frac{1}{n}}(x)\}.$$

Clearly the RHS is $\mathcal{F}_0^t$-measurable. So we are done. □

Definition 2.84. We will say that φ is *right-continuously invertible* if φ is right-continuous, invertible, and has the property that for any decreasing[28] sequence (t_n) in $\mathbb{T}^+$ converging to a value t and any sequence (x_n) in X converging to a point x, $\varphi(t_n,\omega)^{-1}(x_n) \to \varphi(t,\omega)^{-1}(x)$ as $n \to \infty$ for all $\omega \in \Omega$.

Definition 2.85. We will say that φ is *continuously invertible* if φ is continuous, invertible, and has the property that the map $(t,x) \mapsto \varphi(t,\omega)^{-1}(x)$ is continuous for all $\omega \in \Omega$.

Now we will say that a σ-locally compact metrisable space Y *respects inverses* if for any sequence of homeomorphisms $f_n\colon Y \to Y$ converging uniformly on compact sets to a homeomorphism $f\colon Y \to Y$, we have that f_n^{-1} converges uniformly on compact sets to f^{-1}. As in Appendix B, if either Y is compact or every point in Y has a neighbourhood that is contained in a compact connected set, then Y respects inverses.

Lemma 2.86. *Suppose X is σ-locally compact and respects inverses. (A) If φ is right-continuous and $\varphi(t,\omega)$ is a homeomorphism for all t and ω, then φ is right-continuously invertible. (B) If φ is continuous and $\varphi(t,\omega)$ is a homeomorphism for all t and ω, then φ is continuously invertible.*

[27] See [Cra02b, Lemma 1.1] or [New15a, Lemma 16(A)]

[28] Once again, here a "decreasing sequence" need *not* be strictly decreasing.

Proof. (A) Fix a decreasing sequence (t_n) in $\mathbb{T}^+$ converging to a value $t \in \mathbb{T}^+$. By Lemma B.7, the right-continuity of φ means precisely that for every $\omega \in \Omega$, $\varphi(t_n,\omega)$ converges uniformly on compact sets to $\varphi(t,\omega)$. Hence, since X respects inverses, $\varphi(t_n,\omega)^{-1}$ converges unifomly on compact sets to $\varphi(t,\omega)^{-1}$; so by Lemma B.7, for any sequence (x_n) in X converging to a point x, $\varphi(t_n,\omega)^{-1}(x_n)$ converges to $\varphi(t,\omega)^{-1}(x)$. (B) is similar, replacing a "decreasing sequence (t_n)" with any convergent sequence (t_n). ☐

With this, we can now prove Example 2.57:

Proposition 2.87. *For any Lipschitz 1-periodic $b:\mathbb{R}\to\mathbb{R}$ and any $\sigma\in\mathbb{R}$, the RDS on $\mathbb{S}^1$ generated by the SDE $d\phi_t = b(\phi_t)dt + \sigma dW_t$ is continuously invertible.*

Proof. Let φ be the RDS on $\mathbb{S}^1$ generated by the SDE $d\phi_t = b(\phi_t)dt + \sigma dW_t$. By Lemma 2.86, it is sufficient just to show that $\varphi(t,\omega)$ is a homeomorphism for all $t \geq 0$ and $\omega \in \Omega$; but since any continuous injective self-map of $\mathbb{S}^1$ is a homeomorphism, it is sufficient to show that $\varphi(t,\omega)$ is injective for all t and ω. Fix $t \geq 0$, $\omega \in \Omega$ and $x_0 \in \mathbb{S}^1$, and let $x_1 = \varphi(t,\omega)x_0$. Let $u:[0,t]\to\mathbb{R}$ be a continuous function such that $\pi(u(s)) = \varphi(t-s,\omega)x_0$ for all $s\in[0,t]$. Define $\bar{\omega}\in\Omega$ by $\bar{\omega}(s) = \omega(t-s)-\omega(t)$ for all $s\in\mathbb{R}$. Then we have that for all $\tau\in[0,t]$,

$$\begin{aligned}
u(\tau) &= u(t) + \int_0^{t-\tau} b(u(t-s))\,ds + \sigma\omega(t-\tau) \\
&= u(t) + \int_\tau^t b(u(s))\,ds + \sigma\omega(t-\tau) \\
&= u(0) + \int_\tau^t b(u(s))\,ds + \sigma\omega(t-\tau) + u(t) - u(0) \\
&= u(0) + \int_\tau^t b(u(s))\,ds + \sigma\omega(t-\tau) - \left(\int_0^t b(u(s))\,ds + \sigma\omega(t)\right) \\
&= u(0) + \int_0^\tau -b(u(s))\,ds + \sigma\bar{\omega}(\tau).
\end{aligned}$$

So, letting φ' denote the RDS on $\mathbb{S}^1$ generated by the SDE $d\phi_t = -b(\phi_t)dt + \sigma dW_t$, we have that

$$x_0 = \pi(u(t)) = \varphi'(t,\bar{\omega})\pi(u(0)) = \varphi'(t,\bar{\omega})x_1.$$

Recall that the point $x_0\in\mathbb{S}^1$ was arbitrary; hence $\varphi(t,\omega)$ is injective. ☐

Now recall that if φ is invertible, then for each $x \in X$ and $t \in \mathbb{T}^+$ we may define a probability measure $\bar{\varphi}_x^t$ on X by $\bar{\varphi}_x^t(A) = \mathbb{P}(\omega : \varphi(t,\omega)^{-1}(x)\in A)$.

Lemma 2.88. *Suppose φ is right-continuously invertible. Then for any open set $U\subset X$ the following are equivalent:*

(i) U is invariant;

(ii) U is crudely invariant (i.e. for each $t\in\mathbb{T}^+$, for $\mathbb{P}$-almost all $\omega\in\Omega$, $\varphi(t,\omega)U\subset U$);

(iii) $\bar{\varphi}_x^t(U) = 0$ for all $x\in X\setminus U$ and $t\in\mathbb{T}^+$.

(Observe that (iii) is the same as saying that $X\setminus U$ is very crudely invariant under the inverse RDS $\bar{\varphi}$.)

Proof. Let U be an open set, and let $G := X \setminus U$. Note that for any t and ω, $\varphi(t,\omega)U \subset U$ if and only if $\varphi(t,\omega)^{-1}(G) \subset G$. Hence the statement is proved by going through the proofs of Lemmas 2.74 and 2.75, replacing $\varphi(t,\omega)$ with $\varphi(t,\omega)^{-1}$ and φ_x^t with $\bar{\varphi}_x^t$. □

Remark 2.89. Even when φ is right-continuously invertible, a *very* crudely invariant open set $U \subset X$ need not be invariant. Indeed, if $\varphi(t,\omega)$ is bijective for all t and ω and the probability measure φ_x^t is atomless for all x and t, then it is easy to show that the complement of every finite set is very crudely invariant but not invariant.

Recall that we say that φ *has minimal dynamics* if the only closed invariant sets are X and $\varnothing$. In this case, every stationary probability measure of the Markov transition probabilities $(\varphi_x^t)_{x\in X,\, t\in\mathbb{T}^+}$ has full support.

Definition 2.90. Suppose φ is right-continuously invertible. We will say that φ *has reverse-minimal dynamics* (on X) if the only open invariant sets are X and $\varnothing$.

Note that (by characterisation (iii) in Lemma 2.88) this is the same as saying that X is minimal according to the Markov transition probabilities $(\bar{\varphi}_x^t)_{x\in X,\, t\in\mathbb{T}^+}$.

Obviously if φ is right-continuously invertible and has reverse-minimal dynamics, then every stationary probability measure of the Markov transition probabilities $(\bar{\varphi}_x^t)_{x\in X,\, t\in\mathbb{T}^+}$ has full support. But moreover, we have the following:

Lemma 2.91. *Suppose that X is infinite, that φ is right-continuously invertible, and that φ either has minimal dynamics or has reverse-minimal dynamics. Let ρ be a probability measure that is stationary under either the Markov transition probabilities $(\varphi_x^t)_{x\in X,\, t\in\mathbb{T}^+}$ or the Markov transition probabilities $(\bar{\varphi}_x^t)_{x\in X,\, t\in\mathbb{T}^+}$. Then ρ is atomless.*

Proof. Suppose for a contradiction that ρ is not atomless. As in the proof of Lemma 2.58, let P be the set of points of maximal mass according to ρ. Recall that P is crudely invariant; and therefore (since φ is a right-continuous RDS) P is invariant. So if φ has minimal dynamics then $P = X$; but this cannot be the case, since P is finite. On the other hand, since P is a finite invariant set, $\mathbb{S}^1 \setminus P$ is clearly also invariant, and therefore if φ has reverse-minimal dynamics then $\mathbb{S}^1 \setminus P = X$; but P is not empty, so this cannot be the case. □

Chapter 3. Measurable Dynamics, and Clusters of Trajectories

Overview

When we speak of "synchronising behaviour" in random dynamical systems, we refer broadly to the phenomenon that under a typical noise realisation, many trajectories of the RDS become very close to each other after a long time. Now if we wish to somehow "quantify" such synchronising behaviour, we can either consider *rates of mutual approach* of different trajectories, or we can consider the *scale* of the synchronising behaviour—that is, *how much* of the phase space will contract into a small region after a long time. In this section (and indeed, throughout this thesis), we focus on the latter consideration.

Suppose we have a RDS φ on a phase space X (which, for the moment, we assume to be equipped with some separable metric, so that we can measure the distance between trajectories). As in Remark 2.63, suppose we endow X with some initial distribution of mass ρ; in this case, let us take ρ to be an ergodic probability measure of the Markov transition probabilities associated to φ. We then simulate the forward-time running of the RDS, and see how the distribution of mass evolves over time. "Synchronising behaviour" corresponds to significant proportions of the mass becoming clustered into very small regions of space after a long time. Now if we wait a very long time, and then look to see how the mass has become distributed, we will observe one of the following scenarios:

(i) there is no obvious indication of any real synchronising behaviour;

(ii) virtually all of the mass has separated out into n tiny clusters (for some $n \in \mathbb{N}$), each of mass approximately equal to $\frac{1}{n}$.

If we start the process again, keeping the same initial mass distribution ρ but allowing the noise realisation to be different, we will observe the same scenario (with the same n if scenario (ii) occurs).

The above has essentially been demonstrated by Le Jan[1] in the context of a composition of random diffeomorphisms on a compact smooth manifold. It is known (e.g. [FGS14]) that the arguments can be extended well beyond this context. In fact, one of the goals of this chapter is to prove that the same phenomenon holds true for any RDS (*even if discontinuous*) on a Borel subset of a complete separable metric space satisfying the measurability requirements in Section 2.2.

When scenario (ii) occurs, we will refer to the number of clusters n as the *ρ-clustering number of φ*; and when scenario (i) occurs, we will say that *the ρ-clustering number of φ*

[1]See Lemme 1 and part (a) of the proof of Proposition 2 in [LeJ87]. Proposition 3 of [LeJ87] describes a stronger form of clustering that occurs when there is local asymptotic stability; see also Theorem 4.52 of this thesis.

is ∞. In the case that the ρ-clustering number of φ is 1, we will say that φ *is statistically synchronising with respect to* ρ.

Remark 3.1. Our above description of the phenomenon is, of course, very crude: the "scenarios" described above are actually statements about the *asymptotic* behaviour (with the clusters having "infinitesimal diameter" asymptotically), and may be formalised as in Corollary 3.9. It is well-known that there exist dynamical systems where it takes a remarkably long time for the "asymptotic picture" to start to develop within finite-time simulations; and it seems that this is particularly likely to arise with *random* dynamical systems, for the following reason: Due to the strong law of large numbers, if a model of noise allows for "freak events" with positive probability, then with *full* probability such "freak events" will happen infinitely often. It may well happen that such events, when they occur, have a "freak effect" on the phase space dynamics, while the overall effect of the more "normal" behaviour of the noise does nothing to counteract this. Accordingly, these "freak events" may be what determine the asymptotic behaviour of the RDS; and yet in such cases, since these events are so rare, one can expect it to take a very long time for the dynamics to begin to resemble the asymptotic dynamics. Such a scenario is particularly likely to occur for systems affected by Gaussian white noise, since the tails of the Gaussian distribution decay extraordinarily fast. Accordingly, there is a place for studying "intermediate time-scale dynamics" (as opposed to *asymptotic* dynamics) of RDS. In some cases, one possible way to do this is to study the asymptotic behaviour of the (not necessarily memoryless) RDS obtained when a very small perturbation is made to either the probability distribution $\mathbb{P}$ of the underlying noise or the action φ of the noise, in such a manner that sufficiently "extreme" behaviour of the noise now either has zero probability or no longer has an "extreme" effect on the dynamics. (In particular, this can serve as one motivation for the study of "bounded noise" RDS.) Nonetheless, it is out of the scope of this particular thesis to study "intermediate time-scale dynamics".

The major goal of this chapter of the thesis is to prove the following remarkable fact: The clustering number of a RDS is purely a "*measurable* dynamics" property; that is to say, given a RDS on a standard measurable space (X, Σ), the clustering number exists and is the same under all separable metrics[2] whose Borel σ-algebra coincides with Σ.

A further goal of this section is to prove that for a monotone RDS admitting an ergodic distribution, the associated clustering number is always equal to 1; in fact, provided the phase space is the real line (equipped with its usual ordering), and the RDS has appropriate continuity properties, the RDS will admit a "pullback-attracting random fixed point".

3.1 Statistical equilibria and clustering numbers

Let $(\Omega, \mathcal{F}, (\mathcal{F}_s^{s+t})_{s\in\mathbb{T}, t\in\mathbb{T}^+}, (\theta^t)_{t\in\mathbb{T}}, \mathbb{P})$ be a noise space (in accordance with our formalism in Section 2.2), let (X, Σ) be a standard measurable space, and let φ be a RDS on (X, Σ) over $(\Omega, \mathcal{F}, (\mathcal{F}_s^{s+t})_{s\in\mathbb{T}, t\in\mathbb{T}^+}, (\theta^t)_{t\in\mathbb{T}}, \mathbb{P})$ (in accordance with our formalism in Section 2.2).

[2]As in Remark 2.2, assuming the axiom of choice, *every* metric whose Borel σ-algebra coincides with Σ is separable.

For any measurable space $(E, \mathcal{E})$, we denote the set of probability measures on $(E, \mathcal{E})$ by $\mathcal{M}_{(E,\mathcal{E})}$, which we equip with the evaluation σ-algebra $\mathfrak{K}_{(E,\mathcal{E})}$. For convenience, we will drop the subscripts when considering the space of probability measures on (X, Σ); that is to say, we will just write $\mathcal{M}$ to denote the set of probability measures on (X, Σ), and we will just write $\mathfrak{K}$ for the evaluation σ-algebra on $\mathcal{M}$.

Hence, for example, an element $Q \in \mathcal{M}_{(\mathcal{M},\mathfrak{K})}$ is a *probability measure on the space of probability measures on* (X, Σ).

Given any separable metrisable topological space $(E, \mathcal{T})$, we write $\mathcal{N}_{\mathcal{T}}$ for the associated topology of weak convergence on $\mathcal{M}_{(E,\sigma(\mathcal{T}))}$. Recall that the Borel σ-algebra of $\mathcal{N}_{\mathcal{T}}$ is precisely $\mathfrak{K}_{(E,\sigma(\mathcal{T}))}$.

Recall that for any $A \subset X$, $\Delta_A := \{(x,x) : x \in A\}$. We will use the notations introduced in Section 2.10 for the n-point motions and the image-measure RDS, and their associated Markov transition probabilities.

Definition 3.2. For any $n \in \mathbb{N}$, let $\mathcal{K}_n \subset \mathcal{M}$ denote the set of probability measures ρ of the form $\rho = \frac{1}{n}\sum_{i=1}^{n} \delta_{x_i}$ for *distinct* points $x_1, \ldots, x_n \in X$. (In particular, $\mathcal{K}_1$ denotes the set of Dirac masses on X.) Let $\mathcal{K}_\infty \subset \mathcal{M}$ denote the set of atomless probability measures on X.

Lemma 3.3. *For all* $n \in \mathbb{N} \cup \{\infty\}$, $\mathcal{K}_n$ *is* $\mathfrak{K}$*-measurable.*

Proof. Fix a separable metrisable topology on X generating Σ, and let $\mathcal{U}$ be a countable base for this topology. For any open $V \subset X$, let $\mathcal{U}_V$ be the set of members of $\mathcal{U}$ that are contained in V. It is easy to show that for any finite n, a probability measure ρ on X belongs to $\mathcal{K}_n$ if and only if there exist mutually disjoint sets $V_1, \ldots, V_n \in \mathcal{U}$ such that for each $i \in \{1, \ldots, n\}$,

- $\rho(V_i) = \frac{1}{n}$, and
- for all $U \in \mathcal{U}_{V_i}$, $\rho(U)$ is equal to either 0 or $\frac{1}{n}$.

So then, writing $\mathfrak{U}_n$ to denote the collection of all mutually disjoint subcollections of $\mathcal{U}$ of size n—that is,

$$\mathfrak{U}_n := \{\mathcal{V} \subset \mathcal{U} : |\mathcal{V}| = n,\ V \cap \tilde{V} = \varnothing \text{ for all distinct } V, \tilde{V} \in \mathcal{V}\}$$

—we can express $\mathcal{K}_n$ as

$$\mathcal{K}_n = \bigcup_{\mathcal{V} \in \mathfrak{U}_n} \bigcap_{V \in \mathcal{V}} \left(\{\rho \in \mathcal{M} : \rho(V) = \tfrac{1}{n}\} \cap \bigcap_{U \in \mathcal{U}_V} \{\rho \in \mathcal{M} : \rho(U) \in \{0, \tfrac{1}{n}\}\} \right).$$

Hence $\mathcal{K}_n \in \mathfrak{K}$.

It remains to show that $\mathcal{K}_\infty \in \mathfrak{K}$. It is not hard to show that a probability measure ρ belongs to $\mathcal{K}_\infty$ if and only if $\rho \otimes \rho(\Delta_X) = 0$; one way to show this is as follows:

$$\begin{aligned}
\rho \in \mathcal{K}_\infty \;&\overset{\text{def}}{\iff}\; \rho(\{x\}) = 0 \;\forall\, x \in X \\
&\iff \rho(\{x\}) = 0 \text{ for } \rho\text{-almost all } x \in X \\
&\iff \int_X \mathbb{1}_{\Delta_X}(x,y)\,\rho(dy) \;=\; 0 \quad \text{for } \rho\text{-almost all } x \in X \\
&\iff \int_X \int_X \mathbb{1}_{\Delta_X}(x,y)\,\rho(dy)\,\rho(dx) \;=\; 0 \\
&\iff \rho \otimes \rho(\Delta_X) \;=\; 0.
\end{aligned}$$

So then, since the map $\rho \mapsto \rho \otimes \rho(\Delta_X)$ is measurable (Lemma A.15), $\mathcal{K}_\infty \in \mathfrak{K}$. □

Now for any $Q \in \mathcal{M}_{(\mathcal{M},\mathfrak{K})}$ and any $n \in \mathbb{N}$, we define the probability measure $E_n(Q)$ on X^n by

$$E_n(Q)(A) \;:=\; \int_{\mathcal{M}} \rho^{\otimes n}(A)\,Q(d\rho)$$

for all $A \in \Sigma^{\otimes n}$. (This is well-defined by Lemma A.15, and the monotone convergence theorem gives that this is a probability measure.)

Observe that for any $m, n \in \mathbb{N}$ with $m \leq n$, the image measure of $E_n(Q)$ under the projection $(x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_m)$ is precisely $E_m(Q)$.

Lemma 3.4. *Suppose we have $Q \in \mathcal{M}_{(\mathcal{M},\mathfrak{K})}$ and $n \in \mathbb{N} \cup \{\infty\}$ such that $Q(\mathcal{K}_n) = 1$. Then for any $A \in \Sigma$,*

$$E_2(Q)(\Delta_A) \;=\; \begin{cases} \frac{1}{n}E_1(Q)(A) & n < \infty \\ 0 & n = \infty. \end{cases}$$

Proof. First suppose that $n < \infty$. For any $\rho \in \mathcal{K}_n$, writing $\rho = \frac{1}{n}\sum_{x \in P} \delta_x$ where $|P| = n$, we have that

$$\begin{aligned}
\rho \otimes \rho(\Delta_A) \;&=\; \int_A \rho(\{x\})\,\rho(dx) \\
&=\; |A \cap P| \cdot \tfrac{1}{n^2} \\
&=\; n\rho(A) \cdot \tfrac{1}{n^2} \\
&=\; \frac{\rho(A)}{n}\,;
\end{aligned}$$

so

$$E_2(Q)(\Delta_A) \;=\; \int_{\mathcal{M}} \frac{\rho(A)}{n}\,Q(d\rho) \;=\; \tfrac{1}{n}E_1(Q)(A).$$

Now we have seen in the proof of Lemma 3.3 that for every $\rho \in \mathcal{K}_\infty$, $\rho \otimes \rho(\Delta_X) = 0$. Hence, if $n = \infty$ then $E_2(Q)(\Delta_X) = 0$, and therefore $E_2(Q)(\Delta_A) = 0$ for any $A \in \Sigma$. □

The following lemma is a useful link between the image-measure RDS and the n-point motions.

Lemma 3.5. *For any $n \in \mathbb{N}$ and any probability measure Q on $(\mathcal{M}, \mathfrak{K})$ that is stationary under the Markov transition probabilities $(\varphi_\rho^t)_{\rho \in \mathcal{M},\, t \in \mathbb{T}^+}$, $E_n(Q)$ is stationary under the Markov transition probabilities $(\varphi_{\mathbf{x}}^t)_{\mathbf{x} \in X^n,\, t \in \mathbb{T}^+}$.*

Proof. We need to show that for each $t \in \mathbb{T}^+$, $\varphi^{t*}_{(n)} E_n(Q) = E_n(Q)$. For any $t \in \mathbb{T}^+$ and $A \in \Sigma^{\otimes n}$, we have

$$\begin{aligned}
\left(\varphi^{t*}_{(n)} E_n(Q)\right)(A) &= \int_\Omega E_n(Q)\big(\,\varphi^{\times n}(t,\omega)^{-1}(A)\,\big)\,\mathbb{P}(d\omega) \\
&= \int_\Omega \int_{\mathcal{M}} \rho^{\otimes n}\left(\,\varphi^{\times n}(t,\omega)^{-1}(A)\,\right) Q(d\rho)\,\mathbb{P}(d\omega) \\
&= \int_\Omega \int_{\mathcal{M}} (\varphi(t,\omega)_*\rho)^{\otimes n}(A)\, Q(d\rho)\,\mathbb{P}(d\omega) \\
&= \int_\Omega \int_{\mathcal{M}} \rho^{\otimes n}(A)\ \varphi(t,\omega)_{**}Q(d\rho)\ \mathbb{P}(d\omega) \\
&= \int_{\mathcal{M}} \rho^{\otimes n}(A)\ \varphi^{t*}_{*}Q(d\rho) \\
&= \int_{\mathcal{M}} \rho^{\otimes n}(A)\, Q(d\rho) \\
&\qquad \text{(since } Q \text{ is stationary under } (\varphi^t_\rho)_{\rho\in\mathcal{M},\,t\in\mathbb{T}^+}) \\
&= E_n(Q)(A).
\end{aligned}$$

So we are done. □

We will use the notation $(E^n, \mathcal{T}^{\otimes n})$ to denote the n-fold product of a topological space $(E, \mathcal{T})$. We now state the central theorem of Chapter 3.[3]

Theorem 3.6. *Let ρ be a probability measure on X that is stationary under the Markov transition probabilities $(\varphi^t_x)_{x\in X,\,t\in\mathbb{T}^+}$. Then there exists a probability measure Q_ρ on $(\mathcal{M}, \mathfrak{K})$ that is ergodic with respect to the Markov kernel $(\varphi^t_{\tilde\rho})_{\tilde\rho\in\mathcal{M}}$ for every $t \in \mathbb{T}^+ \setminus \{0\}$, such that $E_1(Q_\rho) = \rho$ and for any separable metrisable topology $\mathcal{T}$ on X generating Σ, the following statements hold:*

(a) φ^t_ρ converges in $\mathcal{N}_{\mathcal{N}_\mathcal{T}}$ to Q_ρ as $t \to \infty$;

(b) for all $r \in \mathbb{N}$, $\varphi^{t*}_{(r)}(\rho^{\otimes r})$ converges in $\mathcal{N}_{\mathcal{T}^{\otimes r}}$ to $E_r(Q_\rho)$ as $t \to \infty$.

Moreover, if ρ is ergodic with respect to the Markov transition probabilities $(\varphi^t_x)_{x\in X,\,t\in\mathbb{T}^+}$, then there exists $n \in \mathbb{N} \cup \{\infty\}$ such that $Q_\rho(\mathcal{K}_n) = 1$.

(Note that by Lemma 3.5, for each $r \in \mathbb{N}$, $E_r(Q_\rho)$ is stationary under the Markov transition probabilities $(\varphi^t_{\mathbf{x}})_{\mathbf{x}\in X^r,\,t\in\mathbb{T}^+}$.)

Now observe that (as in Remark 2.13(II), but applied to the image-measure RDS) for any $t \in \mathbb{T}^+$, φ^t_ρ is precisely the law of the measure-valued random variable $\omega \mapsto \varphi(t,\omega)_*\rho$. So then, given any separable metrisable topology $\mathcal{T}$ on X generating Σ, we may regard Q_ρ as the *limiting distribution* of the Markov process $(\omega \mapsto \varphi(t,\omega)_*\rho)_{t\in\mathbb{T}^+}$ whose state space is the topological space $(\mathcal{M}, \mathcal{N}_\mathcal{T})$.

Definition 3.7. Let ρ be a stationary probability measure of the Markov transition probabilities $(\varphi^t_x)_{x\in X,\,t\in\mathbb{T}^+}$. We refer to the measure Q_ρ described in Theorem 3.6 as the *statistical equilibrium associated to ρ*.

[3] In Theorem 3.6, property (a) generalises [LeJ87, Lemme 1(b)], while property (b) generalises [Bax91, Proposition 2.6]; the final statement about the case that ρ is ergodic generalises part (a) of the proof of [LeJ87, Proposition 2].

Definition 3.8. Let ρ be an ergodic probability measure of the Markov transition probabilities $(\varphi_x^t)_{x\in X,\,t\in\mathbb{T}^+}$, and let $n \in \mathbb{N} \cup \{\infty\}$ be such that $Q_\rho(\mathcal{K}_n) = 1$. Then we refer to n as the *ρ-clustering number of φ*. In the case that $n = 1$, we will say that *φ is statistically synchronising with respect to ρ.*

(We will in fact see that φ is statistically synchronising with respect to ρ if and only if there exists an $\mathcal{F}^0_{-\infty}$-measurable random fixed point of φ whose law is ρ, formalising Remark 2.40.)

We now provide a more "geometric" interpretation of the clustering number (formalising the crude description given in the overview).

Let d be a separable metric on X whose Borel σ-algebra coincides with Σ. For any integer $n \geq 2$, any $\delta > 0$ and any $0 < \gamma < \delta$, we will say that a probability measure ρ on X is *(n,γ,δ)-clustered (according to d)* if there exist points $x_1,\ldots,x_n \in X$ such that $\min\{d(x_i,x_j) : i \neq j\} > \delta$ and for each $1 \leq i \leq n$, $\rho(B_\gamma(x_i)) > \frac{1-\gamma}{n}$.

Corollary 3.9. *Let ρ be an ergodic probability measure of the Markov transition probabilities $(\varphi_x^t)_{x\in X,\,t\in\mathbb{T}^+}$, and let n be the ρ-clustering number of φ. For any separable metric d on X whose Borel σ-algebra coincides with Σ, we have:*

(A) If $n = \infty$, then for all $\varepsilon > 0$ there exists $\delta > 0$ and $T \in \mathbb{T}^+$ such that for all $t \geq T$,

$$\mathbb{P}(\,\omega \in \Omega : \textit{for all } x \in X,\ \varphi(t,\omega)_*\rho(B_\delta(x)) \leq \varepsilon\,) \ > \ 1-\varepsilon.$$

(B) If $n = 1$, then for all $\varepsilon > 0$ there exists $T \in \mathbb{T}^+$ such that for all $t \geq T$,

$$\mathbb{P}(\,\omega \in \Omega : \textit{there exists } x \in X \textit{ s.t. } \varphi(t,\omega)_*\rho(B_\varepsilon(x)) > 1-\varepsilon\,) \ > \ 1-\varepsilon.$$

(C) If $2 \leq n < \infty$, then for all $\varepsilon > 0$ there exists $\delta > 0$ such that for all $0 < \gamma < \delta$ there exists $T \in \mathbb{T}^+$ such that for all $t \geq T$,

$$\mathbb{P}(\,\omega \in \Omega : \varphi(t,\omega)_*\rho \textit{ is } (n,\gamma,\delta)\textit{-clustered}\,) \ > \ 1-(\varepsilon+\gamma).$$

Proof. Fix a separable metric d on X whose Borel σ-algebra coincides with Σ, and let $\mathcal{T}$ be the induced topology. For any $r \in (0,1]$, let $\mathcal{J}_r \subset \mathcal{M}$ be the set of probability measures $\tilde{\rho}$ on X with the property that there exists $x \in X$ such that $\tilde{\rho}(\{x\}) \geq r$. We start by proving the following claim:

Claim 1. For any $r \in (0,1]$, if $\mathcal{J}_r$ is a Q_ρ-null set then the following holds: for all $\varepsilon > 0$ there exists $\delta > 0$ and $T \in \mathbb{T}^+$ such that for all $t \geq T$,

$$\mathbb{P}(\,\omega \in \Omega : \text{for all } x \in X,\ \varphi(t,\omega)_*\rho(B_\delta(x)) \leq r\,) \ > \ 1-\varepsilon.$$

Proof of Claim 1: Let S be a countable dense subset of X. For each $k \in \mathbb{N}$, let $\mathcal{J}_r^k \subset \mathcal{M}$ be the set of probability measures $\tilde{\rho}$ on X with the property that there exists $x \in S$ such that $\tilde{\rho}(B_{\frac{1}{4^k}}(x)) \geq r$. (Obviously, $\mathcal{J}_r^k$ is decreasing in k.) We first show that

(I) $\bigcap_{k=1}^\infty \mathcal{J}_r^k = \mathcal{J}_r$, and

(II) for each k, the $\mathcal{N}_{\mathcal{T}}$-closure of $\mathcal{J}_r^{k+1}$ is contained in $\mathcal{J}_r^k$.

To see that (I) holds: It is clear that $\mathcal{J}_r \subset \mathcal{J}_r^k$ for each k; in the other direction, suppose we have a probability measure $\tilde{\rho}$ belonging to $\bigcap_{k=1}^\infty \mathcal{J}_r^k$. Let $(x_k)_{k\in\mathbb{N}}$ be a sequence in S such that for each k, $\tilde{\rho}(B_{\frac{1}{4^k}}(x_k)) \geq r$. We now consider $\mathbb{N}$ equipped with the structure of a simple graph[4] where the set of edges E is as follows: for all distinct $k_1, k_2 \in \mathbb{N}$,

$$\{k_1, k_2\} \in E \iff B_{\frac{1}{4^{k_1}}}(x_{k_1}) \cap B_{\frac{1}{4^{k_2}}}(x_{k_2}) \neq \varnothing.$$

Since $\tilde{\rho}(X)$ is finite, it is clear that every induced subgraph of $\mathbb{N}$ must have finitely many connected components; hence in particular, every infinite induced subgraph of $\mathbb{N}$ must contain a connected infinite induced subgraph. So let $(G_j)_{j\in\mathbb{N}}$ be a sequence of connected infinite induced subgraphs of $\mathbb{N}$ such that for each $j \in \mathbb{N}$, $G_{j+1} \subset G_j \cap [j, \infty)$. For each $j \in \mathbb{N}$, let $U_j := \bigcup_{k\in G_j} B_{\frac{1}{4^k}}(x_k)$. Since $G_{j+1} \subset G_j$ for each j, we have that $U_{j+1} \subset U_j$ for each j; and since $\tilde{\rho}(U_j) \geq r$ for all j, it follows that $\tilde{\rho}\left(\bigcap_{j=1}^\infty U_j\right) \geq r$. Now since G_j is connected and $\min G_j \geq j$ for every $j \in \mathbb{N}$, it follows that $\operatorname{diam} U_j \to 0$ as $j \to \infty$. (This is due to the convergence of the series $\sum_k \frac{1}{4^k}$.) Hence $\bigcap_{j=1}^\infty U_j$ is a singleton. So $\tilde{\rho}$ belongs to $\mathcal{J}_r$.

To see that (II) holds: Fix $k \in \mathbb{N}$, and let $(\tilde{\rho}_j)_{j\in\mathbb{N}}$ be a sequence in $\mathcal{J}_r^{k+1}$ converging in $\mathcal{N}_{\mathcal{T}}$ to a measure $\tilde{\rho} \in \mathcal{M}$. Let $(x_j)_{j\in\mathbb{N}}$ be a sequence in S such that for each j, $\tilde{\rho}_j(B_{\frac{1}{4^{k+1}}}(x_j)) \geq r$. Let $U_j := \bigcup_{i=j}^\infty B_{\frac{1}{4^{k+1}}}(x_i)$ for each j. (Obviously U_j is decreasing in j.) First suppose for a contradiction that $\bigcap_{j=1}^\infty \overline{U_j}$ is empty. Then we may cover X by open sets V with the property that for sufficiently large $j \in \mathbb{N}$, $V \cap U_j = \varnothing$. Since X is separable, this cover admits a countable subcover $\{V_i\}_{i\in\mathbb{N}}$. For each $m \in \mathbb{N}$, let $W_m := \bigcup_{i=1}^m V_i$. For each m, we have that for sufficiently large j, $W_m \cap U_j = \varnothing$ and therefore $\tilde{\rho}_j(W_m) \leq 1 - r$. Consequently, $\tilde{\rho}(W_m) \leq 1 - r$ for each m; but since W_m increases to X as $m \to \infty$, this then implies $\tilde{\rho}(X) \leq 1 - r$, giving a contradiction. So then, $\bigcap_{j=1}^\infty \overline{U_j}$ is non-empty; so fix a point $x \in \bigcap_{j=1}^\infty \overline{U_j}$. It is clear that for infinitely many j, $B_{\frac{3}{4^{k+1}}}(x)$ contains $B_{\frac{1}{4^{k+1}}}(x_j)$ and therefore $\tilde{\rho}_j(B_{\frac{3}{4^{k+1}}}(x)) \geq r$. It follows that $\tilde{\rho}\left(\overline{B_{\frac{3}{4^{k+1}}}(x)}\right) \geq r$. Hence in particular, there obviously exists $\tilde{x} \in S$ such that $\tilde{\rho}(B_{\frac{1}{4^k}}(\tilde{x})) \geq r$. So $\tilde{\rho} \in \mathcal{J}_r^k$, as required.

Now then, since (I) holds, we have that $Q_\rho(\mathcal{J}_r^k) \to 0$ as $k \to \infty$. Since (II) holds, we have that for each k,

$$\limsup_{t\to\infty} \varphi_\rho^t(\mathcal{J}_r^{k+1}) \;\leq\; \limsup_{t\to\infty} \varphi_\rho^t\left(\overline{\mathcal{J}_r^{k+1}}\right) \;\leq\; Q_\rho\left(\overline{\mathcal{J}_r^{k+1}}\right) \;\leq\; Q_\rho(\mathcal{J}_r^k).$$

Combining these, we have that

$$\lim_{k\to\infty} \limsup_{t\to\infty} \varphi_\rho^t(\mathcal{J}_r^{k+1}) \;=\; 0.$$

[4] A *simple graph* is a set G equipped with a set E of 2-element subsets of G (called the *set of edges*). Given $x, y \in G$, it is said that x *is connected to* y if either $x = y$ or there exists $n \in \mathbb{N}$ and a list $(x_0, \ldots, x_n) \in X^{n+1}$ such that $x_0 = x$, $x_n = y$ and $\{x_{i-1}, x_i\} \in E$ for all $1 \leq i \leq n$. This defines an equivalence relation on G; the equivalence classes are called *connected components*. When there is only one connected component (namely, the whole of G), we say that G is *connected*. An *induced subgraph* of G is a set $H \subset G$ equipped with the set of edges $E_H := \{P \subset H : P \in E\}$.

So then, for every $\varepsilon > 0$ there exists $k \in \mathbb{N}$ and $T \in \mathbb{T}^+$ such that for all $t \geq T$, $\varphi^t_\rho(\mathcal{J}^{k+1}_r) < \varepsilon$; the statement that $\varphi^t_\rho(\mathcal{J}^{k+1}_r) < \varepsilon$ is precisely the statement that

$$\mathbb{P}(\,\omega \in \Omega : \text{there exists } x \in S \text{ s.t. } \varphi(t,\omega)_*\rho(B_\delta(x)) \geq r\,) \; < \; \varepsilon$$

where $\delta := \frac{1}{4^{k+1}}$, which implies that

$$\mathbb{P}(\,\omega \in \Omega : \text{for all } x \in S,\ \varphi(t,\omega)_*\rho(B_\delta(x)) \leq r\,) \; > \; 1-\varepsilon,$$

which (due to Lemma A.13) is precisely the same as saying that

$$\mathbb{P}(\,\omega \in \Omega : \text{for all } x \in X,\ \varphi(t,\omega)_*\rho(B_\delta(x)) \leq r\,) \; > \; 1-\varepsilon.$$

This completes the proof of Claim 1.

We now prove part (A): If $n = \infty$ then (by definition) $Q_\rho(\mathcal{K}_\infty) = 1$ and therefore $Q_\rho(\mathcal{J}_r) = 0$ for every $r \in (0,1]$. So then, for every $\varepsilon \in (0,1]$, applying Claim 1 with $r := \varepsilon$ gives that there exists $\delta > 0$ and $T \in \mathbb{T}^+$ such that for all $t \geq T$,

$$\mathbb{P}(\,\omega \in \Omega : \text{for all } x \in X,\ \varphi(t,\omega)_*\rho(B_\delta(x)) \leq \varepsilon\,) \; > \; 1-\varepsilon$$

as required. (The case that $\varepsilon > 1$ is obviously an automatic tautology.)

Now assuming $n < \infty$, let $\bar{\mathcal{K}}_n \subset \mathcal{M}$ be the set of probability measures $\tilde{\rho}$ taking the form $\tilde{\rho} = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$ for some $(x_1,\ldots,x_n) \in X^n$ (where the points $x_1,\ldots,x_n$ are *not* necessarily distinct). Moreover, for any $\gamma > 0$, let $\mathcal{I}^\gamma_n \subset \mathcal{M}$ be the set of probability measures $\tilde{\rho}$ for which there exist points $x_1,\ldots,x_n \in X$ with the following property: for any distinct $i_1,\ldots,i_m \in \{1,\ldots,n\}$,

$$\tilde{\rho}\left(\bigcup_{k=1}^m B_\gamma(x_{i_k})\right) \; > \; \frac{m(1-\gamma)}{n}.$$

(Due to Lemma A.13, for any countable dense $S \subset X$, it is always possible to choose the points $x_1,\ldots,x_n$ to belong to S.) Obviously, if $n \geq 2$ and there exists $\delta > \gamma$ such that $\tilde{\rho}$ is (n,γ,δ)-clustered, then $\tilde{\rho} \in \mathcal{I}^\gamma_n$.

Since $Q_\rho(\mathcal{K}_n) = 1$, we obviously have in particular that $Q_\rho(\bar{\mathcal{K}}_n) = 1$, and therefore:

Claim 2. For every $\gamma > 0$ there exists $T \in \mathbb{T}^+$ such that for all $t \geq T$, $\varphi^t_\rho(\mathcal{I}^\gamma_n) > 1-\gamma$.

Proof of Claim 2: Fix $\gamma > 0$. We first show that $\mathcal{I}^\gamma_n$ contains an $\mathcal{N}_\mathcal{T}$-open set U containing $\bar{\mathcal{K}}_n$. Let $(\tilde{\rho}_j)_{j\in\mathbb{N}}$ be an $\mathcal{N}_\mathcal{T}$-convergent sequence whose limit $\tilde{\rho}$ belongs to $\bar{\mathcal{K}}_n$; we need to show that for all j sufficiently large, $\tilde{\rho}_j$ belongs to $\mathcal{I}^\gamma_n$. But this is clear: writing $\tilde{\rho} = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$, we have that for any distinct $i_1,\ldots,i_m \in \{1,\ldots,n\}$,

$$\tilde{\rho}\left(\bigcup_{k=1}^m B_\gamma(x_{i_k})\right) \; \geq \; \frac{m}{n}$$

and therefore, for all j sufficiently large

$$\tilde{\rho}_j\left(\bigcup_{k=1}^m B_\gamma(x_{i_k})\right) \; > \; \frac{m(1-\gamma)}{n}$$

as required.

So then, letting U be an open set with $\bar{\mathcal{K}}_n \subset U \subset \mathcal{I}_n^\gamma$, we have

$$\liminf_{t\to\infty} \varphi_\rho^t(\mathcal{I}_n^\gamma) \;\geq\; \liminf_{t\to\infty} \varphi_\rho^t(U) \;\geq\; Q_\rho(U) \;\geq\; Q_\rho(\bar{\mathcal{K}}_n) \;=\; 1,$$

and therefore, for all t sufficiently large, $\varphi_\rho^t(\mathcal{I}_n^\gamma) > 1-\gamma$. This completes the proof of Claim 2.

We now prove part (B): Suppose $n=1$, and fix $\varepsilon > 0$. By Claim 2, there exists $T \in \mathbb{T}^+$ such that for all $t \geq T$, $\varphi_\rho^t(\mathcal{I}_1^\varepsilon) > 1-\varepsilon$, i.e.

$$\mathbb{P}(\,\omega \in \Omega : \text{there exists } x \in X \text{ s.t. } \varphi(t,\omega)_*\rho(B_\varepsilon(x)) > 1-\varepsilon\,) \;>\; 1-\varepsilon$$

as required.

We now prove part (C): Suppose $2 \leq n < \infty$. Given any $\delta > 0$ and any $0 < \gamma < \min(\delta, \frac{1}{3})$, if a probability measure $\tilde{\rho} \in \mathcal{I}_n^\gamma$ is *not* (n,γ,δ)-clustered then there exist $x_1, x_2 \in X$ such that $d(x_1,x_2) \leq \delta$ and $\tilde{\rho}(B_\gamma(x_1) \cup B_\gamma(x_2)) > \frac{2(1-\gamma)}{n} > \frac{4}{3n}$, from which it folllows that $\tilde{\rho}(B_{2\delta}(x_1)) > \frac{4}{3n}$. Hence a sufficient condition for a probability measure $\tilde{\rho}$ to be (n,γ,δ)-clustered is that $\tilde{\rho} \in \mathcal{I}_n^\gamma$ and for all $x \in X$, $\tilde{\rho}(B_{2\delta}(x)) \leq \frac{4}{3n}$.

Now fix $\varepsilon > 0$. Obviously $Q_\rho(\mathcal{J}_{\frac{4}{3n}}) = 0$, and so on the basis of Claim 1, let $\delta \in (0, \frac{1}{3})$ and $T' \in \mathbb{T}^+$ be such that for all $t \geq T'$,

$$\mathbb{P}(\,\omega \in \Omega : \text{for all } x \in X,\ \varphi(t,\omega)_*\rho(B_{2\delta}(x)) \leq \tfrac{4}{3n}\,) \;>\; 1-\varepsilon.$$

For $0 < \gamma < \delta$, by Claim 2 there exists $T'' \in \mathbb{T}^+$ such that for all $t \geq T''$,

$$\mathbb{P}(\,\omega \in \Omega : \varphi(t,\omega)_*\rho \in \mathcal{I}_n^\gamma\,) \;>\; 1-\gamma,$$

and so for all $t \geq T := \max(T', T'')$,

$$\mathbb{P}(\,\omega \in \Omega : \varphi(t,\omega)_*\rho \in \mathcal{I}_n^\gamma \text{ and for all } x \in X,\ \varphi(t,\omega)_*\rho(B_{2\delta}(x)) \leq \tfrac{4}{3n}\,) \;>\; 1-(\varepsilon+\gamma)$$

which implies

$$\mathbb{P}(\,\omega \in \Omega : \varphi(t,\omega)_*\rho \text{ is } (n,\gamma,\delta)\text{-clustered}\,) \;>\; 1-(\varepsilon+\gamma).$$

So we are done. □

We also have a further way in which to understand statistical synchronisation:

Corollary 3.10. *Let ρ be an ergodic probability measure of the Markov transition probabilities $(\varphi_x^t)_{x\in X,\, t\in\mathbb{T}^+}$. Fix a separable metric d on X whose Borel σ-algebra coincides with Σ, and for each $t \in \mathbb{T}^+$, define the function*

$$\begin{aligned} r_t : \Omega \times X \times X &\to [0,\infty) \\ (\omega, x, y) &\mapsto d(\,\varphi(t,\omega)x\,,\, \varphi(t,\omega)y\,). \end{aligned}$$

The following statements are equivalent:

(i) φ *is statistically synchronising with respect to* ρ*;*

(ii) as a stochastic process over the probability space $(\Omega \times X \times X, \mathcal{F} \otimes \Sigma \otimes \Sigma, \mathbb{P} \otimes \rho \otimes \rho)$*,* r_t *converges in probability to* 0 *as* $t \to \infty$*.*

Proof. First suppose that φ is statistically synchronising with respect to ρ. So $E_2(Q_\rho)(\Delta_X) = 1$. Fix $\varepsilon > 0$. Let

$$U_\varepsilon \; := \; \{(u,v) \in X \times X : d(u,v) < \varepsilon\}.$$

Since U_ε is a neighbourhood of Δ_X, we have that $\varphi^{t*}_{(2)}(\rho \otimes \rho)(U_\varepsilon) \to 1$ as $t \to \infty$ (by statement (ii) in Theorem 3.6). But $\varphi^{t*}_{(2)}(\rho \otimes \rho)(U_\varepsilon)$ is precisely equal to

$$\mathbb{P} \otimes \rho \otimes \rho(\,(\omega,x,y) \,:\, r_t(\omega,x,y) < \varepsilon\,).$$

So r_t converges in probability to 0 as $t \to \infty$.

Now suppose that φ is *not* statistically synchronising with respect to ρ. So $E_2(Q_\rho)(\Delta_X) < 1$. For each $\varepsilon > 0$, let

$$G_\varepsilon \; := \; \{(u,v) \in X \times X : d(u,v) \le \varepsilon\}.$$

It is clear that G_ε decreases as ε decreases, with the intersection $\bigcap_{\varepsilon > 0} G_\varepsilon$ being Δ_X. Hence there must exist $\varepsilon > 0$ such that $c := E_2(Q_\rho)(G_\varepsilon) < 1$. Since G_ε is closed, we have (by statement (ii) in Theorem 3.6) that for all t sufficiently large, $\varphi^{t*}_{(2)}(\rho \otimes \rho)(G_\varepsilon) \le c$. But $\varphi^{t*}_{(2)}(\rho \otimes \rho)(G_\varepsilon)$ is precisely

$$\mathbb{P} \otimes \rho \otimes \rho(\,(\omega,x,y) \,:\, r_t(\omega,x,y) \le \varepsilon\,).$$

So it follows in particular that r_t does not converge in probability to 0 as $t \to \infty$. □

Let us now mention the "deterministic" case of Theorem 3.6. Suppose Ω is a singleton $\{\omega\}$; then writing $f^t := \varphi(t,\omega)$, we have that for any probability measure ρ on X, φ^t_ρ is precisely equal to $\delta_{f^t_*\rho}$. Hence, for any (f^t)-invariant probability measure ρ, Q_ρ is simply equal to δ_ρ. So then, the final statement in Theorem 3.6 reduces to a simple statement about the atoms of an ergodic measure of an autonomous dynamical system, which we can easily prove directly.

Proposition 3.11. *Let* $(f^t)_{t \in \mathbb{T}^+}$ *be an autonomous dynamical system on* (X, Σ)*, and let* ρ *be an* (f^t)*-ergodic probability measure. Then either:*

(i) ρ *is atomless; or*

(ii) ρ *can be expressed in the form* $\rho = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$ *for distinct points* $x_1, \ldots, x_n \in X$ *forming the locus of a periodic orbit of* (f^t)*.*

In the case that $\mathbb{T}^+ = [0,\infty)$*,* ρ *must be either atomless or a Dirac mass at a fixed point of* (f^t)*.*[5]

[5]The proof that we present for this last statement is adapted from the answer to the MathOverflow question [MO15c] provided by Arnaud Chéritat.

Remark 3.12. The last statement of Proposition 3.11 does not extend to statistical equilibria of RDS in general: it is perfectly possible to have (under some metric d) a continuous RDS in continuous time admitting an ergodic distribution for which the associated clustering number is finite and strictly more than 1. For example, for any $n \in \mathbb{N}$, if $b: \mathbb{R} \to \mathbb{R}$ is a Lipschitz periodic function with least period $\frac{1}{n}$ then (for any $\sigma \neq 0$) the RDS on $\mathbb{S}^1$ generated by the SDE $d\phi_t = b(\phi_t)dt + \sigma dW_t$ has a ρ-clustering number of exactly n (where ρ is the unique stationary probability measure). See Corollary 5.27 for the proof.

Proof of Proposition 3.11. Suppose ρ is not atomless. Let $m := \max\{\rho(\{x\}) : x \in X\}$, and let $P := \{x \in X : \rho(\{x\}) = m\}$. Since ρ is invariant under (f^t), we have that for all $t \in \mathbb{T}^+$,

$$\rho(f^t(P)) = \rho(f^{-t}(f^t(P))) \geq \rho(P)$$

(where $f^{-t}(\cdot)$ denotes the preimage under f^t), and therefore (due to the definition of P), $f^t(P) = P$. Hence, due to the ergodicity of ρ, we have that $\rho(P) = 1$; in other words, writing $P =: \{x_1, \ldots, x_n\}$, we have that $\rho = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$. Once again, due to the ergodicity of ρ, the only invariant proper subset of P is $\varnothing$. Hence, fixing $x \in P$, the set $\{f^t(x) : t \geq T\}$ must be equal to P for every $T \in \mathbb{T}^+$; in other words, x is a periodic point of (f^t) whose trajectory is equal to P.

Now suppose that $\mathbb{T}^+ = [0, \infty)$. Suppose once again that ρ is not atomless; so ρ is a uniform distribution on set P of size n, forming the locus of a periodic trajectory of (f^t). To show that $n = 1$, it is sufficient to show that for every $t > 0$, the elements of P are t-periodic. Fix $t > 0$, and let $g := f^{\frac{t}{n!}}$. Obviously there exist $r \in \{1, \ldots, n\}$ and $x \in P$ such that $g^r(x) = x$. But since r divides $n!$, it follows that $f^t(x) = x$, and so the points of P are t-periodic. ☐

So Theorem 3.6 is not a particularly "deep" statement in the deterministic case; it particular, it says nothing about synchronisation. However, in the more general non-deterministic case, it is very common for an atomless ergodic probability measure of the Markov transition probabilities to have a finite clustering number (implying real synchronising behaviour).

3.2 Synchronisation and pullback-attracting random fixed points in monotone RDS

We now mention an important case where statistical synchronisation is guaranteed.

Theorem 3.13. *Suppose there exists a Borel linear order $\leq$ on (X, Σ) with respect to which φ is monotone. Then φ is statistically synchronising with respect to every ergodic probability measure of the Markov transition probabilities $(\varphi_x^t)_{x \in X, t \in \mathbb{T}^+}$.*

In the case that $\leq$ is just the standard ordering on $\bar{\mathbb{R}}$ or a Borel subset thereof, we actually find that there is a "(crudely) pullback-attracting" random fixed point. Although this is very much a "topology-specific" concept, for the sake of completeness we will now describe it here.

Standing Assumption. *For the rest of Section 3.2 (except the Open Question 3.22), let X be a Borel-measurable subset of $\bar{\mathbb{R}}$, equipped with the induced topology from $\bar{\mathbb{R}}$, and let Σ be the Borel σ-algebra. We equip X with the usual order $\leq$, and assume that φ is monotone with respect to $\leq$. Given any $A \subset X$, we write* $\inf A$ *and* $\sup A$ *to denote the infimum and supremum of A as a subset of $\bar{\mathbb{R}}$.*

(So $\inf A$ and $\sup A$ always exist, but might not be elements of X.)

Given a set $A \subset X$, a *left-accumulation point of* A is a point $x \in X$ that can be expressed as the limit of a strictly increasing sequence in A, and a *right-accumulation point of* A is a point $x \in X$ that can be expressed as the limit of a strictly decreasing sequence in A. An *isolated point of* X is a point $x \in X$ such that the singleton $\{x\}$ is open in X; this is equivalent to saying that x is neither a left-accumulation point of X nor a right-accumulation point of X.

Definition 3.14. Let $A \subset X$ be a convex set, and let $q: \Omega \to X$ be a random fixed point of φ. We say that q *is pullback-attracting over* A if $\mathbb{P}$-almost every $\omega \in \Omega$ has the property that for all $x \in A$, $\varphi(t, \theta^{-t}\omega)x \to q(\omega)$ as $t \to \infty$. We say that q *is crudely pullback-attracting over* A if for every unbounded countable set $S \subset \mathbb{T}^+$ there is a $\mathbb{P}$-full set $\Omega_S \subset \Omega$ such that for every $\omega \in \Omega_S$ and $x \in A$, $\varphi(t, \theta^{-t}\omega)x \to q(\omega)$ as t tends to ∞ in S.

Obviously, if $\mathbb{T} = \mathbb{Z}$ (or, more generally, if φ has left-continuous pullback trajectories), then any random fixed point that is crudely pullback-attracting over a convex set A is, in fact, pullback-attracting over A. We also have the following:

Lemma 3.15. *Suppose that φ is a right-continuous RDS. Let $q: \Omega \to X$ be a random fixed point that is crudely pullback-attracting over a convex set $A \subset X$. Suppose the following statements both hold:*

(a) if $\max A$ *exists and is a right-accumulation point of X, then* $\max A$ *is also a left-accumulation point of A;*

(b) if $\min A$ *exists and is a left-accumulation point of X, then* $\min A$ *is also a right-accumulation point of A.*

Then q is pullback-attracting over A.

(Note that statements (a) and (b) cover the case that A is open in X and the case that A is connected and not a singleton.)

Proof. Since the case that $\mathbb{T} = \mathbb{Z}$ is immediate, assume that $\mathbb{T} = \mathbb{R}$. Let us work with the metric $d(x, y) = \arctan|x - y|$. Fix any $\omega \in \Omega$ with the property that for all $x \in X_\rho$, $\varphi(t, \theta^{-t}\omega)x \to q(\omega)$ in $\mathbb{Q}$. Fix any $x \in X_\rho$ and $\varepsilon > 0$, and let $G := \bar{B}_\varepsilon(q(\omega))$; we will show that for all $t \in \mathbb{R}$ sufficiently large, $\varphi(t, \theta^{-t}\omega)x \in G$. We consider separately the following cases:

(I) x is in the interior (relative to X) of A;

(II) $x = \max A$, and x is a right-accumulation point of X and a left-accumulation point of A;

(III) $x = \min A$, and x is a left-accumulation point of X and a right-accumulation point of A.

First consider case (I). Let U be a neighbourhood of x such that $\bar{U} \subset A$ and $\min \bar{U} =: a$ and $\max \bar{U} =: b$ both exist. Let $T > 0$ be such that for all $t \in \mathbb{Q}$ with $t \geq T$, $\varphi(t,\theta^{-t}\omega)a$ and $\varphi(t,\theta^{-t}\omega)b$ both belong to G. Since φ is monotone and G is convex, we have that for all $t \in \mathbb{Q}$ with $t \geq T$, $\bar{U} \subset \varphi(t,\theta^{-t}\omega)^{-1}(G)$. Now fix any $t \in \mathbb{R}$ with $t > T$. Since the map $\tau \mapsto \varphi(\tau,\theta^{-t}\omega)x$ is right-continuous, we can choose $\delta \in (0, t-T)$ such that $t - \delta \in \mathbb{Q}$ and $\varphi(\delta,\theta^{-t}\omega)x \in U$. So then,

$$\begin{aligned} \varphi(t,\theta^{-t}\omega)x &= \varphi(t-\delta,\theta^{\delta-t}\omega)\varphi(\delta,\theta^{-t}\omega)x \\ &\in \varphi(t-\delta,\theta^{\delta-t}\omega)U \\ &\subset G \quad (\text{since } t-\delta \in \mathbb{Q} \text{ and } t-\delta > T). \end{aligned}$$

Now consider case (II). Take any $a \in A$ that is not an extreme point of A, and let $U := X \cap (a,x)$. Let $T > 0$ be such that (i) for all $t \in \mathbb{Q}$ with $t \geq T$, $\varphi(t,\theta^{-t}\omega)x \in G$, and (ii) for all $t \in \mathbb{R}$ with $t \geq T$, $\varphi(t,\theta^{-t}\omega)a \in G$. (Such a time T exists, by case (I).) Note that for all $t \in \mathbb{Q}$ with $t \geq T$, $U \subset \varphi(t,\theta^{-t}\omega)^{-1}(G)$. Now fix any $t \in \mathbb{R}$ with $t > T$. Suppose for a contradiction that $x \notin \varphi(t,\theta^{-t}\omega)^{-1}(G)$. Since $a \in \varphi(t,\theta^{-t}\omega)^{-1}(G)$ and $\varphi(t,\theta^{-t}\omega)^{-1}(G)$ is convex, it follows that $\sup \varphi(t,\theta^{-t}\omega)^{-1}(G) \leq x$. But moreover, since φ is continuous in space, we have that $\varphi(t,\theta^{-t}\omega)^{-1}(G)$ is closed in X and therefore $\sup \varphi(t,\theta^{-t}\omega)^{-1}(G) < x$. So fix any $y \in A$ with $\sup \varphi(t,\theta^{-t}\omega)^{-1}(G) < y < x$. Now U is a neighbourhood of y, and therefore, just as in case (I), we have that $\varphi(t,\theta^{-t}\omega)y \in G$; but this contradicts the fact that $\sup \varphi(t,\theta^{-t}\omega)^{-1}(G) < y$.

Case (III) is similar to case (II). □

Remark 3.16. Note that if a random fixed point $q\colon\Omega \to X$ is crudely pullback-attracting over some non-empty convex set $A \subset X$, then q has a modification $\tilde{q}$ that is $\mathcal{F}^0_{-\infty}$-measurable: fixing any $y \in A$, let $\tilde{q}(\omega)$ be equal to the limit of the sequence $\varphi(n,\theta^{-n}\omega)y$ if this limit exists, and some arbitrary constant otherwise. Hence in particular, by Lemma 2.39, $q_*\mathbb{P}$ is ergodic with respect to Markov transition probabilities $(\varphi^t_x)_{x\in X,\, t\in\mathbb{T}^+}$. Using the dominated convergence theorem, it is not hard to show that for every probability measure m on X with A being an m-full set, $\varphi^{t*}m$ converges weakly to $q_*\mathbb{P}$ as $t \to \infty$.

Remark 3.17. Suppose φ is a right-continuous RDS, and suppose we have a function $q\colon\Omega \to X$ such that for $\mathbb{P}$-almost every $\omega \in \Omega$, for all $x \in X$, $\varphi(t,\theta^{-t}\omega)x \to q(\omega)$ as $t \to \infty$. Let $\tilde{q}\colon\Omega \to X$ be as in Remark 3.16 (where we may take any $y \in X$). It is easy to show that $\tilde{q}$ is a random fixed point agreeing with q outside a null set, and that $\tilde{q}$ is pullback-attracting over X. Moreover, one can show that $\tilde{q}$ is a "strong" random fixed point, in the following sense: if we let

$$\hat{\Omega} := \{\omega \in \Omega : \text{for all } n \in \mathbb{N}_0 \text{ and } x \in X,\ \varphi(t,\theta^{-(n+t)}\omega)x \to \tilde{q}(\theta^{-n}\omega) \text{ as } t \to \infty\},$$

then $\hat{\Omega}$ is a $\mathbb{P}$-full set with the properties that

(i) $\theta^t(\hat{\Omega}) = \hat{\Omega}$ for all $t \in \mathbb{T}$, and

(ii) $\varphi(t,\omega)\tilde{q}(\omega) = \tilde{q}(\theta^t\omega)$ for all $t \in \mathbb{T}^+$ and $\omega \in \hat{\Omega}$.

(Moreover, $\hat{\Omega} \in \mathcal{F}_{-\infty}^{-\infty}$.)

Now recall that for any probability measure ρ on X, X_ρ denotes the smallest ρ-full-measure convex set (referred to as the "convex core of ρ").

Theorem 3.18. *Let ρ be a stationary probability measure of the Markov transition probabilities $(\varphi_x^t)_{x\in X,\, t\in\mathbb{T}^+}$.*

(A) If ρ is (φ_x^t)-ergodic then there is a random fixed point $q:\Omega \to X$ with law $q_*\mathbb{P} = \rho$ which is crudely pullback-attracting over X_ρ.

[The remaining statements simply concern conditions for ergodicity and unique ergodicity; they can be derived as a consequence of part (A).]

(B) If ρ is (φ_x^t)-ergodic then the only (φ_x^t)-stationary probability measure $\tilde{\rho}$ with $\tilde{\rho}(X_\rho) = 1$ is ρ.

(C) Hence, for any two distinct (φ_x^t)-ergodic probability measures ρ_1 and ρ_2, the sets X_{ρ_1} and X_{ρ_2} are disjoint.

(D) The following statements are equivalent:

(i) ρ is (φ_x^t)-ergodic;

(ii) every downward-inclusive crudely invariant set is either ρ-null or ρ-full;

(iii) every upwards-inclusive crudely invariant set is either ρ-null or ρ-full;

(iv) there are no non-empty convex crudely invariant proper subsets of X_ρ.

(E) Suppose that φ is a right-continuous RDS, and that ρ is not a Dirac mass and X_ρ is connected. Then ρ is (φ_x^t)-ergodic if and only if there are no deterministic fixed points in X_ρ. (In this case, the random fixed point q in part (A) is pullback-attracting over X_ρ.)

Example 3.19. (This example is the same as that studied in [AM14]). Let $X = (0,1)$, with Σ being the Borel σ-algebra. Fix any $c \in (0,\frac{1}{2})$. Let $I = \{0,1\}$ (with $\mathcal{I}$ being the discrete σ-algebra), let $\nu = \frac{1}{2}(\delta_0 + \delta_1)$, and define the functions $f_0, f_1 : (0,1) \to (0,1)$ by

$$f_0(x) = \begin{cases} \frac{1}{2(1-c)}\,x & x \in (0,1-c] \\ 1 - \frac{1}{2c}(1-x) & x \in [1-c,1) \end{cases}$$

$$f_1(x) = \begin{cases} \frac{1}{2c}\,x & x \in (0,c] \\ 1 - \frac{1}{2(1-c)}(1-x) & x \in [c,1). \end{cases}$$

In other words, f_0 and f_1 are piecewise-affine order-preserving homeomorphisms, with the point $(1-c,\frac{1}{2})$ being the only corner point of the graph of f_0, and the point $(c,\frac{1}{2})$ being the only corner point of the graph of f_1. Let φ be the RDS generated by the random map $(I,\mathcal{I},\nu,(f_i)_{i\in I})$. It is easy to show that the Lebesgue measure l on $(0,1)$ is stationary under the Markov transition probabilities $(\varphi_x^n)_{x\in(0,1),\, n\in\mathbb{N}_0}$. It is also clear that φ has no deterministic fixed points. Hence, by Theorem 3.18 (parts (A), (B) and (E)), l

is the unique (φ_x^n)-ergodic probability measure and φ admits a random fixed point with law l which is pullback-attracting over the whole of $(0,1)$.[6]

Remark 3.20 (cf. [AM14, Theorem 6.3])**.** Let q be any version of the random fixed point in Example 3.19. Then for any $n \in \mathbb{N}$, for any non-empty $E \in \mathcal{F}_{-n}^0$ and any $E' \subset E$ with $E \setminus E'$ being $\mathbb{P}$-null, the image $q(E')$ is a Lebesgue-full subset of $(0,1)$. To see this: Without loss of generality, assume $E' \in \mathcal{F}$. Let $\alpha_{-n+1}, \alpha_{-n+2}, \ldots, \alpha_0 \in I$ be such that the cylinder set $C := I^{\mathbb{Z}_{\leq -n}} \times \{(\alpha_{-n+1}, \alpha_{-n+2}, \ldots, \alpha_0)\} \times I^{\mathbb{N}}$ is contained in E. Let $\tilde{\Omega} \in \mathcal{F}$ be a $\mathbb{P}$-full set such that $\varphi(n, \theta^{-n}\omega)q(\theta^{-n}\omega) = q(\omega)$ for all $\omega \in \tilde{\Omega}$, and let $C' = C \cap \tilde{\Omega} \cap E'$. So $q(E') \supset q(C') \supset f_{\alpha_0} \circ \ldots \circ f_{\alpha_{-n+1}}(q(\theta^{-n}(C')))$. Since q has a modification that is $\mathcal{F}_{-\infty}^0$-measurable and $\mathcal{F}_{-\infty}^0$ is independent of $\mathcal{F}_0^n$, we must have that for every Lebesgue-positive measure set $S \in \mathcal{B}((0,1))$, $q^{-1}(S)$ has a positive-measure intersection with $\theta^{-n}(C)$ and therefore with $\theta^{-n}(C')$. But since (by the "measurable image theorem", [New15a, Exercise 104(B)]) the set $q(\theta^{-n}(C'))$ is universally measurable with respect to $\mathcal{B}((0,1))$, it follows that $q(\theta^{-n}(C'))$ is a Lebesgue-full set. Since the maps $f_{\alpha_{-n+1}}, \ldots, f_0$ are surjective and piecewise linear, they map Lebesgue-full sets onto Lebesgue-full sets,[7] and so $q(E')$ is a Lebesgue-full set.

Example 3.21 (adapted from [CF98])**.** Let $(\Omega, \mathcal{F}, (\mathcal{F}_s^{s+t})_{s\in\mathbb{R}, t\geq 0}, (\theta^t)_{t\in\mathbb{R}}, \mathbb{P})$ be as in Example 2.6, with $d = 1$. Let $X = \mathbb{R}$. Consider a RDS φ generated by an equation of the form

$$dx_t = f(x_t)dt + \sigma d\omega(t)$$

where $\sigma \neq 0$ and $f \in C^1(\mathbb{R}, \mathbb{R})$, satisfying the integrability condition that for an antiderivative F of f,

$$\int_{-\infty}^{\infty} e^{\frac{F(x)}{\sigma^2}} dx < \infty.$$

As in [CF98, Remark 3.7], there exists a unique stationary probability measure of the Markov transition probabilities (φ_x^t). By Theorem 3.18(A) (together with the last statement in part (E) of Theorem 3.18), there exists a random fixed point $q:\Omega \to \mathbb{R}$ that is pullback-attracting over $\mathbb{R}$. To illustrate: the deterministic ODE

$$dx_t = (\alpha x - x^3)dt \tag{3.1}$$

exhibits a supercritical pitchfork bifurcation as α crosses from negative to positive, but the Wiener-driven SDE

$$dx_t = (\alpha x - x^3)dt + \sigma dW_t \tag{3.2}$$

(where $\sigma \neq 0$) has a globally pullback-attracting random fixed point for all values of α. Hence, in this scenario, noise destroys the pitchfork bifurcation. (See also Example 6.7.)

Open Question 3.22. *Is Theorem 3.18(B) specific to the case that X is a Borel-ordered subspace of $\bar{\mathbb{R}}$, or does it hold for monotone RDS on any standard measurable space (X, Σ) equipped with a Borel linear order $\leq$?*

[6] The same result is obtained in [AM14], by a different method. As one of the steps within this method, it is proved that φ is synchronising in *forward* time (i.e. synchronising in the sense of Definition 4.6).

[7] More generally, a function $f: I \to J$ (where I and J are intervals) is said to have the *Luzin N property* if the image of any Lebesgue-null set is Lebesgue-null; if f is surjective and has the Luzin N property, then the image of any Lebesgue-full set is Lebesgue-full.

(If the answer is that Theorem 3.18(B) holds in the general case, then parts (C) and (D) of Theorem 3.18 also extend to the general case, since they follow from part (B) without any reference to the special structure of $\bar{\mathbb{R}}$.)

Remark 3.23. The notion of pullback-attracting random fixed points (and more general pullback-attracting "random invariant sets") is not specific to monotone systems on subsets of $\bar{\mathbb{R}}$. In particular, for a general RDS φ on a metric space (X,d) (with $\mathcal{B}(X)$ being standard), a random fixed point $q:\Omega \to X$ is said to be *globally pullback-attracting* if $\mathbb{P}$-almost every $\omega \in \Omega$ has the property that for every non-empty bounded $B \subset X$, $\sup_{x\in B} d(\varphi(t,\theta^{-t}\omega)x, q(\omega)) \to 0$ as $t \to \infty$.

Now let us make some comments about the physical significance of pullback-attraction. Pullback-attraction does not directly represent "synchronisation", since synchronisation concerns the mutual approach of *forward-time* trajectories, while pullback-attraction just concerns dynamics in the past. Nonetheless, it is true that "almost sure pullback-attraction towards a singleton implies (forward-time) synchronisation in probability". To be precise: suppose we have a random fixed point $q:\Omega \to X$ and a set $A \in \mathcal{B}(X)$ such that for $\mathbb{P}$-almost all $\omega \in \Omega$, $\sup_{x\in A} d(\varphi(t,\theta^{-t}\omega)x, q(\omega)) \to 0$ as $t \to \infty$; then the stochastic process[8] $\mathrm{diam}(\varphi(t,\cdot)A)$ converges in probability to 0 as $t \to \infty$. (This is due to the fact that almost sure convergence implies convergence in probability, combined with the (θ^t)-invariance of $\mathbb{P}$.)

The rest of Chapter 3 is now devoted to proving Theorems 3.6, 3.13 and 3.18.

3.3 Random measures and measure-valued stochastic processes

Convergence of stochastic processes

Let $(E,\mathcal{E})$ be any measurable space; although an "E-valued random variable" is normally written as an $(\mathcal{F},\mathcal{E})$-measurable function $Y:\Omega \to E$, one can alternatively write it as an Ω-indexed family $(x_\omega)_{\omega\in\Omega}$ of elements of E such that the map $\omega \mapsto x_\omega$ is $(\mathcal{F},\mathcal{E})$-measurable. Likewise, we have two possible notational conventions for a "stochastic process" taking values in E: one is to regard an "E-valued stochastic process" as being a $\mathbb{T}^+$-indexed family $(Y_t)_{t\in\mathbb{T}^+}$ of $(\mathcal{F},\mathcal{E})$-measurable functions $Y_t:\Omega \to E$, while the other is to regard an "E-valued stochastic process" as being a $(\mathbb{T}^+ \times \Omega)$-indexed family $(x_{t,\omega})_{t\in\mathbb{T}^+,\omega\in\Omega}$ of elements of E such that the map $\omega \mapsto x_{t,\omega}$ is $(\mathcal{F},\mathcal{E})$-measurable for each $t \in \mathbb{T}^+$.

Although, for random variables and for stochastic processes, the former convention is the more standard, we will often use the latter convention in this section.

Now let E be a separable metrisable topological space. Let $(x_\omega)_{\omega\in\Omega}$ be an E-valued random variable, and let $(x_{t,\omega})_{t\in\mathbb{T}^+,\omega\in\Omega}$ be an E-valued stochastic process.

[8] It is clear that the map $\omega \mapsto \mathrm{diam}(\varphi(t,\omega)A)$ is measurable for each t if φ is spatially continuous; in general, the map $\omega \mapsto \mathrm{diam}(\varphi(t,\omega)A)$ is universally measurable for each t. (This is a fairly straightforward consequence of the measurable projection theorem.)

Definition 3.24. We say that the stochastic process $(x_{t,\omega})$ *converges almost surely* to the random variable (x_ω) if $\mathbb{P}$-almost every $\omega \in \Omega$ has the property that $x_{t,\omega} \to x_\omega$ as $t \to \infty$.

Definition 3.25. We say that the stochastic process $(x_{t,\omega})$ *converges via countable subnets* to the random variable (x_ω) if for every unbounded countable set $S \subset \mathbb{T}^+$ there is a $\mathbb{P}$-full set $\Omega_S \subset \Omega$ such that for every $\omega \in \Omega_S$, $x_{t,\omega} \to x_\omega$ as t tends to ∞ in S.

Obviously if $\mathbb{T}^+ = \mathbb{N}_0$ then almost sure convergence and convergence via countable subnets are the same. If $\mathbb{T}^+ = [0,\infty)$, then the notion of almost sure convergence is strictly stronger than the notion of convergence via countable subnets; note that convergence via countable subnets is preserved under modification (whereas almost sure convergence is only preserved under indistinguishability).

Random probability measures

We use the notation that for any probability measure ρ on X and any ρ-integrable function $g\colon X \to \mathbb{R}$, $\rho(g) := \int_X g(x)\,\rho(dx)$. So for any $A \in \Sigma$, $\rho(A) = \rho(\mathbb{1}_A)$.

Recall that we write $\pi_\Omega\colon \Omega \times X \to \Omega$ and $\pi_X\colon \Omega \times X \to X$ to denote, respectively, the projections $(\omega,x) \mapsto \omega$ and $(\omega,x) \mapsto x$.

A "random probability measure on X" simply means an $(\mathcal{M},\mathfrak{K})$-valued random variable, that is, an Ω-indexed family $(\mu_\omega)_{\omega\in\Omega}$ of probability measures μ_ω on (X,Σ) such that the map $\omega \mapsto \mu_\omega(A)$ is measurable for all $A \in \Sigma$.

Given a sub-σ-algebra $\mathcal{G}$ of $\mathcal{F}$, we say that a random probability measure (μ_ω) is *$\mathcal{G}$-measurable* if the map $\omega \mapsto \mu_\omega$ is $(\mathcal{G},\mathfrak{K})$-measurable (i.e. if the map $\omega \mapsto \mu_\omega(A)$ is $\mathcal{G}$-measurable for all $A \in \Sigma$).

We say that two random probability measures (μ_ω) and $(\tilde{\mu}_\omega)$ on X are *equivalent* if $\mathbb{P}(\omega \in \Omega : \mu_\omega = \tilde{\mu}_\omega) = 1$.

We now introduce some notations: We write $\mathcal{L}^0(\Omega,\mathcal{F};\mathcal{M})$ for the set of all random probability measures on X. Given any sub-σ-algebra $\mathcal{G}$ of $\mathcal{F}$, we write $\mathcal{L}^0(\Omega,\mathcal{G};\mathcal{M})$ for the set of all random probability measures on (X,Σ) that are $\mathcal{G}$-measurable. We write $L^0(\mathbb{P};\mathcal{M})$ for the set of all equivalence classes of random probability measures on X. Likewise, given any sub-σ-algebra $\mathcal{G}$ of $\mathcal{F}$, we write $L^0(\mathbb{P}|_\mathcal{G};\mathcal{M})$ for the set of all equivalence classes of $\mathcal{G}$-measurable random probability measures on X.

Given a random probability measure (μ_ω) on X, we define the "mean probability measure" $\mathbb{E}_\omega\mu_\omega$ on X by

$$(\mathbb{E}_\omega\mu_\omega)(A) \;:=\; \int_\Omega \mu_\omega(A)\,\mathbb{P}(d\omega)$$

for all $A \in \Sigma$. Note that for any bounded measurable $g\colon X \to \mathbb{R}$,

$$(\mathbb{E}_\omega\mu_\omega)(g) \;=\; \int_\Omega \mu_\omega(g)\,\mathbb{P}(d\omega).$$

Also note that for any random variable $q\colon\Omega \to X$, $\mathbb{E}_\omega \delta_{q(\omega)}$ is precisely the law $q_*\mathbb{P}$ of q.

We have a "dominated convergence theorem for random probability measures":

Lemma 3.26. *Fix a separable metrisable topology $\mathcal{T}$ on X generating Σ. Let $(\mu_\omega)_{\omega\in\Omega}$ be a random probability measure, and let $(\mu_\omega^n)_{n\in\mathbb{N},\omega\in\Omega}$ be an $\mathcal{M}$-valued stochastic process converging almost surely to $(\mu_\omega)_{\omega\in\Omega}$ in $\mathcal{N}_\mathcal{T}$. Then $\mathbb{E}_\omega\mu_\omega^n \to \mathbb{E}_\omega\mu_\omega$ in $\mathcal{N}_\mathcal{T}$.*

Proof. Given any bounded continuous function $g\colon X \to \mathbb{R}$, we have that $\mu_n^\omega(g) \to \mu_\omega(g)$ for $\mathbb{P}$-almost all $\omega \in \Omega$, and so by the dominated convergence theorem

$$(\mathbb{E}_\omega\mu_\omega^n)(g) = \int_\Omega \mu_\omega^n(g)\,\mathbb{P}(d\omega) \;\to\; \int_\Omega \mu_\omega(g)\,\mathbb{P}(d\omega) = (\mathbb{E}_\omega\mu_\omega)(g)$$

as required. □

Now we say that a probability measure μ on $(\Omega\times X, \mathcal{F}\otimes\Sigma)$ is $\mathbb{P}$-*compatible* if $\pi_{\Omega*}\mu = \mathbb{P}$, that is, $\mu(E\times X) = \mathbb{P}(E)$ for all $E\in\mathcal{F}$. We write $\mathcal{M}^\mathbb{P}$ for the set of $\mathbb{P}$-compatible probability measures on $(\Omega\times X, \mathcal{F}\otimes\Sigma)$.

Likewise, given any sub-σ-algebra $\mathcal{G}$ of $\mathcal{F}$, we say that a probability measure μ on $(\Omega\times X, \mathcal{G}\otimes\Sigma)$ is $\mathbb{P}|_\mathcal{G}$-*compatible* if $\mu(E\times X) = \mathbb{P}(E)$ for all $E\in\mathcal{G}$. We write $\mathcal{M}^{\mathbb{P}|_\mathcal{G}}$ for the set of $\mathbb{P}|_\mathcal{G}$-compatible probability measures on $(\Omega\times X, \mathcal{G}\otimes\Sigma)$.

We will soon see that "random probability measures (up to equivalence) are in one-to-one correspondence with compatible probability measures".

Disintegrations

Given any random probability measure (μ_ω) on X, we may define a $\mathbb{P}$-compatible probability measure μ on $\Omega\times X$ by

$$\mu(A) \;:=\; \int_\Omega \mu_\omega(A_\omega)\,\mathbb{P}(d\omega)$$

where, for any $A\in\mathcal{F}\otimes\Sigma$ and any $\omega\in\Omega$, A_ω denotes the ω-section of A, that is

$$A_\omega := \{x\in X : (\omega,x)\in A\}.$$

(Using the monotone convergence theorem, it is easy to check that μ is indeed a probability measure.)

It is easy to show (using Corollary A.7) that for any bounded measurable $g\colon\Omega\times X\to\mathbb{R}$,

$$\int_{\Omega\times X} g(\omega,x)\,\mu(d(\omega,x)) \;=\; \int_\Omega\int_X g(\omega,x)\,\mu_\omega(dx)\,\mathbb{P}(d\omega).$$

We refer to μ as the *integrated form of* (μ_ω). Note that $\pi_{X*}\mu = \mathbb{E}_\omega\mu_\omega$.

By Corollary A.6, to show that a probability measure μ on $\Omega\times X$ is the integrated form of a random probability measure (μ_ω), it is sufficient to show that for every $E \in \mathcal{F}$ and $A \in \Sigma$,

$$\mu(E \times A) = \int_E \mu_\omega(A)\,\mathbb{P}(d\omega).$$

Now given a $\mathbb{P}$-compatible probability measure μ on $\Omega \times X$, we refer to any random probability measure (μ_ω) whose integrated form is equal to μ as a (*version of the*) *disintegration of* μ (*with respect to* $\mathbb{P}$). And more generally: let $\mathcal{G}$ be any sub-σ-algebra of $\mathcal{F}$; then, given a $\mathbb{P}|_{\mathcal{G}}$-compatible probability measure ν on $(\Omega \times X, \mathcal{G} \otimes \Sigma)$, a (*version of the*) *disintegration of* ν *with respect to* $\mathbb{P}|_{\mathcal{G}}$ is a $\mathcal{G}$-measurable random probability measure (μ_ω) whose integrated form agrees with ν on $\mathcal{G} \otimes \Sigma$.

It is clear that if two random probability measures (μ_ω) and $(\tilde{\mu}_\omega)$ are equivalent then they share the same integrated form. We now give the "disintegration theorem", which essentially states that *every* $\mathbb{P}$*-compatible probability measure admits a disintegration, and this disintegration is unique up to equivalence* (and more generally: every $\mathbb{P}|_{\mathcal{G}}$-compatible probability measure admits a disintegration with respect to $\mathbb{P}|_{\mathcal{G}}$, and this disintegration is unique up to equivalence).

Lemma 3.27 (Disintegration Theorem)**.** *Fix any sub-σ-algebra $\mathcal{G}$ of $\mathcal{F}$. For any $(\mu_\omega)_{\omega\in\Omega} \in \mathcal{L}^0(\Omega, \mathcal{G}; \mathcal{M})$, let $[(\mu_\omega)_{\omega\in\Omega}] \in L^0(\mathbb{P}|_{\mathcal{G}}; \mathcal{M})$ denote the equivalence class of $\mathcal{G}$-measurable random probability measures represented by $(\mu_\omega)_{\omega\in\Omega}$, and let μ denote the integrated form of $(\mu_\omega)_{\omega\in\Omega}$. Then the map*

$$\begin{aligned} L^0(\mathbb{P}|_{\mathcal{G}}; \mathcal{M}) &\to \mathcal{M}^{\mathbb{P}|_{\mathcal{G}}} \\ [(\mu_\omega)_{\omega\in\Omega}] &\mapsto \mu|_{\mathcal{G}\otimes\Sigma} \end{aligned}$$

serves as a bijection between $L^0(\mathbb{P}|_{\mathcal{G}}; \mathcal{M})$ and $\mathcal{M}^{\mathbb{P}|_{\mathcal{G}}}$.

Remark 3.28. Really, it suffices just to state the case that $\mathcal{G} = \mathcal{F}$, namely, to state that the map

$$\begin{aligned} L^0(\mathbb{P}; \mathcal{M}) &\to \mathcal{M}^{\mathbb{P}} \\ [(\mu_\omega)_{\omega\in\Omega}] &\mapsto \mu \end{aligned}$$

serves as a bijection between $L^0(\mathbb{P}; \mathcal{M})$ and $\mathcal{M}^{\mathbb{P}}$; the case of a more general sub-σ-algebra $\mathcal{G}$ then follows by redefining the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ to be $(\Omega, \mathcal{G}, \mathbb{P}|_{\mathcal{G}})$.

We now give a proof of Lemma 3.27, adapted from [Bil95, Theorem 33.3] (which specifically considers conditional distributions of random variables).

Proof of Lemma 3.27. As in Remark 3.28, we assume without loss of generality that $\mathcal{G} = \mathcal{F}$.

Surjectivity: Fix $\mu \in \mathcal{M}^{\mathbb{P}}$. First suppose that X is finite or countable. For each $x \in X$, let p_x be the finite measure on Ω given by $p_x(E) = \mu(E \times \{x\})$ for all $E \in \mathcal{F}$; since μ is $\mathbb{P}$-compatible, we have that for every $\mathbb{P}$-null set $E \in \mathcal{F}$, $p_x(E) \leq \mu(E \times \{x\}) = 0$. So p_x is absolutely continuous with respect to $\mathbb{P}$ for all $x \in X$. So for each $x \in X$, let $h_x : \Omega \to [0,1]$ be a version of the density of p_x with respect to $\mathbb{P}$. Now it is clear that

$$\sum_{x\in X} p_x = \mathbb{P},$$

and therefore

$$\sum_{x\in X} h_x \overset{\mathbb{P}\text{-a.s.}}{=} 1.$$

So let let $\tilde{\Omega} \in \mathcal{F}$ be a $\mathbb{P}$-full set such that for all $\omega \in \tilde{\Omega}$, $\sum_{x\in X} h_x(\omega) = 1$. For each $\omega \in \tilde{\Omega}$, let μ_ω be the probability measure on X given by $\mu_\omega(\{x\}) = h_x(\omega)$ for all $x \in X$; and fixing an arbitrary probability measure c on X, let $\mu_\omega := c$ for all $\omega \in \Omega \setminus \tilde{\Omega}$. It is clear that for every $A \in \Sigma$, the map $\omega \mapsto \mu_\omega(A)$ is measurable; so $(\mu_\omega)_{\omega\in\Omega} \in \mathcal{L}^0(\Omega,\mathcal{F};\mathcal{M})$. Now for each $E \in \mathcal{F}$ and $A \subset X$, we have

$$\int_E \mu_\omega(A)\,\mathbb{P}(d\omega) \;=\; \sum_{x\in A}\int_E h_\omega(x)\,\mathbb{P}(d\omega) \;=\; \sum_{x\in A} p_x(E) \;=\; \mu(E\times A).$$

So μ is the integrated form of (μ_ω), as required.

Now suppose that X is uncountable. By the Borel isomorphism theorem, we may assume without loss of generality that $(X,\Sigma) = ([0,1],\mathcal{B}([0,1]))$. For each $a \in [0,1]\cap\mathbb{Q}$, let P_a be the finite measure on Ω given by $P_a(E) = \mu(E\times[0,a])$ for all $E \in \mathcal{F}$; once again, it is clear that P_a is absolutely continuous with respect to $\mathbb{P}$ for all a. So for each $a \in [0,1)\cap\mathbb{Q}$, let $H_a\colon\Omega \to [0,1]$ be a version of the density of P_a with respect to $\mathbb{P}$; and set $H_1(\omega) := 1$ for all $\omega \in \Omega$. For each $a,b \in [0,1]\cap\mathbb{Q}$ with $a \le b$, there exists a $\mathbb{P}$-full set $\Omega_{a,b} \in \mathcal{F}$ such that $H_a(\omega) \le H_b(\omega)$ for all $\omega \in \Omega_{a,b}$. So let

$$\Omega' \;:=\; \bigcap_{\substack{a,b\in[0,1]\cap\mathbb{Q}\\ a\le b}} \Omega_{a,b}.$$

By construction, for all $\omega \in \Omega'$ the map $a \mapsto H_a(\omega)$ from $[0,1]\cap\mathbb{Q}$ to $[0,1]$ is increasing. Now for each $a \in [0,1)\cap\mathbb{Q}$ and $\omega \in \Omega'$, let $H_{a+}(\omega) := \inf\{H_b(\omega) : b \in (a,1]\cap\mathbb{Q}\}$; obviously $P_b(\Omega')$ decreases to $P_a(\Omega')$ as b decreases to a, and so by the monotone convergence theorem we have that

$$\int_{\Omega'} H_{a+}(\omega)\,\mathbb{P}(d\omega) \;=\; P_a(\Omega') \;=\; \int_{\Omega'} H_a(\omega)\,\mathbb{P}(d\omega),$$

implying in particular that there is a $\mathbb{P}$-full subset $\Omega'_a \in \mathcal{F}$ of Ω' such that $H_{a+} = H_a$ on Ω'_a. So let

$$\tilde{\Omega} \;:=\; \bigcap_{a\in[0,1)\cap\mathbb{Q}} \Omega'_a.$$

For each $\omega \in \tilde{\Omega}$ and $x \in [0,1]\setminus\mathbb{Q}$, let $H_x(\omega) := \inf\{H_a(\omega) : a \in (x,1]\cap\mathbb{Q}\}$. It is clear that the map $x \mapsto H_x(\omega)$ from $[0,1]$ to $[0,1]$ is increasing and right-continuous for all $\omega \in \tilde{\Omega}$. So for each $\omega \in \tilde{\Omega}$, let μ_ω be the probability measure on $[0,1]$ given by $\mu_\omega([0,x]) = H_x(\omega)$ for all $x \in [0,1]$; and fixing an arbitrary probability measure c on $[0,1]$, let $\mu_\omega := c$ for all $\omega \in \Omega\setminus\tilde{\Omega}$. For each $E \in \mathcal{F}$, let $\mathcal{D}_E$ be the collection of all Borel subsets A of $[0,1]$ with the properties that the map $\omega \mapsto \mu_\omega(A)$ is measurable and $\int_E \mu_\omega(A)\,\mathbb{P}(d\omega) = \mu(E\times A)$. By construction, $\mathcal{D}_E$ contains $[0,a]$ for every $a \in [0,1]\cap\mathbb{Q}$; moreover, by the monotone convergence theorem, $\mathcal{D}_E$ is a λ-system on $[0,1]$. Hence, by the π-λ theorem (Lemma A.5), $\mathcal{D}_E$ is equal to the whole of $\mathcal{B}([0,1])$. This is true for every $E \in \mathcal{F}$, and therefore μ is the integrated form of (μ_ω).

Injectivity: Let (μ_ω) and (μ'_ω) be random probability measures sharing the same integrated form μ. For each $A \in \Sigma$, we have that

$$\int_E \mu_\omega(A)\,\mathbb{P}(d\omega) \;=\; \int_E \mu'_\omega(A)\,\mathbb{P}(d\omega) \quad \forall\, E \in \mathcal{F}$$

and therefore $\mu_\omega(A) = \mu'_\omega(A)$ for $\mathbb{P}$-almost all $\omega \in \Omega$. Now let $\mathcal{C}$ be a countable π-system generating Σ. Then $\mathbb{P}$-almost every $\omega \in \Omega$ has the property that for all $A \in \mathcal{C}$, $\mu_\omega(A) = \mu'_\omega(A)$; and hence, by Corollary A.6, $\mu_\omega = \mu'_\omega$ for $\mathbb{P}$-almost all $\omega \in \Omega$. □

We have mentioned that taking the "expectation" of a random probability measure can be achieved by taking the X-projection of the integrated form of the random probability measure. We will now see that (as in [Arn98, p23]) taking the "conditional expectation given $\mathcal{G}$" of a random probability measure can be achieved by taking a $\mathbb{P}|_{\mathcal{G}}$-disintegration of the $(\mathcal{G} \otimes \Sigma)$-restriction of the integrated form of the random probability measure:

Lemma 3.29. *Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. Let (μ_ω) and (μ'_ω) be random probability measures on X, with (μ'_ω) being $\mathcal{G}$-measurable. Suppose that the integrated forms of (μ_ω) and (μ'_ω) agree on $\mathcal{G} \otimes \Sigma$; then for any bounded measurable $g\colon X \to \mathbb{R}$,*

$$\mathbb{E}[\,\tilde{\omega} \mapsto \mu_{\tilde{\omega}}(g)\,|\,\mathcal{G}\,](\omega) \;\overset{\mathbb{P}\text{-a.s.}}{=}\; \mu'_\omega(g).$$

Proof. Fix any $E \in \mathcal{G}$. Writing μ and μ' for the integrated forms of (μ_ω) and (μ'_ω) respectively, we have that

$$\int_E \mu_\omega(g)\,\mathbb{P}(d\omega) \;=\; \int_{\Omega\times E} g(x)\,\mu(d(\omega,x)) \;=\; \int_{\Omega\times E} g(x)\,\mu'(d(\omega,x)) \;=\; \int_E \mu'_\omega(g)\,\mathbb{P}(d\omega)$$

as required. □

We now show that the converse of Lemma 3.29 holds; in fact, we will prove a slightly stronger version of the converse. Let us say that a collection $\{g_\alpha\}_{\alpha\in I}$ of bounded measurable functions $g_\alpha\colon X \to \mathbb{R}$ is *measure-determining* if for any probability measures ρ_1 and ρ_2 on X,

$$\rho_1(g_\alpha) = \rho_2(g_\alpha)\ \forall\, \alpha \in I \quad\Longrightarrow\quad \rho_1 = \rho_2.$$

Note that a countable measure-determining set does exist: take $\{\mathbb{1}_A\}_{A\in\mathcal{C}}$ for some countable π-system $\mathcal{C}$ generating Σ.

Lemma 3.30. *Let $\{g_i\}_{i\in\mathbb{N}}$ be a countable measure-determining set of bounded measurable functions $g_i\colon X \to \mathbb{R}$, and let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. Let (μ_ω) and (μ'_ω) be random probability measures on X, and suppose that for each $i \in \mathbb{N}$,*

$$\mathbb{E}[\,\tilde{\omega} \mapsto \mu_{\tilde{\omega}}(g_i)\,|\,\mathcal{G}\,](\omega) \;\overset{\mathbb{P}\text{-a.s.}}{=}\; \mu'_\omega(g_i).$$

Then the integrated forms of (μ_ω) and (μ'_ω) agree on $\mathcal{G} \otimes \Sigma$.

Proof. On the basis of Lemma 3.27, let $(\hat{\mu}_\omega)$ be a $\mathcal{G}$-measurable random probability measure whose integrated form agrees with the integrated form of (μ_ω) on $\mathcal{G} \otimes \Sigma$. By Lemma 3.29, we have that for each $i \in \mathbb{N}$, for $\mathbb{P}$-almost all $\omega \in \Omega$,

$$\hat{\mu}_\omega(g_i) \;=\; \mathbb{E}[\,\tilde{\omega} \mapsto \mu_{\tilde{\omega}}(g_i)\,|\,\mathcal{G}\,](\omega) \;=\; \mu'_\omega(g_i).$$

So let $\tilde{\Omega} \subset \Omega$ be a $\mathbb{P}$-full set such that $\hat{\mu}_\omega(g_i) = \mu'_\omega(g_i)$ for all $\omega \in \tilde{\Omega}$ and $i \in \mathbb{N}$. Then, since $\{g_i\}_{i\in\mathbb{N}}$ is measure-determining, $\hat{\mu}_\omega = \mu'_\omega$ for all $\omega \in \tilde{\Omega}$. Thus $(\hat{\mu}_\omega)$ and (μ'_ω) have the same integrated form, and therefore in particular, the integrated form of (μ'_ω) agrees with the integrated form of (μ_ω) on $\mathcal{G} \otimes \Sigma$. □

Definition 3.31. Let $\mathcal{G}$ be a sub-σ-algebra of $\mathcal{F}$. We will say that a $\mathbb{P}$-compatible probability measure μ on $(\Omega \times X, \mathcal{F} \otimes \Sigma)$ is *$\mathcal{G}$-measurable* if there exists a disintegration of μ with respect to $\mathbb{P}$ which is $\mathcal{G}$-measurable.[9]

Lemma 3.32. *Let $\mathcal{G}_1$ and $\mathcal{G}_2$ be independent sub-σ-algebras of $\mathcal{F}$, and let μ be a $\mathbb{P}$-compatible probability measure that is $\mathcal{G}_1$-measurable. Then*

$$\mu|_{\mathcal{G}_2 \otimes \Sigma} = \mathbb{P}|_{\mathcal{G}_2} \otimes \pi_{X*}\mu .$$

Proof. Let (μ_ω) be a $\mathcal{G}_1$-measurable disintegration of μ. For any $E \in \mathcal{G}_2$ and $A \in \Sigma$, we have

$$\begin{aligned} \mu(E \times A) &= \int_\Omega \mathbb{1}_E(\omega)\mu_\omega(A)\,\mathbb{P}(d\omega) \\ &= \int_\Omega \mathbb{1}_E(\tilde{\omega})\,\mathbb{P}(d\tilde{\omega}) \int_\Omega \mu_\omega(A)\,\mathbb{P}(d\omega) \quad \text{(by Lemma A.10)} \\ &= \mathbb{P}(E)\pi_{X*}\mu(A) \end{aligned}$$

as required. □

Measure-valued martingales

As in Remark A.19, given a separable metrisable topology $\mathcal{T}$ on X generating Σ, a countable set $\{g_i\}_{i\in\mathbb{N}}$ of bounded continuous functions $g_i : X \to \mathbb{R}$ is said to be *convergence-determining* (according to $\mathcal{T}$) if for any sequence (ρ_n) in $\mathcal{M}$ and any $\rho \in \mathcal{M}$,

$$\rho_n(g_i) \to \rho(g_i) \text{ as } n \to \infty \text{ for every } i \in \mathbb{N} \quad \Longrightarrow \quad \rho_n \to \rho \text{ in } \mathcal{N}_\mathcal{T} \text{ as } n \to \infty.$$

As in Theorem A.16, such a set of functions $\{g_i\}_{i\in\mathbb{N}}$ does exist. Obviously any convergence-determining set of functions is also measure-determining.

The following result can be regarded as "Lévy's upward theorem for measures".

Theorem 3.33. *Let $(\mathcal{G}_t)_{t\in\mathbb{T}^+}$ be a filtration on $(\Omega, \mathcal{F})$, and write $\mathcal{G}_\infty := \sigma(\mathcal{G}_t : t \in \mathbb{T}^+)$. Let $(\mu_\omega)_{\omega\in\Omega}$ be a $\mathcal{G}_\infty$-measurable random probability measure on X, and let $(\mu^t_\omega)_{t\in\mathbb{T}^+,\omega\in\Omega}$ be an $\mathcal{M}$-valued stochastic process such that for each $t \in \mathbb{T}^+$, the random probability measure $(\mu^t_\omega)_{\omega\in\Omega}$ is $\mathcal{G}_t$-measurable and the integrated form of $(\mu^t_\omega)_{\omega\in\Omega}$ agrees with the integrated form of $(\mu_\omega)_{\omega\in\Omega}$ on $\mathcal{G}_t \otimes \Sigma$. Then for every separable metrisable topology $\mathcal{T}$ on X generating Σ, $(\mu^t_\omega)_{t\in\mathbb{T}^+,\omega\in\Omega}$ converges via countable subnets to $(\mu_\omega)_{\omega\in\Omega}$ in the narrow topology $\mathcal{N}_\mathcal{T}$.*

Remark 3.34. We emphasise that for each unbounded countable $S \subset \mathbb{T}^+$, the exceptional $\mathbb{P}$-null set on which the convergence fails will generally depend on the topology $\mathcal{T}$. (Indeed, for any unbounded countable $S \subset \mathbb{T}^+$, if the exceptional null set can be chosen independently of $\mathcal{T}$, then outside this exceptional null set we will have *strong* convergence as $t \to \infty$ in S, by Lemma 2.3.)

[9] If $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space and $\mathcal{G}$ contains all $\mathbb{P}$-null sets, then this is equivalent to saying that every disintegration of μ with respect to $\mathbb{P}$ is $\mathcal{G}$-measurable.

Proof of Theorem 3.33. Fix a separable metrisable topology $\mathcal{T}$ on X generating Σ. Let $\{g_i\}_{i\in\mathbb{N}}$ be a convergence-determining set of bounded continuous functions $g_i\colon X\to\mathbb{R}$. For each $i\in\mathbb{N}$ and $t\in\mathbb{T}^+$, Lemma 3.29 gives that

$$\mathbb{E}[\,\tilde{\omega}\mapsto\mu_{\tilde{\omega}}(g_i)\,|\,\mathcal{G}_t\,](\omega)\ \overset{\mathbb{P}\text{-a.s.}}{=}\ \mu^t_\omega(g_i).$$

So then, Lévy's upward theorem[10] gives that for each i, the stochastic process $(\mu^t_\omega(g_i))_{t\in\mathbb{T}^+,\omega\in\Omega}$ is a modification of a stochastic process converging almost surely to the random variable $(\mu_\omega(g_i))_{\omega\in\Omega}$. Hence in particular, $(\mu^t_\omega(g_i))_{t\in\mathbb{T}^+,\omega\in\Omega}$ converges via countable subnets to $(\mu_\omega(g_i))_{\omega\in\Omega}$. So for each $i\in\mathbb{N}$ and each unbounded countable $S\subset\mathbb{T}^+$, let $\Omega^{(i)}_S\subset\Omega$ be a $\mathbb{P}$-full set such that for each $\omega\in\Omega^{(i)}_S$, $\mu^t_\omega(g_i)\to\mu_\omega(g_i)$ as $t\to\infty$ in S. Then, defining $\Omega_S:=\bigcap_{i=1}^\infty\Omega^{(i)}_S$ for any unbounded countable $S\subset\mathbb{T}^+$, we have that for each $\omega\in\Omega_S$, $\mu^t_\omega\to\mu_\omega$ in $\mathcal{N}_\mathcal{T}$ as $t\to\infty$ in S. □

The following "extension theorem" is based on the martingale convergence theorem.

Theorem 3.35. *Let $(\mathcal{G}_t)_{t\in\mathbb{T}^+}$ be a filtration on $(\Omega,\mathcal{F})$, and write $\mathcal{G}_\infty:=\sigma(\mathcal{G}_t:t\in\mathbb{T}^+)$. Let*

$$\dot{\mu}:\bigcup_{t\in\mathbb{T}^+}(\mathcal{G}_t\otimes\Sigma)\ \to\ [0,1]$$

be a function with the property that for each $t\in\mathbb{T}^+$, $\dot{\mu}|_{\mathcal{G}_t\otimes\Sigma}$ is a $\mathbb{P}|_{\mathcal{G}_t}$-compatible probability measure. Then there exists a unique probability measure μ on $(\Omega\times X,\mathcal{G}_\infty\otimes\Sigma)$ agreeing with $\dot{\mu}$ on $\bigcup_{t\in\mathbb{T}^+}(\mathcal{G}_t\otimes\Sigma)$; the measure μ is itself $\mathbb{P}|_{\mathcal{G}_\infty}$-compatible.

Proof. Existence: For each $n\in\mathbb{N}$, let $(\mu^n_\omega)_{\omega\in\Omega}$ be a disintegration of $\dot{\mu}|_{\mathcal{G}_n\otimes\Sigma}$ with respect to $\mathbb{P}|_{\mathcal{G}_n}$. Fix a compact metrisable topology $\mathcal{T}$ on X generating Σ, and let $\{g_i\}_{i\in\mathbb{N}}$ be a convergence-determining set of continuous functions $g_i\colon X\to[0,1]$. For each $i,n\in\mathbb{N}$, Lemma 3.29 gives that

$$\mathbb{E}[\,\tilde{\omega}\mapsto\mu^{n+1}_{\tilde{\omega}}(g_i)\,|\,\mathcal{G}_{n+1}\,](\omega)\ \overset{\mathbb{P}\text{-a.s.}}{=}\ \mu^n_\omega(g_i).$$

Thus, for each i, the stochastic process $(\mu^n_\omega(g_i))_{n\in\mathbb{N},\omega\in\Omega}$ is a (uniformly bounded) martingale, and therefore converges almost surely. So let $\tilde{\Omega}\in\mathcal{F}$ be a $\mathbb{P}$-full set such that for every $\omega\in\tilde{\Omega}$ and $i\in\mathbb{N}$, the sequence $(\mu^n_\omega(g_i))_{n\in\mathbb{N}}$ is convergent. By Corollary A.18, for each $\omega\in\tilde{\Omega}$, μ^n_ω converges in the narrow topology some probability measure μ_ω as $n\to\infty$; fixing an arbitrary probability measure c on X, we can then define $\mu_\omega:=c$ for all $\omega\in\Omega\setminus\tilde{\Omega}$. Let $\bar{\mu}$ be the integrated form of the random probability measure $(\mu_\omega)_{\omega\in\Omega}$. Fix any $n\in\mathbb{N}$; for each $i\in\mathbb{N}$, the conditional dominated convergence theorem gives that

$$\begin{aligned}\mathbb{E}[\,\tilde{\omega}\mapsto\mu_{\tilde{\omega}}(g_i)\,|\,\mathcal{G}_n\,](\omega)\ &\overset{\mathbb{P}\text{-a.s.}}{=}\ \lim_{m\to\infty}\mathbb{E}[\,\tilde{\omega}\mapsto\mu^m_{\tilde{\omega}}(g_i)\,|\,\mathcal{G}_n\,](\omega)\\ &\overset{\mathbb{P}\text{-a.s.}}{=}\ \lim_{m\to\infty}\mu^n_\omega(g_i)\quad\text{(by Lemma 3.29)}\\ &=\ \mu^n_\omega(g_i).\end{aligned}$$

[10][Nev65, Proposition IV.5.6] gives a combined statement of the martingale convergence theorem and Lévy's upward theorem for separable stochastic processes; by [Nev65, Proposition III.4.3], every $\bar{\mathbb{R}}$-valued stochastic process has a separable modification.

Hence, by Lemma 3.30, $\dot{\mu}|_{\mathcal{G}_n\otimes\Sigma} = \bar{\mu}|_{\mathcal{G}_n\otimes\Sigma}$. This is true for every n, so the measure $\mu := \bar{\mu}|_{\mathcal{G}_\infty\otimes\Sigma}$ fulfils the properties described in the statement of the theorem.

Uniqueness: It is easy to see that $\mathcal{G}_\infty \otimes \Sigma$ is precisely the σ-algebra on $\Omega \times X$ generated by $\bigcup_{t\in\mathbb{T}^+}(\mathcal{G}_t \otimes \Sigma)$ (which is itself a π-system, as in Remark A.3). Hence Corollary A.6 gives the uniqueness of the measure μ. □

3.4 Invariant measures

Recall that for each $t \in \mathbb{T}^+$, we define the map $\Theta^t : \Omega \times X \to \Omega \times X$ by $\Theta^t(\omega, x) = (\theta^t\omega, \varphi(t,\omega)x)$.

Lemma 3.36 ([Arn98, Lemma 1.4.4]). *Let μ be a $\mathbb{P}$-compatible probability measure on $\Omega \times X$, with disintegration $(\mu_\omega)_{\omega\in\Omega}$. Then for any $t \in \mathbb{T}^+$, $\Theta^t_*\mu$ is $\mathbb{P}$-compatible, with disintegration $(\varphi(t,\theta^{-t}\omega)_*\mu_{\theta^{-t}\omega})_{\omega\in\Omega}$.*

Proof. Fix $t \in \mathbb{T}^+$. Let ν be the integrated form of $(\varphi(t,\theta^{-t}\omega)_*\mu_{\theta^{-t}\omega})_{\omega\in\Omega}$. For any $E \in \mathcal{F}$ and $A \in \Sigma$, we have

$$\begin{aligned}
\Theta^t_*\mu(E \times A) &= \int_\Omega \int_X \mathbb{1}_{E\times A}(\theta^t\omega, \varphi(t,\omega)x)\, \mu(d(\omega,x)) \\
&= \int_\Omega \int_X \mathbb{1}_E(\theta^t\omega)\ \mathbb{1}_A(\varphi(t,\omega)x)\ \mu_\omega(dx)\, \mathbb{P}(d\omega) \\
&= \int_\Omega \int_X \mathbb{1}_E(\omega)\ \mathbb{1}_A(\varphi(t,\theta^{-t}\omega)x)\ \mu_{\theta^{-t}\omega}(dx)\, \mathbb{P}(d\omega) \\
&= \int_E \varphi(t,\theta^{-t}\omega)_*\mu_{\theta^{-t}\omega}(A)\, \mathbb{P}(d\omega) \\
&= \nu(E \times A).
\end{aligned}$$

Hence $\Theta^t_*\mu = \nu$. □

As a particular case of this, we will now prove Remark 2.22.

Corollary 3.37. *Let ρ be a probability measure on X. The following are equivalent:*

- *(i)* *$(\Omega \times X, \mathcal{F} \otimes \Sigma, \mathbb{P} \otimes \rho, (\Theta^t)_{t\in\mathbb{T}^+})$ is a measure-preserving dynamical system;*
- *(ii)* *there exists $r \in \mathbb{T}^+ \setminus \{0\}$ such that $(\Omega \times X, \mathcal{F}^\infty_{-r} \otimes \Sigma, \mathbb{P}|_{\mathcal{F}^\infty_{-r}} \otimes \rho, (\Theta^t)_{t\in\mathbb{T}^+})$ is a measure-preserving dynamical system;*
- *(iii)* *ρ is crudely incompressible.*

Proof. It is clear that (i)⇒(ii). Now suppose that (ii) holds. Fix any $t \in \mathbb{T}^+$ with $0 \le t \le r$. Given any $A \in \Sigma$, Lemma 3.36 gives that for all $E \in \mathcal{F}^\infty_{-t}$,

$$\int_E \varphi(t,\theta^{-t}\omega)_*\rho(A)\, \mathbb{P}(d\omega) = \Theta^t_*(\mathbb{P} \otimes \rho)(E \times A) = \mathbb{P} \otimes \rho(E \times A) = \mathbb{P}(E)\rho(A);$$

since this is true for every $E \in \mathcal{F}^\infty_{-t}$ and the map $\omega \mapsto \varphi(t,\theta^{-t}\omega)_*\rho(A)$ is itself $\mathcal{F}^\infty_{-t}$-measurable, it follows that $\varphi(t,\theta^{-t}\omega)_*\rho(A) = \rho(A)$ for $\mathbb{P}$-almost all $\omega \in \Omega$. This is true for each $A \in \Sigma$; so let $\mathcal{C}$ be a countable π-system generating Σ, and let $\tilde{\Omega} \subset \Omega$ be a $\mathbb{P}$-full set such that for every $\omega \in \tilde{\Omega}$ and $A \in \mathcal{C}$, $\varphi(t, \theta^{-t}\omega)_*\rho(A) = \rho(A)$. Then (by Lemma A.6) $\varphi(t, \theta^{-t}\omega)_*\rho = \rho$ for every $\omega \in \tilde{\Omega}$. Thus we have shown that for every $t \in \mathbb{T}^+$ with $0 \leq t \leq r$, $\varphi(t, \theta^{-t}\omega)_*\rho$ is $\mathbb{P}$-almost surely equal to ρ. Now consider $t \in \mathbb{T}^+$ with $t > r$, and let $n := \lfloor \frac{t}{r} \rfloor$. For each $i \in \{1, \dots, n\}$, we have that $\varphi(r, \theta^{-ir}\omega)_*\rho = \rho$ for $\mathbb{P}$-almost all $\omega \in \Omega$; hence $\varphi(nr, \theta^{-nr}\omega)_*\rho = \rho$ for $\mathbb{P}$-almost all $\omega \in \Omega$. But we also have that $\varphi(t-nr, \theta^{-t}\omega)_*\rho = \rho$ for $\mathbb{P}$-almost all $\omega \in \Omega$. So then, $\varphi(t, \theta^{-t}\omega)_*\rho = \varphi(nr, \theta^{-nr}\omega)_*\varphi(t - nr, \theta^{-t}\omega)_*\rho = \rho$ for $\mathbb{P}$-almost all $\omega \in \Omega$. Thus φ is crudely incompressible.

Finally, the fact that (iii)⇒(i) follows immediately from Lemma 3.36, with $\mu := \mathbb{P} \otimes \rho$. □

Definition 3.38. A probability measure μ on $(\Omega \times X, \mathcal{F} \otimes \Sigma)$ is said to be *φ-invariant* if μ is both $\mathbb{P}$-compatible and invariant under the dynamical system $(\Theta^t)_{t \in \mathbb{T}^+}$; and μ is said to be *φ-ergodic* if μ is both $\mathbb{P}$-compatible and ergodic with respect to the dynamical system $(\Theta^t)_{t \in \mathbb{T}^+}$.

Remark 3.39. Recall that an invariant probability measure of a dynamical system is ergodic if and only if it is an extreme point of the convex set of invariant measures of that dynamical system. Now observe that for any $(\Theta^t)_{t \in \mathbb{T}^+}$-invariant measure μ, $\pi_{\Omega *}\mu$ is (θ^t)-invariant. Consequently, since $\mathbb{P}$ is (θ^t)-ergodic, it is not hard to show (as in [Cra02b, Lemma 6.19(i)]) that a φ-invariant probability measure is φ-ergodic if and only if it is an extreme point of the convex set of φ-invariant probability measures.

Definition 3.40. A random probability measure (μ_ω) is said to be *φ-invariant* (resp. *φ-ergodic*) if its integrated form is φ-invariant (resp. φ-ergodic).

Using Lemma 3.36, we can now characterise φ-invariant probability measures in terms of their disintegrations.

Lemma 3.41. *For any random probability measure (μ_ω) on X, the following are equivalent:*

(i) (μ_ω) is φ-invariant;

(ii) for each $t \in \mathbb{T}^+$, for $\mathbb{P}$-almost all $\omega \in \Omega$, $\mu_\omega = \varphi(t, \theta^{-t}\omega)_*\mu_{\theta^{-t}\omega}$;

(iii) for each $t \in \mathbb{T}^+$, for $\mathbb{P}$-almost all $\omega \in \Omega$, $\mu_{\theta^t\omega} = \varphi(t, \omega)_*\mu_\omega$;

(iv) the map $\omega \mapsto \mu_\omega$ is a random fixed point of the image-measure RDS φ_*.

Proof. The equivalence of (i) and (ii) follows immediately from Lemma 3.36. The equivalence of (ii) and (iii) is due to the (θ^t)-invariance of $\mathbb{P}$. The equivalence of (iii) and (iv) is automatic from the definitions. □

Remark 3.42. Observe that for any measurable function $q: \Omega \to X$, the random probability measure $(\delta_{q(\omega)})$ is φ-invariant if and only if q is a random fixed point of φ. (In fact, it is not hard to show that if q is a random fixed point then $(\delta_{q(\omega)})$ is φ-ergodic.)

The following is adapted from part (a) of the proof of [LeJ87, Proposition 2].

Lemma 3.43. *Let (μ_ω) be a φ-ergodic random probability measure. Then there exists $n \in \mathbb{N} \cup \{\infty\}$ such that for $\mathbb{P}$-almost all $\omega \in \Omega$, $\mu_\omega \in \mathcal{K}_n$.*

To prove Lemma 3.43, we use the following simple observation:

Lemma 3.44. *Suppose we have a value $c \in [0,1]$ and a probability measure ρ on X, such that for ρ-almost every $x \in X$, $\rho(\{x\}) = c$. Then either $c = 0$ and $\rho \in \mathcal{K}_\infty$, or $c = \frac{1}{n}$ and $\rho \in \mathcal{K}_n$ for some $n \in \mathbb{N}$.*

Proof of Lemma 3.44. If $c = 0$ then there clearly cannot exist a singleton of strictly positive measure under ρ; so $\rho \in \mathcal{K}_\infty$. If $c > 0$, then the finite set P of points x satisfying $\rho(\{x\}) = c$ is a ρ-full measure set; so letting $n := |P|$, it is clear that $c = \frac{1}{n}$ and $\rho \in \mathcal{K}_n$. □

Proof of Lemma 3.43. Define the function $h\colon \Omega \times X \to [0,1]$ by $h(\omega,x) = \mu_\omega(\{x\})$. Note that h is measurable, since it can be expressed as

$$h(\omega,x) \;=\; \int_X \mathbb{1}_{\Delta_X}(x,y)\,\mu_\omega(dy).$$

Now for each $t \in \mathbb{T}^+$, let $\Omega_t \subset \Omega$ be a $\mathbb{P}$-full set such that for each $\omega \in \Omega_t$, $\varphi(t,\omega)_*\mu_\omega = \mu_{\theta^t\omega}$. Then for all $(\omega,x) \in \Omega_t \times X$, we have

$$\begin{aligned} h(\Theta^t(\omega,x)) &= \mu_{\theta^t(\omega)}(\,\{\varphi(t,\omega)x\}\,) \\ &= \mu_\omega\left(\,\varphi(t,\omega)^{-1}(\,\{\varphi(t,\omega)x\}\,)\,\right) \\ &\geq \mu_\omega(\{x\}) \\ &= h(\omega,x). \end{aligned}$$

So then, letting μ be the integrated form of (μ_ω), we have that $h \circ \Theta^t \overset{\mu\text{-a.s.}}{\geq} h$ for each $t \in \mathbb{T}^+$. Hence, since μ is Θ-ergodic, there exists $c \in [0,1]$ such that $h^{-1}(\{c\})$ is a μ-full set. So for $\mathbb{P}$-almost every $\omega \in \Omega$, μ_ω has the property that $\mu_\omega(\{x\}) = c$ for μ_ω-almost all $x \in X$. The result then follows by Lemma 3.44. □

The following is an extension of [Arn98, Theorem 1.8.4(iv)].

Lemma 3.45. *Suppose there exists a Borel linear order $\leq$ on (X,Σ) with respect to which φ is monotone. Let (μ_ω) be φ-ergodic random probability measure. Then for $\mathbb{P}$-almost all $\omega \in \Omega$, μ_ω is a Dirac mass.*

We first prove the following:

Lemma 3.46. *Let $\leq$ be a Borel linear order on (X,Σ). Suppose we have a value $c \in [0,1]$ and a probability measure ρ on X, such that for ρ-almost every $x \in X$, $\rho(v \in X : v \leq x) = c$. Then ρ is a Dirac mass (and $c = 1$).*

Lemma 3.46 is fairly clear in the case $(X,\Sigma,\leq) = (\mathbb{R},\mathcal{B}(\mathbb{R}),\leq)$; nonetheless, we can prove it for more general Borel linear orders:

Proof of Lemma 3.46. We write "$x < y$" to mean "$x \leq y$ and $x \neq y$". For any $x \in X$, let $I_x := \{v \in X : v \leq x\}$ and let $I'_x := I_x \setminus \{x\} = \{v \in X : v < x\}$. For any $x \in X$ with $\rho(\{x\}) > 0$, we have that $\rho(I'_x) = 0$ (since for all $v \in I'_x$, $\rho(I_v)$ is strictly less than $\rho(I_x)$). So then, we have that either

(a) ρ is atomless; or

(b) there exists a unique $x^* \in X$ with $\rho(\{x^*\}) > 0$, in which case $\rho(I'_{x^*}) = 0$ and $\rho(\{x^*\}) = c$.

Now suppose for a contradiction that ρ is not a Dirac mass (in which case, in case (b), $c \neq 1$). Define a probability measure $\tilde{\rho}$ on X and a value $\tilde{c} \in [0,1]$ by

$$\tilde{\rho}(A) := \begin{cases} \rho(A) & \text{in case (a)} \\ \frac{\rho(A \setminus \{x^*\})}{1-c} & \text{in case (b)} \end{cases} \qquad \text{and} \qquad \tilde{c} = \begin{cases} c & \text{in case (a)} \\ 0 & \text{in case (b).} \end{cases}$$

We first show that $\tilde{\rho}$ is atomless and for $\tilde{\rho}$-almost all $x \in X$, $\tilde{\rho}(I'_x) = \tilde{c}$. In case (a), this is clear. In case (b), it is clear that $\tilde{\rho}$ is atomless. Also, in case (b), for any $x \in X$ with the property that $\rho(I_x) = c$, we have that $\tilde{\rho}(I'_x) = \tilde{\rho}(I_x) = 0$; by assumption, the set of points x that do *not* have this property is a ρ-null set, and is therefore clearly also a $\tilde{\rho}$-null set.

Now then, since $\tilde{\rho}$ is atomless, we have that $\tilde{\rho} \otimes \tilde{\rho}(\Delta_X) = 0$, and therefore

$$\begin{aligned} 1 &= \tilde{\rho} \otimes \tilde{\rho}(\,(x,y) : x \prec y\,) + \tilde{\rho} \otimes \tilde{\rho}(\,(x,y) : y \prec x\,) \\ &= 2\tilde{\rho} \otimes \tilde{\rho}(\,(x,y) : x \prec y\,) \\ &= 2 \int_X \int_X \chi_{x \prec y}\, \tilde{\rho}(dx)\, \tilde{\rho}(dy) \\ &= 2 \int_X \tilde{\rho}(I'_y)\, \tilde{\rho}(dy) \\ &= 2\tilde{c}. \end{aligned}$$

Hence $\tilde{c} = \frac{1}{2}$.

(Of course, this rules out case (b); but we shall soon rule out every situation.)

Now define the following 6 subsets of $X \times X \times X$:

$$\begin{aligned} J_1 &:= \{\,(x,y,z) : x \prec y \prec z\,\} \\ J_2 &:= \{\,(x,y,z) : x \prec z \prec y\,\} \\ J_3 &:= \{\,(x,y,z) : y \prec x \prec z\,\} \\ J_4 &:= \{\,(x,y,z) : y \prec z \prec x\,\} \\ J_5 &:= \{\,(x,y,z) : z \prec x \prec y\,\} \\ J_6 &:= \{\,(x,y,z) : z \prec y \prec x\,\}. \end{aligned}$$

Note that the sets $J_1, \ldots, J_6$ are mutually disjoint. Now J_1 is $(\Sigma \otimes \Sigma \otimes \Sigma)$-measurable, since, writing $I' := \{\,(x,y) : x \prec y\,\} \subset X \times X$, we can express J_1 as

$$J_1 = (I' \times X) \cap (X \times I').$$

Likewise the sets $J_2,\ldots,J_6$ are also $(\Sigma\otimes\Sigma\otimes\Sigma)$-measurable. So then,

$$\begin{aligned}
\tilde\rho\otimes\tilde\rho\otimes\tilde\rho\left(\bigcup_{i=1}^{6} J_i\right) &= 6\tilde\rho\otimes\tilde\rho\otimes\tilde\rho(J_1)\\
&= 6\int_X\int_X\int_X \chi_{x<y<z}\,\tilde\rho(dx)\,\tilde\rho(dy)\,\tilde\rho(dz)\\
&= 6\int_X\int_X \left(\chi_{y<z}\int_X \chi_{x<y}\,\tilde\rho(dx)\right)\tilde\rho(dy)\,\tilde\rho(dz)\\
&= 6\int_X\int_X \chi_{y<z}\,\tilde\rho(I'_y)\,\tilde\rho(dy)\,\tilde\rho(dz)\\
&= 6\int_X\int_{I'_z} \tfrac{1}{2}\,\tilde\rho(dy)\,\tilde\rho(dz)\\
&= 6\int_X \tfrac{1}{2}\,\tilde\rho(I'_z)\,\tilde\rho(dz)\\
&= 6\int_X \tfrac{1}{4}\,\tilde\rho(dz)\\
&= \tfrac{3}{2},
\end{aligned}$$

which is absurd. Hence ρ is a Dirac mass (and therefore it is also clear that $c=1$). □

Proof of Lemma 3.45. Let $I := \{(v,x)\in X\times X : v\le x\}$, and for each $x\in X$, let $I_x := \{v\in X : v\le x\}$. Since φ is monotone, we have that for all t, ω and x, $\varphi(t,\omega)I_x\subset I_{\varphi(t,\omega)x}$.

Now define the function $h\colon\Omega\times X\to[0,1]$ by $h(\omega,x)=\mu_\omega(I_x)$. Note that h is measurable, since it can be expressed as

$$h(\omega,x) = \int_X \mathbb{1}_I(v,x)\,\mu_\omega(dv).$$

For each $t\in\mathbb{T}^+$, let $\Omega_t\subset\Omega$ be a $\mathbb{P}$-full set such that for each $\omega\in\Omega_t$, $\varphi(t,\omega)_*\mu_\omega=\mu_{\theta^t\omega}$. Then for all $(\omega,x)\in\Omega_t\times X$, we have

$$\begin{aligned}
h(\Theta^t(\omega,x)) &= \mu_{\theta^t(\omega)}(I_{\varphi(t,\omega)x})\\
&= \mu_\omega(\varphi(t,\omega)^{-1}(I_{\varphi(t,\omega)x}))\\
&\ge \mu_\omega(I_x)\\
&= h(\omega,x).
\end{aligned}$$

So then, letting μ be the integrated form of (μ_ω), we have that $h\circ\Theta^t \overset{\mu\text{-a.s.}}{\ge} h$ for each $t\in\mathbb{T}^+$. Hence, since μ is Θ-ergodic, there exists $c\in[0,1]$ such that $h^{-1}(\{c\})$ is a μ-full set. So for $\mathbb{P}$-almost every $\omega\in\Omega$, μ_ω has the property that $\mu_\omega(I_x)=c$ for μ_ω-almost all $x\in X$. The result then follows by Lemma 3.46. □

3.5 Markov invariant measures and the proofs of the main results

Definition 3.47. A *Markov invariant measure of* φ is a φ-invariant probability measure μ on $(\Omega\times X,\mathcal{F}\otimes\Sigma)$ that is $\mathcal{F}^0_{-\infty}$-measurable.

Remark 3.48. As in Remark 2.38, any φ-invariant measure that is $\mathcal{F}_{-\infty}^{r}$-measurable for some $r \in \mathbb{T}$ is in fact a Markov invariant measure: if, for some $r \in \mathbb{T}^+$, (μ_ω) is an $\mathcal{F}_{-\infty}^{r}$-measurable disintegration of a φ-invariant measure μ, then $(\varphi(r,\theta^{-r}\omega)_*\mu_{\theta^{-r}\omega})$ is an $\mathcal{F}_{-\infty}^{0}$-measurable disintegration of the same probability measure μ.

Now let $\mathcal{I}_M$ denote the set of Markov invariant measures of φ, and let $\mathcal{S}$ denote the set of probability measures on X that are stationary under the Markov transition probabilities $(\varphi_x^t)_{x\in X,\, t\in\mathbb{T}^+}$.

The following theorem is a generalisation of [KS12, Theorem 4.2.9].[11]

Theorem 3.49. *The map $\mu \mapsto \pi_{X*}\mu$ serves as a bijection between $\mathcal{I}_M$ and $\mathcal{S}$. The inverse map can be constructed as follows: for any $\mu \in \mathcal{I}_M$, letting $\rho := \pi_{X*}\mu$ and letting (μ_ω) be any disintegration of μ, we have that for any separable metrisable topology $\mathcal{T}$ on X generating Σ the $\mathcal{M}$-valued stochastic process $(\varphi(t,\theta^{-t}\omega)_*\rho)_{t\in\mathbb{T}^+,\,\omega\in\Omega}$ converges via countable subnets to (μ_ω) in the narrow topology $\mathcal{N}_{\mathcal{T}}$. If the topology $\mathcal{T}$ is such that φ has left-continuous pullback trajectories, then this convergence can be strengthened to almost sure convergence.*

For any $\mu \in \mathcal{I}_M$, μ is the only (Θ^t)-invariant probability measure on $(\Omega \times X, \mathcal{F}\otimes\Sigma)$ whose restriction to $\mathcal{F}_0^\infty \otimes \Sigma$ coincides with $\mathbb{P}|_{\mathcal{F}_0^\infty} \otimes \pi_{X*}\mu$. Hence μ is φ-ergodic if and only if $\pi_{X*}\mu$ is ergodic with respect to the Markov transition probabilities $(\varphi_x^t)_{x\in X,\, t\in\mathbb{T}^+}$.

So then, for any stationary probability measure ρ of the Markov transition probabilities $(\varphi_x^t)_{x\in X,\, t\in\mathbb{T}^+}$ there exists a unique Markov invariant measure μ whose X-projection coincides with ρ. Via disintegration, we can re-express this fact as follows: for any stationary probability measure ρ of the Markov transition probabilities $(\varphi_x^t)_{x\in X,\, t\in\mathbb{T}^+}$ there exists an $\mathcal{F}_{-\infty}^{0}$-measurable φ-invariant random probability measure (μ_ω) such that $E_\omega\mu_\omega = \rho$; and such a random probability measure is unique up to equivalence.

Note that by Corollary 3.37, for any crudely incompressible probability measure ρ, the unique Markov invariant measure whose X-projection coincides with ρ is $\mathbb{P}\otimes\rho$.

Proof of Theorem 3.49. For any $\mu \in \mathcal{I}_M$, we have (by Lemma 3.32 and Remark 3.48) that $\mu|_{\mathcal{F}_0^\infty\otimes\Sigma} = \mathbb{P}|_{\mathcal{F}_0^\infty} \otimes \pi_{X*}\mu$; hence, by Lemma 2.21(i), $\pi_{X*}\mu \in \mathcal{S}$.

We next establish that the map $\mu \mapsto \pi_{X*}\mu$ from $\mathcal{I}_M$ to $\mathcal{S}$ is surjective, by constructing explicitly a (right-)inverse. Fix any $\rho \in \mathcal{S}$. For each $t \in \mathbb{T}^+$, define the probability measure μ^t on $(\Omega \times X, \mathcal{F}_{-t}^\infty \otimes \Sigma)$ by

$$\mu^t(A) \;=\; \mathbb{P}\otimes\rho(\Theta^{-t}(A)) \quad \forall\ A \in \mathcal{F}_{-t}^\infty \otimes \Sigma.$$

Since $\mathbb{P}$ is (θ^t)-invariant, it is clear that μ^t is $\mathbb{P}|_{\mathcal{F}_{-t}^\infty}$-compatible for each t. Now recall from Section 2.6 that for each $s \in \mathbb{T}^+$, Θ^t is $(\mathcal{F}_0^\infty \otimes \Sigma, \mathcal{F}_{-s}^\infty \otimes \Sigma)$-measurable; so then, given any

[11] Working only in the context of a spatially continuous RDS on a Polish space, Theorem 4.2.9 of [KS12] asserts that the map $\mu \mapsto \pi_{X*}\mu$ from $\mathcal{I}_M$ to $\mathcal{S}$ is bijective, with the inverse being as in (1.7) for any unbounded increasing (t_n).

$s, t \in \mathbb{T}^+$ with $s \leq t$, we have that for all $A \in \mathcal{F}_{-s}^{\infty} \otimes \Sigma$,

$$\begin{aligned}\mu^t(A) &= \mathbb{P} \otimes \rho(\Theta^{-t}(A)) \\ &= \mathbb{P} \otimes \rho(\Theta^{s-t}(\Theta^{-s}(A))) \\ &= \mathbb{P} \otimes \rho((\Theta^{-s}(A)) \\ &\qquad \text{(since, by Lemma 2.21(i), } \mathbb{P}|_{\mathcal{F}_0^\infty} \otimes \rho \text{ is } \Theta^{t-s}\text{-invariant)} \\ &= \mu^s(A).\end{aligned}$$

Hence, by Theorem 3.35 (with $\mathcal{G}_t = \mathcal{F}_{-t}^{\infty}$ and $\mathcal{G}_\infty = \mathcal{F}$), there exists a unique probability measure μ on $(\Omega \times X, \mathcal{F} \otimes \Sigma)$ which agrees with μ^t on $\mathcal{F}_{-t}^{\infty} \otimes \Sigma$ for every $t \in \mathbb{T}^+$; and μ is itself $\mathbb{P}$-compatible. Fixing any $\tau \in \mathbb{T}^+$: for all $t \in \mathbb{T}^+$, as in Section 2.6 we have that Θ^τ is $(\mathcal{F}_{-t}^{\infty} \otimes \Sigma, \mathcal{F}_{-(t+\tau)}^{\infty} \otimes \Sigma)$-measurable; and so for any $A \in \mathcal{F}_{-t}^{\infty} \otimes \Sigma$, noting that A is obviously also in $\mathcal{F}_{-(t+\tau)}^{\infty} \otimes \Sigma$, we have that

$$\begin{aligned}\Theta_*^\tau \mu(A) &= \mu^t(\Theta^{-\tau}(A)) \\ &= \mathbb{P} \otimes \rho(\Theta^{-(t+\tau)}(A)) \\ &= \mathbb{P} \otimes \rho(\Theta^{-t}(A)) \\ &\qquad \text{(since, by Lemma 2.21(i), } \mathbb{P}|_{\mathcal{F}_0^\infty} \otimes \rho \text{ is } \Theta^{\tau}\text{-invariant)} \\ &= \mu^t(A).\end{aligned}$$

Hence $\Theta_*^\tau \mu$ is equal to μ. This is true for any $\tau \in \mathbb{T}^+$, and therefore μ is φ-invariant. Now by Lemma 3.36, for each $t \in \mathbb{T}^+$ the integrated form of $(\varphi(t, \theta^{-t}\omega)_*\rho)_{\omega \in \Omega}$ agrees with μ^t on $\mathcal{F}_{-t}^{\infty} \otimes \Sigma$. Hence, by Theorem 3.33, letting (μ_ω) be any disintegration of μ, we have that for any separable metrisable topology $\mathcal{T}$ on X generating Σ the $\mathcal{M}$-valued stochastic process $(\varphi(t, \theta^{-t}\omega)_*\rho)_{t \in \mathbb{T}^+, \omega \in \Omega}$ converges via countable subnets to (μ_ω) in $\mathcal{N}_\mathcal{T}$; in the case that φ has left-continuous pullback trajectories under $\mathcal{T}$, we have that for any bounded continuous $g\colon X \to \mathbb{R}$ the map $t \mapsto g(\varphi(t, \theta^{-t}\omega)x)$ is left-continuous for all x and ω, and therefore (by the dominated convergence theorem) the map $t \mapsto \int_X g(\varphi(t, \theta^{-t}\omega)x)\, \rho(dx)$ is left-continuous for all ω; hence, in this case, the map $t \mapsto \varphi(t, \theta^{-t}\omega)_*\rho$ is left-continuous in $\mathcal{N}_\mathcal{T}$ for all ω, and so $(\varphi(t, \theta^{-t}\omega)_*\rho)_{t \in \mathbb{T}^+, \omega \in \Omega}$ converges almost surely to (μ_ω) in $\mathcal{N}_\mathcal{T}$.

In any case, the measure μ is $\mathcal{F}_{-\infty}^0$-measurable, since, fixing any Polish topology $\mathcal{T}$ on X generating Σ and an arbitrary probability measure c on X, the random probability measure

$$\mu_\omega = \begin{cases} \mathcal{N}_\mathcal{T}\text{-}\lim\limits_{n \to \infty} \varphi(n, \theta^{-n}\omega)_*\rho & \text{if this limit exists} \\ c & \text{otherwise} \end{cases}$$

is a disintegration of μ. Hence $\mu \in \mathcal{I}_M$. By construction, $\mu|_{\mathcal{F}_0^\infty \otimes \Sigma} = \mathbb{P}|_{\mathcal{F}_0^\infty} \otimes \rho$, and therefore (by Lemma 3.32) $\pi_{X*}\mu = \rho$. This completes the proof of the surjectivity of the map $\mu \mapsto \pi_{X*}\mu$ from $\mathcal{I}_M$ to $\mathcal{S}$.

We next show that for any $\tilde{\mu} \in \mathcal{I}_M$, $\tilde{\mu}$ is the only (Θ^t)-invariant probability measure on $(\Omega \times X, \mathcal{F} \otimes \Sigma)$ whose restriction to $\mathcal{F}_0^\infty \otimes \Sigma$ coincides with $\mathbb{P}|_{\mathcal{F}_0^\infty} \otimes \pi_{X*}\tilde{\mu}$; by Lemma 3.32, this implies the injectivity of the map $\mu \mapsto \pi_{X*}\mu$ from $\mathcal{I}_M$ to $\mathcal{S}$. Fix any $\tilde{\mu} \in \mathcal{I}_M$; let $\rho := \pi_{X*}\tilde{\mu} \in \mathcal{S}$, and let μ^t (for each t) and μ be as constructed above. Let μ' be any

(Θ^t)-invariant probability measure with the property that $\mu'|_{\mathcal{F}_0^\infty\otimes\Sigma} = \mathbb{P}|_{\mathcal{F}_0^\infty} \otimes \rho$. Then for any $t \in \mathbb{T}^+$, for all $A \in \mathcal{F}_{-t}^\infty \otimes \Sigma$,

$$\mu'(A) \;=\; \mu'(\Theta^{-t}(A)) \;=\; \mathbb{P}\otimes\rho(\Theta^{-t}(A)) \;=\; \mu^t(A),$$

and therefore $\mu' = \mu$; but since $\tilde{\mu}$ itself has the property that $\tilde{\mu}|_{\mathcal{F}_0^\infty\otimes\Sigma} = \mathbb{P}|_{\mathcal{F}_0^\infty} \otimes \rho$, we can conclude that $\mu' = \tilde{\mu}$, as required.

Now for any $\mu \in \mathcal{I}_M$ that is φ-ergodic, it is clear that $\mathbb{P}|_{\mathcal{F}_0^\infty} \otimes \pi_{X*}\mu$ is ergodic with respect to (Θ^t), and therefore by Lemma 2.21(ii), $\pi_{X*}\mu$ is ergodic with respect to the Markov transition probabilities $(\varphi_x^t)_{x\in X,\, t\in\mathbb{T}^+}$. Conversely, for any $\mu \in \mathcal{I}_M$ that is *not* φ-ergodic, μ can be expressed as a non-trivial convex combination of two (Θ^t)-invariant probability measures ν_1 and ν_2 that are distinct from μ; since μ is the only (Θ^t)-invariant probability measure on $(\Omega\times X, \mathcal{F}\otimes\Sigma)$ whose restriction to $\mathcal{F}_0^\infty\otimes\Sigma$ coincides with $\mathbb{P}|_{\mathcal{F}_0^\infty}\otimes\pi_{X*}\mu$, we have that $\nu_1|_{\mathcal{F}_0^\infty\otimes\Sigma}$ and $\nu_2|_{\mathcal{F}_0^\infty\otimes\Sigma}$ are distinct from $\mathbb{P}|_{\mathcal{F}_0^\infty}\otimes\pi_{X*}\mu$. Thus $\mathbb{P}|_{\mathcal{F}_0^\infty}\otimes\pi_{X*}\mu$ is *not* ergodic with respect to (Θ^t), and therefore, by Lemma 2.21(ii), $\pi_{X*}\mu$ is not ergodic with respect to the Markov transition probabilities $(\varphi_x^t)_{x\in X,\, t\in\mathbb{T}^+}$. □

Remark 3.50. For any $\mathbb{P}$-compatible probability measure μ, we can define the *initial observation time of* μ to be the infimum of the set of values $r \in \bar{\mathbb{T}}$ such that μ is $\mathcal{F}_{-\infty}^r$-measurable. We have established (in Remark 3.48) that for any φ-invariant measure μ, the initial observation time of μ cannot be a strictly positive finite value. Now suppose we have a Markov invariant measure μ for which the initial observation time is strictly negative; then there exists $r \in \mathbb{T}^+ \setminus \{0\}$ such that μ is $\mathcal{F}_{-\infty}^{-r}$-measurable, and so by Lemma 3.32, the restriction of μ to $\mathcal{F}_{-r}^\infty\otimes\Sigma$ coincides with $\mathcal{P}|_{\mathcal{F}_{-r}^\infty}\otimes\pi_{X*}\mu$; hence, by Corollary 3.37, $\pi_{X*}\mu$ is crudely incompressible, and therefore $\mu = \mathbb{P}\otimes\pi_{X*}\mu$. So then, we have the following simple classification of Markov invariant measures: for any Markov invariant measure μ, either

(a) $\mu = \mathbb{P}\otimes\rho$ for some crudely incompressible probability measure ρ (in which case the initial observation time of μ is obviously $-\infty$); or

(b) $\pi_{X*}\mu$ is *not* crudely incompressible, and the initial observation time of μ is 0.

In view of Theorem 3.49, we are now in a position to prove Theorem 3.6:

Theorem 3.51. *Let ρ be a stationary probability measure of the Markov transition probabilities $(\varphi_x^t)_{x\in X,\, t\in\mathbb{T}^+}$. Let (μ_ω) be a disintegration of the unique Markov invariant measure μ satisfying $\pi_{X*}\mu = \rho$, and let $Q_\rho \in \mathcal{M}_{(\mathcal{M},\mathfrak{K})}$ be the image measure of $\mathbb{P}$ under the map $\omega \mapsto \mu_\omega$. (This does not depend on which version (μ_ω) of the disintegration is chosen.) Then Q_ρ fulfils the properties described in Theorem 3.6.*

Proof. Since $\omega \mapsto \mu_\omega$ is a random fixed point of φ_* and has a modification that is $\mathcal{F}_{-\infty}^0$-measurable, Lemma 2.39 gives that Q_ρ is ergodic with respect to the Markov kernel $(\varphi_{\tilde{\rho}}^t)_{\tilde{\rho}\in\mathcal{M}}$ for every $t \in \mathbb{T}^+ \setminus \{0\}$. Now fix a separable metrisable topology $\mathcal{T}$ on X generating Σ. For each $t \in \mathbb{T}^+$, φ_ρ^t is the law of the random variable $\omega \mapsto \varphi(t,\theta^{-t}\omega)_*\rho$; so then, for any unbounded increasing sequence $(t_n)_{n\in\mathbb{N}}$ in $\mathbb{T}^+$, since $(\varphi(t_n,\theta^{-t_n}\omega)_*\rho)_{n\in\mathbb{N},\,\omega\in\Omega}$ converges almost surely to $(\mu_\omega)_{\omega\in\Omega}$ in $\mathcal{N}_\mathcal{T}$, Lemma A.20 gives that $\varphi_\rho^{t_n}$ converges to Q_ρ in $\mathcal{N}_{\mathcal{N}_\mathcal{T}}$. Fixing any $r \in \mathbb{N}$, we have that $(\varphi(t,\theta^{-t}\omega)_*\rho)^{\otimes r} = \varphi^{\times r}(t,\theta^{-t}\omega)_*(\rho^{\otimes r})$ for all

t and ω; and so, for any unbounded increasing $(t_n)_{n\in\mathbb{N}}$ in $\mathbb{T}^+$, Lemma A.21 gives that $(\varphi^{\times r}(t_n,\theta^{-t_n}\omega)_*(\rho^{\otimes r}))_{n\in\mathbb{N},\,\omega\in\Omega}$ converges almost surely to $(\mu_\omega^{\otimes r})_{\omega\in\Omega}$ in $\mathcal{N}_{\mathcal{T}^{\otimes r}}$, and therefore

$$\begin{aligned}\varphi_{(r)}^{t_n*}(\rho^{\otimes r}) &= \mathbb{E}_\omega(\varphi^{\times r}(t_n,\theta^{-t_n}\omega)_*(\rho^{\otimes r})) \\ &\overset{n\to\infty}{\to} \mathbb{E}_\omega(\mu_\omega^{\otimes r}) \quad \text{(by Lemma 3.26)} \\ &= E_r(Q).\end{aligned}$$

Now if ρ is ergodic with respect to the Markov transition probabilities $(\varphi_x^t)_{x\in X,\,t\in\mathbb{T}^+}$, then (μ_ω) is φ-ergodic (by Theorem 3.49), and so by Lemma 3.43, there exists $n\in\mathbb{N}\cup\{\infty\}$ such that $\mu_\omega\in\mathcal{K}_n$ for $\mathbb{P}$-almost all $\omega\in\Omega$, and therefore $Q_\rho(\mathcal{K}_n)=1$. □

Theorem 3.13 then follows from Lemma 3.45 and Theorem 3.51.

Remark 3.52. For any probability measure ρ that is ergodic with respect to the Markov transition probabilities $(\varphi_x^t)_{x\in X,\,t\in\mathbb{T}^+}$, φ is statistically synchronising with respect to ρ if and only if the unique (up to equivalence) $\mathcal{F}_{-\infty}^0$-measurable φ-invariant random probability measure (μ_ω) satisfying $\mathbb{E}_\omega\mu_\omega=\rho$ is $\mathbb{P}$-almost everywhere a Dirac mass. In other words, φ is statistically synchronising if and only if there exists an $\mathcal{F}_{-\infty}^0$-measurable random fixed point $q:\Omega\to X$ such that $q_*\mathbb{P}=\rho$.

We now prove Theorem 3.18.

Proof of Theorem 3.18. (A) Let us work with the metric $d(x,y)=\arctan|x-y|$. Suppose ρ is (φ_x^t)-ergodic. On the basis of Theorem 3.13 and Remark 3.52, let $q:\Omega\to X$ be an $\mathcal{F}_{-\infty}^0$-measurable random fixed point of φ such that $q_*\mathbb{P}=\rho$. We will show that q is crudely pullback-attracting over X_ρ. Fix any unbounded countable $S\subset\mathbb{T}^+$. By Theorem 3.49, for $\mathbb{P}$-almost every $\omega\in\Omega$, $\varphi(t,\theta^{-t}\omega)_*\rho$ converges weakly to $\delta_{q(\omega)}$ as t tends to ∞ in S; so fix any sample point ω for which this is the case. Fix $x\in X_\rho$ and $\varepsilon>0$, and let $V:=B_\varepsilon(q(\omega))$. Obviously $\varphi(t,\theta^{-t}\omega)_*\rho(V)\to 1$ as t tends to ∞ in S. Let $A^-:=\{y\in X_\rho : y\le x\}$ and $A^+:=\{y\in X_\rho : y\ge x\}$. It is clear that $\rho(A^-)$ and $\rho(A^+)$ are both positive. Let $\gamma:=\min(\rho(A^-),\rho(A^+))$. Let $T\in S$ be such that for all $t\in S$ with $t\ge T$, $\varphi(t,\theta^{-t}\omega)_*\rho(V)>1-\gamma$. Fix any $t\in S$ with $t\ge T$. Since $\rho(\varphi(t,\theta^{-t}\omega)^{-1}(V))>1-\gamma$, we have that $\varphi(t,\theta^{-t}\omega)^{-1}(V)$ intersects both A^- and A^+; but since φ is monotone, we also have that $\varphi(t,\theta^{-t}\omega)^{-1}(V)$ is convex. Therefore, $x\in\varphi(t,\theta^{-t}\omega)^{-1}(V)$.

(B) Suppose ρ is (φ_x^t)-ergodic. Let $\tilde{\rho}$ be a probability measure on X with $\tilde{\rho}(X_\rho)=1$. For any bounded continuous $g:X\to\mathbb{R}$, the dominated convergence theorem gives that

$$\begin{aligned}\int_X g(y)\,\varphi^{n*}\tilde{\rho}(dy) &= \int_{\Omega\times X} g(\varphi(n,\theta^{-n}\omega)x)\,\mathbb{P}\otimes\tilde{\rho}(d(\omega,x)) \\ &\to \int_{\Omega\times X} g(q(\omega))\,\mathbb{P}\otimes\tilde{\rho}(d(\omega,x)) \quad \text{as } n\to\infty \text{ in } \mathbb{N} \\ &= \int_\Omega g(q(\omega))\,\mathbb{P}(d\omega) \\ &= \int_X g(y)\,\rho(dy).\end{aligned}$$

Thus $\varphi^{n*}\tilde{\rho}$ converges weakly to ρ as $n\to\infty$ in $\mathbb{N}$. Hence in particular, if $\tilde{\rho}$ is (φ_x^t)-stationary then $\tilde{\rho}=\rho$.

(C) Take any two distinct (φ_x^t)-ergodic probability measures ρ_1 and ρ_2. Since X_{ρ_1} is crudely invariant (by Lemma 2.52), $\rho_2(X_{\rho_1})$ must be equal to either 0 or 1; but $\rho_2(X_{\rho_1})$ cannot be equal to 1, otherwise (by part (B)) we would have that $\rho_2 = \rho_1$. So $\rho_2(X_{\rho_1}) = 0$; and likewise, $\rho_1(X_{\rho_2}) = 0$. Hence, by Lemma 2.51, we have that either X_{ρ_1} and X_{ρ_2} are disjoint or $X_{\rho_1} = X_{\rho_2}$. But it is clear that $X_{\rho_1} \neq X_{\rho_2}$, since $\rho_1(X_{\rho_1}) = 1$ but $\rho_1(X_{\rho_2}) = 0$. So X_{ρ_1} and X_{ρ_2} are disjoint.

(D) It is clear that (i)⇒(ii) and (i)⇒(iii). We next show that (ii) and (iii) together imply (iv). Suppose (iv) does not hold, and let A be a non-empty convex crudely invariant proper subset of X_ρ. Let A^- be the smallest downward-inclusive set containing A, and let A^+ be the smallest upward-inclusive set containing A. Since φ is monotone and A is crudely invariant, it clearly follows that A^- and A^+ are crudely invariant. Since $A = A^- \cap A^+$ and A is a proper subset of X_ρ, we have that at least one of the sets A^- and A^+ does not contain the whole of X_ρ. First suppose that $X_\rho \not\subset A^-$. Then $X_\rho \cap A^-$ is a convex proper subset of X_ρ, and so $\rho(A^-) = \rho(X_\rho \cap A^-) < 1$. But also, $X_\rho \setminus A^-$ is a convex proper subset of X_ρ, and so $\rho(X_\rho \setminus A^-) < 1$, and so $\rho(A^-) > 0$. So then, A^- is a downward-inclusive crudely invariant set that is neither ρ-null nor ρ-full, and so (ii) does not hold. Likewise, if we suppose that $X_\rho \not\subset A^+$, then A^+ is an upward-inclusive crudely invariant set that is neither ρ-null nor ρ-full, and so (iii) does not hold.

It remains to show that (iv)⇒(i). Suppose that (i) does not hold, i.e. that ρ is not ergodic. Let ρ_0 be a (φ_x^t)-ergodic probability measure such that $\rho_0(X_\rho) = 1$. Obviously $X_{\rho_0} \subset X_\rho$. By part (B), it follows that X_{ρ_0} is a proper subset of X_ρ (since ρ_0 is the only (φ_x^t)-stationary probability measure assigning full probability to X_{ρ_0}). Hence (iv) does not hold.

(E) Suppose there is a deterministic fixed point p in X_ρ. If we assume that p is not the maximum of X_ρ, we have that the set $X \cap [-\infty, p]$ is a measurable invariant set that is neither ρ-null nor ρ-full, and therefore ρ is not (φ_x^t)-ergodic. Likewise, if we assume that p is not the minimum of X_ρ, we have that the set $X \cap [p, \infty]$ is a measurable invariant set that is neither ρ-null nor ρ-full, and therefore ρ is not (φ_x^t)-ergodic. Since ρ is not a Dirac mass, p cannot be both the maximum and the minimum of X_ρ; so we conclude that ρ is not (φ_x^t)-ergodic.

Conversely, suppose ρ is not (φ_x^t)-ergodic. Let ρ_0 be a (φ_x^t)-ergodic probability measure such that $\rho_0(X_\rho) = 1$. As in part (D), we have that X_{ρ_0} is a proper subset of X_ρ. Let A be a connected component of $X_\rho \setminus X_{\rho_0}$. First suppose that the elements of A are greater than the elements of X_{ρ_0}, and let $p = \sup X_{\rho_0} = \inf A$. Since X_{ρ_0} is crudely invariant and φ is continuous in space, it is clear that p is crudely subinvariant; and therefore, since φ has right-continuous trajectories, p is subinvariant (by Lemma 2.46). Now for each $n \in \mathbb{N}$, it is clear that $\rho([p, p + \frac{1}{n}] \cap A) > 0$; so let ρ_n be a (φ_x^t)-ergodic probability measure such that $\rho_n([p, p + \frac{1}{n}] \cap A) > 0$, and let $p_n := \inf X_{\rho_n}$. By part (C), X_{ρ_n} and X_{ρ_0} are disjoint, and therefore $p_n \geq p$. Hence in particular, $p_n \in [p, p + \frac{1}{n}]$. For the same reason that p is subinvariant, we have that p_n is superinvariant for each n. Let $\tilde{\Omega}$ be a $\mathbb{P}$-full set such that for all $\omega \in \tilde{\Omega}$, $n \in \mathbb{N}$ and $t \in \mathbb{T}^+$, $\varphi(t, \omega)p_n \geq p_n$. Since φ is continuous in space and $p_n \to p$

as $n \to \infty$, we have that for all $\omega \in \tilde{\Omega}$ and $t \in \mathbb{T}^+$, $\varphi(t,\omega)p \geq p$. So p is superinvariant. But p is also subinvariant. Hence p is a deterministic fixed point. Now if we suppose instead that the elements of A are less than the elements of X_{ρ_0}, then we can similarly show that $\inf X_{\rho_0}$ is a deterministic fixed point.

If ρ is (φ_x^t)-ergodic, Lemma 3.15 gives that the random fixed point q in part (A) is pullback-attracting over X_ρ. □

Chapter 4. Sample-Pathwise Concepts of Synchronisation and Stability

In Chapter 3, we considered a notion of synchronising behaviour which, philosophically, is based around *convergence in probability.* (See, in particular, Corollary 3.9.) We will now consider notions of synchronisation that are based around the notion of mutual convergence of distinct trajectories under individual realisations ω of the noise. Unlike in Chapter 3, the notions that we are consider are *not* measurable invariants, and in many cases, are not even *topological* invariants, but depend on a given metric (or at least, a given uniform structure). Nonetheless, in the case that the phase space X is equipped with a *compact* metrisable topology, all the notions of synchronisation that we shall consider *are* topological invariants, due to the following lemma:

Lemma 4.1. *Fix a compact metrisable topology on X, and suppose we have a $\mathbb{T}^+$-indexed family $(A_t)_{t\in\mathbb{T}^+}$ of sets $A_t \subset X$. Suppose there exists a metrisation of the topology on X in which* $\operatorname{diam}(A_t) \to 0$ *as $t \to \infty$. Then it holds that in* every *metrisation of the topology on X,* $\operatorname{diam}(A_t) \to 0$ *as $t \to \infty$.*

Proof. Fix any metrisation d of the topology on X; we will show that the following statements are equivalent:

(i) $\operatorname{diam}(A_t) \to 0$ as $t \to \infty$;

(ii) for every neighbourhood U of Δ_X in $X \times X$, there exists $T \in \mathbb{T}^+$ such that for all $t \geq T$, $A_t \times A_t \subset U$.

(Since (ii) makes no reference to the metric d, this will complete the proof.) Let us equip $X\times X$ with the metric $d_1(\,(x_1,x_2)\,,\,(y_1,y_2)\,) = d(x_1,y_1) + d(x_2,y_2)$. Since Δ_X is compact, every neighbourhood U of Δ_X contains some ε-neighbourhood of Δ_X. But it is easy to check that for every $\varepsilon > 0$, the ε-neighbourhood U_ε of Δ_X is precisely

$$U_\varepsilon \;=\; \{(x_1,x_2) \in X \times X \,:\, d(x_1,x_2) < \varepsilon\}$$

and therefore, for any $A \subset X$,

$$\operatorname{diam}(A) < \varepsilon \quad\Longleftrightarrow\quad A \times A \,\subset\, U_\varepsilon.$$

Hence it is clear that (i) and (ii) are equivalent. □

Standing Assumption. *Throughout the rest of Chapter 4, we fix a separable metric d on X whose Borel σ-algebra coincides with Σ, and we assume that φ is a right-continuous RDS on the metric space (X,d).*

Since we work with a fixed metric d, we will usually write $\mathcal{B}(X)$ instead of Σ; nonetheless, we are still assuming that $\mathcal{B}(X)$ is standard (which is equivalent to saying that X is a Borel-measurable subset of the d-completion of X). We still write $\mathcal{M}$ for the set of Borel probability measures on X. Recall that φ is called an *open-mapping RDS* if $\varphi(t,\omega)U$ is open for every t, ω and open $U \subset X$.

Before discussing synchronising behaviour, it will be useful first to briefly discuss "random equivalence relations". (Specifically, this will become relevant when considering pairs of initial conditions whose subsequent trajectories synchronise.)

4.1 Borel and random equivalence relations

A *Borel equivalence relation on* $(X, \mathcal{B}(X))$ is an equivalence relation $\sim$ on X such that the set $\{(x,y) \in X \times X : x \sim y\}$ is $\mathcal{B}(X \times X)$-measurable. Note that if $\sim$ is a Borel equivalence relation then every equivalence class of $\sim$ is $\mathcal{B}(X)$-measurable.

Lemma 4.2. *Let $\sim$ be a Borel equivalence relation on $(X, \mathcal{B}(X))$. Given any probability measure ρ on X, the following statements are equivalent:*

(i) one of the equivalence classes of $\sim$ is a ρ-full measure set;

(ii) $\rho \otimes \rho(\,(x,y) : x \sim y\,) = 1$.

Proof. Suppose (i) holds, and let $\tilde{X}$ be a ρ-full equivalence class of $\sim$. Then

$$\begin{aligned} \rho \otimes \rho(\,(x,y) : x \sim y\,) &= \int_X \rho(y : x \sim y)\,\rho(dx) \\ &= \int_{\tilde{X}} \rho(y : x \sim y)\,\rho(dx) \\ &= \int_{\tilde{X}} \rho(\tilde{X})\,\rho(dx) \\ &= 1. \end{aligned}$$

Now suppose (i) does not hold. Then for every $x \in X$, $\rho(y : x \sim y) < 1$ and therefore

$$\rho \otimes \rho(\,(x,y) : x \sim y\,) = \int_X \rho(y : x \sim y)\,\rho(dx) < 1. \qquad \square$$

Now we define a *random equivalence relation on* $(X, \mathcal{B}(X))$ to be an Ω-indexed family $(\sim_\omega)_{\omega \in \Omega}$ of equivalence relations $\sim_\omega$ on X such that the set $\{(\omega, x, y) \in \Omega \times X \times X : x \sim_\omega y\}$ is $(\mathcal{F} \otimes \mathcal{B}(X \times X))$-measurable.

Note that in this case, $\sim_\omega$ is a Borel equivalence relation for every $\omega \in \Omega$.

Given a sub-σ-algebra $\mathcal{G}$ of $\mathcal{F}$, we will say that a random equivalence relation $(\sim_\omega)$ is $\mathcal{G}$*-measurable* if the set $\{(\omega, x, y) \in \Omega \times X \times X : x \sim_\omega y\}$ is $(\mathcal{G} \otimes \mathcal{B}(X \times X))$-measurable.

Definition 4.3. Given a random equivalence relation $(\sim_\omega)$ on $(X, \mathcal{B}(X))$, we define the $\mathbb{P}$*-almost-sure projection of* $(\sim_\omega)$ to be the equivalence relation $\sim$ on X given by

$$x \sim y \iff \mathbb{P}(\omega : x \sim_\omega y) = 1.$$

Note that this is indeed an equivalence relation on X, and moreover, that it is a Borel equivalence relation on $(X, \mathcal{B}(X))$.

Lemma 4.4. *Let $(\sim_\omega)$ be a random equivalence relation on $(X, \mathcal{B}(X))$, with $\sim$ being the $\mathbb{P}$-almost-sure projection of $(\sim_\omega)$. Given any probability measure ρ on X, the following statements are equivalent:*

(i) one of the equivalence classes of $\sim$ is a ρ-full measure set;

(ii) for $\mathbb{P}$-almost every $\omega \in \Omega$, one of the equivalence classes of $\sim_\omega$ is a ρ-full set.

Proof. By Lemma 4.2, (i) is equivalent to saying that $\rho \otimes \rho(\,(x,y) \,:\, x \sim y\,) = 1$, i.e. that $(\rho\otimes\rho)$-every $(x,y) \in X \times X$ has the property that for $\mathbb{P}$-almost all $\omega \in \Omega$, $x \sim_\omega y$; by Fubini's theorem, this is equivalent to saying that for $\mathbb{P}$-almost all $\omega \in \Omega$, $\rho\otimes\rho(\,(x,y) \,:\, x \sim_\omega y\,) = 1$, which (by Lemma 4.2 again) is equivalent to (ii). □

4.2 Synchronisation

Given any sample point $\omega \in \Omega$ and any points $x, y \in X$, we will say that *x and y synchronise under ω*, and write $x \sim_\omega y$, if $d(\varphi(t,\omega)x, \varphi(t,\omega)y) \to 0$ as $t \to \infty$.

It is clear that for each $\omega \in \Omega$, $\sim_\omega$ is an equivalence relation on X. By Lemma 4.1 (applied to the set $A_t := \varphi(t,\omega)\{x,y\}$), if X is compact then for every ω the equivalence relation $\sim_\omega$ is preserved under any topology-preserving change of metric.

Lemma 4.5. *$(\sim_\omega)_{\omega\in\Omega}$ is an $\mathcal{F}_0^\infty$-measurable random equivalence relation on $(X, \mathcal{B}(X))$.*

Proof. Let D be a countable dense subset of $\mathbb{T}^+$. Since φ has right-continuous trajectories, it is clear that $x \sim_\omega y$ if and only if $d(\varphi(t,\omega)x, \varphi(t,\omega)y) \to 0$ as t tends to ∞ in D; hence we can write

$$\mathfrak{R} \;=\; \bigcap_{i=1}^{\infty} \bigcup_{j=1}^{\infty} \bigcap_{\substack{t\in D\\ t\geq j}} \{(\omega,x,y) \,:\, d(\varphi(t,\omega)x, \varphi(t,\omega)y) \leq \tfrac{1}{i}\}.$$

Hence the set $\{(\omega,x,y) : x \sim_\omega y\}$ is $(\mathcal{F}_0^\infty \otimes \mathcal{B}(X \times X))$-measurable. □

Now let $\sim$ be the $\mathbb{P}$-almost-sure projection of $(\sim_\omega)$, that is

$$x \sim y \quad \Longleftrightarrow \quad \mathbb{P}(\,\omega \,:\, d(\varphi(t,\omega)x, \varphi(t,\omega)y) \to 0 \text{ as } t \to \infty\,) \;=\; 1.$$

Given $x, y \in X$, we will say that *x and y synchronise almost surely* if $x \sim y$.

Definition 4.6. We say that *φ is synchronising* if the whole of X is one equivalence class of $\sim$. Given a probability measure ρ on X, we will say that φ is *ρ-almost everywhere synchronising* if $\sim$ admits a ρ-full measure equivalence class.

For example: It is easy to show that the RDS in Example 2.34 is synchronising, and that the RDS is Example 2.37 is synchronising for $\alpha < 0$. The RDS generated by (3.2) in Example 3.21 is also synchronising, as will be explained in Example 6.7.

Lemma 4.7. *Let ρ be an ergodic probability measure of the Markov transition probabilities $(\varphi_x^t)_{x\in X,\, t\in\mathbb{T}^+}$ such that φ is ρ-almost everywhere synchronising. Then φ is statistically synchronising with respect to ρ.*

Proof. Since φ is ρ-almost everywhere synchronising, the stochastic process $(r_t)_{t\in\mathbb{T}^+}$ in Corollary 3.10 converges almost surely to 0. Hence (r_t) converges in probability to 0, and so φ is statistically synchronising with respect to ρ. □

The following is adapted from [New15b, Proposition 2.1.4].

Lemma 4.8. *If φ is synchronising then there is at most one stationary probability measure of the Markov transition probabilities $(\varphi_x^t)_{x\in X,\, t\in\mathbb{T}^+}$; and if such a stationary probability measure ρ exists, we have that for every probability measure m on X, $\varphi^{t*}m \to \rho$ weakly as $t \to \infty$.*[1]

Proof. Suppose φ is synchronising, and suppose we have a stationary probability measure ρ of (φ_x^t). Fix an arbitrary point $p \in X$; we will show that φ_p^t converges weakly to ρ as $t \to \infty$. Let $g : X \to \mathbb{R}$ be any bounded Lipschitz function. Since φ is synchronising, we have that for $(\mathbb{P}\otimes\rho)$-almost all $(\omega, x) \in \Omega\times X$, $x \sim_\omega p$ and so $g(\varphi(t,\omega)x) - g(\varphi(t,\omega)p) \to 0$ as $t \to \infty$. Hence the dominated convergence theorem gives that

$$\underbrace{\int_{\Omega\times X} g(\varphi(t,\omega)x)\,(\mathbb{P}\otimes\rho)(d(\omega,x))}_{ⓐ} - \underbrace{\int_{\Omega} g(\varphi(t,\omega)p)\,\mathbb{P}(d\omega)}_{ⓑ} \;\to\; 0 \quad \text{as } t \to \infty.$$

Observe, however, that

$$ⓐ \;=\; \int_X g(z)\,\rho(dz)$$

since $\mathbb{P}|_{\mathcal{F}_0^\infty} \otimes \rho$ is (Θ^t)-invariant, and that

$$ⓑ \;=\; \int_X g(z)\,\varphi_p^t(dz).$$

Thus we have shown that $\varphi_p^t \to \rho$ weakly as $t \to \infty$ for every $p \in X$. It follows, by the dominated convergence theorem, that $\varphi^{t*}m \to \rho$ weakly as $t \to \infty$ for every probability measure m on X. □

Recall that, given a probability measure ρ on X, a *continuity set of* ρ is a set $A \subset X$ with the property that $\rho(A^\circ) = \rho(\bar{A})$ (i.e. $\rho(\partial A) = 0$). (As in Appendix A, weak convergence of probability measures can be characterised by convergence on measurable continuity sets.) The following is a generalisation of [New15c, Theorem 4.5(II)].

Proposition 4.9. *Let ρ be a probability measure on X that is ergodic with respect to the Markov transition probabilities $(\varphi_x^t)_{x\in X,\, t\in\mathbb{T}^+}$, and let $A \subset X$ be a continuity set of ρ. Then $\mathbb{P}$-almost every $\omega \in \Omega$ has the property that for any equivalence class C of $\sim_\omega$ with $\rho(C) > 0$, for any $x \in C$,*

$$\frac{1}{n}\sum_{i=0}^{n-1} \mathbb{1}_A(\varphi(i,\omega)x) \;\to\; \bar{\rho}(A) \quad \text{as } n \to \infty \qquad \text{if } \mathbb{T} = \mathbb{Z},$$

$$\frac{1}{T}\int_0^T \mathbb{1}_A(\varphi(t,\omega)x)\,dt \;\to\; \bar{\rho}(A) \quad \text{as } T \to \infty \qquad \text{if } \mathbb{T} = \mathbb{R}$$

(where $\bar{\rho}$ denotes the completion of ρ).

[1] When a Feller-continuous family of Markov transition probabilities (μ_x^t) on a separable metric space X admits a probability measure ρ with the property that $\mu^{t*}m \to \rho$ weakly as $t \to \infty$ for every probability measure m on X, it is sometimes said that ρ *is strongly mixing with respect to* (μ_x^t).

Two immediate corollaries are:

Corollary 4.10. *Suppose that φ is synchronising, and that there exists a probability measure ρ on X that is stationary under the Markov transition probabilities $(\varphi_x^t)_{x\in X,\,t\in\mathbb{T}^+}$. Given any $x \in X$ and any continuity set $A \subset X$ of ρ, for $\mathbb{P}$-almost all $\omega \in \Omega$ we have*

$$\frac{1}{n}\sum_{i=0}^{n-1} \mathbb{1}_A(\varphi(i,\omega)x) \;\to\; \bar{\rho}(A) \quad \text{as } n \to \infty \qquad \text{if } \mathbb{T} = \mathbb{Z},$$
$$\frac{1}{T}\int_0^T \mathbb{1}_A(\varphi(t,\omega)x)\,dt \;\to\; \bar{\rho}(A) \quad \text{as } T \to \infty \qquad \text{if } \mathbb{T} = \mathbb{R}.$$

Corollary 4.11. *Let ρ_1 and ρ_2 be distinct ergodic probability measures of the Markov transition probabilities $(\varphi_x^t)_{x\in X,\,t\in\mathbb{T}^+}$. Then $\mathbb{P}$-almost every $\omega \in \Omega$ has the property that for any equivalence class C of $\sim_\omega$, if $\rho_1(C) > 0$ then $\rho_2(C) = 0$.*

(Corollary 4.11 follows from Prpoosition 4.9 by considering a continuity set $A \in \mathcal{B}(X)$ of ρ_1 such that $\rho_2(A) \neq \rho_1(A)$.)

Proof of Proposition 4.9. For convenience, write λ for the counting measure on $\mathbb{Z}$ if $\mathbb{T} = \mathbb{Z}$, or the Lebesgue measure on $\mathbb{R}$ if $\mathbb{T} = \mathbb{R}$; and for any $B \subset \mathbb{R}$, we write $\mathbb{T}_B := \mathbb{T} \cap B$. Given any $S \subset X$ and $\delta > 0$, $B_\delta(S)$ denotes the δ-neighbourhood of S. Note that $\bar{\rho}(A) = \rho(\bar{A})$.

For each $\delta > 0$, let

$$D_\delta \;:=\; \big(B_\delta(A) \setminus \bar{A}\big) \cup \big(B_\delta(X \setminus A) \cap \bar{A}\big).$$

Note that the set D_δ decreases as δ decreases. Also note that for any $x \in D_\delta$, $\max(d(x,A), d(x, X\setminus A)) < \delta$; hence $\bigcap_{\delta>0} D_\delta \subset \partial A$. So then, since $\rho(\partial A) = 0$, we have that $\rho(D_\delta) \to 0$ as $\delta \to 0$.

On the basis of Lemma 2.15, let $\tilde{\Omega} \subset \Omega$ be a $\mathbb{P}$-full set with the property that for every $\omega \in \tilde{\Omega}$, for ρ-almost all $y \in X$ the following two statements hold:

(i) $\frac{1}{T}\int_{\mathbb{T}_{[0,T)}} \mathbb{1}_{\bar{A}}(\varphi(t,\omega)y)\,\lambda(dt) \;\to\; \bar{\rho}(A)$ as $T \to \infty$;

(ii) for each $n \in \mathbb{N}$, $\frac{1}{T}\int_{\mathbb{T}_{[0,T)}} \mathbb{1}_{D_{\frac{1}{n}}}(\varphi(t,\omega)y)\,\lambda(dt) \;\to\; \rho(D_{\frac{1}{n}})$ as $T \to \infty$.

Fix any $\omega \in \tilde{\Omega}$ for which $\sim_\omega$ admits an equivalence class C with $\rho(C) > 0$. Fix such an equivalence class C, and fix any $x \in C$. Since $\rho(C) > 0$, we may fix a point $y \in C$ such that statements (i) and (ii) hold. Now we need to show that

$$\frac{1}{T}\int_{\mathbb{T}_{[0,T)}} \mathbb{1}_A(\varphi(t,\omega)x)\,\lambda(dt) \;\to\; \bar{\rho}(A) \quad \text{as} \quad T \to \infty.$$

Fix $\varepsilon > 0$, and let $N \in \mathbb{N}$ be such that $\rho(D_{\frac{1}{N}}) < \varepsilon$. Let $T_0 \in \mathbb{T}^+$ be such that for all

$t \in \mathbb{T}_{[T_0,\infty)}$, $d(\varphi(t,\omega)x, \varphi(t,\omega)y) < \frac{1}{N}$. For each $T > T_0$, we have that

$$\begin{aligned}
&\left|\frac{1}{T}\int_{\mathbb{T}_{[T_0,T)}} \mathbb{1}_A(\varphi(t,\omega)x)\,\lambda(dt) \;-\; \bar{\rho}(A)\right| \\
&= \left|\frac{1}{T}\int_{\mathbb{T}_{[T_0,T)}} \mathbb{1}_{\bar{A}}(\varphi(t,\omega)y)\,\lambda(dt) \;-\; \bar{\rho}(A) \;+\; \frac{1}{T}\int_{\mathbb{T}_{[T_0,T)}} \mathbb{1}_A(\varphi(t,\omega)x) - \mathbb{1}_{\bar{A}}(\varphi(t,\omega)y)\,\lambda(dt)\right| \\
&\le \left|\frac{1}{T}\int_{\mathbb{T}_{[T_0,T)}} \mathbb{1}_{\bar{A}}(\varphi(t,\omega)y)\,\lambda(dt) \;-\; \bar{\rho}(A)\right| + \frac{1}{T}\int_{\mathbb{T}_{[T_0,T)}} |\mathbb{1}_A(\varphi(t,\omega)x) - \mathbb{1}_{\bar{A}}(\varphi(t,\omega)y)|\,\lambda(dt) \\
&\le \underbrace{\left|\frac{1}{T}\int_{\mathbb{T}_{[T_0,T)}} \mathbb{1}_{\bar{A}}(\varphi(t,\omega)y)\,\lambda(dt) \;-\; \bar{\rho}(A)\right|}_{\to 0 \text{ as } T\to\infty} + \underbrace{\frac{1}{T}\int_{\mathbb{T}_{[T_0,T)}} \mathbb{1}_{D_{\frac{1}{N}}}(\varphi(t,\omega)y)\,\lambda(dt)}_{\to \rho(D_{\frac{1}{N}}) \text{ as } T\to\infty}.
\end{aligned}$$

So, since $\rho(D_{\frac{1}{N}}) < \varepsilon$, we have that for all T sufficiently large,

$$\left|\frac{1}{T}\int_{\mathbb{T}_{[0,T)}} \mathbb{1}_A(\varphi(t,\omega)x)\,\lambda(dt) \;-\; \bar{\rho}(A)\right| \;<\; \varepsilon$$

as required. □

The following is a slight generalisation of part (b) of the proof of [LeJ87, Proposition 2].

Lemma 4.12. *Let ρ be an ergodic probability measure of the Markov transition probabilities $(\varphi_x^t)_{x\in X,\,t\in\mathbb{T}^+}$, and suppose there exists a $\mathbb{P}$-positive measure set $E \in \mathcal{F}$ such that for each $\omega \in E$, $\sim_\omega$ admits an equivalence class of strictly positive measure according to ρ. Then the ρ-clustering number of φ is finite.*

Proof. Let $\mathfrak{R} := \{(\omega,x,y) : x \sim_\omega y\}$, and let $\delta := \mathbb{P}\otimes\rho\otimes\rho(\mathfrak{R})$. We have

$$\begin{aligned}
\delta \;&=\; \int_\Omega\int_X \rho(y\in X : y\sim_\omega x)\,\rho(dx)\,\mathbb{P}(d\omega) \\
&\ge\; \int_E\int_X \rho(y\in X : y\sim_\omega x)\,\rho(dx)\,\mathbb{P}(d\omega) \\
&>\; 0.
\end{aligned}$$

Fix any $\varepsilon > 0$, and let $G_\varepsilon := \{(x,y) : d(x,y) \le \varepsilon\}$. For each $(\omega,x,y) \in \mathfrak{R}$ we have that $\mathbb{1}_{G_\varepsilon}(\varphi(t,\omega)x, \varphi(t,\omega)y) \to 1$ as $t\to\infty$, and therefore (by the dominated convergence theorem),

$$\lim_{t\to\infty}\int_{\mathfrak{R}} \mathbb{1}_{G_\varepsilon}(\varphi(t,\omega)x,\varphi(t,\omega)y)\,\mathbb{P}\otimes\rho\otimes\rho(d(\omega,x,y)) \;=\; \delta.$$

Hence in particular,

$$\limsup_{t\to\infty}\,\varphi_{(2)}^{t*}(\rho\otimes\rho)(G_\varepsilon) \;=\; \limsup_{t\to\infty}\int_{\Omega\times X\times X} \mathbb{1}_{G_\varepsilon}(\varphi^{\times 2}(t,\omega)(x,y))\,\mathbb{P}\otimes\rho\otimes\rho(d(\omega,x,y)) \;\ge\; \delta,$$

and so, letting Q_ρ be the statistical equilibrium associated to ρ, $E_2(Q_\rho)(G_\varepsilon) \ge \delta$. Since ε was arbitrary, it follows that $E_2(Q_\rho)(\Delta_X) \ge \delta$. So $E_2(Q_\rho)(\Delta_X) > 0$, and therefore the ρ-clustering number of φ is finite. □

Remark 4.13. One can also obtain Lemma 4.12 as a fairly direct consequence of Corollary 3.9(A).

Lemma 4.14. *Suppose that φ is invertible, and that φ is synchronising. Let ρ be a probability measure that is ergodic with respect to* either *the Markov transition probabilities* $(\varphi_x^t)_{x\in X,\, t\in\mathbb{T}^+}$ or *the Markov transition probabilities* $(\bar{\varphi}_x^t)_{x\in X,\, t\in\mathbb{T}^+}$. *Then ρ is either atomless or a Dirac mass at a deterministic fixed point.*

Proof. If φ is synchronising then any finite invariant subset of X must be a singleton. Hence the result follows immediately from Lemma 2.58. □

4.3 Concepts of local stability

Overview

Suppose we have a physical process whose time-evolution is "theoretically" governed by some mathematical model. In practice, the process will always be subject to small perturbations from various sources not accounted for in the model; and moreover, if the model includes an assumption on what the initial state of the process is, in practice the initial state will most likely not be *exactly* as is assumed in the model. Heuristically, we regard the process as being "stable" (with respect to the model) if, in spite of these practical considerations, the time-evolution of the physical process will still not "deviate too seriously" away from the time-evolution predicted by the model.

In the context of *synchronisation* of different processes, stability is a highly relevant issue. If, under some (deterministic or stochastic) mathematical model, two physical processes are predicted to synchronise with each other, this synchronisation will never actually be achieved in practice if the processes are easily "knocked off course" by small unaccounted-for perturbations.

Perhaps the simplest approach to examining "stability" is the following: assume that the model for the evolution of the process really is accurate, but imagine that at some time, we perturb the process within an instant from its current state x to a new state $x+\delta$ that is close to x; what effect will this perturbation have on the subsequent trajectory?

From this point of view, we now mention two basic notions of stability for processes governed by *deterministic* models:

- *Lyapunov stability.* Heuristically, this is the notion that if a process is subjected to a small perturbation at some given moment in time, provided we know that the perturbation is small enough, the subsequent evolution of the process will never deviate too far away from how the process would have evolved if the perturbation had not occurred.

- *Asymptotic stability.* In the sense that we shall use the term,[2] heuristically, this is the notion that if a process is subjected to a small perturbation at some

[2] Asymptotic stability is often defined as the combination of both Lyapunov stability and local pointwise attractivity; the definition that we shall use (which is similar in principle to that used in

given moment in time, there is a definite time-scale within which—provided the perturbation was sufficiently small—the subsequent evolution of the process will return to being "practically the same" as if the perturbation had not occurred.

To illustrate: Imagine we have a fixed metal dome that is modelled as rigid and frictionless, and we have a "point particle" placed on the surface of this dome. Imagine that the particle is "intended" to start with zero velocity at the very *top* of the dome, and is modelled as being subject only to gravity and the contact force from the dome. Provided the curvature at the top of the dome is finite, the classical laws of mechanics dictate that the particle will forever *remain* at the top of the dome. However, if the initial position of the particle is not quite perfectly at the top of the dome, then the particle will slide down the dome, and will eventually be in a completely different location from if it had started perfectly at the top of the dome. So a stationary particle at the top of the dome is *not* Lyapunov stable.

By contrast, imagine we have a fixed upright metal cup that is modelled as rigid and frictionless, and we have a "point particle" placed on the inside of this cup. Imagine that the particle is "intended" to start with zero velocity at the very *bottom* of the cup, and is modelled as being subject only to gravity and the contact force from the cup. Once again, the classical laws of mechanics dictate that the particle will forever remain at the bottom of the cup. Now if the initial position of the particle is not quite perfectly at the bottom of the cup but is very close, and the initial speed of the particle is not exactly zero but is very small, then the particle *will* still forever move around very close to the bottom of the cup. Thus, a stationary particle at the bottom of the cup *is* Lyapunov stable; but since there are no dissipative forces present, a stationary particle at the bottom of the cup is *not* asymptotically stable. However, if we now imagine that the cup is filled with air and our model incorporates air resistance, then any particle which does not escape the cup will eventually settle towards being stationary at the bottom of the cup. Hence a stationary particle at the bottom of the cup *is* now asymptotically stable.

We will soon show that, provided basic continuity requirements are satisfied, asymptotic stability always implies Lyapunov stability.

So far, our description of stability has been within a deterministic setting, and we will soon go on to formalise the above notions within the deterministic setting. In the case of a *noise-influenced* process, each possible realisation ω of the noise gives rise to a law specifying the time-evolution of the process;[3] and hence, for *each* ω, we can consider the notions of Lyapunov and asymptotic stability.

[FGS14]) is slightly different from this, but "very nearly coincides" with this; see the Appendix of [New15b] for details.

[3] In reality, many stochastic models for the evolution of a system do not actually assign to each noise-realisation ω a law for the evolution of a system; instead, they provide an *equivalence class* of such assignments, where two such assignments are "equivalent" if they agree on almost every noise-realisation ω. (For example, this is the case for a typical multiplicative-noise SDE with more than one diffusion term.) In such a situation, every representative of the equivalence class serves as an equally valid model; so one can just fix an arbitrary representative of the equivalence class.

We now formalise the notions of Lyapunov and asymptotic stability, and show that (under mild conditions) asymptotic stability implies Lyapunov stability. This is an important fact to prove, since, when we go on to consider stability in random dynamical systems, we will only work with asymptotic stability.

Stability in non-autonomous dynamical systems

(In the following discussion of stability in non-autonomous dynamical systems, the separability of the metric d and the standardness of $\mathcal{B}(X)$ is not relevant.)

A *non-autonomous dynamical system* on the metric space (X,d) is a family $(f_s^{s+t})_{s\in\mathbb{T},\,t\in\mathbb{T}^+}$ of continuous functions $f_s^{s+t}\colon X \to X$ such that

(i) $f_s^s = \mathrm{id}_X$ for all $s \in \mathbb{T}$;

(ii) $f_{t_0}^{t_2} = f_{t_1}^{t_2} \circ f_{t_0}^{t_1}$ for all $t_0 \le t_1 \le t_2$ in $\mathbb{T}$.

Definition 4.15. We will say that a non-autonomous dynamical system (f_s^{s+t}) is *right-continuous* if for any decreasing[4] sequence (t_n) in $\mathbb{T}^+$ converging to a value t and any sequence (x_n) in X converging to a point x, $f_s^{s+t_n}(x_n) \to f_s^{s+t}(x)$ as $n \to \infty$ for all $s \in \mathbb{T}^+$.

Note that if φ is a right-continuous RDS, then $(\varphi(t,\theta^s\omega))_{s\in\mathbb{T},\,t\in\mathbb{T}^+}$ is a right-continuous non-autonomous dynamical system for every $\omega \in \Omega$.

Definition 4.16. We will say that a non-autonomous dynamical system (f_s^{s+t}) is *càdlàg* if (f_s^{s+t}) is right-continuous and for each $s \in \mathbb{T}$ and $t \in \mathbb{T}^+ \setminus \{0\}$ there exists a function[5] $g_s^{s+t}\colon X \to X$ such that for any strictly increasing sequence (t_n) in $\mathbb{T}^+$ converging to t and any sequence (x_n) in X converging to a point x, $f_s^{s+t_n}(x_n) \to g_s^{s+t}(x)$ as $n \to \infty$.

Note that if φ is a càdlàg RDS, then $(\varphi(t,\theta^s\omega))_{s\in\mathbb{T},\,t\in\mathbb{T}^+}$ is a càdlàg non-autonomous dynamical system for every $\omega \in \Omega$.

Now recall that a family $(f_\alpha)_{\alpha\in I}$ of functions $f_\alpha\colon X \to X$ is said to be *equicontinuous* at a point $x \in X$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that for all $\alpha \in I$, $f_\alpha(B_\delta(x)) \subset B_\varepsilon(f_\alpha(x))$.

Definition 4.17. We will say that a non-autonomous dynamical system $(f_s^{s+t})_{s\in\mathbb{T},\,t\in\mathbb{T}^+}$ is *finite-time stable* if for every $s \in \mathbb{T}$ and $T \in \mathbb{T}^+$, the family of functions $(f_s^{s+t})_{0\le t\le T}$ is equicontinuous at every point in X.

The following is [New15b, Lemma A1].

Lemma 4.18. *Every càdlàg non-autonomous dynamical system is finite-time stable.*

Proof. Let $(f_s^{s+t})_{s\in\mathbb{T},\,t\in\mathbb{T}^+}$ be a càdlàg non-autonomous dynamical system, and suppose for a contradiction that we have $s \in \mathbb{T}$, $T \in \mathbb{T}^+$ and $x \in X$ such that the family $(f_s^{s+t})_{0\le t\le T}$ is not equicontinuous at x. Then there exist $\varepsilon > 0$, a sequence (x_n) in X converging to x, and a sequence (t_n) in $\mathbb{T}^+ \cap [0,T]$, such that $d(f_s^{s+t_n}(x_n), f_s^{s+t_n}(x)) > \varepsilon$ for all

[4] Here, a "decreasing sequence" need *not* be strictly decreasing.

[5] For our purposes here, we will not need the function g_s^{s+t} to be continuous.

n. Let $(n_i)_{i\in\mathbb{N}}$ be an unbounded increasing sequence in $\mathbb{N}$ such that *either* $(t_{n_i})_{i\in\mathbb{N}}$ is a decreasing sequence (with $t^* := \inf_i t_{n_i}$) *or* $(t_{n_i})_{i\in\mathbb{N}}$ is a strictly increasing sequence (with $t^* := \sup_i t_{n_i}$). If (t_{n_i}) is decreasing then $f_s^{s+t_{n_i}}(x_{n_i}) \to f_s^{s+t^*}(x)$ and $f_s^{s+t_{n_i}}(x) \to f_s^{s+t^*}(x)$ as $i \to \infty$, so $d(f_s^{s+t_{n_i}}(x_{n_i}), f_s^{s+t_{n_i}}(x)) \to 0$ as $i \to \infty$, contradicting the fact that $d(f_s^{s+t_n}(x_n), f_s^{s+t_n}(x)) > \varepsilon$ for all n. If (t_{n_i}) is strictly increasing then $f_s^{s+t_{n_i}}(x_{n_i}) \to g_s^{s+t^*}(x)$ and $f_s^{s+t_{n_i}}(x) \to g_s^{s+t^*}(x)$ as $i \to \infty$, so once again, $d(f_s^{s+t_{n_i}}(x_{n_i}), f_s^{s+t_{n_i}}(x)) \to 0$ as $i \to \infty$, contradicting the fact that $d(f_s^{s+t_n}(x_n), f_s^{s+t_n}(x)) > \varepsilon$ for all n. So in either case, we have a contradiction. □

Definition 4.19. Given a point $x \in X$ and a value $s_0 \in \mathbb{T}$, we will say that a non-autonomous dynamical system $(f_s^{s+t})_{s\in\mathbb{T}, t\in\mathbb{T}^+}$ is *Lyapunov stable at x at time s_0* if the family of functions $(f_{s_0}^{s_0+t})_{t\in\mathbb{T}^+}$ is equicontinuous at x.

Definition 4.20. Given a point $x \in X$ and a value $s_0 \in \mathbb{T}$, we will say that a non-autonomous dynamical system $(f_s^{s+t})_{s\in\mathbb{T}, t\in\mathbb{T}^+}$ is *asymptotically stable at x at time s_0* if there exists a neighbourhood U of x such that $\operatorname{diam}(f_{s_0}^{s_0+t}(U)) \to 0$ as $t \to \infty$.

Example 4.21. Let $X = \mathbb{R}$ and let $\mathbb{T} = \mathbb{R}$. (A) Suppose $f_s^{s+t}(x) = xe^{-t}$. Then for every $x \in \mathbb{R}$ and $s_0 \in \mathbb{R}$, (f_s^{s+t}) is asymptotically stable at x at time s_0. (B) Suppose $f_s^{s+t}(x) = x$. Then for every $x \in \mathbb{R}$ and $s_0 \in \mathbb{R}$, (f_s^{s+t}) is Lyapunov stable at x at time s_0, but not asymptotically stable at x at time s_0. (C) Suppose $f_s^{s+t}(x) = xe^{\operatorname{sgn}(x)t}$. Then for every $x < 0$ and $s_0 \in \mathbb{R}$, (f_s^{s+t}) is asymptotically stable at x at time s_0; but for every $x \geq 0$ and $s_0 \in \mathbb{R}$, (f_s^{s+t}) is not Lyapunov stable at x at time s_0.

The following is essentially [New15b, Theorem A11(II)].

Lemma 4.22. *Let $(f_s^{s+t})_{s\in\mathbb{T}, t\in\mathbb{T}^+}$ be a non-autonomous dynamical system that is finite-time stable. For any $x \in X$ and $s_0 \in \mathbb{T}$, if $(f_s^{s+t})_{s\in\mathbb{T}, t\in\mathbb{T}^+}$ is asymptotically stable at x at time s_0, then $(f_s^{s+t})_{s\in\mathbb{T}, t\in\mathbb{T}^+}$ is Lyapunov stable at x at time s_0.*

Proof. Let $x \in X$ and $s_0 \in \mathbb{T}$ be such that $(f_s^{s+t})_{s\in\mathbb{T}, t\in\mathbb{T}^+}$ is asymptotically stable at x at time s_0, and fix $\varepsilon > 0$. Let $r > 0$ be such that $\operatorname{diam}(f_{s_0}^{s_0+t}(B_r(x))) \to 0$ as $t \to \infty$. Let $T \in \mathbb{T}^+$ be such that for all $t > T$, $f_{s_0}^{s_0+t}(B_r(x)) \subset B_\varepsilon(f_{s_0}^{s_0+t}(x))$. On the basis of the fact that $(f_s^{s+t})_{s\in\mathbb{T}, t\in\mathbb{T}^+}$ is finite-time stable, let $\tilde{r} > 0$ be such that for all $0 \leq t \leq T$, $f_{s_0}^{s_0+t}(B_{\tilde{r}}(x)) \subset B_\varepsilon(f_{s_0}^{s_0+t}(x))$. Then setting $\delta := \min(r, \tilde{r})$, we have that $f_{s_0}^{s_0+t}(B_\delta(x)) \subset B_\varepsilon(f_{s_0}^{s_0+t}(x))$ for all $t \in \mathbb{T}^+$. So we are done. □

4.4 Asymptotic stability in RDS

(Most of the content of this section is taken from Section 2.2 of [New15b].)

Typically, local stability of trajectories of RDS is investigated by considering *Lyapunov exponents*. Specifically, given a differentiable RDS[6] on a Riemannian manifold and an ergodic probability measure ρ for the associated Markov transition probabilities, provided the partial derivatives of the RDS are "reasonably well controlled", there exists a value

[6] This means that the RDS is differentiable in space, with the derivatives depending continuously (or right-continuously) on time.

$\lambda \in [-\infty, \infty)$ (called the *maximal Lyapunov exponent*) which, loosely speaking, is a measure of infinitesimal-scale repulsivity common to the trajectories of ρ-almost all initial conditions almost surely. If $\lambda = 0$, one cannot normally make any conclusions about either Lyapunov or asymptotic stability. However, if $\lambda < 0$ then one can usually conclude that (at any one given time) the trajectories of ρ-almost all initial conditions are almost surely asymptotically stable. (See [New15b, Remark 2.2.12] and the references mentioned therein for further details.)

Naturally then, when considering stability in RDS, we will specifically consider *asymptotic* stability; for càdlàg RDS, this automatically implies Lyapunov stability (by Lemmas 4.22 and 4.18).

Definition 4.23. Given a sample point $\omega \in \Omega$ and a set $A \subset X$, we will say that A *contracts under* ω if $\mathrm{diam}(\varphi(t,\omega)A) \to 0$ as $t \to \infty$.

(By Lemma 4.1, if X is compact then this notion is not specific to our choice d of metrisation of the topology of X.)

Note that for any sample point ω, if $A_1, A_2 \subset X$ are sets which contract under ω and $A_1 \cap A_2 \neq \varnothing$, then $A_1 \cup A_2$ contracts under ω. (More generally, if $A_1, A_2 \subset X$ are sets which contract under ω, and A_1 and A_2 belong to the same equivalence class of the synchronisation equivalence relation $\sim_\omega$, then $A_1 \cup A_2$ contracts under ω.)

Definition 4.24. Given a sample point $\omega \in \Omega$ and a point $x \in X$, we will say that x *is asymptotically stable under* ω, or that the pair (ω, x) *is asymptotically stable*, if there exists a neighbourhood U of x such that U contracts under ω.

In other words, x is asymptotically stable under ω if and only if the non-autonomous dynamical system $(\varphi(t,\theta^s\omega))_{s\in\mathbb{T},\,t\in\mathbb{T}^+}$ is asymptotically stable at x at time 0. More generally: given any $r \in \mathbb{T}$, the statement that the non-autonomous dynamical system $(\varphi(t,\theta^s\omega))_{s\in\mathbb{T},\,t\in\mathbb{T}^+}$ is asymptotically stable at x at time r is precisely the statement that x is asymptotically stable under $\theta^r\omega$.

Now let $O \subset \Omega \times X$ denote the set of all asymptotically stable pairs (ω, x). For each $x \in X$, let $O_x := \{\omega \in \Omega : (\omega, x) \in O\}$, that is O_x is the set of sample points ω under which x is asymptotically stable.

For any $A \subset X$, let $E_A \subset \Omega$ denote the set of sample points under which A contracts. Obviously, for any sets $A_1 \subset A_2 \subset X$, we have that $E_{A_2} \subset E_{A_1}$. Note that for any $x \in X$,

$$O_x = \bigcup_{n=1}^{\infty} E_{B_{\frac{1}{n}}(x)}.$$

Lemma 4.25. *(A) For every open $U \subset X$, $E_U \in \mathcal{F}_0^\infty$. (B) $O \in \mathcal{F}_0^\infty \otimes \mathcal{B}(X)$. (C) For all $t \in \mathbb{T}^+$, $\Theta^{-t}(O) \subset O$. If φ is an open-mapping RDS, then for all $t \in \mathbb{T}^+$, $\Theta^{-t}(O) = O$.*

Proof. (A) Fix a non-empty open set $U \subset X$. Let $S \subset U$ be a countable set that is dense in U. Then (by the spatial continuity of φ) E_U is precisely the set of sample points ω for which $\sup_{x,y\in S} d(\varphi(t,\omega)x, \varphi(t,\omega)y) \to 0$ as $t \to \infty$. But now letting D

be a countable dense subset of $\mathbb{T}^+$, since a pointwise supremum of right-continuous functions is right lower semicontinuous, E_U is precisely the set of sample points ω for which $\sup_{x,y\in S} d(\varphi(t,\omega)x, \varphi(t,\omega)y) \to 0$ as t tends to ∞ in D; that is,

$$E_U = \bigcap_{i=1}^{\infty} \bigcup_{j=1}^{\infty} \bigcap_{t\in D\cap[j,\infty)} \bigcap_{x,y\in S} \{\omega : d(\varphi(t,\omega)x, \varphi(t,\omega)y) < \tfrac{1}{i}\}.$$

So $E_U \in \mathcal{F}_0^\infty$.

(B) Now let $\mathcal{U}$ be a countable base for the topology on X. It is clear that, given any $V \in \mathcal{U}$, every pair $(\omega, x) \in E_V \times V$ is asymptotically stable; conversely, given any asymptotically stable pair (ω, x), there exists a neighbourhood $V \in \mathcal{U}$ of x such that V contracts under ω, i.e. such that $(\omega, x) \in E_V \times V$. Thus, we have that $O = \bigcup_{V\in\mathcal{U}} E_V \times V$, and therefore $O \in \mathcal{F}_0^\infty \otimes \mathcal{B}(X)$.

(C) Given any $t \in \mathbb{T}^+$ and $(\omega, x) \in \Theta^{-t}(O)$, there exists a neighbourhood U of $\varphi(t,\omega)x$ such that U contracts under $\theta^t\omega$, and therefore $\varphi(t,\omega)^{-1}(U)$ contracts under ω; since $\varphi(t,\omega)$ is continuous, $\varphi(t,\omega)^{-1}(U)$ is a neighbourhood of x, and so x is asymptotically stable under ω. Hence $\Theta^{-t}(O) \subset O$. One can similarly show that if φ is an open-mapping RDS then for $t \in \mathbb{T}^+$, $\Theta^t(O) \subset O$ (i.e. $O \subset \Theta^{-t}(O)$); and of course, combining this with the fact that $\Theta^{-t}(O) \subset O$ for all $t \in \mathbb{T}^+$ gives that $\Theta^{-t}(O) = O$ for all $t \in \mathbb{T}^+$. □

Now then, for each $x \in X$, let

$$P_0(x) := \mathbb{P}(O_x) = \mathbb{P}(\theta^{-s}(O_x)) \quad \text{(for any } s \in \mathbb{T})$$

and let $P_r(x) = \mathbb{P}(E_{B_r(x)})$ for all $r > 0$. It is clear that $P_r(x)$ increases as r decreases, with $P_0(x) = \sup_{r>0} P_r(x) = \lim_{r\to 0} P_r(x)$.

Definition 4.26. We will say that a point $x \in X$ is *almost surely stable* if $P_0(x) = 1$. We will say that x is *potentially stable* if $P_0(x) > 0$.

Definition 4.27. Let $A \subset X$ be a set that is invariant under φ. We will say that φ *is everywhere stable in* A if every $x \in A$ is almost surely stable. We will say that φ *is uniformly stable in* A if $P_r(\cdot) \to 1$ uniformly on A as $r \to 0$.

In the case that the invariant set A is compact, there is no difference between everywhere stability and uniform stability:

Lemma 4.28. *Let $K \subset X$ be a compact set. If $P_0(x) = 1$ for all $x \in K$, then $P_r(\cdot) \to 1$ uniformly on K as $r \to 0$.*

Proof. Fix $\varepsilon > 0$. Every $x \in K$ has a neighbourhood U such that $\mathbb{P}(E_U) > 1 - \varepsilon$. So let $\mathcal{U}$ be a collection of open sets covering K, such that for each $U \in \mathcal{U}$, $\mathbb{P}(E_U) > 1 - \varepsilon$. Since K is compact, there exists $\delta > 0$ such that for every $x \in K$ there exists $U_x \in \mathcal{U}$ such that $B_\delta(x) \subset U_x$. So then, for all $r \in (0, \delta)$ and $x \in K$, $P_r(x) = \mathbb{P}(E_{B_\delta(x)}) > 1 - \varepsilon$. □

Lemma 4.29. *Let ρ be an ergodic probability measure of the Markov transition probabilities $(\varphi_x^t)_{x\in X, t\in\mathbb{T}^+}$. Then $\mathbb{P} \otimes \rho(O)$ is equal to either 0 or 1.*

Proof. Follows immediately from Lemma 4.25 (parts (B) and (C)) and the fact that $\mathbb{P}|_{\mathcal{F}_0^\infty} \otimes \rho$ is (Θ^t)-ergodic (Lemma 2.21). □

Now let $U_{ps} \subset X$ be the set of potentially stable points, and let A_s be the set of almost surely stable points.

Proposition 4.30. *(A) $U_{ps} \subset X$ is open, and $X \setminus U_{ps}$ is invariant under φ. If φ is an open-mapping RDS then A_s is very crudely invariant under φ (that is, $\varphi_x^t(A_s) = 1$ for all $x \in A_s$ and $t \in \mathbb{T}^+$). (B) For any stationary probability measure ρ of the Markov transition probabilities $(\varphi_x^t)_{x\in X,\, t\in\mathbb{T}^+}$, $\rho(U_{ps} \setminus A_s) = 0$.*

Lemma 4.31. *For any $x \in X$ and $t \in \mathbb{T}^+$, $P_0(x) \geq \int_X P_0(y)\,\varphi_x^t(dy)$. If φ is an open-mapping RDS then the inequality becomes equality.*

Proof of Lemma 4.31. Note that $P_0(y) = \mathbb{P}(\theta^{-t}(O_y))$ for all y and t. Now fix any $x \in X$ and $t \in \mathbb{T}^+$.

$$\begin{aligned}
\int_X P_0(y)\,\varphi_x^t(dy) &= \int_\Omega P_0(\varphi(t,\omega)x)\,\mathbb{P}(d\omega) \\
&= \int_\Omega \mathbb{P}(\theta^{-t}(O_{\varphi(t,\omega)x}))\,\mathbb{P}(d\omega) \\
&= \int_\Omega \int_\Omega \mathbb{1}_O(\theta^t\tilde{\omega}, \varphi(t,\omega)x)\,\mathbb{P}(d\tilde{\omega})\,\mathbb{P}(d\omega) \\
&= \int_\Omega \mathbb{1}_O(\theta^t\omega, \varphi(t,\omega)x)\,\mathbb{P}(d\omega) \\
&\qquad \text{(using Lemma A.10, since the map } (\omega,y) \mapsto \mathbb{1}_O(\theta^t\omega, y) \\
&\qquad \text{is } (\mathcal{F}_t^\infty \otimes \mathcal{B}(X))\text{-measurable, due to Lemma 4.25(B))} \\
&= \int_\Omega \mathbb{1}_{\Theta^{-t}(O)}(\omega, x)\,\mathbb{P}(d\omega) \\
&\leq \int_\Omega \mathbb{1}_O(\omega, x)\,\mathbb{P}(d\omega) \quad \text{(by Lemma 4.25(C))} \\
&= P_0(x).
\end{aligned}$$

If φ is an open-mapping RDS then the "$\leq$" in the penultimate line becomes "$=$". □

Proof of Proposition 4.30. (A) For any $x \in U_{ps}$ and $r > 0$ with $P_r(x) > 0$, we clearly have that $B_r(x) \subset U_{ps}$. So U_{ps} is open. Now fix any $x \in X \setminus U_{ps}$ and $t \in \mathbb{T}^+$. Since $P_0(x) = 0$, Lemma 4.31 gives that $P_0(y) = 0$ for φ_x^t-almost all $y \in X$, i.e. $\varphi_x^t(X \setminus U_{ps}) = 1$. So $X \setminus U_{ps}$ is very crudely invariant under φ; and therefore, since $X \setminus U_{ps}$ is closed, Lemma 2.75 gives that $X \setminus U_{ps}$ is invariant. Now suppose that φ is an open-mapping RDS, and fix any $x \in A_s$ and $t \in \mathbb{T}^+$. Since $P_0(x) = 1$, Lemma 4.31 gives that $P_0(y) = 1$ for φ_x^t-almost all $y \in X$, i.e. $\varphi_x^t(A_s) = 1$. So A_s is very crudely invariant.

(B) For any ergodic probability measure ρ' of the Markov transition probabilities $(\varphi_x^t)_{x\in X,\, t\in\mathbb{T}^+}$, Lemma 4.29 gives that either $\mathbb{P}\otimes\rho'(O) = 0$ or $\mathbb{P}\otimes\rho'(O) = 1$. In the former case we have that $\rho'(U_{ps}) = \rho'(A_s) = 0$, and in the latter case we have that $\rho'(U_{ps}) = \rho'(A_s) = 1$; so in either case, $\rho'(U_{ps} \setminus A_s) = 0$. Now for any stationary probability measure ρ of the Markov transition probabilities $(\varphi_x^t)_{x\in X,\, t\in\mathbb{T}^+}$, ρ admits an integral representation via ergodic probability measures (as in Appendix C), and therefore $\rho(U_{ps} \setminus A_s) = 0$. □

Definition 4.32. Let ρ be a stationary probability measure of the Markov transition probabilities $(\varphi_x^t)_{x\in X,\, t\in\mathbb{T}^+}$. We will say that φ *is stable with respect to* ρ if $\mathbb{P}\otimes\rho(O) = 1$.

Note that, by Fubini's theorem, the following statements are equivalent:

(i) φ is stable with respect to ρ;

(ii) ρ-almost every $x \in X$ is almost surely stable;

(iii) $\mathbb{P}$-almost every sample point $\omega \in \Omega$ has the property that ρ-almost every $x \in X$ is asymptotically stable under ω.

For any $s \in \mathbb{T}$, since θ^s is $\mathbb{P}$-preserving, we have that these are equivalent to:

(iv) $\mathbb{P}$-almost every sample point $\omega \in \Omega$ has the property that ρ-almost every $x \in X$ is asymptotically stable under $\theta^s\omega$.

The following proposition can be interpreted, crudely, as saying that if φ is stable with respect to ρ, then φ is "ρ-almost uniformly stable in X".

Proposition 4.33. *Let ρ be a stationary probability measure of the Markov transition probabilities $(\varphi_x^t)_{x\in X,\, t\in\mathbb{T}^+}$ such that φ is stable with respect to ρ. For every $\varepsilon > 0$ there exists $\delta > 0$ such that for all $t \in \mathbb{T}^+$, the set*

$$\{\, (\omega, x) \,:\, B_\delta(\varphi(t,\omega)x) \text{ contracts under } \theta^t\omega \,\}$$

contains a set of measure greater than $1-\varepsilon$ according to $\mathbb{P}\otimes\rho$.

Proof. Let $S \subset X$ be a countable dense set, and let

$$\tilde{O} \;:=\; \bigcup_{n=1}^{\infty}\bigcup_{y\in S} E_{B_{\frac{2}{n}}(y)} \times B_{\frac{1}{n}}(y).$$

Note that the collection $\mathcal{U} := \{B_{\frac{1}{n}}(y) : n \in \mathbb{N}, y \in S\}$ is a base for the topology on X, and therefore (as in the proof of Lemma 4.25(B)),

$$\begin{aligned} O \;&=\; \bigcup_{V\in\mathcal{U}} E_V \times V \\ &=\; \bigcup_{n=1}^{\infty}\bigcup_{y\in S} E_{B_{\frac{1}{n}}(y)} \times B_{\frac{1}{n}}(y) \\ &\subset\; \bigcup_{n=1}^{\infty}\bigcup_{y\in S} E_{B_{\frac{2}{n}}(y)} \times B_{\frac{1}{n}}(y) \;=\; \tilde{O}. \end{aligned}$$

So $\mathbb{P}\otimes\rho(\tilde{O}) = 1$.

Now fix $\varepsilon > 0$, let $N \in \mathbb{N}$ be such that

$$\mathbb{P}\otimes\rho\left(\bigcup_{n=1}^{N}\bigcup_{y\in S} E_{B_{\frac{2}{n}}(y)} \times B_{\frac{1}{n}}(y)\right) \;>\; 1-\varepsilon,$$

and let $\delta = \frac{1}{N}$. Fix any $t \in \mathbb{T}^+$. Since $\mathbb{P}|_{\mathcal{F}_0^\infty} \otimes \rho$ is Θ^t-invariant, we have that

$$\begin{aligned}
&\mathbb{P} \otimes \rho \left(\bigcup_{n=1}^{N} \bigcup_{y \in S} \Theta^{-t} \left(E_{B_{\frac{2}{n}}(y)} \times B_{\frac{1}{n}}(y) \right) \right) \\
= \ &\mathbb{P} \otimes \rho \left(\Theta^{-t} \left(\bigcup_{n=1}^{N} \bigcup_{y \in S} E_{B_{\frac{2}{n}}(y)} \times B_{\frac{1}{n}}(y) \right) \right) \\
> \ &1 - \varepsilon.
\end{aligned}$$

So it remains to show that for any $n \in \{1, \ldots, N\}$ and $y \in S$, for any $(\omega, x) \in \Theta^{-t}\left(E_{B_{\frac{2}{n}}(y)} \times B_{\frac{1}{n}}(y) \right)$, $B_\delta(\varphi(t,\omega)x)$ contracts under $\theta^t\omega$. Fix such n, y and (ω, x). We know that $d(y, \varphi(t,\omega)x) < \frac{1}{n}$, and that $B_{\frac{2}{n}}(y)$ contracts under $\theta^t\omega$. For any $z \in B_\delta(\varphi(t,\omega)x)$, we have that

$$d(y,z) \ \le \ d(y, \varphi(t,\omega)x) + d(\varphi(t,\omega)x, z) \ < \ \tfrac{1}{n} + \delta \ \le \ \tfrac{2}{n}.$$

So $B_\delta(\varphi(t,\omega)x) \subset B_{\frac{2}{n}}(y)$, and therefore $B_\delta(\varphi(t,\omega)x)$ contracts under $\theta^t\omega$. □

Now recall that two distinct ergodic probability measures are mutually singular. A further statement can be made in the case that φ is stable with respect to one of the two measures.

Lemma 4.34. *Let ρ be an ergodic probability measure of the Markov transition probabilities $(\varphi_x^t)_{x \in X, t \in \mathbb{T}^+}$ such that φ is stable with respect to ρ. For any ergodic probability measure ρ' of $(\varphi_x^t)_{x \in X, t \in \mathbb{T}^+}$ distinct from ρ, $(\operatorname{supp}\rho) \cap (\operatorname{supp}\rho')$ has empty interior relative to* $\operatorname{supp}\rho$.

Proof. Let ρ' be an ergodic probability measure of (φ_x^t) such that $(\operatorname{supp}\rho) \cap (\operatorname{supp}\rho')$ has non-empty interior relative to $\operatorname{supp}\rho$. So $(\operatorname{supp}\rho) \cap (\operatorname{supp}\rho')$ has ρ-positive measure. Since φ is stable with respect to ρ, ρ-almost every $x \in X$ is almost surely stable. So pick a point $x \in (\operatorname{supp}\rho) \cap (\operatorname{supp}\rho')$ that is almost surely stable. For $\mathbb{P}$-almost every $\omega \in \Omega$, the $\sim_\omega$-equivalence class of x contains a neighbourhood of x, and therefore has both ρ-positive measure and ρ'-positive measure. By Corollary 4.11, it follows that $\rho = \rho'$. □

We go on to consider "sets admitting stable trajectories". For any $A \subset X$, define the set

$$\begin{aligned}
O_A \ &:= \{\omega \in \Omega : \textit{at least one point } x \in A \textit{ is asymptotically stable under } \omega\} \\
&= \bigcup_{x \in A} O_x.
\end{aligned}$$

It is easy to see that $O_A = O_{\bar{A}}$.

Lemma 4.35. *For any $A \subset X$, O_A is $\mathcal{F}_0^\infty$-measurable.*

Proof. Let $\mathcal{U}$ be a countable base for the topology on X, and let

$$\mathcal{U}_A \ := \ \{U \in \mathcal{U} : U \cap A \ne \varnothing\}.$$

It is clear that for any $\omega \in \Omega$, the existence of a point $x \in A$ that is asymptotically stable under ω is equivalent to the existence of an open set $U \in \mathcal{U}_A$ that contracts under ω. In other words,

$$O_A \ = \ \bigcup_{U \in \mathcal{U}_A} E_U,$$

and so Lemma 4.25(A) gives the result. □

Definition 4.36. We will say that a set $A \subset X$ *admits stable trajectories* if $\mathbb{P}(O_A) > 0$.

Proposition 4.37. *(A) A set $A \subset X$ admits stable trajectories if and only if there exists $x \in A$ that is potentially stable. (B) Suppose φ is an open-mapping RDS. Then a closed invariant set $G \subset X$ admits stable trajectories if and only if $\mathbb{P}(O_G) = 1$.*

The proof of part (A) uses the following fact:

Lemma 4.38. *$\mathbb{P}$-almost every $\omega \in \Omega$ has the property that for any $x \in X$, if (ω, x) is asymptotically stable then x is potentially stable.*

Proof of Lemma 4.38. Let $\mathcal{U}$ be a countable base for the topology on X, and let

$$\mathcal{U}_0 := \{U \in \mathcal{U} : \mathbb{P}(E_U) = 0\}.$$

Let

$$\tilde{\Omega} := \Omega \setminus \bigcup_{U \in \mathcal{U}_0} E_U.$$

Now fix any $\omega \in \tilde{\Omega}$ and $x \in X$. If (ω, x) is asymptotically stable then there exists $U \in \mathcal{U}$ with $x \in U$ such that $\omega \in E_U$, and hence $U \notin \mathcal{U}_0$; so $\mathbb{P}(E_U) > 0$ and therefore x is potentially stable. □

Proof of Proposition 4.37. (A) For every $x \in A$, since $O_x \subset O_A$, it is clear that if x is potentially stable (i.e. $\mathbb{P}(O_x) > 0$) then A admits stable trajectories. Conversely: Let $\tilde{\Omega}$ be a $\mathbb{P}$-full set with the property described in Lemma 4.38. If $\mathbb{P}(O_A) > 0$ then $O_A \cap \tilde{\Omega}$ is non-empty. So take any $\omega \in O_A \cap \tilde{\Omega}$, and let $x \in A$ be a point that is asymptotically stable under ω; then, since $\omega \in \tilde{\Omega}$, x is potentially stable.

(B) Let $\hat{\Omega}$ be a $\mathbb{P}$-full set such that $\varphi(t, \omega)G \subset G$ for all $t \in \mathbb{T}^+$ and $\omega \in \Omega$. For any $\omega \in \hat{\Omega} \cap O_G$ and $t \in \mathbb{T}^+$, if we let $x \in G$ be a point that is asymptotically stable under ω, then $\varphi(t, \omega)x$ is a point in G that is asymptotically stable under $\theta^t\omega$, and so $\theta^t\omega \in O_G$. So $\theta^t(\hat{\Omega} \cap O_G) \subset O_G$ for all $t \in \mathbb{T}^+$; hence, by Lemma 2.8, $\mathbb{P}(O_G) \in \{0, 1\}$. So then, G admits stable trajectories if and only if $\mathbb{P}(O_G) = 1$. □

Lemma 4.39. *Let ρ be an ergodic probability measure of the Markov transition probabilities $(\varphi_x^t)_{x \in X, t \in \mathbb{T}^+}$. Then φ is stable with respect to ρ if and only if* $\operatorname{supp} \rho$ *admits stable trajectories.*

Proof. It is clear that if φ is stable with respect to ρ then $\operatorname{supp} \rho$ admits stable trajectories. Now suppose $\operatorname{supp} \rho$ admits stable trajectories. So $\mathbb{P}(O_{\operatorname{supp} \rho}) > 0$; and obviously, for each $\omega \in O_{\operatorname{supp} \rho}$ there is a ρ-positive measure set of points that are asymptotically stable under ω. Hence $\mathbb{P} \otimes \rho(O) > 0$, and so by Lemma 4.29, $\mathbb{P} \otimes \rho(O) = 1$. □

The following result will play a key role when we come to study "stable synchronisation" on compact spaces.

Proposition 4.40. *Let $C \subset X$ be a compact invariant set, and suppose that C contains only one minimal set K. Then φ is uniformly stable in C if and only if K admits stable trajectories; and in this case, for every $x \in C$,*

$$\mathbb{P}(\omega : d(\varphi(t, \omega)x, K) \to 0 \text{ as } t \to \infty) = 1.$$

In the proof of Proposition 4.40 (and also later on), we will use the following elementary lemma (which, heuristically, will play the role of a "strong Markov" property) in conjunction with Corollary 2.81 (which, heuristically, will generate a "random time at which to apply the strong Markov property"):

Lemma 4.41. *Let D be a countable set, and let $\leq$ be a linear order on D. Suppose we have, for each $s \in D$ and $n \in \mathbb{N}$, events $R_{n,s}, S_{n,s} \in \mathcal{F}$ with the following properties:*

- *for all n and s, $S_{n,s}$ is independent of $\sigma(R_{n,t} : t \leq s)$;*
- *for all n, $\mathbb{P}(\bigcup_{s\in D} R_{n,s}) = 1$;*
- $\inf_{s\in D} \mathbb{P}(S_{n,s}) \to 1$ *as $n \to \infty$.*

Then

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty}\bigcup_{s\in D} R_{n,s} \cap S_{n,s}\right) = 1.$$

Proof. We write "$s < t$" to mean "$s \leq t$ and $s \neq t$". First fix $n \in \mathbb{N}$. Since $\mathbb{P}(\bigcup_{s\in D} R_{n,s}) = 1$ and D is countable, we must have that for all $\varepsilon > 0$ there exist $t_1 < \ldots < t_m$ in D such that $\mathbb{P}(\bigcup_{i=1}^{m} R_{n,t_i}) > 1 - \varepsilon$, and so

$$\begin{aligned}
\mathbb{P}\left(\bigcup_{s\in D} R_{n,s} \cap S_{n,s}\right) &\geq \mathbb{P}\left(\bigcup_{i=1}^{m} R_{n,t_i} \cap S_{n,t_i}\right) \\
&\geq \mathbb{P}\left(\bigcup_{i=1}^{m}\left(R_{n,t_i} \setminus \bigcup_{j=1}^{i-1} R_{n,t_j}\right) \cap S_{n,t_i}\right) \quad \left(\text{where } \bigcup_{j=1}^{0} R_{n,t_j} := \varnothing\right) \\
&= \sum_{i=1}^{m} \mathbb{P}\left(R_{n,t_i} \setminus \bigcup_{j=1}^{i-1} R_{n,t_j}\right)\mathbb{P}(S_{n,t_i}) \\
&\geq \left(\sum_{i=1}^{m} \mathbb{P}\left(R_{n,t_i} \setminus \bigcup_{j=1}^{i-1} R_{n,t_j}\right)\right)\inf_{s\in D} \mathbb{P}(S_{n,s}) \\
&= \mathbb{P}\left(\bigcup_{i=1}^{m} R_{n,t_i}\right)\inf_{s\in D} \mathbb{P}(S_{n,s}) \\
&\geq (1-\varepsilon)\inf_{s\in D} \mathbb{P}(S_{n,s}).
\end{aligned}$$

This is true for all ε, and so

$$\mathbb{P}\left(\bigcup_{s\in D} R_{n,s} \cap S_{n,s}\right) \geq \inf_{s\in D} \mathbb{P}(S_{n,s}).$$

The desired result then follows from the fact that $\inf_{s\in D} \mathbb{P}(S_{n,s}) \to 1$ as $n \to \infty$. □

Proof of Proposition 4.40. Suppose K admits stable trajectories. We first show that there exists at least one point in K that is almost surely stable. Since K is compact and is minimal according to the Markov transition probabilities $(\varphi_x^t)_{x\in X,\, t\in\mathbb{T}^+}$, there is an ergodic probability measure ρ of the Markov transition probabilities $(\varphi_x^t)_{x\in X,\, t\in\mathbb{T}^+}$ such that $\operatorname{supp}\rho = K$. By Lemma 4.39, φ is stable with respect to ρ, i.e. ρ-almost every point $x \in K$ is almost surely stable.

Now let $p \in K$ be an almost surely stable point. Fix any $x \in C$, and for each $n \in \mathbb{N}$ and $s \in \mathbb{Q} \cap \mathbb{T}^+$ let

$$\begin{aligned} R_{n,s} &= \left\{ \omega \in \Omega : \varphi(s,\omega)x \in B_{\frac{1}{n}}(p) \right\} \\ S_{n,s} &= \theta^{-s}\left(E_{B_{\frac{1}{n}}(p)} \right). \end{aligned}$$

Note that for every n and s, $\sigma(R_{n,t} : t \leq s) \subset \mathcal{F}_0^s$ and (by Lemma 4.25(A)) $S_{n,s} \in \mathcal{F}_s^\infty$. Corollary 2.81 gives that $\mathbb{P}\left(\bigcup_s R_{n,s}\right) = 1$ for all n. Obviously $\mathbb{P}(S_{n,s}) = P_{\frac{1}{n}}(p)$ for all n and s, and so $\mathbb{P}(S_{n,s}) \to 1$ as $n \to \infty$ uniformly in s. So then, Lemma 4.41 gives that

$$\mathbb{P}\left(\bigcup_n \bigcup_s R_{n,s} \cap S_{n,s} \right) = 1.$$

Now it is clear that $R_{n,s} \cap S_{n,s} \subset O_x$ for all n and s, and so $\mathbb{P}(O_x) = 1$, i.e. x is almost surely stable. Since $x \in C$ was arbitrary, it follows that φ is everywhere stable in C, and therefore, since C is compact, φ is uniformly stable in C. Now let $\tilde{\Omega}$ be a $\mathbb{P}$-full set such that $\varphi(t,\theta^s\omega)K \subset K$ for all $\omega \in \tilde{\Omega}$, $t \in \mathbb{T}^+$ and $s \in \mathbb{Q} \cap \mathbb{T}^+$. Fix any $x \in C$, and still let $R_{n,s}$ and $S_{n,s}$ be as above for all $n \in \mathbb{N}$ and $s \in \mathbb{Q} \cap \mathbb{T}^+$. For any n and s, for any $\omega \in R_{n,s} \cap S_{n,s} \cap \tilde{\Omega}$, we have that

$$d(\,\varphi(t,\omega)x\,,\,\varphi(t-s,\theta^s\omega)p\,) \to 0 \ \text{ as } t \to \infty$$

and therefore

$$d(\varphi(t,\omega)x, K) \to 0 \ \text{ as } t \to \infty.$$

Hence the set $\{\omega : d(\varphi(t,\omega)x, K) \to 0 \text{ as } t \to \infty\}$ is a $\mathbb{P}$-full set, as required. □

Remark 4.42. As a consequence of Proposition 4.40, if X is compact and there is only one stationary probability measure ρ of the Markov transition probabilities $(\varphi_x^t)_{x \in X, t \in \mathbb{T}^+}$, then φ is uniformly stable in X if and only if φ is stable with respect to ρ.

4.5 Stable synchronisation

Recall that for each $\omega \in \Omega$, we define the equivalence relation $\sim_\omega$ on X by

$$x \sim_\omega y \quad \Longleftrightarrow \quad d(\,\varphi(t,\omega)x\,,\,\varphi(t,\omega)y\,) \to 0 \ \text{ as } t \to \infty.$$

Definition 4.43. We will say that φ is *pointwise-stably synchronising* if φ is both synchronising and everywhere stable in X.

Definition 4.44. We will say that φ is (*uniformly*) *stably synchronising* if φ is both synchronising and uniformly stable in X.

Recall that due to Lemma 4.28, if X is compact, then φ is uniformly stable in X if and only if φ is everywhere stable in X. Hence in particular, if X is compact, then φ is stably synchronising if and only if φ is pointwise-stably synchronising.

Example 4.45. Suppose Ω is a singleton $\{\omega\}$, $\mathbb{T} = \mathbb{Z}$ and $X = \mathbb{S}^1$. Let $f : \mathbb{S}^1 \to \mathbb{S}^1$ be an orientation-preserving homeomorphism possessing exactly one fixed point, and define φ by $\varphi(n,\omega) = f^n$. Then φ is synchronising, since every trajectory converges to the fixed point of f; but φ is not (pointwise-)stably synchronsing, since the fixed point of f does not have a neighbourhood that contracts under ω. Indeed, the fixed point of f is not even a Lyapunov stable fixed point. From a practical point of view: ironically, although f is in theory a "synchronising" dynamical system, any (sufficiently accurate) practical implementation of this dynamical system will appear to be "chaotic". For, no matter how close two trajectories start, they will eventually move very close to the unstable fixed point, from which point what happens next will be entirely dictated by unaccounted-for sources of perturbation or random inaccuracies.

Lemma 4.46. *Given a sample point $\omega \in \Omega$ and an open set $U \subset X$, the following statements are equivalent:*

(i) U can be expressed as a union of open sets that contract under ω and are contained in the same equivalence class of $\sim_\omega$;

(ii) there exists an increasing sequence $(V_k)_{k\in\mathbb{N}}$ of open subsets V_k of U such that $\bigcup_{k=1}^{\infty} V_k = U$ and for each k, V_k contracts under ω.

Proof. It is clear that (ii)⇒(i). Now suppose that (i) holds. Since the union of an arbitrary collection of open sets is equal to the union of some countable subcollection thereof, we may write $U = \bigcup_{r=1}^{\infty} W_r$ for some sequence $(W_r)_{r\in\mathbb{N}}$ of open sets that contract under ω. Since all members of the collection $\{W_r\}_{r\in\mathbb{N}}$ belong to the same equivalence class of $\sim_\omega$, it follows that any finite union of members of this collection contracts under ω; so taking $V_k := \bigcup_{r=1}^{k} W_r$, we have that (ii) holds. □

Definition 4.47. When the equivalent statements in Lemma 4.46 hold, we will say that *U is σ-contracting under ω.*

Remark 4.48. Note that if U is σ-contracting under ω then any compact $G \subset U$ contracts under ω. If X is σ-locally compact then the converse statement also holds: if every compact $G \subset U$ contracts under ω then (since U can be expressed as the union of the interiors of countably many compact subsets of U) U is σ-contracting.

Proposition 4.49. *φ is pointwise-stably synchronising if and only if there is a $\mathbb{P}$-full set $\tilde{\Omega} \subset \Omega$ and a $\tilde{\Omega}$-indexed family $(U(\omega))_{\omega\in\tilde{\Omega}}$ of dense open subsets of X such that*

(i) for each $\omega \in \tilde{\Omega}$, $U(\omega)$ is σ-contracting under ω, and

(ii) for each $x \in X$, the set $\{\omega \in \tilde{\Omega} : x \in U(\omega)\}$ is a $\mathbb{P}$-full set.

Proof. First suppose that φ is pointwise-stably synchronising. Let S be a countable dense subset of X. Let $\tilde{\Omega}$ be a $\mathbb{P}$-full set such that for all $\omega \in \tilde{\Omega}$,

(a) for all $x, y \in S$, $x \sim_\omega y$, and

(b) every $x \in S$ has a neighbourhood that contracts under ω.

For each $\omega \in \tilde{\Omega}$, let $U(\omega)$ be the union of all open sets that contract under ω. Since every non-empty open set intersects S, we have in particular that all open sets contracting under ω are contained in the same equivalence class that contains S; hence U is σ-contracting. Obviously, for any $x \in X$, since the set $\{\omega \in \Omega : x \text{ is asymptotically stable under } \omega\}$ is a $\mathbb{P}$-full set, it follows that the set $\{\omega \in \tilde{\Omega} : x \in U(\omega)\}$ is a $\mathbb{P}$-full set.

Now in the converse direction, suppose we have $\tilde{\Omega}$ and $(U(\omega))_{\omega\in\tilde{\Omega}}$ as in the statement of the proposition. For any two points $x, y \in X$, we have that for $\mathbb{P}$-almost all ω, x and y both belong to $U(\omega)$ and therefore $x \sim_\omega y$. So φ is synchronising. Moreover, for any $x \in X$, for any $\omega \in \tilde{\Omega}$ with $x \in U(\omega)$, since $U(\omega)$ can be expressed as a union of open sets that contract under ω, we have in particular that x has a neighbourhood that contracts under ω. So φ is everywhere stable in X. □

Definition 4.50. We say that φ is *globally contractive* if for $\mathbb{P}$-almost all $\omega \in \Omega$, every non-empty bounded subset of X contracts under ω.

Obviously if φ is globally contractive then φ is stably synchronising.

Remark 4.51. As a consequence of Remark 2.33, if φ is globally contractive then φ has at most one random fixed point $q : \Omega \to X$ (up to $\mathbb{P}$-almost sure equality). Moreover, such a random fixed point, if it exists, must have a modification which is $\mathcal{F}^0_{-\infty}$-measurable.[7] Hence, if φ is globally contractive, then the existence of a random fixed point is equivalent to the existence of a stationary probability measure for the Markov transition probabilities (φ^t_x). When φ is globally contractive and a random fixed point $q : \Omega \to X$ does exist, q is sometimes said to be a *globally forward-attracting* random fixed point. This is in contrast to the notion of a *globally pullback-attracting* random fixed point (defined in Remark 3.23), which concerns dynamics in the *past*. (Note, however, that both of these serve as *globally weakly attracting* random fixed points, as defined in Remark 2.33.)

We now go on to consider "ρ-almost stable synchronisation". The following important result is essentially Proposition 3 of [LeJ87]:

Theorem 4.52. *Let ρ be an ergodic probability measure of the Markov transition probabilities $(\varphi^t_x)_{x\in X, t\in\mathbb{T}^+}$ such that φ is stable with respect to ρ. Let n be the ρ-clustering number of φ. Then $n < \infty$, and for $\mathbb{P}$-almost every $\omega \in \Omega$ there exist mutually disjoint sets open sets $U_1, \ldots, U_n \subset X$ such that the following holds: for each $i \in \{1, \ldots, n\}$, $\rho(U_i) = \frac{1}{n}$ and U_i is σ-contracting under ω; but for any distinct $i, j \in \{1, \ldots, n\}$, for any $x \in U_i$ and $y \in U_j$, $x \not\sim_\omega y$.*

As an immediate corollary of Theorem 4.52 and Lemma 4.7, we have:

Corollary 4.53. *Let ρ be an ergodic probability measure of the Markov transition probabilities $(\varphi^t_x)_{x\in X, t\in\mathbb{T}^+}$. Then the following statements are equivalent:*

(i) φ is both stable with respect to ρ, and ρ-almost everywhere synchronising;

[7] To see this: Fix any $c \in X$. Since $d(\varphi(n,\omega)c, q(\theta^n\omega)) \to 0$ as $n \to \infty$ for $\mathbb{P}$-almost all $\omega \in \Omega$, it follows that the stochastic process $d(\varphi(n, \theta^{-n}\cdot)c, q(\cdot))$ converges in probability to 0, and therefore there exists an increasing sequence (m_n) in $\mathbb{N}$ such that $\varphi(m_n, \theta^{-m_n}\omega)c \to q(\omega)$ for $\mathbb{P}$-almost all $\omega \in \Omega$.

(ii) *φ is both stable with respect to ρ, and statistically synchronising with respect to ρ;*

(iii) *for $\mathbb{P}$-almost every $\omega \in \Omega$ there is an open ρ-full set that is σ-contracting under ω.*

Obviously, (i) is the same as saying that there exists a ρ-full set $A \in \mathcal{B}(X)$ such that any point $x \in A$ is almost surely stable and any two points $x, y \in A$ synchronise almost surely.

Definition 4.54. Let ρ be an ergodic probability measure of the Markov transition probabilities $(\varphi_x^t)_{x \in X, t \in \mathbb{T}^+}$. We say that φ is *ρ-almost stably synchronising* if the equivalent statements in Corollary 4.53 hold.

Example 4.55. Let $X = \mathbb{S}^1$, which we identify with $\mathbb{R}/\mathbb{Z}$ (with $\pi : \mathbb{R} \to \mathbb{S}^1$ denoting the associated projection). Let φ be the RDS such that $\varphi(t,\omega)\pi(0) = \pi(0)$ for all t and ω, and on $\mathbb{S}^1 \setminus \{\pi(0)\}$, φ agrees with the projection onto $\mathbb{S}^1$ of the RDS in Example 3.19. Then φ is not synchronising, but φ *is* l-almost stably synchronising (where l is the Lebesgue measure on $\mathbb{S}^1$).

Let us now give a proof of Theorem 4.52. We start with the following lemma:

Lemma 4.56. *Fix any $n \in \mathbb{N}$, and let (μ_ω) be a φ-invariant random probability measure such that for $\mathbb{P}$-almost all $\omega \in \Omega$, μ_ω takes the form*

$$\mu_\omega = \frac{1}{n} \sum_{y \in A(\omega)} \delta_y$$

for some set $A(\omega) \subset X$ with $|A(\omega)| = n$. Then for $\mathbb{P}$-almost every $\omega \in \Omega$ for which μ_ω takes this form, for any distinct $y_1, y_2 \in A(\omega)$, y_1 and y_2 belong to different equivalence classes of $\sim_\omega$.

Proof. The statement is vacuously true if $n = 1$, so assume $n \geq 2$. Let $\tilde{\Omega} \in \mathcal{F}$ be a $\mathbb{P}$-full set such that for all $\omega \in \tilde{\Omega}$, μ_ω takes the form described in Lemma 4.56. For each $k \in \mathbb{N}$ let

$$E_k := \{\omega \in \Omega : \text{there exists } x \in X \text{ s.t. } \mu_\omega(B_{\frac{1}{k}}(x)) > \tfrac{1}{n}\}.$$

Note that E_k is measurable: for any countable dense set $S \subset X$, due to Lemma A.13 we have that

$$E_k = \bigcup_{x \in S} \{\omega \in \Omega : \mu_\omega(B_{\frac{1}{k}}(x)) > \tfrac{1}{n}\}.$$

Now for any k, for any $\omega \in \tilde{\Omega} \cap E_k$, there must exist disinct points $y_1, y_2 \in A(\omega)$ such that $d(y_1, y_2) < \frac{2}{k}$. Hence it is clear that $\tilde{\Omega} \cap \bigcap_k E_k = \emptyset$. So then, $\mathbb{P}\left(\bigcap_k E_k\right) = 0$, and therefore $\mathbb{P}(E_k) \to 0$ as $k \to \infty$. So letting

$$E := \bigcap_{k=1}^{\infty} \bigcup_{i=0}^{\infty} \bigcap_{j=i}^{\infty} \theta^{-j}(E_k),$$

we have that $\mathbb{P}(E) = 0$. Now for any $j, k \in \mathbb{N}$, for any $\omega \in \theta^{-j}(\tilde{\Omega})$, if there exist disinct points $y_1, y_2 \in A(\theta^j\omega)$ such that $d(y_1, y_2) < \frac{1}{k}$, then $\omega \in \theta^{-j}(E_k)$. Accordingly, for any $\omega \in \bigcap_{j=0}^{\infty} \theta^{-j}(\tilde{\Omega})$, if

$$\min(\, d(y_1, y_2) : \text{distinct } y_1, y_2 \in A(\theta^j\omega)\,) \to 0 \quad \text{as } j \to \infty$$

then $\omega \in E$. Since (μ_ω) is φ-invariant, we have that for $\mathbb{P}$-almost all $\omega \in \bigcap_{j=0}^{\infty} \theta^{-j}(\tilde{\Omega})$, for every $j \in \mathbb{N}_0$, $A(\theta^j\omega) = \varphi(j,\omega)A(\omega)$. So then, since $\mathbb{P}(E) = 0$, we have that for $\mathbb{P}$-almost every $\omega \in \bigcap_{j=0}^{\infty} \theta^{-j}(\tilde{\Omega})$,

$$\min(\, d(y_1,y_2) \,:\, \text{distinct } y_1,y_2 \in \varphi(j,\omega)A(\omega)\,) \;\not\to\; 0 \quad \text{as } j \to \infty$$

and hence in particular, there do not exist distinct $y_1, y_2 \in A(\omega)$ such that $d(\varphi(j,\omega)y_1, \varphi(j,\omega)y_2) \to 0$ as $j \to \infty$. Thus we have shown that for $\mathbb{P}$-almost all $\omega \in \tilde{\Omega}$, any distinct points $y_1, y_2 \in A(\omega)$ must belong to different equivalence classes of $\sim_\omega$. □

Proof of Theorem 4.52. For $\mathbb{P}$-almost every $\omega \in \Omega$, there exist open sets intersecting $\operatorname{supp}\rho$ which contract under ω (and are therefore contained in equivalence classes of $\sim_\omega$). Hence, by Lemma 4.12, $n < \infty$.

To prove the rest of the theorem, it is sufficient to find, for each $\varepsilon > 0$, a $\mathbb{P}$-full set $\Omega_\varepsilon \in \mathcal{F}$ of sample points ω with the property that there exist open sets $W_1, \ldots, W_n \subset X$ such that for each $i \in \{1,\ldots,n\}$, $\rho(W_i) > \frac{1}{n} - \varepsilon$ and W_i contracts under ω, and moreover such that the sets $W_1, \ldots, W_n$ are contained in distinct equivalence classes of $\sim_\omega$.

(The fact that this is sufficient for the desired result to be true is visually intuitively clear; still, one way to prove it rigorously is as follows: Let $\Omega^* := \bigcap_{j=1}^{\infty} \Omega_{\frac{1}{j}}$. Clearly, $\mathbb{P}(\Omega^*) = 1$; now fix any $\omega \in \Omega^*$. For each integer $j \geq n(n+1)$, let $W_1^{(j)}, \ldots, W_n^{(j)} \subset X$ be open sets contracting under ω, each of measure greater than $\frac{1}{n} - \frac{1}{j}$ according to ρ, and belonging to distinct equivalence classes of $\sim_\omega$. Note in particular that $\rho(W_i^{(j)}) > \frac{1}{n+1}$ for each $j \geq n(n+1)$ and $i \in \{1,\ldots,n\}$. Hence, it is not hard to see that for each $j \geq n(n+1)$ and $i \in \{1,\ldots,n\}$, there exists a unique $\pi_j(i) \in \{1,\ldots,n\}$ such that $W_i^{(n(n+1))} \cap W_{\pi_j(i)}^{(j)} \neq \varnothing$. So for each $i \in \{1,\ldots,n\}$, define $U_i := \bigcup_{j=n(n+1)}^{\infty} W_{\pi_j(i)}^{(j)}$. Then U_i is σ-contracting under ω for each i, and the sets $U_1, \ldots, U_n$ are contained in distinct equivalence classes of $\sim_\omega$. Moreover, it is clear that $\rho(U_i) \geq \frac{1}{n}$ for each i, and therefore $\rho(U_i) = \frac{1}{n}$ for each i.)

Let μ be the unique Markov invariant measure whose X-projection coincides with ρ, and let (μ_ω) be a disintegration of μ. On the basis of Lemma 4.56, let $\tilde{\Omega} \in \mathcal{F}$ be a $\mathbb{P}$-full set such that for all $\omega \in \Omega$, μ_ω takes the form

$$\mu_\omega \;=\; \frac{1}{n} \sum_{y \in A(\omega)} \delta_y$$

for some set $A(\omega) \subset X$ with $|A(\omega)| = n$ such that the elements of $A(\omega)$ belong to distinct equivalence classes of $\sim_\omega$. Now let $\mathcal{U}$ be a countable base for the topology on X that is closed under finite unions, and let $\mathfrak{U}$ denote the collection of all mutually disjoint subcollections of $\mathcal{U}$ of size n. For each $\varepsilon > 0$, let

$$\mathfrak{U}^\varepsilon \;:=\; \{\, \mathcal{V} \in \mathfrak{U} \,:\, \forall\, W \in \mathcal{V},\ \rho(W) > \tfrac{1}{n} - \varepsilon \,\},$$

let

$$\tilde{\Omega}_\varepsilon \;:=\; \bigcup_{\mathcal{V} \in \mathfrak{U}^\varepsilon} \bigcap_{W \in \mathcal{V}} \left(E_W \,\cap\, \{\omega \in \Omega : \mu_\omega(W) = \tfrac{1}{n}\} \right),$$

and let $\Omega_\varepsilon := \tilde{\Omega}_\varepsilon \cap \tilde{\Omega}$. Note that every $\omega \in \Omega_\varepsilon$ has the desired property given further above: since $\omega \in \tilde{\Omega}_\varepsilon$, there exist mutually disjoint open sets $W_1, \ldots, W_n$ contracting under ω, each of measure greater than $\frac{1}{n} - \varepsilon$ according to ρ and measure equal to $\frac{1}{n}$ according to μ_ω; and so, since $\omega \in \tilde{\Omega}$, it follows that the sets $W_1, \ldots, W_n$ are contained in distinct equivalence classes of $\sim_\omega$. So then, it only remains to show that $\mathbb{P}(\Omega_\varepsilon) = 1$ for every $\varepsilon > 0$.

Since $O \in \mathcal{F}_0^\infty \otimes \mathcal{B}(X)$, we have that $\mu(O) = \mathbb{P} \otimes \rho(O) = 1$, and so for $\mathbb{P}$-almost all $\omega \in \Omega$, $\mu_\omega(x : (\omega, x) \in O) = 1$. Let $\hat{\Omega} \in \mathcal{F}$ be a $\mathbb{P}$-full set such that for every $\omega \in \hat{\Omega}$ the following statements hold:

(i) for all $j \in \mathbb{N}_0$, $\theta^{-j}\omega \in \tilde{\Omega}$ and $\varphi(j, \theta^{-j}\omega)_* \mu_{\theta^{-j}\omega} = \mu_\omega$;

(ii) $\varphi(j, \theta^{-j}\omega)_* \rho$ converges weakly to μ_ω as j tends to ∞ in $\mathbb{N}$;

(iii) $\mu_\omega(x : (\omega, x) \in O) = 1$.

Note that statement (i) implies that for every $j \in \mathbb{N}$, $\varphi(j, \theta^{-j}\omega)$ maps $A(\theta^{-j}\omega)$ bijectively into $A(\omega)$. Also note that statement (iii) is equivalent to saying that every point in $A(\omega)$ is asymptotically stable under O.

Now fix any ε. We will show that

$$\hat{\Omega} \subset \bigcup_{k=1}^{\infty} \bigcap_{j=k}^{\infty} \theta^j(\Omega_\varepsilon). \tag{4.1}$$

Fix any $\omega \in \hat{\Omega}$. Let us write $y_1, \ldots, y_n$ for the elements of $A(\omega)$. Let $V_1, \ldots, V_n$ be open neighbourhoods of $y_1, \ldots, y_n$ respectively, that are mutually disjoint and each contract under ω. Since $\varphi(j, \theta^{-j}\omega)_* \rho$ converges weakly to μ_ω as j tends to ∞ in $\mathbb{N}$, and $\mu_\omega(V_i) = \frac{1}{n}$ for each $1 \le i \le n$, there must exist $k \in \mathbb{N}$ such that for all $j \ge k$, $\rho(\varphi(j, \theta^{-j}\omega)^{-1}(V_i)) > \frac{1}{n} - \varepsilon$ for each i. Now fix any $j \ge k$. For each $1 \le i \le n$, since the base $\mathcal{U}$ is closed under finite unions, $\varphi(j, \theta^{-j}\omega)^{-1}(V_i)$ contains a set $W_i \in \mathcal{U}$, itself containing $\varphi(j, \theta^{-j}\omega)^{-1}(\{y_i\})$, such that $\rho(W_i) > \frac{1}{n} - \varepsilon$. Moreover, the sets $W_1, \ldots, W_n$ are obviously mutually disjoint, and each contract under $\theta^{-j}\omega$; and since each W_i contains exactly one element of $A(\theta^{-j}\omega)$, we have that $\mu_{\theta^{-j}\omega}(W_i) = \frac{1}{n}$ for each i. Hence $\theta^{-j}\omega \in \tilde{\Omega}_\varepsilon$. And by assumption, $\theta^{-j}\omega \in \tilde{\Omega}$; so $\theta^{-j}\omega \in \Omega_\varepsilon$. Thus we have proved (4.1).

So then, for any $\eta > 0$, letting $k \in \mathbb{N}$ be sufficiently large that $\mathbb{P}\left(\bigcap_{j=k}^{\infty} \theta^j(\Omega_\varepsilon)\right) > 1 - \eta$, we have that

$$\mathbb{P}(\Omega_\varepsilon) = \mathbb{P}(\theta^k(\Omega_\varepsilon)) > 1 - \eta.$$

Since η was arbitrary, we are done. □

We now mention the case of monotone RDS on a one-dimensional phase space. (We still assume that φ is right-continuous.)

We will say that a set $A \subset \mathbb{R}$ is *endpoint-complete* if the following statements both hold:

(i) if A is bounded above then $\max A$ exists;

(ii) if A is bounded below then $\min A$ exists.

In other words, A is endpoint-complete if and only if the convex hull of A (relative to $\mathbb{R}$) is a closed subset of $\mathbb{R}$.

Lemma 4.57. *Suppose X is a Borel-measurable subset of $\mathbb{R}$, with d being the standard metric. Suppose that φ is monotone (with respect to the usual order $\leq$). Let ρ be any probability measure on X with $X_\rho = X$. The following statements are equivalent:*

- *(i) $\mathbb{P}$-almost every $\omega \in \Omega$ has the property that for any $a, b \in X$ with $a \leq b$, $[a,b] \cap X$ contracts under ω;*
- *(ii) φ is pointwise-stably synchronising;*
- *(iii) φ is synchronising;*
- *(iv) φ is ρ-almost everywhere synchronising.*

If X is endpoint-complete, then these are in turn equivalent to:

- *(v) φ is globally contractive;*
- *(vi) φ is stably synchronising.*

Proof. It is clear that (i)⇒(ii)⇒(iii)⇒(iv). We now show that (iv)⇒(i). Suppose that (iv) holds. Fix any $\omega \in \Omega$ with the property that $\sim_\omega$ has a ρ-full equivalence class. Let $\tilde{X}$ be the ρ-full equivalence class of $\sim_\omega$. Now fix any $a, b \in X$ with $a \leq b$, and let $I := [a,b] \cap X$. Since X_ρ is the whole of X, the sets $(-\infty, a] \cap X$ and $[b, \infty) \cap X$ both have positive measure under ρ; hence these sets both have non-empty intersection with $\tilde{X}$. So there exist $x_1 \leq a$ and $x_2 \geq b$ such that $x_1 \sim_\omega x_2$. But since φ is monotone, we also have that $\varphi(t,\omega)I \subset [\varphi(t,\omega)x_1, \varphi(t,\omega)x_2]$ for all $t \in \mathbb{T}^+$. It follows, therefore, that I contracts under ω, as required.

Obviously (v)⇒(vi)⇒(ii). Finally, if X is endpoint-complete then (i)⇒(v), since it is clear that every bounded set $B \subset X$ is contained in some closed interval $[a,b]$ with $a, b \in X$. □

Example 4.58. The RDS in Example 2.34 is globally contractive, as is the RDS in Example 2.37 for $\alpha < 0$. The RDS in Example 3.19 (which has been shown to be synchronising in [AM14]) is pointwise-stably synchronising, but not uniformly stably synchronising. The RDS generated by (3.2) in Example 3.21 is globally contractive; see Example 6.7.

Proposition 4.59. *Suppose X is a Borel-measurable subset of $\mathbb{R}$, with d being the standard metric. Suppose that φ is monotone (with respect to the usual order $\leq$). Suppose moreover that there exists an ergodic probability measure ρ of the Markov transition probabilities $(\varphi_x^t)_{x \in X,\, t \in \mathbb{T}^+}$ such that $X_\rho = X$. Then φ is synchronising if and only if φ is stable with respect to ρ.*

Proof. If φ is synchronising then, in particular, φ is pointwise-stably synchronising (by Lemma 4.57); so φ is everywhere stable in X, and therefore φ is stable with respect to ρ. Conversely: we know by Theorem 3.13 that φ is statistically synchronising with respect to ρ; so if, in addition, φ is stable with respect to ρ, then by Corollary 4.53 we have that φ is ρ-almost everywhere synchronising, and therefore (by Lemma 4.57) φ is synchronising. □

4.6 Contractibility

Much of the remainder of this thesis will be devoted to *tests* for synchronising behaviour. All of these tests will involve the notion of "contractibility" of pairs of points, and so we will introduce this notion now.

Note that for any invariant set $A \subset X$, $A \times A$ is invariant under the two-point motion $\varphi^{\times 2}$.

Definition 4.60. Given points $x, y, p \in X$, we will say that (x,y) *is contractible towards* p if every neighbourhood of (p,p) in $X \times X$ is accessible from (x,y) under the two-point motion $\varphi^{\times 2}$. This is the same as saying that for every $\varepsilon > 0$,

$$\mathbb{P}(\,\omega : \exists\, t \in \mathbb{T}^+ \text{ s.t. } \varphi(t,\omega)x, \varphi(t,\omega)y \in B_\varepsilon(p)\,) \;>\; 0.$$

Proposition 4.61. *The set* $\mathfrak{C} = \{\,(p,x,y) \in X \times X \times X : (x,y)$ *is contractible toward* $p\,\}$ *is* $\mathcal{B}(X \times X \times X)$*-measurable.*

Proof. Given any metrisation $\tilde{d}$ of the product topology on $X \times X$, we have that

$$\mathfrak{C} \;=\; \{\,(p,x,y) : \tilde{d}((p,p)\,,\, G_{(x,y)}) \;=\; 0\,\}.$$

Hence the result follows from Lemma 2.77 applied to $\varphi^{\times 2}$. □

Definition 4.62. Given points $x, y \in X$ and a set $A \subset X$, we will say that (x,y) *is contractible towards* A if every neighbourhood of Δ_A in $X \times X$ is accessible from (x,y) under $\varphi^{\times 2}$.

Lemma 4.63. *For any* $x, y \in X$ *and* $A \subset X$*,* (x,y) *is contractible towards* A *if and only if there exists* $p \in A$ *such that* (x,y) *is contractible towards* p*.*

Proof. (x,y) is contractible towards A if and only if $G_{(x,y)}$ intersects every neighbourhood of Δ_A; but since $G_{(x,y)}$ is closed, this is equivalent to saying that $G_{(x,y)}$ intersects Δ_A. Obviously this is the same as saying that there exists $p \in A$ such that $(p,p) \in G_{(x,y)}$, which is equivalent to saying that there exists $p \in A$ such that every neighbourhood of (p,p) is accessible from (x,y) under $\varphi^{\times 2}$. So we are done. □

Definition 4.64. Given points $x, y \in X$, we say that (x,y) *is generally contractible* if (x,y) is contractible towards X (i.e. if there exists $p \in X$ such that (x,y) is contractible towards p).

Note that, given a closed invariant set $G \subset X$ and any points $x, y \in G$, if (x,y) is generally contractible then (x,y) is contractible towards G. (The reason is that, since $G \times G$ is a closed invariant set under $\varphi^{\times 2}$, (x,y) cannot be contractible towards any point $p \in X \setminus G$.)

Lemma 4.65. *Given a compact invariant set* $K \subset X$ *and points* $x, y \in K$*,* (x,y) *is generally contractible if and only if for every* $\varepsilon > 0$*,*

$$\mathbb{P}(\,\omega : \exists\, t \in \mathbb{T}^+ \textit{ s.t. } d(\varphi(t,\omega)x, \varphi(t,\omega)y) < \varepsilon\,) \;>\; 0. \tag{4.2}$$

Proof. The "only if" direction is clear (and has nothing to do with the set K). For the "if" direction: For each $\varepsilon > 0$, let

$$U^{K,\varepsilon} := \{(u,v) \in K \times K : d(u,v) < \varepsilon\}.$$

Suppose that (4.2) holds for every $\varepsilon > 0$. Since $K \times K$ is invariant under $\varphi^{\times 2}$, it follows that there is a positive-measure set of sample points ω with the property that at some time t, $\varphi^{\times 2}(t,\omega)(x,y) \in U^{K,\varepsilon}$. Now for every neighbourhood U of Δ_K, as in the proof of Lemma 4.1 there exists $\varepsilon > 0$ such that $U^{K,\varepsilon} \subset U$; and so U is accessible from (x,y). □

The following concept has been considered in [BS88] and [Bax91].

Definition 4.66. Let $K \subset X$ be a compact invariant set. We will say that φ *is contractible*[8] *on* K if for all distinct $x, y \in K$,

$$\mathbb{P}(\,\omega : \exists\, t \in \mathbb{T}^+ \text{ s.t. } d(\varphi(t,\omega)x, \varphi(t,\omega)y) < d(x,y)\,) \;>\; 0.$$

Remark 4.67. Suppose there exists a separable metrisable topology on Ω whose Borel σ-algebra coincides with $\mathcal{F}$, such that $\mathbb{P}$ has full support and for all $t \in \mathbb{T}^+$ and $x \in X$, the map $\omega \mapsto \varphi(t,\omega)x$ is continuous. Then, as in Remark 2.73, in order to show that φ is contractible on a compact invariant set K, it is sufficient to show that for each pair of distinct points $x, y \in K$ there exists a sample point $\omega \in \Omega$ and a time $t \in \mathbb{T}^+$ such that $d(\varphi(t,\omega)x, \varphi(t,\omega)y) < d(x,y)$.

Proposition 4.68 (cf. [BS88, Proposition 4.1])**.** *Let $K \subset X$ be a compact invariant set. The following statements are equivalent:*

(i) every pair $(x,y) \in K \times K$ is generally contractible;

(ii) φ is contractible on K;

(iii) $(K \times K) \setminus \Delta_K$ contains no non-empty closed[9] *invariant sets (under $\varphi^{\times 2}$);*

(iv) given any two points $x, y \in K$, for $\mathbb{P}$-almost every $\omega \in \Omega$ there exists an unbounded increasing sequence (t_n) in $\mathbb{T}^+ \cap \mathbb{Q}$ such that $d(\varphi(t_n,\omega)x, \varphi(t_n,\omega)y) \to 0$ as $n \to \infty$.

Proof. (i)⇒(ii) is clear. We next show (ii)⇒(iii). Suppose (iii) does not hold, and let $C \subset (K \times K) \setminus \Delta_K$ be a non-empty compact invariant set. Let $(x,y) \in C$ be a point which minimises the function $(u,v) \mapsto d(u,v)$ on $C \times C$. Then it is clear that (x,y) is not generally contractible; so (ii) does not hold. Now suppose that (iii) holds. To show that (iv) holds, it is sufficient to show that for each $x, y \in K$ and $k \in \mathbb{N}$, for $\mathbb{P}$-almost all $\omega \in \Omega$ there exist arbitrarily large times t at which $d(\varphi(t,\omega)x, \varphi(t,\omega)y) < \frac{1}{k}$. So fix x, y and k, and let $U := \{(u,v) \in X \times X : d(u,v) < \frac{1}{k}\}$. Then $(K \times K) \setminus U$ is a compact set containing no non-empty closed invariant sets, and so by Lemma 2.79, for $\mathbb{P}$-almost all $\omega \in \Omega$ there exist arbitrarily large times t at which $\varphi^{\times 2}(t,\omega)(x,y) \notin (K \times K) \setminus U$. But since $K \times K$ is itself invariant, it follows that for $\mathbb{P}$-almost all $\omega \in \Omega$ there exist arbitrarily large times t at which $\varphi^{\times 2}(t,\omega)(x,y) \in U$, as required. Finally, (iv)⇒(i) follows immediately from Lemma 4.65. □

[8] In [New15b], the term "two-point contractible" is used.

[9] that is, closed in $X \times X$; the statement that $(K \times K) \setminus \Delta_K$ contains no invariant sets that are closed relative to $(K \times K) \setminus \Delta_K$ would be much stronger.

Lemma 4.69. *If K is a compact invariant set on which φ is contractible, then K contains only one minimal set.*

Proof. Let K be a compact invariant set containing two distinct minimal sets M_1 and M_2. Note that M_1 and M_2 are mutually disjoint. Let (x,y) be a point in $M_1 \times M_2$ which minimises the function $(u,v) \mapsto d(u,v)$ on $M_1 \times M_2$. Then x and y are distinct points, and (x,y) is not generally contractible. □

Now in one of the synchronisation tests that we will present later on (Theorem 6.1), one of the conditions involved is contractibility on a certain compact minimal set. In some situations, it may not be easy to verify directly that φ is contractible on a given set K; but it may be easier to verify that φ is "contractible on a full measure open subset" of K—in which case, if K is minimal, φ must be contractible on the whole of K. To be precise:

Proposition 4.70. *Let ρ be a stationary probability measure of the Markov transition probabilities $(\varphi_x^t)_{x \in X,\, t \in \mathbb{T}^+}$ such that $\operatorname{supp}\rho$ is a compact minimal set. Suppose there exists $A \subset \operatorname{supp}\rho$ such that:*

- $\rho(A) = 1$*;*
- *the interior of A relative to $\operatorname{supp}\rho$ is non-empty;*
- *every pair in $A \times A$ is generally contractible.*

Then φ is contractible on $\operatorname{supp}\rho$.

Proposition 4.70 is adapted from [New15b, Proposition 3.1.2].

Proof. Let $K := \operatorname{supp}\rho$. Fix any $x,y \in K$; we will show that (x,y) is generally contractible. This is equivalent to showing that $G_{(x,y)} \cap \Delta_K \neq \emptyset$. Since $G_{(u,v)} \cap \Delta_K \neq \emptyset$ for all $(u,v) \in A \times A$, it is sufficient just to show that $G_{(x,y)}$ contains at least one point in $A \times A$.

By Lemma 2.82, since $K \times K$ is compact, the image of $G_{(x,y)}$ under the projection $(u,v) \mapsto u$ is precisely G_x; but this is itself equal to K, since K is minimal. Now let D be a countable dense subset of $\mathbb{T}^+$, and let

$$B \;:=\; \{x \in K : \text{for all } t \in D,\ \varphi_x^t(A) = 1\}.$$

Since ρ is stationary and $\rho(A) = 1$, we have that $\rho(B) = 1$—and so, in particular, B is non-empty. So let us fix a point $(u,v) \in G_{(x,y)}$ with $u \in B$. Let $U \subset X$ be an open set such that $U \cap K$ is a non-empty subset of A; since K is minimal, U is acessible from v. Since the map $t \mapsto \varphi_v^t(U)$ is right lower semicontinuous, there must exist $t^* \in D$ such that $\varphi_v^{t^*}(U) > 0$. Since K is invariant, it follows that $\varphi_v^{t^*}(A) > 0$. So then, there exists a $\mathbb{P}$-positive-measure set of sample points ω such that $\varphi(t^*,\omega)u$ and $\varphi(t^*,\omega)v$ are both in A. Hence $G_{(u,v)}$ has non-trivial intersection with $A \times A$, and therefore $G_{(x,y)}$ has non-trivial intersection with $A \times A$. □

4.7 Deterministic-rate synchronisation

In this section, we consider the issue of whether one can give an upper bound on how long one has to wait in order to observe synchronisation of trajectories of a RDS.

Definition 4.71. Given $x, y \in X$, we say that *x and y synchronise at a deterministic rate* if there exists a function $h\colon \mathbb{T}^+ \to [0,\infty]$, with $h(t) \to 0$ as $t \to \infty$, such that for $\mathbb{P}$-almost every $\omega \in \Omega$, for all $t \in \mathbb{T}^+$, $d(\varphi(t,\omega)x, \varphi(t,\omega)y) \le h(t)$. Given such a function h, we say that *x and y synchronise at least as quickly as h.*

Now when two deterministic processes do not synchronise in the absence of noise, it is normal that noise can cause the processes to synchronise, but such synchronisation cannot be expected to occur at any deterministic rate. This is indicated by the following:

Proposition 4.72. *Suppose that (as in Remark 2.73) there exists a separable metrisable topology on Ω whose Borel σ-algebra coincides with $\mathcal{F}$, such that $\mathbb{P}$ has full support and the map $\omega \mapsto \varphi(t,\omega)x$ is continuous for each t and x. Suppose we have $x, y \in X$ and $\omega_0 \in \Omega$ such that $x \not\sim_{\omega_0} y$. Then x and y do not synchronise at a deterministic rate.*

Proof. Fix any function $h\colon \mathbb{T}^+ \to [0,\infty]$ such that $h(t) \to 0$ as $t \to \infty$. Let $t \in \mathbb{T}^+$ be such that $d(\varphi(t,\omega_0)x, \varphi(t,\omega_0)y) > h(t)$. Then by Remark 2.73 applied to the two-point motion $\varphi^{\times 2}$, $\mathbb{P}(\omega : d(\varphi(t,\omega)x, \varphi(t,\omega)y) > h(t)) > 0$. □

Now if we wish to be able to say that two physical processes will synchronise faster than some given deterministic rate, we may need to take into account that the processes will inevitably be subject to small perturbations not accounted for in the model. This motivates the following definitions:

Definition 4.73. Given $x \in X$, we say that *x is asymptotically stable at a deterministic rate* if there exists a neighbourhood U of x and a function $h\colon \mathbb{T}^+ \to [0,\infty]$, with $h(t) \to 0$ as $t \to \infty$, such that for $\mathbb{P}$-almost every $\omega \in \Omega$, for all $t \in \mathbb{T}^+$, $\mathrm{diam}(\varphi(t,\omega)U) \le h(t)$.

Definition 4.74. We will say that *φ is globally contracting at a deterministic rate* if for every bounded $B \subset X$ there exists a function $h\colon \mathbb{T}^+ \to [0,\infty]$, with $h(t) \to 0$ as $t \to \infty$, such that for $\mathbb{P}$-almost every $\omega \in \Omega$, for all $t \in \mathbb{T}^+$, $\mathrm{diam}(\varphi(t,\omega)B) \le h(t)$.

Note that if $X = \mathbb{R}^d$ (with the usual metric) and every $x \in \mathbb{R}^d$ is asymptotically stable at a deterministic rate, then φ is globally contracting at a deterministic rate.[10]

Remark 4.75. It is easy to see that if X is compact and φ is invertible, then there must exist at least one point in X that is not asymptotically stable at a deterministic rate. Consequently, one can show that if φ also has reverse-minimal dynamics, then there cannot exist a point in X that is asymptotically stable at a deterministic rate.

Example 4.76. In Example 2.34, φ is globally contracting at a deterministic rate. (For the deterministic-rate synchronisation of x and y, take $h(n) = \frac{d(x,y)}{2^n}$.) In Example 2.37 with $\alpha < 0$, φ is globally contracting at a deterministic rate. (For the deterministic-rate

[10] It is not hard to show that this statement generalises to whenever (X,d) has the Heine-Borel property (namely, that every closed bounded set is compact) and is path-connected.

synchronisation of x and y, take $h(t) = d(x,y)e^{\alpha t}$.) In Example 3.19, there do not exist distinct points $x, y \in (0,1)$ that synchronise at a deterministic rate. (This is easy to show using the fact that sufficiently close to 0, f_0 is a linear contraction towards 0 and f_1 is a linear expansion away from 0.) For the RDS generated by (3.2) in Example 3.21, if $\alpha \leq 0$ then φ is globally contracting at a deterministic rate, but if $\alpha > 0$ then there do not exist distinct points $x, y \in \mathbb{R}$ that synchronise at a deterministic rate. (See Example 6.7.) In Example 4.45, for every $x, y \in \mathbb{S}^1$, x and y synchronise at a deterministic rate; but this is of virtually no practical relevance, since the fixed point towards which all trajectories converge is not even Lyapunov stable under f.

Remark 4.77. One may be tempted to assume that having synchronisation at a deterministic rate (in the sense of Definition 4.71) is *inherently* more practically useful than having almost sure synchronisation in a model that cannot provide a strict upper bound on the time taken for the synchronisation to be observed. However, (even assuming that there are no issues concerning local stability) this is not so. Suppose, for example, that we have one system in which the distance between the trajectories of two given initial conditions is predicted to decay almost surely at least as quickly as some function h. And suppose we have a second system (with the same state space) in which the distance between the trajectories of the same initial conditions is predicted to decay, with probability greater than $1 - 2^{-100}$, at least as quickly as h. In practice, it is far more likely for *either* system to undergo some catastrophe not accounted for in the model (e.g. theft, or an earthquake) than for someone to toss 100 consecutive heads on a fair coin! (Nonetheless, it should still be said that for many systems where noise-induced synchronisation theoretically occurs, the synchronisation will take a long time to be observed, especially if the noise intensity is small; see Remark 3.1.)

Chapter 5. Synchronisation in Orientation-Preserving RDS on the Circle

Before presenting general criteria for synchronising behaviour in RDS (Chapter 6), we first look specifically at the case of orientation-preserving RDS on $\mathbb{S}^1$. Here, we have a geometrically intuitive characterisation of stable synchronisation, as well as fairly weak sufficient conditions for stable synchronisation that do not involve local stability. The content of this chapter is based on [New15c].

Let $\mathbb{S}^1$ be the unit circle, which we identify with the quotient of the additive group $(\mathbb{R},+)$ by the subgroup $\mathbb{Z}$. Let $\pi:\mathbb{R}\to\mathbb{S}^1$ denote the natural projection; a *lift* of a point $x\in\mathbb{S}^1$ is a point $x'\in\mathbb{R}$ such that $\pi(x')=x$, and a lift of a set $A\subset\mathbb{S}^1$ is a set $B\subset\mathbb{R}$ such that $\pi(B)=A$. Let l denote the (normalised) Lebesgue measure on $\mathbb{S}^1$. Define the metric d on $\mathbb{S}^1$ by

$$d(x,y) \;=\; \min\{|x'-y'| : x' \text{ is a lift of } x,\ y' \text{ is a lift of } y\}.$$

Note that under this metric, for any connected $J\subset\mathbb{S}^1$,

$$\operatorname{diam} J \;=\; \min\!\big(l(J),\tfrac{1}{2}\big).$$

The following basic fact is sufficiently clear that we do not write out a proof; nonetheless, it will be useful to state it explicitly.

Lemma 5.1. *(A) For any probability measure ρ on $\mathbb{S}^1$, the following statements are equivalent:*

- *ρ is atomless;*
- *for any sequence (J_n) of connected subsets of $\mathbb{S}^1$, if $l(J_n)\to 0$ as $n\to\infty$ then $\rho(J_n)\to 0$ as $n\to\infty$;*
- *for any sequence (J_n) of connected subsets of $\mathbb{S}^1$, if $l(J_n)\to 1$ as $n\to\infty$ then $\rho(J_n)\to 1$ as $n\to\infty$.*

(B) For any probability measure ρ on $\mathbb{S}^1$, the following statements are equivalent:

- *ρ has full support;*
- *for any sequence (J_n) of connected subsets of $\mathbb{S}^1$, if $\rho(J_n)\to 0$ as $n\to\infty$ then $l(J_n)\to 0$ as $n\to\infty$;*
- *for any sequence (J_n) of connected subsets of $\mathbb{S}^1$, if $\rho(J_n)\to 1$ as $n\to\infty$ then $l(J_n)\to 1$ as $n\to\infty$.*

Define the *anticlockwise distance function* $d_+:\mathbb{S}^1\times\mathbb{S}^1\to[0,1)$ by

$$d_+(x,y) \;=\; \min\{r\geq 0 : \pi(x'+r)=y\}$$

where x' may be any lift of x. Obviously d_+ is not symmetric, but rather satisfies the relation

$$d_+(y,x) = 1 - d_+(x,y).$$

It is clear that for all $x, y \in \mathbb{S}^1$,

$$d(x,y) = \begin{cases} d_+(x,y) & \text{if } d_+(x,y) \leq \frac{1}{2} \\ d_+(y,x) & \text{if } d_+(x,y) \geq \frac{1}{2}. \end{cases}$$

Note that d_+ is continuous on the set $\{(x,y) \in \mathbb{S}^1 \times \mathbb{S}^1 : x \neq y\}$. For any interval $I \subset \mathbb{R}$ of positive length less than 1, letting $x_1 := \pi(\inf I)$, $x_2 := \pi(\sup I)$ and $J := \pi(I)$, we have that

$$l(\varphi(t,\omega)J) = d_+(\varphi(t,\omega)x_1, \varphi(t,\omega)x_2) \tag{5.1}$$

for all t and ω.

Standing Assumption. *Throughout Chapter 5, we assume that $X = \mathbb{S}^1$, equipped with the metric d given above, and that φ is a right-continuous RDS. We also assume that $\varphi(t,\omega)$ is an orientation-preserving homeomorphism for all $t \in \mathbb{T}^+$ and $\omega \in \Omega$.*

By Lemmas 2.83 and 2.86, it follows that φ is right-continuously invertible; and if φ is continuous then φ is continuously invertible.

Recall that, as in Section 2.9, $(\bar{\varphi}^t_x)_{x \in X, t \in \mathbb{T}^+}$ denotes the family of "time-reversed" Markov transition probabilities associated to φ. As in Chapter 4, for each $\omega \in \Omega$, $\sim_\omega$ denotes the equivalence relation

$$x \sim_\omega y \iff d(\,\varphi(t,\omega)x\,,\,\varphi(t,\omega)y\,) \to 0 \text{ as } t \to \infty.$$

5.1 Stable synchronisation in terms of crack points

Definition 5.2 (c.f. [Kai93]). Given a point $r \in \mathbb{S}^1$ and a sample point $\omega \in \Omega$, we will say that r is a *crack point* of ω if the following equivalent statements hold:

- for every open $U \subset \mathbb{S}^1$ with $r \in U$, $l(\varphi(t,\omega)U) \to 1$ as $t \to \infty$;
- for every closed $G \subset \mathbb{S}^1$ with $r \notin G$, $l(\varphi(t,\omega)G) \to 0$ as $t \to \infty$;
- for every $A \subset \mathbb{S}^1$ with $r \notin \bar{A}$, $\mathrm{diam}(\varphi(t,\omega)A) \to 0$ as $t \to \infty$.

It is clear that any sample point admits at most one crack point. If a sample point ω admits a crack point, then we will say that ω is *contractive*.

Now if a sample point ω admits a crack point r, then it is clear that all points in $\mathbb{S}^1 \setminus \{r\}$ are equivalent under $\sim_\omega$. Hence, we have that either

(a) the equivalence relation $\sim_\omega$ has two equivalence classes, namely $\{r\}$ and $\mathbb{S}^1 \setminus \{r\}$; or

(b) the equivalence relation $\sim_\omega$ has one equivalence class (the whole of $\mathbb{S}^1$).

In case (a), we say that r is a *repulsive crack point* of ω.

Definition 5.3. Let $\Omega_c \subset \Omega$ be the set of contractive sample points, and let $\tilde{r}:\Omega_c \to \mathbb{S}^1$ denote the function mapping a contractive sample point ω to its crack point $\tilde{r}(\omega)$.

Lemma 5.4. *Ω_c is $\mathcal{F}_0^\infty$-measurable, and $\tilde{r}:\Omega_c \to \mathbb{S}^1$ is measurable with respect to the σ-algebra $\mathcal{F}_c$ of $\mathcal{F}_0^\infty$-measurable subsets of Ω_c. For all $t \in \mathbb{T}$, $\theta^t(\Omega_c) = \Omega_c$; and $\tilde{r}(\theta^t\omega) = \varphi(t,\omega)\tilde{r}(\omega)$ for all $\omega \in \Omega_c$ and $t \in \mathbb{T}^+$.*

Proof. Let R be a countable dense subset of $\mathbb{S}^1$. For any connected $J \subset \mathbb{S}^1$, it is clear (by considering rational times) that

$$\{\omega \in \Omega : l(\varphi(t,\omega)J) \to 0 \text{ as } t \to \infty\} \in \mathcal{F}_0^\infty. \tag{5.2}$$

So then, in order to show that $\Omega_c \in \mathcal{F}_0^\infty$, it suffices to prove the following statement: a sample point $\omega \in \Omega$ is contractive if and only if for every $n \in \mathbb{N}$ there is a connected open set $U_n \subset \mathbb{S}^1$ with endpoints in R such that $1-\frac{1}{n} < l(U_n) < 1$ and $l(\varphi(t,\omega)U_n) \to 0$ as $t \to \infty$. Now the "only if" direction is obvious. For the "if" direction: suppose that for every $n \in \mathbb{N}$ there exists a connected open set $U_n \subset \mathbb{S}^1$ with endpoints in R such that $1-\frac{1}{n} < l(U_n) < 1$ and $l(\varphi(t,\omega)U_n) \to 0$ as $t \to \infty$; and let $U := \bigcup_{n=1}^\infty U_n$. Since U_n is connected for all n and $l(U_n) \to 1$ as $n \to \infty$, we clearly have that either $U = \mathbb{S}^1$ or $\mathbb{S}^1 \setminus \{U\}$ is a singleton. Now suppose, for a contradiction, that $U = \mathbb{S}^1$. Then, since $\mathbb{S}^1$ is compact, there is a finite subset $\{n_1,\ldots,n_k\}$ of $\mathbb{N}$ such that $\mathbb{S}^1 = \bigcup_{i=1}^k U_{n_i}$; but since $l(\varphi(t,\omega)U_{n_i}) \to 0$ as $t \to \infty$ for each i, we then have that $l(\varphi(t,\omega)\mathbb{S}^1) \to 0$ as $t \to \infty$, which is absurd. So then, we must have that $\mathbb{S}^1 \setminus U$ is equal to a singleton $\{r\}$. We now show that r is a crack point. Fix any closed $G \subset \mathbb{S}^1$ with $r \notin G$. Take n such that $l(U_n) > 1 - d(r,G)$; then $G \subset U_n$ and so $l(\varphi(t,\omega)G) \to 0$ as $t \to \infty$. Hence r is a crack point of ω.

Thus we have shown that Ω_c is $\mathcal{F}_0^\infty$-measurable. Now for any non-empty closed connected $K \subset \mathbb{S}^1$, it is clear that a sample point $\omega \in \Omega_c$ belongs to $\tilde{r}^{-1}(K)$ if and only if for every closed connected $G \subset \mathbb{S}^1 \setminus K$ with boundary in R, $l(\varphi(t,\omega)G) \to 0$ as $t \to \infty$. So by (5.2) and the countability of R, $\tilde{r}^{-1}(K) \in \mathcal{F}_c$ for every closed connected $K \subset \mathbb{S}^1$. Hence $\tilde{r}$ is $\mathcal{F}_c$-measurable.

Now for any $\omega \in \Omega$, $r \in \mathbb{S}^1$ and $t \in \mathbb{T}^+$, we obviously have that if $U \subset \mathbb{S}^1$ is a neighbourhood of r then $\varphi(t,\omega)U$ is a neighbourhood of $\varphi(t,\omega)r$, and that if $V \subset \mathbb{S}^1$ is a neighbourhood of $\varphi(t,\omega)r$ then $\varphi(t,\omega)^{-1}(V)$ is a neighbourhood of r; so then, it is easy to see that

$$r \text{ is a crack point of } \omega \iff \varphi(t,\omega)r \text{ is a crack point of } \theta^t\omega.$$

So then, for any $\omega \in \Omega$ and $t \in \mathbb{T}^+$, we have that $\omega \in \Omega_c \Leftrightarrow \theta^t\omega \in \Omega_c$ (so $\theta^{-t}(\Omega_c) = \Omega_c$); and obviously $\tilde{r}(\theta^t\omega) = \varphi(t,\omega)\tilde{r}(\omega)$ for all $\omega \in \Omega_c$. □

Corollary 5.5. *$\mathbb{P}(\Omega_c)$ is equal to either 0 or 1. In the case that $\mathbb{P}(\Omega_c) = 1$, either:*

(a) for every $x \in \mathbb{S}^1$, $\mathbb{P}(\omega \in \Omega_c : \tilde{r}(\omega) = x) = 0$; or

(b) there exists a deterministic fixed point $p \in \mathbb{S}^1$ such that $\mathbb{P}(\omega \in \Omega_c : \tilde{r}(\omega) = p) = 1$.

Proof. The fact that $\mathbb{P}(\Omega_c) \in \{0,1\}$ follows immediately from Lemma 5.4 and the ergodicity of $\mathbb{P}$ under (θ^t) (Lemma 2.8). In the case that $\mathbb{P}(\Omega_c) = 1$, define the function $r{:}\,\Omega \to \mathbb{S}^1$ by

$$r(\omega) = \begin{cases} \tilde{r}(\omega) & \omega \in \Omega_c \\ k & \omega \in \Omega \setminus \Omega_c \end{cases}$$

where k is an arbitrary constant. By Lemma 5.4, r is an $\mathcal{F}_0^\infty$-measurable random fixed point, and therefore Lemma 2.61 gives that either case (a) or case (b) holds. □

We now characterise stable synchronisation in terms of crack points.

Theorem 5.6. *φ is stably synchronising if and only if $\mathbb{P}(\Omega_c) = 1$ and case (a) of Corollary 5.5 holds. In this case, for $\mathbb{P}$-almost every $\omega \in \Omega_c$, $\tilde{r}(\omega)$ is a repulsive crack point of ω.*

Remark 5.7. If $\mathbb{P}(\Omega_c) = 1$ and case (b) of Corollary 5.5 holds, then φ is synchronising if and only if for $\mathbb{P}$-almost all $\omega \in \Omega_c$, p is a non-repulsive crack point of ω.

Most of the rest of Section 5.1 is devoted to proving Theorem 5.6.

The following lemma is similar in principle to [LeJ87, Lemme 1(a)].

Lemma 5.8. *Let ρ be an atomless[1] probability measure that is stationary under the Markov transition probabilities $(\bar{\varphi}_x^t)_{x\in X,\,t\in\mathbb{T}^+}$. For any connected $J \subset \mathbb{S}^1$, for $\mathbb{P}$-almost all $\omega \in \Omega$, $\rho(\varphi(t,\omega)J)$ is convergent as $t \to \infty$.*

Proof. Fix a connected $J \subset \mathbb{S}^1$, and for each t and ω let $h_t(\omega) = \rho(\varphi(t,\omega)J)$. Note that for each boundary point x of J, the map $t \mapsto \varphi(t,\omega)x$ is right-continuous for all ω. Hence, since ρ is atomless, the map $t \mapsto h_t(\omega)$ is right-continuous for all ω. So if we can show that $(h_t)_{t\in\mathbb{T}^+}$ is a martingale with respect to the filtration $(\mathcal{F}_0^t)_{t\in\mathbb{T}^+}$, then the martingale convergence theorem will give the desired result. Fix any $s,t \in \mathbb{T}^+$. We have that

$$\begin{aligned}
\mathbb{E}[h_{s+t}|\mathcal{F}_0^s](\omega) &= \mathbb{E}[\,\tilde{\omega} \mapsto \rho(\varphi(s+t,\tilde{\omega})J)\,|\,\mathcal{F}_0^s\,](\omega) \\
&= \mathbb{E}[\,\tilde{\omega} \mapsto \rho(\varphi(t,\theta^s\tilde{\omega})(\varphi(s,\tilde{\omega})J))\,|\,\mathcal{F}_0^s\,](\omega) \\
&= \mathbb{E}[\,\tilde{\omega} \mapsto \rho(\varphi(t,\theta^s\tilde{\omega})(\varphi(s,\omega)J))\,] \\
&\qquad \text{(by Lemma A.11, since } \mathcal{F}_0^s \text{ and } \mathcal{F}_s^{s+t} \text{ are independent)} \\
&= \rho(\varphi(s,\omega)J) \\
&\qquad \text{(by equation (2.10) with } \theta^s\omega \text{ in place of } \omega\text{)} \\
&= h_s(\omega).
\end{aligned}$$

So we are done. □

Lemma 5.9. *Suppose that φ is synchronising, and that there exists an atomless probability measure that is stationary under the Markov transition probabilities $(\bar{\varphi}_x^t)_{x\in X,\,t\in\mathbb{T}^+}$. Then $\mathbb{P}(\Omega_c) = 1$.*

[1] The condition that ρ is atomless can in fact be dropped, although the proof then becomes significantly longer, as it is harder to justify that the martingale $(h_t)_{t\in\mathbb{T}^+}$ almost surely has right-continuous sample paths. In any case, we will not need this for our purposes.

Proof. Let ρ be an atomless stationary probability measure of the Markov transition probabilities $(\bar{\varphi}_x^t)$. First, let us fix any connected set $J \subset \mathbb{S}^1$ with $0 < l(J) < 1$, and write $\partial J =: \{x, y\}$. For $\mathbb{P}$-almost every $\omega \in \Omega$, we have that $d(\varphi(t,\omega)x, \varphi(t,\omega)y) \to 0$ as $t \to \infty$, and so for any unbounded increasing sequence (t_n) in $\mathbb{T}^+$ with $l(\varphi(t_n,\omega)J)$ convergent, the limit of $l(\varphi(t_n,\omega)J)$ is equal to either 0 or 1. But also, we know by Lemma 5.8 that for $\mathbb{P}$-almost every $\omega \in \Omega$, $\rho(\varphi(t,\omega)J)$ is convergent as $t \to \infty$. Combining these facts, we have (using Lemma 5.1(A)) that for $\mathbb{P}$-almost every $\omega \in \Omega$, $l(\varphi(t,\omega)J)$ converges to either 0 or 1 as $t \to \infty$.

Now then, fix an arbitrary $k \in \mathbb{R}$, and for each $v \in [0,1]$, let $J_v := \pi([k, k+v])$. Let $\Omega' \subset \Omega$ be a $\mathbb{P}$-full set such that for each $\omega \in \Omega'$ and $v \in [0,1] \cap \mathbb{Q}$, $l(\varphi(t,\omega)J_v)$ converges to either 0 or 1 as $t \to \infty$. For each $\omega \in \Omega'$, let

$$\begin{aligned} c(\omega) &:= \sup\{v \in [0,1] : l(\varphi(t,\omega)J_v) \to 0 \text{ as } t \to \infty\} \\ &= \inf\{v \in [0,1] : l(\varphi(t,\omega)J_v) \to 1 \text{ as } t \to \infty\}. \end{aligned}$$

It is easy to see that for each $\omega \in \Omega'$, $\pi(k + c(\omega))$ is a crack point of ω. So we are done. □

Lemma 5.10. *Suppose φ is stably synchronising. Then the Markov transition probabilities then the Markov transition probabilities $(\bar{\varphi}_x^t)_{x \in X,\, t \in \mathbb{T}^+}$ admit at least one stationary probability measure that is atomless.*

Proof. First suppose that φ does *not* have a deterministic fixed point. Since $\mathbb{S}^1$ is compact, there exists at least one probability measure ρ that is ergodic with respect to $(\bar{\varphi}_x^t)$; and by Lemma 4.14, such a probability measure must be atomless.

Now suppose that φ *does* have a deterministic fixed point p. Let $p' \in \mathbb{R}$ be a lift of p, and for each $v \in [0,1]$, let $J_v := \pi([p', p'+v])$. Define the function $h : \Omega \to [0,1]$ by

$$h(\omega) = \sup\{v \in [0,1) : l(\varphi(t,\omega)J_v) \to 0 \text{ as } t \to \infty\}.$$

For any $c \in [0,1)$ and $\omega \in \Omega$, $h(\omega) > c$ if and only if there exists $v \in (c,1) \cap \mathbb{Q}$ such that $l(\varphi(t,\omega)J_v) \to 0$ as $t \to \infty$. Hence h is $\mathcal{F}_0^\infty$-measurable. Now since φ is everywhere stable in $\mathbb{S}^1$, we know that for $\mathbb{P}$-almost every $\omega \in \Omega$ there exists a neighbourhood U of p such that $l(\varphi(t,\omega)U) \to 0$ as $t \to \infty$. Hence $h(\omega) \in (0,1)$ for $\mathbb{P}$-almost all $\omega \in \Omega$.

Now define the function $q : \Omega \to \mathbb{S}^1$ by

$$q(\omega) = \pi(p' + h(\omega)).$$

Since h is $\mathcal{F}_0^\infty$-measurable, q is $\mathcal{F}_0^\infty$-measurable. Given any $t \in \mathbb{T}^+$ and $\omega \in \Omega$, we have that for all $v \in [0,1)$,

$$l(\varphi(s,\omega)J_v) \to 0 \text{ as } s \to \infty \quad \Longleftrightarrow \quad l(\varphi(s,\theta^t\omega)\,(\varphi(t,\omega)J_v)\,) \to 0 \text{ as } s \to \infty$$

and therefore $q(\theta^t\omega) = \varphi(t,\omega)q(\omega)$. Hence, by Corollary 2.60, $q_*\mathbb{P}$ is ergodic with respect to $(\bar{\varphi}_x^t)$. Moreover, since $h(\omega) \in (0,1)$ for $\mathbb{P}$-almost all $\omega \in \Omega$, $q_*\mathbb{P}$ is not equal to δ_p. Since φ is synchronising, φ cannot have more than one deterministic fixed point, and so $q_*\mathbb{P}$ is not a Dirac mass at a deterministic fixed point. Therefore (by either Lemma 4.14 or Lemma 2.61), $q_*\mathbb{P}$ is atomless. □

Combining Lemmas 5.9 and 5.10 gives that if φ is stably synchronising then $\mathbb{P}(\Omega_c) = 1$.

Lemma 5.11. *Suppose $\mathbb{P}(\Omega_c) = 1$. Then φ is stably synchronising if and only if case (a) in the statement of Corollary 5.5 holds.*

Proof. For any $x, y \in \mathbb{S}^1$ and $\omega \in \Omega_c$, if $\tilde{r}(\omega) \neq x$ and $\tilde{r}(\omega) \neq y$ then $x \sim_\omega y$. Hence it is clear that in case (a) in the statement of Corollary 5.5, φ is synchronising. For any $x \in \mathbb{S}^1$ and $\omega \in \Omega_c$, if $\tilde{r}(\omega) \neq x$ then there obviously exists a neighbourhood U of x such that $\operatorname{diam}(\varphi(t,\omega)U) \to 0$ as $t \to \infty$. Hence, in case (a) in the statement of Corollary 5.5, φ is everywhere stable (and therefore uniformly stable, since $\mathbb{S}^1$ is compact). Thus we have seen that in case (a) in the statement of Corollary 5.5, φ is stably synchronising.

Now if there exists $p \in \mathbb{S}^1$ such that $\mathbb{P}(\omega \in \Omega_c : \tilde{r}(\omega) = p) > 0$, then p is not almost surely stable, and so φ is not stably synchronising. □

Combining Lemma 5.11 with the fact that if φ is stably synchronising then $\mathbb{P}(\Omega_c) = 1$ yields all of Theorem 5.6, except the final assertion that if φ is stably synchronising then $\tilde{r}(\omega)$ is almost surely repulsive.

The following statement is not specific to orientation-preserving RDS on the circle, but generalises to any right-continuous RDS on a metric space (X, d) with $\mathcal{B}(X)$ standard.

Lemma 5.12. *Let $q : \Omega \to \mathbb{S}^1$ be a $\mathcal{F}_0^\infty$-measurable random fixed point, and suppose that $q_*\mathbb{P}$ is atomless. Let ρ be any stationary probability measure of the Markov transition probabilities $(\varphi_x^t)_{x \in X, t \in \mathbb{T}^+}$. For $\mathbb{P}$-almost every $\omega \in \Omega$, $\rho(x \in \mathbb{S}^1 : x \sim_\omega q(\omega)) = 0$.*

Proof. Define the function $\Theta_{[2]} : \Omega \times \mathbb{S}^1 \times \mathbb{S}^1 \to \Omega \times \mathbb{S}^1 \times \mathbb{S}^1$ by

$$\Theta_{[2]}(\omega, x, y) \;=\; (\theta^1\omega, \varphi(1,\omega)x, \varphi(1,\omega)y).$$

Define the probability measure $\mathfrak{p}$ on the measurable space $(\Omega \times \mathbb{S}^1 \times \mathbb{S}^1\,,\, \mathcal{F}_0^\infty \otimes \mathcal{B}(\mathbb{S}^1 \times \mathbb{S}^1))$ by

$$\mathfrak{p}(A) \;:=\; \mathbb{P} \otimes \rho(\,(\omega, x) \in \Omega \times \mathbb{S}^1 \,:\, (\omega, x, q(\omega)) \in A\,).$$

For any $A \in \mathcal{F}_0^\infty \otimes \mathcal{B}(\mathbb{S}^1 \times \mathbb{S}^1)$, since q is $\mathcal{F}_0^\infty$-measurable, the set $\{(\omega, x) : (\omega, x, q(\omega)) \in A\}$ is $(\mathcal{F}_0^\infty \otimes \mathcal{B}(\mathbb{S}^1))$-measurable. With this, we have

$$\begin{aligned}
\mathfrak{p}(\Theta_{[2]}^{-1}(A)) &= \mathbb{P} \otimes \rho(\,(\omega, x) \in \Omega \times \mathbb{S}^1 \,:\, (\theta^1\omega, \varphi(1,\omega)x, \varphi(1,\omega)q(\omega)) \in A\,) \\
&= \mathbb{P} \otimes \rho(\,(\omega, x) \in \Omega \times \mathbb{S}^1 \,:\, (\theta^1\omega, \varphi(1,\omega)x, q(\theta^1\omega)) \in A\,) \\
&= \mathbb{P} \otimes \rho(\,\Theta^{-1}\{(\omega, x) \in \Omega \times \mathbb{S}^1 \,:\, (\omega, x, q(\omega)) \in A\}\,) \\
&= \mathbb{P} \otimes \rho(\,(\omega, x) \in \Omega \times \mathbb{S}^1 \,:\, (\omega, x, q(\omega)) \in A\,) \\
&\qquad \text{(since } \mathbb{P}|_{\mathcal{F}_0^\infty} \otimes \rho \text{ is } \Theta^1\text{-invariant)} \\
&= \mathfrak{p}(A).
\end{aligned}$$

So $\mathfrak{p}$ is $\Theta_{[2]}$-invariant. Now since $q_*\mathbb{P}$ is atomless, we have that

$$\begin{aligned}
\mathfrak{p}(\Omega \times \Delta_X) &= \mathbb{P} \otimes \rho(\,(\omega, x) \in \Omega \times \mathbb{S}^1 \,:\, q(\omega) = x\,) \\
&= \int_{\mathbb{S}^1} \mathbb{P}(\omega \in \Omega : q(\omega) = x)\, \rho(dx) \\
&= 0.
\end{aligned}$$

So letting $U_\varepsilon := \{(x,y) \in \mathbb{S}^1 \times \mathbb{S}^1 : d(x,y) < \varepsilon\}$ for each $\varepsilon > 0$, we have that $\mathfrak{p}(\Omega \times U_\varepsilon) \to 0$ as $\varepsilon \to 0$. Hence the set

$$\begin{aligned} K &:= \{ (\omega, x, y) \in \Omega \times \mathbb{S}^1 \times \mathbb{S}^1 \,:\, d(\varphi(n,\omega)x, \varphi(n,\omega)y) \to 0 \text{ as } n \to \infty \} \\ &= \bigcap_{n=1}^{\infty} \bigcup_{i=1}^{\infty} \bigcap_{j=i}^{\infty} \Theta_{[2]}^{-j}(\Omega \times U_{\frac{1}{n}}) \end{aligned}$$

is a $\mathfrak{p}$-null set. Therefore (by definition of $\mathfrak{p}$), the set

$$L \;:=\; \{ (\omega, x) \in \Omega \times \mathbb{S}^1 \,:\, d(\varphi(n,\omega)x, \varphi(n,\omega)q(\omega)) \to 0 \text{ as } n \to \infty \}$$

is a $(\mathbb{P} \otimes \rho)$-null set. So (with Fubini's theorem) we are done. □

Hence we can complete the proof of Theorem 5.6: since $\mathbb{S}^1$ is compact, there must exist a stationary probability measure ρ of the Markov transition probabilities (φ_x^t); so applying Lemma 5.12 with q being the function r in the proof of Corollary 5.5, we have that for $\mathbb{P}$-almost every $\omega \in \Omega_c$ the equivalence relation $\sim_\omega$ has more than one equivalence class, and so $\tilde{r}(\omega)$ is repulsive.

We mention a further relevant fact, which we will not prove here:

Proposition 5.13. *Suppose φ is stably synchronising, and let $r\colon\Omega \to \mathbb{S}^1$ be a measurable function agreeing with $\tilde{r}$ $\mathbb{P}$-almost everywhere in Ω_c. There exists an $\mathcal{F}_{-\infty}^0$-measurable random fixed point $a\colon\Omega \to \mathbb{S}^1$ such that every φ-invariant probability measure on $(\Omega \times \mathbb{S}^1, \mathcal{F} \otimes \mathcal{B}(\mathbb{S}^1))$ has a disintegration (μ_ω) taking the form*

$$\mu_\omega \;=\; \lambda\delta_{a(\omega)} + (1-\lambda)\delta_{r(\omega)}$$

for some $\lambda \in [0,1]$.

We may regard the pair of random fixed points (a, r) as *the attractor-repeller pair* of φ. For a proof of Proposition 5.13, see [New15c, Theorems 5.10, 5.13].

For an important example of a stably synchronising RDS on $\mathbb{S}^1$, see Section 5.3.

Finally, we introduce briefly the notion of a "crack set":

Definition 5.14. Fix a sample point $\omega \in \Omega$ and a non-empty finite set $R \subset \mathbb{S}^1$. We say that R is a *crack set of* ω if the following statements hold:

(i) each connected component of $\mathbb{S}^1 \setminus R$ is σ-contracting under ω;

(ii) any two distinct connected components of $\mathbb{S}^1 \setminus R$ are contained in distinct equivalence classes of $\sim_\omega$.

And we say that a crack set R is *repulsive* if for any $x \in \mathbb{S}^1 \setminus R$, $d(\varphi(t,\omega)x, \varphi(t,\omega)R) \not\to 0$ as $t \to \infty$.

Observe that a point $r \in \mathbb{S}^1$ is a crack point of ω if and only if the singleton $\{r\}$ is a crack set of ω, and that r is a repulsive crack point if and only if $\{r\}$ is a repulsive crack set. Also note that if R is a crack set of ω with at least two elements, then R is precisely the set of all boundary points of equivalence classes of $\sim_\omega$. Hence, any given sample point possesses at most one crack set.

Proposition 5.15. *Let φ' be a right-continuous RDS on $\mathbb{S}^1$ over $(\Omega, \mathcal{F}, (\mathcal{F}_s^{s+t}), (\theta^t), \mathbb{P})$ such that $\varphi'(t,\omega)$ is an orientation-preserving homeomorphism for all t and ω; and suppose we have a deterministic semiconjugacy $h: \mathbb{S}^1 \to \mathbb{S}^1$ from φ to φ', with a strictly increasing lift $H: \mathbb{R} \to \mathbb{R}$. For any $\omega \in \Omega$ and $r \in \mathbb{S}^1$, if r is a crack point of ω under φ' then $h^{-1}(\{r\})$ is a crack set of ω under φ. If, in addition, r is repulsive (under φ') then $h^{-1}(\{r\})$ is repulsive (under φ).*

Proof. For any $x \in \mathbb{S}^1$, the number of elements of $h^{-1}(\{x\})$ is precisely the degree k of h,[2] and each connected component of $\mathbb{S}^1 \setminus h^{-1}(\{x\})$ is mapped homeomorphically into $\mathbb{S}^1 \setminus \{x\}$ by h. It is not hard to show that

$$\inf\{d(x,y) : x, y \in \mathbb{S}^1,\ h(x) = h(y)\} \ > \ 0. \tag{5.3}$$

It is also not hard to show that for all $\varepsilon > 0$ there exists $\delta(\varepsilon) \in (0, \frac{1}{2})$ such that for any connected $J \subset \mathbb{S}^1$ of length less than δ, the length of every connected component of $h^{-1}(J)$ is less than ε.

Now suppose r is a crack point of ω under φ'. Let I be any connected component of $\mathbb{S}^1 \setminus h^{-1}(\{r\})$, and let G be a compact subset of I. Fix any $\varepsilon > 0$. Obviously $h(G)$ is a compact subset of $\mathbb{S}^1 \setminus \{r\}$, so let J be a compact connected subset of $\mathbb{S}^1 \setminus \{r\}$ containing $h(G)$. Note that G is contained in a connected component K of $h^{-1}(J)$. For all $t \in \mathbb{T}^+$, we have that $\varphi(t,\omega)K \subset h^{-1}(\varphi'(t,\omega)J)$. Now let $T \in \mathbb{T}^+$ be such that for all $t \geq T$, $l(\varphi'(t,\omega)J) < \delta(\varepsilon)$. Then, since $\varphi(t,\omega)K$ is connected (for any t), we have that for all $t \geq T$, $l(\varphi(t,\omega)K) < \varepsilon$ and therefore $\operatorname{diam}(\varphi(t,\omega)G) < \varepsilon$. Since ε was arbitrary, we have that G contracts under ω. Hence I is σ-contracting under ω. Now given distinct connected components I_1 and I_2 of $\mathbb{S}^1 \setminus h^{-1}(\{r\})$, if we take $x \in I_1$ and $y \in I_2$ such that $h(x) = h(y)$, then $h(\varphi(t,\omega)x) = h(\varphi(t,\omega)y)$ for all t, and therefore by (5.3), $x \nsim_\omega y$. Hence I_1 and I_2 belong to distinct equivalence classes of $\sim_\omega$. Thus we have shown that $h^{-1}(\{r\})$ is a crack set of ω under φ.

Suppose that the crack set $h^{-1}(\{r\})$ is not repulsive, and let $x \in \mathbb{S}^1 \setminus h^{-1}(\{r\})$ be such that $d(\varphi(t,\omega)x, \varphi(t,\omega)(h^{-1}(\{r\}))) \to 0$ as $t \to \infty$. Since h is uniformly continuous, it follows that $d(\varphi'(t,\omega)h(x), \varphi'(t,\omega)r) \to 0$ as $t \to \infty$. Since $h(x) \neq r$, it follows that r is not repulsive. □

5.2 A test for stable synchronisation

The aim of this section is to present weak and easily verifiable sufficient conditions for stable synchronisation. Heuristically, the conditions that we shall give demonstrate that "sufficient flexibility" in how the noise can effect the system is guaranteed to lead to stable synchronisation. An application shall be presented in detail in Section 5.3.

Some additional results which will not be presented here are included in Section 2 of [New15c].

When we wish to say that φ is contractible on $\mathbb{S}^1$, we will just say that "φ is contractible".

[2] that is, the unique integer k for which the map $y \mapsto H(y) - ky$ is periodic.

Definition 5.16. We say that φ is *two-way contractible*[3] if for any distinct $x, y \in \mathbb{S}^1$,

$$\mathbb{P}(\,\omega : \exists\, t \in \mathbb{T}^+ \text{ s.t. } d_+(\varphi(t,\omega)x, \varphi(t,\omega)y) < d_+(x,y)\,) \;>\; 0.$$

By reversing the order of inputs, this is equivalent to saying that for any distinct $x, y \in \mathbb{S}^1$,

$$\mathbb{P}(\,\omega : \exists\, t \in \mathbb{T}^+ \text{ s.t. } d_+(\varphi(t,\omega)x, \varphi(t,\omega)y) > d_+(x,y)\,) \;>\; 0.$$

We can also define two-way contractibility in terms of connected subsets of $\mathbb{S}^1$: φ is two-way contractible if and only if for every connected set $J \subset \mathbb{S}^1$ with $0 < l(J) < 1$,

$$\mathbb{P}(\,\omega : \exists\, t \in \mathbb{T}^+ \text{ s.t. } l(\varphi(t,\omega)J) < l(J)\,) \;>\; 0.$$

Again, this is equivalent to saying that for every connected set $J \subset \mathbb{S}^1$ with $0 < l(J) < 1$,

$$\mathbb{P}(\,\omega : \exists\, t \in \mathbb{T}^+ \text{ s.t. } l(\varphi(t,\omega)J) > l(J)\,) \;>\; 0.$$

Obviously, if φ is two-way contractible then φ is contractible.

Remark 5.17. Suppose there exists a separable metrisable topology on Ω whose Borel σ-algebra coincides with $\mathcal{F}$, such that $\mathbb{P}$ has full support and for all $t \in \mathbb{T}^+$ and $x \in \mathbb{S}^1$, the map $\omega \mapsto \varphi(t,\omega)x$ is continuous. Then, as in Remarks 2.73 and 4.67, in order to show that φ is two-way contractible, it is sufficient to show that for each pair of distinct points $x, y \in \mathbb{S}^1$ there exists a sample point $\omega \in \Omega$ and a time $t \in \mathbb{T}^+$ such that $d_+(\varphi(t,\omega)x, \varphi(t,\omega)y) < d_+(x,y)$.

We will not need the following proposition elsewhere, but it is worth stating nonetheless:

Proposition 5.18. *If φ is two-way contractible then for any $x, y \in \mathbb{S}^1$ and $\varepsilon > 0$ there exists $t \in \mathbb{T}^+$ such that*

$$\mathbb{P}(\,\omega : d_+(\varphi(t,\omega)x, \varphi(t,\omega)y) < \varepsilon\,) \;>\; 0.$$

For the proof, see [New15c, Proposition 2.5].

The following theorem (the main result of this section) generalises results in [DKN07, Section 5.1].

Theorem 5.19. *The following statements are equivalent:*

(i) φ is two-way contractible and has no deterministic fixed points;

(ii) φ is contractible and has reverse-minimal dynamics;

and when these hold, φ is stably synchronising.

[3] In [New15c], the term "compressible" is used; however, since we already use the term "incompressible" to describe probability measures on the phase space X, we use the term "two-way contractible" here in order to avoid confusion. (The term "two-point contractible" also reflects the meaning more clearly.)

Observe in particular that if φ has reverse-minimal dynamics then contractibility, two-way contractibility, synchronisation and stable synchronisation are all equivalent.

Before proving Theorem 5.19, it is worth mentioning that in continuous time, if φ is continuous then reverse-minimal dynamics and minimal dynamics are the same:

Proposition 5.20. *If $\mathbb{T} = \mathbb{R}$ and φ is continuously invertible, then the following are equivalent:*

(i) φ has reverse-minimal dynamics on $\mathbb{S}^1$;

(ii) φ has minimal dynamics on $\mathbb{S}^1$.

Proof. We first show that (i)⇒(ii). Suppose we have a closed invariant non-empty proper subset G of $\mathbb{S}^1$; we need to show that there exists an open invariant non-empty proper subset U of $\mathbb{S}^1$. Firstly, if G is a singleton $\{p\}$ then $U := \mathbb{S}^1 \setminus \{p\}$ is clearly invariant. Now consider the case that G is not a singleton, and let V be a connected component of $\mathbb{S}^1 \setminus G$; we will show that $U := \mathbb{S}^1 \setminus \bar{V}$ is invariant. (Note that U is non-empty, i.e. $\bar{V} \neq \mathbb{S}^1$, since G is not a singleton.) Fix any ω with the property that $\varphi(t,\omega)G \subset G$ for all $t \in \mathbb{T}^+$. Since $\partial V \subset G$, we have that for all t, $\varphi(t,\omega)\partial V \subset G$ and therefore in particular $\varphi(t,\omega)\partial V \cap V = \varnothing$. Now since φ is a continuous RDS, we can define continuous functions $a, b\colon [0,\infty) \to \mathbb{R}$ with $a < b$ such that $[a(t), b(t)]$ is a lift of $\varphi(t,\omega)\bar{V}$ for all t. (So $\{a(t), b(t)\}$ projects onto $\varphi(t,\omega)\partial V$ for all t.) For all t, since $\varphi(t,\omega)\partial V \cap V = \varnothing$, we have that $a(t), b(t) \notin (a(0), b(0))$. Therefore (due to the intermediate value theorem), $a(t) \leq a(0)$ for all t and $b(t) \geq b(0)$ for all t. Hence $\bar{V} \subset \varphi(t,\omega)\bar{V}$ for all t. Since $\varphi(t,\omega)$ is bijective for all t, it follows that $\varphi(t,\omega)U \subset U$ for all t. So U is invariant.

Now, in order to show that (ii)⇒(i), first observe that a set $A \subset \mathbb{S}^1$ is invariant if and only if $\mathbb{P}$-almost every $\omega \in \Omega$ has the property that for all $t \in \mathbb{T}^+$,

$$\varphi(t,\omega)^{-1}(X \setminus A) \;\subset\; X \setminus A.$$

Hence the fact that (ii)⇒(i) follows from the fact that (i)⇒(ii), except with the family of functions $(\varphi(t,\omega))_{t\in\mathbb{T}^+,\omega\in\Omega}$ replaced by the family of functions $(\varphi(t,\omega)^{-1})_{t\in\mathbb{T}^+,\omega\in\Omega}$. □

Proof of Theorem 5.19

To prove Theorem 5.19, we will first prove that (i)⇒(ii)⇒stable synchronisation, and then, using material developed along the way, we will prove that (ii)⇒(i).

Proof that (i)⇒(ii). Suppose φ is two-way contractible and has no deterministic fixed points. Suppose for a contradiction that φ does not have reverse-minimal dynamics, and let U be an open invariant non-empty proper subset of $\mathbb{S}^1$. Let V be a maximal-length connected component of U. Since there are no deterministic fixed points, $\mathbb{S}^1 \setminus U$ is not a singleton and so $l(V) < 1$. Hence, since φ is two-way contractible, there is a positive-measure set of sample points $\omega \in \Omega$ for each of which, for some $t_\omega \in \mathbb{T}^+$, $l(\varphi(t_\omega,\omega)V) > l(V)$. However, $\varphi(t,\omega)V$ is connected for all t and ω, and so if $l(\varphi(t_\omega,\omega)V) > l(V)$ then $\varphi(t_\omega,\omega)V$ cannot be a subset of U. This contradicts the fact that U is invariant. □

We now start working towards the proof that (ii)⇒stable synchronisation.

Lemma 5.21. *Suppose that φ is contractible, and that there exists a stationary probability measure ρ of the Markov transition probabilities $(\bar{\varphi}_x^t)_{x\in X,\,t\in\mathbb{T}^+}$ that is atomless and has full support. Then φ is synchronising.*

(We will soon prove that under these same conditions, φ is *stably* synchronising.)

Proof. Fix any distinct $x,y \in \mathbb{S}^1$. Let $J \subset \mathbb{S}^1$ be a connected set with $\partial J = \{x,y\}$. By Proposition 4.68 and Lemma 5.8, there is a $\mathbb{P}$-full set of sample points ω with the properties that

(a) there exists an unbounded increasing sequence (t_n) in $\mathbb{T}^+$ such that

$$d(\varphi(t_n,\omega)x, \varphi(t_n,\omega)y) \to 0 \quad \text{as } n \to \infty\,;$$

(b) $\rho(\varphi(t,\omega)J)$ is convergent as $t \to \infty$.

Fix any ω with both these properties, and let (t_n) be as in (a). For any n, $d(\varphi(t_n,\omega)x, \varphi(t_n,\omega)y)$ is precisely the smaller of $l(\varphi(t_n,\omega)J)$ and $1-l(\varphi(t_n,\omega)J)$. Hence there must exist a subsequence (t_{m_n}) of (t_n) such that either $l(\varphi(t_{m_n},\omega)J) \to 0$ as $n \to \infty$ or $l(\varphi(t_{m_n},\omega)J) \to 1$ as $n \to \infty$. Since ρ is atomless, Lemma 5.1(A) then gives that either $\rho(\varphi(t_{m_n},\omega)J) \to 0$ as $n \to \infty$ or $\rho(\varphi(t_{m_n},\omega)J) \to 1$ as $n \to \infty$. Since $\rho(\varphi(t,\omega)J)$ is convergent as $t \to \infty$, it follows that either $\rho(\varphi(t,\omega)J) \to 0$ as $t \to \infty$ or $\rho(\varphi(t,\omega)J) \to 1$ as $t \to \infty$. Since ρ has full support, Lemma 5.1(B) then gives that either $l(\varphi(t,\omega)J) \to 0$ as $t \to \infty$ or $l(\varphi(t,\omega)J) \to 1$ as $t \to \infty$. Hence $d(\varphi(t,\omega)x, \varphi(t,\omega)y) \to 0$ as $t \to \infty$. □

Lemma 5.22. *Under the hypotheses of Lemma 5.21, for any connected $J \subset \mathbb{S}^1$,*

$$\mathbb{P}(\,\omega : l(\varphi(t,\omega)J) \to 0 \text{ as } t \to \infty\,) \;=\; 1 - \rho(J).$$

Proof. Fix any connected $J \subset \mathbb{S}^1$. As in the proof of Lemma 5.21, we have that for $\mathbb{P}$-almost every $\omega \in \Omega$, either

$$\rho(\varphi(t,\omega)J) \to 0 \quad \text{and} \quad l(\varphi(t,\omega)J) \to 0 \quad \text{as } t \to \infty.$$

or

$$\rho(\varphi(t,\omega)J) \to 1 \quad \text{and} \quad l(\varphi(t,\omega)J) \to 1 \quad \text{as } t \to \infty.$$

So then, letting E denote the set of sample points ω for which the latter scenario holds, the dominated convergence theorem gives that as $t \to \infty$,

$$\int_\Omega \rho(\varphi(t,\omega)J)\,\mathbb{P}(d\omega) \;\to\; \int_\Omega \mathbb{1}_E(\omega)\,\mathbb{P}(d\omega) \;=\; \mathbb{P}(E).$$

But we also know that for any t,

$$\int_\Omega \rho(\varphi(t,\omega)J)\,\mathbb{P}(d\omega) \;=\; \rho(J).$$

Hence $\mathbb{P}(E) = \rho(J)$, i.e. the probability of the latter scenario is $\rho(J)$ and the probability of the former scenario is $1 - \rho(J)$, as required. □

Combining Lemmas 5.21 and 5.22, we have:

Corollary 5.23. *Under the hypotheses of Lemma 5.21, φ is stably synchronising.*

Proof. We already know (from Lemma 5.21) that φ is synchronising. Now fix any $x \in X$. Let $(U_n)_{n\in\mathbb{N}}$ be a nested sequence of connected neighbourhoods of x such that $\bigcap_n U_n = \{x\}$. For each n,

$$\begin{aligned} P_0(x) &= \mathbb{P}(\,\omega : \exists \,\text{open}\; U \ni x \;\text{ s.t. }\; l(\varphi(t,\omega)U) \to 0 \text{ as } t \to \infty\,) \\ &\geq \mathbb{P}(\,\omega : l(\varphi(t,\omega)U_n) \to 0 \text{ as } t \to \infty\,) \\ &= 1 - \rho(U_n). \end{aligned}$$

But since ρ is atomless, $\rho(U_n) \to 0$ as $n \to \infty$. Hence $P_0(x) = 1$. So φ is everywhere stable (and therefore uniformly stable). □

Now since $\mathbb{S}^1$ is compact, there exists at least one stationary probability measure ρ of the Markov transition probabilities $(\bar{\varphi}_x^t)_{x\in X,\, t\in\mathbb{T}^+}$. If φ has reverse-minimal dynamics, then it is clear that ρ has full support, and Lemma 2.91 gives that ρ is atomless. Combining this with Corollary 5.23 completes the proof that (ii)⇒stable synchronisation.

Finally, to show that (ii)⇒(i), we use the following corollary of Lemma 5.22:

Corollary 5.24. *Under the hypotheses of Lemma 5.21, φ is two-way contractible.*

Proof. For any connected $J \subset \mathbb{S}^1$ with $0 < l(J) < 1$, since ρ has full support, $\rho(J) < 1$. Hence, by Lemma 5.22, there is a positive-measure set of sample points ω such that $l(\varphi(t,\omega)J) \to 0$ as $t \to \infty$. So in particular, φ is two-way contractible. □

Combining Corollary 5.24 with the fact that reverse-minimality implies the existence of a $(\bar{\varphi}_x^t)$-stationary probability measure that is atomless and has full support yields that (ii)⇒(i).

5.3 Example: Additive-noise SDE on $\mathbb{S}^1$

We now demonstrate an application of Theorem 5.19: we will see that for a "generic" vector field on $\mathbb{S}^1$, the system resulting from an additive superposition of Gaussian white noise over this vector field is stably synchronising. Specifically, stable synchronisation occurs when the vector field has no subperiodicity. We will also see what happens when the vector field does have subperiodicity.

Recall that the Wiener measure in Example 2.6 has full support (with respect to the topology of uniform convergence on compact sets).

Theorem 5.25. *Given any Lipschitz periodic function $b\colon \mathbb{R} \to \mathbb{R}$ with least period 1, and any $\sigma \in \mathbb{R} \setminus \{0\}$, the RDS on $\mathbb{S}^1$ generated by the SDE $d\phi_t = b(\phi_t)dt + \sigma dW_t$ is stably synchronising.*

Lemma 5.26. *Let $b\colon \mathbb{R} \to \mathbb{R}$ be a continuous periodic function, and let $k > 0$ be a value that is* not *a period of b. Then there exists $a \in \mathbb{R}$ such that $b(a+k) < b(a)$.*

Proof of Lemma 5.26. Let $m > 0$ be a period of b. To prove the result, we assume that there exists $a' \in [0, m)$ such that $b(a'+k) > b(a')$ and show that this implies the existence of a point $a \in [0, m)$ such that $b(a+k) < b(a)$. Note that since b is m-periodic,

$$\int_0^m b(x+k)\,dx \;=\; \int_0^m b(x)\,dx.$$

But since b is continuous, there exists $\varepsilon \in (0, m-a')$ such that

$$\int_{a'}^{a'+\varepsilon} b(x+k)\,dx \;>\; \int_{a'}^{a'+\varepsilon} b(x)\,dx.$$

Hence it is clear that there exists $a \in [0, m)$ such that $b(a+k) < b(a)$. □

Proof of Theorem 5.25. Let φ be the RDS generated by the SDE $d\phi_t = b(\phi_t)dt + \sigma dW_t$. We first show that φ has no deterministic fixed points: for any $p \in \mathbb{S}^1$, if we take $\omega(t) = \frac{k-b(p)}{\sigma}t$ for some arbitrary $k \neq 0$, then the function $u\colon t \mapsto \varphi(t,\omega)p$ satisfies the differential equation $\dot{u} = b(u) - b(p) + k$ (to be interpreted in the obvious way) and so it is not the case that $\varphi(t,\omega)p = p$ for all t; so by Remark 2.73, p is not a deterministic fixed point. We next show that φ is two-way contractible. Assume without loss of generality that $\sigma > 0$. Fix a connected set $J \subset \mathbb{S}^1$ with $0 < l(J) < 1$, and let $[c_1, c_2] \subset \mathbb{R}$ be a lift of $\bar{J}$ (so $c_2 - c_1 = l(J)$). Since b is continuous and periodic but not $l(J)$-periodic, by Lemma 5.26 there exists $a \in \mathbb{R}$ such that $b(a + l(J)) < b(a)$; obviously, we can choose a to be larger than c_1. Pick any $0 < k < b(a) - b(a + l(J))$, and let $\varepsilon > 0$ be such that for all $x_1 \in (a-\varepsilon, a+\varepsilon)$ and $x_2 \in (a + l(J) - \varepsilon, a + l(J) + \varepsilon)$,

$$b(x_1) \;>\; b(x_2) + k.$$

Let $M := \max_{x \in \mathbb{R}} |b(x)|$, and pick any $\delta > 0$ with

$$\delta \;<\; \min\left(\frac{k\varepsilon}{4M}, \frac{\varepsilon}{2}\right).$$

Let $\eta > 0$ be a value sufficiently large sufficiently that

$$\frac{M}{\sigma\eta} \;<\; \frac{\delta}{a - c_1}.$$

Let $\omega \in \Omega$ be a sample point such that

$$\omega(t) \;=\; \begin{cases} \eta t & t \in [0, \frac{a-c_1}{\sigma\eta}] \\ \frac{a-c_1}{\sigma} & t \in [\frac{a-c_1}{\sigma\eta}, \infty). \end{cases}$$

Let $u_1, u_2\colon [0,\infty) \to \mathbb{R}$ be lifts of the functions $t \mapsto \varphi(t,\omega)\pi(c_1)$ and $t \mapsto \varphi(t,\omega)\pi(c_2)$ respectively, such that $u_1(0) = c_1$ and $u_2(0) = c_2$. For each $t \in (0, \frac{a-c_1}{\sigma\eta})$, we have that

$$\dot{u}_1(t) \;=\; b(u_1(t)) + \sigma\eta \;\in\; \left(\sigma\eta(1 - \tfrac{M}{\sigma\eta}), \sigma\eta(1 + \tfrac{M}{\sigma\eta})\right) \;\subset\; \left(\sigma\eta(1 - \tfrac{\delta}{a-c_1}), \sigma\eta(1 + \tfrac{\delta}{a-c_1})\right)$$

and likewise

$$\dot{u}_2(t) \;\in\; \left(\sigma\eta(1 - \tfrac{\delta}{a-c_1}), \sigma\eta(1 + \tfrac{\delta}{a-c_1})\right).$$

Hence, we have that

$$u_1(\tfrac{a-c_1}{\sigma\eta}) \;\in\; (a-\delta, a+\delta)$$

and

$$u_2(\tfrac{a-c_1}{\sigma\eta}) \;\in\; (a+l(J)-\delta, a+l(J)+\delta).$$

Now for each $t \in (\frac{a-c_1}{\sigma\eta}, \infty)$, we have that

$$\dot{u}_1(t) \;=\; b(u_1(t)) \quad \text{and} \quad \dot{u}_2(t) \;=\; b(u_2(t)).$$

Suppose for a contradiction that there exists $t \in (\frac{a-c_1}{\sigma\eta}, \frac{a-c_1}{\sigma\eta} + \frac{\varepsilon}{2M})$ such that

$$u_1(t) \;\notin\; (a-\varepsilon, a+\varepsilon).$$

Let

$$t_1 \;:=\; \min\{t \geq \tfrac{a-c_1}{\sigma\eta} : u_1(t) \notin (a-\varepsilon, a+\varepsilon)\}.$$

Obviously $u(t_1)$ is equal to either $a-\varepsilon$ or $a+\varepsilon$. So (by the mean value theorem), there exists $t_2 \in (\frac{a-c_1}{\sigma\eta}, t_1)$ such that

$$\begin{aligned}
|b(u_1(t_2))| \;=\; |\dot{u}_1(t_2)| \;&>\; \frac{\varepsilon-\delta}{t_2 - \frac{a-c_1}{\sigma\eta}} \\
&>\; \frac{\varepsilon-\delta}{\left(\frac{\varepsilon}{2M}\right)} \\
&>\; \frac{\frac{1}{2}\varepsilon}{\left(\frac{\varepsilon}{2M}\right)} \\
&=\; M,
\end{aligned}$$

contradicting the fact that $M = \max_{x\in\mathbb{R}} |b(x)|$. So then, we have that

$$u_1(t) \;\in\; (a-\varepsilon, a+\varepsilon) \quad \forall\; t \in (\tfrac{a-c_1}{\sigma\eta}, \tfrac{a-c_1}{\sigma\eta} + \tfrac{\varepsilon}{2M}).$$

Likewise, we have that

$$u_2(t) \;\in\; (a+l(J)-\varepsilon, a+l(J)+\varepsilon) \quad \forall\; t \in (\tfrac{a-c_1}{\sigma\eta}, \tfrac{a-c_1}{\sigma\eta} + \tfrac{\varepsilon}{2M}).$$

So then,

$$\dot{u}_2(t) - \dot{u}_1(t) \;<\; -k \quad \forall\; t \in (\tfrac{a-c_1}{\sigma\eta}, \tfrac{a-c_1}{\sigma\eta} + \tfrac{\varepsilon}{2M})$$

and therefore

$$\begin{aligned}
u_2(\tfrac{a-c_1}{\sigma\eta} + \tfrac{\varepsilon}{2M}) - u_1(\tfrac{a-c_1}{\sigma\eta} + \tfrac{\varepsilon}{2M}) \;&<\; u_2(\tfrac{a-c_1}{\sigma\eta}) - u_1(\tfrac{a-c_1}{\sigma\eta}) - \frac{k\varepsilon}{2M} \\
&=\; \Big(u_2(\tfrac{a-c_1}{\sigma\eta}) - (a+l(J))\Big) + \Big(a - u_1(\tfrac{a-c_1}{\sigma\eta})\Big) + l(J) - \frac{k\varepsilon}{2M} \\
&<\; \Big|u_2(\tfrac{a-c_1}{\sigma\eta}) - (a+l(J))\Big| + \Big|a - u_1(\tfrac{a-c_1}{\sigma\eta})\Big| + l(J) - \frac{k\varepsilon}{2M} \\
&<\; 2\delta + l(J) - \frac{k\varepsilon}{2M} \\
&<\; 2\delta + l(J) - 2\delta \\
&=\; l(J).
\end{aligned}$$

Hence

$$l(\varphi(\tfrac{a-c_1}{\sigma\eta} + \tfrac{\varepsilon}{2M}, \omega)J) \;<\; l(J).$$

So (by Remark 5.17), φ is two-way contractible.

Since φ has no deterministic fixed points and is two-way contractible, Theorem 5.19 gives that φ is stably synchronising. □

As an example, consider the SDE

$$d\phi_t \;=\; (a + \varepsilon\cos(2\pi\phi_t))\,dt \;+\; \sigma dW_t$$

where $\varepsilon \neq 0$. In the deterministic case where $\sigma = 0$, we have the following: for $|a| < |\varepsilon|$, there is one repelling fixed point and one attracting fixed point, whose basin of attraction is the whole circle minus the repelling fixed point; for $|a| = |\varepsilon|$, there is exactly one fixed point, which attracts every orbit but is not Lyapunov stable; and for $|a| > |\varepsilon|$, there are no fixed points, and all orbits move periodically round the circle with the same periodicity. (So the system exhibits a *saddle-node bifurcation* as a increases past $|\varepsilon|$ or decreases past $-|\varepsilon|$.) However, when noise is incorporated—i.e. when $\sigma \neq 0$—Theorem 5.25 gives that the associated RDS is stably synchronising for *all* values of a. Hence we can say that for $|a| \geq |\varepsilon|$, the addition of noise has the effect of "creating" synchronisation, i.e. the phenomenon of "noise-induced synchronisation" occurs. In terms of random attractors and repellers: by Proposition 5.13, when noise is incorporated, we have that for *all* values of a there is one repelling random fixed point (namely, the crack point) and one attracting random fixed point. (So noise destroys the saddle-node bifurcation.)

We now consider the case that the least period of b is not 1. Obviously if b is a constant function then there cannot be synchronisation, since (under any realisation of the noise) any two trajectories stay the same distance apart. If the least period of b is $\frac{1}{n}$ for some $n \geq 2$, then the RDS is not contractible on $\mathbb{S}^1$, since any two trajectories starting at distance $\frac{1}{n}$ apart will remain at distance $\frac{1}{n}$ apart; nonetheless, there will still be some local synchronisation:

Corollary 5.27. *Let $b\colon\mathbb{R} \to \mathbb{R}$ be a Lipschitz periodic function with least period $\frac{1}{n}$ (for some $n \in \mathbb{N}$), and fix any $\sigma \in \mathbb{R} \setminus \{0\}$. Let φ be the RDS on $\mathbb{S}^1$ generated by the SDE $d\phi_t = b(\phi_t)dt + \sigma dW_t$. Then for $\mathbb{P}$-almost every $\omega \in \Omega$ there exists $p \in \mathbb{S}^1$ such that the set $\{p + \pi(\frac{k}{n})\}_{k=0,\dots,n-1}$ is a repulsive crack set of ω. Consequently, there is a unique stationary probability measure ρ for the Markov transition probabilities $(\varphi_x^t)_{x\in X,\, t\in\mathbb{T}^+}$, and the ρ-clustering number of φ is precisely n.*

Lemma 5.28. *Let $b\colon\mathbb{R} \to \mathbb{R}$ be a Lipschitz $\frac{1}{n}$-periodic function (for some $n \in \mathbb{N}$), and fix any $\sigma \in \mathbb{R}$. Let φ be the RDS on $\mathbb{S}^1$ generated by the SDE $d\phi_t = b(\phi_t)dt + \sigma dW_t$. Let φ' be the RDS on $\mathbb{S}^1$ generated by the SDE $d\phi_t = nb(\frac{1}{n}\phi_t)dt + n\sigma dW_t$. Then, for any $\omega \in \Omega$, $y \in \mathbb{S}^1$ and $k \in \mathbb{Z}$, letting $u\colon[0,\infty) \to \mathbb{R}$ be a lift of the map $t \mapsto \varphi'(t,\omega)y$, we have that*

$$\varphi(t,\omega)\pi(\tfrac{1}{n}u(0) + \tfrac{k}{n}) \;=\; \pi(\tfrac{1}{n}u(t) + \tfrac{k}{n}) \qquad (5.4)$$

for all $t \geq 0$. Hence the map $h\colon x \mapsto nx$ is a deterministic semiconjugacy from φ to φ'.

Proof of Lemma 5.28. Fix $k \in \mathbb{Z}$. Let $v(t) := \frac{1}{n}u(t) + \frac{k}{n}$ for all t. Then for all $t \geq 0$,

$$\begin{aligned} v(t) &= \frac{1}{n}\left(u(0) + \int_0^t nb(\tfrac{1}{n}u(s))\,ds + n\sigma\omega(t)\right) + \frac{k}{n} \\ &= v(0) + \int_0^t b(v(s) - \tfrac{k}{n})\,ds + \sigma\omega(t) \\ &= v(0) + \int_0^t b(v(s))\,ds + \sigma\omega(t). \end{aligned}$$

This proves (5.4). Now for any $x \in \mathbb{S}^1$, if we fix a lift $x' \in \mathbb{R}$ of x and take

$$y := h(x),\ u(0) := nx',\ k := 0,$$

then applying h to both sides of (5.4) gives that $h(\varphi(t,\omega)x) = \varphi'(t,\omega)y$. So h is a deterministic semiconjugacy from φ to φ'. □

Proof of Corollary 5.27. Let φ, φ' and h be as in Lemma 5.28. By Theorem 5.25, φ' is stably synchronising. Hence, by Theorem 5.6, $\mathbb{P}$-almost every $\omega \in \Omega$ admits a repulsive crack point $r(\omega)$ under φ'. By Proposition 5.15 and Lemma 5.28, it follows that $h^{-1}(\{r(\omega)\})$ is a repulsive crack set of ω under φ. Obviously, $h^{-1}(\{r(\omega)\})$ takes the form $\{p + \pi(\frac{k}{n})\}_{k=0,\ldots,n-1}$ for some $p \in \mathbb{S}^1$.

Now recall, from the proof of Theorem 5.25, the construction of a sample point ω taking the trajectory of c_1 into the arc with lift $(a-\delta, a+\delta)$; this construction demonstrates in general that every (deterministic) non-empty open set is accessible from every point in $\mathbb{S}^1$. So φ has minimal dynamics on $\mathbb{S}^1$. Now let ρ be an ergodic probability measure of (φ_x^t). Since $\mathbb{S}^1$ is minimal, ρ must have full support. Since $\mathbb{P}$-almost every $\omega \in \Omega$ admits a crack set, it is clear that φ is stable with respect to ρ. But since every stationary probability measure of (φ_x^t) must have full support, it then follows by Lemma 4.34 that ρ is the only stationary probability measure of (φ_x^t).[4]

Now let k be the ρ-clustering number of φ. Since φ is stable with respect to ρ, Theorem 4.52 gives that $k < \infty$. For $\mathbb{P}$-almost every $\omega \in \Omega$, letting $U_1, \ldots, U_k$ be as described in Theorem 4.52, we have that $U_1, \ldots, U_k$ are contained in distinct equivalence classes of $\sim_\omega$ and (since ρ has full support) $\bigcup_{i=1}^k U_i$ is dense in $\mathbb{S}^1$. Hence (by a simple "pigeonhole principle" argument) we have that $k = n$. □

As an example, consider the SDE

$$d\phi_t = (a + \varepsilon\cos(2\pi n\phi_t))\,dt + \sigma dW_t$$

where $\varepsilon \neq 0$ and $n \geq 2$. In the deterministic case where $\sigma = 0$, we have the following: for $|a| < |\varepsilon|$ there are n repelling fixed points and n attracting fixed points, with the basin of attraction of each attracting fixed point being the open interval connecting two consecutive repelling fixed points; for $|a| = |\varepsilon|$, there are exactly n fixed points, with

[4] Alternatively: it is not hard to show, using the *strong Markov property*, that for any continuous RDS φ on $\mathbb{S}^1$ with $\mathbb{T} = \mathbb{R}$, the interiors of the supports of two distinct (φ_x^t)-ergodic probability measures must be mutually disjoint. Hence in particular, if φ has minimal dynamics, then there is only one (φ_x^t)-stationary probability measure.

heteroclinic connections between consecutive fixed points; and for $|a| > |\varepsilon|$, there are no fixed points, and all orbits move periodically round the circle with the same periodicity. However, when noise is incorporated—i.e. when $\sigma \neq 0$—the conclusions of Corollary 5.27 hold: the addition of noise does not have the effect of causing "global" synchronisation, but synchronisation within intervals of length $\frac{1}{n}$ does occur. (Thus one can still say that for $|a| \geq |\varepsilon|$, the phenomenon of "noise-induced synchronisation" occurs.) In terms of random attractors and repellers: when noise is incorporated, for all values of a there is a "random repeller" consisting of n points (namely, the crack set) and a "random attractor" consisting of n points (namely, the support of the unique Markov invariant measure). Dynamically, this scenario is somewhat analogous to the dynamics exhibited by the discrete-time dynamical system f on $\mathbb{S}^1$ given by the lift

$$F(x) \;=\; x + \varepsilon\cos(2\pi n x) + \tfrac{1}{n}$$

where $\varepsilon \neq 0$ is small. (This dynamical system also has an n-point repeller, namely the periodic orbit $\left\{\frac{4k+3}{4n}\right\}_{k=0}^{n-1}$, and an n-point attractor, namely the periodic orbit $\left\{\frac{4k+1}{4n}\right\}_{k=0}^{n-1}$.)

Chapter 6. General Synchronisation Tests

So far, we have seen criteria for synchronisation in monotone RDS and in orientation-preserving RDS on the circle. In this chapter, we present two tests for synchronisation in RDS on more general phase spaces.

Standing Assumption. *Throughout Chapter 6, we fix a separable metric d on X whose Borel σ-algebra coincides with Σ, and we assume that φ is a right-continuous RDS on the metric space (X, d).*

(Recall once again that the condition that the Borel σ-algebra of d is standard is equivalent to the condition that X is a Borel subset of the d-completion of X.)

As in Chapters 4 and 5, for each $\omega \in \Omega$, $\sim_\omega$ denotes the equivalence relation

$$x \sim_\omega y \quad \Longleftrightarrow \quad d(\,\varphi(t,\omega)x\,,\,\varphi(t,\omega)y\,) \to 0 \text{ as } t \to \infty.$$

6.1 Necessary and sufficient conditions for stable synchronisation on compact spaces

Recall that if X is compact then φ admits at least one minimal set $K \subset X$. Also recall that every compact minimal set can be expressed as the support of an ergodic probability measure of the Markov transition probabilities $(\varphi_x^t)_{x \in X,\, t \in \mathbb{T}^+}$. The following result is the main result of [New15b].

Theorem 6.1. *Suppose X is compact. Then φ is stably synchronising if and only if the following conditions hold:*

(i) there is a unique minimal set $K \subset X$;

(ii) φ is contractible on the unique minimal set K;

(iii) the unique minimal set K admits stable trajectories.

Now if X is compact and φ is synchronising then (by Lemma 4.8) there is a unique stationary probability measure of the Markov transition probabilities $(\varphi_x^t)_{x \in X,\, t \in \mathbb{T}^+}$. Hence Theorem 6.1 can be re-expressed as follows:

Corollary 6.2. *Suppose X is compact, and let ρ be a stationary probability measure of the Markov transition probabilities $(\varphi_x^t)_{x \in X,\, t \in \mathbb{T}^+}$. Then φ is stably synchronising if and only if the following conditions hold:*

(i) $\operatorname{supp} \rho$ *is the only minimal set;*

(ii) φ is contractible on $\operatorname{supp} \rho$*;*

(iii) $\operatorname{supp} \rho$ *admits stable trajectories.*

Note that a sufficient condition for $\operatorname{supp}\rho$ to be the only minimal set is that ρ is the only (φ_x^t)-stationary probability measure.

Proof of Theorem 6.1. If φ is synchronising then obviously φ is contractible on X. So Lemma 4.69 gives that (i) holds; and therefore, obviously, (ii) holds. If φ is stably synchronising then (iii) also holds.

Now suppose that (i), (ii) and (iii) hold. By Proposition 4.40, (i) and (iii) imply that φ is uniformly stable; so we just need to establish that φ is synchronising. Let $C \subset X \times X$ be any non-empty closed set that is invariant under the two-point motion $\varphi^{\times 2}$. By the final assertion in Proposition 4.40, it is clear that C has non-empty intersection with $K \times K$. Now by (ii), Proposition 4.68 gives that any non-empty closed invariant subset of $K \times K$ has non-empty intersection with Δ_K. Hence C has non-empty intersection with Δ_K. Now it is clear that for any closed $G \subset X$, G is invariant if and only if Δ_G is invariant under $\varphi^{\times 2}$; so then, since K is minimal, it is clear that Δ_K is minimal under $\varphi^{\times 2}$. Hence, since C has non-empty intersection with Δ_K, it follows that $\Delta_K \subset C$. Now recall that C was an arbitrary non-empty closed invariant set under $\varphi^{\times 2}$. So then, we have seen that Δ_K is contained in every non-empty closed invariant set under $\varphi^{\times 2}$. So Δ_K is the only minimal set under $\varphi^{\times 2}$.

Now fix any $x, y \in X$. Fix a point $p \in K$, and for each $n \in \mathbb{N}$ and $s \in \mathbb{Q} \cap \mathbb{T}^+$, let

$$\begin{aligned} R_{n,s} &= \left\{\omega \in \Omega : (\varphi(s,\omega)x, \varphi(s,\omega)y) \in B_{\frac{1}{n}}(p) \times B_{\frac{1}{n}}(p)\right\} \\ S_{n,s} &= \theta^{-s}\left(E_{B_{\frac{1}{n}}(p)}\right). \end{aligned}$$

Note that for every n and s, $\sigma(R_{n,t} : t \leq s) \subset \mathcal{F}_0^s$ and $S_{n,s} \in \mathcal{F}_s^\infty$. Since K is the only minimal set under $\varphi^{\times 2}$, Corollary 2.81 gives that $\mathbb{P}(\bigcup_s R_{n,s}) = 1$ for all n. Obviously $\mathbb{P}(S_{n,s}) = P_{\frac{1}{n}}(p)$ for all n and s, and so $\mathbb{P}(S_{n,s}) \to 1$ as $n \to \infty$ uniformly in s. So then, Lemma 4.41 gives that

$$\mathbb{P}\left(\bigcup_n \bigcup_s R_{n,s} \cap S_{n,s}\right) = 1. \tag{6.1}$$

Now for any n and s, for any $\omega \in R_{n,s} \cap S_{n,s}$, we clearly have that $x \sim_\omega y$. Hence (6.1) gives that for $\mathbb{P}$-almost all $\omega \in \Omega$, $x \sim_\omega y$. Since x and y were arbitrary, φ is synchronising. □

Example 6.3. This example is taken from Section 4 of [New15b] (which is itself an extension of the "no subperiodicity" case of the example in [LeJ87]). Let $X = \mathbb{S}^1$, which we identify with $\mathbb{R}/\mathbb{Z}$ in the obvious manner. Recall that for any continuous function $f : \mathbb{S}^1 \to \mathbb{S}^1$ there exists $k \in \mathbb{Z}$ (called the *degree of* f) such that for any lift $F : \mathbb{R} \to \mathbb{R}$ of f, the map $y \mapsto F(y) - ky$ is 1-periodic. If f is an orientation-preserving homeomorphism, the degree of f is 1. Following the terminology of [Kai93], if f has degree 1 then a *subperiod of* f is a value $\alpha \in (0,1)$ such that the map $y \mapsto F(y) - y$ is α-periodic. Now fix any smooth function $f : \mathbb{S}^1 \to \mathbb{S}^1$ of degree 1, and let $F : \mathbb{R} \to \mathbb{R}$ be a lift of f. Let $I = [0,1)$, with $\mathcal{I}$ being the Borel σ-algebra of I, and let ν be the Lebesgue measure on I. For each $\alpha \in I$, define $f_\alpha : \mathbb{S}^1 \to \mathbb{S}^1$ by

$$f_\alpha(\pi(x)) = \pi(F(x + \alpha) - \alpha)$$

(where $\pi: \mathbb{R} \to \mathbb{S}^1$ denotes the natural projection). Let φ be the RDS generated by the random map $(I, \mathcal{I}, \nu, (f_\alpha)_{\alpha \in I})$. It is easy to show that the Lebesgue measure l on $\mathbb{S}^1$ is stationary under the Markov transition probabilities $(\varphi_x^n)_{x \in \mathbb{S}^1, n \in \mathbb{N}_0}$. Define $\lambda \in [-\infty, \infty)$ by

$$\lambda := \int_0^1 \log |F'(y)| \, dy.$$

If l is ergodic with respect to the Markov transition probabilities (φ_x^n), then λ is precisely the "Lyapunov exponent" associated to l; by [LeJ87, Lemme 3], if $\lambda < 0$ then φ is stable with respect to l. If l is *not* (φ_x^n)-ergodic but $\lambda < 0$, then (due to the existence of an ergodic decomposition of l) there exists a (φ_x^n)-ergodic probability measure ρ for which the associated Lyapunov exponent is negative; and so once again, φ is stable with respect to ρ. Hence, in either case, if $\lambda < 0$ then $\mathbb{S}^1$ admits stable trajectories. (If f is a diffeomorphism, then due to the strict form of Jensen's inequality, we automatically have that $\lambda < 0$.) Now one can show that $\mathbb{S}^1$ is minimal if and only if f is not a rational rotation; and one can show that φ is contractible on $\mathbb{S}^1$ if and only if f has no subperiods. (Note that any rational rotation *must* have a subperiod.) So then, applying Theorem 6.1, we have the following: *if f has no subperiods and $\lambda < 0$ then φ is stably synchronising.*

6.2 Necessary and sufficient conditions for ρ-almost stable synchronisation

If we have a $(\varphi_x^t)_{x \in X, t \in \mathbb{T}^+}$-ergodic probability measure ρ on X such that φ is stable with respect to ρ, a natural question to ask is whether φ is ρ-almost stably synchronising. To phrase the issue another way: In Theorem 4.52, when do we have that $n = 1$? We will now answer this question.

As in Chapter 4, define the equivalence relation $\sim$ on X by

$$x \sim y \iff \mathbb{P}(\omega : x \sim_\omega y) = 1.$$

Definition 6.4. Let ρ be a probability measure on X. A *ρ-full-length rectangle* is a set $A \subset X \times X$ taking the form $A = A_1 \times A_2$ where $A_1, A_2 \in \mathcal{B}(X)$ with $\rho(A_1) > 0$ and $\rho(A_2) = 1$.

Now let ρ be an ergodic probability measure of the Markov transition probabilities $(\varphi_x^t)_{x \in X, t \in \mathbb{T}^+}$. For any $x \in \operatorname{supp} \rho$, it is clear that either $\rho(G_x) = 0$ or $G_x = \operatorname{supp} \rho$.

Definition 6.5. Given an ergodic probability measure ρ of the Markov transition probabilities $(\varphi_x^t)_{x \in X, t \in \mathbb{T}^+}$ and a point $x \in \operatorname{supp} \rho$, we will say that x is *ρ-transitive* if $G_x = \operatorname{supp} \rho$.

Let A_ρ denote the set of ρ-transitive points. By Lemma 2.76, A_ρ is a ρ-full set.

For any $p \in X$, we write $\mathfrak{C}_p \subset X \times X$ for the set of pairs that are contractible towards p. For any $A \subset X$, we write $\mathfrak{C}_A \subset X \times X$ for the set of pairs that are contractible towards A.

The following is the main result of [New16].

Theorem 6.6. *Let ρ be an ergodic probability measure of the Markov transition probabilities $(\varphi_x^t)_{x\in X,\, t\in\mathbb{T}^+}$, and suppose that φ is stable with respect to ρ.*[1] *The following statements are equivalent:*

(i) there is a non-ρ-null set $R \subset X$ such that for each $p \in R$, the set $\mathfrak{C}_p$ contains a ρ-full-length rectangle;

(ii) the set $\mathfrak{C}_{A_\rho}$ contains a ρ-full-length rectangle;

(iii) φ is ρ-almost stably synchronising;

(iv) there is a ρ-full set $A \subset X$ such that for all $x, y \in A$ and $p \in \operatorname{supp}\rho$, (x, y) is contractible towards p.

Condition (ii) is likely to be the most useful in practice for testing whether we have ρ-almost stable synchronisation.

The "non-trivial" part is showing that (ii)⇒(iii). Our proof generalises a technique in [Hom13, Proof of Theorem 1.1].

Proof. Suppose (i) holds; then since A_ρ is a ρ-full set, $A_\rho \cap R \neq \varnothing$, and so there exists $p \in A_\rho$ such that $\mathfrak{C}_p$ contains a ρ-full-length rectangle, implying (ii).

Now suppose that (ii) holds. Let Q_ρ be the statistical equilibrium associated to ρ. Let $(\Theta_{[2]}^t)_{t\in\mathbb{T}^+}$ be the skew product flow associated to $\varphi^{\times 2}$, that is

$$\Theta_{[2]}^t(\omega, x, y) \;:=\; (\theta^t\omega, \varphi(t,\omega)x, \varphi(t,\omega)y).$$

By Lemma 2.21(A), $\mathbb{P}|_{\mathcal{F}_0^\infty} \otimes E_2(Q_\rho)$ is $(\Theta_{[2]}^t)_{t\in\mathbb{T}^+}$-invariant. For each $\varepsilon > 0$, let $U_\varepsilon := \{(x,y) \in X \times X : d(u,v) < \varepsilon\}$. Obviously $\Omega \times \Delta_X = \bigcap_{k=1}^\infty \Omega \times U_{\frac{1}{k}}$, and so writing

$$\begin{aligned} Z &:= \{\, (\omega, x, y) : d(\varphi(j,\omega)x, \varphi(j,\omega)y) \to 0 \text{ as } j \to \infty \,\} \\ &= \bigcap_{k=1}^{\infty} \bigcup_{i=0}^{\infty} \bigcap_{j=i}^{\infty} \Theta_{[2]}^{-j}(\Omega \times U_{\frac{1}{k}}), \end{aligned}$$

we have that $\mathbb{P} \otimes E_2(Q_\rho)(Z) \leq E_2(Q_\rho)(\Delta_X)$. But it is also clear that $\Omega \times \Delta_X \subset Z$. Therefore $\mathbb{P} \otimes E_2(Q_\rho)(Z \setminus (\Omega \times \Delta_X)) = 0$. Hence, by Fubini's theorem, the set

$$Y \;:=\; \{\, (x,y) \in (X \times X) \setminus \Delta_X : \mathbb{P}(\omega : x \sim_\omega y) > 0 \,\}$$

is an $E_2(Q_\rho)$-null set. Now let $A_1, A_2 \in \mathcal{B}(X)$ be such that $\rho(A_1) > 0$, $\rho(A_2) = 1$ and $A_1 \times A_2 \subset \mathfrak{C}_{A_\rho}$. We will show that for any $(x,y) \in A_1 \times A_2$, $\mathbb{P}(\omega : x \sim_\omega y) > 0$. Fix any $(x,y) \in A_1 \times A_2$, and let $p \in A_\rho$ be such that (x,y) is contractible towards p. Let $U, V \subset X$ be open sets with $\bar{U} \subset V$, $\rho(U) > 0$ and $\mathbb{P}(E_V) > 0$; and let $t_1 \in \mathbb{T}^+$ be such that $\varphi_p^{t_1}(U) > 0$. Since $\varphi(t_1,\omega)$ is continuous for all ω, let $r > 0$ be such that

$$k_1 := \mathbb{P}(\omega : \varphi(t_1,\omega)\overline{B_r(p)} \subset \bar{U}) \;>\; 0$$

[1] Recall that by Lemma 4.39, this is precisely the same as saying that $\operatorname{supp}\rho$ admits stable trajectories.

and let $t_0 \in \mathbb{T}^+$ be such that

$$k_0 := \mathbb{P}(\,\omega : \varphi(t_0,\omega)x, \varphi(t_0,\omega)y \in B_r(p)\,) \;>\; 0.$$

Then we have that

$$\begin{aligned}&\mathbb{P}(\omega : x \sim_\omega y)\\ &\geq\; \mathbb{P}(\,\omega : \varphi(t_0,\omega)x, \varphi(t_0,\omega)y \in B_r(p) \;\text{ and }\; \varphi(t_1,\theta^{t_0}\omega)\overline{B_r(p)} \subset \bar{U} \;\text{ and }\; \theta^{t_0+t_1} \in E_V\,)\\ &=\; k_0 k_1 \mathbb{P}(E_V)\\ &>\; 0\end{aligned}$$

as required. So in particular, $(A_1 \times A_2) \smallsetminus \Delta_X \subset Y$. Now since $E_1(Q_\rho) = \rho$, we have that $\tilde{\rho}(A_2) = 1$ for Q_ρ-almost all $\tilde{\rho} \in \mathcal{M}$, and therefore

$$E_2(Q_\rho)(A_1 \times A_2) \;=\; \int_{\mathcal{M}} \tilde{\rho}(A_1)\tilde{\rho}(A_2)\, Q_\rho(d\tilde{\rho}) \;=\; \int_{\mathcal{M}} \tilde{\rho}(A_1)\, Q_\rho(d\tilde{\rho}) \;=\; \rho(A_1).$$

Let n be the ρ-clustering number of φ. (Since φ is stable with respect to ρ, $n < \infty$.) We have that

$$E_2(Q_\rho)((A_1 \times A_2) \cap \Delta_X) \;=\; E_2(Q_\rho)(\Delta_{A_1 \cap A_2}) \;=\; \tfrac{1}{n}\rho(A_1)$$

by Lemma 3.4, and therefore

$$E_2(Q_\rho)\,((A_1 \times A_2) \smallsetminus \Delta_X) \;=\; \tfrac{n-1}{n}\rho(A_1).$$

But since $(A_1 \times A_2) \smallsetminus \Delta_X \subset Y$, we have that

$$E_2(Q_\rho)\,((A_1 \times A_2) \smallsetminus \Delta_X) \;=\; 0.$$

Since $\rho(A_1) \neq 0$, it obviously follows that $n = 1$, i.e. (iii) holds.

Now suppose that (iii) holds; we show that (iv) holds. Let A be the ρ-full-measure equivalence class of $\sim$, and (on the basis of Lemma 2.76) let $z \in A$ be a point with the property that for $\mathbb{P}$-almost all $\omega \in \Omega$, for every $T \in \mathbb{T}^+$, $\{\varphi(t,\omega)z : t \geq T\}$ is dense in $\operatorname{supp}\rho$. Fix any $x, y \in A$, and any $p \in \operatorname{supp}\rho$ and $\varepsilon > 0$. Let $T \in \mathbb{T}^+$ be such that the set

$$E \;:=\; \{\,\omega : \forall\, t \geq T,\; \varphi(t,\omega)x, \varphi(t,\omega)y \in B_{\frac{\varepsilon}{2}}(\varphi(t,\omega)z)\,\}$$

has positive measure. For $\mathbb{P}$-almost every $\omega \in E$, there exists $t \geq T$ such that $\varphi(t,\omega)z \in B_{\frac{\varepsilon}{2}}(p)$ and therefore $\varphi(t,\omega)x, \varphi(t,\omega)y \in B_\varepsilon(p)$.

Finally, it is clear that (iv)⇒(i). □

Example 6.7 (Single- and double-well potentials, cf. [FGS14], [New16], [Cal+13])**.** Let $X = \mathbb{R}^d$ (equipped with the Euclidean metric) for some $d \in \mathbb{N}$. Let $V : \mathbb{R}^d \to \mathbb{R}$ be a radially symmetric polynomial of order at most 4, that is

$$V(x) \;=\; \alpha_2|x|^4 + \alpha_1|x|^2 + \alpha_0$$

where $\alpha_0, \alpha_1, \alpha_2 \in \mathbb{R}$. We say that V is a *well potential* if either $\alpha_2 > 0$, or $\alpha_2 = 0$ and $\alpha_1 > 0$. We say that V is a *hill potential* if either $\alpha_2 < 0$, or $\alpha_2 = 0$ and $\alpha_1 < 0$. If $\alpha_2 = \alpha_1 = 0$, then we say that V is a *flat potential*. We can divide the case that V is a well potential into two cases: (a) if either $\alpha_2 > 0$ and $\alpha_1 \geq 0$, or $\alpha_2 = 0$ and $\alpha_1 > 0$, then V has its global minimum at $x = \mathbf{0}$, and we refer to V as a *single-well potential*; (b) if $\alpha_2 > 0$ and $\alpha_1 < 0$ then V has a local maximum at $x = \mathbf{0}$ and its global minimum throughout the ring $|x| = \sqrt{-\frac{\alpha_1}{2\alpha_2}}$, and we refer to V as a *double-well potential*. Now let $(\Omega, \mathcal{F}, (\mathcal{F}_s^{s+t})_{s\in\mathbb{R}, t\geq 0}, (\theta^t)_{t\in\mathbb{R}}, \mathbb{P})$ be as in Example 2.6. Let $b := -\nabla V$. We consider the equation

$$dx_t \;=\; b(x_t)dt + d\omega(t). \tag{6.2}$$

One can check that

$$h \cdot Db(x)h \;=\; -8\alpha_2(x \cdot h)^2 - 4\alpha_2|x|^2 - 2\alpha_1.$$

for all $x, h \in \mathbb{R}^d$ with $|h| = 1$. Hence, if $\alpha_2 \geq 0$, then b satisfies the one-sided Lipschitz condition (2.6) and so (6.2) generates a RDS φ. (In the case that $\alpha_2 = 0$ and $\alpha_1 > 0$, (6.2) describes a *d-dimensional Ornstein-Uhlenbeck process.*) *From now on, suppose that V is a well potential.* By the integrability condition in [FGS14, Section 2.2], the probability measure ρ on $\mathbb{R}^d$ with density proportional to $e^{-2V(\cdot)}$ has the property that for each $t > 0$, ρ is the unique stationary probability measure of the Markov kernel $(\varphi_x^t)_{x\in\mathbb{R}^d}$. Moreover, by [FGS14, Example 4.8], the "maximal Laypunov exponent" associated to ρ is strictly negative. As stated in Section 4 of [FGS14], one deduces that the time-1 discretisation $\mathring{\varphi}_1$ of φ (see Section 2.3) is stable with respect to ρ. By (2.7), it follows that φ is stable with respect to ρ. Now (as with any additive-noise SDE) it is not hard to see that the whole phase space $\mathbb{R}^d$ is a minimal set of φ: fix any $x \in \mathbb{R}^d$ and any non-empty open $U \subset \mathbb{R}^d$. Take any $y \in U$ and, selecting a sufficiently large value $\eta_0 > 0$, take a sample point $\omega_0 \in \Omega$ with

$$\omega_0(t) \;=\; \eta_0 t(y - x) \quad \forall\, t \in [0, \tfrac{1}{\eta_0}].$$

Then we will have that $\varphi(\frac{1}{\eta_0}, \omega_0) \in U$. So $\mathbb{R}^d$ is a minimal set of φ; note that this is precisely the same as saying that every point in $\mathbb{R}^d$ is ρ-transitive under φ. Now it is not hard to see that every $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$ is contractible towards any of the points $k \in \mathbb{R}^d$ at which V is minimal: e.g. taking k of the form $(|k|, \mathbf{0}^{(d-1)})$ and fixing any $\varepsilon > 0$, we can select sufficiently large values $\eta_1, \eta_2 > 0$ that if we take a sample point ω_1 with

$$\omega_1(t) \;=\; \begin{cases} (\eta_1\eta_2 t, \mathbf{0}^{(d-1)}) & t \in [0, \frac{1}{\eta_1}] \\ (\eta_2, \mathbf{0}^{(d-1)}) & t \in [\frac{1}{\eta_1}, \infty), \end{cases}$$

we will have that $\varphi(t, \omega_1)x, \varphi(t, \omega_1)y \in B_\varepsilon(k)$ for all sufficiently large t. So then, φ satisfies hypothesis (ii) of Theorem 6.6 (since $A_\rho = \mathbb{R}^d$ and $\mathfrak{C}_{\mathbb{R}^d} \supset \mathfrak{C}_k = \mathbb{R}^d \times \mathbb{R}^d$), so φ is ρ-almost stably synchronising. By Hörmander's theorem ([Hai11, Theorem 1.3]), φ_x^t is equivalent to the Lebesgue measure[2]—and therefore, equivalent to ρ—for all $x \in \mathbb{R}^d$ and $t > 0$. Hence we conclude that φ is actually pointwise-stably synchronising.[3] (To

[2]Direct application of [Hai11, Theorem 1.3] would require b to have bounded derivatives. Although b does not have bounded derivatives, one can multiply b by test functions ψ which are equal to 1 on arbitrarily large balls around x in order to conclude that φ_x^t is equivalent to the Lebesgue measure.

[3]In the case that V is a single-well potential, it is easy to show by elementary methods (as we will soon see) that φ is in fact globally contractive. In the case that V is a double-well potential, the author expects that by combining the facts that φ has a globally pullback-attracting random fixed point ([FGS14]) and φ is stable with respect to ρ, it will follow that φ is globally contractive.

see this, apply Lemma A.10 with $h(\omega,\tilde{\omega})$ being the characteristic function of the event that $\varphi(1,\omega)x$ and $\varphi(1,\omega)y$ synchronise under $\theta^1\tilde{\omega}$ and are each asymptotically stable under $\theta^1\tilde{\omega}$.) Now if the noise term "$+d\omega(t)$" is removed from the right-hand side of (6.2), the associated (deterministic) flow is globally synchronising in the case that V is a single-well potential, but not in the case that V is a double-well potential. Thus we have "noise-induced synchronisation" in the case that V is a double-well potential. Hence in particular, the *bifurcation* between the dynamics of the single-well potential and the dynamics of the double-well potential is "destroyed" by the presence of noise. We now consider the determinism of the rate of synchronisation. If $\alpha_2 \geq 0$ and $\alpha_1 > 0$ then the one-sided Lipschitz constant of b can be taken to be negative, and so by (2.7), φ is globally contracting at an (exponential) deterministic rate. Now if $\alpha_2 > 0$ and $\alpha_1 = 0$, it is easy to show that for any compact $K \subset (0,\infty)$ there exists $\lambda_K < 0$ such that

$$(b(y)-b(x))\cdot(y-x) \;\leq\; \lambda_K \quad \forall\, x,y \in \mathbb{R}^d \text{ with } |y-x| \in K.$$

Hence, using Grönwall's inequality, one can again show that φ is globally contracting at a deterministic rate. So then, φ is globally contracting at a deterministic rate whenever V is a single-well potential. On the other hand, if V is a double-well potential, then for any distinct $x,y \in \mathbb{R}^d$, x and y do not synchronise at a deterministic rate: we can select a sufficiently large value $\eta > 0$ that if we take a sample point ω with

$$\omega(t) \;=\; \begin{cases} -\frac{1}{2}\eta t(x+y) & t \in [0,\frac{1}{\eta}] \\ -\frac{1}{2}(x+y) & t \in [\frac{1}{\eta},\infty), \end{cases}$$

then $\varphi(t,\omega)x$ and $\varphi(t,\omega)y$ converge to different minimum points of V as $t \to \infty$. (Of course if $d > 1$ then just taking $\omega(t) \equiv 0$ would also work for a generic pair of points $x,y \in \mathbb{R}^d$.)

Bibliography

[AM14] Lluís Alsedà and Michał Misiurewicz. "Random interval homeomorphisms". *Publl. Mat.* 58.suppl. (2014), pp. 15–36. URL: http://projecteuclid.org/euclid.pm/1400505223.

[Ant84] V. A. Antonov. "Modeling of processes of cyclic evolution type. Synchronization by a random signal". *Vestnik Leningrad. Univ. Mat. Mekh. Astronom.* vyp. 2 (1984), pp. 67–76.

[Apo74] Tom M. Apostol. *Mathematical analysis.* Second. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1974, pp. xvii+492.

[Arn98] Ludwig Arnold. *Random dynamical systems.* Springer Monographs in Mathematics. Springer-Verlag, Berlin, 1998. URL: http://dx.doi.org/10.1007/978-3-662-12878-7.

[Bax91] Peter H. Baxendale. "Statistical equilibrium and two-point motion for a stochastic flow of diffeomorphisms". *Spatial stochastic processes.* Vol. 19. Progr. Probab. Birkhäuser Boston, Boston, MA, 1991, pp. 189–218.

[BBD14] Jairo Bochi, Christian Bonatti, and Lorenzo J. Díaz. "Robust vanishing of all Lyapunov exponents for iterated function systems". *Math. Z.* 276.1-2 (2014), pp. 469–503. URL: http://dx.doi.org/10.1007/s00209-013-1209-y.

[Bil95] Patrick Billingsley. *Probability and measure.* Third. Wiley Series in Probability and Mathematical Statistics. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1995.

[Bil99] Patrick Billingsley. *Convergence of probability measures.* Second. Wiley Series in Probability and Statistics: Probability and Statistics. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 1999. URL: http://dx.doi.org/10.1002/9780470316962.

[BPR06] Patrizia Berti, Luca Pratelli, and Pietro Rigo. "Almost sure weak convergence of random probability measures". *Stochastics* 78.2 (2006), pp. 91–97. URL: http://dx.doi.org/10.1080/17442500600745359.

[BS88] P. H. Baxendale and D. W. Stroock. "Large deviations and stochastic flows of diffeomorphisms". *Probab. Theory Related Fields* 80.2 (1988), pp. 169–215. URL: http://dx.doi.org/10.1007/BF00356102.

[Bur83] Theodore Allen Burton. *Volterra integral and differential equations.* Vol. 167. Mathematics in Science and Engineering. Academic Press, Inc., Orlando, FL, 1983.

[Cal+13] Mark Callaway, Thai Son Doan, Jeroen S.W. Lamb, and Martin Rasmussen. "The dichotomy spectrum for random dynamical systems and pitchfork bifurcations with additive noise". arXiv:1310.6166v1 [math.DS]. 2013.

[Car85] Andrew Carverhill. “Flows of stochastic dynamical systems: ergodic theory”. *Stochastics* 14.4 (1985), pp. 273–317. URL: http://dx.doi.org/10.1080/17442508508833343.

[CCK06] Tomás Caraballo, Igor D. Chueshov, and Peter E. Kloeden. “Synchronization of a stochastic reaction-diffusion system on a thin two-layer domain”. *SIAM J. Math. Anal.* 38.5 (2006/07), pp. 1489–1507. URL: http://dx.doi.org/10.1137/050647281.

[CF98] Hans Crauel and Franco Flandoli. “Additive noise destroys a pitchfork bifurcation”. *J. Dynam. Differential Equations* 10.2 (1998), pp. 259–274. URL: http://dx.doi.org/10.1023/A:1022665916629.

[Che+15] Anna Maria Cherubini, Jeroen S.W. Lamb, Martin Rasmussen, and Yuzuru Sato. “A random dynamical systems perspective on stochastic resonance”. arXiv:1510.06973v1 [math.DS]. 2015.

[Chu02] Igor Chueshov. *Monotone random systems theory and applications.* Vol. 1779. Lecture Notes in Mathematics. Springer-Verlag, Berlin, 2002. URL: http://dx.doi.org/10.1007/b83277.

[CKS04] Tomás Caraballo, E. Peter Kloeden, and Björn Schmalfuß. “Exponentially Stable Stationary Solutions for Stochastic EvolutionEquations and Their Perturbation”. *Applied Mathematics and Optimization* 50.3 (2004), pp. 183–207. URL: http://dx.doi.org/10.1007/s00245-004-0802-1.

[CM01] I. Csiszár and F. Matúš. “Convex cores of measures on $\mathbb{R}^d$”. *Studia Sci. Math. Hungar.* 38 (2001), pp. 177–190. URL: http://dx.doi.org/10.1556/SScMath.38.2001.1-4.12.

[Cra02a] Hans Crauel. “Invariant measures for random dynamical systems on the circle”. *Arch. Math. (Basel)* 78.2 (2002), pp. 145–154. URL: http://dx.doi.org/10.1007/s00013-002-8228-y.

[Cra02b] Hans Crauel. *Random probability measures on Polish spaces.* Vol. 11. Stochastics Monographs. Taylor & Francis, London, 2002.

[CS04] Igor Chueshov and Michael Scheutzow. “On the structure of attractors and invariant measures for a class of monotone random systems”. *Dyn. Syst.* 19.2 (2004), pp. 127–144. URL: http://dx.doi.org/10.1080/1468936042000207792.

[CSS05] Igor Chueshov, Michael Scheutzow, and Björn Schmalfuß. “Interacting Stochastic Systems”. Ed. by Jean-Dominique Deuschel and Andreas Greven. Berlin, Heidelberg: Springer Berlin Heidelberg, 2005. Chap. Continuity Properties of Inertial Manifolds for Stochastic Retarded Semilinear Parabolic Equations, pp. 353–375. URL: http://dx.doi.org/10.1007/3-540-27110-4_16.

[Dij05] Jan J. Dijkstra. “On Homeomorphism Groups and the Compact-Open Topology”. *American Mathematical Monthly* 112.12 (2005), pp. 910–912.

[Din13] Seán Dineen. *Probability theory in finance*. Second. Vol. 70. Graduate Studies in Mathematics. A mathematical guide to the Black-Scholes formula. American Mathematical Society, Providence, RI, 2013, pp. xiv+305. URL: http://dx.doi.org/10.1090/gsm/070.

[DKN07] Bertrand Deroin, Victor Kleptsyn, and Andrés Navas. "Sur la dynamique unidimensionnelle en régularité intermédiaire". *Acta Mathematica* 199.2 (2007), pp. 199–262. URL: http://dx.doi.org/10.1007/s11511-007-0020-1.

[Doo94] J. L. Doob. *Measure theory*. Vol. 143. Graduate Texts in Mathematics. Springer-Verlag, New York, 1994, pp. xii+210. URL: http://dx.doi.org/10.1007/978-1-4612-0877-8.

[FGS14] Franco Flandoli, Benjamin Gess, and Michael Scheutzow. "Synchronization by noise". arXiv:1411.1340v1 [math.PR]. 2014.

[FGS15] Franco Flandoli, Benjamin Gess, and Michael Scheutzow. "Synchronization by noise for order-preserving random dynamical systems". arXiv:1503.08737v1 [math.PR]. 2015.

[Fre13] D. H. Fremlin. *Measure Theory, Volume 4: Topological Measure Spaces.* Chapter 47. URL: https://www.essex.ac.uk/maths/people/fremlin/chap47.pdf (version of 04/08/2013).

[FZ15] Chunrong Feng and Huaizhong Zhao. "Random Periodic Processes, Periodic Measures and Ergodicity". arXiv:1408.1897v3 [math.PR]. 2015.

[GGTQ15] Beniamin Goldys, Cecilia González-Tokman, and Anthony Quas. *Random Dynamical Systems and Multiplicative Ergodic Theorems Workshop (15w5059): Final Report*. 2015. URL: http://www.birs.ca/workshops/2015/15w5059/report15w5059.pdf.

[Hai11] Martin Hairer. "On Malliavins proof of Hrmanders theorem". *Bulletin des Sciences Mathmatiques* 135.67 (2011). Special issue in memory of Paul MalliavinConference in Memory of Paul Malliavin, pp. 650 –666. URL: http://www.sciencedirect.com/science/article/pii/S0007449711000765.

[Hom13] Ale Jan Homburg. "Synchronization in iterated function systems". arXiv:1303.6054v1 [math.DS]. 2013.

[IL02] PETER IMKELLER and CHRISTIAN LEDERER. "THE COHOMOLOGY OF STOCHASTIC AND RANDOM DIFFERENTIAL EQUATIONS, AND LOCAL LINEARIZATION OF STOCHASTIC FLOWS". *Stochastics and Dynamics* 02.02 (2002), pp. 131–159. URL: http://www.worldscientific.com/doi/abs/10.1142/S021949370200039X.

[Kai93] Thomas Kaijser. "On stochastic perturbations of iterations of circle maps". *Phys. D* 68.2 (1993), pp. 201–231. URL: http://dx.doi.org/10.1016/0167-2789(93)90081-B.

[Kak43] Shizuo Kakutani. "Notes on infinite product measure spaces. I". *Proc. Imp. Acad. Tokyo* 19 (1943), pp. 148–151.

[Kec95] Alexander S. Kechris. *Classical descriptive set theory*. Vol. 156. Graduate Texts in Mathematics. Springer-Verlag, New York, 1995. URL: http://dx.doi.org/10.1007/978-1-4612-4190-4.

[Kif86] Yuri Kifer. *Ergodic theory of random transformations*. Vol. 10. Progress in Probability and Statistics. Birkhäuser Boston, Inc., Boston, MA, 1986. URL: http://dx.doi.org/10.1007/978-1-4684-9175-3.

[KL14] David Kerr and Hanfeng Li. *Ergodic Theory: Independence and Dichotomies*. 2014. URL: http://www.math.tamu.edu/~kerr/book/ (version of 09/13/2014).

[KN04] V. A. Kleptsyn and M. B. Nalskii. "Contraction of orbits in random dynamical systems on the circle". *Funct. Anal. Appl.* 38.4 (2004), pp. 267–282.

[KS12] Sergei Kuksin and Armen Shirikyan. *Mathematics of Two-Dimensional Turbulence*. Cambridge Tracts in Mathematics. Cambridge: Cambridge University Press, 2012.

[Kun90] Hiroshi Kunita. *Stochastic flows and stochastic differential equations*. Vol. 24. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, 1990.

[LeJ87] Y. Le Jan. "Équilibre statistique pour les produits de difféomorphismes aléatoires indépendants". *Ann. Inst. H. Poincaré Probab. Statist.* 23.1 (1987), pp. 111–120.

[Lin02] Georg Lindgren. *Lectures on Stationary Stochastic Processes*. Lund University, 2002. URL: http://www.maths.lth.se/matstat/staff/georg/Publications/lecture2002.pdf.

[MO10] William (http://mathoverflow.net/users/39719/william). *Lipschitz equivalence of Riemannian metrics*. MathOverflow. 2010. URL: http://mathoverflow.net/q/37468.

[MO14] Julian Newman (http://mathoverflow.net/users/15570/julian newman). *Is it true that all stationary measurable stochastic processes are "measurably stationary"?* MathOverflow. 2014. URL: http://mathoverflow.net/q/181640.

[MO15a] Julian Newman (http://mathoverflow.net/users/15570/julian newman). *Can the integral of a "generic" bounded measurable function be determined by its values on the rationals?* MathOverflow. 2015. URL: http://mathoverflow.net/q/218361.

[MO15b] Julian Newman (http://mathoverflow.net/users/15570/julian newman). *Is every convex subset of a Borel-linearly ordered space measurable?* MathOverflow. 2015. URL: http://mathoverflow.net/q/218778.

[MO15c] Julian Newman (http://mathoverflow.net/users/15570/julian newman). *Can a (non-measurable) autonomous flow have a non-trivial periodic orbit without a minimal period?* MathOverflow. 2015. URL: http://mathoverflow.net/q/220718.

[MS99] Salah-Eldin A. Mohammed and Michael K. R. Scheutzow. "The stable manifold theorem for stochastic differential equations". *Ann. Probab.* 27.2 (1999), pp. 615–652. URL: `http://dx.doi.org/10.1214/aop/1022677380`.

[Nak+05] Hiroya Nakao, Ken-suke Arai, Ken Nagai, Yasuhiro Tsubo, and Yoshiki Kuramoto. "Synchrony of limit-cycle oscillators induced by random external impulses". *Phys. Rev. E* 72.2 (2005). 026220.

[Nev65] Jacques Neveu. *Mathematical foundations of the calculus of probability.* Translated by Amiel Feinstein. Holden-Day, Inc., San Francisco, Calif.-London-Amsterdam, 1965.

[New15a] Julian Newman. *Ergodic Theory for Semigroups of Markov Kernels.* 2015. URL: `http://wwwf.imperial.ac.uk/~jmn07/Ergodic_Theory_for_Semigroups_of_Markov_Kernels.pdf` (version of 04/28/2015).

[New15b] Julian Newman. "Necessary and Sufficient Conditions for Stable Synchronisation in Random Dynamical Systems". arXiv:1408.5599v3 [math.DS]. 2015.

[New15c] Julian Newman. "Synchronisation in Invertible Random Dynamical Systems on the Circle". arXiv:1502.07618v2 [math.DS]. 2015.

[New16] Julian Newman. "Synchronisation of almost all trajectories of a random dynamical system". arXiv:1511.08831v2 [math.DS]. 2016.

[Par05] K. R. Parthasarathy. *Probability measures on metric spaces.* Reprint of the 1967 original. AMS Chelsea Publishing, Providence, RI, 2005, pp. xii+276.

[Pik84] Arkady S. Pikovskii. "Synchronization and stochastization of array of self-excited oscillators by external noise". *Radiophys. Quantum Electron.* 27.5 (1984), pp. 390–395.

[Rue79] David Ruelle. "Ergodic theory of differentiable dynamical systems". *Inst. Hautes Études Sci. Publ. Math.* 50 (1979), pp. 27–58. URL: `http://www.numdam.org/item?id=PMIHES_1979__50__27_0`.

[SH02] Klaus Reiner Schenk-Hoppé. "Sample-Path Stability of Non-Stationary Dynamic Economic Systems". *Annals of Operations Research* 114.1 (2002), pp. 263–280. URL: `http://dx.doi.org/10.1023/A:1021074506105`.

[Sri98] S. M. Srivastava. *A course on Borel sets.* Vol. 180. Graduate Texts in Mathematics. Springer-Verlag, New York, 1998. URL: `http://dx.doi.org/10.1007/978-3-642-85473-6`.

[Sus78] Hector J. Sussmann. "On the Gap Between Deterministic and Stochastic Ordinary Differential Equations". *Ann. Probab.* 6.1 (Feb. 1978), pp. 19–41. URL: `http://dx.doi.org/10.1214/aop/1176995608`.

[Tor+01] R. Toral, C. R. Mirasso, E. Hernández-García, and O. Piro. "Analytical and numerical studies of noise-induced synchronization of chaotic systems". *Chaos* 11.3 (2001), pp. 665–673.

[Ver14] Olivier Verdier. *Variational Equation.* 2014. URL: `https://www.olivierverdier.com/posts/2014/01/28/variational-equation/` (version of 01/28/2014).

[Wil91] David Williams. *Probability with martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge, 1991. URL: http://dx.doi.org/10.1017/CBO9780511813658.

Appendix A. Some Preliminaries from Measure and Probability Theory

Throughout this thesis, familiarity with some of the most foundational concepts and results from probability theory and stochastic analysis is assumed. Nonetheless, there are some topics in measure theory and probability theory that are worth covering explicitly, which we do here. Specifically, we will cover: infinite product σ-algebras; the π-λ theorem and some of its important corollaries (such as the monotone class theorem); results concerning expectations and conditional expectations involving independent σ-algebras; measurability of operations involving integrals; a result concerning the measure of a ball about a variable centre point in a metric space; the formula for changing variables in conditional expectations; and the measurable and topological structure of a space of probability measures (including a fairly detailed exposition of the *narrow topology*).

Infinite product σ-algebras

Given a family $((X_\alpha, \Sigma_\alpha))_{\alpha \in I}$ of measurable spaces $(X_\alpha, \Sigma_\alpha)$, if the Cartesian product $\bigtimes_{\alpha \in I} X_\alpha$ is non-empty then we define the product σ-algebra $\bigotimes_{\alpha \in I} \Sigma_\alpha$ to be the smallest σ-algebra on $\bigtimes_{\alpha \in I} X_\alpha$ with respect to which the map

$$\begin{aligned} \bigtimes_{\alpha \in I} X_\alpha &\to X_{\tilde{\alpha}} \\ (x_\alpha)_{\alpha \in I} &\mapsto x_{\tilde{\alpha}} \end{aligned}$$

is measurable for every $\tilde{\alpha} \in I$. If we also have a family $(\mu_\alpha)_{\alpha \in I}$ of probability measures μ_α on $(X_\alpha, \Sigma_\alpha)$, then there exists a unique probability measure $\bigotimes_{\alpha \in I} \mu_\alpha$ on the product space $\left(\bigtimes_{\alpha \in I} X_\alpha , \bigotimes_{\alpha \in I} \Sigma_\alpha \right)$ such that for any $\alpha_1, \ldots, \alpha_n \in I$, for any $A_1 \in \Sigma_{\alpha_1}, \ldots, A_n \in \Sigma_{\alpha_n}$,

$$\bigotimes_{\alpha \in I} \mu_\alpha \left(\{ (x_\alpha)_{\alpha \in I} : x_{\alpha_i} \in A_i \ \forall\, 1 \leq i \leq n \} \right) = \prod_{i=1}^{n} \mu_{\alpha_i}(A_i).$$

(See e.g. [Kak43].)

It is easy to show that for a family $(X_\alpha)_{\alpha \in I}$ of second-countable topological spaces X_α indexed by a *countable* set I,

$$\mathcal{B}\left(\bigtimes_{\alpha \in I} X_\alpha \right) = \bigotimes_{\alpha \in I} \mathcal{B}(X_\alpha)$$

where $\bigtimes_{\alpha \in I} X_\alpha$ is equipped with the product topology.

The π-λ theorem

A π-*system* is a collection of sets that is closed under pairwise intersections. A λ-*system* (or *Dynkin system*) on a set Ω is a collection of subsets of Ω that includes Ω itself and is closed under both countable disjoint unions and complements relative to Ω. For example, given a σ-algebra $\mathcal{F}$ on Ω and two probability measures μ_1 and μ_2 on $(\Omega, \mathcal{F})$, it is clear that $\{E \in \mathcal{F} : \mu_1(E) = \mu_2(E)\}$ is a λ-system on Ω.

Remark A.1. If $(\Omega, \mathcal{F})$ is a measurable space with $\mathcal{F}$ being countably generated, then there exists a countable π-system generating $\mathcal{F}$: for any countable $\mathcal{A} \subset \mathcal{F}$ with $\sigma(\mathcal{A}) = \mathcal{F}$, the smallest π-system containing $\mathcal{A}$, namely the set

$$\mathcal{C} := \{A_1 \cap \ldots \cap A_n : n \in \mathbb{N},\, A_1, \ldots, A_n \in \mathcal{A}\},$$

is a countable π-system generating $\mathcal{F}$.[1] Note that the Borel σ-algebra of a second-countable topological space is countably generated (since any countable base for a topology is also a generator for the Borel σ-algebra thereof), and therefore is generated by a countable π-system.

Remark A.2. For any family $((X_\alpha, \Sigma_\alpha))_{\alpha \in I}$ of measurable spaces $(X_\alpha, \Sigma_\alpha)$, if the Cartesian product $\bigtimes_{\alpha \in I} X_\alpha$ is non-empty then the product σ-algebra $\bigotimes_{\alpha \in I} \Sigma_\alpha$ is generated by the π-system

$$\mathcal{C} := \left\{ \left(\bigtimes_{\alpha \in J} A_\alpha \right) \times \left(\bigtimes_{\alpha \in I \setminus J} X_\alpha \right) : J \subset I \text{ finite, } A_\alpha \in \Sigma_\alpha \, \forall \alpha \in J \right\}.$$

Remark A.3. Let $(\Omega, \mathcal{F})$ be a measurable space, and let $\{\mathcal{F}_\alpha : \alpha \in I\}$ be a collection of sub-σ-algebras of $\mathcal{F}$ that is totally ordered by inclusion. Then $\bigcup_{\alpha \in I} \mathcal{F}_\alpha$ is a π-system.

Remark A.4. We mention another important example of a λ-system: Let $(\Omega, \mathcal{F})$ and (X, Σ) be measurable spaces, and let $(\mu_\omega)_{\omega \in \Omega}$ be a family of probability measures on X. Then it is easy to show that the set $\mathcal{D} := \{A \in \Sigma : \omega \mapsto \mu_\omega(A) \text{ is measurable}\}$ is a λ-system on X.

Lemma A.5 (π-λ theorem)**.** *Let $\mathcal{D}$ be a λ-system on a set Ω, and let $\mathcal{C} \subset \mathcal{D}$ be a π-system. Then the σ-algebra on Ω generated by $\mathcal{C}$ is contained in $\mathcal{D}$.*

For a proof, see [Wil91, Lemma A1.3]. We now give three important immediate corollaries. (For "generalisations" of the first two of these corollaries, see Exercise 6 and Lemma 7 of [New15a].)

Corollary A.6. *Let $(\Omega, \mathcal{F})$ be a measurable space, and let μ_1 and μ_2 be probability measures on $(\Omega, \mathcal{F})$. If there exists a π-system $\mathcal{C} \subset \mathcal{F}$ generating $\mathcal{F}$ such that $\mu_1(E) = \mu_2(E)$ for all $E \in \mathcal{C}$, then $\mu_1 = \mu_2$.*

Proof. Since $\{E \in \mathcal{F} : \mu_1(E) = \mu_2(E)\}$ is a λ-system on Ω containing the π-system $\mathcal{C}$, we have by Lemma A.5 that μ_1 and μ_2 agree on the whole of $\mathcal{F}$. □

Corollary A.7 (Monotone class theorem, [Wil91, Theorem 3.14])**.** *Let $(\Omega, \mathcal{F})$ be a measurable space, and let $\mathcal{H}$ be a set of functions from Ω to $\mathbb{R}$ such that:*

(a) the constant function $\omega \mapsto 1$ is in $\mathcal{H}$;

[1] The author is grateful to Nathaniel Eldredge for first drawing his attention to this very useful fact.

(b) *for any* $c_1, c_2 \in \mathbb{R}$ *and* $g_1, g_2 \in \mathcal{H}$, $c_1 g_1 + c_2 g_2 \in \mathcal{H}$;

(c) *for any uniformly bounded, increasing sequence of functions* $g_n \in \mathcal{H}$, *the pointwise limit* $g_\infty := \lim_n g_n$ *is in* $\mathcal{H}$;

(d) *there exists a* π*-system* $\mathcal{C}$ *generating* $\mathcal{F}$ *such that for all* $E \in \mathcal{C}$, $\mathbb{1}_E \in \mathcal{H}$.

Then $\mathcal{H}$ *includes all bounded measurable functions* $g : \Omega \to \mathbb{R}$.

Proof. Let $\mathcal{D} := \{E \in \mathcal{F} : \mathbb{1}_E \in \mathcal{H}\}$. By properties (a), (b) and (c), $\mathcal{D}$ is a λ-system; and property (d) states that $\mathcal{C} \subset \mathcal{D}$. Hence by Lemma A.5, $\mathbb{1}_E \in \mathcal{H}$ for every $E \in \mathcal{F}$. Property (b) then gives that $\mathcal{H}$ includes all bounded simple functions. Property (c) then gives that $\mathcal{H}$ includes all bounded measurable functions. □

Corollary A.8. *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space, let* $\mathcal{C} \subset \mathcal{F}$ *be a* π*-system, and let* $G \in \mathcal{F}$ *be an event that is independent (under* $\mathbb{P}$*) of every member of* $\mathcal{C}$. *Then* G *is independent of* $\sigma(C)$.

Proof. Let $\mathcal{D} := \{E \in \mathcal{F} : \mathbb{P}(E \cap G) = \mathbb{P}(E)\mathbb{P}(G)\}$. We know that $\mathcal{C} \subset \mathcal{D}$, and it is easy to see that $\mathcal{D}$ is a λ-system. Therefore, by Lemma A.5, $\sigma(\mathcal{C}) \subset \mathcal{D}$, i.e. G is independent of $\sigma(\mathcal{C})$. □

As an important special case of Corollary A.8, we have the following:

Corollary A.9. *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space, let* $\{\mathcal{F}_\alpha : \alpha \in I\}$ *be a collection of sub-*σ*-algebras of* $\mathcal{F}$ *that is totally ordered by inclusion, and let* $G \in \mathcal{F}$ *be an event that is independent (under* $\mathbb{P}$*) of* $\mathcal{F}_\alpha$ *for every* $\alpha \in I$. *Then* G *is independent of* $\sigma(\mathcal{F}_\alpha : \alpha \in I)$.

Proof. As in Remark A.3, $\mathcal{C} := \bigcup_{\alpha \in I} \mathcal{F}_\alpha$ is a π-system. Hence Corollary A.8 gives the desired result. □

Results about independent σ-algebras

Lemma A.10. *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space, let* $\mathcal{G}_1$ *and* $\mathcal{G}_2$ *be independent sub-*σ*-algebras of* $\mathcal{F}$ *(under* $\mathbb{P}$*), and let* $h : \Omega \times \Omega \to \mathbb{R}$ *be a bounded* $(\mathcal{G}_1 \otimes \mathcal{G}_2)$*-measurable function. Then*

$$\int_\Omega h(\omega, \omega)\, \mathbb{P}(d\omega) = \int_{\Omega \times \Omega} h(\omega, \tilde{\omega})\, \mathbb{P} \otimes \mathbb{P}(d(\omega, \tilde{\omega})).$$

Proof. In the case that $h = \mathbb{1}_{G_1 \times G_2}$ for some $G_1 \in \mathcal{G}_1$ and $G_2 \in \mathcal{G}_2$, we have

$$\begin{aligned}
\int_\Omega \mathbb{1}_{G_1 \times G_2}(\omega, \omega)\, \mathbb{P}(d\omega) &= \mathbb{P}(G_1 \cap G_2) \\
&= \mathbb{P}(G_1)\mathbb{P}(G_2) \\
&= \mathbb{P} \otimes \mathbb{P}(G_1 \times G_2) \\
&= \int_{\Omega \times \Omega} \mathbb{1}_{G_1 \times G_2}(\omega, \tilde{\omega})\, \mathbb{P} \otimes \mathbb{P}(d(\omega, \tilde{\omega}))
\end{aligned}$$

as required. Now $\{G_1 \times G_2 : G_1 \in \mathcal{G}_1, G_2 \in \mathcal{G}_2\}$ is a π-system generating $\mathcal{G}_1 \otimes \mathcal{G}_2$, so Corollary A.7 gives the desired result for general h. □

Corollary A.11. *Assume the hypotheses of Lemma A.10. Then*

$$\mathbb{E}[\,\omega \mapsto h(\omega,\omega)\,|\,\mathcal{G}_1\,] \overset{\mathbb{P}\text{-a.s.}}{=} \mathbb{E}[\,\omega \mapsto h(\,\cdot\,,\omega)\,].$$

The intuition behind Corollary A.11 is quite clear: if a random quantity H is determined by two statistically independent pieces of information, then the *conditional* expectation of H *given* the knowledge of the first piece of information is simply equal to the mean of H averaged over the set of possibilities for the second piece of information, with the first piece of information being taken to be as it was observed.

Proof of Corollary A.11. Fix any $G \in \mathcal{G}_1$. Defining $\hat{h}\colon \Omega \times \Omega \to \mathbb{R}$ by $\hat{h}(\tilde{\omega},\omega) = \mathbb{1}_G(\tilde{\omega})h(\tilde{\omega},\omega)$, we have

$$\begin{aligned}
\int_G h(\omega,\omega)\,\mathbb{P}(d\omega) &= \int_\Omega \hat{h}(\omega,\omega)\,\mathbb{P}(d\omega) \\
&= \int_\Omega \int_\Omega \hat{h}(\tilde{\omega},\omega)\,\mathbb{P}(d\omega)\,\mathbb{P}(d\tilde{\omega}) \quad \text{(by Lemma A.10)} \\
&= \int_G \mathbb{E}[\,\omega \mapsto h(\tilde{\omega},\omega)\,]\,\mathbb{P}(d\tilde{\omega})
\end{aligned}$$

as required. □

Partial integrals are measurable

The following fundamental result is a particular case of [New15a, Lemma 8].

Lemma A.12. *Let $(I,\mathcal{I})$, $(\Omega,\mathcal{F})$ and (X,Σ) be measurable spaces, and suppose we have a family $(\rho_\omega)_{\omega\in\Omega}$ of probability measures ρ_ω on X such that the mapping $\omega \mapsto \rho_\omega(A)$ is $\mathcal{F}$-measurable for all $A \in \Sigma$. For any bounded measurable function $g\colon \Omega \times I \times X \to \mathbb{R}$, the function $\overline{g}\colon \Omega \times I \to \mathbb{R}$ given by*

$$\overline{g}(\omega,\alpha) = \int_X g(\omega,\alpha,x)\,\rho_\omega(dx)$$

is measurable.

In most applications, either g will not depend on α or g will not depend on ω, and in many cases ρ_ω will also not depend on ω (i.e. the integrator will just be a deterministic measure, rather than a random measure).

Proof of Lemma A.12. In the case that $g = \mathbb{1}_{E\times B\times A}$ for some $E \in \mathcal{F}$, $B \in \mathcal{I}$ and $A \in \Sigma$, the function $\bar{g}$ is given by $\bar{g}(\omega,\alpha) = \mathbb{1}_{E\times B}(\omega,\alpha)\rho_\omega(A)$, which is clearly a measurable function. Since $\{E \times B \times A : E \in \mathcal{F}, B \in \mathcal{I}, A \in \Sigma\}$ is a π-system generating $\mathcal{F} \otimes \mathcal{I} \otimes \Sigma$, we apply Corollary A.7 to give the desired result for general g. (Condition (c) of Corollary A.7 is satisfied due to the dominated convergence theorem and the fact that a pointwise limit of real-valued measurable functions is measurable.) □

Measures of balls

Lemma A.13. *Let (X,d) be a metric space, and let ρ be a Borel probability measure on X. Let $\delta\colon X \to (0,\infty)$ be a lower semicontinuous function. Then the map*

$$x \mapsto \rho(B_{\delta(x)}(x))$$

is lower semicontinuous.

(In particular, taking δ to be a constant, we have that the map $x \mapsto \rho(B_\delta(x))$ is lower semicontinuous.)

Proof. Fix an arbitrary convergent sequence (x_n) in X and value $l \in [0,1]$ with the property that $\rho(B_{\delta(x_n)}(x_n)) \leq l$ for all n; writing $x := \lim_n x_n$, we will show that $\rho(B_{\delta(x)}(x)) \leq l$. Let $\tilde{\delta} := \liminf_{n\to\infty} \delta(x_n)$; so $\delta(x) \leq \tilde{\delta}$. For each integer $j > \frac{1}{2\tilde{\delta}}$, let $n_j \in \mathbb{N}$ be such that $\delta(x_{n_j}) > \tilde{\delta} - \frac{1}{2j}$ and $d(x, x_{n_j}) < \frac{1}{2j}$; for any $y \in B_{\tilde{\delta}-\frac{1}{j}}(x)$, we have that

$$d(y, x_{n_j}) \;\leq\; d(y,x) + d(x, x_{n_j}) \;<\; \tilde{\delta} - \frac{1}{2j} \;<\; \delta(x_{n_j}),$$

and so $y \in B_{\delta(x_{n_j})}(x_{n_j})$. So then, for each $j > \frac{1}{2\tilde{\delta}}$, $B_{\tilde{\delta}-\frac{1}{j}}(x)$ is contained in $B_{\delta(x_{n_j})}(x_{n_j})$, and therefore $\rho(B_{\tilde{\delta}-\frac{1}{j}}(x)) \leq l$. Consequently, we have that $\rho(B_{\tilde{\delta}}(x)) \leq l$, and therefore $\rho(B_{\delta(x)}(x)) \leq l$. □

Transformation of conditional expectations

The following result is the conditional-expectation version of the "change-of-variables formula" $[\int g \circ Y \, d\mathbb{P} = \int g \, d(Y_*\mathbb{P})]$.

Lemma A.14. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $(S, \mathcal{S})$ be a measurable space, let $g\colon S \to \mathbb{R}$ be a measurable function, and let $\mathcal{E}$ be a sub-σ-algebra of $\mathcal{S}$. Given a random variable $Y\colon \Omega \to S$ satisfying $\mathbb{E}_{(\mathbb{P})}[|g(Y)|] < \infty$ (and therefore $\mathbb{E}_{(Y_*\mathbb{P})}[|g|] < \infty$), we have that*

$$\mathbb{E}_{(\mathbb{P})}[g(Y)|Y^{-1}\mathcal{E}] \overset{\mathbb{P}\text{-a.s.}}{=} \mathbb{E}_{(Y_*\mathbb{P})}[g|\mathcal{E}](Y).$$

(That is to say, for any version $h\colon S \to \mathbb{R}$ of the conditional expectation $\mathbb{E}_{(Y_*\mathbb{P})}[g|\mathcal{E}]$, the function $h \circ Y$ is a version of the conditional expectation $\mathbb{E}_{(\mathbb{P})}[g(Y)|Y^{-1}\mathcal{E}]$.)

Proof. Let $h\colon S \to \mathbb{R}$ be a version of $\mathbb{E}_{(Y_*\mathbb{P})}[g|\mathcal{E}]$. For any $A \in Y^{-1}\mathcal{E}$, writing $A = Y^{-1}(E)$ for some $E \in \mathcal{E}$, we have

$$\int_A h(Y(\omega))\,\mathbb{P}(d\omega) \;=\; \int_E h(x)\,Y_*\mathbb{P}(dx) \;=\; \int_E g(x)\,Y_*\mathbb{P}(dx) \;=\; \int_A g(Y(\omega))\,\mathbb{P}(d\omega)$$

as required. □

Spaces of probability measures

For convenience, given a probability space (X, Σ, ρ) and a ρ-integrable function $g\colon X \to \mathbb{R}$, we will sometimes write $\rho(g)$ as a shorthand for $\int_X g(x)\,\rho(dx)$.

Given a measurable space (X, Σ), we write $\mathcal{M}_{(X,\Sigma)}$ for the set of probability measures on (X, Σ), which we equip with its "evaluation σ-algebra"

$$\mathfrak{K}_{(X,\Sigma)} \;:=\; \sigma(\rho \mapsto \rho(A) : A \in \Sigma).$$

So for any measurable space $(\Omega, \mathcal{F})$, a function $p : \Omega \to \mathcal{M}_{(X,\Sigma)}$ is measurable if and only if the map $\omega \mapsto p(\omega)(A)$ is measurable for all $A \in \Sigma$. In this case, we also have that for every bounded measurable function $g\colon X \to \mathbb{R}$, the map $\omega \mapsto p(\omega)(g)$ is measurable (since we can approximate g by simple functions). Moreover, given a probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$, we can define the "mean probability measure" $\bar{p}$ on X by $\bar{p}(A) := \int_\Omega p(\omega)(A)\,\mathbb{P}(d\omega)$. (This is indeed a probability measure, by the monotone convergence theorem.) For any bounded measurable $g\colon X \to \mathbb{R}$, we have that $\bar{p}(g) = \int_\Omega p(\omega)(g)\,\mathbb{P}(d\omega)$. (To see this, just approximate g by a uniformly bounded sequence of simple functions, and apply the dominated convergence theorem.)

Lemma A.15. *Given measurable spaces* (X_1, Σ_1) *and* (X_2, Σ_2), *the map*

$$\begin{array}{rcl} \mathcal{M}_{(X_1,\Sigma_1)} \times \mathcal{M}_{(X_2,\Sigma_2)} & \to & \mathcal{M}_{(X_1\times X_2, \Sigma_1\otimes\Sigma_2)} \\ (\rho_1, \rho_2) & \mapsto & \rho_1 \otimes \rho_2 \end{array}$$

is measurable (with respect to the respective evaluation σ*-algebras).*

Proof. Let $\mathcal{D} := \{B \in \Sigma_1 \otimes \Sigma_2 \,:\, \text{the map } (\rho_1, \rho_2) \mapsto \rho_1 \otimes \rho_2(B) \text{ is measurable}\}$. As in Remark A.4, $\mathcal{D}$ is a λ-system on $X_1 \times X_2$. It is also clear that for any $A_1 \in \Sigma_1$ and $A_2 \in \Sigma_2$, the map $(\rho_1, \rho_2) \mapsto \rho_1 \otimes \rho_2(A_1 \times A_2) = \rho_1(A_1)\rho_2(A_2)$ is measurable; so $\mathcal{D}$ contains the π-system $\{A_1 \times A_2 : A_1 \in \Sigma_1, A_2 \in \Sigma_2\}$. Hence, by Lemma A.5, the map $(\rho_1, \rho_2) \mapsto \rho_1 \otimes \rho_2(B)$ is measurable for all $B \in \Sigma_1 \otimes \Sigma_2$, as required. □

We now go on to consider Borel probability measures on separable metric spaces.

Recall that, given two topological spaces X and Y, a function $f\colon X \to Y$ is called a *topological embedding* (*of* X *into* Y) if f is continuous and injective, and the map $f^{-1}\colon f(X) \to X$ is continuous (where $f(X)$ is equipped with the induced topology from Y). In the case that X and Y are metrisable, a function $f\colon X \to Y$ is a topological embedding if and only if the following holds: for any sequence (x_n) in X and any point $x \in X$,

$$x_n \to x \;\Leftrightarrow\; f(x_n) \to f(x).$$

Given two topological spaces X and Y, a function $f\colon X \to Y$ is called a *closed embedding* (*of* X *into* Y) if f is a topological embedding and $f(X)$ is a closed subset of Y. This implies that for every closed $G \subset X$, $f(G)$ is closed in Y. In the case that X and Y are metrisable, it is easy to check that a function $f\colon X \to Y$ is a closed embedding if and only if the following holds: f is continuous, and for every divergent sequence (x_n) in X, the sequence $(f(x_n))$ is divergent in Y.

Recall that a topological space (or a topology) is said to be *Polish* if it is both separable and completely metrisable.

Theorem A.16. *Let X be a separable metrisable topological space (with $\mathcal{T}$ denoting the topology on X). Then there exists a separable metrisable topology $\mathcal{N}_{\mathcal{T}}$ on $\mathcal{M}_{(X,\mathcal{B}(X))}$ characterised as follows: fixing any metrisation d of $\mathcal{T}$, a sequence (ρ_n) converges in $\mathcal{N}_{\mathcal{T}}$ to ρ if and only if the equivalent statements*

(i) $\rho(U) \leq \liminf_{n\to\infty} \rho_n(U)$ *for every open $U \subset X$;*

(ii) $\rho(G) \geq \limsup_{n\to\infty} \rho_n(G)$ *for every closed $G \subset X$;*

(iii) $\rho_n(g) \to \rho(g)$ *for every bounded d-Lipschitz $g : X \to \mathbb{R}$;*

(iv) $\rho_n(g) \to \rho(g)$ *for every bounded continuous $g : X \to \mathbb{R}$;*

(v) $\rho_n(A) \to \rho(A)$ *for every Borel-measurable continuity set A of ρ;*[2]

hold. There exists a countable set $\{g_i\}_{i\in\mathbb{N}}$ of continuous functions $g_i : X \to [0,1]$ such that the map $\rho \mapsto (\rho(g_i))_{i\in\mathbb{N}}$ serves as a topological embedding of $\mathcal{M}_{(X,\mathcal{B}(X))}$ (equipped with the topology $\mathcal{N}_{\mathcal{T}}$) into $[0,1]^{\mathbb{N}}$ (equipped with the product topology). The topology $\mathcal{N}_{\mathcal{T}}$ is compact if and only if $\mathcal{T}$ is compact, and $\mathcal{N}_{\mathcal{T}}$ is Polish if and only if $\mathcal{T}$ is Polish. The Borel σ-algebra of $\mathcal{N}_{\mathcal{T}}$ is precisely $\mathfrak{K}_{(X,\mathcal{B}(X))}$.

A proof is given in Section 0.6 of [New15a], except for characterisation (v) of the topology $\mathcal{N}_{\mathcal{T}}$, which can be found in [Par05, Theorem II.6.1].

Definition A.17. The topology $\mathcal{N}_{\mathcal{T}}$ is called the *narrow topology* or the *topology of weak convergence*. When the topology $\mathcal{T}$ on X is implicitly assumed from the context, we will say that "*μ_n converges weakly to μ*" to mean that μ_n converges to μ in $\mathcal{N}_{\mathcal{T}}$.

Note that for any metric space Y and any function $p : Y \to \mathcal{M}_{(X,\mathcal{B}(X))}$ that is continuous with respect to the narrow topology,

(i) the map $y \mapsto p(y)(U)$ is lower semicontinuous for every open $U \subset X$;

(ii) the map $y \mapsto p(y)(G)$ is upper semicontinuous for every closed $G \subset X$;

(iii) the map $y \mapsto \int_X g(x)\, p(y)(dx)$ is continuous for every bounded continuous function $g : X \to \mathbb{R}$.

The following corollary of Theorem A.16 is clear:

Corollary A.18. *The separable metrisable space $(X, \mathcal{T})$ is compact if and only if the topological embedding $\rho \mapsto (\rho(g_i))_{i\in\mathbb{N}}$ of $\mathcal{M}_{(X,\mathcal{B}(X))}$ into $[0,1]^{\mathbb{N}}$ described in Theorem A.16 is a closed embedding.*

[2] A *continuity set* of ρ is a set $A \subset X$ satisfying $\rho(\partial A) = 0$.

Remark A.19. A set of functions $\{g_i\}_{i\in\mathbb{N}}$ with the property described in Theorem A.16 is said to be *convergence-determining*, since this property is precisely the property that for any sequence (ρ_n) of probability measures on X and any probability measure ρ on X,

$$\rho_n \overset{\mathcal{N}_\mathcal{T}}{\to} \rho \text{ as } n\to\infty \quad\Longleftrightarrow\quad \forall\, i\in\mathbb{N},\ \rho_n(g_i)\to\rho(g_i) \text{ as } n\to\infty.$$

Lemma A.20 ("Almost sure convergence implies convergence in distribution")**.** *Let X be a separable metrisable topological space. Let $(\Omega,\mathcal{F},\mathbb{P})$ be a probability space, let $q\colon\Omega\to X$ be a measurable function, and let (q_n) be a sequence of measurable functions $q_n\colon\Omega\to X$ such that $q_n(\omega)\to q(\omega)$ as $n\to\infty$ for $\mathbb{P}$-almost all $\omega\in\Omega$. Then $q_{n*}\mathbb{P}$ converges in the narrow topology to $q_*\mathbb{P}$ as $n\to\infty$.*

Proof. Fix any bounded continuous $g\colon X\to\mathbb{R}$. For $\mathbb{P}$-almost all $\omega\in\Omega$, $g(q_n(\omega))\to g(q(\omega))$ as $n\to\infty$, and so by the dominated convergence theorem, we have

$$q_{n*}\mathbb{P}(g) = \int_\Omega g(q_n(\omega))\,\mathbb{P}(d\omega) \overset{n\to\infty}{\to} \int_\Omega g(q(\omega))\,\mathbb{P}(d\omega) = q_*\mathbb{P}(g)$$

as required. □

We now give the "topological version" of Lemma A.15.

Lemma A.21. *Given separable metrisable topological spaces X_1 and X_2, the map*

$$\begin{aligned} \mathcal{M}_{(X_1,\mathcal{B}(X_1))}\times\mathcal{M}_{(X_2,\mathcal{B}(X_2))} &\to \mathcal{M}_{(X_1\times X_2,\mathcal{B}(X_1\times X_2))} \\ (\rho_1,\rho_2) &\mapsto \rho_1\otimes\rho_2 \end{aligned}$$

is continuous (with respect to the respective narrow topologies).

For a proof, see [Bil99, Theorem 2.8(ii)].

Lemma A.22. *Let $(X,\mathcal{T})$ be a separable metrisable topological space. The map $x\mapsto\delta_x$ serves as a closed embedding of $(X,\mathcal{T})$ into $(\mathcal{M}_{(X,\mathcal{B}(X))},\mathcal{N}_\mathcal{T})$.*

Proof. Given a sequence (x_n) in X converging to x, we clearly have that $\delta_{x_n}(g)\to\delta_x(g)$ for any continuous $g\colon X\to\mathbb{R}$; so in particular, $x\mapsto\delta_x$ is an $\mathcal{N}_\mathcal{T}$-continuous mapping. So to complete the proof, it remains just to show that if (x_n) is a divergent sequence in X then (δ_{x_n}) is a divergent sequence in $\mathcal{M}_{(X,\mathcal{B}(X))}$. Suppose for a contradiction that (x_n) is a divergent sequence but δ_{x_n} converges in $\mathcal{N}_\mathcal{T}$ to a probability measure ρ. Since every point in X is *not* the limit of the sequence (x_n), we can cover X by open sets U for which $\{n\in\mathbb{N} : x_n\notin U\}$ is infinite; note that for every such set U, $\liminf_{n\to\infty}\delta_{x_n}(U)=0$, and therefore $\rho(U)=0$. Now X is second-countable, and therefore we can find a countable subcover for the open cover of X that we have constructed. Thus we can cover X by countably many ρ-null sets, giving a contradiction. □

Appendix B. Uniform convergence on compact sets

The purpose of this Appendix is to present some of the basic facts concerning "uniform convergence on compact sets" that are assumed in this thesis.

Let Y be a metrisable topological space.

Lemma B.1. *Let K be a compact metrisable topological space and let $(f_n)_{n\in\mathbb{N}\cup\{\infty\}}$ be a family of continuous functions $f_n\colon K\to Y$. For any two metrisations d_1 and d_2 of the topology on Y, f_n converges uniformly to f_∞ under d_1 as $n\to\infty$ if and only if f_n converges uniformly to f_∞ under d_2 as $n\to\infty$.*

Proof. Let $C := \{\frac{1}{2^n}\}_{n\in\mathbb{N}} \cup \{0\}$, and define $F : C\times K \to Y$ by $F(0,x) = f_\infty(x)$ and $F(\frac{1}{2^n},x) = f_n(x)$ for all $x\in K$ and $n\in\mathbb{N}$. Let d_K be a metrisation of the topology on K, and let d be the metrisation of the product topology on $C\times K$ given by

$$d((t_1,x_1),(t_2,x_2)) \;=\; \max(|t_2-t_1|, d_K(x_1,x_2)).$$

Obviously $C\times K$ is compact, and therefore under any given metrisation of the topology on Y, F is continuous if and only if F is uniformly continuous. Hence, to show the desired result, it is sufficient to show that under any given metrisation d_Y of the topology on Y, F is uniformly continuous if and only if f_n converges uniformly to f_∞ as $n\to\infty$. Fix such a metrisation d_Y. It is obvious that if F is uniformly continuous then f_n converges uniformly to f_∞ as $n\to\infty$. Conversely, suppose that f_n converges uniformly to f_∞ as $n\to\infty$, and fix any $\varepsilon>0$. Let $N\in\mathbb{N}$ be such that for all $n\geq N$ and $x\in K$, $d_Y(f_n(x),f_\infty(x))<\frac{\varepsilon}{3}$. Now since K is compact, f_n is uniformly continuous for all $n\in\mathbb{N}\cup\{\infty\}$. So let $\delta_1>0$ be such that for all $x,y\in K$ with $d_K(x,y)<\delta_1$, $d_Y(f_\infty(x),f_\infty(y))<\frac{\varepsilon}{3}$; and let $\delta_2>0$ be such that for all $x,y\in K$ with $d_K(x,y)<\delta_2$ and all $n\in\{1,\ldots,N-1\}$, $d_Y(f_n(x),f_n(y))<\varepsilon$. Now set $\delta := \min(\frac{1}{2^N},\delta_1,\delta_2)$. Then it is easy to show that for any (t_1,x_1) and (t_2,x_2) in $C\times K$ with $d((t_1,x_1),(t_2,x_2))<\delta$, $d_Y(F(t_1,x_1),F(t_2,x_2))<\varepsilon$. So we are done. □

Given a compact metrisable space K and a metrisation d_Y of the topology on Y, we may define a metric d_{K,d_Y} on the set $C(K,Y)$ of continuous functions $f: K\to Y$ by $d_{K,d_Y}(f_1,f_2) = \max_{x\in K} d_Y(f_1(x),f_2(x))$. It is easy to show that this is indeed a metric, and that convergence in this metric precisely coincides with uniform convergence. Hence, by Lemma B.1, the topology induced by d_{K,d_Y} on $C(K,Y)$ is independent of the metric d_Y. We refer to this topology as the *topology of uniform convergence* or the *uniform topology.*

Lemma B.2. *Let K be a compact metrisable space. If Y is separable then $C(K,Y)$ is separable. If d_Y is complete then d_{K,d_Y} is complete. Hence if Y is Polish then $C(K,Y)$ is Polish.*

For a proof, see [Kec95, Theorem 4.19].

Lemma B.3. *Let K be a compact metrisable space. For any subbase $\mathcal{V}$ for the topology on Y, the collection of sets*

$$\mathcal{U} := \{\{f \in C(K,Y) : f(G) \subset U\} : U \in \mathcal{V},\ \textit{closed}\ G \subset K\}$$

is a subbase for the uniform topology on $C(K,Y)$.

Proof. Throughout this proof, we work with a metrisation d_Y of the topology on Y. First, fix any open $U \subset Y$ and closed $G \subset K$; for any $f \in C(K,Y)$ with $f(G) \subset U$, since $f(G)$ is compact, we have that U is a uniform neighbourhood of $f(G)$. Hence it is clear that $\{f \in C(K,Y) : f(G) \subset U\}$ is an open set in the uniform topology.

Now if we fix any subbase $\mathcal{V}$ for the topology on Y, letting $\tilde{\mathcal{V}}$ be the π-system generated by $\mathcal{V}$ (that is, the set of all finite intersections of members of $\mathcal{V}$), it is clear that π-system generated by $\mathcal{U}$ contains the set

$$\{\{f \in C(K,Y) : f(G) \subset U\} : U \in \tilde{\mathcal{V}},\ \text{closed}\ G \subset K\}.$$

Hence we may assume without loss of generality that $\mathcal{V}$ is a *base* for the topology on Y. Fix any $f_0 \in C(K,Y)$ and $\varepsilon > 0$; we need to find $W_1, \ldots, W_n \in \mathcal{U}$ such that $f_0 \in \bigcap_{i=1}^n W_i$ and for any $f \in \bigcap_{i=1}^n W_i$, $\max_{x \in K} d_Y(f_0(x), f(x)) < \varepsilon$. Since $f_0(K)$ is compact, there exist $U_1, \ldots, U_m \in \mathcal{V}$ such that $\operatorname{diam} U_j < \varepsilon$ for each $1 \le j \le m$ and $f_0(K) \subset \bigcup_{j=1}^m U_j$. Moreover (due to the Lebesgue number lemma) there exist $V_1, \ldots, V_n \in \mathcal{V}$ such that $f_0(K) \subset \bigcup_{i=1}^n V_i$ and for each $1 \le i \le n$ there exists $1 \le j_i \le m$ such that $\overline{V_i} \subset U_{j_i}$. It is clear that taking

$$W_i := \left\{f \in C(K,Y) : f\left(f_0^{-1}(\overline{V_i})\right) \subset U_{j_i}\right\}$$

for each $1 \le i \le n$ fulfils our requirement. □

Now we say that a topological space is *σ-locally compact* if it is both locally compact and σ-compact. It is not hard to show that for a metrisable topological space X the following are equivalent:

(i) X is σ-locally compact;

(ii) X is both locally compact and separable;

(iii) there exists a sequence (K_n) of compact subsets of X such that $\bigcup_n K_n^\circ = X$.

It is easy to see that if X is σ-locally compact then every closed subset of X is σ-locally compact. But also, if X is σ-locally compact then every *open* subset of X is σ-locally compact. To see this: Let $U \subset X$ be an open set, and let (K_n) be an increasing sequence of compact subsets of X such that $\bigcup_n K_n^\circ = X$; define the sequence $(\tilde{K}_n)$ of compact subsets of U by $\tilde{K}_n := \{x \in K_n : d(x, X \setminus U) \ge \frac{1}{n}\}$. Then it is clear that $\bigcup_n \tilde{K}_n^\circ = U$.

Proposition B.4. *Let X be a σ-locally compact metrisable space. Then there is a metrisable topology on $C(X,Y)$ such that a sequence (f_n) in $C(X,Y)$ converges to a function $f \in C(X,Y)$ if and only if for every compact $K \subset X$, $f_n|_K$ converges uniformly to $f|_K$. Given any sequence $(K_n)_{n \in \mathbb{N}}$ of non-empty compact subsets of X with $\bigcup_n K_n^\circ = X$,* *and any metrisation* d_{prod} *of the product topology on* $\times_{n\in\mathbb{N}} C(K_n, Y)$*, if we define the function*

$$\begin{aligned} H : C(X,Y) &\to \bigtimes_{n\in\mathbb{N}} C(K_n,Y) \\ f &\mapsto (f|_{K_1}, f|_{K_2}, f|_{K_3}, \ldots) \end{aligned}$$

then an exemplary metrisation of the above topology on $C(X,Y)$ *is*

$$(f_1, f_2) \mapsto d_{\mathrm{prod}}(H(f_1), H(f_2)).$$

We refer to the topology described in Proposition B.4 as the *topology of uniform convergence on compact sets.* Note that if K is compact then this is simply the topology of uniform convergence.

Proof of Proposition B.4. It is clear that H is injective, and therefore $(f_1, f_2) \mapsto d_{\mathrm{prod}}(H(f_1), H(f_2))$ is a metric on $C(X,Y)$. Hence, to show the desired results, it is clearly sufficient just to show that if a sequence (f_n) in $C(X,Y)$ converges to $f \in C(X,Y)$ uniformly on K_i for all $i \in \mathbb{N}$, then (f_n) converges to f uniformly on any compact $K \subset X$. But this is clear, since for any compact $K \subset X$ there must exist a finite set $S \subset \mathbb{N}$ such that $K \subset \bigcup_{i\in S} K_i^\circ$. So we are done. □

From now on, fix a σ-locally compact metrisable space X. We always assume that $C(X,Y)$ is equipped with the topology of uniform convergence on compact sets.

Corollary B.5. *If* Y *is separable (resp. completely metrisable, Polish) then* $C(X,Y)$ *is separable (resp. completely metrisable, Polish). In any case, for any subbase* $\mathcal{V}$ *for the topology on* Y *the collection of sets*

$$\mathcal{U} := \{\{f \in C(X,Y) : f(K) \subset U\} : U \in \mathcal{V},\ \text{compact } K \subset X\}$$

is a subbase for the topology on $C(X,Y)$*.*

Proof. It is clear from Proposition B.4 and Lemma B.2 that if Y is separable then $C(X,Y)$ is separable. Now it is easy to show that the function H in Proposition B.4 clearly maps $C(X,Y)$ onto a *closed* subset of $\times_n C(K_n, Y)$ (using the fact that every point in X has a neighbourhood entirely contained in one of the compact sets K_n). Hence, by Proposition B.4 and Lemma B.2, if Y is completely metrisable then $C(X,Y)$ is completely metrisable. Finally, fix a sequence $(K_n)_{n\in\mathbb{N}}$ of non-empty compact sets such that $K_i^\circ \subset K_{i+1}^\circ$ for all $i \in \mathbb{N}$ and $\bigcup_n K_n^\circ = X$. By Proposition B.4 and Lemma B.3, the collection of sets

$$\{\{f \in C(X,Y) : f(G) \subset U\} : U \in \mathcal{V},\ \text{closed } G \subset K_n,\ n \in \mathbb{N},\}$$

is a subbase for the topology on $C(X,Y)$. But this collection is precisely equal to $\mathcal{U}$. □

The following lemma in a sense justifies taking the topology of uniform convergence on compact sets as the "natural" topology on $C(X,Y)$.

Lemma B.6. *Suppose we have a metric space* T *and a function* $h{:}T \to C(X,Y)$*. Then* h *is continuous if and only if the map* $(t,x) \mapsto h(t)(x)$ *is continuous.*

Proof. Fix a metrisation d_Y of the topology on Y. First we suppose h is continuous. Fix a sequence $(t_n, x_n)_{n\in\mathbb{N}}$ converging in $T \times X$ to a point (t, x), and fix any $\varepsilon > 0$. Let $K \subset X$ be a compact set containing a neighbourhood of x. Let $N_1 \in \mathbb{N}$ be such that for all $n \geq N_1$, $x_n \in K$ and $d_Y(h(t)(x_n), h(t)(x)) < \frac{\varepsilon}{2}$. Let N_2 be such that for all $n \geq N_2$, $d_{K,d_Y}(h(t_n)|_K, h(t)|_K) < \frac{\varepsilon}{2}$. Then for every $n \geq \max(N_1, N_2)$ we have that

$$d_Y(h(t_n)(x_n), h(t)(x)) \;\leq\; d_Y(h(t_n)(x_n), h(t)(x_n)) + d_Y(h(t)(x_n), h(t)(x)) \;<\; \varepsilon.$$

Now suppose that the map $(t, x) \mapsto h(t)(x)$ is continuous. Fix a sequence $(t_n)_{n\in\mathbb{N}}$ in T converging to point t_∞ and a compact set $K \subset X$. We know that the map $(t, x) \mapsto h(t)(x)$ is uniformly continuous on $\{t_n\}_{n\in\mathbb{N}\cup\{\infty\}} \times K$. It immediately follows that as $n \to \infty$, $h(t_n)$ converges to $h(t_\infty)$ uniformly on K. □

Corollary B.7. *Let (f_n) be a sequence of continuous functions $f_n\colon X \to Y$, and let $f\colon X \to Y$ be another continuous function. The following statements are equivalent:*

(i) f_n converges to f uniformly on compact sets;

(ii) for every convergent sequence (x_n) in Y converging to a point x, $f_n(x_n) \to f(x)$ as $n \to \infty$.

Proof. Follows from Lemma B.6 with $T := \mathbb{N} \cup \{\infty\}$, $h(n) := f_n$ for $n < \infty$, and $h(\infty) = f$. □

Now let Homeo(X) be the set of homeomorphisms from X to itself, equipped with the induced topology from $C(X, X)$.

Lemma B.8. *If either (a) X is compact or (b) every point in X has a neighbourhood contained in a compact connected set, then the map*

$$\begin{aligned} \mathrm{Homeo}(X) &\to \mathrm{Homeo}(X) \\ f &\mapsto f^{-1} \end{aligned}$$

is continuous.

The case that X is compact is quite elementary, and we will soon present it. The other case has been proved in [Dij05].

Proof of Lemma B.8 for X compact.[1] Let (f_n) be a sequence of homeomorphisms $f_n\colon X \to X$ converging uniformly to a homeomorphism $f\colon X \to X$. Since X is compact, we have that f^{-1} is uniformly continuous and therefore $f^{-1} \circ f_n$ converges uniformly to the identity function. By the symmetry of $d(\cdot,\cdot)$, it follows that $f_n^{-1} \circ f$ converges uniformly to the identity function; and therefore (by right-composing with f^{-1}) f_n^{-1} converges uniformly to f^{-1}. □

Lemma B.9. *Suppose Y is separable. Then the Borel σ-algebra of $C(X, Y)$ is precisely the "evaluation σ-algebra" $\sigma(f \mapsto f(x) : x \in X)$.*

[1] Our proof will actually give the following more general fact: If (X, d) is a metric space and (f_n) is a sequence of bijective functions $f_n\colon X \to X$ converging uniformly to a bijective function $f\colon X \to X$ with f^{-1} being uniformly continuous, then f_n^{-1} converges uniformly to f^{-1}.

Proof. It is clear that for each $x \in X$ the evaluation map $f \mapsto f(x)$ from $C(X,Y)$ to Y is continuous and therefore Borel-measurable. So the evaluation σ-algebra is contained in $\mathcal{B}(C(X,Y))$.

Now let $(K_n)_{n\in\mathbb{N}}$ be a sequence of compact subsets of X such that $\bigcup_n K_n^\circ = X$ and let $\{U_n\}_{n\in\mathbb{N}}$ be a countable subbase for the topology on Y. We have that

$$\mathcal{B}(C(X,Y)) \;=\; \sigma(\,\{f \in C(X,Y) : f(K_n) \subset U_m\} \,:\, m,n \in \mathbb{N}\,).$$

Fix $m,n \in \mathbb{N}$, and let d_Y be a metrisation of the topology on Y. For each $i \in \mathbb{N}$, let $G_i := \{x \in U_m : d_Y(x, Y \setminus U_m) \geq \frac{1}{i}\}$. For any $f \in C(X,Y)$, since $f(K_n)$ is compact, we have that $f(K_n) \subset U_m$ if and only if U_m is a uniform neighbourhood of $f(K_n)$, which is the same as saying that there exists $i \in \mathbb{N}$ such that $f(K_n) \subset G_i$. So if we let E be a countable dense subset of K_n then we have that

$$\{f \in C(X,Y) : f(K_n) \subset U_m\} \;=\; \bigcup_{i=1}^{\infty} \bigcap_{x\in E} \{f \in C(X,Y) : f(x) \in G_i\}.$$

Hence $\{f \in C(X,Y) : f(K_n) \subset U_m\}$ is a member of the evaluation σ-algebra on $C(X,Y)$. It follows that $\mathcal{B}(C(X,Y))$ is contained in the evaluation σ-algebra. □

Note in particular that if Y is Polish then the evaluation σ-algebra on $C(X,Y)$ is standard.

Corollary B.10. *Suppose Y is separable. For any compact $K \subset X$, the σ-algebra $\sigma(f \mapsto f(x) : x \in K)$ is the Borel σ-algebra of the smallest topology with respect to which the restriction map $f \mapsto f|_K$ is continuous, where $f|_K$ is regarded as a member of $C(K,Y)$ equipped with the uniform topology.*

Appendix C. Ergodic Theory and Markov Processes

The proofs of most of the results in this appendix can be found in [New15a].

We will state results first for dynamical systems (in discrete and continuous time), and then for Markov transition probabilities (in discrete and continuous time). However, a dynamical system is really just the "deterministic case" of a family of Markov transition probabilities.

(Hence many of the results given here for dynamical systems are not proved *separately* for dynamical systems in [New15a], but are special cases of results for Markov transition probabilities that *are* proved explicitly in [New15a].)

As in the main body of the thesis, $\mathbb{T}^+$ denotes either $\mathbb{N}_0$ or $[0,\infty)$. Given a measurable space (X,Σ), $\mathcal{M}_{(X,\Sigma)}$ denotes the set of probability measures on X. As throughout the rest of the thesis, $\mathcal{M}_{(X,\Sigma)}$ is equipped with the σ-algebra characterised by the following property: for any measurable space $(\Omega,\mathcal{F})$, a map $p\colon\Omega\to\mathcal{M}_{(X,\Sigma)}$ is measurable if and only if the map $\omega\mapsto p(\omega)(A)$ is measurable for all $A\in\Sigma$.

C.1 Ergodic theory for measurable maps

Invariant and ergodic measures

Let (X,Σ) be a measurable space, and let $f\colon X\to X$ be a measurable map.

We say that a probability measure ρ on X is *f-invariant* (or *invariant under f*) if $f_*\rho=\rho$ (i.e. $\rho(f^{-1}(A))=\rho(A)$ for all $A\in\Sigma$). In this case, we also say that f *is ρ-preserving*, or that f *is a measure-preserving transformation of* (X,Σ,ρ), or that (X,Σ,ρ,f) *is a measure-preserving dynamical system.*

Note that any convex combination of f-invariant probability measures is f-invariant.

Given an f-invariant probability measure ρ, we will say that a set $A\in\Sigma$ is *ρ-almost invariant* (under f) if the following equivalent statements hold:

(i) $\rho(A\setminus f^{-1}(A))=0$ (i.e. for ρ-almost all $x\in A$, $f(x)\in A$);

(ii) $\rho(f^{-1}(A)\setminus A)=0$ (i.e. for ρ-almost all $x\in X\setminus A$, $f(x)\in X\setminus A$);

(iii) $\rho(A\triangle f^{-1}(A))=0$ (i.e. for ρ-almost all $x\in X$, $x\in A\Leftrightarrow f(x)\in A$).

It is not hard to show that the set $\mathcal{I}^f_\rho$ of all ρ-almost invariant sets $A\in\Sigma$ forms a sub-σ-algebra of Σ.

We will say that a set $A \in \Sigma$ is *strictly invariant* (under f) if $f^{-1}(A) = A$. Again, it is not hard to show that the set $\mathcal{I}^f$ of strictly invariant sets $A \in \Sigma$ forms a sub-σ-algebra of Σ.

Given an f-invariant probability measure ρ, we will say that a measurable function $g: X \to \mathbb{R}$ is ρ-*almost invariant* (under f) if the following equivalent statements hold:

(i) $g(f(x)) = g(x)$ for ρ-almost all $x \in X$;

(ii) $g(f(x)) \geq g(x)$ for ρ-almost all $x \in X$;

(iii) $g(f(x)) \leq g(x)$ for ρ-almost all $x \in X$;

(iv) g is measurable with respect to $\mathcal{I}^f_\rho$.

We will say that a probability measure ρ on X is *ergodic with respect to* f (or f-*ergodic*) if the following equivalent statements hold:

(i) ρ is f-invariant, and $\rho(A) \in \{0, 1\}$ for every ρ-almost invariant set $A \in \Sigma$;

(ii) ρ is f-invariant, and $\rho(A) \in \{0, 1\}$ for every strictly invariant set $A \in \Sigma$;

(iii) ρ is f-invariant, and for every measurable ρ-almost invariant $g : X \to \mathbb{R}$ there exists $c \in \mathbb{R}$ such that $g(x) = c$ for ρ-almost all $x \in X$;

(iv) ρ is f-invariant, and the only f-invariant probability measure that is absolutely continuous with respect to ρ is ρ itself;

(v) ρ is an extreme point of the convex set of f-invariant probability measures (that is to say, ρ is f-invariant and cannot be expressed as a non-trivial convex combination of two distinct f-invariant probability measures).

In this case, we will also say that f *is an ergodic* (*measure-preserving*) *transformation of* (X, Σ, ρ), or that (X, Σ, ρ, f) *is an ergodic* (*measure-preserving*) *dynamical system*.

It is known that any two distinct ergodic probability measures are mutually singular.

Birkhoff's ergodic theorem

Let f be a measure-preserving transformation of a probability space (X, Σ, ρ), and let $g: X \to \mathbb{R}$ be a ρ-integrable function. Then

$$\frac{1}{n} \sum_{i=0}^{n-1} g(f^i(x)) \;\to\; \mathbb{E}_{(\rho)}[g|\mathcal{I}^f](x) \quad \text{as } n \to \infty$$

for ρ-almost all $x \in X$.[1] (This statement also holds with $\mathcal{I}^f$ replaced by $\mathcal{I}^f_\rho$, as it can be shown that $\mathcal{I}^f$ and $\mathcal{I}^f_\rho$ agree modulo ρ-null sets.)

[1] Obviously here we *fix*, independently of x, a version $\mathbb{E}_{(\rho)}[g|\mathcal{I}^f]: X \to \mathbb{R}$ of the conditional expectation of g given $\mathcal{I}^f$.

In particular, if ρ is ergodic then

$$\frac{1}{n}\sum_{i=0}^{n-1} g(f^i(x)) \to \int_X g\,d\rho \quad \text{as } n \to \infty$$

for ρ-almost all $x \in X$.

Ergodic decomposition

Let (X, Σ) be a standard measurable space (meaning, as in Section 2.2, that Σ can be expressed as the Borel σ-algebra of a Polish topology on X).

Let ρ be a probability measure on X, and let $\mathcal{I}$ be a sub-σ-algebra of Σ. As a special case of the disintegration theorem (Lemma 3.27), one can show the following: There exists (unique up to ρ-almost everywhere equality) a measurable function $\rho(|\mathcal{I}) : X \to \mathcal{M}_{(X,\Sigma)}$ with the property that for every $A \in \Sigma$, the map $x \mapsto \rho(|\mathcal{I})(x)(A)$ is a version of the conditional probability $\rho(A|\mathcal{I})$. The function $\rho(|\mathcal{I})$ is referred to as (*a version of*) *the conditional distribution of* ρ *given* $\mathcal{I}$.

Now let $f : X \to X$ be a measurable map, and let ρ be an f-invariant probability measure. Fix a version $\rho(|\mathcal{I}^f)$ of the conditional distribution of ρ given $\mathcal{I}^f$. As an equation to be evaluated at each set $A \in \Sigma$, we have (trivially, from the definition of a conditional distribution) the following integral representation of ρ:

$$\rho = \int_X \rho(|\mathcal{I}^f)(x)\,\rho(dx). \tag{C.1}$$

It is not hard to show that for ρ-almost all $x \in X$, the probability measure $\rho(|\mathcal{I}^f)(x)$ is f-invariant; moreover, using Birkhoff's ergodic theorem, one can show that for ρ-almost all $x \in X$, $\rho(|\mathcal{I}^f)(x)$ is *ergodic* with respect to f.[2] Hence equation (C.1) is referred to as an *ergodic decomposition of* ρ.

(Once again, in all the above we can replace the σ-algebra $\mathcal{I}^f$ with $\mathcal{I}^f_\rho$.)

Note, in particular, that as a consequence we have the following: if a measurable map on a standard measurable space admits an invariant probability measure, then it admits an ergodic probability measure; moreover, if it admits an invariant probability measure assigning full measure to some set A, then it admits an ergodic probability measure assigning full measure to the same set A.

Continuous maps

Let (X, d) be a separable metric space. Recall that for any Borel probability measure ρ on X, the *support of* ρ (denoted $\operatorname{supp}\rho$) is defined as the smallest closed ρ-full measure

[2] It is clear that for *each* strictly invariant $A \in \Sigma$, for ρ-almost all $x \in X$, $\rho(|\mathcal{I}^f)(x)$ assigns trivial measure to A. However, this does *not* automatically imply that $\rho(|\mathcal{I}^f)(x)$ is ergodic for ρ-almost all $x \in X$, since there may be *uncountably many* strictly invariant sets, and $\mathcal{I}^f$ need not even be countably generated.

subset of X.[3] Note that this is precisely the set of points in X all of whose open neighbourhoods have strictly positive measure according to ρ.

Let $f\colon X \to X$ be a continuous map.

We say that a set $A \subset X$ is (*forward-*)*invariant* (under f) if $f(A) \subset A$. Obviously, an arbitrary intersection of invariant sets is invariant. Note that for any f-invariant probability measure ρ, if $A \in \mathcal{B}(X)$ is invariant then A is ρ-almost invariant, since $A \setminus f^{-1}(A)$ is empty (and is therefore obviously a ρ-null set).

Obviously, for any $x \in X$, the smallest invariant set containing x is the locus of its trajectory, $\{f^n(x) : n \in \mathbb{N}_0\}$. Now, for any $x \in X$, let $G_x := \overline{\{f^n(x) : n \in \mathbb{N}_0\}}$. So any *closed* invariant set containing the point x must contain the set G_x. But moreover, observe that G_x is itself invariant: for any $y \in G_x$, letting $(m_n)_{n\in\mathbb{N}}$ be a sequence in $\mathbb{N}_0$ such that $f^{m_n}(x) \to y$ as $n \to \infty$, we will have (due to the continuity of f) that $f^{m_n+1}(x) \to f(y)$ as $n \to \infty$.

So then, G_x is the smallest closed invariant set containing x.

Now it is easy to check that for any closed invariant $G \subset X$, the following two statements are equivalent:

(i) the only closed invariant proper subset of G is $\varnothing$;

(ii) G is non-empty, and for all $x \in G$, $G_x = G$.

When these statements hold, we will say that G is *minimal* (with respect to f). Since the intersection of two closed invariant sets is a closed invariant set, it is clear that any two disinct minimal sets must be mutually disjoint.

Proposition. *Every non-empty compact invariant set contains at least one minimal set.*

Hence in particular, if X is compact, then: (a) there exists at least one minimal set; and (b) if there is only one minimal set K, then K is the smallest non-empty closed invariant set.

Our proof is taken from [New15b, Proposition 1.2.6] (which is, in turn, loosely adapted from the solution to Exercicse 3.3.4 of [KH95]).

Proof. Given non-empty compact sets $C_2 \subset C_1 \subset X$, let $d_H(C_1, C_2) := \max_{x\in C_1} d(x, C_2)$. For any non-empty compact invariant $C \subset X$, we write $I(C)$ for the set of non-empty closed invariant subsets of C, and we write

$$m(C) := \sup_{\tilde{C}\in I(C)} d_H(C, \tilde{C}).$$

Now fix a non-empty compact invariant set $C_0 \subset X$; we will show that C_0 contains a minimal set K. Let $C_0 \supset C_1 \supset C_2 \supset \ldots$ be a nested sequence of non-empty closed invariant

[3] This exists due to the existence of a countable base for the topology of X.

sets, with $d_H(C_n, C_{n+1}) \geq \frac{n}{n+1} m(C_n)$ for all $n \in \mathbb{N}_0$. Cantor's intersection theorem gives that $K := \bigcap_{n=0}^{\infty} C_n$ is non-empty; and obviously K is invariant. Now since C_0 is totally bounded, we must have that $d_H(C_n, C_{n+1}) \to 0$ as $n \to \infty$, and so $m(C_n) \to 0$ as $n \to \infty$. It is easy to see that $m(\cdot)$ is monotone, so it follows that $m(K) = 0$. Hence K is minimal. □

Now it is not hard to show that for any f-invariant probability measure ρ, $\operatorname{supp} \rho$ is invariant.[4] The *Krylov-Bogolyubov theorem* gives a kind of "partial converse": for any non-empty compact invariant $G \subset X$, there exists an f-invariant probability measure ρ such that $\rho(G) = 1$ (i.e. such that $\operatorname{supp} \rho \subset G$). Since G is compact, by restricting f to G we can obtain an ergodic decomposition of ρ; and so there must exist at least one f-ergodic probability measure $\tilde{\rho}$ such that $\tilde{\rho}(G) = 1$.

Note, in particular, that for any compact minimal $K \subset X$, there must exist at least one ergodic probability measure ρ such that $\operatorname{supp} \rho = K$. Also note that if X is compact and f admits only one invariant probability measure ρ, then $\operatorname{supp} \rho$ is the smallest non-empty closed invariant set.

C.2 Ergodic theory for dynamical systems

An (*autonomous*) *dynamical system* on a set X is a $\mathbb{T}^+$-indexed family of $(f^t)_{t \in \mathbb{T}^+}$ of functions $f^t \colon X \to X$ such that the "flow equations"

(i) $f^0 = \mathrm{id}_X$;

(ii) $f^{s+t} = f^t \circ f^s$ for all $s, t \in \mathbb{T}^+$.

Sometimes an individual function $h \colon X \to X$ is called a dynamical system on X, because it naturally generates the discrete-time dynamical system $(h^n)_{n \in \mathbb{N}_0}$. (Indeed, we made reference to this use of terminology in Section C.1 when we mentioned "measure-preserving dynamical systems".)

Given a measurable space (X, Σ), an (*autonomous*) *dynamical system on* (X, Σ) is a dynamical system (f^t) on the set X with the additional property that f^t is (Σ, Σ)-measurable for all $t \in \mathbb{T}^+$. We will say that a dynamical system (f^t) on (X, Σ) is *measurable* if the map $(t, x) \mapsto f^t(x)$ is $(\mathcal{B}(\mathbb{T}^+) \otimes \Sigma, \Sigma)$-measurable. (Obviously, if $\mathbb{T}^+ = \mathbb{N}_0$ then any dynamical system on (X, Σ) is measurable.)

Invariant and ergodic measures

Let $(f^t)_{t \in \mathbb{T}^+}$ be a dynamical system on a measurable space (X, Σ).

We say that a probability measure ρ on X is (f^t)-*invariant* (or *invariant under* (f^t)) if ρ is f^t-invariant for every $t \in \mathbb{T}^+$. In this case, we also say that (f^t) *is ρ-preserving*, or that $(X, \Sigma, \rho, (f^t))$ *is a measure-preserving dynamical system.* Note that if $\mathbb{T}^+ = \mathbb{N}_0$, then a probability measure ρ on X is (f^t)-invariant if and only if ρ is f^1-invariant.

[4] In fact, it is not hard to show that $f(\operatorname{supp} \rho)$ is a dense subset of $\operatorname{supp} \rho$. Hence in particular, if $\operatorname{supp} \rho$ is compact then $f(\operatorname{supp} \rho) = \operatorname{supp} \rho$.

Given an (f^t)-invariant probability measure ρ, we will say that a set $A \in \Sigma$ is *ρ-almost invariant* (under (f^t)) if A is ρ-almost invariant under f^t for every $t \in \mathbb{T}^+$. It is easy to show that if $\mathbb{T}^+ = \mathbb{N}_0$, then a set $A \in \Sigma$ is ρ-almost invariant under (f^t) if and only if A is ρ-almost invariant under f^1. Let $\mathcal{I}_\rho^{(f^t)}$ denote the set of ρ-almost invariant sets, that is,

$$\mathcal{I}_\rho^{(f^t)} = \bigcap_{t \in \mathbb{T}^+} \mathcal{I}_\rho^{f^t}.$$

Obviously $\mathcal{I}_\rho^{(f^t)}$ is a sub-σ-algebra of Σ.

We will say that a set $A \in \Sigma$ is *strictly invariant* (under (f^t)) if A is strictly invariant under f^t for all $t \in \mathbb{T}^+$. It is easy to show that if $\mathbb{T}^+ = \mathbb{N}_0$, then a set $A \in \Sigma$ is strictly invariant under (f^t) if and only if A is strictly invariant under f^1. Let $\mathcal{I}^{(f^t)}$ denote the set of strictly invariant sets, that is,

$$\mathcal{I}^{(f^t)} = \bigcap_{t \in \mathbb{T}^+} \mathcal{I}^{f^t}.$$

Obviously $\mathcal{I}^{(f^t)}$ is a sub-σ-algebra of Σ.

Given an (f^t)-invariant probability measure ρ, we will say that a measurable function $g: X \to \mathbb{R}$ is *ρ-almost invariant* (under (f^t)) if the following equivalent statements hold:

(i) g is ρ-almost invariant under f^t for all $t \in \mathbb{T}^+$;

(ii) g is measurable with respect to $\mathcal{I}_\rho^{(f^t)}$.

Note that if $\mathbb{T}^+ = \mathbb{N}_0$, then a measurable function $g: X \to \mathbb{R}$ is ρ-almost invariant under (f^t) if and only if g is ρ-almost invariant under f^1.

We will say that a probability measure ρ on X is *ergodic with respect to* (f^t) (or (f^t)*-ergodic*) if the following equivalent statements hold:

(i) ρ is (f^t)-invariant, and $\rho(A) \in \{0, 1\}$ for every ρ-almost invariant set $A \in \Sigma$;

(ii) ρ is (f^t)-invariant, and for every measurable ρ-almost invariant $g : X \to \mathbb{R}$ there exists $c \in \mathbb{R}$ such that $g(x) = c$ for ρ-almost all $x \in X$;

(iii) ρ is (f^t)-invariant, and the only (f^t)-invariant probability measure that is absolutely continuous with respect to ρ is ρ itself;

(iv) ρ is an extreme point of the convex set of (f^t)-invariant probability measures.

In this case, we will also say that $(X, \Sigma, \rho, (f^t))$ is an *ergodic* (*measure-preserving*) *dynamical system*.

Once again, any two distinct ergodic probability measures are mutually singular.

Note that if ρ is an (f^t)-invariant probability measure and there exists $\tau \in \mathbb{T}^+$

such that ρ is ergodic with respect to the map f^τ, then ρ is ergodic with respect to the dynamical system (f^t). Note also that if $\mathbb{T}^+ = \mathbb{N}_0$ then a probability measure ρ on X is ergodic with respect to (f^t) if and only if ρ is ergodic with respect to f^1.

Now if (f^t) is *measurable*, then for any (f^t)-invariant probability measure ρ, the following are equivalent:

(i) ρ is (f^t)-ergodic;

(ii) $\rho(A) \in \{0,1\}$ for every strictly invariant set $A \in \Sigma$.

Birkhoff's ergodic theorem for semiflows

Assume $\mathbb{T}^+ = [0,\infty)$. Let $(X,\Sigma,\rho,(f^t))$ be a measure-preserving dynamical system, with (f^t) measurable, and let $g\colon X \to \mathbb{R}$ be a ρ-integrable function.

Then for ρ-almost all $x \in X$, the map $t \mapsto g(f^t(x))$ is locally integrable and

$$\frac{1}{T}\int_0^T g(f^t(x))\,dt \;\to\; \mathbb{E}_{(\rho)}[g|\mathcal{I}^{(f^t)}](x) \quad \text{as } T \to \infty.$$

(This statement also holds with $\mathcal{I}^{(f^t)}$ replaced by $\mathcal{I}_\rho^{(f^t)}$, as it can be shown that due to the measurability of (f^t), $\mathcal{I}^{(f^t)}$ and $\mathcal{I}_\rho^{(f^t)}$ agree modulo ρ-null sets.)

In particular, if ρ is ergodic then

$$\frac{1}{T}\int_0^T g(f^t(x))\,dt \;\to\; \int_X g\,d\rho \quad \text{as } T \to \infty$$

for ρ-almost all $x \in X$.

Ergodic decomposition

Let (f^t) be a measurable dynamical system on a standard measurable space (X,Σ). Then for any (f^t)-invariant probability measure ρ on X, we have (as in Section C.1) the integral representation

$$\rho \;=\; \int_X \rho(|\mathcal{I}^{(f^t)})(x)\,\rho(dx)$$

and one can show that $\rho(|\mathcal{I}^{(f^t)})(x)$ is (f^t)-ergodic for ρ-almost every $x \in X$.

(Once again, we can also replace $\mathcal{I}^{(f^t)}$ with $\mathcal{I}_\rho^{(f^t)}$.)

Spatially continuous dynamical systems

Let (X,d) be a separable metric space. Let (f^t) be a dynamical system on X such that f^t is continuous for all $t \in \mathbb{T}^+$.

We say that a set $A \subset X$ is *(forward-)invariant* (under (f^t)) if A is invariant under f^t for all $t \in \mathbb{T}^+$. In the case that $\mathbb{T}^+ = \mathbb{N}_0$, A is invariant under (f^t) if and only if A is invariant under f^1. Note that, once again, an arbitrary intersection of invariant sets is invariant.

For any $x \in X$, let $G_x := \overline{\{f^t(x) : t \in \mathbb{T}^+\}}$. Once again, it is easy to show that G_x is the smallest closed invariant set containing x.

We say that a set $G \subset X$ is *minimal* (with respect to (f^t)) if the following equivalent statements hold:

(i) G is closed and invariant, and the only closed invariant proper subset of G is $\varnothing$;

(ii) G is a non-empty closed invariant set, and for all $x \in G$, $G_x = G$.

Note that if a closed set $G \subset X$ is invariant under (f^t) and there exists $\tau \in \mathbb{T}^+$ such that G is minimal with respect to f^τ, then G is minimal with respect to (f^t). Also note that if $\mathbb{T}^+ = \mathbb{N}_0$ then a set $G \subset X$ is minimal with respect to (f^t) if and only if G is minimal with respect to f^1.

Exactly the same proof as in Section C.1 gives that every non-empty compact invariant set contains at least one minimal set.

Once again, the support of any (f^t)-invariant probability measure is invariant. If (f^t) is measurable as a dynamical system on $(X, \mathcal{B}(X))$,[5] then every non-empty compact invariant contains the support of at least one (f^t)-ergodic probability measure, and every compact minimal set is equal to the support of at least one (f^t)-ergodic probability measure. If, in addition, X is compact and (f^t) admits only one invariant probability measure ρ, then $\operatorname{supp}\rho$ is the smallest non-empty closed invariant set.

C.3 Ergodic theory for Markov kernels

Let (X, Σ) be a measurable space. A *Markov kernel* (or *family of one-step transition probabilities*) *on* X is an X-indexed family $(\mu_x)_{x\in X}$ of probability measures on X, such that the map $x \mapsto \mu_x(A)$ is measurable for all $A \in \Sigma$.[6] Note that for any measurable function $f : X \to X$, $(\delta_{f(x)})_{x\in X}$ is a Markov kernel on X. We refer to the Markov kernel $(\delta_x)_{x\in X}$ as the *identity kernel*.

Let (μ_x) be a Markov kernel on X. For any probability measure ρ on X, we define the probability measure $\mu^*\rho$ on X by

$$\mu^*\rho(A) \ := \ \int_X \mu_x(A)\,\rho(dx).$$

Note that if $(\mu_x) = (\delta_{f(x)})$ for some measurable $f : X \to X$, then for any probability measure ρ on X, $\mu^*\rho = f_*\rho$.

[5] A sufficient condition for this is that the map $t \mapsto f^t(x)$ is right-continuous for all $x \in X$.

[6] We prefer the "$\mu_x(A)$" notation to the (perhaps more common) "$P(x, A)$" notation, as it allows us to consider individual probability measures μ_x without unnecessary cumbersomeness of notation.

We say that a probability measure ρ on X is *stationary under* (μ_x) (or (μ_x)-*stationary*) if $\mu^*\rho = \rho$. (In particular, given any measurable $f: X \to X$, a probability measure ρ on X is stationary under $(\delta_{f(x)})$ if and only if ρ is f-invariant. Note that *every* probability measure on X is stationary under the identity kernel.)

Given a (μ_x)-stationary probability measure ρ, for any $A \in \Sigma$ with $\rho(A) = 1$, it is clear that $\mu_x(A) = 1$ for ρ-almost all $x \in X$. (This does *not*, however, imply that μ_x is absolutely continuous with respect to ρ for ρ-almost all $x \in X$, since the collection of all ρ-full-measure members of Σ is not generally countable. Indeed, as a simple counter-example: the Lebesgue measure on $[0,1]$ is stationary with respect to the identity kernel on $[0,1]$, and yet there does not exist $x \in [0,1]$ such that δ_x is absolutely continuous with respect to the Lebesgue measure.)

Note that any convex combination of (μ_x)-stationary probability measures is (μ_x)-stationary.

Given a (μ_x)-stationary probability measure ρ, we will say that a set $A \in \Sigma$ is *ρ-almost invariant* (according to (μ_x)) if the following equivalent statements hold:

(i) for ρ-almost all $x \in A$, $\mu_x(A) = 1$;

(ii) for ρ-almost all $x \in X \setminus A$, $\mu_x(A) = 0$;

(iii) for ρ-almost all $x \in X$, $\mu_x(A) = \mathbb{1}_A(x)$.

It is not hard to show that the set $\mathcal{I}_\rho^{(\mu_x)}$ of all ρ-almost invariant sets $A \in \Sigma$ forms a sub-σ-algebra of Σ.

Note that, given a measurable map $f: X \to X$ and an f-invariant probability measure ρ, a set $A \in \Sigma$ is ρ-almost invariant according to the Markov kernel $(\delta_{f(x)})$ if and only if it is ρ-almost invariant under f.

Given a probability measure ρ on X and a ρ-integrable function $g: X \to \mathbb{R}$, we write $\rho(g)$ as a shorthand for $\int_X g(x)\,\rho(dx)$. Given a (μ_x)-stationary probability measure ρ, we will say that a bounded measurable function $g: X \to \mathbb{R}$ is *ρ-almost invariant* (according to (μ_x)) if the following equivalent statements hold:

(i) $\mu_x(\, y \in X \,:\, g(y) = g(x)\,) = 1$ for ρ-almost all $x \in X$;

(ii) $\mu_x(g) = g(x)$ for ρ-almost all $x \in X$;

(iii) $\mu_x(g) \geq g(x)$ for ρ-almost all $x \in X$;

(iv) $\mu_x(g) \leq g(x)$ for ρ-almost all $x \in X$;

(v) g is measurable with respect to $\mathcal{I}_\rho^{(\mu_x)}$.

We will say that a probability measure ρ on X is *ergodic with respect to* (μ_x) (or (μ_x)-*ergodic*) if the following equivalent statements hold:

(i) ρ is (μ_x)-stationary, and $\rho(A) \in \{0,1\}$ for every ρ-almost invariant set $A \in \Sigma$;

(ii) ρ is (μ_x)-stationary, and for every bounded measurable ρ-almost invariant function $g : X \to \mathbb{R}$ there exists $c \in \mathbb{R}$ such that $g(x) = c$ for ρ-almost all $x \in X$;

(iii) ρ is (μ_x)-stationary, and the only (μ_x)-stationary probability measure that is absolutely continuous with respect to ρ is ρ itself;

(iv) ρ is an extreme point of the convex set of (μ_x)-stationary probability measures.

Note that, given a measurable map $f : X \to X$, a probability measure ρ on X is ergodic with respect to f if and only if it is ergodic with respect to $(\delta_{f(x)})$.

Once again, any two distinct ergodic probability measures are mutually singular.

Ergodic decompositions and continuity of Markov kernels will be considered in the next section.

C.4 Ergodic theory for semigroups of Markov kernels

Let (X, Σ) be a measurable space. A *family of Markov transition probabilities* or a *semigroup of Markov kernels* on X is an $(X \times \mathbb{T}^+)$-indexed family $(\mu_x^t)_{x \in X,\, t \in \mathbb{T}^+}$ of probability measures μ_x^t on X such that the following hold:

(i) the map $x \mapsto \mu_x^t(A)$ is measurable for each $A \in \Sigma$ and $t \in \mathbb{T}^+$;

(ii) $\mu_x^0 = \delta_x$ for all $x \in X$ (i.e. $(\mu_x^0)_{x \in X}$ is the identity kernel);

(iii) for all $x \in X$, $s, t \in \mathbb{T}^+$ and $A \in \Sigma$, the "Chapman-Kolmogorov relation"

$$\mu_x^{s+t}(A) \;=\; \int_X \mu_y^t(A)\, \mu_x^s(dy)$$

is satisfied.

Obviously $(\mu_x^t)_{x \in X}$ is a Markov kernel on X for each $t \in \mathbb{T}^+$. So, using the notation introduced in Section C.3, point (iii) can be expressed slightly more succinctly as

$$\mu_x^{s+t} \;=\; \mu^{t*}\mu_x^s$$

for all $x \in X$ and $s, t \in \mathbb{T}^+$. Note that for any Markov kernel (μ_x) on X, there is a unique discrete-time family of Markov transition probabilities $(\mu_x^n)_{x \in X,\, n \in \mathbb{N}_0}$ such that $(\mu_x^1) = (\mu_x)$; this can be constructed explicitly by the recursive relation

$$\mu_x^0 \;=\; \delta_x\,; \quad \mu_x^{n+1}(A) \;=\; \int_X \mu_y^n(A)\, \mu_x(dy) \;\text{ for } n \geq 0.$$

It is easy to check that for any dynamical system (f^t) on (X, Σ), $(\delta_{f^t(x)})_{x \in X,\, t \in \mathbb{T}^+}$ is a family of Markov transition probabilities.

Wherever we do not include subscripts after (μ_x^t), assume that (μ_x^t) refers to the whole family of Markov transition probabilities $(\mu_x^t)_{x \in X,\, t \in \mathbb{T}^+}$.

We say that a family of Markov transition probabilities (μ_x^t) is *measurable* if the map $(x,t) \mapsto \mu_x^t(A)$ is $(\Sigma \otimes \mathcal{B}(\mathbb{T}^+), \mathcal{B}([0,1]))$-measurable for every $A \in \Sigma$. Note that if $\mathbb{T}^+ = \mathbb{N}_0$ then every family of Markov transition probabilities is measurable. Also note that for any measurable dynamical system (f^t) on (X,Σ), $(\delta_{f^t(x)})$ is measurable.

Sometimes, for convenience, we just use the terms *kernel* and *semigroup* to refer, respectively, to a Markov kernel or semigroup of Markov kernels.

Stationary and ergodic measures

Let (μ_x^t) be a family of Markov transition probabilities on a measurable space (X,Σ). We say that a probability measure ρ on X is *stationary under* (μ_x^t) (or (μ_x^t)*-stationary*) if ρ is stationary under the kernel $(\mu_x^t)_{x \in X}$ for each $t \in \mathbb{T}^+$. In the case that $\mathbb{T}^+ = \mathbb{N}_0$, a probability measure ρ is stationary under (μ_x^t) if and only if ρ is stationary under the kernel $(\mu_x^1)_{x \in X}$.

Given a (μ_x^t)-stationary probability measure ρ, we will say that a set $A \in \Sigma$ is ρ-*almost invariant* (according to (μ_x^t)) if A is ρ-almost invariant according to the kernel $(\mu_x^t)_{x \in X}$ for each $t \in \mathbb{T}^+$. In the case that $\mathbb{T}^+ = \mathbb{N}_0$, A is ρ-almost invariant according to (μ_x^t) if and only if A is ρ-almost invariant according to the kernel $(\mu_x^1)_{x \in X}$.

Let $\mathcal{I}_\rho^{(\mu_x^t)}$ denote the set of ρ-almost invariant sets, that is,

$$\mathcal{I}_\rho^{(\mu_x^t)} = \bigcap_{t \in \mathbb{T}^+} \mathcal{I}_\rho^{(\mu_x^t)_{x \in X}}.$$

Obviously $\mathcal{I}_\rho^{(\mu_x^t)}$ is a sub-σ-algebra of Σ.

Given a (μ_x^t)-stationary probability measure ρ, we will say that a measurable function $g\colon X \to \mathbb{R}$ is ρ-*almost invariant* (according to (μ_x^t)) if the following equivalent statements hold:

(i) g is ρ-almost invariant according to the kernel $(\mu_x^t)_{x \in X}$ for each $t \in \mathbb{T}^+$;

(ii) g is measurable with respect to $\mathcal{I}_\rho^{(\mu_x^t)}$.

Note that if $\mathbb{T}^+ = \mathbb{N}_0$, then a measurable function $g\colon X \to \mathbb{R}$ is ρ-almost invariant according to (μ_x^t) if and only if g is ρ-almost invariant according to the kernel (μ_x^1).

We will say that a probability measure ρ on X is *ergodic with respect to* (μ_x^t) (or (μ_x^t)*-ergodic*) if the following equivalent statements hold:

(i) ρ is (μ_x^t)-stationary, and $\rho(A) \in \{0,1\}$ for every ρ-almost invariant set $A \in \Sigma$;

(ii) ρ is (μ_x^t)-stationary, and for every bounded measurable ρ-almost invariant function $g : X \to \mathbb{R}$ there exists $c \in \mathbb{R}$ such that $g(x) = c$ for ρ-almost all $x \in X$;

(iii) ρ is (μ_x^t)-stationary, and the only (μ_x^t)-stationary probability measure that is absolutely continuous with respect to ρ is ρ itself;

(iv) ρ is an extreme point of the convex set of (μ_x^t)-stationary probability measures.

Once again, any two distinct ergodic probability measures are mutually singular.

Note that if ρ is a (μ_x^t)-stationary probability measure and there exists $\tau \in \mathbb{T}^+$ such that ρ is ergodic with respect to the kernel $(\mu_x^\tau)_{x\in X}$, then ρ is ergodic with respect to the semigroup (μ_x^t). Note also that if $\mathbb{T}^+ = \mathbb{N}_0$ then a probability measure ρ on X is ergodic with respect to (μ_x^t) if and only if ρ is ergodic with respect to $(\mu_x^1)_{x\in X}$.

Ergodic decomposition

Let (μ_x^t) be a measurable family of Markov transition probabilities on a standard measurable space (X,Σ). Then for any (μ_x^t)-stationary probability measure ρ on X, we have the integral representation

$$\rho = \int_X \rho(|\mathcal{I}_\rho^{(\mu_x^t)})(y)\,\rho(dy)$$

and one can show that $\rho(|\mathcal{I}_\rho^{(\mu_x^t)})(y)$ is (μ_x^t)-ergodic for ρ-almost every $y \in X$.

Feller-continuous Markov transition probabilities

Let (X,d) be a separable metric space. We will say that a Markov kernel (μ_x) on X is *Feller-continuous* if the map $x \mapsto \mu_x$ is continuous with respect to the narrow topology. We will say that a family of Markov transition probabilities (μ_x^t) on X is Feller-continuous if the kernel $(\mu_x^t)_{x\in X}$ is Feller-continuous for every $t \in \mathbb{T}^+$; in the case that $\mathbb{T}^+ = \mathbb{N}_0$, this is equivalent to saying that the kernel $(\mu_x^1)_{x\in X}$ is Feller-continuous.

We say that a set $A \in \mathcal{B}(X)$ is *forward-invariant* according to a Markov kernel (μ_x) on X if $\mu_x(A) = 1$ for every $x \in A$. We say that $A \in \mathcal{B}(X)$ is *forward-invariant* according to a family of Markov transition probabilities (μ_x^t) on X if A is forward-invariant according to the kernel $(\mu_x^t)_{x\in X}$ for every $t \in \mathbb{T}^+$; if $\mathbb{T}^+ = \mathbb{N}_0$ then this is equivalent to saying that A is forward-invariant according to the kernel $(\mu_x^1)_{x\in X}$.

For either a kernel (μ_x) or a semigroup (μ_x^t), it is not hard to show (using the fact that there is a countable base for the topology of X) that an arbitrary intersection of closed forward-invariant sets is forward-invariant.

Now let (μ_x^t) be a Feller-continuous family of Markov transition probabilities on X, and for any $x \in X$, let

$$G_x := \overline{\bigcup_{t\in\mathbb{T}^+} \operatorname{supp}\mu_x^t}.$$

Fix $x \in X$. For any open $U \subset X$, it is easy to see that $U \cap G_x \neq \varnothing$ if and only if there exists $t \in \mathbb{T}^+$ such that $\mu_x^t(U) > 0$. In other words, G_x is precisely the set of points y such that for every neighbourhood U of y there exists $t \in \mathbb{T}^+$ such that $\mu_x^t(U) > 0$. Moreover, we have the following:

Proposition. *G_x is the smallest closed forward-invariant set containing x.*

The proof is taken from [New15b, Lemma 1.2.3].

Proof. It is clear that any closed forward-invariant set containing x must contain $\operatorname{supp}\mu_x^t$ for every t, and therefore must contain G_x. So it remains to show that G_x is itself forward-invariant. Fix any $y \in G_x$, and suppose for a contradiction that there exists $t \in \mathbb{T}^+$ such that $\mu_y^t(G_x) < 1$. Since G_x is closed, the map $\xi \mapsto \mu_\xi^t(G_x)$ is upper semicontinuous, and so there exists a neighbourhood V of y such that $\mu_\xi^t(G_x) < 1$ for all $\xi \in V$. Since $y \in G_x$, there exists $s \in \mathbb{T}^+$ such that $\mu_x^s(V) > 0$. Hence

$$\mu_x^{s+t}(X \setminus G_x) \;=\; \int_X \mu_\xi^t(X \setminus G_x)\,\mu_x^s(d\xi) \;\geq\; \int_V \mu_\xi^t(X \setminus G_x)\,\mu_x^s(d\xi) \;>\; 0.$$

But it is clear that $\mu_x^{s+t}(G_x) = 1$, since by definition $\operatorname{supp}\mu_x^{s+t} \subset G_x$. So we have a contradiction. □

We say that a set $G \subset X$ is *minimal* according to a Feller-continuous family of Markov transition probabilities (μ_x^t) if the following equivalent statements hold:

(i) G is closed and forward-invariant, and the only closed forward-invariant proper subset of G is $\varnothing$;

(ii) G is a non-empty closed forward-invariant set, and for all $x \in G$, $G_x = G$;

(iii) G is a non-empty closed forward-invariant set, and for all $x \in G$ and open $U \subset X$ with $U \cap G \neq \varnothing$, there exists $t \in \mathbb{T}^+$ such that $\mu_x^t(U) > 0$.

We say that a set $G \subset X$ is *minimal* according to a Feller-continuous Markov kernel (μ_x) is the following equivalent statements hold:

(i) G is closed and forward-invariant, and the only closed forward-invariant proper subset of G is $\varnothing$;

(ii) G is non-empty, closed, and forward-invariant according to the unique discrete-time family of Markov transition probabilities $(\mu_x^n)_{x \in X,\, n \in \mathbb{N}_0}$ with $(\mu_x^1) = (\mu_x)$.

Note that, given a Feller-continuous family of Markov transition probabilities (μ_x^t), if a closed set $G \subset X$ is forward-invariant according to (μ_x^t) and there exists $\tau \in \mathbb{T}^+$ such that G is minimal according to the kernel $(\mu_x^\tau)_{x \in X}$, then G is minimal according to (μ_x^t).

Given a Feller-continuous kernel (μ_x) or semigroup (μ_x^t), any two distinct minimal sets are mutually disjoint, and every non-empty compact forward-invariant set contains at least one minimal set; the proof is exactly the same as in Section C.1. Once again, the support of any stationary probability measure is forward-invariant. Given a semigroup (μ_x^t) that is both Feller-continuous and measurable,[7] every non-empty compact forward-invariant contains the support of at least one ergodic probability measure, and every compact minimal set is equal to the support of at least one ergodic probability measure. If, in addition, X is compact and (μ_x^t) admits only one stationary probability measure ρ, then $\operatorname{supp}\rho$ is the smallest non-empty closed forward-invariant set.

[7] A sufficient condition for a Feller-continuous family of Markov transition probabilities (μ_x^t) to be measurable is that the map $t \mapsto \mu_x^t$ is right-continuous with respect to the narrow topology for all $x \in X$.

C.5 Discrete-time Markov processes

Let (X,Σ) be a measurable space. Let $(\Omega,\mathcal{F},(\mathcal{F}_n)_{n\in\mathbb{N}_0},\mathbb{P})$ be a filtered probability space. Let $(\mu_x)_{x\in X}$ be a Markov kernel on X.

We say that a sequence $(M_n)_{n\in\mathbb{N}_0}$ of functions $M_n\colon\Omega\to X$ is a (*homogeneous*) *Markov process with respect to the filtration* $(\mathcal{F}_n)$*, with transition probabilities* $(\mu_x)_{x\in X}$ if the following hold:

(i) for each $n\in\mathbb{N}_0$, M_n is $(\mathcal{F}_n,\Sigma)$-measurable;

(ii) for each $n\in\mathbb{N}_0$ and $A\in\Sigma$,

$$\mathbb{P}(M_{n+1}^{-1}(A)|\mathcal{F}_n)(\omega) = \mu_{M_n(\omega)}(A) \quad \mathbb{P}\text{-a.s.}$$

It follows that for any $n\in\mathbb{N}_0$ and any bounded measurable $g\colon X\to\mathbb{R}$,

$$\mathbb{E}[g(M_{n+1})|\mathcal{F}_n](\omega) = \int_X g(x)\,\mu_{M_n(\omega)}(dx) \quad \mathbb{P}\text{-a.s.}$$

(Just approximate g by simple functions, and use the dominated and conditional dominated convergence theorems.)

Remark. Given a sequence $(M_n)_{n\in\mathbb{N}_0}$ of $(\mathcal{F},\Sigma)$-measurable functions $M_n\colon\Omega\to X$, if there exists a filtration of sub-σ-algebras of $\mathcal{F}$ with respect to which (M_n) is a Markov process with transition probabilities (μ_x), then in particular (M_n) must be a Markov process with respect to its natural filtration $\tilde{\mathcal{F}}_n := \sigma(M_r : 0\le r\le n)$, with the same transition probabilities (μ_x). This property can be characterised purely by the *law* of the stochastic process (M_n) (that is, the image measure of $\mathbb{P}$ under the map $\omega\mapsto(M_n(\omega))_{n\in\mathbb{N}_0}$ from Ω to $X^{\mathbb{N}_0}$).[8]

Proposition. *Let* (M_n) *be a Markov process (with respect to any filtration on* $(\Omega,\mathcal{F})$*), with transition probabilities* (μ_x)*. Then for all* $n\in\mathbb{N}_0$*,* $M_{n+1*}\mathbb{P} = \mu^*(M_{n*}\mathbb{P})$*.*

Proof. For any $A\in\Sigma$,

$$\mathbb{P}(M_{n+1}^{-1}(A)) = \int_\Omega \mu_{M_n(\omega)}(A)\,\mathbb{P}(d\omega) = \int_X \mu_x(A)\,M_{n*}\mathbb{P}(dx) = \mu^*(M_{n*}\mathbb{P})(A)$$

as required. □

Now let $(\mu_x^n)_{x\in X,\,n\in\mathbb{N}_0}$ be the unique discrete-time semigroup of Markov kernels with $\mu_x^1 = \mu_x$ for all x.

Proposition. *Let* (M_n) *be a Markov process with respect to the filtration* $(\mathcal{F}_n)$*, with transition probabilities* (μ_x)*. Then for any* $n,r\in\mathbb{N}_0$ *and* $A\in\Sigma$*,*

$$\mathbb{P}(M_{n+r}^{-1}(A)|\mathcal{F}_n)(\omega) = \mu^r_{M_n(\omega)}(A) \quad \mathbb{P}\text{-}a.s.$$

[8] Specifically, (M_n) is a Markov process with respect to its natural filtration if and only if its law is a "Markov measure" as in Section 4 of [New15a] (not to be confused with a "Markov invariant measure" of a RDS, as introduced in Section 3.5 of this thesis).

(Therefore, we may also describe (M_n) as a "Markov process with (n-step) transition probabilities $(\mu_x^n)_{x\in X,\, n\in\mathbb{N}_0}$".)

Proof. We prove the statement by induction on r. The statement is clear for all $n \in \mathbb{N}_0$ with $r = 0$. Now fix $k \in \mathbb{N}_0$ such that the statement is true for all $n \in \mathbb{N}_0$ with $r = k$. For any $n \in \mathbb{N}_0$ and $A \in \Sigma$, we have that for $\mathbb{P}$-almost all ω,

$$\begin{aligned} \mu_{M_n(\omega)}^{k+1}(A) &= \int_X \mu_x^k(A)\, \mu_{M_n(\omega)}(dx) \\ &= \mathbb{E}[\mu_{M_{n+1}}^k(A)|\mathcal{F}_n](\omega) \\ &= \mathbb{E}[\,\mathbb{P}(M_{n+k+1}^{-1}(A)|\mathcal{F}_{n+1})\,|\mathcal{F}_n](\omega) \\ &= \mathbb{P}(M_{n+k+1}^{-1}(A)|\mathcal{F}_n)(\omega). \end{aligned}$$

Thus the statement is true for all $n \in \mathbb{N}_0$ with $r = k+1$. Hence the result follows by induction. □

The ergodic theorem for discrete-time Markov processes

Let ρ be a (μ_x)-stationary probability measure, and let $g\colon X \to \mathbb{R}$ be a ρ-integrable function. Let $(M_n)_{n\in\mathbb{N}_0}$ be a Markov process (with respect to any filtration on $(\Omega,\mathcal{F})$) with transition probabilities (μ_x), and suppose moreover that $M_{0*}\mathbb{P} = \rho$ (from which it follows that $M_{n*}\mathbb{P} = \rho$ for every $n \in \mathbb{N}_0$). Then

$$\frac{1}{n}\sum_{i=0}^{n-1} g(M_i(\omega)) \;\to\; \mathbb{E}[\,g(M_0)\,|\,M_0^{-1}\mathcal{I}_\rho^{(\mu_x)}\,](\omega) \quad \text{as } n \to \infty$$

for $\mathbb{P}$-almost all $\omega \in \Omega$. Hence in particular, if ρ is ergodic then

$$\frac{1}{n}\sum_{i=0}^{n-1} g(M_i(\omega)) \;\to\; \int_X g\, d\rho \quad \text{as } n \to \infty$$

for $\mathbb{P}$-almost all $\omega \in \Omega$.

The above result is obtained by applying Birkhoff's ergodic theorem to the left-shift map on the sequence space $X^{\mathbb{N}_0}$, equipped with the law $\boldsymbol{\mu}_\rho$ of (M_n). The main technicality is to show that every $\boldsymbol{\mu}_\rho$-almost invariant set in this sequence space depends, up to modification, on only the first coordinate. For details, see Section 4 of [New15a].

C.6 Continuous-time Markov processes

Let (X,Σ) be a measurable space. Let $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\in[0,\infty)},\mathbb{P})$ be a filtered probability space, and let $\mathcal{F}_\infty := \sigma(\mathcal{F}_t : t \geq 0)$. Let $(\mu_x^t)_{x\in X,\, t\in[0,\infty)}$ be a semigroup of Markov kernels on X.

We say that a $[0,\infty)$-indexed family $(M_t)_{t\geq 0}$ of functions $M_t\colon \Omega \to X$ is a (*homogeneous*) *Markov process with respect to the filtration* $(\mathcal{F}_t)$, *with transition probabilities* $(\mu_x^t)_{x\in X,\, t\in[0,\infty)}$ if the following hold:

(i) for each $t \geq 0$, M_t is $(\mathcal{F}_t, \Sigma)$-measurable;

(ii) for each $s, t \geq 0$ and $A \in \Sigma$,

$$\mathbb{P}(M_{s+t}^{-1}(A)|\mathcal{F}_s)(\omega) \;=\; \mu^t_{M_s(\omega)}(A) \quad \mathbb{P}\text{-a.s.}$$

Remark. Given a family $(M_t)_{t\geq 0}$ of measurable functions $M_t : \Omega \to X$, if there exists a filtration on $(\Omega, \mathcal{F})$ with respect to which (M_t) is a Markov process with transition probabilities (μ^t_x), then (M_t) must in particular be a Markov process with respect to its natural filtration $\tilde{\mathcal{F}}_t := \sigma(M_s : 0 \leq s \leq t)$, with the same transition probabilities (μ^t_x). (Once again, this property is characterised purely by the law of the stochastic process (M_t), that is, the image measure of $\mathbb{P}$ under the map $\omega \mapsto (M_t(\omega))_{t\geq 0}$ from Ω to $X^{[0,\infty)}$.)

Now for any $t \geq 0$ and any probability measure ρ on X, in keeping with the notation introduced in Section C.3, we define the probability measure $\mu^{t*}\rho$ on X by

$$\mu^{t*}\rho(A) \;:=\; \int_X \mu^t_x(A)\,\rho(dx).$$

As in discrete time, it is easy to show that if (M_t) is a Markov process with transition probabilities (μ^t_x) then for any $s, t \geq 0$, $M_{s+t*}\mathbb{P} = \mu^{t*}(M_{s*}\mathbb{P})$.

The ergodic theorem for continuous-time Markov processes

Let ρ be a (μ^t_x)-stationary probability measure, and let $g : X \to \mathbb{R}$ be a ρ-integrable function. Let $(M_t)_{t\geq 0}$ be a Markov process (with respect to any filtration on $(\Omega, \mathcal{F})$) with transition probabilities (μ^t_x). Suppose moreover that the map $(t, \omega) \mapsto M_t(\omega)$ is jointly measurable, and $M_{0*}\mathbb{P} = \rho$ (from which it follows that $M_{t*}\mathbb{P} = \rho$ for every $t \geq 0$).

Then for $\mathbb{P}$-almost all $\omega \in \Omega$, the map $t \mapsto g(M_t(\omega))$ is locally integrable and

$$\frac{1}{T}\int_0^T g(M_t(\omega))\,dt \;\to\; \mathbb{E}[\,g(M_0)\,|\,M_0^{-1}\mathcal{I}_\rho^{(\mu^t_x)}\,](\omega) \quad \text{as } T \to \infty.$$

In particular, if ρ is ergodic then for $\mathbb{P}$-almost all $\omega \in \Omega$,

$$\frac{1}{T}\int_0^T g(M_t(\omega))\,dt \;\to\; \int_X g\,d\rho \quad \text{as } T \to \infty.$$

A proof in the case that $(M_t)_{t\geq 0}$ has right-continuous sample paths in some separable metrisable topology generating Σ can be found in Section 4 of [New15a]. The more general case is obtained by combining [New15a, Corollary 71] (where it is shown that the almost-invariant sets of the time-shift dynamical system on $X^{[0,\infty)}$ are determined modulo null sets by their 0-coordinate) with the general ergodic theorem for stationary stochastic processes (see e.g. [Lin02, Theorem 5.5][9] with two-sided time replaced by one-sided time and $x(t) := g \circ M_t$).

[9] The statement as appears in [Lin02, Theorem 5.5] requires the additional assumption that the map $(t, \omega) \mapsto x(t)(\omega)$ is jointly measurable. The proof also omits some non-trivial steps: The stationarity of the stochastic process $x_n := \int_{n-1}^n x(t)\,dt$ is justified by [MO14]. (More precisely, this directly covers the case that $x(t)$ is essentially bounded; the general case is then obtained by "capping" $x(t)$ within $[-N, N]$ and letting $N \to \infty$.) The fact that the limit of the finite-time averages is $\mathcal{J}$-measurable modulo null sets relies on this limit being $\mathcal{B}_{\mathbb{R}}$-measurable modulo null sets; this is justified by [MO15a] together with the stationarity of $(x(t))_{t\in\mathbb{R}}$. (Again, this only directly covers the bounded case, but can then be extended to cover the unbounded case.)