UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

Extracting coherent sets in aperiodically driven flows from generators of Mather semigroups

Chemnitz, Robin; Engel, Maximilian; Koltai, Péter

DOI
10.48550/arXiv.2403.19274
10.3934/dcdsb.2024149

Publication date
2025

Document Version
Submitted manuscript

Published in
Discrete and Continuous Dynamical Systems - Series B

License
CC BY-NC-ND

Link to publication

Citation for published version (APA):
Chemnitz, R., Engel, M., & Koltai, P. (2025). Extracting coherent sets in aperiodically driven flows from generators of Mather semigroups. *Discrete and Continuous Dynamical Systems - Series B*, *30*(6), 1952-1995. https://doi.org/10.48550/arXiv.2403.19274, https://doi.org/10.3934/dcdsb.2024149

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:15 Oct 2026

arXiv:2403.19274v1 [math.DS] 28 Mar 2024

Extracting coherent sets in aperiodically driven flows from generators of Mather semigroups

Robin Chemnitz[1], Maximilian Engel[1,2], and Péter Koltai[3]

[1]Freie Universität Berlin
[2]University of Amsterdam
[3]Universität Bayreuth

Abstract

Coherent sets are time-dependent regions in the physical space of nonautonomous flows that exhibit little mixing with their neighborhoods, robustly under small random perturbations of the flow. They thus characterize the global long-term transport behavior of the system. We propose a framework to extract such time-dependent families of coherent sets for nonautonomous systems with an ergodic driving dynamics and (small) Brownian noise in physical space. Our construction involves the assembly and analysis of an operator on functions over the augmented space of the associated skew product that, for each fixed state of the driving, propagates distributions on the corresponding physical-space fibre according to the dynamics. This time-dependent operator has the structure of a semigroup (it is called the Mather semigroup), and we show that a spectral analysis of its generator allows for a trajectory-free computation of coherent families, simultaneously for all states of the driving. Additionally, for quasi-periodically driven torus flows, we propose a tailored Fourier discretization scheme for this generator and demonstrate our method by means of three examples of two-dimensional flows.

1 Introduction

Understanding the transport and mixing behavior of complicated nonautonomous flows is still an outstanding problem on the interface of dynamical systems theory and its many applications. We will consider in this work the problem of finding time-evolving families of sets in the state space of the dynamics that mix little with their surroundings. This way we provide a macroscopic view on prevalent transport-related characteristics of aperiodically driven flows. In the classical theory of deterministic dynamical systems, invariant manifolds organize the state space in terms of transport: they form impenetrable barriers to purely advective transport. For nonautonomous systems, this led to the concept of lobe dynamics [RKW90], of general transport barriers such as Lagrangian Coherent Structures [HY00, SLM05, RBBV+07], and related objects [HBV13, MB14, HHFH16, BOR18].

1.1 Background

Small noise, metastability, and coherence. In a physically motivated modeling step, small random perturbations can be added to the (otherwise deterministic) state evolution, while the questions remain: What are the characteristic state-space objects governing mixing and the typical timescales associated to them? In a natural approach to this question, subsets of the state space can be sought that mix particularly little with their surroundings. For autonomous systems, such sets are called *almost-invariant* [DJ99] or *metastable* [Dav82a, Dav82b, SS13, BDH16]. Their presence implies the existence of persistent patterns, such as slowly decaying concentration fields of quantities passively transported by the flow [LH04]. The corresponding objects in nonautonomous systems are called *coherent sets*, and this term loosely refers to time-varying sets that have little dynamical interaction with their neighborhood. The concept has initially been developed in a finite-time setting for systems subject to noise [FSM10, Fro13, FPG14]; it has later been generalized separately to deterministic dynamics [Fro15] and to the periodically-forced infinite-time setting [FK17]. We note that the deterministic and noisy notions can be linked in the vanishing-diffusion limit [Fro15, KK20, SKJ21].

Escape rates. If finite-time considerations are replaced by infinite-time ones, the notions of metastability and coherence need to be adapted as well. *Escape rates* become a natural measure of dynamical persistence of a set, and the notion has been extensively studied for autonomous *open systems* [PY79, CMT98, LMD03, DY05, DT17, DWY12, FS10, BY11, BFGTM14] and for time-homogeneous (Markovian) systems alike [CMS97, FJK13] The nonautonomous generalization of escape rates from a family of time-dependent sets was considered in [Bal14] for deterministic flows in terms of time-varying flux. It was used to extract coherent sets in [FK17, FKS20] for specific driving dynamics, and has appeared in the random dynamical systems literature for open systems [AFGTV23] to measure persistence.

Practical characterization and computation: Transfer operators. In computational approaches to metastability and coherence a functional view is prevailing. Sets are relaxed to (signed) distributions and these in turn are evolved by *transfer operators* associated with the dynamics. Metastable sets of autonomous systems can be connected to the dominant eigenspectrum of transfer operators [Dav82a, Dav82b, DJ99, DHFS00, BEGK02, DW05, HS06, SS13], while finite-time coherent sets can be extracted from a singular value decomposition of associated Markov operators [FSM10, Fro13, FPG14, DJM16]. Intuitively, the eigen-(singular) values here are commensurate with the metastability (coherence) of the associated sets: The closer these values are to 1, the less dynamical leakage do the corresponding metastable (coherent) sets exhibit.

The connection to escape rates arises if we consider instead of one-step quantities (such as eigen- or singular values) of transfer operators an iterated quantity: *decay rates*. Naturally, for autonomous systems the decay of a function is governed by its decomposition into eigenfunctions of the transfer operator (assuming it has one). However, nonautonomous systems require a generalization of this concept, which is predominantly formalized in the framework of transfer operator *cocycles* and multiplicative ergodic theory, where the spectra and eigenfunctions are

replaced by Lyapunov spectra (exponents) and associated Lyapunov filtrations (in special cases, Oseledets splittings) [FLS10, FLQ10, GT18]. While our considerations below are intimately related to Lyapunov spectra and Oseledets splittings, we defer rigorous elaborations on this connection to future studies. The purpose of the present study is to highlight other aspects. A description of linear stability in nonautonomous dynamical systems that more broadly mimics the autonomous distinction into stable and unstable directions is given by the Sacker–Sell (or dichotomy) spectrum [SS78] which typically contains the Lyapunov spectrum and allows for constructions of dynamically relevant subsets such as invariant manifolds [Pöt12, PR16]. We will make use of this spectrum within this work.

Augmented space. The same way as time-augmentation transforms a nonautonomous system into an equivalent autonomous one, we ask whether metastability as a persistence concept for the skew-product associated to a nonautonomous system reveals coherent sets of the latter. This research thread was initiated in [FK17] for periodic driving and in [FKS20] for finite-time systems, and here we consider the natural next step, where a general ergodic flow is driving a stochastic differential equation (SDE) on physical space. We set up the transfer operator of this skew-product system, a so-called Mather (evolution) semigroup [CL99, Section 6.2], and use its spectrum to extract coherent sets. More precisely, we consider the generator of this semigroup, since this has advantages over the numerical approximation of the entire semigroup—as we elaborate below. A phenomenologically "dual" construction in [GD20] considers the evolution semigroup of Koopman operators and extracts so-called "coherent patterns", with the crucial difference to our construction that they apply diffusion both in the driving and in physical space. The spectral properties of evolution semigroups (also called Howland semigroup [How74] if the parameter space is the real line representing time) and Mather semigroups in relation to the dynamical properties of the driving system (the "base") are comprehensively elaborated in [CL99].

1.2 Contributions

The setting. To describe our approach and contributions more precisely, we consider a nonautonomous SDE on the physical space $M \subset \mathbb{R}^d$ (or $M = \mathbb{T}^d$) with additive noise (and reflecting boundary conditions on ∂M, unless $M = \mathbb{T}^d$) that is driven by some base dynamics on a parameter space Θ,

$$\begin{aligned} d\theta_t &= \Psi(\theta_t)dt, \\ dx_t &= v(\theta_t, x_t)dt + \varepsilon dw_t, \end{aligned} \tag{1}$$

i.e., $x_t \in M$ and $\theta_t \in \Theta$. Here, Θ is a compact smooth manifold equipped with a continuous vector field Ψ. We assume that the unique solution to the ODE on Θ governed by Ψ is an invertible ergodic flow $\phi^t : \Theta \to \Theta$ and that $v : \Theta \times M \to \mathbb{R}^d$ is a smooth, divergence-free vector field.

Ensembles of states in M, which will be generalized to be viewed as functions $f \in L^2(M; \mathbb{C})$, are evolved by (1) as described by the transfer operator cocycle $\mathcal{P}_\theta^t$, acting on $L^2(M; \mathbb{C})$ equipped with the usual norm $\|\cdot\|_2$. This means, if $\theta_0 = \theta$ and the random physical-space initial condition

is f-distributed (we write $x_0 \sim f$), then $x_t \sim \mathcal{P}_\theta^t f$. We note that $\mathcal{P}_\theta^t$ is the solution operator of the Fokker–Planck (or advection-diffusion) equation associated with the SDE for x_t. The transfer operators satisfy the cocycle identity $\mathcal{P}_{\phi^s\theta}^t \circ \mathcal{P}_\theta^s = \mathcal{P}_\theta^t$.

In general, nonautonomous transfer operators depend on an initial time s and a time duration t. For a cocycle, the initial time is replaced by the parameter $\theta \in \Theta$ describing the state of the driving system at initial time. Thus, Θ can be viewed as a compactification of the set $\mathbb{R}_{\geq 0}$ of initial times s. This compactification allows for the study of asymptotics as $t \to \infty$ without (numerically) emulating this limit. Such asymptotics, like the exponential decay rate of $\|\mathcal{P}_\theta^t f\|_2$ for a distribution $f \in L^2(M, \mathbb{C})$, are commonly studied using multiplicative ergodic theory. In this work, we take a different approach and study the long-term behavior of the transfer operator cocycle by constructing an augmented generator that captures the nonautonomous dynamics in a single time-independent operator. An analogous approach has previously been applied to study periodically driven flows in [FK17]. The extension of this approach to aperiodically driven flows introduces significant complexity to the analysis and requires a different theoretical framework which we will elaborate on in the following.

Our results. Our central object of study is the linear operator $\mathbf{M}^t$ acting on a function space $\mathfrak{F}$, which is either $L^2(\Theta, L^2(M, \mathbb{C}))$ or $C(\Theta, L^2(M, \mathbb{C}))$, describing the evolution of some $\mathbf{f} \in \mathfrak{F}$ fibre-wise according to $\mathcal{P}$:

$$\big[\mathbf{M}^t \mathbf{f}\big](\theta) = \mathcal{P}_{\phi^{-t}\theta}^t \mathbf{f}(\phi^{-t}\theta).$$

In Section 2.1, we introduce a set of assumptions (I)–(IV) for the transfer operators $\mathcal{P}_\theta^t$, guaranteeing that $(\mathbf{M}^t)_{t\geq 0}$ is a strongly continuous semigroup, which is called the *Mather semigroup*. The study of the Mather semigroup and its generator, in particular their spectrum and spectral subspaces, provides insight into the long-term behavior of the transfer operator cocycle $\mathcal{P}$. We establish a spectral mapping relation in Theorem 2.6. In Theorem 3.5, we ensure that the operators $\mathcal{P}_\theta^t$ of the Fokker–Planck equation related to (1), indeed, form a well-defined linear cocycle satisfying assumptions (I)–(IV).

The purpose of the current work is to study in theory and by several numerical examples how spectral objects of $\mathbf{M}^t$ and its generator $\mathbf{G}$ give rise to families of coherent sets $A_\theta^t \subset M$, $\theta \in \Theta$, $t \geq 0$. To measure coherence of such a family, we will consider two measures. The first is the escape rate, which is an asymptotic quantity related to the tail behavior of the survival probabilities $\mathbb{P}(x_s \in A_\theta^s,\ \forall s \in [0, t])$ for $t \to \infty$. The second is the cumulative survival probability, which is an integral quantity of the survival probabilities, and, in particular, includes its finite-time characteristics.

The spectral objects of $\mathbf{M}^t$ which we extract coherent sets from are eigenfunctions, approximate eigenfunctions and spectral subspaces. We provide rigorous bounds on the measure of coherence of the resulting family A_θ^t for eigenfuctions (Theorem 4.2) and approximate eigenfunctions (Proposition 4.4) and additionally derive heuristic methods that can be applied in numerical computations.

To illustrate the central idea of how we construct coherent sets, assume that $\mathbf{f}$ is an eigenfunction of the Mather semigroup to a real eigenvalue $\lambda \in \mathbb{R}$, i.e. $\mathcal{P}_\theta^t \mathbf{f}(\theta) = e^{\lambda t}\mathbf{f}(\phi^t\theta)$ for all $t \geq 0$ and $\theta \in \Theta$. Define the family of sets $A_\theta^t := \{\mathbf{f}(\phi^t\theta) \geq 0\} \subset M$ which is the non-negative

support of the fibres of $\mathbf{f}$. Since the transfer operator $\mathcal{P}_\theta^t$ maps $\mathbf{f}(\theta)$ to (a multiple of) $\mathbf{f}(\phi^t\theta)$, particles that were initialized in the set A_θ^0 have a high probability to lie inside A_θ^t after time t. A schematic representation of this setting is found in Figure 1.

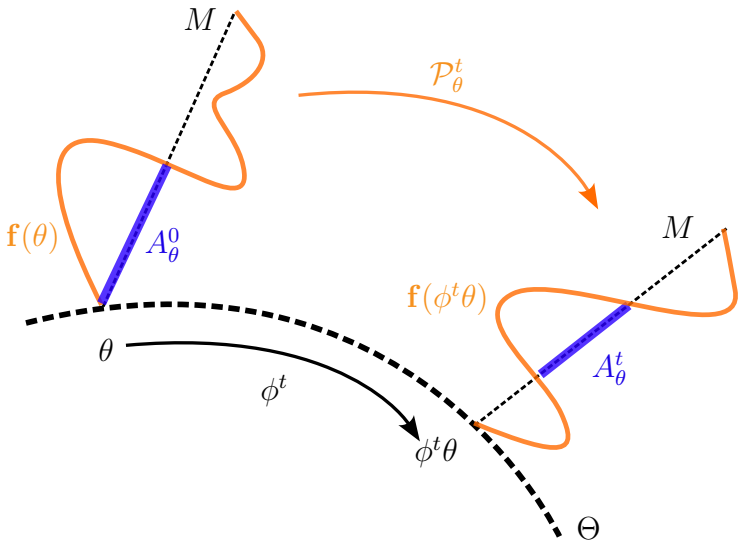

Figure 1: Sketch of the temporal evolution of the coherent set $A_\theta^t = \{\mathbf{f}(\phi^t\theta) \geq 0\} \subset M$, where $\mathbf{f}$ is an eigenfunction of the Mather semigroup with real eigenvalue. The distribution $\mathbf{f}(\theta)$ is mapped to (a multiple of) $\mathbf{f}(\phi^t\theta)$ under the cocycle $\mathcal{P}_\theta^t$ over the driving $\phi^t : \Theta \to \Theta$.

To evaluate the applicability of our methods, we numerically compute coherent sets and measure their coherence for three example dynamics: translated gyres, oscillating gyres and alternating shears. We restrict our considerations to the case where both the parameter space Θ and the physical space M are given by tori $\mathbb{T}^d$. In order to extract coherent sets from respective eigenfuncions, we discretize the Mather semigroup and its generator $\mathbf{G}$ by Galerkin projection onto selected Fourier modes. Based on the infinite-dimensional matrix representation of the augmented generator $\mathbf{G}$ in Proposition 5.2, we introduce an informed selection process to boost accuracy even in our numerically challenging multidimensional setting, where functions over $\Theta \times M$ need to be discretized. For simplicity, the driving dynamics on Θ is assumed to be a quasi-periodic torus rotation. The extension of our methods to general domains and driving dynamics is discussed in the outlook. For all three examples, we compute coherent sets whose survival probabilities decay close to the theoretical bounds. An implementation of our method in Matlab, including all examples presented below, is available under

`https://github.com/RobinChemnitz/MatherCoherent`

Outline. The rest of the paper is structured as follows. Section 2 formally introduces the objects which we will work with in the later sections; namely linear cocycles, their Sacker–Sell spectrum, and the Mather semigroup. In Section 3, we study the SDE (1) through its

transfer operators $\mathcal{P}_\theta^t$. The central result of this section is that the operators $\mathcal{P}_\theta^t$ form a linear cocycle that enables the construction of the associated Mather semigroup and its generator. In Section 4, we introduce coherent sets and derive methods to extract coherent sets from spectral objects of the Mather semigroup and its generator. In Section 5, we represent the augmented generator $\mathbf{G}$ in Fourier coordinates and construct a discretized generator using a Galerkin approximation onto carefully selected Fourier modes. We extract coherent sets from eigenfunctions of the discretized generator and test their coherence experimentally for three examples in Section 6. We conclude with a discussion of our results as well as an outlook.

2 Preliminaries

2.1 Continuous-time cocycles

In this section, we introduce the notion of a linear cocycle, its Sacker–Sell spectrum, and the Mather semigroup in a general setting.

Let Θ be a compact metric space equipped with its Borel σ-algebra and let $\mathcal{H}$ be a Hilbert space. We denote the space of bounded linear operators on $\mathcal{H}$ by $\mathcal{L}(\mathcal{H})$. Let $\phi^t : \Theta \to \Theta$, where $t \in \mathbb{R}$, be a continuous and invertible flow on Θ. That is, $(\phi^t)_{t\in\mathbb{R}}$ is a one-parameter group on Θ such that the map $(\theta, t) \mapsto \phi^t\theta$ is continuous. Let μ be an ergodic measure of ϕ, that is $\mu(A) \in \{0, 1\}$ for any set $A \subset \Theta$ that is ϕ^t-invariant for all $t \geq 0$. Since Θ is compact at least one such measure exists, which can be shown using the Banach–Alaoglu theorem. W.l.o.g. we assume $\operatorname{supp}(\mu) = \Theta$, i.e. any open set $U \subset \Theta$ has $\mu(U) > 0$. There are two cases that can occur. Either $\Theta \cong S^1$ consists of a single periodic orbit, or the set of aperiodic points with respect to ϕ is dense in Θ. Throughout, we assume that the set of aperiodic points is dense in Θ.

A linear cocycle in $\mathcal{H}$ with driving ϕ is a map

$$\begin{aligned} \Phi : \Theta \times \mathbb{R}_{\geq 0} &\to \mathcal{L}(\mathcal{H}) \\ (\theta, t) &\mapsto \Phi_\theta^t, \end{aligned}$$

with the following defining cocycle properties:

(i) $\Phi_\theta^0 = \mathrm{Id}, \quad \forall \theta \in \Theta;$

(ii) $\Phi_\theta^{s+t} = \Phi_{\phi^s\theta}^t \circ \Phi_\theta^s, \quad \forall \theta \in \Theta,\ \forall s, t \in \mathbb{R}_{\geq 0}.$

We introduce additional assumptions which we impose on (and verify for specific) continuous-time cocycles Φ throughout this work.

(I) *compact.* The operators $\Phi_\theta^t \in \mathcal{L}(\mathcal{H})$ are compact $\forall \theta \in \Theta,\ t > 0$;

(II) *norm-continuous in θ.* For each $t \geq 0$, the map $\theta \mapsto \Phi_\theta^t \in \mathcal{L}(\mathcal{H})$ is norm-continuous;

(III) *strongly-continuous.* For each $x \in \mathcal{H}$, the map $(\theta, t) \mapsto \Phi_\theta^t x \in \mathcal{H}$ is continuous.

(IV) *exponentially bounded.* There are constants $K > 0$ and $L > 0$ such that $\|\Phi_\theta^t\| \leq Ke^{Lt}$, $\forall t \geq 0$.

Assumptions (I) and (II) are necessary to obtain a nice representation of the Sacker–Sell spectrum, cf. Proposition 2.4. Assumptions (III) and (IV) are needed to define the Mather semigroup, cf. Section 2.3.

2.2 The Sacker–Sell spectrum

This section introduces the Sacker–Sell spectrum, also called dichotomy spectrum, of a continuous-time linear cocycle Φ satisfying conditions (I)–(IV). For the most part, we follow [CL99, Chapter 6.1].

Definition 2.1. *We say that a cocycle Φ has an* exponential dichotomy at $\lambda \in \mathbb{R}$ *if there are constants $\beta > 0$ and $C > 0$ and a strongly continuous, projection-valued function $\Pi : \Theta \to \mathcal{L}(\mathcal{H})$ with complemented subspaces*

$$S(\theta) := \operatorname{ran}(\Pi(\theta)), \qquad U(\theta) := \ker(\Pi(\theta)),$$

that together satisfy the following properties for every $\theta \in \Theta$ and $t \geq 0$:

(i) $\Pi(\phi^t\theta)\Phi_\theta^t = \Phi_\theta^t\Pi(\theta)$;

(ii) The restriction $\Phi_\theta^t|_U : U(\theta) \to U(\phi^t\theta)$ is invertible;

(iii) $\|\Phi_\theta^t|_S\| \leq Ce^{(\lambda-\beta)t}$, $\qquad \left\| \left(\Phi_\theta^t|_U\right)^{-1} \right\|^{-1} \geq C^{-1}e^{(\lambda+\beta)t}$,
where $\Phi_\theta^t|_S$ is defined analogously to $\Phi_\theta^t|_U$.

The letters S and U stand for the stable and unstable bundles, respectively.

Remark 2.2. One can verify that the first property is equivalent to $\Phi_\theta^t S(\theta) \subset S(\phi^t\theta)$ and $\Phi_\theta^t U(\theta) \subset U(\phi^t\theta)$. Since $\Phi_\theta^t|_U$ is assumed to be invertible and thereby surjective, we can even conclude $\Phi_\theta^t U(\theta) = U(\phi^t\theta)$. The third property implies $\|\Phi_\theta^t u\| \geq C^{-1}e^{(\lambda+\beta)t}$ for all $u \in U(\theta)$ and $t \geq 0$. If $U(\theta)$ was infinite-dimensional, this would imply that Φ_θ^t is not compact. Hence, for any cocycle Φ satisfying the compactness assumption (I), that admits an exponential dichotomy, the unstable bundle $U(\theta)$ is finite-dimensional.

Definition 2.3. *The Sacker–Sell spectrum $\Sigma(\Phi)$ is defined as*

$$\Sigma(\Phi) := \{\lambda \in \mathbb{R} \mid \Phi \text{ does not have an exponential dichotomy at } \lambda\}.$$

By definition, $\Sigma(\Phi) \subset \mathbb{R}$ is a closed set. For cocycles Φ satisfying assumptions (I)–(IV), the Sacker–Sell spectrum is of the following form. For a proof, see [CL99, Theorem 8.12].

Proposition 2.4. *The Sacker–Sell spectrum $\Sigma(\Phi)$ consists of a finite or countably infinite number of closed segments*

$$\Sigma(\Phi) = \bigcup_{k=1}^{N} [r_k^-, r_k^+],$$

where $N \in \mathbb{N}_0 \cup \{\infty\}$ and $-\infty \leq \ldots < r_2^- \leq r_2^+ < r_1^- \leq r_1^+$.

Assume $\dim(\mathcal{H}) = \infty$. If $N < \infty$, then $r_N^- = -\infty$. If $N = \infty$, we find

$$\lim_{k\to\infty} r_k^{\pm} = -\infty.$$

Note that $r_k^- = r_k^+$ is allowed, such that the segment $[r_k^-, r_k^+]$ consists of a single point. In general, the cases $\Sigma(\Phi) = \emptyset$ and $\Sigma(\Phi) = (-\infty, r_1^+]$ can occur. In Section 2.5 we further specify the structure of the Sacker–Sell spectrum in the special case where Φ is autonomous, i.e. the operators Φ_θ^t do not depend on θ.

Remark 2.5. In the literature, e.g. [CL99], exponential dichotomies are commonly only defined at $\lambda = 0$. A cocycle Φ has an exponential dichotomy at λ if and only if the cocycle defined by $\Psi_\theta^t := e^{-\lambda t}\Phi_\theta^t$ has an exponential dichotomy at 0. This type of rescaling argument is used on multiple occasions throughout this work.

2.3 The Mather semigroup

In this section we consider the evolution of elements under a cocycle Φ for all $\theta \in \Theta$ simultaneously. Let $\mathfrak{F}$ be either the Bochner space $L^2(\Theta, \mathcal{H})$ or $C(\Theta, \mathcal{H})$, where the latter is the space of continuous functions equipped with the supremum norm $\|\cdot\|_\infty$. Functions in $\mathfrak{F}$ are denoted by bold letters. For each $t \geq 0$, define the linear operator $\mathbf{M}^t : \mathfrak{F} \to \mathfrak{F}$ by

$$\big[\mathbf{M}^t \mathbf{f}\big](\theta) = \Phi_{\phi^{-t}\theta}^t \mathbf{f}(\phi^{-t}\theta). \tag{2}$$

The operators $(\mathbf{M}^t)_{t\geq 0}$ satisfy the semigroup property, i.e. $\mathbf{M}^t \circ \mathbf{M}^s = \mathbf{M}^{s+t}$, for all $s, t \geq 0$. The semigroup $(\mathbf{M}^t)_{t\geq 0}$ is called the *Mather semigroup*. The action of the Mather semigroup can be interpreted as evolving $\mathbf{f}$ fibre-wise through the cocycle Φ. Hence, the θ-fibre of $\mathbf{f}$ is mapped to the $\phi^t\theta$-fibre of $\mathbf{M}^t\mathbf{f}$. Different fibres of $\mathbf{f}$ do not interact.

In this section, we state our results for the Mather semigroup constructed over either $L^2(\Theta, \mathcal{H})$ or $C(\Theta, \mathcal{H})$. In Section 5, we discretize the operators $\mathbf{M}^t$ using a Galerkin projection onto Fourier modes, which requires us to consider the Mather semigroup over the Hilbert space $L^2(\Theta, \mathcal{H})$. In Section 4, we need higher regularity than L^2, for which considering $C(\Theta, \mathcal{H})$ is necessary. We note that the chosen function space only plays a role in the theoretical study, since a computational implementation of our methods is unable to distinguish continuous and measurable functions. For the study of the Mather semigroup over different function spaces in a more general setting, see [BV19].

By [CL99, Theorem 6.20 and Lemma 6.33], $(\mathbf{M}^t)_{t\in\mathbb{R}_{\geq 0}}$ is a strongly continuous semigroup, given that Φ is strongly continuous and exponentially bounded, which is satisfied with assumptioins (III) and (IV). Hence, the Mather semigroup has a generator defined pointwise by

$$\mathbf{G} = \lim_{t\to 0} \frac{\mathbf{M}^t - \mathrm{Id}}{t}.$$

The (potentially unbounded) operator $\mathbf{G}$ is closed and has a dense domain $\mathcal{D}(\mathbf{G}) \subset \mathfrak{F}$, see e.g. [Paz12, Corollary 2.5]. We call $\mathbf{G}$ the *augmented generator*. The following theorem characterizes the spectrum of $\mathbf{M}^t$ and its generator $\mathbf{G}$ in terms of the Sacker–Sell spectrum $\Sigma(\Phi)$. A schematic representation of the spectra $\sigma(\mathbf{G})$ and $\sigma(\mathbf{M}^t)$ is given in Figure 2. We note that even when the spectrum of $\mathbf{M}^t$ and $\mathbf{G}$ is non-empty, their point spectrum may be empty, i.e. they might not possess eigenfunctions.

Theorem 2.6. *The spectra of $\mathbf{M}^t$, where $t > 0$, and $\mathbf{G}$ are given by*

$$\sigma(\mathbf{M}^t) \setminus \{0\} = \{e^{\lambda t + \eta i} \mid \lambda \in \Sigma(\Phi),\ \eta \in [0, 2\pi)\},$$
$$\sigma(\mathbf{G}) = \{\lambda + \eta i \mid \lambda \in \Sigma(\Phi),\ \eta \in \mathbb{R}\}.$$

In other words, $\sigma(\mathbf{M}^t)$ consists of annuli around the center, and $\sigma(\mathbf{G})$ consists of bands parallel to the imaginary axis.

Proof. The spectral mapping theorem [CL99, Theorem 6.30, Theorem 6.37] states that for all $t > 0$ we find

$$\sigma(\mathbf{M}^t) \setminus \{0\} = e^{t\sigma(\mathbf{G})}. \tag{3}$$

Additionally, it states that $\sigma(\mathbf{M}^t)$ is invariant under rotations around the center and that $\sigma(\mathbf{G})$ is invariant under translations by imaginary values. The cocycle Φ has an exponential dichotomy at 0 if and only if $\sigma(\mathbf{M}^t) \cap S^1 = \emptyset$. For $\mathfrak{F} = L^2(\Theta, \mathcal{H})$, this follows from [LS91, Assertion 1.4]. For $\mathfrak{F} = C(\Theta, \mathcal{H})$, the statement is shown in [CL99, Theorem 6.41]. By rescaling, the cocycle Φ has an exponential dichotomy at $\lambda \in \mathbb{R}$ if and only if $\sigma(\mathbf{M}^t) \cap e^{\lambda t} S^1 = \emptyset$. Since the Sacker–Sell spectrum $\Sigma(\Phi)$ consists exactly of those $\lambda \in \mathbb{R}$ for which Φ does not have an exponential dichotomy, we conclude

$$\sigma(\mathbf{M}^t) \setminus \{0\} = \{e^{\lambda t + \eta i} \mid \lambda \in \Sigma(\Phi),\ \eta \in [0, 2\pi)\}.$$

The characteriziation of the spectrum of $\mathbf{G}$ follows directly from (3) together with the fact that $\sigma(\mathbf{G})$ is invariant under translations by imaginary values. □

2.4 Spectral projections

Theorem 2.6 provides a direct link between the spectrum of the Mather semigroup $\mathbf{M}^t$ and the Sacker–Sell spectrum $\Sigma(\Phi)$. In this section, we show that this connection extends to the level of the corresponding subspaces, in the sense that the spectral projections of $\mathbf{M}^t$ can be described in terms of the projection-valued functions that make up the exponential dichotomies of Φ, cf. (4). In the following, we give a brief introduction to spectral projections.

In a finite-dimensional vector space V, each eigenvalue λ of a linear operator $T : V \to V$ has a corresponding generalized eigenspace E_λ. These generalized eigenspaces decompose the vector space V, and the spectrum of the restriction $T|_{E_\lambda}$ is the set $\{\lambda\}$. For an infinite-dimensional Banach space $\mathbb{B}$ and a closed linear operator $T : \mathbb{B} \to \mathbb{B}$, a similar decomposition is obtained using Riesz-projections [RN12], also called *spectral projections*. For a more recent introduction,

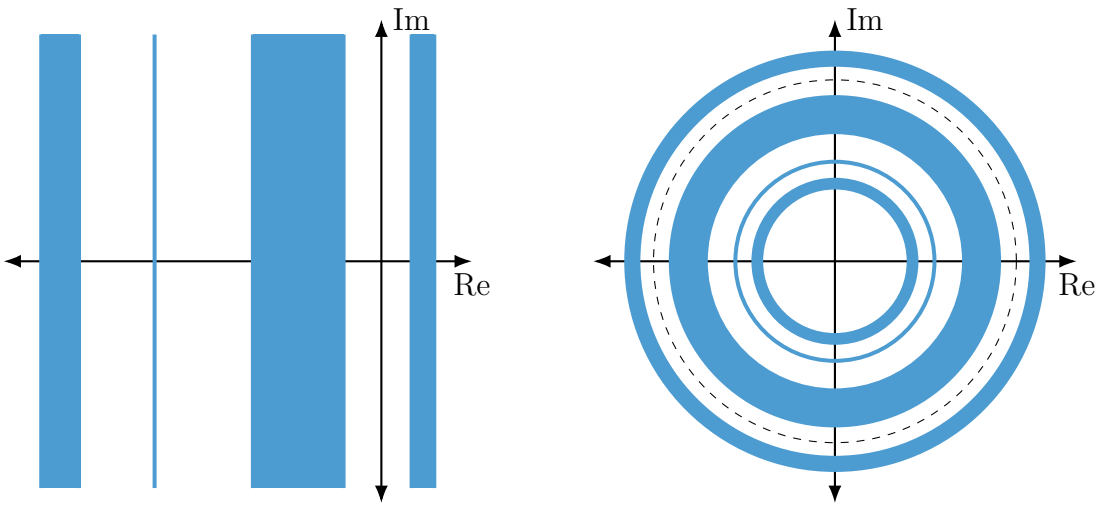

Figure 2: A schematic representation of the spectrum of the generator $\mathbf{G}$ (left), and the spectrum of the Mather operator $\mathbf{M}^1$ (right). The unit circle S^1 is marked as a dashed circle. In this example, the Sacker–Sell spectrum consists of four intervals, one of which is a single point.

we refer to [vN22, Chapter 6]. Let $\Lambda \subset \sigma(T)$ be an isolated part of the spectrum of T. We call Λ a *spectral set*. Consider a closed contour $\Gamma \subset \rho(T)$ around Λ with positive orientation. The Riesz-projection $\mathfrak{P}_\Lambda : \mathbb{B} \to \mathbb{B}$ is defined by

$$\mathfrak{P}_\Lambda = -\frac{1}{2\pi i} \oint_\Gamma (T - z)^{-1} dz.$$

The linear operator $\mathfrak{P}_\Lambda : \mathbb{B} \to \mathbb{B}$ is a bounded projection that splits the Banach space into its range $\operatorname{ran}(\mathfrak{P}_\Lambda) \subset \mathbb{B}$ and its kernel $\ker(\mathfrak{P}_\Lambda) \subset \mathbb{B}$. The subspaces $\operatorname{ran}(\mathfrak{P}_\Lambda)$ and $\ker(\mathfrak{P}_\Lambda)$ are invariant under T and the spectrum of their restrictions is given by $\sigma(T|_{\operatorname{ran}(\mathfrak{P}_\Lambda)}) = \Lambda$ and $\sigma(T|_{\ker(\mathfrak{P}_\Lambda)}) = \sigma(T) \setminus \Lambda$. Hence, the subspace $\operatorname{ran}(\mathfrak{P}_\Lambda)$ can be considered as an infinite-dimensional version of the generalized eigenspace of the spectral set $\Lambda \subset \sigma(T)$. We call $\operatorname{ran}(\mathfrak{P}_\Lambda) \subset \mathbb{B}$ the *spectral subspace* of Λ.

We use Riesz-projections to describe a connection between the spectrum of the Mather semigroup $\mathbf{M}^t$ and the stable/unstable bundles of the cocycle Φ. Let $[r^-, r^+]$ be an isolated segment of the Sacker–Sell spectrum $\Sigma(\Phi)$, and let $\gamma^- < r^- \leq r^+ < \gamma^+$ be such that $[\gamma^-, \gamma^+] \cap \Sigma(\Phi) = [r^-, r^+]$. By Theorem 2.6, the annulus $\Lambda^t := \{e^{\lambda t + \eta i} \mid \lambda \in [r^-, r^+],\ \eta \in [0, 2\pi)\}$ is an isolated component of the spectrum of $\mathbf{M}^t$, and the rings of radius $e^{\gamma^- t}$ and $e^{\gamma^+ t}$ separate Λ^t from the rest of the spectrum of $\mathbf{M}^t$. Let $\mathbb{D} \subset \mathbb{C}$ be the unit disk. The Riesz-projection onto the spectral sets $\sigma(\mathbf{M}^t) \cap e^{\gamma^- t}\mathbb{D}$, respectively $\sigma(\mathbf{M}^t) \cap e^{\gamma^+ t}\mathbb{D}$, are given by

$$\mathfrak{P}_{e^{\gamma^- t}} := -\frac{1}{2\pi i} \oint_{e^{\gamma^- t} S^1} (\mathbf{M}^t - z)^{-1} dz, \quad \mathfrak{P}_{e^{\gamma^+ t}} := -\frac{1}{2\pi i} \oint_{e^{\gamma^+ t} S^1} (\mathbf{M}^t - z)^{-1} dz.$$

In particular, the projection onto the spectral set Λ^t is given by

$$\mathfrak{P}_{\Lambda^t} = \mathfrak{P}_{e^{\gamma^+ t}} - \mathfrak{P}_{e^{\gamma^- t}}.$$

Since γ^- and γ^+ lie outside of $\Sigma(\Phi)$, the cocycle Φ has exponential dichotomies at γ^- and γ^+. Therefore, there are strongly continuous projections $\Pi_-(\theta)$, $\Pi_+(\theta)$ that split the space $\mathcal{H}$ into the respective stable and unstable bundles $S_\pm(\theta) := \operatorname{ran}(\Pi_\pm(\theta))$ and $U_\pm(\theta) := \ker(\Pi_\pm(\theta))$. Note that $S_-(\theta) \subset S_+(\theta)$ and $U_+(\theta) \subset U_-(\theta)$. In particular, for each $\theta \in \Theta$, we obtain a splitting $\mathcal{H} = S(\theta) \oplus E(\theta) \oplus U(\theta)$, where $S(\theta) := S_-(\theta)$, $E(\theta) := S_+(\theta) \cap U_-(\theta)$, and $U(\theta) := U_+(\theta)$. The operator $\Pi(\theta) := \Pi_+(\theta) - \Pi_-(\theta)$ is a projection onto $E(\theta)$ along $S(\theta) \oplus U(\theta)$. Vectors $s \in S(\theta)$ decay uniformly faster than the rate γ^- under Φ_θ^t, vectors $u \in U(\theta)$ grow uniformly faster than rate the γ^+, and vectors $x \in E(\theta)$ decay/grow uniformly at rate between γ^- and γ^+.

The connection between the spectral projection $\mathfrak{P}_{\Lambda^t}$ and the dichotomy projections $\Pi_\pm(\theta)$ is of the following form: The projections $\mathfrak{P}_{e^{\gamma^\pm t}}$ are given by

$$[\mathfrak{P}_{e^{\gamma^\pm t}}\mathbf{f}](\theta) = \Pi_\pm(\theta)\mathbf{f}(\theta).$$

For $\mathfrak{F} = L^2(\Theta, \mathcal{H})$, this is shown in [LS91, Assertion 1.4]. For $\mathfrak{F} = C(\Theta, \mathcal{H})$, the statement is found in [CL99, Theorem 6.38]. We conclude that the projection $\mathfrak{P}_{\Lambda^t}$ is given by

$$[\mathfrak{P}_{\Lambda^t}\mathbf{f}](\theta) = \Pi(\theta)\mathbf{f}(\theta).$$

Therefore, the spectral subspace of Λ^t is given by

$$\operatorname{ran}(\mathfrak{P}_{\Lambda^t}) = \{\mathbf{f} \in \mathfrak{F} \mid \mathbf{f}(\theta) \in E(\theta),\ \forall \theta \in \Theta\}. \tag{4}$$

In the case $\mathfrak{F} = L^2(\Theta, \mathcal{H})$, functions $\mathbf{f} \in \mathfrak{F}$ are only defined up to μ-null sets. In that case, (4) is to be understood as an μ-a.e. statement. We conclude that the spectral subspace of an annulus Λ^t is made up of functions $\mathbf{f} \in \mathfrak{F}$ whose fibres decay/grow uniformly at rate between r_- and r_+.

2.5 The autonomous case

In this section we study the special case where the operators Φ_θ^t do not depend on $\theta \in \Theta$, but only on $t \geq 0$. We will encounter such cocycles in examples of Section 6. Since Φ_θ^t does not depend on θ, we write $\Phi^t \in \mathcal{L}(\mathcal{H})$ for the time-t-operator. By the cocycle property, $(\Phi^t)_{t\geq 0}$ is a one-parameter semigroup, i.e. $\Phi^0 = \mathrm{Id}$, and for all $s, t \geq 0$ we find $\Phi^{s+t} = \Phi^t \circ \Phi^s$. Since Φ is a cocycle satifying assumptions (I)–(IV), $(\Phi^t)_{t\geq 0}$ is a strongly continuous, exponentially bounded semigroup that is compact for $t > 0$. For the general study of one-parameter semigroups, we refer to standard textbooks [ENB00, Paz12].

By [ENB00, Proposition V.1.15], the cocycle Φ has an exponential dichotomy at 0 if and only if $\sigma(\Phi^t) \cap S^1 = \emptyset$ for one, and therefore all, $t > 0$. Consequently, Φ has an exponential dichotomy at $\lambda \in \mathbb{R}$ if and only if $\sigma(\Phi^t) \cap e^{\lambda t} S^1 = \emptyset$ for one, and therefore all, $t > 0$. Given that Φ has an exponential dichotomy at $\lambda \in \mathbb{R}$, the corresponding projection-valued function $\Pi : \Theta \to \mathcal{L}(\mathcal{H})$ is constant. For $t > 0$, the operator Φ^t is compact. By the Riesz–Schauder theorem, cf. [vN22, Theorem 7.11], the spectrum $\sigma(\Phi^t)$ consists of discrete points $(z_k)_{1\leq k\leq N}$ for $N \in \mathbb{N} \cup \{\infty\}$ that may only accumulate at 0. Additionally, each nonzero element of the spectrum is an eigenvalue, i.e. lies in the point spectrum. We conclude that the Sacker–Sell spectrum of Φ is given by

$$\Sigma(\Phi) = \big\{ \log(|z_k|) \mid 1 \leq k \leq N \big\}.$$

Hence, $\Sigma(\Phi)$ consists only of discrete points. Let $\ldots < \lambda_2 < \lambda_1$ be the discrete values such that $\Sigma(\Phi) = \{\lambda_k \mid 1 \leq k \leq N\}$. For each λ_k there is a vector $x \in \mathcal{H}$ and a complex value $z_k \in \mathbb{C}$ with $|z_k| = \lambda_k$ such that for each $t \geq 0$ we find $\Phi^t x = e^{z_k t} x$.

Theorem 2.6 implies that the spectrum of $\mathbf{M}^t$ consists of circles around the origin at radii $e^{\lambda_k t}$ and the spectrum of $\mathbf{G}$ consists of lines parallel to the imaginary axis at real parts λ_k. Furthermore, the existence of eigenfunctions is guaranteed. Let $x \in \mathcal{H}$ such that $\Phi^t x = e^{z_k t} x$ for all $t \geq 0$. Then, the function $\mathbf{f} \in \mathfrak{F}$ that admits the constant value x, i.e. $\mathbf{f}(\theta) = x$ for all $\theta \in \Theta$, is an eigenfunction of the Mather operators $\mathbf{M}^t$ and of the generator $\mathbf{G}$ with eigenvalue $e^{z_k t}$ and z_k, respectively.

3 The transfer-operator cocycle

We consider the evolution of particles on a compact domain advected by a nonautonomous vector field and subject to small diffusion. The goal of our analysis is to characterize and numerically compute *coherent sets*, which are time-dependent regions in space that only mix slowly with their surrounding. This problem has been studied for finite-time horizons in [DJM16, FKS20] and for periodically driven vector fields in [FK17]. In this work, we apply a similar approach to find coherent sets in vector fields that are driven by an ergodic base dynamics.

Let Θ be a compact smooth manifold equipped with a continuous vector field Ψ. Let $M \subset \mathbb{R}^d$ be an open, bounded domain with piecewise C^4 boundary (such that $\overline{M} \subset \mathbb{R}^d$ is compact) or a torus $M = \mathbb{T}^d \cong [0,1]^d_{/\sim}$. We call Θ the *parameter space* and M the *physical space*. We consider the nonautonomous SDE

$$\begin{aligned} d\theta_t &= \Psi(\theta_t)dt, \\ dx_t &= v(\theta_t, x_t)dt + \varepsilon\, dw_t, \end{aligned} \tag{5}$$

where $\varepsilon > 0$ and w_t is a d-dimensional standard Wiener process. We impose reflecting boundary conditions, unless $M = \mathbb{T}^d$. We assume that Ψ is a continuous vector field on Θ that generates an invertible flow $\phi^t : \Theta \to \Theta$, for which the aperiodic points are dense, and fix an ergodic measure μ on Θ. The vector field $v : \Theta \times \overline{M} \to \mathbb{R}^d$ is assumed to be divergence-free and smooth in both variables. For all $\theta \in \Theta$, the vector field $v(\theta, \cdot)$ should have no flow across the boundary of M, i.e. $v(\theta, x) \cdot n(x) = 0$ for all $x \in \partial M$, where $n(x)$ is the outer normal unit vector in x.

We study how distributions of particles evolve in time under the SDE (5) on the domain M with reflecting boundary conditions. For a formal introduction to the Fokker–Planck equation and its unique solvability, we refer to [Sta22, Chapter 2]. In the following, we consider particle distributions represented by functions $f \in L^2(M, \mathbb{R})$. The temporal evolution of an initial distribution $f_0 \in L^2(M, \mathbb{R})$ under the SDE (5) with starting parameter $\theta_0 \in \Theta$ is governed by the Fokker–Planck equation

$$\partial_t f(t,x) = \frac{1}{2}\varepsilon^2 \Delta_x f(t,x) - \operatorname{div}_x\big(f(t,\cdot)v(\phi^t\theta_0,\cdot)\big)(x), \quad f(0,\cdot) = f_0, \quad \frac{\partial f(t,\cdot)}{\partial n}\Big|_{\partial M} = 0. \tag{6}$$

The third term is the reflecting boundary condition, which requires that the derivative of $f(t,\cdot)$ in the direction of the outer normal unit vector on ∂M vanishes for all $t > 0$. We use

the convention $f(t) := f(t,\cdot) \in L^2(M,\mathbb{R})$. The Fokker–Planck equation can be written as a nonautonomous abstract Cauchy problem (NACP) on the space $C(\mathbb{R}_{\geq 0}, L^2(M,\mathbb{R}))$

$$\begin{aligned} \partial_t f(t) &= G(\phi^t\theta_0) f(t), \\ f(0) &= f_0. \end{aligned} \tag{7}$$

The generator $G(\theta)$ is the unbounded operator defined on a dense domain $\mathcal{D}(G(\theta)) \subset L^2(M,\mathbb{R})$, and is given by

$$\begin{aligned} \big[G(\theta)f\big](x) &= \frac{1}{2}\varepsilon^2 \Delta_x f(x) - \operatorname{div}_x\left(v(\theta,\cdot) f\right)(x) \\ &= \frac{1}{2}\varepsilon^2 \Delta_x f(x) - \nabla_x f(x) \cdot v(\theta, x). \end{aligned}$$

The simplification of the divergence term is possible since v was assumed to be divergence-free. The implementation of reflecting boundary conditions is discussed below.

Definition 3.1. *A classical solution to the NACP (7) is a function $f \in C(\mathbb{R}_{\geq 0}, L^2(M,\mathbb{R})) \cap C^1(\mathbb{R}_{>0}, L^2(M,\mathbb{R}))$ with $f(0) = f_0$ that, for all $t > 0$ satisfies $\partial_t f(t) = G(\phi^t\theta_0) f(t)$ and $f(t) \in \mathcal{D}(G(\phi^t\theta_0))$.*

The reflecting boundary condition can be encoded in the domain of the operator. We choose, independently of θ,

$$\mathcal{D}(G) = \mathcal{D}(G(\theta)) := \Big\{ f \in H^2(M) \,\Big|\, \frac{\partial f}{\partial n}\Big|_{\partial M} = 0 \Big\}, \tag{8}$$

where $H^2(M) = W^{2,2}(M)$ the usual Sobolev space. In the case $M = \mathbb{T}^d$ there is no boundary and the domain of the generator is simply $H^2(M)$. Note that $\mathcal{D}(G) \subset L^2(M,\mathbb{R})$ is dense. With this choice of the domain of G, any classical solution to the NACP (7) naturally satisfies the boundary condition, and hence, is a solution to the Fokker–Planck equation (6) and vice versa. Therefore, we call both (6) and (7) the Fokker–Planck equation and use whichever formulation is suitable in the situation at hand. The following theorem asserts the existence of a unique classical solution to the Fokker–Planck equation.

Theorem 3.2. *For any $f_0 \in L^2(M,\mathbb{R})$ and $\theta_0 \in \Theta$, the NACP (7) with domain $\mathcal{D}(G) = \{f \in H^2(M) \mid \frac{\partial f}{\partial n}|_{\partial M} = 0\}$ has a unique classical solution.*

Proof. See [Sta22, Theorem 3.12]. For further details on parabolic PDEs generating NACPs we refer to the classical textbooks [Lun12, Tan17]. □

For a starting parameter $\theta \in \Theta$ and a time $t > 0$, we define $\mathcal{P}_\theta^t : L^2(M,\mathbb{R}) \to L^2(M,\mathbb{R})$ to be the solution operator, also called Perron–Frobenius operator or Kolmogorov forward operator, for time t of the Fokker–Planck equation, i.e. $\mathcal{P}_\theta^t f_0 = f(t)$. By convention, we set $\mathcal{P}_\theta^0 = \mathrm{Id}$. It is well-known, that the operators $\mathcal{P}_\theta^t$ are Markov operators, i.e.

$$f_0 \geq 0 \quad \Rightarrow \quad \mathcal{P}_\theta^t f_0 \geq 0, \quad \text{and} \quad \int_M [\mathcal{P}_\theta^t f_0](x)\, dx = \int_M f_0(x)\, dx.$$

Since v is divergence-free, the constant function $\mathbb{1} \in L^2(M, \mathbb{R})$ is an invariant solution to the Fokker–Planck equation, i.e. $\mathcal{P}_\theta^t \mathbb{1} = \mathbb{1}$ for all $\theta \in \Theta$ and $t \geq 0$. Due to diffusion, $\mathcal{P}_\theta^t f$ converges to a constant function as $t \to \infty$ for any $f \in L^2(M, \mathbb{R})$. Hence, functions that integrate to 0 converge to the 0-function under $\mathcal{P}$. Proposition 3.3 below gives a quantitative formulation of this fact.

We can extend $\mathcal{P}_\theta^t$ by linearity to act on complex-valued functions. Define the Hilbert space

$$\mathcal{H} := L^2(M, \mathbb{C}).$$

By construction, the operators $\mathcal{P}_\theta^t$ are real, i.e. they map real-valued functions to real-valued functions. We also define the space of functions whose integral vanishes

$$\mathcal{H}_0 := \{f \in L^2(M, \mathbb{C}) \mid \textstyle\int_M f(x)dx = 0\}.$$

Since the operators $\mathcal{P}_\theta^t$ leave the integral of a function invariant, the subspace $\mathcal{H}_0$ is forward invariant, i.e. $\mathcal{P}_\theta^t \mathcal{H}_0 \subset \mathcal{H}_0$. From now on, we restrict our attention to the subspace $\mathcal{H}_0$ and study the solution operators $\mathcal{P}_\theta^t$ as operators on $\mathcal{H}_0$.

Proposition 3.3. *There is a constant $\varrho < 0$ such that the transfer operators $\mathcal{P}_\theta^t : \mathcal{H}_0 \to \mathcal{H}_0$ satisfy*

$$\left\|\mathcal{P}_\theta^t\right\| \leq e^{\varrho t},$$

for all $\theta \in \Theta$ and $t \geq 0$.

Proof. See Appendix A. □

We address the regularity of the operators $\mathcal{P}_\theta^t$ in both θ and t. The Perron–Frobenius transfer operator of a deterministic dynamic is strongly continuous with respect to the vector field and to time, but not norm continuous. In the case of additive noise, [KLP19] showed that $\mathcal{P}_\theta^t$ is not only norm continuous in θ, but even Fréchet-differentiable. The following lemma characterizes the regularity of $\mathcal{P}_\theta^t$ in both θ and t simultaneously.

Lemma 3.4. *Fix $t \geq 0$ and $\theta_1, \theta_2 \in \Theta$. Let $\delta > 0$ such that*

$$\|v(\phi^s \theta_2, x) - v(\phi^s \theta_1, x)\|_e \leq \delta,$$

in the Euclidean norm, for all $s \in [0, t]$ and $x \in M$. Then, the following bound in operator norm holds:

$$\left\|\mathcal{P}_{\theta_1}^t - \mathcal{P}_{\theta_2}^t\right\| \leq \frac{\delta}{\varepsilon}\sqrt{t}.$$

Proof. See Appendix A. □

We verify that the solution operators $\mathcal{P}$ form a cocycle over ϕ satisfying assumptions (I)–(IV). By invariance of the subspace $\mathbb{C}\mathbb{1}$, analogous assertions hold when $\mathcal{P}$ is considered as a cocycle over $\mathcal{H}$ instead of $\mathcal{H}_0$.

Theorem 3.5. *The Perron–Frobenius operators $\mathcal{P}_\theta^t : \mathcal{H}_0 \to \mathcal{H}_0$ of the Fokker–Planck equation (6) form a well-defined linear cocycle over the driving ϕ that satisfies assumptions (I)–(IV) introduced in Section 2.1.*

Proof. To apply the results of [Sta22], we restrict ourselves to the time interval $[0, T]$ for now and fix a starting parameter $\theta \in \Theta$. By [Sta22, Theorem 3.12], the solution operator to the Fokker–Planck equation (6) is given by a two-parameter evolution family $P_{s,t} : \mathcal{H} \to \mathcal{H}$, where $P_{s,t}$ is the solution operator from time $s \geq 0$ to time $t \geq s$. By definition, we find $\mathcal{P}_\theta^t = P_{0,t}|_{\mathcal{H}_0}$, as well as $\mathcal{P}_{\phi^s\theta}^t = P_{s,s+t}|_{\mathcal{H}_0}$. The evolution property of P yields

$$\mathcal{P}_\theta^{s+t} = P_{0,s+t}|_{\mathcal{H}_0} = P_{s,s+t}|_{\mathcal{H}_0} \circ P_{0,s}|_{\mathcal{H}_0} = \mathcal{P}_{\phi^s\theta}^t \circ \mathcal{P}_\theta^s.$$

This holds for all $\theta \in \Theta$ and $0 \leq s \leq t \leq T$. Since $T > 0$ was arbitrary, the operators $\mathcal{P}$ form a well-defined cocycle over ϕ.

We show that the cocycle $\mathcal{P}$ satisfies assumptions (I)–(IV) defined in Section 2.1.

(I) *compact.* Compactness of $\mathcal{P}_\theta^t$ for $t > 0$ follows from [Sta22, Theorem 3.12].

(II) *norm-continuous in θ.* Fix $t \geq 0$ and let $\theta_n \to \theta \in \Theta$. The vector field v is smooth in both variables and the flow ϕ is continuous. Hence, as $\theta_n \to \theta$ the smallest number $\delta_n > 0$ that satisfies

$$\|v(\phi^s\theta_n, x) - v(\phi^s\theta, x)\|_e \leq \delta_n,$$

for all $s \in [0, t]$ and $x \in M$, approaches 0. The statement of Lemma 3.4 implies that $\left\|\mathcal{P}_{\theta_n}^t - \mathcal{P}_\theta^t\right\| \to 0$. This shows that $\theta \mapsto \mathcal{P}_\theta^t$ is norm-continuous.

(III) *strongly continuous.* Fix $f \in \mathcal{H}_0$. Let $\theta_n \to \theta \in \Theta$ and $t_n \to t \geq 0$. We compute

$$\left\|\mathcal{P}_{\theta_n}^{t_n} f - \mathcal{P}_\theta^t f\right\| \leq \left\|\mathcal{P}_{\theta_n}^{t_n} f - \mathcal{P}_\theta^{t_n} f\right\| + \left\|\mathcal{P}_\theta^{t_n} f - \mathcal{P}_\theta^t f\right\|. \tag{9}$$

The sequence t_n is bounded from above by some value $T > 0$. As $\theta_n \to \theta$, the smallest number $\delta_n > 0$ that satisfies

$$\|v(\phi^s\theta_n, x) - v(\phi^s\theta, x)\|_e \leq \delta_n,$$

for all $s \in [0, T]$ and $x \in M$, approaches 0. By Lemma 3.4, we find $\left\|\mathcal{P}_{\theta_n}^{t_n} - \mathcal{P}_\theta^{t_n}\right\| \to 0$ as $n \to \infty$. In particular, the first summand in (9) tends to 0 as $n \to \infty$.

Since $f(s) = \mathcal{P}_\theta^s f$ is the solution to the NACP (7) with starting parameter θ, and $f(s)$ is continuous in s, the second summand in (9) vanishes as well as $n \to \infty$. Hence, $(\theta, t) \mapsto \mathcal{P}_\theta^t f$ is continuous.

(IV) *exponentially bounded.* By Proposition 3.3, we find $\|\mathcal{P}_\theta^t\| \leq e^{\varrho t}$, for some $\varrho < 0$. Hence, the cocycle $\mathcal{P}$ is exponentially bounded with $K = 1$ and $L = \varrho$.

This finishes the proof. □

Since the Perron–Frobenius cocycle $\mathcal{P}$ satisfies assumptions (I)–(IV), the theory of the previous sections is applicable. Let $\mathfrak{F}$ be either $L^2(\Theta, \mathcal{H}_0)$ or $C(\Theta, \mathcal{H}_0)$. The Mather semigroup $\mathbf{M}^t : \mathfrak{F} \to \mathfrak{F}$ defined by (2) is a strongly continuous semigroup with closed generator $\mathbf{G}$. Since $\mathcal{P}$ is the strongly continuous cocycle that is the solution to the NACP (7), we apply [CL99, Example 6.22, Proposition 6.23][1] to compute the augmented generator $\mathbf{G}$:

$$[\mathbf{G}\mathbf{f}](\theta) = G(\theta)\mathbf{f}(\theta) - [\mathbf{d}\mathbf{f}](\theta),$$

where $\mathbf{d}$ is the generator of the translation semigroup on $\mathfrak{F}$ in the direction of Ψ. The operator is defined on its maximal domain and given by

$$[\mathbf{d}\mathbf{f}](\theta) = \partial_t \mathbf{f}(\phi^t\theta)|_{t=0} = \partial_\theta \mathbf{f}(\theta) \cdot \Psi(\theta).$$

We note that for $\mathfrak{F} = L^2(\Theta, \mathcal{H}_0)$ the derivative is in the weak sense. The generator $\mathbf{G}$ is defined on

$$\mathcal{D}(\mathbf{G}) := \{\mathbf{f} \in \mathfrak{F} \mid \mathbf{f} \in \mathcal{D}(\mathbf{d}),\ \mathbf{f} : \Theta \to \mathcal{D}(G),\ G(\cdot)\mathbf{f}(\cdot) - \mathbf{d}\mathbf{f} \in \mathfrak{F}\}. \tag{10}$$

The key property of this domain which will be used later is the fact that for any function $\mathbf{f} \in \mathcal{D}(\mathbf{G})$ the fibres $\mathbf{f}(\theta)$ lie in $\mathcal{D}(G)$. An explicit form of the augmented generator is given by

$$[\mathbf{G}\mathbf{f}](\theta, x) = \frac{1}{2}\varepsilon^2 \Delta_x \mathbf{f}(\theta, x) - \nabla_x \mathbf{f}(\theta, x) \cdot v(\theta, x) - \partial_\theta \mathbf{f}(\theta, x) \cdot \Psi(\theta). \tag{11}$$

By Theorem 2.6, the spectrum of $\mathbf{G}$ is given by

$$\sigma(\mathbf{G}) = \{\lambda + \eta i \mid \lambda \in \Sigma(\mathcal{P}),\ \eta \in \mathbb{R}\}.$$

Since $\|\mathcal{P}_\theta^t\| \leq e^{\varrho t}$ for all $\theta \in \Theta$ and $t \geq 0$, we find $\Sigma(\mathcal{P}) \subset (-\infty, \varrho]$. In general, the existence of eigenvalues of $\mathbf{G}$ cannot be guaranteed. The existence of so-called approximate eigenfunctions is discussed in Section 4.2.

Remark 3.6. The space $L^2(\Theta, \mathcal{H})$ is isomorphic to $L^2(\Theta \times M, \mathbb{C})$. The Mather operators $\mathbf{M}^t$, interpreted as operators on the space $L^2(\Theta \times M, \mathbb{C})$ are, in fact, transfer operators of the dynamics on $\Theta \times M$ defined in (5). Hence, the Mather semigroup on $L^2(\Theta, \mathcal{H}_0)$ can be interpreted as the semigroup of transfer operators of the autonomous (skew product) SDE (5) on $\Theta \times M$ in which the noise only acts in the dimensions of M.

4 Coherent sets

Given a nonautonomous particle flow like (5), there are regions in space that experience more mixing than other regions. Due to the nonautonomous nature of the system, these regions are in general time-dependent. We call a time-evolving region that experiences little mixing a *coherent set*. We characterize coherent sets by the property that particles that start inside a

[1] In [CL99], the statement is formulated for the Mather semigroup on $C(\Theta, \mathcal{H}_0)$. However, the proof is directly transferable to $L^2(\Theta, \mathcal{H}_0)$.

coherent set stay inside the coherent set for a long time. Given a starting parameter $\theta \in \Theta$, consider a family of sets $A_\theta^\bullet := (A_\theta^t)_{t\geq 0}$, with $A_\theta^t \subset M$. We say that $A_\theta^\bullet$ is coherent if the survival probability

$$\mathbb{P}_{x_0\sim A_\theta^0}(x_s \in A_\theta^s,\ \forall s \in [0,t])$$

decays slowly, or alternatively we call $A_\theta^\bullet$ a coherent set. The initial condition $x_0 \sim A_\theta^0$ asserts that x_0 is uniformly distributed on the set A_θ^0. This is not a rigorous definition of a coherent set until we have made more precise what slow decay means. To this end, we introduce two quantities that quantify the coherence of $A_\theta^\bullet$. The first notion is the *escape rate*[2]

$$E(A_\theta^\bullet) := \liminf_{t\to\infty} \frac{1}{t} \log\left(\mathbb{P}_{x_0\sim A_\theta^0}(x_s \in A_\theta^s,\ \forall s \in [0,t])\right), \tag{12}$$

which considers the asymptotic decay rate of the survival probability. By definition, the escape rate is non-positive. An escape rate close to 0 corresponds to a highly coherent set. The escape rate $E(A_\theta^\bullet)$ has been studied for periodically driven systems in [FK17]. Note that the escape rate does not guarantee coherence on any finite time horizon. It might happen that the survival probability quickly approaches 0, before transitioning to an exponential decay at rate $E(A_\theta^\bullet)$. For this reason, we introduce another quantity of coherence, the *cumulative survival probability*

$$C(A_\theta^\bullet) := \int_0^\infty \mathbb{P}_{x_0\sim A_\theta^0}(x_s \in A_\theta^s,\ \forall s \in [0,t])\, dt.$$

The goal of this section is to elaborate how coherent sets $A_\theta^\bullet$ with an escape rate $E(A_\theta^\bullet)$ close to 0 or a high cumulative survival probability $C(A_\theta^\bullet)$ can be extracted from spectral objects associated with $\mathbf{M}^t$ or the cocycle $\mathcal{P}$. Ultimately, we want to compute coherent sets $A_\theta^\bullet$ for every starting parameter $\theta \in \Theta$ simultaneously, since $\mathbf{M}^t$ allows to simultaneously evolve functions f_θ by all $\mathcal{P}_\theta^t$, for all $\theta \in \Theta$, respectively. Hence, we are looking for a family of sets $A_\bullet^\bullet := (A_\theta^t)_{\theta\in\Theta, t\geq 0}$ such that $A_\theta^\bullet := (A_\theta^t)_{t\geq 0}$ is coherent for each $\theta \in \Theta$. In some cases it is possible to choose the sets A_θ^t such that $A_\theta^t = A_{\phi^t\theta}^0$ for each $\theta \in \Theta$ and $t \geq 0$. Then, the family $A_\bullet^\bullet$ is fully characterized by the sets A_θ^0 for $\theta \in \Theta$. If this is the case, we omit the superscript 0 and simply write $A_\bullet = (A_\theta)_{\theta\in\Theta}$. Such a representation is particularly nice in practice, since once the sets A_θ are determined for all $\theta \in \Theta$, one obtains a coherent set for any starting parameter θ.

In [FK17] it was shown that for fixed $\theta \in \Theta$, coherent sets can be extracted from functions $f \in \mathcal{H}_0$ that decay slowly under $\mathcal{P}$. This construction requires the point-wise evaluation of $\mathcal{P}_\theta^t f$ in x. Since we work with L^2-functions, point-wise evaluations are a priori not well-defined. However, we know that $\mathcal{P}_\theta^t f \in \mathcal{D}(G) \subset H^2(M)$ such that there is a unique continuous representative of $\mathcal{P}_\theta^t f$. We state an even stronger regularity result. To apply a suited Sobolev embedding, the dimension of the physical space M should not exceed 3.

Proposition 4.1. *Assume that the dimension d of M is at most 3. Let $f_0 \in \mathcal{D}(G)$ and let $\theta \in \Theta$. Then $f : \mathbb{R}_{\geq 0} \to \mathcal{H}_0$, defined by $f(t) = [\mathcal{P}_\theta^t f_0]$, is represented by a continuous function $f \in C(\mathbb{R}_{\geq 0} \times M, \mathbb{C})$.*

[2]In the literature, the escape rate is sometimes defined using the lim sup. Hence, our definition corresponds to the "pessimistic" escape rate.

Proof. See [FKS20, Theorem A.4]. The assumption $d \leq 3$ is needed such that the Sobolev embedding $W^{2,2}(M) \subset C^\alpha(M)$ holds for $\alpha = 2 - \frac{d}{2}$, where $C^\alpha(M)$ is the usual Hölder space. □

The following theorem is almost identical to [FK17, Theorem 19] with the difference that the latter result requires slightly stronger assumptions.

Theorem 4.2. *Let $f_0 \in \mathcal{D}(G)$ be a real function with $\int_M f_0(x)dx = 0$ and let $\theta \in \Theta$. Define the sets*

$$A_\theta^t := \{x \in M \mid [\mathcal{P}_\theta^t f_0](x) \geq 0\}. \tag{13}$$

Then, for any $t \geq 0$, we find

$$\mathbb{P}_{x_0 \sim A_\theta^0}(x_s \in A_\theta^s,\ \forall s \in [0,t]) \geq \frac{1}{2} \|f_0\|_\infty^{-1} \left|A_\theta^0\right|^{-1} \left\|\mathcal{P}_\theta^t f_0\right\|_1.$$

Proof. See Appendix A. □

Recall that $\mathcal{D}(G) \subset H^2(M)$, cf. (8), such that $\|f_0\|_\infty$ is finite. Note that the theorem above bounds the decay of the survival probability by the L^1-norm of $\mathcal{P}_\theta^t f$. Since we work on $\mathcal{H}_0 \subset L^2(M, \mathbb{C})$, we need an additional estimation to express the decay in terms of the L^2-norm of $\mathcal{P}_\theta^t f$. Bounding the L^1-norm $\|f\|_1$ from below by $\|f\|_\infty^{-1} \|f\|_2^2$ via Hölder, we obtain

$$\mathbb{P}_{x_0 \sim A_\theta^0}(x_s \in A_\theta^s,\ \forall s \in [0,t]) \geq \frac{1}{2} \|f\|_\infty^{-2} \left|A_\theta^0\right|^{-1} \left\|\mathcal{P}_\theta^t f\right\|_2^2. \tag{14}$$

The bound on the the L^1-norm in terms of the ∞-norm and the square of the L^2-norm is in general rather crude. Hence, (14) is expected to underestimate the survival probability of $A_\theta^\bullet$. In Section 6, we compute and evaluate coherent sets for some examples. Discretizing the system introduces errors that decrease the coherence of the computed coherent sets. The results of Section 6 indicate that the discretization error and the crude bound on the L^1-norm seem to cancel, in the sense that the simulated survival probabilities decay similarly to the right-hand side of (14).

Remark 4.3. Theorem 4.2 establishes a connection between the decay rate of a function $f \in C(M, \mathbb{R})$ under $\mathcal{P}$ and the escape rate of the corresponding coherent set $A_\theta^\bullet$ defined in (13). We find

$$\begin{aligned} E(A_\theta^\bullet) &\geq \liminf_{t\to\infty} \frac{1}{t} \log\left(\frac{1}{2} \|f\|_\infty^{-1} \left|A_\theta^0\right|^{-1} \left\|\mathcal{P}_\theta^t f\right\|_1\right) \\ &= \liminf_{t\to\infty} \frac{1}{t} \log\left(\left\|\mathcal{P}_\theta^t f\right\|_1\right). \end{aligned}$$

This quantity is called the *Lyapunov exponent* (in the L^1-norm) of f under $\mathcal{P}$. The study of Lyapunov exponents to characterize coherent sets has been pioneered in [FLS10].

4.1 Coherent sets from eigenfunctions

In the previous works [FK17, FKS20], coherent sets were extracted via Theorem 4.2 from eigenfunctions of an augmented generator, as these provided functions that decayed slowly. In our setting, the augmented generator $\mathbf{G}$ and the Mather operators $\mathbf{M}^t$ do in general not possess eigenfunctions. However, in Section 5 we construct finite-dimensional approximations of the linear operator $\mathbf{G}$ and the Mather operators $\mathbf{M}^t$. Naturally, these matrices admit eigenfunctions. In the following paragraph we assume the existence of an eigenfunction $\mathbf{f}$ of the Mather semigroup and its generator over $C(\Theta, \mathcal{H}_0)$ and show how to construct coherent sets from $\mathbf{f}$. Hence, this subsection should be understood as the derivation of a heuristic method that can be applied to eigenfunctions of the discrete generator.

Assume there was an eigenfunction $\mathbf{f} \in C(\Theta, \mathcal{H}_0)$ of the augmented generator $\mathbf{G}$ with eigenvalue $z = \lambda + \eta i \in \mathbb{C}$. Consequently, $\mathbf{f}$ is an eigenfunction of $\mathbf{M}^t$ with eigenvalue e^{zt} for every $t \geq 0$. In particular, we find

$$\mathcal{P}_\theta^t \mathbf{f}(\theta) = e^{zt} \mathbf{f}(\phi^t \theta), \tag{15}$$

for all $\theta \in \Theta$ and $t \geq 0$. By Theorem 3.2, we find $\mathbf{f}^R(\theta) \in \mathcal{D}(G)$ for all $\theta \in \Theta$. Let $\mathbf{f}^R := \mathrm{Re}(\mathbf{f})$ be the real part of $\mathbf{f}$. Since $\mathcal{P}$ is a real cocycle, (15) implies

$$\mathcal{P}_\theta^t \mathbf{f}^R(\theta) = e^{\lambda t} \mathrm{Re}(e^{\eta t i} \mathbf{f}(\phi^t \theta)).$$

We define a coherent family $A_\bullet^\bullet = (A_\theta^t)_{\theta \in \Theta, t \geq 0}$ by

$$A_\theta^t := \{x \in M \mid \mathrm{Re}(e^{\eta t i} \mathbf{f}(\phi^t \theta, x)) \geq 0\}. \tag{16}$$

The estimate (14) yields

$$\mathbb{P}_{x_0 \sim A_\theta^0}(x_s \in A_\theta^s,\ \forall s \in [0,t]) \geq \frac{1}{2} e^{2\lambda t} \left\|\mathbf{f}^R(\theta)\right\|_\infty^{-2} \left|A_\theta^0\right|^{-1} \left\|\mathrm{Re}(e^{\eta t i} \mathbf{f}(\phi^t \theta))\right\|_2^2. \tag{17}$$

Assuming that the real part of the fibres $e^{\eta t i} \mathbf{f}(\phi^t \theta)$ is uniformly bounded from below, we conclude that the survival probability of $A_\theta^\bullet$ decays at most at rate $e^{2\lambda t}$, up to constants. Hence, we expect

$$E(A_\theta^\bullet) \geq 2\lambda, \quad C(A_\theta^\bullet) \gtrsim \int_0^\infty e^{2\lambda t} dt = -\frac{1}{2\lambda}. \tag{18}$$

The results of Section 6 show that $-\frac{1}{2\lambda}$ seems to be a good estimate of the cumulative survival probability that we computed in numerical examples.

Using (16) to define a coherent set enables us to derive theoretical bounds on the survival probability, but the resulting coherent sets might not provide much insight into the dynamical properties of the underlying vector field. The definition simply divides the space M into two regions of roughly the same size. Considering the function $-\mathbf{f}$ shows that the complement $B_\theta^t := M \setminus A_\theta^t$ satisfies the same bounds on the survival probability. Hence, the method (16) is unable to identify multiple smaller coherent regions. It was, however, extended in [FKS20, Proposition 4.6] to use linear combinations of multiple eigenfunctions to obtain multiple coherent sets from them. Here, in Section 5.3 we propose alternative methods to identify coherent sets from the fibres of an eigenfunction $\mathbf{f}$.

4.2 Coherent sets from approximate eigenfunctions

In general, the augmented generator $\mathbf{G}$ does not possess eigenfunctions. However, the existence of *approximate eigenfunctions* is guaranteed. While less applicable in practice, we show how approximate eigenfunctions can be utilized to construct coherent sets. We briefly introduce the approximate spectrum of an operator.

Let $\mathbb{B}$ be a Banach space and let $T : \mathcal{D}(T) \to \mathbb{B}$ be a closed operator, where $\mathcal{D}(T) \subset \mathbb{B}$ is the domain of T. The *approximate spectrum* of T is defined by

$$\sigma_{\mathrm{ap}}(T) = \{z \in \mathbb{C} \mid \exists x_n \in \mathcal{D}(T) \text{ with } \|x_n\| = 1 : \ \|(T - z)x_n\| \to 0, \text{ as } n \to \infty\}.$$

We call the sequence $(x_n)_{n\in\mathbb{N}}$, and any element of it, an *approximate eigenfunction* to the approximate eigenvalue z. We find $\partial\sigma(T) \subset \sigma_{\mathrm{ap}}(T) \subset \sigma(T)$, where $\partial\sigma(T)$ is the boundary of the spectrum of T. For a proof of this fact, see [ENB00, Chapter IV, Proposition 1.10].

Recall Theorem 2.6, which characterized the spectrum of the Mather operators $\mathbf{M}^t$ and augmented generator $\mathbf{G}$. In particular, the spectrum of $\mathbf{G}$ is given by

$$\sigma(\mathbf{G}) = \{\lambda + \eta i \mid \lambda \in \Sigma(\mathcal{P}),\ \eta \in \mathbb{R}\}. \tag{19}$$

As mentioned at the end of Section 3, the Sacker–Sell spectrum $\Sigma(\mathcal{P})$ is contained in $(-\infty, \varrho]$ for the constant $\varrho < 0$ from Proposition 3.3. Assume that $\Sigma(\mathcal{P}) \neq \emptyset$, and let $\lambda < 0$ be its maximum value. By (19), we find $\lambda \in \partial\sigma(\mathbf{G})$, and, therefore, $\lambda \in \sigma_{\mathrm{ap}}(\mathbf{G})$. This holds independently of whether we consider the Mather semigroup and its generator on $L^2(\Theta, \mathcal{H}_0)$ or $C(\Theta, \mathcal{H}_0)$.

Since $\mathbf{G}$ is real and $\lambda \in \mathbb{R}$, we find

$$\begin{aligned}\|(\mathbf{G} - \lambda)\mathbf{f}_n\|_\infty &\geq \max\{\|\mathrm{Re}((\mathbf{G} - \lambda)\mathbf{f}_n)\|_\infty\,,\ \|\mathrm{Im}((\mathbf{G} - \lambda)\mathbf{f}_n)\|_\infty\} \\ &= \max\{\left\|(\mathbf{G} - \lambda)\mathbf{f}_n^R\right\|_\infty\,,\ \left\|(\mathbf{G} - \lambda)\mathbf{f}_n^I\right\|_\infty\},\end{aligned}$$

where $\mathbf{f}_n^R$ and $\mathbf{f}_n^I$ are the real and imaginary parts of $\mathbf{f}_n$. Note that both $\mathbf{f}_n^R$ and $\mathbf{f}_n^I$ are real functions with $\|\mathbf{f}_n\|_\infty \leq \left\|\mathbf{f}_n^R\right\|_\infty + \left\|\mathbf{f}_n^I\right\|_\infty$. Hence, w.l.o.g. we can choose an approximate eigenfunction $(\mathbf{f}_n)_{n\in\mathbb{N}}$ such that $\mathbf{f}_n \in C(\Theta, L^2(M, \mathbb{R}))$ are real functions. In particular, since $\mathbf{f}_n \in \mathcal{D}(\mathbf{G})$, we have that $\mathbf{f}_n(\theta) \in \mathcal{D}(G)$ for every $\theta \in \Theta$, cf. (10).

Proposition 4.4. *For any $T > 0$ and $\delta > 0$, there is an $n \in \mathbb{N}$ such that the family of sets defined by*

$$A_\theta^t := \{x \in M \mid [\mathcal{P}_\theta^t \mathbf{f}_n(\theta)](x) \geq 0\} \tag{20}$$

satisfies the following estimate for all $\theta \in \Theta$ and $t \in [0, T]$:

$$\mathbb{P}_{x_0 \sim A_\theta^0}(x_s \in A_\theta^s,\ \forall s \in [0, t]) \geq \frac{1}{2} e^{2\lambda t} \|\mathbf{f}_n(\theta)\|_\infty^{-2} \left|A_\theta^0\right|^{-1} \left(\left\|\mathbf{f}_n(\phi^t\theta)\right\|_2 - \delta\right)^2. \tag{21}$$

Proof. The unbounded, closed operator $(\mathbf{G}-\lambda)$ generates the strongly continuous semigroup $e^{-\lambda}\mathbf{M}^t$. Since $\mathbf{f}_n \in \mathcal{D}(\mathbf{G} - \lambda)$, we apply [Paz12, Theorem 2.4] to compute

$$e^{-\lambda t}\mathbf{M}^t\mathbf{f}_n - \mathbf{f}_n = \int_0^t \mathbf{M}^s(\mathbf{G} - \lambda)\mathbf{f}_n \, ds.$$

We use the fact $\|\mathbf{M}^s\| \leq e^{\varrho s} \leq 1$, by Proposition 3.3, to obtain

$$\left\|e^{-\lambda t}\mathbf{M}^t\mathbf{f}_n - \mathbf{f}_n\right\|_\infty \leq t\,\|(\mathbf{G}-\lambda)\mathbf{f}_n\|_\infty\,.$$

As $n \to \infty$, the right-hand side approaches 0. For fixed $T > 0$ and $\delta > 0$ there is an $n \in \mathbb{N}$ such that

$$\left\|e^{-\lambda t}\mathcal{P}_\theta^t\mathbf{f}_n(\theta) - \mathbf{f}_n(\phi^t\theta)\right\|_2 \leq \delta$$

for all $\theta \in \Theta$ and $t \in [0,T]$. Rearranging the inequality yields

$$\left\|\mathcal{P}_\theta^t\mathbf{f}_n(\theta)\right\|_2 \geq e^{\lambda t}\big(\left\|\mathbf{f}_n(\phi^t\theta)\right\|_2 - \delta\big).$$

Now, the statement follows from (14). □

Note that the bound (21) bound is almost identical to (17), the bound on the survival probability of a coherent set extracted from an eigenfunction of $\mathbf{G}$, up the error term δ.

In conclusion, a coherent sets $A_\theta^\bullet$ extracted from an approximate eigenfunction is expected to have a similar survival probability as a coherent set extracted from an eigenfunction of $\mathbf{G}$. However, the bounds on the survival probability of $A_\theta^\bullet$ only hold up to a finite time-horizon T. Hence, we cannot make any statement about the asymptotic decay rate $E(A_\theta^\bullet)$, but we expect the cumulative survival probability to be of order $-\frac{1}{2\lambda}$. Another drawback of extracting coherent sets from an approximate eigenfunction versus from an eigenfunction is that the definition of $A_\theta^\bullet$ in (20) requires knowledge of the cocycle evaluations $\mathcal{P}_\theta^t\mathbf{f}_n(\theta)$, unlike the definition (16) of a coherent set extracted from an eigenfunction $\mathbf{f}$, which only required knowledge of the fibres of $\mathbf{f}$.

4.3 Coherent sets from spectral projections

In the previous section we studied coherent sets extracted from approximate eigenfunctions of the augmented generator $\mathbf{G}$. However, we were only able to derive bounds on the decay rate up to a finite time horizon $[0,T]$. In this section we show how coherent sets can be extracted from spectral subspaces, given the existence of exponential dichotomies. The advantage of this approach is that we are able to obtain a uniform decay rate for $t \to \infty$. In particular, we are able to derive a bound on the asymptotic escape rate $E(A_\theta^\bullet)$.

By Proposition 2.4 the Sacker–Sell spectrum $\Sigma(\mathcal{P})$ consists of the union of closed segments $[r_k^-, r_k^+] \subset \mathbb{R}$ for $1 \leq k \leq N$ and $N \in \mathbb{N}_0 \cup \{\infty\}$. Assume that $\Sigma(\mathcal{P}) \neq \emptyset$ and that $r_1^- > -\infty$, i.e. the interval $[r_1^-, r_1^+]$ is an isolated segment of $\Sigma(\mathcal{P})$. Fix $\lambda < r_1^-$ close to r_1^-. There is an exponential dichotomy at λ that splits the space $\mathcal{H}_0$ into a stable bundle $S(\theta)$ and an unstable bundle $U(\theta)$. There is a constants $\beta > 0$ and $C > 0$ such that

$$\left\|\mathcal{P}_\theta^t f\right\|_2 \geq C^{-1}e^{(\lambda+\beta)t}\,\|f\|_2\,,$$

for all $\theta \in \Theta$, $t \geq 0$ and $f \in U(\theta)$. The next lemma shows that w.l.o.g. we can choose $f \in U(\theta)$ to be a real function.

Lemma 4.5. *If $f \in U(\theta)$, then both $f^R := \mathrm{Re}(f)$ and $f^I := \mathrm{Im}(f)$ lie in $U(\theta)$ as well.*

Proof. See Appendix A. □

Let $f \in U(\theta)$ be a real function with $\|f\|_2 = 1$. By the definition of an exponential dichotomy, we find $f \in \mathrm{ran}(\mathcal{P}^t_{\phi^{-t}\theta})$ for some $t > 0$, and thereby, $f \in \mathcal{D}(G)$. Define a coherent set by

$$A^t_\theta := \{x \in M \mid [\mathcal{P}^t_\theta f](x) > 0\}.$$

By (14), the survival probability of this coherent set satisfies

$$\mathbb{P}_{x_0 \sim A^0_\theta}(x_s \in A^s_\theta,\ \forall s \in [0,t]) \geq \frac{1}{2}\|f\|_\infty^{-2}\left|A^0_\theta\right|^{-1} C^{-1} e^{(\lambda+\beta)t}.$$

This bound provides an estimate on the escape rate defined in (12):

$$E(A^\bullet_\theta) \geq \lambda + \beta.$$

Since $\lambda < r_1^-$ was arbitrary, we find coherent sets with escape rate arbitrarily close to r_1^-.

In Theorem 2.6, it was shown that for any $t \geq 0$ the spectrum of the Mather operators $\mathbf{M}^t$ contains the isolated annulus $\Lambda^t := \{e^{t\lambda+\eta i} \mid \lambda \in [r_1^-, r_1^+],\ \eta \in [0, 2\pi)\}$. By the contents of Section 2.4, in particular (4), the range of the Riesz-projection $\mathfrak{P}_{\Lambda^t}$ consists of functions with fibres in $U(\theta)$. Hence, given a function $\mathbf{f} \in \mathrm{ran}(\mathfrak{P}_{\Lambda^t})$, the nonzero fibres $\mathbf{f}(\theta) \in U(\theta)$ induce coherent sets with escape rate at least $\lambda + \beta$ (note that due to the negative sign a larger escape rate corresponds to higher coherence).

5 Numerical implementation

5.1 The augmented generator in Fourier space

The theoretical results of the previous two section were valid for domains M that were either open, bounded subsets of $\mathbb{R}^d$ with piecewise C^4 boundary, or a torus $\mathbb{T}^d$. In the following, to enable the use of Fourier series, we consider the case $\Theta = \mathbb{T}^{d_d}$ and $M = \mathbb{T}^{d_p}$. The dimensions d_d and d_p stand for the driving dimension and the physical dimension, respectively. For the ergodic driving dynamic we choose a quasi-periodic rotation on the torus, i.e.

$$\partial_t \theta_t = \Psi(\theta_t) = \alpha,$$

where $\alpha \in \mathbb{R}^{d_d}$ is a constant vector whose entries are rationally independent. We derive a finite-dimensional approximation of the Mather semigroup $\mathbf{M}^t$ and its generator $\mathbf{G}$.

Note that the space $L^2(\Theta, \mathcal{H})$ is isomorphic to $L^2(\Theta \times M, \mathbb{C}) = L^2(\mathbb{T}^{d_d} \times \mathbb{T}^{d_p}, \mathbb{C})$ which is a Hilbert space with inner product

$$\langle \mathbf{f}, \mathbf{g} \rangle = \int_{\mathbb{T}^{d_d}} \int_{\mathbb{T}^{d_p}} \overline{\mathbf{f}(\theta, x)} \mathbf{g}(\theta, x)\, dx\, d\theta.$$

An orthonormal basis, also called Hilbert basis, is given by the Fourier modes

$$F_{m,n}(\theta, x) := e^{2\pi i(\theta \cdot m + x \cdot n)},$$

for $m \in \mathbb{Z}^{d_d}$ and $n \in \mathbb{Z}^{d_p}$. We represent functions $\mathbf{f} \in L^2(\mathbb{T}^{d_d} \times \mathbb{T}^{d_p}, \mathbb{C})$ with respect to this basis, i.e.

$$\mathbf{f} = \sum_{m,n} \widehat{\mathbf{f}}(m,n) F_{m,n},$$

where the summation is over all $(m,n) \in \mathbb{Z}^{d_d} \times \mathbb{Z}^{d_p}$. The Fourier coefficients $\widehat{\mathbf{f}}(m,n) \in \mathbb{C}$ can be computed by

$$\widehat{\mathbf{f}}(m,n) = \langle F_{m,n}, \mathbf{f} \rangle = \int_{\mathbb{T}^{d_d}} \int_{\mathbb{T}^{d_p}} e^{-2\pi i(\theta \cdot m + x \cdot n)} \mathbf{f}(\theta, x)\, dx\, d\theta.$$

Analogously, we can represent the vector field $v : \mathbb{T}^{d_d} \times \mathbb{T}^{d_p} \to \mathbb{R}^{d_p}$ in Fourier coordinates. Each of the d_p components v_j has its individual Fourier coefficients $\widehat{v_j}(m,n)$. We write $\widehat{v}(m,n) \in \mathbb{C}^{d_p}$ for the vector of Fourier coefficients.

Lemma 5.1. *Let $v : \mathbb{T}^{d_d} \times \mathbb{T}^{d_p} \to \mathbb{R}^{d_p}$ be a divergence-free vector field. For any $(m,n) \in \mathbb{Z}^{d_d} \times \mathbb{Z}^{d_p}$, we find*

$$\begin{aligned} \widehat{v}(m,n) &= \overline{\widehat{v}(-m,-n)}, \\ n \cdot \widehat{v}(m,n) &= 0. \end{aligned}$$

Proof. See Appendix A. □

Similar to how we write functions $\mathbf{f} \in L^2(\mathbb{T}^{d_d} \times \mathbb{T}^{d_p}, \mathbb{C})$ in terms of their Fourier coefficients, we can be represent the augmented generator $\mathbf{G}$ as an infinite-dimensional matrix with rows and columns indexed by $\mathbb{Z}^{d_d} \times \mathbb{Z}^{d_p}$, respectively. Recall that we defined the Mather semigroup $(\mathbf{M}^t)_{t \geq 0}$ and its generator $\mathbf{G}$ over functions $\mathbf{f}$ with fibres $\mathbf{f}(\theta)$ in the Hilbert space $\mathcal{H}_0$ of L^2-functions that integrate to 0. This corresponds to only considering those Fourier modes $F_{m,n}$ for which $n \in \mathbb{Z}^{d_p}$ is nonzero. To avoid additional case distinctions, in this section we include functions that do not integrate to 0. We discuss this choice further in Remark 5.3.

Proposition 5.2. *Let $\mathbf{f} \in \mathcal{D}(\mathbf{G})$. We find*

$$\widehat{\mathbf{G}\mathbf{f}}(m,n) = \sum_{m',n'} \Gamma(m,n\,;\,m',n') \widehat{\mathbf{f}}(m',n').$$

The infinite-dimensional matrix $\Gamma \in \mathbb{C}^{(\mathbb{Z}^{d_d} \times \mathbb{Z}^{d_p}) \times (\mathbb{Z}^{d_d} \times \mathbb{Z}^{d_p})}$ is given by

$$\Gamma(m,n\,;\,m',n') := \begin{cases} \frac{1}{2}\varepsilon^2 (2\pi \|n\|_e)^2 - 2\pi i\big(m \cdot \alpha + n \cdot \widehat{v}(0,0)\big), & (m,n) = (m',n') \\ -2\pi i (n \cdot \widehat{v}(m-m', n-n')), & (m,n) \neq (m',n'). \end{cases} \tag{22}$$

Proof. Let $\mathbf{f} \in \mathcal{D}(\mathbf{G})$. Since the operator $\mathbf{G}$ is closed, we find

$$\mathbf{G}\mathbf{f} = \sum_{m',n'} \widehat{\mathbf{f}}(m',n') \mathbf{G} F_{m',n'}\,.$$

We conclude

$$\widehat{\mathbf{G}\mathbf{f}}(m,n) = \sum_{m',n'} \widehat{\mathbf{G}F_{m',n'}}(m,n)\widehat{\mathbf{f}}(m',n').$$

In the following we show $\widehat{\mathbf{G}F_{m',n'}}(m,n) = \Gamma(m,n\,;\,m',n')$, which will complete the proof. Using (11), we explicitly compute $\mathbf{G}F_{m',n'}$.

$$\begin{aligned}\mathbf{G}F_{m',n'}(\theta,x) &= \frac{1}{2}\varepsilon^2\Delta_x F_{m',n'}(\theta,x) - \nabla_x F_{m',n'}(\theta,x)\cdot v(\theta,x) - \partial_\theta F_{m',n'}(\theta,x)\cdot\Psi(\theta)\\ &= \left(\frac{1}{2}\varepsilon^2(2\pi\left\|n'\right\|_e)^2 - 2\pi i(n'\cdot v(\theta,x)) - 2\pi i(m'\cdot\alpha)\right)F_{m',n'}(\theta,x).\end{aligned}$$

From this formula, we can extract the Fourier coefficients, which are made up of three summands

$$\begin{aligned}\widehat{\mathbf{G}F_{m',n'}}(m,n) &= \langle F_{m,n}, \mathbf{G}F_{m',n'}\rangle\\ &= \frac{1}{2}\varepsilon^2(2\pi\left\|n'\right\|_e)^2\langle F_{m,n}, F_{m',n'}\rangle\\ &\quad - 2\pi i(m'\cdot\alpha)\langle F_{m,n}, F_{m',n'}\rangle\\ &\quad - 2\pi i\langle F_{m,n}, (n'\cdot v)F_{m',n'}\rangle.\end{aligned} \tag{23}$$

Since the Fourier functions form an orthonormal basis, $\langle F_{m,n}, F_{m',n'}\rangle$ is 1 if $(m,n) = (m',n')$ and 0 otherwise. Therefore, the first two summands are 0 unless $(m,n) = (m',n')$. In that case, we can replace m' and n' by m and n. We compute the third summand

$$\begin{aligned}-2\pi i\langle F_{m,n}, (n'\cdot v)F_{m',n'}\rangle &= -2\pi i\langle F_{m-m',n-n'}, n'\cdot v\rangle\\ &= -2\pi i(n'\cdot\langle F_{m-m',n-n'}, v\rangle)\\ &= -2\pi i(n'\cdot\widehat{v}(m-m',n-n'))\\ &= -2\pi i(n\cdot\widehat{v}(m-m',n-n')),\end{aligned} \tag{24}$$

where in the last line we used Lemma 5.1. Inserting (24) into (23) shows that $\widehat{\mathbf{G}F_{m',n'}}(m,n) = \Gamma(m,n\,;\,m',n')$ as defined in (22). □

Remark 5.3. An off-diagonal entry of $\Gamma(m,n\,;\,m',n')$ is nonzero if and only if $n\cdot\widehat{v}(m-m',n-n') \neq 0$. By Lemma 5.1, this is equivalent to $n'\cdot\widehat{v}(m-m',n-n') \neq 0$. Hence, if either $n=0$ or $n'=0$, then $\Gamma(m,n\,;\,m',n') = 0$. This shows that Fourier modes $F_{m,n}$ with $n=0$ do not "communicate" with other Fourier modes under Γ. This justifies including these Fourier modes in Γ, even though in Section 3 the augmented generator $\mathbf{G}$ was constructed over functions $\mathbf{f}$ whose fibres $\mathbf{f}(\theta)$ lie in $\mathcal{H}_0$. In particular, any Fourier mode $F_{m,0}$ is an eigenfunction of $\mathbf{G}$ with purely imaginary eigenvalue (essentially owing to the constant function in physical space being invariant under $\mathcal{P}_\theta^t$ and Fourier modes on the base being eigenfunctions of the ergodic torus rotation). We include these Fourier modes to visualize the spectral gap between ϱ and 0 in the examples of Section 6.

5.2 The discrete augmented generator

We discretize the generator $\mathbf{G}$ by a Galerkin projection onto a chosen set of Fourier modes, e.g. those with m, n close to 0. Let $S_d \subset \mathbb{Z}^{d_d}$ and $S_p \subset \mathbb{Z}^{d_p}$ be a selection of Fourier modes in the driving space and in physical space, respectively. To have the discrete operator mapping real-valued functions to real-valued ones again, Lemma 5.1 suggests to choose the sets S_d and S_p such that $m \in S_d$ if and only if $-m \in S_d$ and $n \in S_p$ if and only if $-n \in S_p$. We project onto the Fourier modes given by

$$S := S_d \times S_p \subset \mathbb{Z}^{d_d} \times \mathbb{Z}^{d_p}.$$

A natural choice for S_d, respective S_p, are all modes in $\mathbb{Z}^{d_d}$, respective $\mathbb{Z}^{d_p}$, that are within a given radius of the origin, with respect to either the Euclidean or ℓ^∞-norm. In two of the examples of Section 6, we fix constants $K \in \mathbb{N}$ and $r > 0$ and define

$$S_d := \{m \in \mathbb{Z}^{d_d} \mid \|m\|_\infty \leq K\}, \quad S_p := \{n \in \mathbb{Z}^{d_p} \mid \|n\|_e \leq r\}. \tag{25}$$

Given a set $S \subset \mathbb{Z}^{d_d} \times \mathbb{Z}^{d_p}$, we obtain the discrete generator $\Gamma_S \in \mathbb{C}^{S \times S}$ as the restriction of the matrix Γ from Proposition 5.2 to the Fourier modes in S. We can write Γ_S as the sum of two diagonal and a skew-Hermitian matrix

$$\Gamma_S = D + R + A. \tag{26}$$

The matrices $D, R \in \mathbb{C}^{S \times S}$ are diagonal and given by

$$D = \operatorname{diag}\left(\frac{1}{2}\varepsilon^2 (2\pi \|n\|_e)^2 \mid (m,n) \in S\right),$$
$$R = \operatorname{diag}\left(-2\pi i (m \cdot \alpha) \mid (m,n) \in S\right).$$

Note that D is real while R is purely imaginary. The matrix $A \in \mathbb{C}^{S \times S}$ is given by

$$A(m,n\,;\,m',n') = -2\pi i (n \cdot \widehat{v}(m-m', n-n')). \tag{27}$$

Using Lemma 5.1, one can verify that A is skew-hermitian, i.e., $A(m',n';m,n) = -\overline{A(m,n\,;\,m',n')}$.

The matrix $\Gamma_S \in \mathbb{C}^{S \times S}$ is the Galerkin projection onto the subspace

$$V_S := \operatorname{span}\{F_{m,n} \mid (m,n) \in S\}.$$

The orthogonal projection $\Pi_S : L^2(\mathbb{T}^{d_d} \times \mathbb{T}^{d_p}, \mathbb{C}) \to L^2(\mathbb{T}^{d_d} \times \mathbb{T}^{d_p}, \mathbb{C})$ onto the subspace S can be expressed as a concatenation $\Pi_S = P_S^* P_S$. Define the linear operator $P_S : L^2(\mathbb{T}^{d_d} \times \mathbb{T}^{d_p}, \mathbb{C}) \to \mathbb{C}^S$ by

$$(P_S \mathbf{f})_{m,n} = \widehat{\mathbf{f}}(m,n),$$

where $(m,n) \in S$. Let $\widehat{\mathbf{f}}_S \in \mathbb{C}^S$. The adjoint $P_S^* : \mathbb{C}^S \to L^2(\mathbb{T}^{d_d} \times \mathbb{T}^{d_p}, \mathbb{C})$ is given by

$$P_S^* \widehat{\mathbf{f}}_S = \sum_{(m,n) \in S} \widehat{\mathbf{f}}_S(m,n) F_{m,n}.$$

Note that $P_S^* P_S$ is the identity on $\mathbb{C}^S$. The matrix Γ_S is a Galerkin projection of augmented generator $\mathbf{G}$ in the sense that

$$\Gamma_S = P_S \mathbf{G} P_S^*.$$

For any $t \geq 0$ we obtain a finite-dimensional approximation of the Mather operator $\mathbf{M}^t$ given by $\exp(t\Gamma_S)$. In general, projection and exponentiation will not commute, hence $P_S \mathbf{M}^t P_S^* \neq \exp(t\Gamma_S)$.

5.3 Coherent sets

Let $\Gamma_S \in \mathbb{C}^{S \times S}$ be a finite-dimensional approximation of the augmented generator $\mathbf{G}$, like introduced in the previous section. Let $\widehat{\mathbf{f}}_S \in \mathbb{C}^S$ be an eigenfunction of Γ_S with eigenvalue $z = \lambda + \eta i \in \mathbb{C}$. Consequently, $\widehat{\mathbf{f}}_S$ is an eigenfunction of the discrete Mather operator $\exp(t\Gamma_S)$ with eigenvalue e^{tz}. Recall the decomposition $\Gamma_S = D + R + A$ from (26). Since $A + R$ is skew-Hermitian and D is diagonal with real, non-positive entries, all eigenvalues of Γ_S have non-positive real part. The unit vector $e_{(0,0)} \in \mathbb{C}^S$ corresponding to the Fourier mode $F_{0,0}$ is in the kernel of Γ_S. In fact, all unit vector $e_{(m,0)} \in \mathbb{C}^S$ are eigenfunctions of Γ_S with purely imaginary eigenvalues. A finite-dimensional vector $\widehat{\mathbf{f}}_S \in \mathbb{C}^S$ represents a function $\mathbf{f} := P_S^* \widehat{\mathbf{f}}_S \in L^2(\mathbb{T}^{d_d} \times \mathbb{T}^{d_p}, \mathbb{C})$. For the computation of coherent sets, we are only interested in those functions $\mathbf{f}$ whose fibres $\mathbf{f}(\theta)$ integrate to 0. Hence, we only consider eigenfunctions $\widehat{\mathbf{f}}_S$ of Γ_S to eigenvalues with negative real part $\lambda < 0$. We extract coherent sets from $\mathbf{f}$ as if it was an eigenfunction of $\mathbf{G}$.

In Section 4, we presented methods to compute coherent sets from from functions $\mathbf{f}$ that are either (approximate) eigenfunction of the generator $\mathbf{G}$ or that lie in the spectral set corresponding to an annulus in $\sigma(\mathbf{M}^t)$ with large radius. In the finite-dimensional case, any approximate eigenvalue is an eigenvalue and every spectral set contains at least one eigenfunction. Hence, heuristically, all three methods of Section 4 apply to the function $\mathbf{f} = P_S^* \widehat{\mathbf{f}}_S$. In all three methods, we defined a coherent set $A_\theta^\bullet$ as the positive support of $\mathrm{Re}(\mathcal{P}_\theta^t \mathbf{f}(\theta))$. We treat $\mathbf{f}$ as if it was an eigenfunction of $\mathbf{G}$ and assume that $\mathcal{P}_\theta^t \mathbf{f}(\theta) = e^{zti} \mathbf{f}(\phi^t \theta)$. Hence, we define coherent sets as the positive level sets $\mathrm{Re}(e^{\eta t i} \mathbf{f}(\phi^t \theta))$. In practice, other methods to extract almost-invariant sets from an eigenfunction have also proven successful [Fro05]. In the numerical example of Section 6 we implement three different methods to extract coherent sets from an eigenfunction $\mathbf{f}$ of the generator $\mathbf{G}$.

(CS1) $A_\theta^t := \{x \in M \mid \mathrm{Re}(e^{\eta t i} \mathbf{f}(\phi^t \theta, x)) > 0\}$.

(CS2) Fix $q > 0$. Define $A_\theta^t := \{x \in M \mid \|\mathrm{Re}(e^{\eta t i} \mathbf{f}(\phi^t \theta))\|_1^{-1} |\mathrm{Re}(e^{\eta t i} \mathbf{f}(\phi^t \theta, x))| > q\}$.

(CS3) Fix $q > 0$. Define $A_\theta^t := \{x \in M \mid \|\mathbf{f}(\phi^t \theta)\|_1^{-1} |\mathbf{f}(\phi^t \theta, x)| > q\}$.

Note that due to the normalization applied, the choice $q = 1$ is natural in (CS2) and (CS3). The first method is the one described in (16). The second method is almost identical, but instead of identifying the positive support of $\mathrm{Re}(e^{\eta t i} \mathbf{f}(\phi^t \theta, x))$ as part of the coherent set, we choose the regions in which $e^{\eta t i} \mathbf{f}(\phi^t \theta, x)$ has a large absolute real part. Thus, the resulting sets A_θ^t consist of multiple smaller regions, rather than one large region. Additionally, unlike

for the first method, the complement $B_\theta^t = M \setminus A_\theta^t$ is in general not coherent. The third method is similar to the second, but instead of the absolute real part, the absolute value in $\mathbb{C}$ is considered. Note that the factor $e^{\eta t i}$ is omitted, as it does not influence the absolute value. The third method is special in the sense that the resulting coherent sets satisfy $A_\theta^t = A_{\phi^t\theta}^0$. Hence, the resulting coherent set is of the form $A_\bullet = (A_\theta)_{\theta\in\Theta}$ with

$$A_\theta := \{x \in M \mid \|\mathbf{f}(\theta)\|_1^{-1} |\mathbf{f}(\theta, x)| > q\}.$$

The third method is to be used with caution. It may happen that $|\mathbf{f}(\theta, x)| = 1$ for all $\theta \in \Theta$ and $x \in M$ such that A_θ is either empty or the entire set M, thus rendering the method futile. In Section 6, we present two examples for which the third method has proven successful.

Given an eigenfunction $\widehat{\mathbf{f}}_S$ of Γ_S and having chosen one of the three methods described above, we can determine whether a given point $x \in M$ lies inside or outside the set A_θ^t by evaluating $\mathbf{f}(\phi^t\theta)$ in the point x. For fixed starting parameter $\theta \in \Theta$, this allows us to estimate the survival probability of the coherent set $A_\theta^\bullet$ experimentally. At time $t = 0$, we initialize a large number of particles that are uniformly distributed in M. As time progresses, the parameter θ evolves according to Ψ, while the particles evolve according to the nonautonomous vector field v. In each time frame $t > 0$ of the simulation, we evaluate $\mathbf{f}(\phi^t\theta)$ in the positions of each of the particles and determine whether they lie inside or outside of A_θ^t. We obtain an estimate of the survival probability of $A_\theta^\bullet$ up to time t by computing the fraction of particles that have remained inside the coherent set up to time t divided by the number of particles that started inside A_θ^0.

5.4 Computational aspects

In this section, we discuss aspects such as how to use sparsity of Γ_S to shorten the computational time it takes to construct Γ_S and to compute its eigenvalues with largest real part.

Recall the representation $\Gamma_S = D + R + A$ as the sum of two diagonal, and one skew-Hermitian matrix, cf. (26)–(27). The matrix D is diagonal with real, nonpositive entries; R is diagonal with purely imaginary entries; and A is skew-Hermitian. Therefore, the off-diagonal entries of Γ_S are skew-Hermitian, i.e. they satisfy $\Gamma_S(m, n\,;\, m', n') = \overline{\Gamma_S(m', n'\,;\, m, n)}$ for all $(m, n) \neq (m', n') \in S$. Hence, it suffices to compute the upper-right half of Γ_S. From (27), it follows that an off-diagonal entry $\Gamma_S(m, n\,;\, m', n')$ is nonzero if and only if

$$n' \cdot \widehat{v}(m - m', n - n') = n \cdot \widehat{v}(m - m', n - n') \neq 0.$$

The fact that n can be exchanged for n' follows from Lemma 5.1. If v only has a small number of nonzero Fourier modes, then Γ_S only has a small number of entries in each column, which leads to a high degree of sparsity. The vector field v is assumed to be smooth, and smooth functions are well-approximated by a small number of Fourier modes. Hence, we set $\widehat{v}(m, n) = 0$ whenever $|\widehat{v}_i(m, n)|$ is below a given threshold $err > 0$ for all $1 \leq i \leq d_p$. In the examples of the next section, we used a threshold of $err = 10^{-4}$. Using this threshold, no vector field considered in Section 6 had more than 150 nonzero Fourier modes.

We compute the eigenvalues of Γ_S using the MATLAB function `eigs`, which approximates the k smallest magnitude eigenvalues of a given matrix. This function is based on matrix inverse

iterations which requires a shift if Γ_S is singular. Hence, we perform an eigenvalue shift by 1. We use `eigs` to compute the k smallest magnitude eigenvalues of $\Gamma_S - \mathrm{Id}$ and afterwards add 1 to all computed eigenvalues. This effectively computes the k eigenvalues of Γ_S that are closest to $1 \in \mathbb{C}$. The shift 1 is chosen if the spectral gap of Γ is unknown. To obtain eigenvalues with more diverse real parts, the heuristic of taking the upper bound $|\varrho|$ on the spectral gap from Proposition 3.3 suggests itself. Note that $|\varrho| \propto \frac{\varepsilon^2}{2}$.

Using these methods, we are able to construct Γ_S and compute 100 eigenvalues around the origin on a conventional computer within minutes for sets S with more than 30.000 entries.

6 Examples

We apply the numerical methods described in the previous section to three different types of nonautonomous dynamics. In all of the following examples, we consider $\Theta = \mathbb{T}^2$ and $M = \mathbb{T}^2$. Fix $\alpha \in \mathbb{R}^2$ that is rationally independent. Throughout, the dynamic on Θ is given by the quasi-periodic torus rotation by α.

The first two examples consist of a 2×2 grid of counter-rotating gyres that is translated in space by an offset that depends on θ. The third example consists of two shears of oscillating strength. The first example is special in the sense that it is dynamically equivalent to an autonomous dynamics. Therefore, we are able to derive the existence of exponential dichotomies and of eigenfunctions of the augmented generator $\mathbf{G}$. For the second example, the existence of eigenfunctions of $\mathbf{G}$ is no longer given. However, our methods are still able to determine coherent sets that have slowly decaying survival probabilities. In the third example, there are no clear coherent structures apparent. However, using our methods we are able to determine coherent sets whose survival probabilities are close to the theoretical bound.

6.1 Translated gyres

Consider the autonomous vector field on $M = \mathbb{T}^2$,

$$v_{\mathrm{aut}}(x) = \begin{pmatrix} \cos(2\pi x_1)\sin(2\pi x_2) \\ -\sin(2\pi x_1)\cos(2\pi x_2) \end{pmatrix}. \tag{28}$$

This vector field corresponds to the stream function $\psi(x) = \frac{1}{2\pi}\cos(2\pi x_1)\cos(2\pi x_2)$, and yields to a 2×2 grid of rotating gyres, cf. Figure 3. We construct a nonautonomous dynamics by translating v_{aut} in space by $\theta \in \mathbb{T}^2$. Define

$$v(\theta, x) := v_{\mathrm{aut}}(x + \theta). \tag{29}$$

The dynamics on the parameter space $\Theta = \mathbb{T}^2$ is given by a quasi-periodic torus rotation by $\alpha \in \mathbb{R}^2$.

Conjugacy to an autonomous flow. We claim that the vector field v is dynamically equivalent to the vector field v_{aut} with a constant drift in direction α, i.e.

$$v_{\mathrm{eff}}(x) = v_{\mathrm{aut}}(x) + \alpha.$$

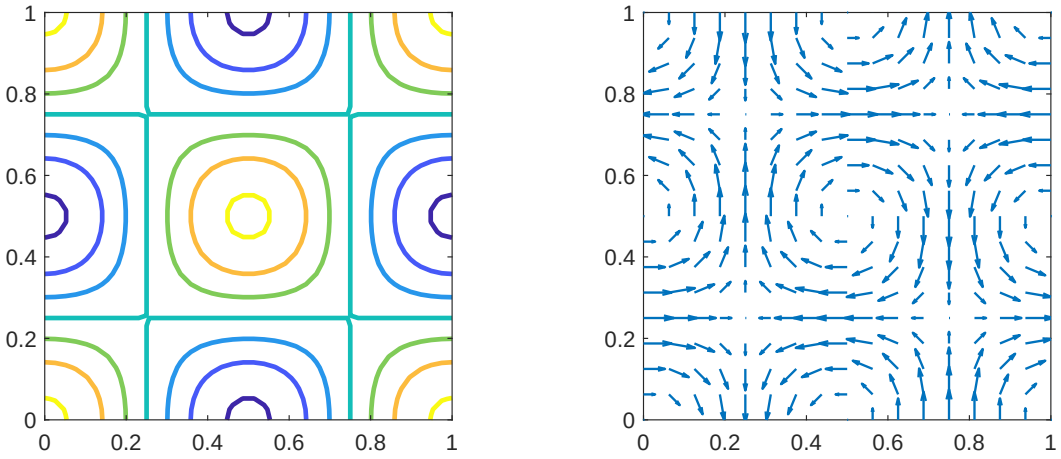

Figure 3: Left: Stream function $\psi(x) = \frac{1}{2\pi}\cos(2\pi x_1)\cos(2\pi x_2)$. Right: The corresponding vector field v_{aut}.

This dynamic equivalence can be understood as follows: As time progresses, the parameter $\theta \in \mathbb{T}^2$ moves constantly in direction α. Hence, the vector field $v(\theta, \cdot)$ is translated at a constant rate in the direction $-\alpha$. By changing the inertial frame of reference, this dynamics is equivalent to the vector field remaining constant and its associated particle motion is superimposed with a translation of constant rate in the opposite direction α.

This heuristic is rigorously described by Proposition 6.1 below. Let $\mathcal{P}_\theta^t$ be the solution operators to the Fokker–Planck equation with the nonautonomous vector field v, and let $\mathcal{P}_{\text{eff}}^t$ be the solution operators to the Fokker–Planck equation with the autonomous vector field v_{eff}. For $\beta \in \mathbb{T}^2$, let $T_\beta : \mathcal{H} \to \mathcal{H}$ be the translation by β, i.e.

$$[T_\beta f](x) = f(x - \beta).$$

Note that T_β is an isometric isomorphism with inverse $T_{-\beta}$.

Proposition 6.1. *The following diagram commutes for every $\theta \in \mathbb{T}^2$ and $t \geq 0$.*

$$\begin{array}{ccc} \mathcal{H} & \xrightarrow{\mathcal{P}_\theta^t} & \mathcal{H} \\ \downarrow{\scriptstyle T_\theta} & & \downarrow{\scriptstyle T_{\phi^t\theta}} \\ \mathcal{H} & \xrightarrow{\mathcal{P}_{\text{eff}}^t} & \mathcal{H} \end{array} \tag{30}$$

Proof. See Appendix A. □

Mather and Sacker–Sell spectra. The fact that the cocycle $\mathcal{P}$ is cohomologous to the autonomous cocycle $\mathcal{P}_{\text{eff}}$ implies that the Mather operators $\mathbf{M}^t$ admit eigenfunctions and that the Sacker–Sell spectrum $\Sigma(\mathcal{P})$ consists of discrete points.

Recall Section 2.5, in which we studied autonomous cocycles, i.e. one-parameter semigroups. The Sacker–Sell spectrum of $\mathcal{P}_{\text{eff}}$ consists of discrete values $\Sigma(\mathcal{P}_{\text{eff}}) = \{\lambda_k \mid 1 \leq k \leq N\}$, for

$N \in \mathbb{N}_0 \cup \{\infty\}$. For each λ_k there is $z_k \in \mathbb{C}$ with $|z_k| = \lambda_k$ and a function $f \in \mathcal{H}_0$ such that $\mathcal{P}^t_{\text{eff}} f = e^{z_k t} f$ for all $t \geq 0$. Let f be such an eigenfunction. Define a function $\mathbf{f} \in L^2(\mathbb{T}^2, \mathcal{H})$ by

$$\mathbf{f}(\theta) = T_{-\theta} f.$$

The commutative diagram (30) yields

$$\begin{aligned}[\mathbf{M}^t \mathbf{f}](\theta) &= \mathcal{P}^t_{\phi^{-t}\theta} \mathbf{f}(\phi^{-t}\theta) = (T_{-\theta} \circ \mathcal{P}^t_{\text{eff}} \circ T_{\phi^{-t}\theta}) T_{-\phi^{-t}\theta} f \\ &= e^{z_k t} T_{-\theta} f = e^{z_k t} \mathbf{f}(\theta).\end{aligned}$$

This shows that $\mathbf{f}$ is an eigenfunction of the Mather semigroup with eigenvalue z_k.

Proposition 6.2. *We find* $\Sigma(\mathcal{P}) = \Sigma(\mathcal{P}_{\mathit{eff}}) = \{\lambda_k \mid 1 \leq k \leq N\}$.

Proof. See Appendix A. □

By Theorem 2.6, the spectrum of the augmented generator $\mathbf{G}$ is given by lines parallel to the imaginary axis with real parts λ_k. The spectrum of the discretized generator Γ_S, which we will construct next, is compatible with this structure seen in Figure 4.

Discrete generator. We construct a numerical approximation $\Gamma_S \in \mathbb{C}^{S \times S}$ of the augmented generator $\mathbf{G}$, as described in Section 5. This requires us to select a set of Fourier modes $S \subset \mathbb{Z}^4$. We use the sparsity structure of Γ to choose a suited set S. Recall that an off-diagonal entry of Γ is only nonzero if $n \cdot \widehat{v}(m - m', n - n') \neq 0$, cf. (22). The vector field v, defined by (29), only has four nonzero Fourier modes $\widehat{v}(m, n)$. Since $m, n \in \mathbb{Z}^2$, we write them as $\widehat{v}(m_1, m_2, n_1, n_2)$. The four nonzero Fourier modes of v are

$$\widehat{v}(+1,+1,+1,+1) = \frac{1}{4} i \begin{pmatrix} -1 \\ +1 \end{pmatrix}, \qquad \widehat{v}(+1,-1,+1,-1) = \frac{1}{4} i \begin{pmatrix} +1 \\ +1 \end{pmatrix},$$
$$\widehat{v}(-1,+1,-1,+1) = \frac{1}{4} i \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \qquad \widehat{v}(-1,-1,-1,-1) = \frac{1}{4} i \begin{pmatrix} +1 \\ -1 \end{pmatrix}.$$

The fact that these are the only nonzero Fourier modes implies that Γ has a block diagonal structure after suitably reordering the indices, i.e. there are sets $\mathcal{I}_k \subset \mathbb{Z}^4$ such that $\Gamma(\mathcal{I}_k; \mathcal{I}_\ell) = \Gamma(\mathcal{I}_\ell; \mathcal{I}_k) = 0$ for any $k \neq \ell$. More precisely, an off-diagonal entry $\Gamma(m, n; m', n')$ that is nonzero requires $m - m' = n - n' = (\pm 1, \pm 1)$, and in particular, $m - n = m' - n'$. The requirement $m_1 - n_1 = m_1' - n_1'$ and $m_2 - n_2 = m_2' - n_2'$ suggests to define for $(k, \ell) \in \mathbb{Z}^2$ the class

$$\mathcal{I}_{k,\ell} := \{(n_1 + k, n_2 + \ell, n_1, n_2) \mid (n_1, n_2) \in \mathbb{Z}^2\} \subset \mathbb{Z}^4. \tag{31}$$

Now for two different classes $\mathcal{I}, \mathcal{J} \subset \mathbb{Z}^4$ we obtain $\Gamma(\mathcal{I}; \mathcal{J}) = \Gamma(\mathcal{J}; \mathcal{I}) = 0$. We note that Γ_S can be decomposed along an even finer set of equivalence classes.

Note that the blocks of Γ are of infinite size. However, after choosing $S \subset \mathbb{Z}^4$, the resulting finite-dimensional matrix $\Gamma_S \in \mathbb{C}^{S \times S}$ inherits a block diagonal structure from Γ. Hence, the spectrum of Γ_S and the corresponding eigenfunctions are given by the spectrum and the

eigenfunctions of each of the individual blocks. It is desirable to choose $S \subset \mathbb{Z}^4$ such that Γ_S contains few diagonal block, each of high resolution, rather than many diagonal blocks of low resolution. We define S in the following way: Fix $K \in \mathbb{N}$ and $r > 0$. Define

$$S := \{(n_1 + k, n_2 + \ell, n_1, n_2) \in \mathbb{Z}^4 \mid k, \ell \in \{-K, \dots, K\},\ \|(n_1, n_2)\|_e \leq r\}. \tag{32}$$

The set S intersects $(2K+1)^2$ equivalence classes of the type (31). The resolution of each of the classes is determined by the radius r.

Results. In the following we present experimental results for the described system. We choose $\alpha = 0.2(1, \sqrt{2})$ as the direction of the quasiperiodic driving and $\varepsilon = 0.03$ as the diffusion constant. The set S is constructed like described in (32) with constants $K = 2$ and $r = 11$. This results in a set $S \subset \mathbb{Z}^4$ with 9425 elements. We compute the discrete generator Γ_S, which is a sparse matrix with 0.04% nonzero entries. In Figure 4 we show the spectrum of Γ_S around the origin, computed as described in Section 5.4. Note that the spectrum of Γ_S is distributed on lines parallel to the imaginary axis. This fits to the description of the spectrum of $\mathbf{G}$, which we showed to consist of discrete lines parallel to the imaginary axis. We pick an eigenvalue $z = -0.089 - 1.041i$ of Γ_S of largest nonzero real part. Hence, the real part is given by $\lambda = -0.089$. Let $\widehat{\mathbf{f}}_S$ be the corresponding eigenfunction. As described in Section 5.3, $\widehat{\mathbf{f}}_S$ represents a function $\mathbf{f} := P_S^* \widehat{\mathbf{f}}_S \in L^2(\mathbb{T}^2 \times \mathbb{T}^2, \mathbb{C})$. For a given $\theta \in \mathbb{T}^2$, we compute a coherent set $A_\theta^\bullet$ based on the real part of the fibres of $\mathbf{f}$, cf. method (CS2). Fix the threshold $q = 1$ and define

$$A_\theta^t := \{x \in \mathbb{T}^2 \mid \left\|\mathrm{Re}(e^{\eta t i}\mathbf{f}(\phi^t\theta))\right\|_1^{-1} \left|\mathrm{Re}(e^{\eta t i}\mathbf{f}(\phi^t\theta, x))\right| > q\}.$$

We estimate the survival probability for the starting parameter $\theta = (0, 0)$. At time $t = 0$, we fill the domain $\mathbb{T}^2$ with a 150×150 grid of particles and let them evolve in time according to the SDE (5) with the nonautonomous vectorfield v. For each time $t > 0$, we obtain an experimental estimate on the survival probability of $A_\theta^\bullet$ up to time t, as described in Section 5.3. In Figure 5, we show snapshots of the simulation for three selected times. In Figure 4, there is a plot of the experimental survival probability up to time $t = 10$. From (17), we expect that the survival probability of $A_\theta^\bullet$ decays at most at rate 2λ, up to constants. The results in Figure 4 agree with this estimate. For t close to 0, the survival probability decays quickly but for later times t, the decay of the survival probability is significantly slower than $e^{2\lambda t}$. The computed cumulative survival probability is $C(A_\theta^\bullet) = 6.070$. This is more than the estimate $-\frac{1}{2\lambda} = 5.682$ from (18). Determining coherent sets based on the absolute value of the fibres $\mathbf{f}(\theta)$ rather than their absolute real part, i.e. based on method (CS3), yields almost identical results. Applying the method (CS1) also yields a coherent set with slowly decaying survival probability (not depicted here).

6.2 Oscillating gyres

This example is defined similarly to the previous one. However, unlike in the previous case, the nonautonomous vector field we construct is not dynamically equivalent to an autonomous one. Like before, consider the autonomous vector field v_{aut}, defined in (28), that consists of a

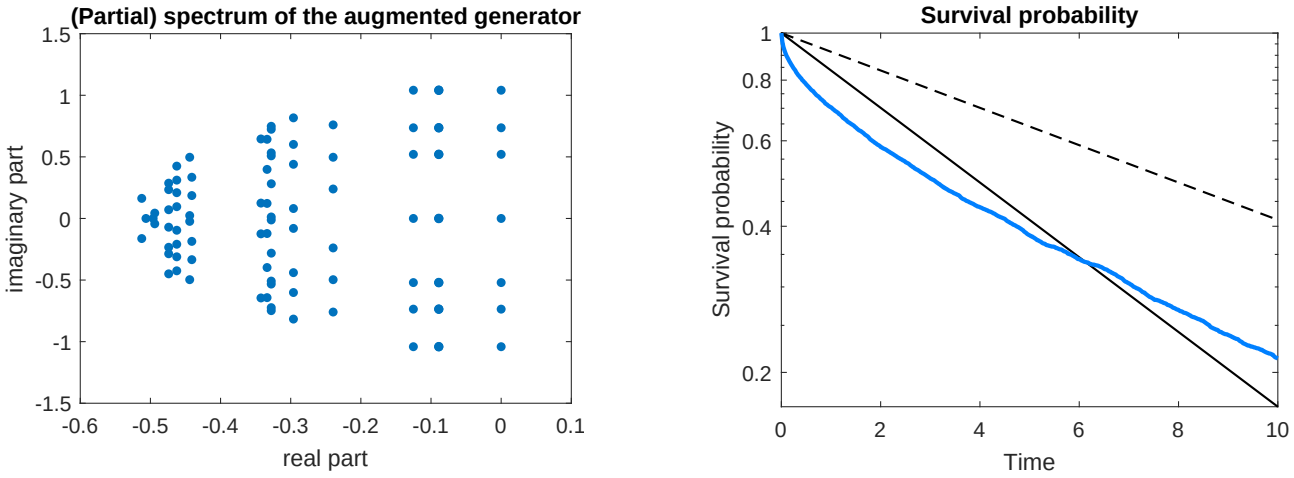

Figure 4: Translated gyre example. Left: 100 eigenvalues of the discrete generator Γ_S around the origin. Right: In blue, the simulated survival probability of the coherent sets extracted from an eigenfunction $\mathbf{f}$ to an eigenvalue z using method (CS3) with threshold $q = 1$. The black lines show the curves $e^{\lambda t}$ (dashed) and $e^{2\lambda t}$ (solid) for comparison, where λ is the real-part of z. The computed cumulative survival probability is $C(A_\theta^\bullet) = 6.070$.

2×2 grid of counter-rotating gyres, cf. Figure 3. We construct a nonautonomous vector field by letting the grid of gyres oscillate in a quasi-periodic manner. Let $\delta > 0$ be the amplitude of the oscillation. Define

$$v(\theta, x) = v_{\text{aut}}\left(x + \delta \begin{pmatrix} \sin(2\pi\theta_1) \\ \cos(2\pi\theta_2) \end{pmatrix}\right). \tag{33}$$

Since $\theta \in \mathbb{T}^2$ evolves constantly in direction $\alpha \in \mathbb{R}^2$, the grid of gyres oscillates in direction of x_1 with a period of α_1^{-1} and in direction x_2 with a period of α_2^{-1}. A trajectory of the oscillation is depicted in Figure 6.

Mather and Sacker–Sell spectra. For δ close to 0, the vector field is almost autonomous. Unlike in the previous example, this does not guarantee the existence of eigenfunctions of the Mather operator $\mathbf{M}^t$, but for δ small enough, we can prove the existence of exponential dichotomies. Let $\mathcal{P}_{\text{aut}}$ be the autonomous cocycle of transfer operators corresponding to the autonomous vector field v_{aut}. From Section 2.5, we know that the Sacker–Sell spectrum $\Sigma(\mathcal{P}_{\text{aut}}) = \{\lambda_k \mid 1 \le k \le N\}$, for $N \in \mathbb{N}_0 \cup \{\infty\}$, consists of discrete points. Let $\lambda \notin \Sigma(\mathcal{P}_{\text{aut}})$. Hence, the spectrum of the Mather operators $\mathbf{M}^t_{\text{aut}}$ of $\mathcal{P}_{\text{aut}}$ satisfies $\sigma(\mathbf{M}^t_{\text{aut}}) \cap e^{\lambda t}S^1 = \emptyset$. In other words, the ring $e^{\lambda t}S^1$ lies in the resolvent set of $\mathbf{M}^t_{\text{aut}}$ for each $t > 0$. By [CL99, Corollary 6.44], the nonautonomous cocycle $\mathcal{P}$ has an exponential dichotomy at λ if for some $t > 0$

$$\sup_{\theta\in\Theta} \left\|\mathcal{P}^t_{\text{aut}} - \mathcal{P}^t_\theta\right\| < \inf_{z\in e^{\lambda t}S^1} \left\|(\mathbf{M}^t_{\text{aut}} - z\mathrm{Id})^{-1}\right\|^{-1}. \tag{34}$$

For fixed $t > 0$, the right-hand side defines a constant that only depends on t and λ. As δ approaches 0, the difference between the vector fields v and v_{aut} tends to 0 uniformly on $\mathbb{T}^2$.

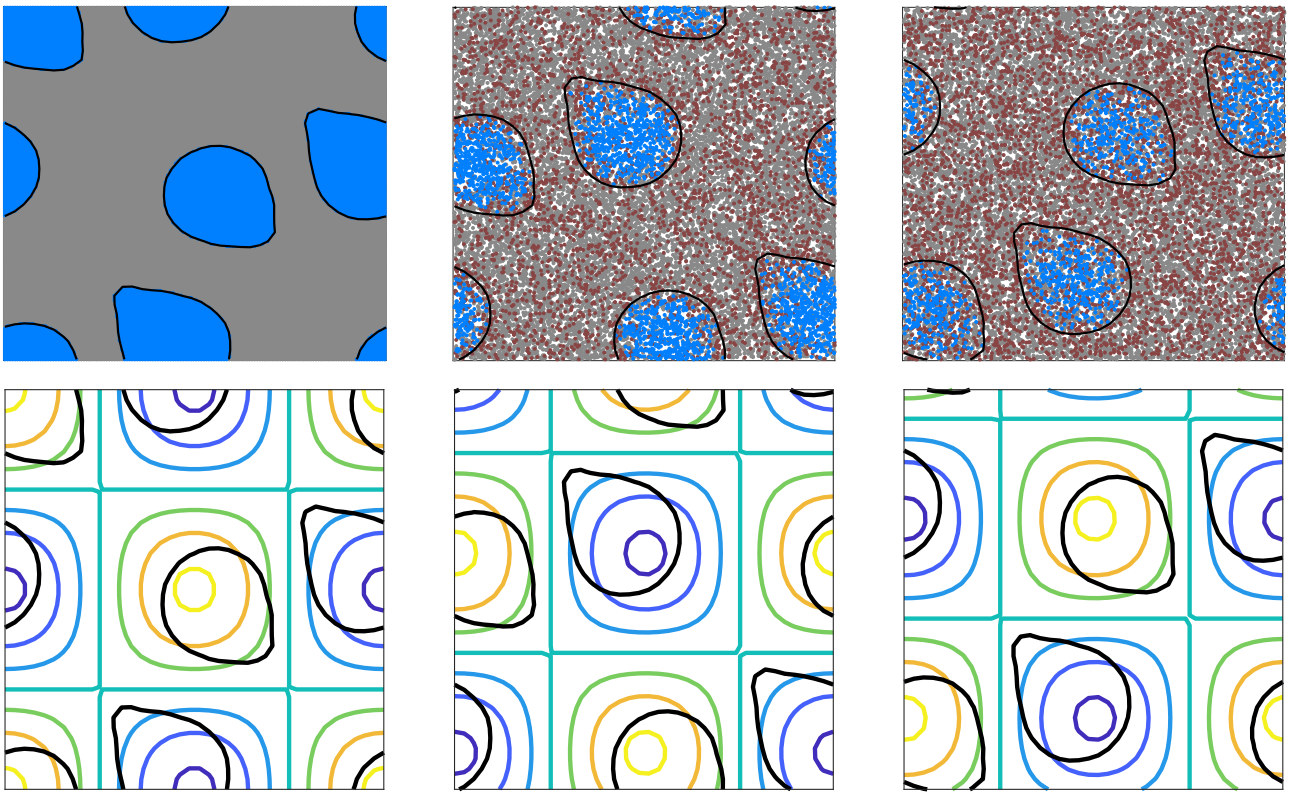

Figure 5: Translated gyre example: Evolution of the coherent sets, marked in black, determined using method (CS3) with threshold $q = 1$. The snapshots are taken at time $t = 0, 5, 10$. The top row depicts a particle simulation. Particles that started inside A_θ^0 and have remained there until time t are colored in blue. Particles that have started inside A_θ^0 but were outside the coherent set at some time $0 < s \leq t$ are colored in red. Other particles are colored in grey. The bottom row locates the coherent set with respect to the position of the gyres, as indicated by the stream function at the respective time. A video of the particle simulation for the time interval $[0, 10]$ can be found at `https://github.com/RobinChemnitz/MatherCoherent`.

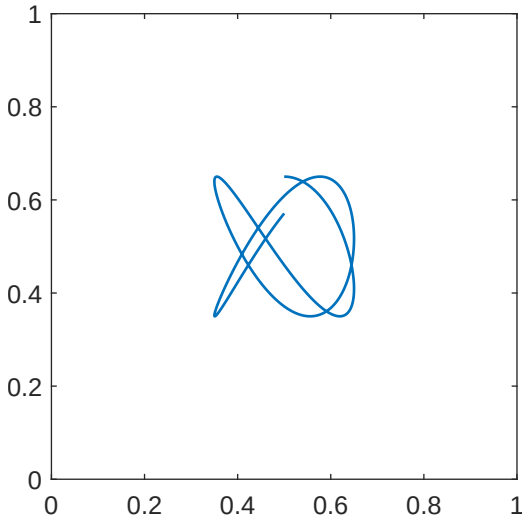

Figure 6: Oscillating gyre example: The trajectory up to time $t = 10$ of the midpoint of the central gyre of the vector field v, defined in (33) for parameters $\alpha = 0.2(1, \sqrt{2})$, $\delta = 0.15$, and strating parameter $\theta = (0, 0)$.

By Lemma 3.4, the left-hand side of (34) vanishes as $\delta \to 0$. We conclude that for sufficiently small δ, the nonautonomous cocycle $\mathcal{P}$ has an exponential dichotomy at λ and the spectrum of the augmented generator $\mathbf{G}$ does not contain the line $\{\lambda + \eta i \mid \eta \in \mathbb{R}\}$, cf. Theorem 2.6. Hence, for small values of δ, we expect the spectrum of $\mathbf{G}$ to consist of thin bands along the imaginary axis, at least for regions close to the origin. In Figure 7, we show the spectrum of the discrete generator Γ_S for $\delta = 0.15$. This value of δ seems to be not small enough such that the spectrum consists of clear lines, nevertheless, the band structure of the spectrum is recognizable in the region close to 0.

Results. In the following the present experimental results for the described system. We choose $\alpha = 0.2(1, \sqrt{2})$ as the direction of the quasiperiodic driving, $\varepsilon = 0.03$ as the diffusion constant, and $\delta = 0.15$ as the strength of the oscillation. We define the set $S \in \mathbb{Z}^4$ of Fourier modes as described in (25).

$$S := \{(m_1, m_2, n_1, n_2) \in \mathbb{Z}^4 \mid \|(m_1, m_2)\|_\infty \leq K,\ \|(n_1, n_2)\|_e \leq r\},$$

where $K = 6$ and $r = 8$. The constant K determines the resolution in the driving and r determines the resolution in physical space. The chosen values of K and r result in a set $S \subset \mathbb{Z}^4$ with 33293 elements. We compute the discrete generator Γ_S, which is a sparse matrix with 0.31% nonzero entries. In Figure 7 we show the spectrum of Γ_S around the origin, computed as described in Section 5.4. We pick a real eigenvalue $z = -0.071$ of Γ_S of largest nonzero real part. Since z is real, we have $\lambda = z$. Let $\widehat{\mathbf{f}}_S$ be the corresponding eigenfunction. As described in Section 5.3, $\widehat{\mathbf{f}}_S$ represents a function $\mathbf{f} := P_S^* \widehat{\mathbf{f}}_S \in L^2(\mathbb{T}^2 \times \mathbb{T}^2, \mathbb{C})$. For a given $\theta \in \mathbb{T}^2$, we compute a coherent set $A_\theta^\bullet$ based on the absolute value of the fibres of $\mathbf{f}$, cf. method (CS2). Fix

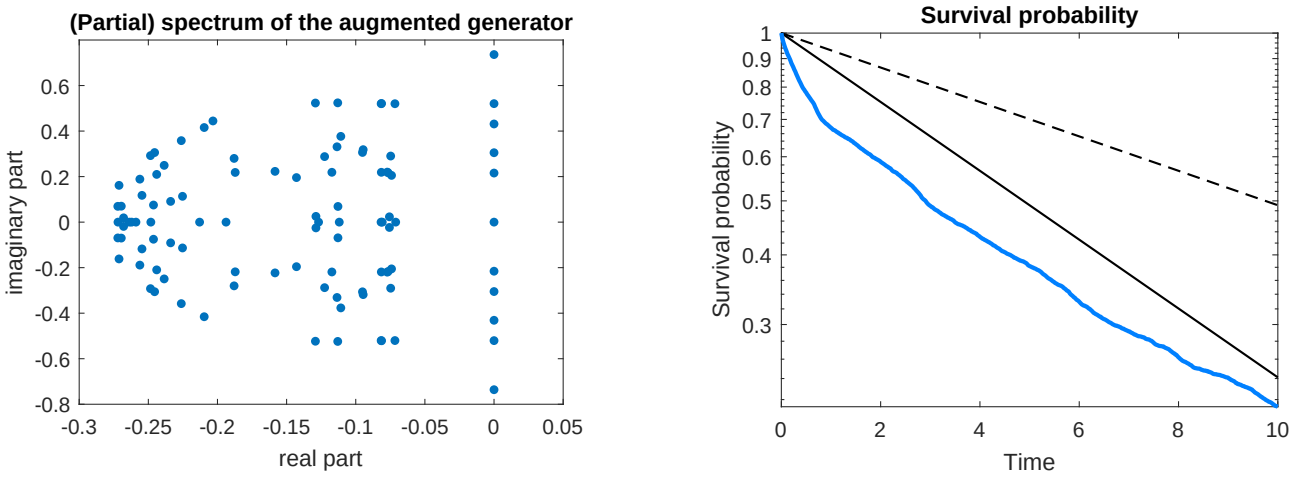

Figure 7: Oscillating gyre example. Left: 100 eigenvalues of the discrete generator Γ_S around the origin. Right: In blue, the simulated survival probability of the coherent sets determined by the absolute value of an eigenfunction $\mathbf{f}$ to an eigenvalue z with threshold $q = 1$. The black lines show the curves $e^{\lambda t}$ (dashed) and $e^{2\lambda t}$ (solid) for comparison, where λ is the real-part of z. The computed cumulative survival probability is $C(A_\theta^\bullet) = 6.158$.

the threshold $q = 1$ and define

$$A_\theta^t := \{x \in \mathbb{T}^2 \mid \left\|\mathbf{f}(\phi^t\theta)\right\|_1^{-1} \left|\mathbf{f}(\phi^t\theta, x)\right| > q\}.$$

Recall that these sets satisfy $A_\theta^t = A_{\phi^t\theta}^0$. Hence, the coherent family can be written as $A_\bullet$. We estimate the survival probability for the starting parameter $\theta = (0, 0)$. At time $t = 0$, we fill the domain $\mathbb{T}^2$ with a 150×150 grid of particles and let them evolve in time according to the SDE (5) with the nonautonomous vectorfield v. For each time $t > 0$, we obtain an experimental estimate on the survival probability of $A_\theta^\bullet$ up to time t, as described in Section 5.3. In Figure 8, we show snapshots of the simulation for three selected time points. In Figure 7, there is a plot of the experimental survival probability up to time $t = 10$. From (17), we expect that the survival probability of $A_\theta^\bullet$ decays at most at rate 2λ, up to constants. The results in Figure 7 agree with this estimate as the estimated survival probability decays very similarly to $e^{2\lambda t}$. The computed cumulative survival probability is $C(A_\theta^\bullet) = 6.158$. This is less than the estimate $-\frac{1}{2\lambda} = 7.042$ from (18). Determining coherent sets based on the absolute real part of the fibres $\mathbf{f}(\theta)$ rather than their absolute value, i.e. based on method (CS2), yields almost identical results. Applying the method (CS1) also yields a coherent set with slowly decaying survival probability (not depicted here).

6.3 Shear

In both of the previous examples, the coherent sets we identified were regions around the center of the four gyres, close to sets that could have been determined from considering the individual vector fields $v(\theta, \cdot)$ for each fixed $\theta \in \mathbb{T}^2$. The example we present next is of different nature.

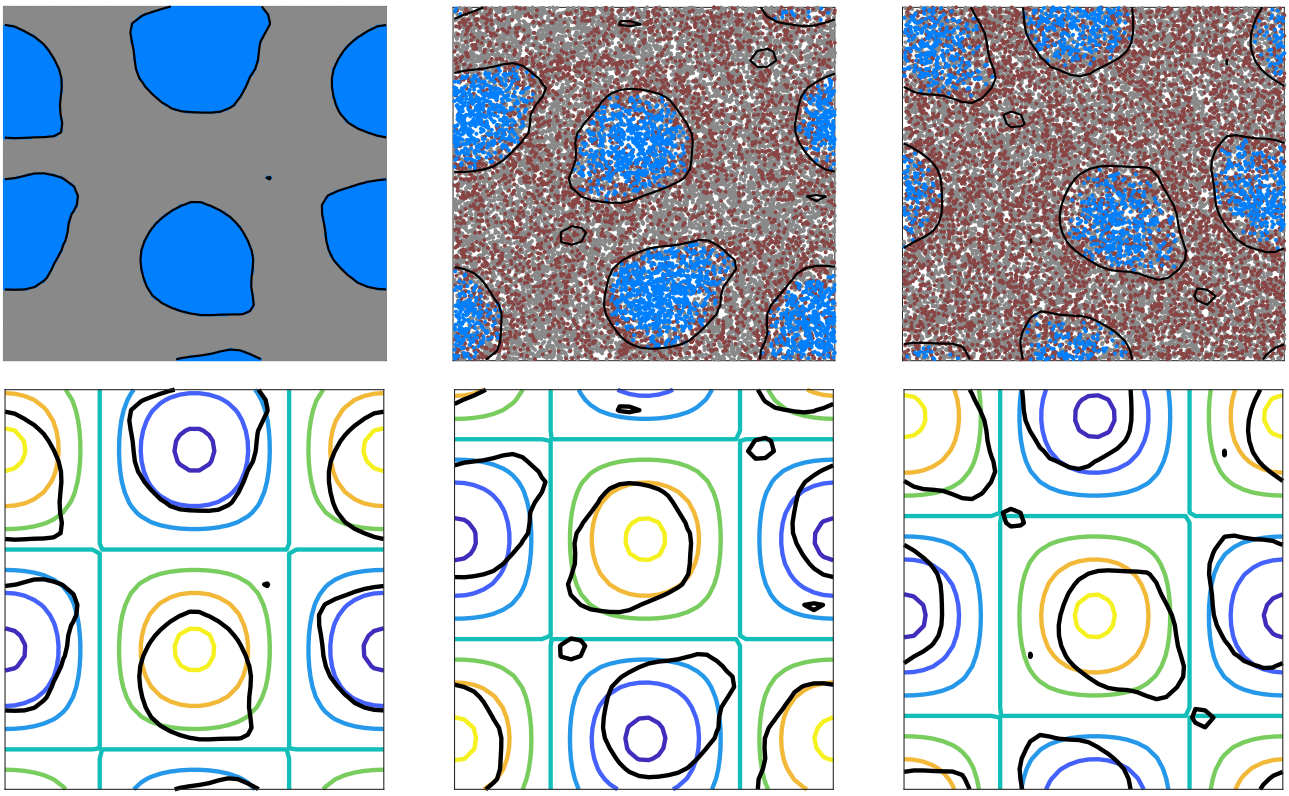

Figure 8: Oscillating gyre example: Evolution of the coherent sets, marked in black, determined using method (CS2) with threshold $q = 1$. The snapshots are taken at time $t = 0, 5, 10$. The top row depicts a particle simulation. Particles that started inside A_θ^0 and have remained there until time t are colored in blue. Particles that have started inside A_θ^0 but were outside the coherent set at some time $0 < s \leq t$ are colored in red. Other particles are colored in grey. The bottom row locates the coherent set with respect to the position of the gyres, as indicated by the stream function at the respective time. A video of the particle simulation for the time interval $[0, 10]$ can be found at `https://github.com/RobinChemnitz/MatherCoherent`.

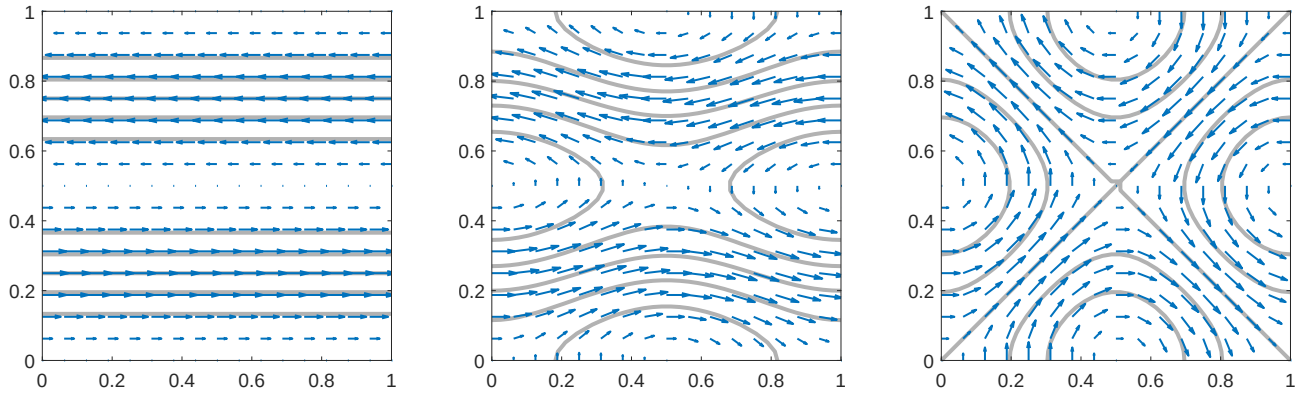

Figure 9: Shear example: The vector field defined in (35) in blue and the corresponding stream function in grey for parameters $\theta = (0.25, 0)$, $(0.25, 0.05)$, $(0.25, 0.25)$. From left to right, the strength of the shear in x-direction stays constant at 1 while the shear in y-direction increases from 0 to 1.

A well-studied dynamical feature that is known to result in chaotic behavior is shear. In this example we overlay two shears of oscillating strength, by considering the nonautonomous vector field

$$v(\theta, x) = \sin(2\pi\theta_1) \begin{pmatrix} \sin(2\pi x_2) \\ 0 \end{pmatrix} + \sin(2\pi\theta_2) \begin{pmatrix} 0 \\ \sin(2\pi x_1) \end{pmatrix}. \tag{35}$$

Figure 9 shows a representation of this vector field for selected values of θ. Determining coherent sets of v is non-trivial, since vector fields with shears are known for their mixing properties [CS89, BCZG23, FK23], while coherent sets are characterized by experiencing little mixing with their surrounding. Despite this feature, which opposes the existence of coherent sets, our methods are able to identify a coherent set whose survival probability decays at a similar rate as the survival probability of the previous example of the oscillating gyres. Note that this comparison is fair in the sense that the velocity field strength is of the same order of magnitude and below we take the same diffusion coefficient ε as in the previous examples.

Results. In the following we present experimental results for the described system. We choose $\alpha = 0.2(1, \sqrt{2})$ as the direction of the quasiperiodic driving and $\varepsilon = 0.03$ as the diffusion constant. We define the set $S \in \mathbb{Z}^4$ of Fourier modes like described in (25):

$$S := \{(m_1, m_2, n_1, n_2) \in \mathbb{Z}^4 \mid \|(m_1, m_2)\|_\infty \leq K,\ \|(n_1, n_2)\|_e \leq r\},$$

where $K = 6$ and $r = 8$. The constant K determines the resolution in the driving and r determines the resolution in physical space. The chosen values of K and r result in a set $S \subset \mathbb{Z}^4$ with 33293 elements. We compute the discrete generator Γ_S, which is a sparse matrix with 0.023% nonzero entries. In Figure 10 we show the spectrum of Γ_S around the origin, computed as described in Section 5.4. Unlike in the previous two examples, there are no recognizable bands of eigenvalues parallel to the imaginary axis. This could indicate the absence of exponential dichotomies for this example. We pick a real eigenvalue $z = -0.097$ of Γ_S with the largest

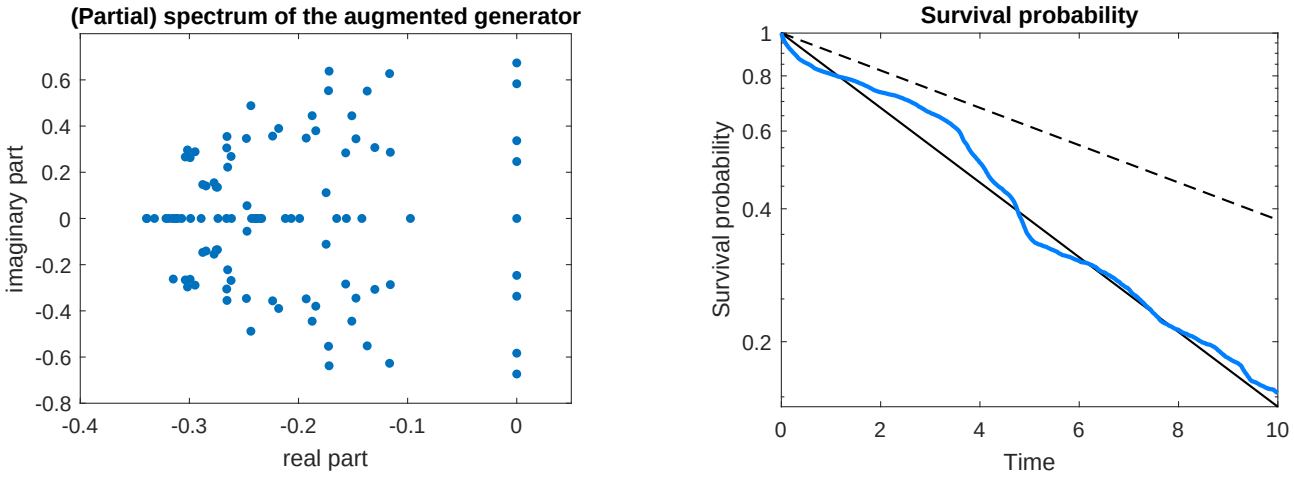

Figure 10: Shear example. Left: 100 eigenvalues of the discrete generator Γ_S around the origin. Right: In blue, the simulated survival probability of the coherent sets determined by positive support of an eigenfunction $\mathbf{f}$ to an eigenvalue z. The black lines show the curves $e^{\lambda t}$ (dashed) and $e^{2\lambda t}$ (solid) for comparison, where λ is the real-part of z. The computed cumulative survival probability is $C(A_\theta^\bullet) = 5.680$.

nonzero real part. Since z is real, we have $\lambda = -0.097$ and $\eta = 0$. Let $\widehat{\mathbf{f}}_S$ be the corresponding eigenfunction. As described in Section 5.3, $\widehat{\mathbf{f}}_S$ represents a function $\mathbf{f} := P_S^* \widehat{\mathbf{f}}_S \in L^2(\mathbb{T}^2 \times \mathbb{T}^2, \mathbb{C})$. For a given $\theta \in \mathbb{T}^2$, we compute a coherent set $A_\theta^\bullet$ based on the positive support of the fibres of $\mathbf{f}$, cf. method (CS1). Define

$$A_\theta^t := \{x \in M \mid \mathrm{Re}(e^{\eta t i} \mathbf{f}(\phi^t \theta, x)) > 0\}.$$

Since $\eta = 0$, these sets satisfy $A_\theta^t = A^0_{\phi^t \theta}$. Hence, the coherent family can be written as $A_\bullet$. We estimate the survival probability for the starting parameter $\theta = (0, 0)$. At time $t = 0$, we fill the domain $\mathbb{T}^2$ with a 150×150 grid of particles and let them evolve in time according to the SDE (5) with the nonautonomous vectorfield v. For each time $t > 0$, we obtain an experimental estimate on the survival probability of $A_\theta^\bullet$ up to time t, as described in Section 5.3. In Figure 11, we show snapshots of the simulation for three selected time instances. In Figure 10, there is a plot of the experimental survival probability up to time $t = 10$. From (17), we expect that the survival probability of $A_\theta^\bullet$ decays at most at rate 2λ, up to constants. The results in Figure 10 agree with this estimate as the estimated survival probability decays very similarly to $e^{2\lambda t}$. The computed cumulative survival probability is $C(A_\theta^\bullet) = 5.680$. This is more than the estimate $-\frac{1}{2\lambda} = 5.133$ from (18). Using method (CS2) or (CS3) did not result in coherent sets with a slowly decaying survival probability.

7 Conclusion

Our numerical implementation to compute coherent sets is mainly based on heuristic derivations and experimental evidence. There are two aspects we would like to discuss further.

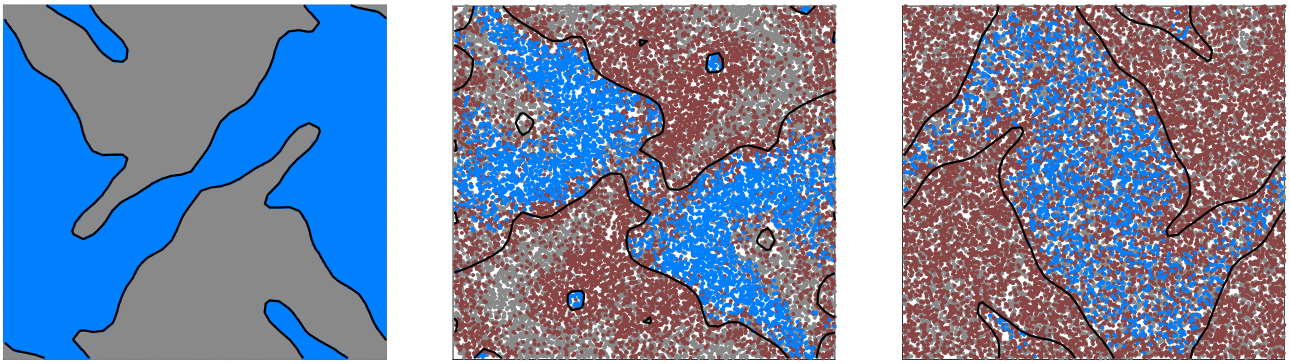

Figure 11: Shear example: Evolution of the coherent sets, marked in black, determined using method (CS1). The snapshots are taken at time $t = 0, 5, 10$. Particles that started inside A_θ^0 and have remained there until time t are colored in blue. Particles that have started inside A_θ^0 but were outside the coherent set at some time $0 < s \leq t$ are colored in red. Other particles are colored in grey. A video of the particle simulation for the time interval $[0, 10]$ can be found at `https://github.com/RobinChemnitz/MatherCoherent`.

Additionally, we provide an outlook on the application of our methods to more general flows.

Discussion. The selection of the set of Fourier modes S, which is a crucial choice for the construction of the discrete generator Γ_S, has so far been made on the basis of experimental evidence. In both examples of Section 6.2 and Section 6.3, we chose $S = S_d \times S_p \subset \mathbb{Z}^{d_d} \times \mathbb{Z}^{d_p}$ with S_d a ball of radius K in the ℓ^∞-norm and S_p a ball of radius r in the Euclidean norm. The radii that performed best in these examples, while keeping the computational effort feasible, were $K = 6$ and $r = 6$. This results in $|S_d| = 169$ and $|S_p| = 197$, i.e. the resolution in the driving and the physical space is similar. The reason we believe that the Euclidean norm works best in the physical space is that the decay of a Fourier mode $F_{m,n}$ due to diffusion scales proportionally to the Euclidean norm of n, cf. (22). Hence, Fourier modes $F_{m,n}$, where n has large Euclidean norm decay fast and are thus expected to contribute less to the eigenfunctions of Γ_S. In the driving, there is no diffusion, which is a possible explanation why the ℓ^∞-norm worked best in our cases. However, these observations have been made only for a small number of examples at a medium resolution $|S| \sim 10^5$. To apply or methods to more complex dynamics at higher resolutions, it is worth to investigate theoretically what choices of S_d and S_p lead to an optimal approximation of the spectrum of $\mathbf{G}$ by the spectrum of Γ_S.

The other aspect we discuss is the extraction of coherent sets from an eigenfunction $\widehat{\mathbf{f}}_S$ of Γ_S with eigenvalue z. In the derivation of our method, we treated $\widehat{\mathbf{f}}_S$ as if it was an eigenfunction of Γ, i.e. as if $P_S^* \widehat{\mathbf{f}}_S$ was an eigenfunction of $\mathbf{G}$. However, even for a high resolution of S, we cannot guarantee that $\|(\Gamma - z)\widehat{\mathbf{f}}_S\|_2$ is small. Hence, $P_S^* \widehat{\mathbf{f}}_S$ is a priori not even an approximate eigenfunction of $\mathbf{G}$. The reason for this is that, unlike in the physical space, there is no diffusion in the driving. Due to diffusion in physical space, an eigenfunction $\widehat{\mathbf{f}}_S$ is expected to have low Fourier coefficients $\widehat{\mathbf{f}}_S(m, n)$ for large n; however, it may have large Fourier coefficients $\widehat{\mathbf{f}}_S(m, n)$ for large values of m, but small values of n. Hence, at the boundary of S, where m is large,

but n is small, there may be a significant discrepancy between $\Gamma_S \widehat{\mathbf{f}}_S$ and $\Gamma \widehat{\mathbf{f}}_S$. Numerically our method still performed well, which could indicate that in practice this problem plays a minor role. Finding theoretical bounds for the survival probability of coherent sets extracted from eigenfunctions of the discrete generator Γ_S is a challenging task that can be the topic of future research.

Outlook. The numerical discretization of the Mather operators $\mathbf{M}^t$ and their generator $\mathbf{G}$ using a Galerkin projection onto Fourier modes required $\Theta = \mathbb{T}^{d_d}$ and $M = \mathbb{T}^{d_p}$. For simplicity, we additionally restricted ourselves to the case that the driving dynamics on Θ is given by a quasiperiodic rotation. In Section 3, the transfer operator cocycle $\mathcal{P}$ and the corresponding Mather semigroup have been constructed for more general domains Θ and M than just tori, and for general ergodic driving. Our results can be viewed as a proof of concept that coherent sets of a nonautonomous dynamics with ergodic driving of the form (5) can be extracted from eigenfunctions of a discretization of the augmented generator $\mathbf{G}$. For domains Θ and M which are not tori, other discretization methods, like an Ulam discretization, are needed. Hybrid methods are possible as well [GD20], such as using a Fourier discretization in Θ and an Ulam discretization in M [FK17]. A significant advantage of the Fourier discretization in the case $\Theta = \mathbb{T}^{d_d}$ and $M = \mathbb{T}^{d_p}$ is that a finite-dimensional vector $\widehat{\mathbf{f}}_S \in \mathbb{C}^S$ represents a function $\mathbf{f} := P_S^* \widehat{\mathbf{f}}_S \in L^2(\Theta \times M, \mathbb{C})$ in an infinite-dimensional space. Hence, coherent sets $A_\bullet^\bullet$ extracted from $\mathbf{f}$, as described in Section 5.3, have a continuous resolution in both θ and physical space.

Recently, methods have been proposed to faithfully approximate (parts of) the spectrum of transfer-type operators, such as the Koopman operator, see [DG19, KPM20, Gia21, CT24, Col23, VG23]. Since some of these methods explicitly aim at approximating continuous spectra and associated spectral projections, their adaptation to the augmented generator $\mathbf{G}$ could be an interesting future avenue.

Acknowledgments

This research has been partially supported by Deutsche Forschungsgemeinschaft (DFG) through grant CRC 1114 "Scaling Cascades in Complex Systems", Project Number 235221301, Project A08 "Characterization and prediction of quasi-stationary atmospheric states". The authors thank G. Froyland for engaging discussions and for suggesting the oscillating gyre example in Section 6.2. The third mentioned author thanks I. Mezić for a suggestion regarding literature which proved valuable for this work.

References

[AFGTV23] J. Atnip, G. Froyland, C. González-Tokman, and S. Vaienti. Equilibrium states for non-transitive random open and closed dynamical systems. *Ergodic Theory and Dynamical Systems*, 43(10):3193–3215, 2023.

[Bal14] S. Balasuriya. Nonautonomous flows as open dynamical systems: characterising escape rates and time-varying boundaries. In *Ergodic Theory, Open Dynamics, and Coherent Structures*, pages 1–30. Springer, 2014.

[BCZG23] A. Blumenthal, M. Coti Zelati, and R. S. Gvalani. Exponential mixing for random dynamical systems and an example of Pierrehumbert. *The Annals of Probability*, 51(4):1559–1601, 2023.

[BDH16] A. Bovier and F. Den Hollander. *Metastability: a potential-theoretic approach*, volume 351. Springer, 2016.

[BEGK02] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein. Metastability and low lying spectra in reversible Markov chains. *Communications in Mathematical Physics*, 228:219–255, 2002.

[BFGTM14] C. Bose, G. Froyland, C. González-Tokman, and R. Murray. Ulam's method for Lasota–Yorke maps with holes. *SIAM Journal on Applied Dynamical Systems*, 13(2):1010–1032, 2014.

[BOR18] S. Balasuriya, N. T. Ouellette, and I. I. Rypina. Generalized Lagrangian coherent structures. *Physica D: Nonlinear Phenomena*, 372:31–51, 2018.

[BV19] L. Barreira and C. Valls. Hyperbolicity via evolution semigroups on l p. *Qualitative theory of dynamical systems*, 18(3):887–908, 2019.

[BY11] L. A. Bunimovich and A. Yurchenko. Where to place a hole to achieve a maximal escape rate. *Israel Journal of Mathematics*, 182:229–252, 2011.

[CL99] C. Chicone and Y. Latushkin. *Evolution semigroups in dynamical systems and differential equations*. Number 70. American Mathematical Soc., 1999.

[CMS97] P. Collet, S. Martínez, and B. Schmitt. The Pianigiani–Yorke measure for topological Markov chains. *Israel Journal of Mathematics*, 97:61–70, 1997.

[CMT98] N. Chernov, R. Markarian, and S. Troubetzkoy. Conditionally invariant measures for Anosov maps with small holes. *Ergodic Theory and Dynamical Systems*, 18(5):1049–1073, 1998.

[Col23] M. J. Colbrook. The mpEDMD algorithm for data-driven computations of measure-preserving dynamical systems. *SIAM Journal on Numerical Analysis*, 61(3):1585–1608, 2023.

[CS89] S. Childress and A. Soward. Scalar transport and alpha-effect for a family of cat's-eye flows. *Journal of Fluid Mechanics*, 205:99–133, 1989.

[CT24] M. J. Colbrook and A. Townsend. Rigorous data-driven computation of spectral properties of Koopman operators for dynamical systems. *Communications on Pure and Applied Mathematics*, 77(1):221–283, 2024.

[Dav82a] E. B. Davies. Metastable states of symmetric Markov semigroups i. *Proceedings of the London Mathematical Society*, 3(1):133–150, 1982.

[Dav82b] E. B. Davies. Metastable states of symmetric markov semigroups ii. *Journal of the London Mathematical Society*, 2(3):541–556, 1982.

[DG19] S. Das and D. Giannakis. Delay-coordinate maps and the spectra of Koopman operators. *Journal of Statistical Physics*, 175(6):1107–1145, 2019.

[DHFS00] P. Deuflhard, W. Huisinga, A. Fischer, and C. Schütte. Identification of almost invariant aggregates in reversible nearly uncoupled Markov chains. *Linear Algebra and its Applications*, 315(1-3):39–59, 2000.

[DJ99] M. Dellnitz and O. Junge. On the approximation of complicated dynamical behavior. *SIAM Journal on Numerical Analysis*, 36(2):491–515, 1999.

[DJM16] A. Denner, O. Junge, and D. Matthes. Computing coherent sets using the Fokker–Planck equation. *Journal of Computational Dynamics*, 3(2):163–177, 2016.

[DT17] M. F. Demers and M. Todd. Equilibrium states, pressure and escape for multimodal maps with holes. *Israel J. Math.*, 221(1):367–424, 2017.

[DW05] P. Deuflhard and M. Weber. Robust Perron cluster analysis in conformation dynamics. *Linear Algebra and its Applications*, 398:161–184, 2005.

[DWY12] M. F. Demers, P. Wright, and L.-S. Young. Entropy, Lyapunov exponents and escape rates in open systems. *Ergodic Theory Dynam. Systems*, 32(4):1270–1301, 2012.

[DY05] M. F. Demers and L.-S. Young. Escape rates and conditionally invariant measures. *Nonlinearity*, 19(2):377, 2005.

[ENB00] K.-J. Engel, R. Nagel, and S. Brendle. *One-parameter semigroups for linear evolution equations*, volume 194. Springer, 2000.

[FJK13] G. Froyland, O. Junge, and P. Koltai. Estimating long-term behavior of flows without trajectory integration: The infinitesimal generator approach. *SIAM Journal on Numerical Analysis*, 51(1):223–247, 2013.

[FK17] G. Froyland and P. Koltai. Estimating long-term behavior of periodically driven flows without trajectory integration. *Nonlinearity*, 30(5):1948, 2017.

[FK23] G. Froyland and P. Koltai. Detecting the birth and death of finite-time coherent sets. *Communications on Pure and Applied Mathematics*, 76(12):3642–3684, 2023.

[FKS20] G. Froyland, P. Koltai, and M. Stahn. Computation and optimal perturbation of finite-time coherent sets for aperiodic flows without trajectory integration. *SIAM Journal on Applied Dynamical Systems*, 19(3):1659–1700, 2020.

[FLQ10] G. Froyland, S. Lloyd, and A. Quas. Coherent structures and isolated spectrum for perron–frobenius cocycles. *Ergodic Theory and Dynamical Systems*, 30(3):729–756, 2010.

[FLS10] G. Froyland, S. Lloyd, and N. Santitissadeekorn. Coherent sets for nonautonomous dynamical systems. *Physica D: Nonlinear Phenomena*, 239(16):1527–1541, 2010.

[FPG14] G. Froyland and K. Padberg-Gehle. Almost-invariant and finite-time coherent sets: directionality, duration, and diffusion. In *Ergodic Theory, Open Dynamics, and Coherent Structures*, pages 171–216. Springer, 2014.

[Fro05] G. Froyland. Statistically optimal almost-invariant sets. *Physica D: Nonlinear Phenomena*, 200(3-4):205–219, 2005.

[Fro13] G. Froyland. An analytic framework for identifying finite-time coherent sets in time-dependent dynamical systems. *Physica D: Nonlinear Phenomena*, 250:1–19, 2013.

[Fro15] G. Froyland. Dynamic isoperimetry and the geometry of Lagrangian coherent structures. *Nonlinearity*, 28(10):3587, 2015.

[FS10] G. Froyland and O. Stancevic. Escape rates and Perron–Frobenius operators: Open and closed dynamical systems. *Discrete Contin. Dyn. Syst. Ser. B*, 14(2):457–472, 2010.

[FSM10] G. Froyland, N. Santitissadeekorn, and A. Monahan. Transport in time-dependent dynamical systems: Finite-time coherent sets. *Chaos*, 20(4):043116, 2010.

[GD20] D. Giannakis and S. Das. Extraction and prediction of coherent patterns in incompressible flows through space–time Koopman analysis. *Physica D: Nonlinear Phenomena*, 402:132211, 2020.

[Gia21] D. Giannakis. Delay-coordinate maps, coherence, and approximate spectra of evolution operators. *Research in the Mathematical Sciences*, 8(1):8, 2021.

[GT18] C. González-Tokman. Multiplicative ergodic theorems for transfer operators: Towards the identification and analysis of coherent structures in non-autonomous dynamical systems. *Contemp. Math*, 709:31–52, 2018.

[HBV13] G. Haller and F. J. Beron-Vera. Coherent Lagrangian vortices: The black holes of turbulence. *Journal of Fluid Mechanics*, 731:R4, 2013.

[HHFH16] G. Haller, A. Hadjighasem, M. Farazmand, and F. Huhn. Defining coherent vortices objectively from the vorticity. *Journal of Fluid Mechanics*, 795:136–173, 2016.

[How74] J. S. Howland. Stationary scattering theory for time-dependent Hamiltonians. *Mathematische Annalen*, 207:315–335, 1974.

[HS06] W. Huisinga and B. Schmidt. Metastability and dominant eigenvalues of transfer operators. *New Algorithms for Macromolecular Simulation*, pages 167–182, 2006.

[HY00] G. Haller and G. Yuan. Lagrangian coherent structures and mixing in two-dimensional turbulence. *Physica D: Nonlinear Phenomena*, 147(3-4):352–370, 2000.

[KK20] D. Karrasch and J. Keller. A geometric heat-flow theory of Lagrangian coherent structures. *Journal of Nonlinear Science*, 30(4):1849–1888, 2020.

[KLP19] P. Koltai, H. C. Lie, and M. Plonka. Fréchet differentiable drift dependence of perron–frobenius and koopman operators for non-deterministic dynamics. *Nonlinearity*, 32(11):4232, 2019.

[KPM20] M. Korda, M. Putinar, and I. Mezić. Data-driven spectral analysis of the Koopman operator. *Applied and Computational Harmonic Analysis*, 48(2):599–629, 2020.

[LH04] W. Liu and G. Haller. Strange eigenmodes and decay of variance in the mixing of diffusive tracers. *Physica D: Nonlinear Phenomena*, 188(1-2):1–39, 2004.

[LMD03] C. Liverani and V. Maume-Deschamps. Lasota–Yorke maps with holes: conditionally invariant probability measures and invariant probability measures on the survivor set. In *Annales de l'IHP Probabilités et statistiques*, volume 39, pages 385–412, 2003.

[LS91] Y. D. Latushkin and A. Stepin. Weighted shift operators, spectral theory of linear extensions, and the multiplicative ergodic theorem. *Mathematics of the USSR-Sbornik*, 70(1):143, 1991.

[Lun12] A. Lunardi. *Analytic semigroups and optimal regularity in parabolic problems*. Springer Science & Business Media, 2012.

[MB14] T. Ma and E. M. Bollt. Differential geometry perspective of shape coherence and curvature evolution by finite-time nonhyperbolic splitting. *SIAM Journal on Applied Dynamical Systems*, 13(3):1106–1136, 2014.

[Paz12] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44. Springer Science & Business Media, 2012.

[Pöt12] C. Pötzsche. Fine structure of the dichotomy spectrum. *Integral Equations Operator Theory*, 73(1):107–151, 2012.

[PR16] C. Pötzsche and E. Russ. Continuity and invariance of the Sacker–Sell spectrum. *J. Dynam. Differential Equations*, 28(2):533–566, 2016.

[PY79] G. Pianigiani and J. A. Yorke. Expanding maps on sets which are almost invariant. decay and chaos. *Transactions of the American Mathematical Society*, 252:351–366, 1979.

[RBBV+07] I. Rypina, M. G. Brown, F. J. Beron-Vera, H. Koçak, M. J. Olascoaga, and I. Udovydchenkov. On the Lagrangian dynamics of atmospheric zonal jets and the permeability of the stratospheric polar vortex. *Journal of the Atmospheric Sciences*, 64(10):3595–3610, 2007.

[RKW90] V. Rom-Kedar and S. Wiggins. Transport in two-dimensional maps. *Archive for Rational Mechanics and Analysis*, 109:239–298, 1990.

[RN12] F. Riesz and B. Nagy. *Functional Analysis*. Dover Books on Mathematics. Dover Publications, 2012.

[SKJ21] N. Schilling, D. Karrasch, and O. Junge. Heat-content and diffusive leakage from material sets in the low-diffusivity limit. *Nonlinearity*, 34(10):7303, 2021.

[SLM05] S. C. Shadden, F. Lekien, and J. E. Marsden. Definition and properties of Lagrangian coherent structures from finite-time Lyapunov exponents in two-dimensional aperiodic flows. *Physica D: Nonlinear Phenomena*, 212(3-4):271–304, 2005.

[SS78] R. J. Sacker and G. R. Sell. A spectral theory for linear differential systems. *J. Differential Equations*, 27(3):320–358, 1978.

[SS13] C. Schütte and M. Sarich. *Metastability and Markov state models in molecular dynamics*, volume 24. American Mathematical Soc., 2013.

[Sta22] M. Stahn. *Augmented Generators for Non-autonomous Flows*. PhD thesis, 2022.

[Tan17] H. Tanabe. *Functional analytic methods for partial differential equations*. CRC Press, 2017.

[VG23] C. Valva and D. Giannakis. Consistent spectral approximation of Koopman operators using resolvent compactification. *arXiv preprint arXiv:2309.00732*, 2023.

[vN22] J. van Neerven. *Functional analysis*, volume 201. Cambridge University Press, 2022.

A Proofs

Proof of Proposition 3.3. We use classical arguments to establish the decay of the evolution family. Since the operators $\mathcal{P}_\theta^t$ are real, it suffices to verify the bound for real functions. We

compute the decay of $\|f(t)\|_2^2$, where $f(t)$ is a solution to the Fokker–Planck equation (6),

$$\begin{aligned}
\partial_t \|f(t)\|_2^2 &= \partial_t \int_M f(t,x)^2 \, dx \\
&= \int_M 2f(t,x)\partial_t f(t,x) \, dx \\
&= \int_M \varepsilon^2 f(t,x)\Delta_x f(t,x) - 2f(t,x)\nabla_x f(t,x)\cdot v(t,x) \, dx \\
&= -\varepsilon^2 \int_M \nabla_x f(t,x)\cdot\nabla_x f(t,x) \, dx - \int_M \nabla_x (f(t,x)^2)\cdot v(t,x) \, dx \\
&= -\varepsilon^2 \|\nabla_x f(t)\|_2^2 - \int_M \operatorname{div}_x (f(t,\cdot)^2 v(t,\cdot)) \, dx.
\end{aligned}$$

By the divergence theorem, the second integral vanishes. Hence, the decay of $\|f(t)\|_2^2$ is given by

$$\partial_t \|f(t)\|_2^2 = -\varepsilon^2 \|\nabla_x f(t)\|_2^2 . \tag{36}$$

The Poincaré–Wirtinger inequaility states that there is a constant $c > 0$, depending only on M, such that for each $f \in \mathcal{H}_0$ we find

$$\|\nabla_x f\|_2 \geq c \|f\|_2 .$$

If $f(0) \in \mathcal{H}_0$, then $f(t) \in \mathcal{H}_0$ for all $t \geq 0$. Hence, we find

$$\partial_t \|f(t)\|_2^2 \leq -\varepsilon^2 c^2 \|f(t)\|_2^2 .$$

By the Grönwall's lemma, we obtain the bound

$$\|f(t)\|_2^2 \leq \|f(0)\|_2^2 \, e^{-\varepsilon^2 c^2 t}$$

Lastly, let $f = f(0) \in \mathcal{H}_0$ have $\|f\|_2 = 1$. By definition of $\mathcal{P}_\theta^t$, we have $\mathcal{P}_\theta^t f = f(t)$, and consequently

$$\left\|\mathcal{P}_\theta^t f\right\|_2 \leq e^{-\frac{\varepsilon^2 c^2}{2} t}.$$

This shows that $\|\mathcal{P}_\theta^t\| \leq e^{\varrho t}$ for $\varrho = -\frac{\varepsilon^2 c^2}{2}$, which completes the proof. □

Proof of Lemma 3.4. Let $f(0) = f \in \mathcal{H}_0$ with $\|f\|_2 = 1$. The functions $f_i(t) := \mathcal{P}_{\theta_i}^t f \in \mathcal{H}_0$, for $i = 1, 2$, are the unique classical solutions to the Fokker–Planck equation with starting parameter θ_i. In particular, we have for $t > 0$

$$\begin{aligned}
\partial_t f_i(t,x) &= \big[G(\phi^t\theta_i) f_i(t)\big](x) \\
&= \frac{1}{2}\varepsilon^2 \Delta_x f_i(t,x) - \nabla_x f_i(t,x)\cdot v(\phi^t\theta_i, x).
\end{aligned} \tag{37}$$

We can rewrite the equation (37) for $f_1(t)$, to be

$$\begin{aligned}
\partial_t f_1(t,x) &= \frac{1}{2}\varepsilon^2 \Delta_x f_1(t,x) - \nabla_x f_1(t,x)\cdot v(\phi^t\theta_2, x) + \nabla_x f_1(t,x)\cdot\big(v(\phi^t\theta_2,x) - v(\phi^t\theta_1,x)\big) \\
&= \big[G(\phi^t\theta_2) f_1(t)\big](x) + g(t),
\end{aligned} \tag{38}$$

where the nonhomogeneous part $g(t) \in \mathcal{H}_0$ is given by

$$g(t,x) := \nabla_x f_1(t,x) \cdot \big(v(\phi^t\theta_2, x) - v(\phi^t\theta_1, x)\big). \tag{39}$$

The nonautonomous generator $G(\phi^t\theta_2)$ generates the operators $\mathcal{P}^t_{\theta_2}$. Since f_1 is a classical solution to (38), the function $\mathcal{P}^t_{\theta_1} f = f_1(t)$ is given by the mild solution

$$\big[\mathcal{P}^t_{\theta_1} f\big](x) = \big[\mathcal{P}^t_{\theta_2} f\big](x) + \int_0^t \big[\mathcal{P}^{t-s}_{\theta_2} g(s)\big](x) ds.$$

Hence, we can bound the distance between the solutions $\mathcal{P}^t_{\theta_1} f$ and $\mathcal{P}^t_{\theta_2} f$.

$$\begin{aligned}
\big\|\mathcal{P}^t_{\theta_1} f - \mathcal{P}^t_{\theta_2} f\big\|_2^2 &= \int_M \left(\int_0^t [\mathcal{P}^{t-s}_{\theta_2} g(s)](x) ds\right)^2 dx \\
&\le \int_M t \int_0^t [\mathcal{P}^{t-s}_{\theta_2} g(s)](x)^2 \, ds \, dx \\
&= t \int_0^t \int_M [\mathcal{P}^{t-s}_{\theta_2} g(s)](x)^2 \, dx \, ds \\
&= t \int_0^t \big\|\mathcal{P}^{t-s}_{\theta_2} g(s)\big\|_2^2 \, ds \\
&< t \int_0^t \|g(s)\|_2^2 \, ds,
\end{aligned} \tag{40}$$

where in the second line we used Jensen's inequality and in the third line the Fubini–Tonelli theorem. We claim that as the distance between θ_1 and θ_2 approaches 0, the integral in the last line tends to 0. By definition of δ and $g(s)$ in (39), we find

$$|g(s,x)| \le \|\nabla_x f_1(s,x)\|_e \, \delta,$$

and consequently $\|g(s)\|_2^2 \le \delta^2 \|\nabla_x f_1(s)\|_2^2$ for $s \in [0,t]$. In the proof of Proposition 3.3, namely equation (36), we showed

$$\|\nabla_x f_1(s)\|_2^2 = -\frac{1}{\varepsilon^2} \partial_s \|f_1(s)\|_2^2 .$$

This yields

$$\|g(s,x)\|_2^2 \le -\frac{\delta^2}{\varepsilon^2} \partial_s \|f_1(s)\|_2^2 .$$

Inserting this bound into (40) yields

$$\begin{aligned}
\big\|\mathcal{P}^t_{\theta_1} f - \mathcal{P}^t_{\theta_2} f\big\|_2^2 &\le -\frac{\delta^2 t}{\varepsilon^2} \int_0^t \partial_s \|f_1(s)\|_2^2 \, ds \\
&= \frac{\delta^2 t}{\varepsilon^2} \big[\, \|f_1(0)\|_2^2 - \|f_1(t)\|_2^2 \,\big] \\
&\le \frac{\delta^2 t}{\varepsilon^2},
\end{aligned}$$

where we used $\|f_1(0)\|_2 = \|f\|_2 = 1$ and that $\|f_1(t)\|_2 \leq \|f_1(0)\|_2$, because $\mathcal{P}_\theta^t$ is a contraction. This estimate holds for any function $f \in \mathcal{H}_0$ with $\|f\|_2 = 1$. We conclude

$$\left\|\mathcal{P}_{\theta_1}^t - \mathcal{P}_{\theta_2}^t\right\| \leq \frac{\delta\sqrt{t}}{\varepsilon}.$$

This completes the proof. □

Proof of Theorem 4.2. Our proof closely follows [FK17, A.6]. Their proof requires that the family of sets $A_\theta^\bullet$ is *sufficiently nice*. We use arguments from [FKS20] to avoid this assumption.

Firstly, by Proposition 4.1, the solution $f(t) = \mathcal{P}_\theta^t f_0$ is represented by a function $f \in C(\mathbb{R}_{\geq 0} \times M, \mathbb{R})$ such that $A_\theta^\bullet$ is well-defined. Fix a time $T \geq 0$ and let $(t_k)_{k\in\mathbb{N}}$ be a dense sequence in $[0,T]$ with $t_1 = T$. Let $x_t(\omega)$ be a random trajectory of the SDE (5), where Ω is the probability space on which the SDE is defined. Define the events

$$\begin{aligned}\mathcal{E}_n &:= \{\omega \in \Omega \mid x_{t_k}(\omega) \in A_\theta^{t_k},\ \forall k \in \{1,\dots,n\}\},\\ \mathcal{E} &:= \{\omega \in \Omega \mid x_t(\omega) \in A_\theta^t,\ \forall t \in [0,T]\}.\end{aligned}$$

Since f is continuous and $(t_k)_{k\in\mathbb{N}}$ is dense in $[0,T]$, we find

$$f(t, x_t(\omega)) \geq 0,\ \forall t \in [0,T] \quad \Longleftrightarrow \quad f(t_k, x_{t_k}(\omega)) \geq 0,\ \forall k \in \mathbb{N}.$$

We conclude $\mathcal{E}_n \downarrow \mathcal{E}$.

Given a probability measure π on M, let $\mathbb{P}_{x_0\sim\pi}$ denote the probability measure on Ω, conditioned on the initial condition $x_0 \sim \pi$. If π is a finite measure that is not a probability measure, we define $\mathbb{P}_{x_0\sim\pi} = \pi(\Omega)\mathbb{P}_{x_0\sim\pi/\pi(\Omega)}$. If π is a finite signed measure, consider the Hahn decomposition $\pi = \pi^+ - \pi^-$, and define $\mathbb{P}_{x_0\sim\pi} = \mathbb{P}_{x_0\sim\pi^+} - \mathbb{P}_{x_0\sim\pi^-}$. Observe that the statement of the theorem does not depend on the scaling of f_0. Hence, w.l.o.g. we may assume $\|f_0\|_1 = 2$. In particular, the positive and negative part of f_0, which we denote by f_0^+, respectively f_0^-, integrate to 1. Hence, f_0^+ and f_0^- are the densities of probability measures ν^+ and ν^- on M. In particular, f_0 is the density of the signed measure $\nu = \nu^+ - \nu^-$.

Since $\mathcal{P}_\theta^T f_0$ is the distribution of particles after time T with initial distribution ν, we find

$$\mathbb{P}_{x_0\sim\nu}(x_T \in A_\theta^T) = \int_{A_\theta^T} [\mathcal{P}_\theta^T f_0](x)\,dx = \frac{1}{2}\left\|\mathcal{P}_\theta^T f_0\right\|_1.$$

The last equality follows from the definition of A_θ^T and the fact that $\mathcal{P}_\theta^T f_0$ integrates to 0. For $n \in \mathbb{N}$, consider the decomposition

$$\begin{aligned}\mathbb{P}_{x_0\sim\nu}(x_T \in A_\theta^T) &= \mathbb{P}_{x_0\sim\nu}(x_{t_k} \in A_\theta^{t_k},\ \forall k \in \{1,\dots,n\})\\ &\quad + \sum_{j=2}^{n} \underbrace{\mathbb{P}_{x_0\sim\nu}(x_{t_j} \notin A_\theta^{t_j},\ x_{t_k} \in A_\theta^{t_k},\ \forall t_k > t_j)}_{=:p_j}.\end{aligned}$$

One can show [FJK13] that $p_j \leq 0$. We conclude

$$\mathbb{P}_{x_0\sim\nu}(\mathcal{E}_n) \geq \frac{1}{2}\left\|\mathcal{P}_\theta^T f_0\right\|_1 .$$

The σ-additivity of $\mathbb{P}_{x_0\sim\nu}$ implies $\lim_{n\to\infty}\mathbb{P}_{x_0\sim\nu}(\mathcal{E}_n) = \mathbb{P}_{x_0\sim\nu}(\mathcal{E})$. This shows that also $\mathbb{P}_{x_0\sim\nu}(\mathcal{E})$ is bounded from below by $\frac{1}{2}\left\|\mathcal{P}_\theta^T f_0\right\|_1$.

It remains to argue that the measure ν is dominated by $\|f\|_\infty |A_\theta^0|\, \mathbb{1}_{A_\theta^0}$, where $\mathbb{1}_{A_\theta^0}$ is the uniform distribution on A_θ^0. This concludes the proof. □

Proof of Lemma 4.5. Let $\Pi : \Theta \to \mathcal{L}(\mathcal{H}_0)$ be the strongly continuous projection-valued function corresponding to the exponential dichotomy at λ. We have $S(\theta) = \mathrm{ran}(\Pi(\theta))$ and $U(\theta) = \ker(\Pi(\theta))$.

Since $f \in U(\theta)$, the inverse $g_t := (\mathcal{P}^t_{\phi^{-t}\theta})^{-1} f \in U(\phi^{-t}\theta)$ is well-defined for each $t \geq 0$. In particular, we find $\mathcal{P}^t_{\phi^{-t}\theta} g_t^R = f^R$ and $\mathcal{P}^t_{\phi^{-t}\theta} g_t^I = f^I$. Now, assume $f^R \notin U(\theta)$. Consider the decompositions

$$\begin{aligned}
f^R &= s_f + u_f, & s_f &:= \Pi(\theta) f^R \in S(\theta), & u_f &:= (\mathrm{Id} - \Pi(\theta)) f^R \in U(\theta),\\
g_t^R &= s_{g_t} + u_{g_t}, & s_{g_t} &:= \Pi(\phi^{-t}\theta) g_t^R \in S(\phi^{-t}\theta), & u_{g_t} &:= (\mathrm{Id} - \Pi(\phi^{-t}\theta)) g_t^R \in U(\phi^{-t}\theta),
\end{aligned}$$

in particular $s_f \neq 0$ by our assumption. By definition of Π, we have $\mathcal{P}^t_{\phi^{-t}\theta} s_f = s_{g_t}$, $\mathcal{P}^t_{\phi^{-t}\theta} u_f = u_{g_t}$ and

$$\begin{aligned}
\|s_f\|_2 &= \left\|\mathcal{P}^t_{\phi^{-t}\theta} s_{g_t}\right\|_2 \leq C e^{(\lambda-\beta)t}\|s_{g_t}\|_2 ,\\
\|u_{g_t}\|_2 &= \left\|(\mathcal{P}^t_{\phi^{-t}\theta})^{-1} u_f\right\|_2 \leq C e^{-(\lambda+\beta)t}\|u_f\|_2 .
\end{aligned}$$

This shows that $\|s_{g_t}\|_2$ grows at least at rate $-(\lambda-\beta)$ in t, while u_{g_t} grows at most at rate $-(\lambda+\beta)$. We conclude

$$\begin{aligned}
\left\|(\mathcal{P}^t_{\phi^{-t}\theta})^{-1} f\right\|_2 &\geq \left\|(\mathcal{P}^t_{\phi^{-t}\theta})^{-1} f^R\right\|_2 = \left\|g_t^R\right\|_2 \geq \|s_{g_t}\|_2 - \|u_{g_t}\|_2\\
&\geq C^{-1} e^{-(\lambda-\beta)t}\|s_f\|_2 - C e^{-(\lambda+\beta)t}\|u_f\|_2 .
\end{aligned} \tag{41}$$

Definition 2.1 implies that $\|(\mathcal{P}^t_\theta|_U)^{-1}\| \leq C e^{-(\lambda+\beta)t}$ for $t \geq 0$ *uniformly* in $\theta \in \Theta$. Since $f \in U(\theta)$, it follows that $e^{(\lambda+\beta)t}\|(\mathcal{P}^t_{\phi^{-t}\theta})^{-1} f\|_2$ is bounded by the constant $C\|f\|_2$ for every $t \geq 0$. This contradicts (41), which predicts that $e^{(\lambda+\beta)t}\|(\mathcal{P}^t_{\phi^{-t}\theta})^{-1} f\|_2$ grows to infinity at least with rate $2\beta > 0$, neglecting additive constants. We conclude $s_f = 0$ and, therefore, $f^R \in U(\theta)$. The proof for $f^I \in U(\theta)$ is analogous. □

Proof of Lemma 5.1. The first equality is standard and holds for any real-valued function.

We express the divergence of v is terms of its Fourier coefficients.

$$\begin{aligned}
\operatorname{div}_x(v)(\theta,x) &= \sum_{j=1}^{d_p} \partial_{x_j} v_j(\theta,x) \\
&= \sum_{j=1}^{d_p} \partial_{x_j} \sum_{m,n} F_{m,n}(\theta,x)\widehat{v_j}(m,n) \\
&= -2\pi i \sum_{j=1}^{d_p} \sum_{m,n} n_j F_{m,n}(\theta,x)\widehat{v_j}(m,n) \\
&= -2\pi i \sum_{m,n} (n\cdot \widehat{v}(m,n)) F_{m,n}(\theta,x).
\end{aligned}$$

This shows that the Fourier coefficients of the divergence of v are given by $n\cdot\widehat{v}(m,n)$. Since the divergence of v is zero as a function, all the Fourier coefficients of the divergence must vanish. We conclude $n\cdot\widehat{v}(m,n)=0$ for all $m,n\in\mathbb{Z}^{d_d}\times\mathbb{Z}^{d_p}$. □

Proof of Proposition 6.1. Fix a fixed starting parameter $\theta\in\mathbb{T}^2$ and a starting distribution $f\in\mathcal{H}_0$. Define $f_1(t) := (T_{\phi^t\theta}\circ\mathcal{P}^t_\theta)f$ and $f_2(t) := (\mathcal{P}^t_{\text{eff}}\circ T_\theta)f$. Both functions satisfy $f_1(0)=f_2(0)=T_\theta f$. We compute the time-derivatives of both functions using the Fokker–Planck equation (6)

$$\begin{aligned}
\partial_t f_1(t,x) &= \partial_t\big[[\mathcal{P}^t_\theta f](x-\theta-\alpha t)\big] \\
&= \frac{1}{2}\varepsilon^2\Delta_x[\mathcal{P}^t_\theta f](t,x-\phi^t\theta) - \nabla_x[\mathcal{P}^t_\theta f](t,x-\phi^t\theta)\cdot\big(v(\phi^t\theta,x-\phi^t\theta)+\alpha\big) \\
&= \frac{1}{2}\varepsilon^2\Delta_x f_1(t,x) - \nabla_x f_1(t,x)\cdot\big(v_{\text{aut}}(x)+\alpha\big). \\
\partial_t f_2(t,x) &= \partial_t[\mathcal{P}^t_{\text{eff}}(T_\theta f)](x) \\
&= \frac{1}{2}\varepsilon^2\Delta_x[\mathcal{P}^t_{\text{eff}}(T_\theta f)](x) - \nabla_x[\mathcal{P}^t_{\text{eff}}(T_\theta f)](x)\cdot v_{\text{eff}}(x) \\
&= \frac{1}{2}\varepsilon^2\Delta_x f_2(t,x) - \nabla_x f_2(t,x)\cdot\big(v_{\text{aut}}(x)+\alpha\big).
\end{aligned}$$

Since f_1 and f_2 solve the same NACP, which is known to have a unique solution, we conclude $f_1(t)=f_2(t)$ for all $t\geq 0$. This concludes the proof. □

Proof of Proposition 6.2. Consider $\lambda\notin\Sigma(\mathcal{P}_{\text{eff}})$, such that there is a strongly continuous projection-valued function $\Pi_{\text{eff}}:\mathbb{T}^2\to\mathcal{L}(\mathcal{H})$ that defines an exponential dichotomy of $\mathcal{P}_{\text{eff}}$ at λ. By Section 2.5, the function Π_{eff} is constant, i.e. does not depend on θ. Define another projection-valued function $\Pi:\mathbb{T}^2\to\mathcal{L}(\mathcal{H}_0)$ by $\Pi(\theta)f = T_{-\theta}\circ\Pi_{\text{eff}}\circ T_\theta$. We show that Π defines an exponential dichotomy of $\mathcal{P}$. Let $\theta\in\Theta$ and $t\geq 0$. Using the commutativity of (30), we

compute

$$\begin{aligned}\mathcal{P}^t_\theta \Pi(\theta) &= T^{-1}_{\phi^t\theta} \circ \mathcal{P}^t_{\text{eff}} \circ T_\theta \circ T_{-\theta} \circ \Pi_{\text{eff}} \circ T_\theta \\ &= T_{-\phi^t\theta} \circ \Pi_{\text{eff}} \circ \mathcal{P}^t_{\text{eff}} \circ T_\theta \\ &= T_{-\phi^t\theta} \circ \Pi_{\text{eff}} \circ T_{\phi^t\theta} \circ T_{-\phi^t\theta} \circ \mathcal{P}^t_{\text{eff}} \circ T_\theta \\ &= \Pi(\phi^t\theta)\mathcal{P}^t_\theta.\end{aligned}$$

Let $S(\theta)$ and $U(\theta)$ be the range and kernel of $\Pi(\theta)$. Analogously, let S_{eff} and U_{eff} be the range and kernel of Π_{eff}. By definition of Π, we find $U(\theta) = T_{-\theta}U_{\text{eff}}$ and $S(\theta) = T_{-\theta}S_{\text{eff}}$. Hence the restriction $\mathcal{P}^t_\theta|_U : U(\theta) \to U(\phi^t\theta)$ is given by

$$\mathcal{P}^t_\theta|_U = (T^{-1}_{\phi^t\theta} \circ \mathcal{P}^t_{\text{eff}} \circ T_\theta)|_U = T^{-1}_{\phi^t\theta} \circ \mathcal{P}^t_{\text{eff}}|_{U_{\text{eff}}} \circ T_\theta. \tag{42}$$

Since $\mathcal{P}^t_{\text{eff}}|_{U_{\text{eff}}}$ is invertible, this shows that $\mathcal{P}^t_\theta|_U$ is invertible.

Let $C > 0$ and $\beta > 0$ be the constants of the exponential dichotomy Π_{eff}. Consider $f \in S(\theta)$, such that $T_\theta f \in S_{\text{eff}}$. We find

$$\left\|\mathcal{P}^t_\theta f\right\|_2 = \left\|T^{-1}_{\phi^t\theta} \circ \mathcal{P}^t_\theta \circ T_\theta f\right\|_2 \leq Ce^{(\lambda-\beta)t} \left\|T_\theta f\right\|_2.$$

Since T_θ is an isometry, this proves $\|\mathcal{P}^t_\theta|_S\| \leq Ce^{(\lambda-\beta)t}$. Now, consider $f \in U(\phi^t\theta)$, such that $T_{\phi^t\theta}f \in U_{\text{eff}}$. From (42), we derive

$$\left\|(\mathcal{P}^t_\theta)^{-1} f\right\|_2 = \left\|T^{-1}_\theta \circ (\mathcal{P}^t_{\text{eff}})^{-1} \circ T_{\phi^t\theta} f\right\|_2 \geq Ce^{-(\lambda+\beta)t} \left\|T_{\phi^t\theta} f\right\|_2.$$

Since $T_{\phi^t\theta}$ is an isometry, this proves $\|\mathcal{P}^t_\theta|_U\|^{-1} \geq C^{-1}e^{(\lambda+\beta)t}$. This shows that Π defines an exponential dichotomy of $\mathcal{P}$ at λ. □