Random Splitting of Fluid Models

Ergodicity, Convergence, and Chaos

by

Omar Melikechi

Department of Mathematics
Duke University

Date: ______________________________
Approved:

____
Jonathan C. Mattingly, Supervisor

____
Alexander Kiselev

____
Jianfeng Lu

____
James Nolen

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Department of Mathematics
in the Graduate School of Duke University
2022

Abstract

Random Splitting of Fluid Models
Ergodicity, Convergence, and Chaos

by

Omar Melikechi

Department of Mathematics
Duke University

Date: ______________________
Approved:

____
Jonathan C. Mattingly, Supervisor

____
Alexander Kiselev

____
Jianfeng Lu

____
James Nolen

An abstract of a dissertation submitted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy in the Department of Mathematics
in the Graduate School of Duke University
2022

Abstract

In this dissertation we study random splitting and apply our results to random splittings of fluid models. Random splitting is loosely defined as follows. Consider the differential equation $\dot{x} = V(x)$ where $\dot{x}$ is a time derivative and the vector field V on $\mathbb{R}^D$ splits as the sum $V = \sum_{j=1}^{n} V_j$. In traditional operator splitting one approximates solutions of $\dot{x} = V(x)$ by composing solutions of $\dot{x} = V_j(x)$ over (typically small) deterministic time steps. Here we take these times to be independent and identically distributed random variables. This turns the aforementioned compositions into a Markov chain, which we call a *random splitting of* V or simply *random splitting*. We prove under relatively mild conditions that these random splittings possess a unique invariant measure (ergodicity), that their trajectories converge on average and almost surely to trajectories of the original system $\dot{x} = V(x)$ (convergence), and that, in certain cases, their top Lyapunov exponent is positive (chaos). After proving these general results, we construct random splittings of four fluid models: the conservative Lorenz-96 and Lorenz-96 equations, and Galerkin approximations of the 2d Euler and 2d Navier-Stokes equations on the torus. We prove these random splittings are ergodic and converge to their deterministic counterparts in a certain sense, and, for conservative Lorenz-96 and 2d Euler, that their top Lyapunov exponent is positive.

To Kevin

Contents

List of Figures

Acknowledgements

Tremendous thanks to my advisor, Jonathan Mattingly, for conceiving of this project, for our numerous conversations, and for his support – as department chair, as an advisor, and as a friend – over the last five years. I also owe tremendous thanks to Andrea Agazzi for his collaboration on this work and for his support and friendship throughout. This dissertation certainly would not be possible without them.

In addition, I thank the Duke University Math Department, especially my committee members, Alexander (Sasha) Kiselev, Jianfeng Lu, and Jim Nolen, as well as those involved in obtaining the grant NSF-DMS-2038056 that partially supported me during this work. I have really enjoyed being a graduate student at Duke and am grateful to everyone in the department for their camaraderie and guidance, mathematical and otherwise. I also thank James Wheeler for his friendship and many interesting conversations as well as my collaborators on projects outside of this work, including Trevor Bowman, David Dunson, Yishu Gong, Siming He, James Johndrow, Sasha Kiselev, Hanyu Song, Tao Tang, Steven Winter, and Alex Young.

I quit my job in May 2015 and drove from Delaware to Tucson with a cursory understanding of calculus and the seemingly impossible goal of earning a math Ph.D. Two weeks later one of my best friends, a friend I had nearly every class with from kindergarten to twelfth grade, committed suicide. I came very close to giving up and going home. To this end, I am forever grateful to the University of Arizona Math Department, especially my professors Bryden Cais, David Glickenstein, Rob Indik, Tyler Kloefkorn, and Klaus Lux, and friends Angel Burr, Craig Gross, and Jimmy Yao, for welcoming me, pushing me, and making me believe this might be possible. And thanks to Kevin for making me laugh all those years. I hope you'd be proud.

Finally, no words can express my gratitude to the friends I grew up with in Dover, Delaware, to my family, Algerian and American, and, above all else, to Mom, Pop, Lilia, and Meagan. Your love means the world.

1

Introduction

Many differential equations model systems whose observed behavior is difficult to deduce from the equation itself. This is famously true of fluid systems whose long-time and chaotic behavior has in many cases eluded mathematical verification for decades. In this dissertation we study deterministic differential equations and the systems they describe by splitting them into simpler, random ones. Two guiding principles underlie this *random splitting* approach. First, splitting into simple parts simplifies by definition; the delicacy is in how to split. If the simple is too simple, important properties of the system are lost. If the simple is not simple enough, the problem remains intractable. Second, randomness turns impossible into improbable, and improbable is far more forgiving than impossible. In particular, deterministic trajectories are predetermined: at any point in space, a deterministic trajectory can move in exactly one direction. With randomness no such certainty exists: at most points, a random splitting trajectory will be able to move in many possible directions. As we will see, the flexibility of uncertainty is crucial to our results.

This dissertation is largely based on *Random splitting of fluid models: Ergodicity and convergence* and *Random splitting of fluid models: Positive Lyapunov exponents* by Andrea Agazzi, Jonathan C. Mattingly, and the author of this dissertation, Omar Melikechi [2, 3]. Both papers are currently submitted and under review.

1.1 Splitting a butterfly

To illustrate the general idea of random splitting, consider the *Lorenz-63 equations*

$$\begin{aligned}\dot{x}_1 &= \sigma(x_2 - x_1)\\ \dot{x}_2 &= (\rho - x_3)x_1 - x_2\\ \dot{x}_3 &= x_1x_2 - \beta x_3,\end{aligned} \tag{1.1}$$

where $x = (x_1, x_2, x_3)$ is in $\mathbb{R}^3$, $\dot{x}$ denotes the derivative of x with respect to time, and σ, ρ, and β are positive constants. Edward Lorenz introduced this toy model for atmospheric convection in 1963 and observed its sensitive dependence on initial conditions, a key feature of what would become known as *chaos* [49, 67]. The vector field V corresponding to (1.1) "splits" as the sum of two vector fields, V_1 and V_2:

$$V(x) = \begin{pmatrix}\sigma(x_2 - x_1)\\ (\rho - x_3)x_1 - x_2\\ x_1x_2 - \beta x_3\end{pmatrix} = \begin{pmatrix}-\sigma & \sigma & 0\\ \rho & -1 & 0\\ 0 & 0 & -\beta\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix} + \begin{pmatrix}0\\ -x_1x_3\\ x_1x_2\end{pmatrix} =: V_1(x) + V_2(x). \tag{1.2}$$

In general there are many ways to decompose a vector field. Here we split V into linear and nonlinear parts, V_1 and V_2, respectively. Solutions, or *flows*, of $\dot{x} = V_1(x)$ dissipate in the x_3 coordinate since $\beta > 0$. Flows of V_2 are rotations in the (x_2, x_3)-plane with angular velocity x_1. In particular, x_1 determines the direction and speed of rotation – clockwise when $x_1 < 0$, counterclockwise when $x_1 > 0$, and faster and faster as $|x_1|$ grows – and hence has a shearing effect. Such nonlinear rotations appear and play a significant role in all examples studied in this work.

The idea of splitting a differential equation as above is called *operator splitting* and has existed since the work of Sophus Lie in 1875 [33]. It is well-established that if $\psi : \mathbb{R}^3 \times \mathbb{R} \to \mathbb{R}^3$ denotes the flow of V and $\varphi^{(i)}$ the flow of V_i for $i = 1, 2$, then the error incurred in approximating $\psi_h := \psi(\cdot, h)$ by the composed flow $\Phi_h := \varphi_h^{(2)} \circ \varphi_h^{(1)}$

Figure 1.1: *Splitting Lorenz-63.* Blue curves in both plots show true trajectories of the Lorenz-63 equations (1.1) starting from $x = (1, 1, 1)$ with the same parameters $\sigma = 10$, $\rho = 28$, and $\beta = 8/3$ used by Lorenz. Note these trajectories settle on the *Lorenz butterfly*, which is the attracting set of the Lorenz-63 dynamics. The orange curves on the left and right are deterministic and random splittings corresponding to (1.2), respectively. Note the deterministic splitting follows a much more apparent pattern than its random counterpart.

is $O(h^2)$. Consequently, for any finite t the error incurred in approximating ψ_t by composing Φ with itself m times, i.e. $\Phi_h^m = \Phi_h \circ \cdots \circ \Phi_h$ with $mh = t$, is $O(h)$ [51]. In particular, the composed flow converges to the true flow at worst linearly in h as $h \to 0$ on any finite time scale. Traditionally the time step h is fixed; the novelty of random splitting is that the time steps of the composed flow Φ are random. A canonical choice in this work is to fix $h > 0$ and choose times τ_i that are independent exponential random variables with mean 1. Hence the collection of times $h\tau_i$ are independent exponential random variables with mean h. Figure 1.1 shows deterministic and random splitting trajectories of (1.1) corresponding to the splitting (1.2).

The role of *random* in random splitting is less obvious but just as significant as that of *splitting*. We elaborate on this in the next section, but for now note the splitting vector fields $\{V_1, V_2\}$ in (1.2) span at most 2 dimensions at any point in

$\mathbb{R}^3$. On the other hand, the 3×3 matrix with columns V_1, V_2, and the Lie bracket[1] $[V_1, V_2]$ has rank 3 almost everywhere in $\mathbb{R}^3$. For instance, evaluating at the point $x_* = (1, 0, 1)$ and row reducing yields

$$\begin{pmatrix} | & | & | \\ V_1(x_*) & V_2(x_*) & [V_1, V_2](x_*) \\ | & | & | \end{pmatrix} \xrightarrow{\text{row reduce}} \begin{pmatrix} -\sigma & * & * \\ 0 & \sigma & * \\ 0 & 0 & \sigma(\beta - \rho) \end{pmatrix},$$

which has rank 3 provided $\beta \neq \rho$ (the $*$ indicates irrelevant entries). Thus appending $[V_1, V_2]$ to the splitting $\{V_1, V_2\}$ adds a new dimension to the dynamics.

1.2 The role of randomness

The long-time behavior of a dynamical system is closely related to its ability to explore its space. For example, trajectories of the Lorenz-63 equations (1.1) converge to the set pictured in blue in Figure 1.1 and come arbitrarily close to every point on that set for almost every initial condition [50]. A significant role of the randomness in random splitting is that it enhances exploration in a manner conducive to mathematical analysis. This was hinted at in the splitting of Lorenz-63. There we saw the Lie bracket $[V_1, V_2]$ added a dimension to $\{V_1, V_2\}$ so that the full dimension, 3, of the space was realized. Thus if the times t_i in composed flows of V_1, V_2, and $[V_1, V_2]$ are taken arbitrarily small, the dynamics can move in any infinitesimal direction in space. This is impossible if the t_i are deterministic as in traditional operator splitting, but if they are chosen randomly from a distribution supported on $(0, \varepsilon)$ there is nonzero probability the t_i can be arbitrarily small. Herein lies the point: *the "random" in random splitting allows infinitesimal access to all splitting vector fields.* In fact, we see shortly that random splitting allows access to the entire Lie algebra generated by the splitting vector fields. This stands in stark contrast to the dynamics of deterministic differential equations and deterministic splitting. The rest of the section aims to

[1] See Section 2.1 for the definition of the Lie bracket.

bridge the gap between the preceding discussion and the more technical results of later chapters. What follows is inspired by discussions in [12] and gives insight into how and why Lie brackets are fundamental to our work.

Consider a smooth d-dimensional manifold $\mathcal{X}$ with $d \geqslant 3$. Vector fields V on $\mathcal{X}$ assign to each point x in $\mathcal{X}$ a vector $V(x)$ in the tangent space $T_x\mathcal{X}$. Intuitively, $V(x)$ specifies a "direction" at x. The *flow* $\varphi : \mathcal{X} \times \mathbb{R} \to \mathcal{X}$ of V is the solution of the differential equation $\dot{x} = V(x)$ on $\mathcal{X}$; in all that follows, we assume vector fields are *complete* so their flows exist for all initial conditions and all times. Setting $\varphi_t(x) := \varphi(x,t)$ we have, by definition,

$$\partial_t \varphi_t(x) = V\left(\varphi_t(x)\right). \tag{1.3}$$

Hence $\partial_t \varphi_t(x)$ gives the instantaneous direction of motion of the dynamics.

Now consider two vector fields V_1 and V_2 on $\mathcal{X}$ with flows $\varphi^{(1)}$ and $\varphi^{(2)}$. When only one vector field is available, we can only move in one infinitesimal direction at any point. What about with two? To answer this, consider the composition $\Phi = \varphi^{(1)} \circ \varphi^{(2)} \circ \varphi^{(1)} : \mathcal{X} \times \mathbb{R}^3 \to \mathcal{X}$. In (1.3) the instantaneous direction of the dynamics was given by the time derivative of the flow. So again we take the time derivative of Φ, which is now multivariate, to obtain the possible directions of motion at x. By the chain rule and basic properties of flows,

$$D_x\Phi_t(x)^{-1} D_t\Phi_t(x) = \begin{pmatrix} | & | & | \\ V_1(x) & D_x\varphi^{(1)}_{-t_1}(x) V_2\left(x^{(1)}\right) & D_x\varphi^{(1)}_{-t_1}(x) D_x\varphi^{(2)}_{-t_2}\left(x^{(1)}\right) V_1\left(x^{(2)}\right) \\ | & | & | \end{pmatrix} \tag{1.4}$$

where $t = (t_1, t_2, t_3)$, $x^{(1)} = \varphi^{(1)}_{t_1}(x)$, $x^{(2)} = \varphi^{(2)}_{t_2} \circ \varphi^{(1)}_{t_1}(x)$, and D_x and D_t denote derivatives with respect to x and t. Note we have pulled the time derivative $D_t\Phi_t(x)$ back by $D_x\Phi_t(x)^{-1}$. This ensures the columns of the $d \times 3$ matrix in (1.4) are tangent vectors at x, which is the point we are interested in, rather than the time-dependent point $\Phi_t(x)$. At $t=0$,

$$D_x\Phi_t(x)^{-1}D_t\Phi_t(x)\Big|_{t=0} = \begin{pmatrix} | & | & | \\ V_1(x) & V_2(x) & V_1(x) \\ | & | & | \end{pmatrix}.$$

Thus if V_1 and V_2 are linearly independent at x, i.e. if the above matrix has rank 2, we can move infinitesimally via Φ in a 2-dimensional space about x. One might suspect this is the best possible outcome with only two vector fields, but it is not. To see why, consider the *Lie bracket* $[V_1, V_2]$ *between* V_1 *and* V_2,

$$[V_1, V_2](x) = \mathscr{L}_{V_1}V_2(x) := \lim_{t\to 0}\frac{D\varphi_{-t}^{(1)}(x)V_2(\varphi_t^{(1)}(x)) - V_2(x)}{t}, \tag{1.5}$$

where $\mathscr{L}_{V_1}V_2$ is the *Lie derivative of* V_2 *with respect to* V_1. The Lie bracket is discussed in Section 2.1; for more on the Lie derivative and proof that $[V_1,V_2] = \mathscr{L}_{V_1}V_2$, see [46, Theorem 9.38]. For now it suffices to know $[V_1,V_2]$ is itself a vector field on $\mathcal{X}$. Rearranging (1.5) and substituting t_1 for t gives

$$D\varphi_{-t_1}^{(1)}(x)^{-1}V_2\left(x^{(1)}\right) = V_2(x) + t_1[V_1,V_2](x) + O(t_1^2). \tag{1.6}$$

Note this is the second column of (1.4). Similarly, the third column satisfies

$$D_x\varphi_{-t_1}^{(1)}(x)D_x\varphi_{-t_2}^{(2)}\left(x^{(1)}\right)V_1\left(x^{(2)}\right) = V_1(x) + t_2[V_2,V_1](x) + O\left(\|t\|^2\right). \tag{1.7}$$

Plugging into (1.4) and ignoring higher order terms,

$$D_x\Phi_t(x)^{-1}D_t\Phi_t(x) \approx \begin{pmatrix} | & | & | \\ V_1(x) & V_2(x) + t_1[V_1,V_2](x) & V_1(x) + t_2[V_2,V_1](x) \\ | & | & | \end{pmatrix}.$$

By elementary column operations this has the same rank as

$$\begin{pmatrix} | & | & | \\ V_1(x) & V_2(x) & [V_1,V_2](x) \\ | & | & | \end{pmatrix} \tag{1.8}$$

provided t_1 and t_2 can be arbitrarily small – again highlighting the significance of randomness – in which case higher order terms are rigorously disposed of using lower semicontinuity of matrix rank. Hence the rank of $D_x\Phi_t(x)^{-1}D_t\Phi_t(x)$ is at least the rank of (1.8) for t_i sufficiently small. So if V_1, V_2, and $[V_1, V_2]$ are linearly independent at x, the composed flow Φ, which corresponds to only two vector fields, can move infinitesimally in 3 dimensions about x. In fact, by further composing $\varphi^{(1)}$ and $\varphi^{(2)}$ and thereby obtaining additional Lie brackets, e.g. $[V_1,[V_2,V_2]]$, $[V_2,[V_1,[V_2,V_1]]]$, etc., a more intricate but fundamentally similar argument shows arbitrary finite compositions of these two flows move infinitesimally in n dimensions about x where n is the dimension of the Lie algebra generated by $\{V_1, V_2\}$ at x [6, 12, 40]. For a formal statement of this fact, see Theorem 2.5 in Chapter 2.

1.3 Models

We consider four fluid models in this work: *conservative Lorenz-96*, *Lorenz-96*, *Galerkin approximations of 2-dimensional Euler on the torus*, and *Galerkin approximations of 2-dimensional Navier-Stokes on the torus*. The conservative models, conservative Lorenz-96 and 2d Euler, are studied in Chapter 6. Their nonconservative counterparts, Lorenz-96 and 2d Navier-Stokes, are studied in Chapter 7. We will construct random splittings for each and apply to them the general results of Chapters 3 and 4. For the conservative models we also apply the results of Chapter 5.

1.3.1 Lorenz-96 and conservative Lorenz-96

Fix $n \geqslant 4$. The *Lorenz-96 equations* [48] are

$$\dot{x} = \sum_{j=1}^{n} \big((x_{j+1} - x_{j-2})x_{j-1} - \nu x_j + F_j\big)e_j \tag{1.9}$$

where x in $\mathbb{R}^n$, $\{e_j\}_{j=1}^n$ is the standard basis of $\mathbb{R}^n$, $\nu, F_j > 0$, and indices are periodized via the identities $x_{-1} := x_{n-1}$, $x_0 := x_n$, and $x_{n+1} := x_1$. The $-\nu x_j$ term represents dissipation in the jth coordinate and F_j is a forcing constant. We also study a variant of Lorenz-96, called *conservative Lorenz-96*, obtained by removing the dissipation and forcing terms:

$$\dot{x} = V(x) := \sum_{j=1}^{n} (x_{j+1} - x_{j-2}) x_{j-1} e_j \,.$$

We sometimes refer to the original Lorenz-96 model (1.9) as *forced Lorenz-96* to emphasize the forcing (though dissipation is equally important). For conservative Lorenz-96, we split V into a collection of simple rotations similar to the one in Lorenz-63 by observing that

$$V(x) = \sum_{j=1}^{n} V_j(x)$$

where $V_j(x) := (x_{j+1} e_j - x_j e_{j+1}) x_{j-1}$. The dynamics given by $\dot{x} = V_j(x)$ are easy to understand on their own; any complex behavior comes from interactions of the rotations. Importantly, each V_j is chosen to conserve, like V, the system's *energy*, which for Lorenz-96 is the square of the usual Euclidean norm, $\|x\|^2 := \sum_{j=1}^{n} x_j^2$.

1.3.2 Navier-Stokes and Euler

Let $\mathbb{T}$ denote the 2-dimensional torus, i.e. $\mathbb{T} := [0, 2\pi]^2$ with periodic boundary conditions. The *incompressible 2d Navier-Stokes equations* on $\mathbb{T}$, which model the flow of an incompressible fluid, are

$$\begin{cases} \partial_t u + (u \cdot \nabla) u = -\nabla p + F + \nu \Delta u \,, \\ \operatorname{div}(u) := \nabla \cdot u = 0, \end{cases}$$

where $u : \mathbb{T} \times \mathbb{R} \to \mathbb{R}^2$ is the fluid velocity, $p : \mathbb{T} \times \mathbb{R} \to \mathbb{R}$ the fluid pressure,

$$(u \cdot \nabla) u = (u_1 \partial_1 u_1 + u_2 \partial_2 u_1, u_1 \partial_1 u_2 + u_2 \partial_2 u_2), \quad \text{and} \quad \Delta u = \partial_1^2 u_1 + \partial_2^2 u_2.$$

Here $u = (u_1, u_2)$ and $\partial_j := \partial_{x_j}$. The viscosity $\nu > 0$ measures the strength of the dissipation introduced by the Laplacian Δ, and $F(x,t)$ is an external driving force whose role is to keep the system from relaxing to the trivial state $u \equiv 0$. The *2d*

Euler equations on $\mathbb{T}$ are obtained from 2d Navier-Stokes by dropping the dissipative and forcing terms:

$$\begin{cases} \partial_t u + (u \cdot \nabla) u = -\nabla p \\ \operatorname{div}(u) := \nabla \cdot u = 0. \end{cases}$$

In this work we consider Nth Galerkin approximations of 2d Euler and 2d Navier-Stokes rather than the full equations themselves. These are obtained by considering the Fourier transform $q(k,t)$ of the vorticity of u and truncating high-frequency modes with $|k| > N$. A precise derivation of the resulting ordinary differential equation and its splitting is given in Chapter 6.

1.4 Related work

This thesis is largely based on [2, 3]. Specifically, results on ergodicity (Chapter 3) and convergence (Chapter 4) as well as the construction of random splittings of all models mentioned above (see also Chapters 6 and 7) can be found in [2]. All results on positive Lyapunov exponents, i.e. chaos, (Chapter 5) are in [3].

Among the simplest fluid models displaying interesting out-of-equilibrium behavior are the 2d Euler and incompressible Navier-Stokes equations described above. By balancing the dissipative effect of Δu, the forcing term allows the system to establish an out-of-equilibrium steady state. Such equilibria often develop fluxes across scales, a phenomenon whose study is an active area of research. Often F is taken to live on only a few scales so that the flux out of those scales can be studied [26, 37, 43, 53]. In practice, the forcing $F(x,t)$ is usually taken to be stochastic in space and time for some stationary distribution which is typically white in time [22, 26, 28, 37]. A common choice in the literature is $F(x,t) = \sum \psi_k(x) \dot{W}_k(t)$ where each $\psi_k(x)$ is a fixed spatial forcing and $\{\dot{W}_k(t)\}$ are mutually independent white-in-time noise terms written here as the formal derivative of a Brownian motion. Stochastic forcing serves multiple purposes in these settings. On one hand, it provides the energetic excitation which keeps the system out of equilibrium and allows for the establishment of a nontrivial statistical steady state. On the other, it provides local agitation which, modulo certain constraints, ensures the existence of a unique statistical steady state to which the system converges for most initial conditions. That is, it guarantees the forcing is sufficiently generic to ensure convergence to a single long time statistical behavior of the system, largely independent of the system's initial configuration.

Random splitting injects randomness while separating in a simple way the various roles served by noise in previous works as mentioned above [22, 26, 28, 37]. In particular, randomness is used primarily to ensure that when the dynamics is sufficiently generic, unique ergodicity[2] holds for a broad class of initial conditions. This will free one to use a much less disruptive class of forcing to keep the system out of equilibrium. More specifically, random splitting has a number of desirable properties:

1. It allows the separation of forcing, which keeps the system out of equilibrium, and stochastic agitation, which ensures the system has a unique long time statistical behavior.

2. It is strongly non-reversible since it is constructed from dynamics which only flow in directions the original dynamics could already move.

3. For the models considered in this work, it preserves the conserved quantities of the original dynamics. This allows properties of the (stochastic) conservative dynamics to be studied directly rather than only as a limit of the forced-dissipated dynamics.

4. Splitting into the composition of simple dynamics isolates particular nonlinear interactions which are relatively intuitive and can be explicitly analyzed.

[2] See Chapter 3.

By balancing between preservation of fundamental macroscopic properties of the original dynamics as in (3) and simplicity of the fundamental building blocks in our model dynamics as in (4), we expect random splitting will provide meaningful physical and dynamical insight into nonequilibrium steady states of models such as 2d Euler and Navier-Stokes.

1.5 Outline

This thesis is organized as follows. Random splitting is formally introduced in Chapter 2 along with the Lie bracket condition and results about real analytic vector fields. In Chapters 3, 4, and 5 we prove under typically mild conditions that random splittings have desirable long-time behavior (ergodicity), that they are close to their deterministic counterparts (convergence), and, in special cases, that their largest Lyapunov exponent is positive (chaos). In Chapter 6 we construct random splittings for the two conservative fluid equations, conservative Lorenz-96 and Galerkin approximations of 2d Euler, and apply the results of Chapters 3, 4, and 5 to prove these random splittings are ergodic in a certain sense, that they converge in the manner of Chapter 4, and that their top Lyapunov exponent is positive. In Chapter 7 we construct random splittings of the nonconservative counterparts of the Chapter 6 equations, namely Lorenz-96 and Galerkin approximations of 2d Navier-Stokes, and again apply the results of Chapters 3 and 4 to prove these random splittings are ergodic and converge. Certain technical aspects of the convergence results in Chapter 4 are given in Appendix A, and those for some of the 2d Euler results in Chapter 6 are given in Appendices B and C.

2

Random splitting

Let $\mathcal{V} := \{V_j\}_{j=1}^n$ be a family of vector fields on $\mathbb{R}^D$. The *flow* $\varphi^{(j)} : \mathbb{R}^D \times \mathbb{R} \to \mathbb{R}^D$ of V_j is the solution of $\dot{x} = V_j(x)$. That is, $\varphi_t^{(j)}(x) := \varphi^{(j)}(x,t)$ satisfies

$$\partial_t \left(\varphi^{(j)}(x,t)\right) = V_j\left(\varphi_t^{(j)}(x)\right).$$

We assume throughout this dissertation that the V_j are *complete* so every $\varphi^{(j)}$ is well-defined for all initial conditions and all time. We also assume the V_j are smooth, i.e. infinitely differentiable, though some results will specify weaker or stronger regularity as appropriate. For the applications considered here, the family $\mathcal{V}$ constitutes a splitting of a vector field V on $\mathbb{R}^D$. That is, $V = \sum_{j=1}^n V_j$. In this case we refer to V as the *true vector field* and to the V_j as *splitting vector fields*. For example, in the Lorenz-63 splitting given in (1.2) the splitting vector fields were $\mathcal{V} = \{V_1, V_2\}$ and the true vector field was $V = V_1 + V_2$.

Given $\mathcal{V}$, define the composition $\Phi : \mathbb{R}^D \times \mathbb{R}^n \to \mathbb{R}^D$ of the $\varphi^{(j)}$ by

$$\Phi(x,t) := \Phi_t(x) := \varphi_{t_n}^{(n)} \circ \cdots \circ \varphi_{t_1}^{(1)}(x).$$

Similarly, define $\Phi^m : \mathbb{R}^D \times \mathbb{R}^{mn} \to \mathbb{R}^D$ by $\Phi^m(x,t) := \Phi_t^m(x) := \varphi_{t_{mn}}^{(n)} \circ \cdots \circ \varphi_{t_1}^{(1)}(x)$ with superscripts cycling in order from 1 to n. Our convention throughout is that the j in $\varphi^{(j)}$ and V_j are implicitly taken mod n if k mod $n \neq 0$ and are n otherwise. For example, if $n = 3$,

$$\varphi^{(6)} \circ \varphi^{(5)} \circ \varphi^{(4)} \circ \varphi^{(3)} \circ \varphi^{(2)} \circ \varphi^{(1)} = \varphi^{(3)} \circ \varphi^{(2)} \circ \varphi^{(1)} \circ \varphi^{(3)} \circ \varphi^{(2)} \circ \varphi^{(1)}.$$

Also, the t in Φ_t^m always belongs to $\mathbb{R}^{mn}$ or, more generally, $t = (t_j)_{j=1}^{\infty}$, so that

$$\Phi_t^m(x) = \varphi_{t_{mn}}^{(n)} \circ \cdots \circ \varphi_{t_1}^{(1)}(x)$$

is a composition of mn flows.

Though our ambient space is $\mathbb{R}^D$, we often restrict attention to subsets affiliated with the family $\mathcal{V}$ called *$\mathcal{V}$-orbits*. Specifically, for each x in $\mathbb{R}^D$ the *$\mathcal{V}$-orbit of x* is

$$\mathcal{X}(x) := \left\{\Phi^m(x,t) : m \geqslant 0, t \in \mathbb{R}^{mn}\right\}. \tag{2.1}$$

This is the set of points that can be reached by $\{\Phi^m\}_{m=0}^{\infty}$ from x in any finite number of steps and over all times. When x is arbitrary or understood, we denote $\mathcal{X}(x)$ by $\mathcal{X}$. Since the V_j are complete, the $\mathcal{V}$-orbits corresponding to different x are identical or disjoint. Hence we have an equivalence relation $x \sim y$ if and only if $\mathcal{X}(x) = \mathcal{X}(y)$, and the $\mathcal{V}$-orbits $\{\mathcal{X}(x) : x \in \mathbb{R}^D\}$ partition $\mathbb{R}^D$. A classic result from geometric control theory ensures every $\mathcal{X}$ is a smooth submanifold of $\mathbb{R}^D$ [40]. In particular, each $\mathcal{X}$ has a Riemannian structure induced by the ambient Euclidean structure on $\mathbb{R}^D$ and an associated volume form v, sometimes called *Lebesgue or Hausdorff measure on $\mathcal{X}$*, which will serve as our reference measure on $\mathcal{X}$.

To introduce randomness, fix $h > 0$ and let $\tau = (\tau_j)_{j=1}^{\infty}$ be a collection of mutually independent real-valued random variables with mean 1 and common distribution ρ. We assume throughout that ρ has a continuous density which, by a slight abuse of notation, is also denoted by ρ, i.e. $\rho(dt) = \rho(t)dt$. We also assume the support of ρ contains an interval $(0, \varepsilon)$ for some $\varepsilon > 0$. A canonical choice is the exponential distribution with mean 1.

Definition 2.1. *Set* $h\tau := (h\tau_j)_{j=1}^{\infty}$. *The* random splitting associated to $\mathcal{V}$ and ρ, *or just* random splitting, *is the sequence* $\{\Phi_{h\tau}^m\}_{m=0}^{\infty}$ *where* $\Phi_{h\tau}^0$ *is the identity and*

$$\Phi_{h\tau}^m := \varphi_{h\tau_{mn}}^{(n)} \circ \cdots \circ \varphi_{h\tau_{(m-1)n+1}}^{(1)}(\Phi_{h\tau}^{m-1}). \tag{2.2}$$

That is, starting from the current step, the next step of the random splitting is obtained by flowing by each V_j *for the random time* $h\tau_j$ *in order from* $j = 1$ *to* n.

Independence of the τ_j and (2.1) imply $\{\Phi_{h\tau}^m\}$ is a Markov chain on $\mathcal{X}$ whenever it starts on $\mathcal{X}$. Its transition kernel $P_h : \mathcal{X} \times \mathcal{B}(\mathcal{X}) \to [0, 1]$ is defined by

$$P_h(x, B) := \mathbb{P}\left(\Phi_{h\tau}(x) \in B\right) = \mathbb{E}\left(\mathbb{1}_B(\Phi_{h\tau}(x))\right). \tag{2.3}$$

Here $\mathcal{B}(\mathcal{X})$ is the Borel σ-algebra on $\mathcal{X}$ and $\mathbb{1}_B$ the indicator function on B. P_h acts on measurable[1] $f : \mathcal{X} \to \mathbb{R}$ by $P_h f(x) := \mathbb{E}f(\Phi_{h\tau}^m(x))$ and on measures μ on $\mathcal{X}$ by

$$\mu P_h(f) := \int_{\mathcal{X}} P_h f(x)\mu(dx).$$

Note we are using the fact that measures μ on $\mathcal{X}$ act on measurable functions f on $\mathcal{X}$ by $\mu(f) := \int f d\mu$. Thus $P_h f$ is a measurable function on $\mathcal{X}$ and μP_h is a measure on $\mathcal{X}$ for all h. The P_h-*invariant* measures will play an important role in our results on ergodicity and chaos.

Definition 2.2. *A measure* μ *on* $\mathcal{X}$ *is* P_h-invariant *if* $\mu P_h = \mu$.

Remark 2.3. *All results in this work remain true if at each step of the random splitting we randomly permute indices in the composition* Φ. *That is, given a current state* x, *the next step is* $\varphi_{h\tau_n}^{(\sigma(n))} \circ \cdots \circ \varphi_{h\tau_1}^{(\sigma(1))}(x)$ *where* σ *is a random permutation of* $\{1, \ldots, n\}$. *This yields both additional randomness and an avenue to higher order approximations of the true dynamics [18, 19, 44, 63, 64]. We forgo this more general setting however to keep exposition more approachable and notationally light.*

[1] *Measurable* will always mean with respect to the Borel σ-algebra.

2.1 The Lie bracket condition

The *Lie bracket condition* defined below is a condition on families of vector fields that implies a certain nondegeneracy of their random splitting dynamics. This condition plays a significant role in our general results on ergodicity and chaos in Chapters 3 and 5, and in verifying these properties for the models considered in Chapters 6 and 7.

Let $\mathcal{V}$ be a family of smooth vector fields as above and fix a d-dimensional $\mathcal{V}$-orbit $\mathcal{X}$. The vector space $\mathfrak{X}(\mathcal{X})$ of smooth vector fields on $\mathcal{X}$ is a Lie algebra when equipped with the *Lie bracket*, $[V, W](f) := V(W(f)) - W(V(f))$, where V and W are in $\mathfrak{X}(\mathcal{X})$ and f belongs to the space $\mathcal{C}^\infty(\mathcal{X})$ of smooth real-valued functions on $\mathcal{X}$. Here V and W are regarded as *derivations* on $\mathcal{X}$, i.e. as linear maps from $\mathcal{C}^\infty(\mathcal{X})$ to $\mathcal{C}^\infty(\mathcal{X})$ satisfying $V(fg) = fV(g) + gV(f)$. Define $\mathrm{Lie}(\mathcal{V})$ to be the smallest subalgebra of the Lie algebra $\mathfrak{X}(\mathcal{X})$ containing $\mathcal{V}$. For each x in $\mathcal{X}$ the collection $\mathrm{Lie}_x(\mathcal{V}) := \{V(x) : V \in \mathrm{Lie}(\mathcal{V})\}$ is a subspace of $T_x\mathcal{X}$.

Definition 2.4. *The* Lie bracket condition *holds at x in $\mathcal{X}$ if* $\mathrm{Lie}_x(\mathcal{V}) = T_x\mathcal{X}$*.*

The Lie bracket condition is called the *weak bracket condition* in [12] and *Condition B* in [6]. Both papers also consider a *strong bracket condition* (*Condition A*) which is used for results on continuous time Markov processes and therefore not needed here. The Lie bracket condition has the following important consequence.

Theorem 2.5. *If the Lie bracket condition holds at x_* then for every neighborhood U of x_* and every $T > 0$ there exists an x in U, an m, and a t in $\mathbb{R}_+^{mn}$ such that $\sum_{j=1}^{mn} t_j \leqslant T$ and $t \mapsto \Phi^m(x_*, t) = x$ is a submersion at t, i.e. $D_t\Phi^m(x_*, t) : T_t\mathbb{R}^{mn} \to T_x\mathcal{X}$ is surjective.*

Here and throughout $\mathbb{R}_+ := (0, \infty)$. A version of Theorem 2.5 appears as Theorem 3.1 in [40]; the equivalent version here is better suited to random splitting and other classes of piecewise deterministic Markov processes. See Theorem 5 in [6] and its subsequent discussion for details. Intuitively, Theorem 2.5 says that if the Lie bracket condition holds at x_* then, because of surjectivity, the random splitting can move in any infinitesimal direction from x_* in arbitrarily small positive times.

2.2 Real analyticity

Many differential equations, including the Lorenz-63 equations and those studied in Chapters 6 and 7, correspond to vector fields that are not only smooth, but analytic[2]. When this is the case, additional results are available. We present certain of these here. Informally, an analytic function is one that equals its Taylor series at all points in its domain. Formally, we have

Definition 2.6. *A smooth function* $f : \mathbb{R} \to \mathbb{R}$ *is* analytic *if for every* x *in* $\mathbb{R}$ *the Taylor series* $\sum f^{(k)}(x)(y-x)^k/k!$ *converges to* $f(y)$ *for all* y *in a neighborhood of* x*. A vector field* $V : \mathbb{R}^D \to \mathbb{R}^D$ *is* analytic *if each of its coordinate functions is analytic.*

All polynomials, the exponential function, and common trigonometric functions such as sine and cosine are analytic. In particular, all vector fields considered in this work are analytic. When this is the case, $\mathcal{V}$-orbits are analytic submanifolds of $\mathbb{R}^D$ [40]. Furthermore, we have the following useful result due to Nagano. See [40, 61] for discussion and proof.

Theorem 2.7. *Suppose the vector fields in* $\mathcal{V}$ *are analytic. If the Lie bracket condition holds at one point in* $\mathcal{X}$*, then it holds at every point in* $\mathcal{X}$*.*

So to check the Lie bracket condition anywhere in $\mathcal{X}$, it suffices in the analytic setting to check it holds at *any* point in $\mathcal{X}$. This will be tremendously useful in later examples where verifying the Lie bracket condition will be significantly easier at some points than at others.

[2] Throughout this thesis *analytic* will mean *real analytic*

2.3 Related work

Random splitting is an example of a random dynamical system [5] or an integrated random function [24], and closely related to piecewise deterministic Markov process when the times are chosen to be exponentially distributed and the ordering of indices follows a Markov process [6, 12]. We continue to describe the system as a random splitting, however, to emphasize its underlying physical motivation, i.e. its relation to the true vector field V, and the corresponding physical structure that plays a central role in the following analysis.

Our splitting into fundamental building blocks is partially motivated by the classical stylized models of dynamics studied in depth at the dawn of the theory of dynamical systems. Examples include the doubling map, quadratic maps, the Henon map, the Smale horseshoe, and extended systems such as coupled map lattices. See [23, 41, 67] and references therein. The form of the decomposition is also motivated by recent progress in proving ergodic properties of piecewise deterministic Markov processes (PDMPs) and their success as modeling and sampling tools. See for example [6, 7, 8, 12, 13, 15, 25, 45, 47, 59].

Random splitting is also inspired by and resembles traditional operator splitting discussed in Chapters 1 and 4. This classical numerical analysis technique is often used in numerical simulations of various ordinary, partial, and stochastic differential equations [4, 14, 18, 19, 34, 44, 51, 63, 64]. Typically the goal is to construct split dynamics that are more computationally tractable than the true dynamics to obtain an efficient and accurate numerical method. A variant of these models was also explored in [69].

Further information about the geometric concepts used throughout this work, such as vector fields on smooth manifolds and Lie brackets, can be found in most differential topology and differential geometry texts. One comprehensive source is

[46]. The discussion about the Lie bracket in the context of piecewise deterministic Markov processes in [12] is especially insightful. See also [6].

3

Ergodicity

Let $\mathcal{V} = \{V_j\}_{j=1}^n$ be a family of smooth vector fields on $\mathbb{R}^D$ and fix a d-dimensional orbit $\mathcal{X}$. Also fix $h > 0$ and let P_h be the transition kernel defined in (5.13). Recall a measure μ on $\mathcal{X}$ is P_h-invariant if $\mu P_h = \mu$. Thus if the random splitting system is distributed according to μ, it will remain so for all time. As we prove for the models in Chapter 7, it is sometimes the case that even when the system starts out of equilibrium, i.e. according to any noninvariant initial measure, its distribution under random splitting converges to an invariant one. A common intuitive example is that of smoke filling a room; if the system is ergodic then, loosely speaking, one can infer properties of the entire system of smoke particles by observing just one particle for a sufficiently long time. The same idea lies at the heart of Monte Carlo methods and is the content of *Birkhoff's ergodic theorem* [16]. Hence invariant measures shed light on the long-time behavior of systems. However, if a transition kernel admits multiple invariant measures it may be unclear which distribution, if any, the system will look like in the long run. This of course is not an issue if the system admits a unique invariant measure, in which case we call the system *uniquely ergodic*.

This nomenclature follows from the fact that the space of P_h-invariant probability measures is convex and the extremal points of this set are precisely the ergodic invariant measures. Thus if there is only one invariant measure, it is necessarily extremal and therefore ergodic. For more on ergodicity of Markov processes and proof of the aforementioned claim see [21, 36].

3.1 Results

Informally, unique ergodicity holds if a system can be sufficiently explored by its dynamics. The rest of this brief chapter is dedicated to making this statement precise for random splitting. We present three related sufficient conditions for P_h to have at most one invariant measure on $\mathcal{X}$. In each case the unique invariant measure, provided it exists, is guaranteed to be absolutely continuous with respect to the volume form on $\mathcal{X}$, which we denote by v. The results presented here are closely related to the analogous results for piecewise determinsitic Markov processes in [6, 12] and will be especially relevant for our study of Lyapunov exponents in Chapter 5 – see the *multiplicative ergodic theorem* discussed therein – and in Chapter 6.

Theorem 3.1. *If there exists x_* in $\mathcal{X}$ such that for all x in $\mathcal{X}$ there is an m and a t in $\mathbb{R}_+^{mn}$ with $\Phi^m(x,t) = x_*$ and $D_t\Phi^m(x,t) : T_t\mathbb{R}_+^{mn} \to T_{x_*}\mathcal{X}$ surjective, then P_h has at most one invariant measure on $\mathcal{X}$. Moreover, if such a measure exists, it is absolutely continuous with respect to the volume form v on $\mathcal{X}$.*

$T_{x_*}\mathcal{X}$ is the tangent space of $\mathcal{X}$ at x_*. The proof of Theorem 3.1 follows from the classical minorization condition [38, 54, 57, 65] given by the following result from [12, Lemma 6.3].

Lemma 3.2. *Let $p \leqslant m$ and let $F : \mathcal{X} \times U \to \mathcal{X}$ be $\mathcal{C}^1$, where U is an open subset of $\mathbb{R}^m$. Suppose τ is a U-valued random variable with continuous density ρ. If for some (x,t) in $\mathcal{X} \times U$ the map $D_tF(x,t)$ is surjective and ρ is bounded below by $c_0 > 0$ on* *a neighborhood of t, then there exists a* $c > 0$ *and neighborhoods* U_x *of* x *and* U_* *of* $x_* := F(x,t)$ *such that*

$$\mathbb{P}\big(F(y,\tau) \in B\big) \geqslant c\,\mathrm{v}(B \cap U_*) \tag{3.1}$$

for all y *in* U_x *and* B *in the Borel* σ*-algebra* $\mathcal{B}(\mathcal{X})$ *of* $\mathcal{X}$.

Remark 3.3. *In our setting,* $U = \mathbb{R}_+^{mn}$, $F = \Phi^m : \mathcal{X} \times \mathbb{R}_+^{mn} \to \mathcal{X}$, *and* $\tau = (\tau_1, \ldots, \tau_{mn})$ *with the* τ_j *independent random variables with mean* h*. In this case, if* $x_* = \Phi^m(x,t)$ *for some* t *with* $D_t\Phi^m(x,t)$ *surjective, then Lemma 3.2 guarantees the existence of a constant* $c > 0$ *and neighborhoods* U_x *of* x *and* U_* *of* x_* *such that, for all* y *in* U_x *and* B *in* $\mathcal{B}(\mathcal{X})$,

$$P^m(y,B) \geqslant c\,\mathrm{v}(B \cap U_*)\,.$$

Proof of Theorem 3.1. The proof is by contradiction. Suppose μ_1 and μ_2 are distinct P_h-invariant probability measures. Assume without loss of generality both μ_i are ergodic and therefore mutually singular [21, 42]. Then there exist disjoint measurable sets A_1 and A_2 partitioning $\mathcal{X}$ such that $\mu_i(B) = \mu_i(B \cap A_i)$ for all B in $\mathcal{B}(\mathcal{X})$. Fix x_i in the support of μ_i so, by definition, μ_i gives positive measure to every neighborhood of x_i. By hypothesis and Remark 3.3 there exist $c_i > 0$, m_i in $\mathbb{N}$, and neighborhoods U_i of x_i and U_* of x_* such that $P_h^{m_i}(x,\cdot) \geqslant c_i\,\mathrm{v}(\cdot \cap U_*)$ for all x in U_i. Therefore

$$\mu_i(B) = \mu_i P_h^{m_i}(B) \geqslant \int_{U_i} P_h^{m_i}(x,B)\mu_i(dx) \geqslant c_i\,\mathrm{v}(B \cap U_*)\mu_i(U_i) \tag{3.2}$$

for all B in $\mathcal{B}(\mathcal{X})$. In particular, $\mu_i(B) = 0$ implies $\mathrm{v}(B \cap U_*) = 0$ since c_i and $\mu_i(U_i)$ are strictly positive. But $\mu_1(A_2 \cap U_*) = \mu_2(A_1 \cap U_*) = 0$ and hence

$$0 < \mathrm{v}(U_*) = \mathrm{v}(A_1 \cap U_*) + \mathrm{v}(A_2 \cap U_*) = 0,$$

a contradiction. Absolute continuity of the P_h-invariant measure μ, provided it exists, follows from uniqueness and the fact that the absolutely continuous part, μ_{ac}, and singular part, μ_s, of μ are P_h-invariant whenever μ is [12, Proposition 2.7]. Specifically, since μ_{ac} and μ_s are P_h-invariant and there can be at most one P_h-invariant probability measure, either μ_{ac} or μ_s is identically zero. Since μ_{ac} is nonzero by (3.2), it follows that $\mu_s = 0$ and hence $\mu = \mu_{ac}$. □

Remark 3.4. *The invariant measure μ in Theorem 3.1, provided it exists, is sometimes called a* stationary measure*. This is because the sequence of random variables generated by the Markov chain starting from an initial condition distributed according to μ will be stationary. This helps distinguish from the invariant measure of the skew flow $(x, \tau) \mapsto (\Psi_{h\tau}(x), \vartheta\tau)$ where the shift ϑ is defined by $\vartheta\tau : \tau = (\tau_1, \tau_2, \cdots) \mapsto (\tau_{n+1}, \tau_{n+2}, \cdots)$.*

The following corollary of Theorem 3.1 highlights the significance of the Lie bracket condition.

Corollary 3.5. *Suppose there is an x_* in $\mathcal{X}$ at which the Lie bracket condition holds and such that for every x in $\mathcal{X}$ there is an m and a t in $\mathbb{R}_+^{mn}$ satisfying $\Phi^m(x, t) = x_*$. Then P_h has at most one invariant measure on $\mathcal{X}$. Furthermore, if such a measure exists, it is absolutely continuous with respect to the volume form on $\mathcal{X}$.*

One benefit of Corollary 3.5 is that it replaces the need to check the surjectivity assumption of Theorem 3.1, which can be challenging in practice, with verification of the Lie bracket condition. The next result provides a further convenience in the analytic setting.

Corollary 3.6. *Suppose the vector fields in $\mathcal{V}$ are analytic and there is an x_* in $\mathcal{X}$ such that for every x in $\mathcal{X}$ there is an m and a t in $\mathbb{R}_+^{mn}$ satisfying $\Phi^m(x, t) = x_*$. If the Lie bracket condition holds at any point in $\mathcal{X}$, then P_h has at most one invariant measure on $\mathcal{X}$ which, provided it exists, is absolutely continuous with respect to the volume form on $\mathcal{X}$.*

Proof. Since the Lie bracket condition holds at one point in $\mathcal{X}$, it also holds at x_* by Nagano's theorem (Theorem 2.7). The result follows by Corollary 3.5. $\square$

3.2 Related work

The results of this chapter largely follow work on piecewise deterministic Markov processes in [6, 12] and geometric control theory in [40, 61, 68]. The former are especially relevant to our results on unique ergodicity, and the latter, especially [40], elaborate considerably on $\mathcal{V}$-orbits and other general geometric concepts that underlie random splitting.

4

Convergence

Throughout this chapter we consider a splitting of the differential equation

$$\dot{x} = V(x) = \sum_{j=1}^{n} V_j(x) \tag{4.1}$$

on $\mathbb{R}^D$ where, as before, V is the *true vector field* with flow ψ and $\mathcal{V} := \{V_j\}_{j=1}^{n}$ are the *splitting vector fields* with flows $\varphi^{(j)}$. For the forthcoming convergence results it suffices to assume the V_j are merely $\mathcal{C}^2$, i.e. twice differentiable with continuous second derivatives. As mentioned in our discussion of the Lorenz-63 equations, it is well established that under quite general conditions the error incurred in approximating the flow ψ_h of V by $\Phi_h = \varphi_h^{(n)} \circ \cdots \circ \varphi_h^{(1)}$ is $O(h^2)$ and hence the error in approximating ψ_t by Φ_h^m for any finite t with $t = mh$ is $O(h)$ [51]. That is, trajectories of the deterministic splitting Φ_h^m converge to the true trajectory ψ_t at worst linearly in h as $h \to 0$ for any finite t. This splitting scheme is therefore called a *first-order method*; higher order methods also exist but will not be considered here [33].

In this chapter we give analogous convergence results for the random splitting associated to (4.1); the pluralized "results" reflects that with randomness comes several different notions of convergence. Specifically, we give two main results. First, as in the deterministic case, the transition kernel P_h converges to the true dynamics linearly in h as $h \to 0$. Second, random splitting converges almost-surely to the true dynamics as $h \to 0$. Precise statements are given in Theorems 4.1 and 4.4, respectively, but first we introduce the relevant setting.

4.1 Preliminaries

We begin with a simple assumption.

Assumption 1. *The $\mathcal{V}$-orbits $\mathcal{X}(x)$ are bounded for each x in $\mathbb{R}^D$.*

Since the vector fields V_j are assumed $\mathcal{C}^2$, Assumption 1 implies the V_j are bounded with bounded first and second derivatives on every $\mathcal{X}$. In particular,

$$C_*(x_0) := \sup_{x \in \mathcal{X}(x_0)} \left\{ \|V_j(x)\|, \|DV_j(x)\|, \|D^2 V_j(x)\| : 1 \leqslant j \leqslant n \right\} < \infty, \tag{4.2}$$

where $\|V_j(x)\|$ is the usual Euclidean norm, $\|DV_j(x)\|$ is the operator norm of $DV_j(x) : \mathbb{R}^D \to \mathbb{R}^D$, and $\|D^2 V_j(x)\|$ is the operator norm of the bilinear map $D^2 V_j(x) : \mathbb{R}^D \times \mathbb{R}^D \to \mathbb{R}^D$. Assumption 1 will hold for the conservative models considered in Chapter 6 but not the nonconservative models in Chapter 7. However, the forthcoming convergence results will still hold in a slightly altered but virtually equivalent sense; see Remark 7.2.

For a positive integer k let $\mathcal{C}^k(\mathcal{X})$ be the space of k-times continuously differentiable functions $f : \mathcal{X} \to \mathbb{R}$. For f in $\mathcal{C}^k(\mathcal{X})$ and $\ell \leqslant k$, the ℓth derivative $D^\ell f(x)$ of f at x is a multilinear operator from $\otimes_1^\ell T_x\mathcal{X}$ to $\mathbb{R}$. The operator norm of $D^\ell f(x)$ is then

$$\|D^\ell f(x)\| := \sup_{\|\eta\|=1} \left\{ |D^\ell f(x)\eta| \right\},$$

where $\eta \in \otimes_1^\ell T_x\mathcal{X}$. Defining $D^0 f(x) := f(x)$, this in turn induces a norm on $\mathcal{C}^k(\mathcal{X})$ given by

$$\|f\|_k := \sup_{x \in \mathcal{X}} \left\{ \|D^\ell f(x)\| : 0 \leqslant \ell \leqslant k \right\}.$$

The corresponding operator norm is denoted $\|\cdot\|_{k\to k}$. More generally, for any k and ℓ define a norm $\|\cdot\|_{k\to\ell}$ on the space of linear operators $L : \mathcal{C}^k(\mathcal{X}) \to \mathcal{C}^\ell(\mathcal{X})$ by

$$\|L\|_{k\to\ell} := \sup_{\|f\|_k=1} \|Lf\|_\ell \,.$$

We make frequent use of the *submultiplicity* of $\|\cdot\|_{k\to\ell}$. Namely, if A and B are bounded linear operators from $\mathcal{C}^j(\mathcal{X})$ to $\mathcal{C}^k(\mathcal{X})$ and from $\mathcal{C}^k(\mathcal{X})$ to $\mathcal{C}^\ell(\mathcal{X})$, respectively, then

$$\|BA\|_{j\to\ell} \leqslant \|B\|_{k\to\ell}\|A\|_{j\to k} \,.$$

The results below are stated in terms of semigroups of the flows ψ and $\varphi^{(j)}$, which are $\mathcal{C}^2$ by assumption. For all $k \leqslant 2$ the semigroup $\{S_t\}_{t\geqslant 0}$ corresponding to ψ acts on f in $\mathcal{C}^k(\mathcal{X})$ via

$$S_t f(x) = e^{tV} f(x) = f(\psi_t(x)) \tag{4.3}$$

and, similarly, the semigroup $\{\widetilde{S}_t^{(j)}\}_{t\geqslant 0}$ corresponding to $\varphi^{(j)}$ is given by

$$\widetilde{S}_t^{(j)} f(x) = e^{tV_j} f(x) = f(\varphi_t^{(j)}(x)). \tag{4.4}$$

In particular, m steps of random splitting corresponds to $\widetilde{S}^m_{h\tau} := \widetilde{S}^{(1)}_{h\tau_1} \cdots \widetilde{S}^{(mn)}_{h\tau_{mn}}$. The transition kernel P_h^m and semigroup composition $\widetilde{S}^m_{h\tau}$ are related via

$$P_h^m f = \mathbb{E}(f(\Phi^m_{h\tau})) = \mathbb{E}(\widetilde{S}^m_{h\tau} f) \,. \tag{4.5}$$

4.2 Results

With the above notation we now present the two main results of this section, Theorems 4.1 and 4.4, which follow from Lemmas 4.2 and 4.5, respectively. Full proofs of both lemmas are in Appendix A, but we discuss the general idea behind each at the end of this section.

Theorem 4.1. *Suppose Assumption 1 holds and fix $t > 0$. For all h sufficiently small and satisfying $mh = t$ for some m in $\mathbb{N}$, there is a constant $C(t)$ depending only on t such that*

$$\|P_h^m - S_t\|_{2\to 0} \leqslant C(t)h. \tag{4.6}$$

Lemma 4.2. *If Assumption 1 holds then there exists a constant C such that*

$$\|P_h - S_h\|_{2\to 0} \leqslant Ch^2 \tag{4.7}$$

for all h sufficiently small.

Recalling from (4.9) that $P_h = \mathbb{E}(\widetilde{S}^1_{h\tau})$, informally Lemma 4.2 states that the average difference between one step of random splitting and the true dynamics is $O(h^2)$ for sufficiently small h, where here and throughout $O(h)$ is with respect to the relevant norm, namely $\|\cdot\|_{2\to 0}$. For any finite time interval $[0, t]$ we can leverage this result to approximate S_t by successive steps of P_h. Specifically, choose h sufficiently small so that (4.7) holds and there exists an integer m with $mh = t$. Then the composition P_h^m corresponds to $O(1/h)$ steps of P_h. Since the difference between P_h and S_h is $O(h^2)$, the difference between P_h^m and S_t is $O(h)$.

Proof of Theorem 4.1. Let h be sufficiently small that (4.7) holds and such that $mh = t$ for some m in $\mathbb{N}$. The quantity of interest can be written as the following telescoping sum:

$$P_h^m - S_t = \sum_{k=1}^{m} P_h^{k-1}(P_h - S_h)S_{h(m-k)}\,. \tag{4.8}$$

For any k and continuous function f with $\|f\|_0 = 1$,

$$\|P_h^k f\|_0 \leqslant \mathbb{E}\left(\left\|f\left(\Phi^k_{h\tau}\right)\right\|_0\right) = 1\,.$$

So $\|P_h^k\|_{0\to 0} = 1$. Similarly, since $mh = t$ implies $h(m-k) \leqslant t$ for $k \geqslant 0$ and $\mathcal{X}$ is bounded by assumption (so ψ and its first and second derivatives are bounded on $\mathcal{X}$, uniformly on $[0, t]$),

$$\|S_{h(m-k)}\|_{2\to 2} \leqslant K(t)$$

for some $K(t)$ depending on t but not h. Hence, by submultiplicity, (4.8), and Lemma 4.2,

$$\|P_h^m - S_t\|_{2\to 0} \leqslant \sum_{k=1}^{m} \|P_h^{k-1}\|_{0\to 0}\|P_h - S_h\|_{2\to 0}\|S_{h(m-k)}\|_{2\to 2} \leqslant K(t)\sum_{k=1}^{m} Ch^2 = C(t)h\,,$$

where $C(t) := K(t)C$, with C the constant from (4.7) in Lemma 4.2. $\square$

Remark 4.3. *Theorem 4.1 had the relation $h = t/m$, while in the almost-sure results below we will take $h = t/m^2$ (note we explicitly write t/m^2, making no reference to the variable h). The reason, loosely speaking, is that the transition kernel depends only on the expectation of the randomness, while the almost-sure results additionally depend on fluctuations of the randomness about its mean. For example, Lemma 4.5 prepares for an application of the Borel-Cantelli lemma by establishing the summability of probabilities of "large" fluctuations over sets of $O(m) = O(1/\sqrt{h})$ cycles. This is discussed in more detail at the end of this section and worked out in full in Appendix A.*

Theorem 4.4. *Suppose Assumption 1 holds and fix $t > 0$. Then for any $\varepsilon > 0$,*

$$\mathbb{P}\left(\limsup_{m\to\infty}\|\widetilde{S}^{m^2}_{t\tau/m^2} - S_t\|_{2\to 0} > \varepsilon\right) = 0\,. \tag{4.9}$$

Lemma 4.5. *Suppose Assumption 1 holds and fix $t > 0$. Then for any $\varepsilon > 0$,*

$$\sum_{m=1}^{\infty} \mathbb{P}\left(\|\widetilde{S}^{m}_{t\tau/m^2} - S_{t/m}\|_{2\to 0} > \tfrac{\varepsilon}{m}\right) < \infty. \tag{4.10}$$

Proof of Theorem 4.4. By the Borel-Cantelli Lemma it suffices to show

$$\sum_{m=1}^{\infty} \mathbb{P}\left(\|\widetilde{S}^{m^2}_{t\tau/m^2} - S_t\|_{2\to 0} > \varepsilon\right) < \infty\,.$$

Consider the telescoping sum

$$\widetilde{S}^{m^2}_{t\tau/m^2} - S_t = \sum_{k=1}^{m} \widetilde{S}^{(k-1)}_{t\tau/m^2}\left(\widetilde{S}^{m}_{t\tau/m^2} - S_{t/m}\right) S_{(m-k)t/m}\,. \tag{4.11}$$

For any k and continuous function f with $\|f\|_0 = 1$,

$$\left\|\widetilde{S}^{k}_{t\tau/m^2} f\right\|_0 = \left\|f\left(\Phi^k_{h\tau}\right)\right\|_0 = 1.$$

So $\|\widetilde{S}^{(k-1)}_{t\tau/m^2}\|_{0\to 0} = 1$. Similarly, since $(m-k)t/m \leqslant t$ for $k \geqslant 0$ and $\mathcal{X}$ is bounded by assumption (so ψ and its first and second derivatives are bounded on $\mathcal{X}$, uniformly on $[0,t]$),

$$\|S_{(m-k)t/m}\|_{2\to 2} \leqslant K(t)$$

for some $K(t)$ depending on t but not h. Hence, by submultiplicity, (4.11), and Lemma 4.5,

$$\left\|\widetilde{S}^{m^2}_{t\tau/m^2} - S_t\right\|_{2\to 0} \leqslant K(t)\sum_{k=1}^{m}\left\|\widetilde{S}^{m}_{t\tau/m^2} - S_{t/m}\right\|_{2\to 0} = K(t)m\left\|\widetilde{S}^{m}_{t\tau/m^2} - S_{t/m}\right\|_{2\to 0},$$

and hence by Lemma 4.5,

$$\sum_{m=1}^{\infty}\mathbb{P}\left(\left\|\widetilde{S}^{m^2}_{t\tau/m^2} - S_t\right\|_{2\to 0} > \varepsilon\right) \leqslant \sum_{m=1}^{\infty}\mathbb{P}\left(\left\|\widetilde{S}^{m}_{t\tau/m^2} - S_{t/m}\right\|_{2\to 0} > \tfrac{\varepsilon}{K(t)m}\right) < \infty. \qquad \square$$

We conclude by sketching the proofs of Lemmas 4.2 and 4.5, which are inspired by ideas from [18, 19] and given in full detail in Appendix A. In what follows we set $\widetilde{S}_{h\tau} := \widetilde{S}^1_{h\tau}$ and define $\widetilde{S}^{(i,j)}_{h\tau} := \widetilde{S}^{(i)}_{h\tau}\cdots\widetilde{S}^{(j)}_{h\tau}$. Consider first Lemma 4.2. Differentiating $\widetilde{S}_{h\tau}$ in h gives

$$\partial_h\widetilde{S}_{h\tau} = \sum_{k=1}^{n}\tau_k e^{h\tau_1}\cdots e^{h\tau_{k-1}}V_k e^{h\tau_k}\cdots e^{h\tau_n} = \sum_{k=1}^{n}\tau_k\widetilde{S}^{(1,k-1)}_{h\tau}V_k\widetilde{S}^{(k,n)}_{h\tau}.$$

Commuting $\widetilde{S}^{(1,k-1)}_{h\tau}$ and V_k via the Lie bracket $[\widetilde{S}^{(1,k-1)}_{h\tau}, V_k] := \widetilde{S}^{(1,k-1)}_{h\tau}V_k - V_k\widetilde{S}^{(1,k-1)}_{h\tau}$ gives

$$\partial_h\widetilde{S}_{h\tau} = \sum_{k=1}^{n}\tau_k V_k\widetilde{S}_{h\tau} + \sum_{k=1}^{n}\tau_k[\widetilde{S}^{(1,k-1)}_{h\tau}, V_k]\widetilde{S}^{(k,n)}_{h\tau} = V\widetilde{S}_{h\tau} + (V_\tau - V)\widetilde{S}_{h\tau} + E_{h\tau}$$

where $V_\tau := \sum_{k=1}^{n}\tau_k V_k$ and $E_{h\tau} := \sum_{k=1}^{n}\tau_k[\widetilde{S}^{(1,k-1)}_{h\tau}, V_k]\widetilde{S}^{(k,n)}_{h\tau}$. So, by variation of constants,

$$\widetilde{S}_{h\tau} - S_h = \int_0^h S_{h-r}(V_\tau - V)\widetilde{S}_{r\tau}dr + \int_0^h S_{h-r}E_{r\tau}dr. \tag{4.12}$$

Loosely speaking, the first integrand is $O(h)$ because

$$\mathbb{E}(V_\tau - V) = \sum_{k=1}^{n}\mathbb{E}(\tau_k - 1)V_k = 0 \tag{4.13}$$

cancels first order terms from the full expression, $S_{h-r}(V_\tau - V)\widetilde{S}_{r\tau}$. On the other hand the second integrand is $O(h)$ because the bracket terms in $E_{h\tau}$ also cancel first order terms (most of the work in the proof is making these two statements precise). Thus, integrating these $O(h)$ terms over the interval $(0, h)$, the difference on the right side of (4.12) is $O(h^2)$ as claimed.

The proof of Lemma 4.5 is structurally similar to the one sketched above in that it again begins with an application of variation of constants. However, in this case our analysis aims to establish a concentration estimate and can therefore not rely solely on the vanishing first moment as in (4.13). Instead, we expect the desired estimate to hold because of the averaging of the independent flow times τ_i in the analog of (4.13). In order to capture such averaging, we cannot limit our analysis to one cycle, but have to consider a variation of constants estimate on $m \gg 1$ such cycles:

$$\widetilde{S}^m_{h\tau} - S_{mh} = \int_0^h S_{m(h-r)}(V_\tau - V)\widetilde{S}^m_{r\tau}dr + \int_0^h S^m_{m(h-r)}E^{(m)}_{r\tau}dr \tag{4.14}$$

where now $E^{(m)}_{r\tau} := \sum_{k=1}^{mn} \tau_k[\widetilde{S}^{(1,k-1)}_{h\tau}, V_k]\widetilde{S}^{(k,n)}_{h\tau}$. Note that the second term contains $\mathcal{O}(m^2)$ commutators, each contributing $\mathcal{O}(h^2)$ as in the previous analysis. On the other hand, once integrated, the difference in the first integral, $\sum_{k=1}^{mn}(\tau_k - 1)V_k$, scales as $\mathcal{O}(\sqrt{m}h)$ by the central limit theorem. In order to have both terms decay faster than $\mathcal{O}(1/m)$ we choose $m \sim \mathcal{O}(1/\sqrt{h})$, whence the relation $h = t/m^2$.

4.3 Related work

As mentioned at the beginning of this chapter, the random splitting scheme considered in this work is a random analogue of a first-order operator splitting method. Higher order operator splitting schemes, i.e. those whose approximations on arbitrary time scales are $O(h^p)$ for $p > 1$, are available in most practical cases. An example of a second-order method is Strang splitting [51]. Higher order can also be obtained by fully randomizing the order [18] or randomly choosing between one ordering and its reverse [44, 63, 64]. Random splitting analogs of such higher order methods are certainly conceivable but are left to future work.

There is a relationship between Theorems 4.1 and 4.4 and the averaging results from Wentzell-Freidlin theory, e.g. [30, Theorem 2.1, Chapter 7]. This theorem builds on local results like Lemmas 4.2 and 4.5. Since our averaging is that of a deterministic, cyclic process, the calculations can be more explicit and more precise. We are able to prove using simple calculations that the local error is $\mathcal{O}(h^2)$ which leads to $\mathcal{O}(h)$ error over order one times. Typical soft averaging results prove a local error of $o(h)$ and then simply conclude that the order one error goes to zero. Of course, more careful calculations are possible in the averaging setting. However, the simple structure of our problems, where the only randomness is in the switching times and not the orderings, allows for the direct proofs presented here.

5

Chaos

There is no universally agreed upon mathematical definition of chaos [67, Chapter 9]. What is agreed upon, however, is that sensitive dependence on initial conditions, i.e. exponential separation of nearby trajectories, is a necessary feature of any system that can reasonably be called chaotic. To illustrate this idea, consider a random splitting $\{\Phi_{h\tau}^m\}_{m=0}^\infty$ associated to a family of vector fields $\mathcal{V}$ as before. Figure 5.1 shows how perturbations of a point x in a $\mathcal{V}$-orbit $\mathcal{X}$ are captured by tangent vectors in $T_x\mathcal{X}$. Indeed, in the Euclidean setting where adding points and tangent vectors is valid, Taylor's theorem gives

$$D\Phi_{h\tau}(x)\eta = \Phi_{h\tau}(x+\eta) - \Phi_{h\tau}(x) + O\left(\|\eta\|^2\right).$$

Thus for h and η sufficiently small, the distance between trajectories of Φ starting from x and the perturbed initial condition $x+\eta$ after one step is approximately $\|D\Phi_{h\tau}(x)\eta\|$. Supposing trajectories separate with constant exponential rate λ, we have $\|D\Phi_{h\tau}(x)\eta\| \approx e^\lambda\|\eta\|$,

$$\|D\Phi_{h\tau}^2(x)\eta\| = \|D\Phi_{h\tau}\left(\Phi_{h\tau}(x)\right) D\Phi_{h\tau}(x)\eta\| \approx e^\lambda\|D\Phi_{h\tau}(x)\eta\| \approx e^{2\lambda}\|\eta\|,$$

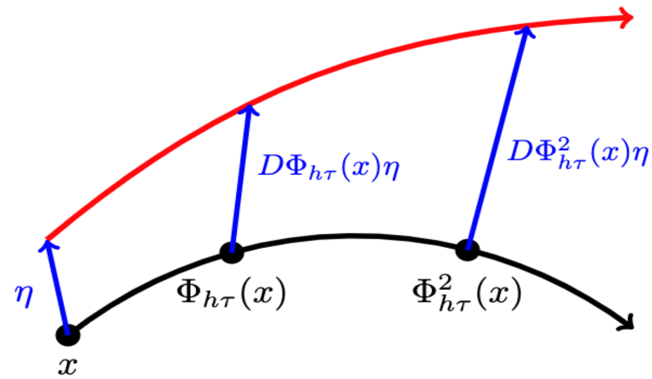


FIGURE 5.1: *Sensitive dependence on initial conditions.*

and, continuing in this manner, $\|D\Phi^m_{h\tau}(x)\eta\| \approx e^{m\lambda}\|\eta\|$. Rearranging and taking m to ∞,

$$\lambda = \lim_{m\to\infty} \tfrac{1}{m} \left(\log\|D\Phi^m_{h\tau}(x)\eta\| - \log\|\eta\|\right) = \lim_{m\to\infty} \tfrac{1}{m} \log\|D\Phi^m_{h\tau}(x)\eta\|. \tag{5.1}$$

The growth rate λ in (5.1), provided it exists, depends on multiple inputs including x, η, and the randomness τ. So a priori there can be infinitely many of them. Quite remarkably, the multiplicative ergodic theorem discussed below guarantees that if the random splitting admits an ergodic invariant measure μ, then there are at most d distinct growth rates $\lambda_1 \geqslant \cdots \geqslant \lambda_d$ that are constant for μ-almost every x and almost every sequence of random times τ. The largest, λ_1 is called the *top Lyapunov exponent.* When it is positive, chaos ensues.

In this chapter we give sufficient conditions for a random splitting to have a positive top Lyapunov exponent. Our proof involves adapting the framework of [9] to the random splitting setting. In particular, we are able to reduce much of the work in proving positivity of the top Lyapunov exponent for conservative systems to the verification of the Lie bracket condition introduced in Section 2.1. A key step in our argument is proving that if the Lie bracket condition holds at a point and the vector fields of the random splitting are real analytic, then a positive power of the

Markov transition kernel of the random splitting is strong Feller on a neighborhood of that point; see Proposition 5.10. We then show via results in [9] that if the Lie bracket condition and some basic integrability criteria hold, and if $d\lambda_1 = \lambda_\Sigma$ where d is the dimension of the state space and λ_Σ is the sum of Lyapunov exponents, then either: (*Alternative 1*) The dynamics are conformal with respect to some Riemannian structure on the state space or (*Alternative 2*) There exist proper subspaces of the tangent spaces to the state space that are invariant under the dynamics. See Theorem 5.2. In particular, if $\lambda_\Sigma = 0$ then either $\lambda_1 = 0$ or one of the two alternatives must hold. We then prove in Section 5.4 that Alternative 1 is ruled out by shearing and Alternative 2 is ruled out when the Lie bracket condition holds at a point in the tangent bundle of the state space.

5.1 Lyapunov exponents

Fix $h > 0$ and consider a random splitting on a d-dimensional $\mathcal{V}$-orbit $\mathcal{X}$ with transition kernel P_h and ergodic P_h-invariant measure μ. Under Integrability Condition 1 below, the multiplicative ergodic theorem guarantees the existence of d numbers $\lambda_1 \geqslant \cdots \geqslant \lambda_d$, called the *Lyapunov exponents* of $\{\Phi^m_{h\tau}\}$, such that for μ-almost every x in $\mathcal{X}$ and every η in $T_x\mathcal{X}$,

$$\lim_{m\to\infty} \tfrac{1}{m} \log\|D\Phi^m_{h\tau}(x)\eta\| = \lambda_k$$

for some k and almost every τ, where D is the derivative in x and $\|\cdot\|$ is the norm on tangent spaces of $\mathcal{X}$ induced by the Euclidean norm on $\mathbb{R}^D$. Moreover, the top Lyapunov exponent, λ_1, and the sum of the Lyapunov exponents, $\lambda_\Sigma := \lambda_1 + \cdots + \lambda_d$, satisfy

$$\lambda_1 = \lim_{m\to\infty} \tfrac{1}{m} \log\|D\Phi^m_{h\tau}(x)\| \quad \text{and} \quad \lambda_\Sigma = \lim_{m\to\infty} \tfrac{1}{m} \log\left|\det\left(D_x\Phi^m_{h\tau}(x)\right)\right|, \tag{5.2}$$

where now $\|\cdot\|$ is the operator norm. The top Lyapunov exponent captures the largest rate of separation between nearby trajectories, e.g. $\{\Phi^m_{h\tau}(x)\}$ and $\{\Phi^m_{h\tau}(y)\}$ for points

y infinitesimally close to x. A positive value corresponds to exponential growth and is a hallmark of chaos. The sum of the Lyapunov exponents captures the overall behavior of volumes under the dynamics with $\lambda_\Sigma > 0$, $\lambda_\Sigma = 0$, and $\lambda_\Sigma < 0$ indicative of expanding, conservative, and contracting dynamics, respectively. It is both possible and in some sense typical to have a system which conserves volumes ($\lambda_\Sigma = 0$) while having some expanding directions ($\lambda_1 > 0$).

Integrability Condition 1. *Integrability condition 1 is*

$$\mathbb{E}\int_{\mathcal{X}} \left(\log^+\|D_x\Phi_{h\tau}(x)\| + \log^+\|D_x\Phi_{h\tau}(x)^{-1}\|\right)\mu(dx) < \infty \tag{5.3}$$

where $\log^+(a) := \max\{\log a, 0\}$ *for* $a > 0$.

The following lemma says that if the derivatives of the V_j are bounded on $\mathcal{X}$, then Integrability Condition 1 always holds. In particular, the Lyapunov exponents exist and satisfy (5.2). Since the V_j are smooth on $\mathbb{R}^D$, this is true whenever $\mathcal{X}$ is bounded.

Lemma 5.1. *Suppose there exists a constant* $C < \infty$ *such that*

$$\sup_{1\leqslant j\leqslant n}\sup_{x\in\mathcal{X}}\|DV_j(x)\| \leqslant C. \tag{5.4}$$

Then the random splitting associated to $\mathcal{V} = \{V_j\}_{j=1}^n$ *satisfies* (5.3).

Proof. Fix t in $\mathbb{R}^n_+$. By the chain rule,

$$D_x\Phi(x,t) = D_x\varphi^{(n)}_{t_n}\big(x^{(n-1)}\big)\cdots D_x\varphi^{(2)}_{t_2}\big(x^{(1)}\big)D_x\varphi^{(1)}_{t_1}\big(x\big)$$

where $x^{(j)} := \varphi^{(j)}_{t_j}\circ\cdots\circ\varphi^{(1)}_{t_1}(x)$. Therefore

$$\|D_x\Phi(x,t)\| \leqslant \prod_{j=1}^n \big\|D_x\varphi^{(j)}_{t_j}\big(x^{(j-1)}\big)\big\|. \tag{5.5}$$

Fix j and set $\varphi = \varphi^{(j)}$. By Cauchy-Schwarz, for any x in $\mathcal{X}$ and unit vector η in $T_x\mathcal{X}$,

$$\partial_s\|D_x\varphi_s(x)\eta\|^2 \leqslant 2\big\|DV_k(\varphi_s(x))\big\|\|D_x\varphi_s(x)\eta\|^2 \leqslant 2C\|D_x\varphi_s(x)\eta\|^2.$$

So by Grönwall's inequality, $\|D_x\varphi_s(x)\eta\| \leqslant \exp(Cs)$. Since j, x, and η were arbitrary,

$$\log\|D_x\Phi(x,t)\| \leqslant \sum_{j=1}^{n} \log e^{Cht_j} = Ch\sum_{j=1}^{n} t_j,$$

which is polynomial in t and hence integrable against the exponential density. The same argument applies to $\log\|D_x\Phi(x,t)^{-1}\|$ since $D_x\Phi(x,t)^{-1} = D_x\Phi(x,-t)$. So (5.3) holds. □

5.2 Conditions for chaos

Assume for the rest of this chapter that all V_j in $\mathcal{V}$ are analytic and that μ is absolutely continuous[1], ergodic, and invariant with respect to P_h for every $h > 0$. Also assume Integrability Condition 1 holds so the Lyapunov exponents $\lambda_1 \geqslant \cdots \geqslant \lambda_d$ on $\mathcal{X}$ exist and are constant for μ-almost every x and almost every τ. Before stating the main result of this section, Theorem 5.2, we need two additional definitions. First, for every m define the *pushforward* of μ by $\Phi^m_{h\tau}$ to be the probability measure $\mu_m := (\Phi^m_{h\tau})_{\#}\mu$ on $\mathcal{X}$ given by

$$\mu_m(B) := (\Phi^m_{h\tau})_{\#}\mu(B) := \mu\left((\Phi^m_{h\tau})^{-1}(B)\right).$$

Second, for probability measures ν and μ on $\mathcal{X}$ define the *relative entropy of ν with respect to μ* (also often called *Kullback-Leibler divergence*) by

$$D_{KL}(\nu \,\|\, \mu) := \begin{cases} \int_{\mathcal{X}} \frac{d\nu}{d\mu}(x)\log\left(\frac{d\nu}{d\mu}(x)\right)\mu(dx) & \text{if } \nu \ll \mu \\ \infty & \text{otherwise,} \end{cases}$$

where $\nu \ll \mu$ means ν is absolutely continuous with respect to μ and $d\nu/d\mu$ is the Radon-Nikodym derivative. Note μ_m is random since it depends on τ, so we can consider the *average relative entropy*, $\mathbb{E}D_{KL}(\mu_m \,\|\, \mu)$. The finiteness of this quantity is important in what follows. Whenever $\mu_m = \mu$ almost surely, as is the case in both

[1] *Absolutely continuous* means absolutely continuous with respect to v on $\mathcal{X}$ unless otherwise specified.

the Lorenz and Euler models, one has that $\mathbb{E}D_{KL}(\mu_m \parallel \mu) = \mathbb{E}D_{KL}(\mu \parallel \mu) = 0$. Another condition guaranteeing $\mathbb{E}D_{KL}(\mu_m \parallel \mu) < \infty$ is that $\log(d\mu/d\mathrm{V})$ is in $L^1(\mu)$ since this and Integrability Condition 1 together imply $\mathbb{E}D_{KL}(\mu_m \parallel \mu) = -m\lambda_\Sigma$ [9, Theorem 4.2]. With this in hand, we have

Theorem 5.2. *Suppose $\mathbb{E}D_{KL}(\mu_m \parallel \mu) < \infty$ for all m and the Lie bracket condition holds at some x_* in the support of μ. If $d\lambda_1 = \lambda_\Sigma$, then there is an open subset U of $\mathcal{X}$ such that $\mu(U) = 1$ and either*

Alternative 1. There is a Riemannian structure[2] $\{g_x : x \in U\}$ on U and an $\alpha : U \to \mathbb{R}_+$ such that for all m, t in $\mathbb{R}_+^{mn}$, x in U, and η and ξ in $T_x\mathcal{X}$,

$$g_{\Phi_t^m(x)}(D_x\Phi_t^m(x)\eta, D_x\Phi_t^m(x)\xi) = \alpha(x)g_x(\eta, \xi). \tag{5.6}$$

That is, Φ_t^m is conformal with respect to $\{g_x : x \in U\}$.

Alternative 2. For all x in U there exist proper linear subspaces $E_x^1, \dots, E_x^p$ of $T_x\mathcal{X}$ such that

$$D_x\Phi_t^m(x)(E_x^i) = E_{\Phi_t^m(x)}^{\sigma(i)} \tag{5.7}$$

for all m, t in $\mathbb{R}_+^{mn}$, and every i, where σ is a permutation possibly depending on m, t, and x.

Theorem 5.2 involves adapting [9, Theorem 6.9] to our setting. To state a version of that result, recall P_h acts on the space $\mathcal{B}_b(\mathcal{X})$ of bounded, measurable functions on $\mathcal{X}$ via

$$P_hf(x) = \mathbb{E}\Big(f\big(\Phi_{h\tau}(x)\big)\Big) = \int_{\mathbb{R}_+^n} f\big(\Phi_{ht}(x)\big)\rho(t)dt, \tag{5.8}$$

where $\rho(t)$ is the probability density function of the n independent times $\tau = (\tau_1, \dots, \tau_n)$. P_h is *strong Feller* if it maps $\mathcal{B}_b(\mathcal{X})$ into $C_b(\mathcal{X})$, the space of bounded,

[2] A *Riemannian structure on* U is a family $\{g_x : x \in U\}$ of inner products $g_x : T_x\mathcal{X} \times T_x\mathcal{X} \to \mathbb{R}$ on $T_x\mathcal{X}$.

continuous functions on $\mathcal{X}$. Similarly, P_h is strong Feller on an open subset U of $\mathcal{X}$ if it maps $\mathcal{B}_b(U)$ into $\mathcal{C}_b(U)$.

Theorem 5.3. *Suppose $\mathbb{E}D_{KL}(\mu_m \parallel \mu) < \infty$ for all m and there is an open subset U_0 of $\mathcal{X}$ containing a point in the support of μ such that P_h is strong Feller on U_0 for every $h > 0$. If $d\lambda_1 = \lambda_\Sigma$, there exists an open set U in $\mathcal{X}$ such that $\mu(U) = 1$ and Alternative 1 or 2 holds.*

Theorem 5.3 is mostly a combination of results in [9], which culminate with Theorem 6.9 therein. It derives in large part from two relationships, one between Lyapunov exponents and relative entropy, and the other between the splitting $\{\Phi^m_{h\tau}\}$ and the *lifted splitting*

$$\widetilde{\Phi}^m_{h\tau}(x, \eta) := \left(\Phi^m_{h\tau}(x), D_x\Phi^m_{h\tau}(x)\eta\right) \tag{5.9}$$

on the projective bundle $P\mathcal{X}$ of $\mathcal{X}$ where, by a slight abuse of notation, we use η to denote both an element of the tangent space $T_x\mathcal{X}$ and its equivalence class[3] in $P_x\mathcal{X}$ whenever $\eta \neq 0$. We also denote the transition kernel of the lifted process by $\widetilde{P}_h$. The assumption $\mathbb{E}D_{KL}(\mu_m \parallel \mu) < \infty$ for all m allows for the comparison of λ_1 and λ_Σ that lies at the heart of Theorem 5.3. In particular, [9, Corollary 5.6] says that if Integrability Condition 1 and the finite average relative entropy condition hold, then $d\lambda_1 = \lambda_\Sigma$ implies there exists a ν in the space $\mathscr{P}_\mu(P\mathcal{X})$ of probability measures on $P\mathcal{X}$ with $\mathcal{X}$-marginal μ whose *regular conditional probability distributions*[4] $\{\nu_x : x \in \mathcal{X}\}$ satisfy

$$\mu\left\{x : \left(\widetilde{\Phi}^m_{h\tau}\right)_{\#}\nu_x = \nu_{\Phi^m_{h\tau}(x)} \text{ for every } m\right\} = 1 \text{ for almost every } \tau. \tag{5.10}$$

The slight difference between Theorem 5.3 and [9, Theorem 6.9] is that in our setting the strong Feller assumption implies any ν satisfying (5.10) has a version such that

[3] Recall the *projective space* $P_x\mathcal{X}$ at x is the space of all lines in the tangent space $T_x\mathcal{X}$.

[4] The ν_x are probability measures on $P\mathcal{X}$ which are well-defined for μ-almost every x and satisfy $\nu_x(P_x\mathcal{X}) = 1$.

$x \mapsto \nu_x$ is continuous[5] on U_0. This is in contrast to Baxendale's setting where the existence of a continuous version of the conditional distribution ν_x requires an additional assumption on the transition kernel of the lifted process for discrete time Markov chains. Specifically, we have the added advantage that h is a continuous variable. This allows us to adapt the statement and proof of [9, Proposition 6.3] despite not having a continuous time Markov semigroup.

Lemma 5.4. *Suppose ν satisfies* (5.10) *and the hypotheses of Theorem 5.3 hold with U_0 as stated there. Then there is a version of $\{\nu_x : x \in \mathcal{X}\}$ such that $x \mapsto \nu_x$ is continuous on U_0.*

Theorem 5.3 follows immediately from Lemma 5.4, proven below, and [9, Theorem 6.9], which we refer the reader to for the details. Therefore the only thing that remains in proving Theorem 5.2 is that if the Lie bracket condition holds at a point, then P_h is strong Feller on a neighborhood of that point for all positive h. This is the content of the next section.

Proof of Lemma 5.4. For any f in $\mathcal{B}_b(P\mathcal{X})$ and probability measure κ on $P\mathcal{X}$ define

$$\kappa\widetilde{P}_h(f) := \int_{P\mathcal{X}} \widetilde{P}_h f(\widetilde{x})\kappa(d\widetilde{x}) = \int_{P\mathcal{X}} \mathbb{E} f\big(\widetilde{\Phi}_{h\tau}(\widetilde{x})\big)\kappa(d\widetilde{x}) = \mathbb{E}\int_{P\mathcal{X}} f(\widetilde{x})\big(\widetilde{\Phi}_{h\tau}\big)_{\#}\kappa(d\widetilde{x}),$$

and $\kappa f : \mathcal{X} \to \mathbb{R}$ by

$$\kappa f(x) := \kappa_x(f) = \int_{P\mathcal{X}} f(\widetilde{x})\kappa_x(d\widetilde{x}).$$

If (5.10) holds, then for every f in $\mathcal{B}_b(P\mathcal{X})$ and μ-almost every x in $\mathcal{X}$,

$$\nu_x\widetilde{P}_h(f) = \mathbb{E}\int_{P\mathcal{X}} f(\widetilde{x})\big(\widetilde{\Phi}_{h\tau}\big)_{\#}\nu_x(d\widetilde{x}) = \mathbb{E}\int_{P\mathcal{X}} f(\widetilde{x})\nu_{\Phi_{h\tau}(x)}(d\widetilde{x}) = P_h\nu f(x). \tag{5.11}$$

[5] The topology on $\mathscr{P}_\mu(P\mathcal{X})$ is that of weak convergence induced by the compact-open topology on $\mathcal{C}_b(P\mathcal{X})$.

Note $P_h\nu f$ is continuous on U_0 since P_h is strong Feller on U_0 by assumption. However, (5.11) holds for μ-*almost every* x in U_0 and we need a version such that $x \mapsto \nu_x$ is continuous for *every* x in U_0. To address this, fix a version $\{\nu_x : x \in \mathcal{X}\}$, set

$$U_1 := \{x \in U_0 : (5.11) \text{ holds for every } f \in \mathcal{C}_b(P\mathcal{X})\},$$

and define a new version $\{\bar{\nu}_x : x \in \mathcal{X}\}$ by

$$\bar{\nu}_x := \begin{cases} \nu_x & \text{if } x \notin U_0 \backslash U_1, \\ \lim_{n\to\infty} \nu_{x_n} & \text{if } x \in U_0 \backslash U_1, \end{cases} \tag{5.12}$$

where $\{x_n\}$ is a sequence in U_1 converging to x and ν_{x_n} a corresponding weakly convergent (sub)sequence in $\mathscr{P}(P\mathcal{X})$. Thus far, our choice of ν_x depends on the sequence $\{x_n\}$. We now show (5.12) is independent of this choice. Suppose x is in $U_0\backslash U_1$ with $\{x_n\}$ any sequence as above. By the dominated convergence theorem,

$$\lim_{h\to 0} \bar{\nu}_x \widetilde{P}_h f = \lim_{h\to 0} \int_{P\mathcal{X}} \widetilde{P}_h f(\tilde{x}) \bar{\nu}_x(d\tilde{x}) = \int_{P\mathcal{X}} f(\tilde{x}) \bar{\nu}_x(d\tilde{x}) = \bar{\nu}_x(f).$$

Therefore,

$$\bar{\nu}_x(f) = \lim_{h\to 0} \bar{\nu}_x \widetilde{P}_h f = \lim_{h\to 0} \lim_{n\to\infty} \nu_{x_n} \widetilde{P}_h f = \lim_{h\to 0} \lim_{n\to\infty} P_h \nu f(x_n) = \lim_{h\to 0} P_h \nu f(x)$$

for all f in $\mathcal{C}_b(P\mathcal{X})$, where the second equality holds by (5.12), the third holds because the x_n are in U_1 and therefore satisfy (5.11), and the fourth holds because P_h is strong Feller on U_0 for all $h > 0$. Thus (5.12) is independent of $\{x_n\}$ which, together with $\mu(U_0\backslash U_1) = 0$, implies $\{\bar{\nu}_x : x \in \mathcal{X}\}$ is a well-defined version of $\{\nu_x : x \in \mathcal{X}\}$, i.e. the two agree almost surely. Furthermore, for f in $\mathcal{C}_b(P\mathcal{X})$ and x in U_0,

$$\lim_{x_n\to x} \bar{\nu}_{x_n}(f) = \lim_{x_n\to x} \lim_{h\to 0} P_h \nu f(x_n) = \lim_{h\to 0} P_h \nu f(x) = \bar{\nu}_x(f),$$

where we again used the strong Feller property. So $x \mapsto \bar{\nu}_x$ is continuous on U_0. □

5.3 Regularity

We begin this section by considering a more general setting than random splitting. Corollary 5.6 and Proposition 5.8 are stated at this heightened level of generality. We then apply these results torandom splitting to prove that if the Lie bracket condition holds at a point, then there is a positive integer m and a neighborhood U of that point such that the transition kernel P_h^m has a density and is strong Feller on U; see Proposition 5.10. Note only the strong Feller part of this result is needed in our study of Lyapunov exponents; the existence of transition densities is included because it is a direct consequence of the coarea formula and of possible independent interest. To avoid notational confusion, $\mathcal{X}$ will always denote a $\mathcal{V}$-orbit.

5.3.1 Transition densities

Let U be an open subset of a smooth d-dimensional manifold $\mathcal{Y}$ with volume form $\mathrm{v}_{\mathcal{Y}}$, and let Ω be a connected, open subset of $\mathbb{R}^p$ with $p \geqslant d$. Assume ρ is a probability measure on Ω that is absolutely continuous with respect to Lebesgue. Any continuous function $\Psi : U \times \Omega \to \mathcal{Y}$ induces an operator P mapping $\mathcal{B}_b(U)$ to itself via

$$Pf(u) := \mathbb{E}\big(f(\Psi(u,\omega))\big) := \int_\Omega (f \circ \Psi)(u,\omega)\rho(\omega)d\omega, \tag{5.13}$$

where, as before, ρ denotes both the measure and its density with respect to Lebesgue. P has a *density* with respect to $\mathrm{v}_{\mathcal{Y}}$ if there exists an integrable $p : U \times \mathcal{Y} \to [0, \infty)$ satisfying

$$Pf(u) = \int_{\mathcal{Y}} f(y)p(u,y)\,\mathrm{v}_{\mathcal{Y}}(dy).$$

That P has a density is a direct corollary of the coarea formula when Ψ is a submersion.

Lemma 5.5 (Coarea formula)**.** *Let $\mathcal{Y}$ and $\mathcal{Z}$ be smooth d and p-dimensional manifolds with volume forms $\mathrm{v}_{\mathcal{Y}}$ and $\mathrm{v}_{\mathcal{Z}}$, respectively. Suppose $F : \mathcal{Z} \to \mathcal{Y}$ is a C^1 submersion, i.e. $DF(z) : T_z\mathcal{Z} \to T_{F(z)}\mathcal{Y}$ is surjective for every z in $\mathcal{Z}$. Then for any $f : \mathcal{Z} \to \mathbb{R}$ measurable with respect to $\mathrm{v}_{\mathcal{Z}}$,*

$$\int_{\mathcal{Z}} f(z)\,\mathrm{v}_{\mathcal{Z}}(dz) = \int_{\mathcal{Y}} \left(\int_{F^{-1}(y)} \frac{f(z)}{\sqrt{\det DF(z)DF(z)^*}} \mathcal{H}^{p-d}(dz) \right) \mathrm{v}_{\mathcal{Y}}(dy) \tag{5.14}$$

as long as either side is finite. Here $\mathcal{H}^{p-d}(dz)$ is $p-d$-dimensional Hausdorff measure on $\mathcal{Z}$.

Proof. See [62, Corollary 2.2] or [27, Theorem 3.2.11]. □

Corollary 5.6. *Suppose Ψ is in $C^1(U \times \Omega, \mathcal{Y})$. If for every u in U,*

$$D_\omega \Psi(u, \omega) : T_\omega \Omega \to T_{\Psi(u,\omega)}\mathcal{Y} \tag{5.15}$$

is surjective for Lebesgue-almost every ω, then P has density p with respect to $\mathrm{v}_{\mathcal{Y}}$ given by

$$p(u, y) = \int_{\Psi_u^{-1}(y)} \frac{\rho(\omega)}{\sqrt{\det D_\omega \Psi(u, \omega) D_\omega \Psi(u, \omega)^*}} \mathcal{H}^{p-d}(d\omega) \tag{5.16}$$

where $\Psi_u^{-1}(y) := \{\omega : \Psi(\omega, u) = y\}$.

The assumption that $D_\omega \Psi(u, \omega)$ is surjective for almost every ω is equivalent to

$$M(u, \omega) := D_\omega \Psi(u, \omega) D_\omega \Psi(u, \omega)^* \tag{5.17}$$

being almost-surely invertible. Since $\mathcal{Y}$ is d-dimensional, M is a symmetric, positive-semidefinite $d \times d$ matrix. Note $p \geqslant d$ is a necessary condition for its invertibility, which was our reason for assuming this above. M is analogous to the Malliavin matrix in Malliavin calculus; for details see the related work section at the end of this chapter and references therein.

Remark 5.7. *The assumption in Corollary 5.6 that for fixed u the map $\Psi(u, \cdot) : \Omega \to \mathcal{Y}$ from noise space to state space is a submersion is a form of hypoellipticity. It allows some regularity of the noise distribution ρ to be transferred to the state space.*

Proof of Corollary 5.6. For f in $\mathcal{B}_b(U)$ we have

$$\begin{aligned} Pf(u) &= \int_\Omega f(\Psi(u, \omega))\rho(\omega) d\omega = \int_{\mathcal{Y}} \left(\int_{\Psi_u^{-1}(y)} \frac{f(\Psi(u, \omega))\rho(\omega)}{\sqrt{\det M(u, \omega)}} \mathcal{H}^{p-d}(d\omega) \right) \mathrm{v}_{\mathcal{Y}}(dy) \\ &= \int_{\mathcal{Y}} f(y) \left(\int_{\Psi_u^{-1}(y)} \frac{\rho(\omega)}{\sqrt{\det M(u, \omega)}} \mathcal{H}^{p-d}(d\omega) \right) \mathrm{v}_{\mathcal{Y}}(dy) = \int_{\mathcal{Y}} f(y) p(u, y)\, \mathrm{v}_{\mathcal{Y}}(dy). \end{aligned}$$

where the second equality is the coarea formula, the third holds because $f(\Psi(u,\omega)) = f(y)$ on $\Psi_u^{-1}(y)$, and the fourth is by definition of $p(u,y)$. One caveat in our application of the coarea formula is that $\det M(u,\omega)$ is nonzero only almost-surely. This is not an issue however since $\{y : \Psi(u,\omega) = y, \det M(u,\omega) = 0\}$ has measure 0 in $\mathcal{Y}$ and hence we can define $p(u,y) = 0$ for y in this set without changing the integral of f in $\mathcal{B}_b(\mathcal{U})$ against p. □

5.3.2 Strong Feller

Let $\mathcal{Y}$, Ω, ρ, Ψ, and $\mathcal{B}_b(\mathcal{Y})$ be as in Section 5.3.1 and recall $\mathcal{C}_b(\mathcal{Y})$ is the space of bounded, continuous functions $f : \mathcal{Y} \to \mathbb{R}$. The transition kernel P defined in (5.13) is *Feller* if it maps $\mathcal{C}_b(\mathcal{Y})$ into $\mathcal{C}_b(\mathcal{Y})$ and *strong Feller* if it maps $\mathcal{B}_b(\mathcal{Y})$ into $\mathcal{C}_b(\mathcal{Y})$. Being Feller is a statement about the dynamics being well-posed in that dependence on initial data is continuous. Strong Feller is a much stronger statement as it implies P has a regularizing effect.

Proposition 5.8. *Suppose U is an open subset of $\mathcal{Y}$ and Ψ is in $C^1(U \times \Omega, \mathcal{Y})$. If for every u in U, $D_\omega \Psi(u,\omega)$ is surjective for Lebesgue-almost every ω, then P is strong Feller on U.*

The next lemma is used to prove Proposition 5.8. Its proof is given at the end of this section.

Lemma 5.9. *Let Ψ be in $C^1(\mathcal{Y} \times \Omega, \mathcal{Y})$ and let K be a compact subset of Ω. For y in $\mathcal{Y}$ set*

$$A_K(y) := \{\omega \in K : \det M(y,\omega) = 0\}.$$

If $M(y,\cdot)$ is almost-surely invertible for every y in $\mathcal{Y}$, then for any y_* in $\mathcal{Y}$ and $\varepsilon > 0$ there exists an open set U in Ω and a $\delta > 0$ such that $\mu(U) < \varepsilon$ and $A_K(y)$ is contained in U for all y in $B_\delta(y_*)$. Moreover, there exists an open neighborhood W

of $K \cap U^c$ such that

$$\inf_{y \in B_\delta(y_*)} \Big\{ \det M(y,\omega) : \omega \in W \Big\} > 0. \tag{5.18}$$

Proof of Proposition 5.8. Fix f in $\mathcal{B}_b(\mathcal{Y})$. The case $f \equiv 0$ is immediate, so assume otherwise; in particular, $C := 6\|f\|_\infty > 0$. Fix y_* in $\mathcal{Y}$ and $\varepsilon > 0$. Since μ is a Borel measure on Ω it is tight. So there exists a compact set K in Ω such that $\mu(K^c) < \varepsilon/C$. By Lemma 5.9 there exists U open and $\delta_1 > 0$ such that $\mu(U) < \varepsilon/C$ and

$$\inf_{y \in B_{\delta_1}(y_*)} \big\{ \det M(y,\omega) : \omega \in W \big\} =: \alpha > 0 \tag{5.19}$$

for some open neighborhood W of $K \cap U^c$. Now since $\mathbb{1}_{K^c} + \mathbb{1}_{K\cap U} + \mathbb{1}_{K \cap U^c} = 1$,

$$\begin{aligned} Pf(y) - Pf(y_*) = \mathbb{E}\Big(\big[f\big(\Psi(y)\big) - f\big(\Psi(y_*)\big)\big]\,\mathbb{1}_{K^c}\Big) + \mathbb{E}\Big(\big[f\big(\Psi(y)\big) - f\big(\Psi(y_*)\big)\big]\,\mathbb{1}_{K\cap U}\Big) \\ + \mathbb{E}\Big(\big[f\big(\Psi(y)\big) - f\big(\Psi(y_*)\big)\big]\,\mathbb{1}_{K\cap U^c}\Big) \end{aligned} \tag{5.20}$$

for any y in $\mathcal{Y}$. By our choice of K,

$$\Big|\mathbb{E}\Big(\big[f\big(\Psi(y)\big) - f\big(\Psi(y_*)\big)\big]\,\mathbb{1}_{K^c}\Big)\Big| \leqslant 2\|f\|_\infty \mu(K^c) < \frac{\varepsilon}{3}. \tag{5.21}$$

And by our choice of U,

$$\Big|\mathbb{E}\Big(\big[f\big(\Psi(y)\big) - f\big(\Psi(y_*)\big)\big]\,\mathbb{1}_{K\cap U}\Big)\Big| \leqslant 2\|f\|_\infty \mu(U) < \frac{\varepsilon}{3}. \tag{5.22}$$

To handle the third expectation (the one with $\mathbb{1}_{K\cap U^c}$), set $L^1 := L^1(\mathrm{v}_{\mathcal{Y}})$ and choose a compactly supported, continuous function g on $\mathcal{Y}$ such that $\|g - f\|_{L^1} < \varepsilon/(18\sqrt{\alpha})$. This can always be done since compactly supported, continuous functions are dense in L^1 [29, Proposition 7.9]. By adding and subtracting $g \circ \Psi$ appropriately,

$$\begin{aligned} \mathbb{E}\Big(\big[f\big(\Psi(y)\big) - f\big(\Psi(y_*)\big)\big]\,\mathbb{1}_{K\cap U^c}\Big) = \mathbb{E}\Big(\big[f\big(\Psi(y)\big) - g\big(\Psi(y)\big)\big]\,\mathbb{1}_{K\cap U^c}\Big) \\ + \mathbb{E}\Big(\big[g\big(\Psi(y)\big) - g\big(\Psi(y_*)\big)\big]\,\mathbb{1}_{K\cap U^c}\Big) \\ + \mathbb{E}\Big(\big[g\big(\Psi(y_*)\big) - f\big(\Psi(y_*)\big)\big]\,\mathbb{1}_{K\cap U^c}\Big). \end{aligned} \tag{5.23}$$

Since $g \circ \Psi$ is continuous, there exists $\delta_2 > 0$ such that for every y in $B_{\delta_2}(y_*)$,

$$\big|g\big(\Psi(y)\big) - g\big(\Psi(y_*)\big)\big| < \frac{\varepsilon}{9}.$$

So for all y in $B_{\delta_2}(y_*)$,

$$\Big|\mathbb{E}\Big(\big[g\big(\Psi(y)\big) - g\big(\Psi(y_*)\big)\big]\,\mathbb{1}_{K\cap U^c}\Big)\Big| < \frac{\varepsilon}{9}. \tag{5.24}$$

For the first and third terms in (5.23), recall μ is absolutely continuous and therefore has an integrable density ρ. Setting $S(y,\hat{y}) := \{\omega \in W : \Psi(y,\omega) = \hat{y}\}$,

$$\begin{aligned}
\mathbb{E}\Big(\big[f\big(\Psi(y)\big) - g\big(\Psi(y)\big)\big]\,\mathbb{1}_{K\cap U^c}\Big) &= \int_\Omega \big[f\big(\Psi(y,\omega)\big) - g\big(\Psi(y,\omega)\big)\big]\,\mathbb{1}_{K\cap U^c}(\omega)\rho(\omega)d\omega \\
&= \int_W \big[f\big(\Psi(y,\omega)\big) - g\big(\Psi(y,\omega)\big)\big]\,\mathbb{1}_{K\cap U^c}(\omega)\rho(\omega)d\omega \\
&= \int_{\mathcal{Y}}\int_{S(y,\hat{y})} \frac{\big[f\big(\Psi(y,\omega)\big) - g\big(\Psi(y,\omega)\big)\big]\,\mathbb{1}_{K\cap U^c}(\omega)\rho(\omega)}{\sqrt{\det M(y,\omega)}} d\omega \, \mathrm{v}_{\mathcal{Y}}(d\hat{y}) \\
&= \int_{\mathcal{Y}} \big(f(\hat{y}) - g(\hat{y})\big)\Big(\int_{S(y,\hat{y})} \frac{\mathbb{1}_{K\cap U^c}(\omega)\rho(\omega)}{\sqrt{\det M(y,\omega)}} d\omega\Big)\, \mathrm{v}_{\mathcal{Y}}(d\hat{y}).
\end{aligned}$$

The second equality holds because W contains $K \cap U^c$ and the third is the coarea formula. By our choice of δ_1 in (5.19) we have that for any y in $B_{\delta_1}(y_*)$ and ω in W,

$$\frac{1}{\sqrt{\det M(y,\omega)}} \leqslant \frac{1}{\sqrt{\alpha}}.$$

So for all y in $B_{\delta_1}(y_*)$ our choice of g implies

$$\Big|\mathbb{E}\Big(\big[f\big(\Psi(y)\big) - g\big(\Psi(y)\big)\big]\,\mathbb{1}_{K\cap U^c}\Big)\Big| \leqslant \frac{1}{\sqrt{\alpha}}\int_{\mathcal{Y}} \big|f(y) - g(y)\big|dy < \frac{\varepsilon}{18}. \tag{5.25}$$

Set $\delta := \min\{\delta_1, \delta_2\}$. Applying the triangle inequality, (5.24), and (5.25) to (5.23),

$$\Big|\mathbb{E}\Big(\big[f\big(\Psi(y)\big) - f\big(\Psi(y_*)\big)\big]\,\mathbb{1}_{K\cap U^c}\Big)\Big| < \frac{\varepsilon}{18} + \frac{\varepsilon}{9} + \frac{\varepsilon}{18} = \frac{\varepsilon}{3} \tag{5.26}$$

for all y in $B_\delta(y_*)$. And applying the triangle inequality, (5.21), (5.22), and (5.26) to (5.20),

$$\big|Pf(y) - Pf(y_*)\big| < \varepsilon$$

for all y in $B_\delta(y_*)$. So Pf is continuous and therefore P is strong Feller. □

Proof of lemma 5.9. Fix y_* in $\mathcal{Y}$ and $\varepsilon > 0$. By assumption $A_K(y)$ has Lebesgue measure zero for all y. In particular, since μ is absolutely continuous with respect to Lebesgue, there exists a neighborhood U of $A_K(y_*)$ such that $\mu(U) < \varepsilon$. Suppose toward a contradiction there is a sequence $\{y_n\}$ converging to y_* such that $A_K(y_n)$ is not contained in U; that is, for each y_n there is an ω_n in $K \cap U^c$ satisfying $\det M(y_n, \omega_n) = 0$. Then since $K \cap U^c$ is compact there is a subsequence $\{\omega_{n_k}\}$ of $\{\omega_n\}$ which converges to some ω_* in $K \cap U^c$. Since $(y, \omega) \mapsto \det M(y, \omega)$ is continuous,

$$0 = \lim_{k\to\infty} \det M(y_{n_k}, \omega_{n_k}) = \det M(y_*, \omega_*).$$

But this implies ω_* is in $A_K(y_*) \cap (K \cap U^c) = \varnothing$, a contradiction.

Consider now the "moreover" part of the lemma. By the preceding argument,

$$\inf_{y\in B_\delta(y_*)} \Big\{ \det M(y, \omega) : \omega \in K \cap U^c \Big\} \geqslant 2\alpha$$

for some $\alpha > 0$. Set $F_y := \det M(y, \cdot)$ and let K' be the closure in Ω of

$$\bigcup_{y\in B_\delta(y_*)} F_y^{-1}(0, \alpha).$$

Straightforward verification shows $K \cap U^c$ and K' are closed and disjoint and are therefore separated by disjoint open sets. By continuity and construction, any such neighborhood W of $K \cap U^c$ satisfies

$$\inf_{y\in B_\delta(y_*)} \Big\{ \det M(y, \omega) : \omega \in W \Big\} \geqslant \alpha > 0. \qquad \square$$

5.3.3 Strong Feller and random splitting

We return now to a general random splitting associated to a family of analytic vector fields $\mathcal{V}$. In this setting the above results yield the following.

Proposition 5.10. *If the Lie bracket condition holds at a point x_* in a d-dimensional $\mathcal{V}$-orbit $\mathcal{X}$, then for some m and open neighborhood U of x_* the map $t \mapsto D_t\Phi^m(x, ht)$*

is a submersion for every x in U, $h > 0$, and almost every t in $\mathbb{R}_+^{mn}$. In particular, the transition kernel P_h^m is strong Feller on U for every $h > 0$ and has transition density $p_{m,h} : U \times \mathcal{X} \to [0, \infty)$ given by

$$p_{m,h}(x, y) = \int_{\{t:\Phi_{ht}^m(x)=y\}} \frac{\rho(t)}{\sqrt{\det M(x, ht)}} \mathcal{H}^{mn-d}(dt), \tag{5.27}$$

for almost every y in $\mathcal{X}$ and $p_{m,h}(x, y) = 0$ otherwise, where

$$M(x, ht) := D_t\Phi^m(x, ht)D_t\Phi^m(x, ht)^*.$$

The proof of Proposition 5.10 uses the following result from [58].

Lemma 5.11. *Let Ω be a connected, open subset of $\mathbb{R}^n$. If $f : \Omega \to \mathbb{R}$ is analytic and not identically 0, then $f^{-1}(0)$ has Lebesgue measure zero in Ω.*

Proof of Proposition 5.10. By Theorem 2.5 there exist m and t_* in $\mathbb{R}_+^{mn}$ such that $t \mapsto \Phi^m(x, t)$ is a submersion at t_*. Define $f : \mathcal{X} \times \mathbb{R}_+^{mn} \to \mathbb{R}$ by

$$f(x, t) := f_t(x) := \det D_t\Phi^m(x, t)D_t\Phi^m(x, t)^* = \det M(x, t).$$

Then $f(x_*, t_*) > 0$ and, since $f_{t_*} : \mathcal{X} \to \mathbb{R}$ is continuous, $U := f_{t_*}^{-1}((0, \infty))$ is an open neighborhood of x_* in $\mathcal{X}$. Now since the vector fields (and hence their flows) are analytic and analyticity is preserved under addition, multiplication, composition, and differentiation, the map $t \mapsto f(x, t)$ is analytic for every x in $\mathcal{X}$. Also for any $h > 0$ and x in U we have $f(x, h(t_*/h)) = f(x, t_*) > 0$ so $t \mapsto f(x, t)$ is not identically 0. And since $\mathbb{R}_+^{mn}$ is connected and open in $\mathbb{R}^{mn}$, Lemma 5.11 implies $M(x, h\tau)$ is almost-surely invertible. Hence $t \mapsto \Phi^m(x, ht)$ is a submersion for every x in U, $h > 0$, and almost every t in $\mathbb{R}_+^{mn}$. This proves the first part of the theorem. The expression for the transition density and the strong Feller property of P_h^m on U then follow immediately from Corollary 5.6 and Proposition 5.8, respectively. □

5.4 Ruling out alternatives

If the hypotheses of Theorem 5.2 are satisfied and $d\lambda_1 = \lambda_\Sigma$, then Alternative 1 or 2 holds on an open set U in $\mathcal{X}$ satisfying $\mu(U) = 1$. In this section we give sufficient conditions under which these alternatives do not hold so that, in particular, $d\lambda_1 \neq \lambda_\Sigma$ whenever the aforementioned hypotheses are true. The primary mechanism for ruling out Alternative 1 is shearing (Proposition 5.12), while Alternative 2 is ruled out when the Lie bracket condition holds at any point in TU (Proposition 5.13).

Proposition 5.12. *Suppose there are indices $i, j, k,$ and ℓ and a constant $C \neq 0$ such that*

$$V_i(x) = Cx_\ell(x_k e_j - x_j e_k), \tag{5.28}$$

where $\{e_j\}$ is the standard basis for $\mathbb{R}^D$. Then Alternative 1 cannot hold.

Proof. The solution of $\dot{x} = V_i(x)$ starting from $x(0)$ is

$$\begin{cases} x_j(t) = x_j(0)\cos(Cx_\ell t) + x_k(0)\sin(Cx_\ell t) \\ x_k(t) = -x_j(0)\sin(Cx_\ell t) + x_k(0)\cos(Cx_\ell t) \\ x_p(t) = x_p(0), \quad p \notin \{j,k\}. \end{cases}$$

Restricting attention to $x_j, x_k,$ and x_ℓ since these are the only coordinates that contribute nontrivially to the flow $\varphi := \varphi^{(i)}$ of V_i, the derivative of φ in the $j, k,$ and ℓ coordinates is

$$D\varphi_t(x) = \begin{pmatrix} \cos(Cx_\ell t) & \sin(Cx_\ell t) & -Ct\left(x_j\sin(Cx_\ell t) - x_k\cos(Cx_\ell t)\right) \\ -\sin(Cx_\ell t) & \cos(Cx_\ell t) & -Ct\left(x_j\cos(Cx_\ell t) + x_k\sin(Cx_\ell t)\right) \\ 0 & 0 & 1 \end{pmatrix}.$$

Evaluating at $t_m := 2\pi m/Cx_\ell$ for any x with $x_\ell \neq 0$ gives

$$D\varphi_{t_m}(x) = \begin{pmatrix} 1 & 0 & 2\pi m\frac{x_k}{x_\ell} \\ 0 & 1 & -2\pi m\frac{x_j}{x_\ell} \\ 0 & 0 & 1 \end{pmatrix} = I + mA \quad \text{where} \quad A = \begin{pmatrix} 0 & 0 & 2\pi\frac{x_k}{x_\ell} \\ 0 & 0 & -2\pi\frac{x_j}{x_\ell} \\ 0 & 0 & 0 \end{pmatrix}. \tag{5.29}$$

Suppose Alternative 1 holds, i.e. there exists a Riemannian structure $\{g_x : x \in U\}$ on an open set U in $\mathcal{X}$ such that (5.30) is valid for all m, t in $\mathbb{R}_+^{mn}$, x in U, and η, ξ in $T_x\mathcal{X}$. In particular,

$$g_{\varphi_t(x)}(D\varphi_t(x)\eta, D\varphi_t(x)\xi) = \alpha(x)g_x(\eta, \xi). \tag{5.30}$$

Consider a metric tensor on the tangent space $T_x\mathcal{X}$, which lifts to a metric tensor $\tilde{g}$ on the ambient space, $\mathbb{R}^D$. This metric tensor can be written as

$$\tilde{g}_x = \sum \lambda_j w_j^\top w_j$$

for some basis $\{w_j\}$ of $T_x\mathbb{R}^D$ and positive constants $\{\lambda_j\}$. Hence, choosing $\eta = \xi$ with $\|\eta\|_{g_x} = 1$, by (5.30) we must have that for any fixed x with $x_\ell \neq 0$,

$$\alpha = g_x(D\varphi_{t_m}\eta, D\varphi_{t_m}\eta) = \sum_j \lambda_j (w_j^\top D\varphi_{t_m}(x)\eta)^2,$$

where $\alpha := \alpha(x)$ is constant. By (5.29) the above is a quadratic function of m so if the coefficient of the largest exponent is nonzero it cannot be constant. This coefficient is

$$\sum_j \lambda_j (w_j^\top A\eta)^2$$

which is a sum of nonnegative terms. So it suffices to show at least one of them is nonzero. This however is the case since the w_j span $T_x\mathcal{X}$ and the vector contained in A is parallel to it, contradicting the assumption of conformal invariance. So Alternative 1 cannot hold. □

Proposition 5.13. *If the Lie bracket condition holds at $\widetilde{x}_*$ in TU, Alternative 2 cannot hold.*

Proof. Analyticity of the vector fields implies the lifted splitting

$$\widetilde{\Phi}_{h\tau}^m(x, \eta) := \big(\Phi_{h\tau}^m(x), D_x\Phi_{h\tau}^m(x)\eta\big), \tag{5.31}$$

which we now regard as a chain on $T\mathcal{X}$ rather than $P\mathcal{X}$, is also analytic. Therefore the Lie bracket condition at $\widetilde{x}_*$ together with an argument essentially identical to the proof of Proposition 5.8 gives the existence of an m such that the transition kernel of the lifted process $\widetilde{P}_h^m$ is strong Feller on a neighborhood $\widetilde{U}$ of $\widetilde{x}_*$ for every $h > 0$. Fix such an h and assume for simplicity $m = 1$, which comes without loss of generality since the Lyapunov exponents of $\{\Phi_{h\tau}^m\}_{m=0}^{\infty}$ are the same as those of $\{\Phi_{h\tau}^{mk}\}_{k=0}^{\infty}$ and both alternatives in Theorem 5.2 hold for all m. Also, by shrinking $\widetilde{U}$ if necessary, assume the projection $\pi(\widetilde{U})$ of $\widetilde{U}$ onto $\mathcal{X}$ is contained in U. If Alternative 2 holds, there exist for every x in U proper subspaces $E_x^1, \ldots, E_x^p$ of $T_x\mathcal{X}$ such that

$$D_x\Phi_t(x)(E_x^i) = E_{\Phi_t(x)}^{\sigma(i)} \tag{5.32}$$

for every t in $\mathbb{R}_+^n$ and i, where σ is a permutation of $\{1, \ldots, p\}$. In particular, setting

$$E_x := \bigcup_{i=1}^{p} E_x^i$$

for each x in U, the map $f : \widetilde{U} \to \mathbb{R}$ given by[6]

$$\widetilde{f}(x, \eta) := \mathbb{1}_{E_x}(\eta) := \begin{cases} 1 & \text{if } \eta \in E_x, \\ 0 & \text{otherwise,} \end{cases}$$

is in $\mathcal{B}_b(\widetilde{U})$ and, since the E_x^i are *proper* subspaces, discontinuous. But by (5.32),

$$\widetilde{P}_h f(x, \eta) = \mathbb{E}\big(\mathbb{1}_{E_{\Phi_{h\tau}^m(x)}}(D_x\Phi_{h\tau}(x)\eta)\big) = \mathbb{E}\left(\mathbb{1}_{D_x\Phi_{h\tau}(x)\eta}(D_x\Phi_{h\tau}(x)\eta)\right) = f(x, \eta),$$

contradicting that $\widetilde{P}_h$ is strong Feller on $\widetilde{U}$. So Alternative 2 cannot hold. □

5.5 Related work

The bulk of this chapter first appeared in [3]. The work of Baxendale [9] builds on ideas in [17, 31, 32, 35, 42] which were also consulted in developing this work. The regularity results and their proofs presented above are reminiscent of the probabilistic understanding, via Malliavin calculus, of Hörmander's classical results on hypoelliptic

[6] The assumption that $\pi(\widetilde{U})$ is contained in U guarantees f is well-defined.

differential operators which require Lie bracket conditions similar to those presented here [39]. In particular, the matrix M in (5.17) is analogous to the controllability Gramian matrix in control theory where ω is a control, and the Malliavin matrix in Malliavin calculus where randomness is Brownian motion. $\sqrt{\det M(y, \omega)}$ is an expression of the tangential Jacobian, so in this setting the coarea formula is just a generalization of the classical change of variable formula [52, 60]. Related results on the hypoellipticty of piecewise deterministic Markov processes, though leading to slightly different statements, can be found in [6, 12].

6

Conservative systems

In this chapter we construct random splittings and apply the above results to conservative Lorenz-96 and Galerkin approximations of 2d Euler on the torus. As we see shortly, the former conserves energy and the latter conserves both energy and enstrophy. Their respective splittings will be constructed so that the splitting vector fields also conserve these quantities.

6.1 Conservative Lorenz-96

Recall from Section 1.3 the *conservative Lorenz-96 equations* are

$$\dot{x} = V(x) := \sum_{j=1}^{n} (x_{j+1} - x_{j-2}) x_{j-1} e_j, \tag{6.1}$$

where $n \geqslant 4$ is fixed, $\{e_j\}_{j=1}^n$ is the standard basis in $\mathbb{R}^n$, and indices are periodized via the identities $x_{-1} := x_{n-1}$, $x_0 := x_n$, and $x_{n+1} := x_1$. We also saw V splits as

$$V(x) = \sum_{j=1}^{n} V_j(x) \quad \text{where} \quad V_j(x) := (x_{j+1} e_j - x_j e_{j+1}) x_{j-1}. \tag{6.2}$$

Similar to V_2 in the splitting of Lorenz-63 in Section 1.1, for each j the flow $\varphi^{(j)}$ of V_j is a rotation in the (x_j, x_{j+1})-plane with angular velocity x_{j-1}. Direct computation shows $\partial_t\|\varphi_t^{(j)}(x)\|^2 = 0$ so each V_j conserves, like V, the system's *energy*, which for Lorenz-96 is defined to be the square of the Euclidean norm, $\|x\|^2 := \sum_{j=1}^n x_j^2$. Throughout this section $\mathcal{V}$ denotes the family of splitting vector fields corresponding to (6.2) and $S^{n-1}(R) := \{x \in \mathbb{R}^n : \|x\| = R\}$ is the sphere of radius R centered at the origin in $\mathbb{R}^n$. By the preceding remarks every $\mathcal{V}$-orbit lies on $S^{n-1}(R)$ for some R. In particular, we have

Proposition 6.1. *All the finite time convergence results of Chapter 4 apply to the random splitting* (6.2) *of conservative Lorenz-96 starting from any initial condition.*

Proof. The splitting vector fields are smooth and Assumption 1 is satisfied since every $\mathcal{V}$-orbit lies on a sphere, so the conclusions of Theorems 4.1 and 4.4 both hold. □

6.1.1 Ergodicity

A complicating feature of the conservative Lorenz-96 equations is that they have fixed points. Specifically, a point x in $\mathbb{R}^n$ is a fixed point of (6.1) if and only if $\sum_{j=1}^n (x_j^2 + x_{j+1}^2)x_{j-1}^2 = 0$. For a 2-sphere embedded in $\mathbb{R}^3$ these are precisely the 6 points of intersection of the sphere with the standard coordinate axes. In higher dimensions, these fixed points lie on submanifolds that in general have dimension greater than 0 and in particular are no longer isolated. Nevertheless, nonfixed points cannot reach fixed points in finite time. In fact, the following result shows there is precisely one $\mathcal{V}$-orbit, called a *generic orbit*, on each sphere that contains all nonfixed points on that sphere. Furthermore, for every $h > 0$ the volume form on any such generic orbit is the unique P_h-invariant measure on that orbit, where P_h is the transition kernel of the random splitting corresponding to (6.2).

Proposition 6.2. *If x is a nonfixed point of conservative Lorenz-96, then*

$$\mathcal{X}_R(x) = \mathcal{X}_R := \left\{ y \in \mathbb{R}^n : \|y\| = R \text{ and } \sum_{k=1}^n (y_k^2 + y_{k+1}^2)y_{k-1}^2 \neq 0 \right\}, \tag{6.3}$$

where $R = \|x\|$. *Furthermore, for all* $h > 0$ *the volume form on* $\mathcal{X}$ *is the unique* P_h*-invariant measure on* $\mathcal{X}$ *where* P_h *is the transition kernel of the random splitting associated to* (6.2).

Proof. Fix $R > 0$ and let x be a nonfixed point with $\|x\| = R$. We first prove x can be mapped via the split dynamics to $x_* := (R/\sqrt{n}, \ldots, R/\sqrt{n})$. Since x is a nonfixed point, i.e. $\sum_{j=1}^{n}(x_j^2 + x_{j+1}^2)x_{j-1}^2 \neq 0$, there exists j such that $x_{j-1} \neq 0$ and x_j or x_{j+1} is nonzero. Now, since $\varphi^{(j)}$ is a rotation in the (x_j, x_{j+1})-plane with angular velocity x_{j-1}, there is a t_j such that both j and $j+1$ coordinates of $\varphi^{(j)}(x, t_j)$ are nonzero. By the same argument there is a t_{j+1} such that the j, $j+1$, and $j+2$ coordinates of $x^{(j+1)} = \varphi^{(j+1)}(\varphi^{(j)}(x, t_j), t_{j+1})$ are nonzero. Continuing this way, we see x can be made to have nonzero coordinates in a finite number of steps.

Now since $\|x\| = R$, there exists an index j such that $|x_j| \geqslant R/\sqrt{n}$. If $j = n$, rotate in the $(n-1, n)$-plane so that the nth coordinate of x becomes $R/\sqrt{n}$. If $j < n$, rotate in the $(j, j+1)$-plane so that the $j+1$ coordinate of x becomes $R/\sqrt{n}$, then rotate in the $(j+1, j+2)$-plane so that the $j+2$ coordinate of x becomes $R/\sqrt{n}$, and so on until the nth-coordinate of x becomes $R/\sqrt{n}$. Such rotations are always possible because all coordinates of x are nonzero by the preceding argument. Thus, whether $j = n$ or $j < n$ we can evolve x via the split dynamics so that its last coordinate, x_n, is $R/\sqrt{n}$. In particular, there now must exist an index $j < n$ such that $|x_j| \geqslant R/\sqrt{n}$. By the same procedure, and without disturbing the last coordinate, we can use rotations to make the $n-1$ coordinate of x equal $R/\sqrt{n}$. Iterating this process maps x to x_* in a finite number of steps. Since x was arbitrary it follows that every nonfixed point with norm R belongs to the same orbit, which is precisely the set $\mathcal{X}_R$ defined in (6.3).

Next we prove there is at most one P_h-invariant measure on $\mathcal{X} := \mathcal{X}_R$. First, since the split dynamics are all rotations the above procedure mapping any x in $\mathcal{X}$ to x_*

can be done using strictly positive times. Furthermore, by direct observation, the matrix of splitting vector fields

$$\begin{pmatrix} | & | & & | \\ V_1(x) & V_2(x) & \cdots & V_{n-1}(x) \\ | & | & & | \end{pmatrix} = \begin{pmatrix} x_2x_n & 0 & \dots & 0 \\ -x_1x_n & x_3x_1 & \dots & 0 \\ 0 & -x_2x_1 & \cdots & \vdots \\ \vdots & \vdots & \ddots & x_nx_{n-2} \\ 0 & 0 & & -x_{n-1}x_{n-2} \end{pmatrix}$$

has rank $n-1$ when all x_j are nonzero. In particular, since $\mathcal{X}$ is an open subset of $S^{n-1}(R)$ and therefore itself an $n-1$-dimensional manifold, the V_j span $T_{x_*}\mathcal{X}$. Hence $\mathrm{Lie}_{x_*}(\mathcal{V}) = T_{x_*}\mathcal{X}$. By Corollary 3.5, P_h has at most one invariant measure on $\mathcal{X}$.

We next show Lebesgue measure, Leb, in $\mathbb{R}^n$ is P_h-invariant. Let $S^{n-1}(R)$ denote the sphere of radius R in $\mathbb{R}^n$ and let $\mathrm{Leb}_t^{(j)} := (\varphi_t^{(j)})_\# \mathrm{Leb}$ be the pushforward of Leb by $\varphi_t^{(j)}$. Since the V_j in (6.2) are divergence free, the continuity equation[1], becomes

$$0 = \partial_t \mathrm{Leb}_t^{(j)} + \mathrm{div}\left(V_j \mathrm{Leb}_t^{(j)}\right) = \partial_t \mathrm{Leb}_t^{(j)} + \nabla \mathrm{Leb}_t^{(j)} \cdot V_j \,.$$

The latter is a transport equation with constant initial condition $\mathrm{Leb}_0^{(j)} \equiv 1$ and hence $\mathrm{Leb}_t^{(j)} = \mathrm{Leb}$ for all t. Because the trajectories of all V_j conserve the energy $\|x\|$, we fiber $\mathbb{R}^n$ using spherical coordinates $(r, \vartheta) \in \mathbb{R}_+ \times S^{n-1}(R)$. In these coordinates, we have that $V_j(r,\vartheta) = 0\,\partial_r + r v_j(\vartheta)\nabla_\vartheta$ and by a change of coordinates of the divergence operator the stationarity equation becomes

$$0 = \mathrm{div}\left(V_j(x)\,\mathrm{v}(x)\right) = u(r)w(\vartheta)\,\mathrm{div}_\vartheta(\mathrm{v}(r,\vartheta)v_j(\vartheta)), \tag{6.4}$$

where div_ϑ denotes the angular terms of the divergence in spherical coordinates, and $u(r), w(\vartheta)$ result from the change of variables. Hence, we can factor the solution $\mathrm{v}(r,\vartheta) = \bar{\mathrm{v}}(\vartheta|r)\cdot\mu_R(dr) = \bar{\mathrm{v}}(\vartheta)\cdot\mu_R(dr)$, where $\bar{\mathrm{v}}(\vartheta|r)$ is the conditional density of Lebesgue measure on a fiber. The measure $\bar{\mathrm{v}}$ solves $w(\vartheta)\,\mathrm{div}_\vartheta(\bar{\mathrm{v}}(\vartheta)v_j(\vartheta)) = 0$ and

[1] This equation should be interpreted as an equation on measures or, equivalently, as holding in the weak sense. In other words, the left and right side are equal when integrated against any compactly supported, smooth test function.

hence is invariant under the flows $\varphi_t^{(j)}$. By rotational symmetry of Leb, we must have that $\bar{\mathrm{v}}(\vartheta)$ is the volume form on $S^{n-1}(R)$. And since $\mathcal{X}$ is a full-measure open subset of $S^{n-1}(R)$, the volume form v on $\mathcal{X}$ is just the restriction of $\bar{\mathrm{v}}$ to $\mathcal{X}$. Thus v is also invariant under the flows and is therefore the unique P_h-invariant measure on $\mathcal{X}$. □

Corollary 6.3. *For all $h > 0$ the volume form on $S^{n-1}(R)$, which we also denote by v, is the unique ergodic P_h-invariant probability measure on $S^{n-1}(R)$ that is absolutely continuous with respect to v.*

Proof. $\mathcal{X}_R$ in (6.3) is the complement of a closed, measure zero subset of $S^{n-1}(R)$. Thus volume form on $\mathcal{X}_R$ agrees with the volume form, also denoted v, on $S^{n-1}(R)$. In particular, v is an ergodic invariant measure on $S^{n-1}(R)$ by Proposition 6.2. Since ergodic invariant measures are mutually singular, see e.g. [36], any other ergodic invariant measure on $S^{n-1}(R)$ must be singular with respect to v. □

6.1.2 Positive top Lyapunov exponent

Throughout this section we refer to orbits from (6.3) as *generic orbits*.

Theorem 6.4. *The top Lyapunov exponent of the conservative Lorenz-96 random splitting* (6.2) *on a generic orbit is positive for every $h > 0$.*

Proof. Fix $R > 0$ and set $\mathcal{X} := \mathcal{X}_R$. We know from Proposition 6.2 that the volume form on any generic orbit $\mathcal{X}$ of the conservative Lorenz-96 splitting is the unique P_h-invariant measure on $\mathcal{X}$ for every $h > 0$. Thus the Lyapunov exponents exist and are almost-surely constant on generic orbits and Integrability Condition 1 trivially holds. Furthermore $\lambda_\Sigma = 0$ (and hence $\lambda_1 \geqslant 0$) since the splitting vector fields conserve Euclidean norm. This establishes the hypotheses of Theorem 5.2 for random splittings on generic orbits of conservative Lorenz-96. And since $\lambda_\Sigma = 0$, Theorem 5.2 says that if $\lambda_1 = 0$ then Alternative 1 or 2 must hold. Alternative 1 is immediately ruled out by Proposition 5.12 upon noting that the splitting vector fields are

$$V_j(x) = x_{j-1}(x_{j+1}e_j - x_j e_{j+1}).$$

So to prove Theorem 6.4, it remains to show Alternative 2 does not hold.

To rule out Alternative 2, it suffices to find a point $\widetilde{x}_* \in T\mathcal{X}$ at which the family $\widetilde{\mathcal{V}} = \{\widetilde{V}_j\}$ satisfies the Lie bracket condition, $\dim(\mathrm{Lie}_{\widetilde{x}}(\widetilde{\mathcal{V}})) = 2n-2$. In the coordinates $(x_1, \eta_1, \ldots, x_n, \eta_n)$, the lifted vector fields of the Lorenz splitting are

$$\widetilde{V}_i(x,\eta) = (0,\ldots,0, x_{i-1}x_{i+1}, \eta_{i-1}x_{i+1} + \eta_{i+1}x_{i-1}, -x_{i-1}x_i, -\eta_{i-1}x_i - \eta_i x_{i-1}, 0, \ldots, 0), \tag{6.5}$$

where, in order from left to right, the nonzero entries correspond to the coordinates x_i, η_i, x_{i+1}, and η_{i+1}. Let $\widetilde{x} = (x,\eta)$ be any point of $T\mathcal{X}$ satisfying

$$x = (a,a,b,b,b,\ldots,b) \quad \text{and} \quad \eta = \begin{cases} (1,-1,1,-1,\ldots,1,-1), & \text{if } n \text{ even} \\ (1,-1,1,-1,\ldots,1,-1,0), & \text{if } n \text{ odd,} \end{cases} \tag{6.6}$$

with $a, b \neq 0$. Note η is perpendicular to x as elements of $\mathbb{R}^n$ and is therefore a well-defined element of $T_x\mathcal{X} = T_x\mathbb{S}^{n-1}(R)$. Consider first the case when n is even. Direct computation via (6.5) shows that for $i = 2, \ldots, n-1$ the vector fields $\widetilde{V}_i$ and $[\widetilde{V}_{i-1}, \widetilde{V}_i]$ evaluated at $\widetilde{x}$ form the $2n \times 2$ matrix

$$\begin{pmatrix} | & | \\ [\widetilde{V}_{i-1}, \widetilde{V}_i](x,\eta) & \widetilde{V}_i(x,\eta) \\ | & | \end{pmatrix} = \begin{pmatrix} \bigstar \\ A_i \\ \mathbf{0} \end{pmatrix},$$

where $\bigstar$ indicates irrelevant entries, $\mathbf{0}$ indicates the rest of the matrix is filled with zeros, and the 2×2 matrix A_i, which comprises the $2i+1$ and $2i+2$ rows of the matrix, is given by

$$A_2 = \begin{pmatrix} 0 & -a^2 \\ 4ab & 0 \end{pmatrix}, \quad A_3 = \begin{pmatrix} a(a^2-b^2) & -ab \\ -(a+b)^2 & b-a \end{pmatrix},$$

$$A_4 = \begin{pmatrix} 0 & -b^2 \\ 4ab & 0 \end{pmatrix}, \quad A_i = \begin{pmatrix} 0 & -b^2 \\ \pm 4b^2 & 0 \end{pmatrix},$$

with the last A_i holding for all $i > 4$. Define the $2n \times 2n-2$ matrix

$$\mathcal{A} := \begin{pmatrix} | & | & | & & | & | & | \\ \widetilde{V}_1 & [\widetilde{V}_1, \widetilde{V}_2] & \widetilde{V}_2 & \cdots & [\widetilde{V}_{n-2}, \widetilde{V}_{n-1}] & \widetilde{V}_{n-1} & \widetilde{V}_n \\ | & | & | & & | & | & | \end{pmatrix} = \left(\begin{array}{ccccccc} \multicolumn{7}{c}{B} \\ \hline 0 & \boxed{A_2} & \star & \star & \star & \star & \star \\ 0 & 0 & \boxed{A_3} & \star & \star & \star & \star \\ 0 & 0 & 0 & \boxed{A_4} & \star & \star & \star \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \boxed{A_{n-1}} & \star \end{array}\right),$$

where B is the $4 \times 2n-2$ matrix

$$B = \begin{pmatrix} ab & -ab^2 & 0 & 0 & 0 & \cdots & 0 & 0 \\ -a-b & -b^2 & 0 & 0 & 0 & \cdots & 0 & 4b^2 \\ -ab & ab^2 & ab & -ab^2 & 0 & \cdots & 0 & 0 \\ a-b & -b^2 & a+b & a^2 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

We claim $\mathcal{A}$ has rank $2n-2$ for certain choices of a and b. First, note A_2, A_4, and A_i, $i > 4$, have rank 2 whenever $a, b \neq 0$. Also A_3 has rank 2 whenever $a, b \neq 0$ and

$$a^3 + ab^2 + 2b^3 \neq 0. \tag{6.7}$$

Row reducing B gives the matrix

$$B' = \begin{pmatrix} ab & \star & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \star & 0 & 0 & 0 & \cdots & 0 & \star \\ 0 & 0 & \star & \star & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & -a^5b^5(2a^2+5ab+2b^2) & 0 & \cdots & 0 & 4a^4b^7(2b-a) \end{pmatrix}.$$

The $\star$ entries, though easily computed and simply expressed, are redacted to emphasize the relevant terms. Suppose that, in addition to $a, b \neq 0$, the relations

$$\begin{cases} 2a^2 + (n-2)b^2 = R^2 \\ 2a^2 + 5ab + 2b^2 = 0 \\ a^3 + ab^2 + 2b^3 \neq 0 \\ 2b - a \neq 0 \end{cases} \tag{6.8}$$

hold. Direct substitution verifies all the above are satisfied when

$$a = -\frac{R}{\sqrt{4n-6}} \quad \text{and} \quad b = \frac{\sqrt{2}R}{\sqrt{2n-3}}. \tag{6.9}$$

The first relation in (6.8) guarantees $x = (a, a, b, \dots, b)$ satisfies $\|x\| = R$ and therefore lies on $\mathcal{X}$, and the third guarantees A_3 has rank 2 by (6.7). The second and fourth guarantee B' has the form

$$B' = \begin{pmatrix} ab & \star & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \star & 0 & 0 & 0 & \cdots & 0 & \star \\ 0 & 0 & \star & \star & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & c \end{pmatrix}$$

for some nonzero constant c. Replacing B with B' in $\mathcal{A}$ and moving the fourth row of B' to the last row of the whole matrix gives a new matrix

$$\left(\begin{array}{c|ccccccc}
ab & & & & & & & \\
\cline{1-1}
0 & & & & \bigstar & & & \\
0 & & & & & & & \\
\cline{2-8}
0 & \boxed{A_2} & \star & \star & \star & \cdots & \star & \star \\
0 & 0 & \boxed{A_3} & \star & \star & \cdots & \star & \star \\
0 & 0 & 0 & \boxed{A_4} & \star & \cdots & \star & \star \\
0 & 0 & 0 & 0 & \boxed{A_5} & \cdots & \star & \star \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & \cdots & \boxed{A_{n-1}} & \star \\
0 & 0 & 0 & 0 & 0 & \cdots & 0 & c
\end{array}\right). \tag{6.10}$$

Since each A_i has rank 2 and $ab, c \neq 0$, this matrix – and hence $\mathcal{A}$ – has rank $2n - 2$.

When n is odd, everything is essentially the same. Only the matrix B is different, but its row reduced form is identical to the B' above, up to the irrelevant $\star$ terms. In particular, the matrix corresponding to $\mathcal{A}$ in the odd case becomes the matrix (6.10) via the exact same procedure detailed above when subjected to the same relations defined in (6.8). Thus the the Lie bracket condition holds for the lifted process at the point $\widetilde{x}$ defined as in (6.6) with a and b as in (6.9) for all $n \geqslant 4$. So by Proposition 5.13 Alternative 2 cannot hold. And by Theorem 5.2 the top Lyapunov exponent of the random splitting of conservative Lorenz-96 is positive. $\square$

6.2 Galerkin approximations of 2d Euler

Recall from Section 1.3 that the 2d Euler equations on the torus $\mathbb{T}$ are

$$\begin{cases} \partial_t u + (u \cdot \nabla)u = -\nabla p \\ \mathrm{div}(u) := \nabla \cdot u = 0 \end{cases} \tag{6.11}$$

where $u : \mathbb{T} \times \mathbb{R} \to \mathbb{R}^2$ is the fluid velocity, $p : \mathbb{T} \times \mathbb{R} \to \mathbb{R}$ the fluid pressure, and

$$(u \cdot \nabla)u = (u_1 \partial_1 u_1 + u_2 \partial_2 u_1, u_1 \partial_1 u_2 + u_2 \partial_2 u_2)\,.$$

To construct a random splitting of (6.11), we first write (6.11) in vorticity form and apply the Fourier transform. This yields an infinite system of ODEs which we truncate to systems of arbitrary finite size, referred to throughout as Galerkin approximations. Finally, we split these Galerkin approximations to obtain the desired random splitting.

6.2.1 Constructing the splitting

The vorticity formulation of (6.11) is obtained by taking the curl of velocity. Specifically, setting $q := \mathrm{curl}(u) := \partial_2 u_1 - \partial_1 u_2$, equation (6.11) becomes

$$\begin{cases} \partial_t q + (\mathcal{K}q \cdot \nabla)q = 0, \\ \mathrm{div}(q) = 0, \end{cases} \tag{6.12}$$

where $\mathcal{K} := \nabla^\perp(-\Delta)^{-1}$ with $\nabla^\perp := (\partial_2, -\partial_1)$. To express (6.12) in Fourier space, set $\mathbb{Z}^2_\infty := \mathbb{Z}^2 \backslash \{(0,0)\}$ and let $\{e_j\}_{j \in \mathbb{Z}^2_\infty}$ be the orthonormal basis of $L^2(\mathbb{T}, \mathbb{R})$ given by $e_j(x) := (2\pi)^{-1} \exp(ix \cdot j)$. Then $q(x,t) = \sum_{j \in \mathbb{Z}^2_\infty} q_j(t) e_j(x)$ where

$$q_j(t) := \langle q, e_j \rangle_{L^2} = \int_{\mathbb{T}} q(x,t) \overline{e}_j(x) dx$$

is the jth Fourier mode of q. Here $\langle \cdot, \cdot \rangle_{L^2}$ is the standard inner product on $L^2(\mathbb{T}, \mathbb{R})$ with $\overline{e}_j$ denoting the complex conjugate of e_j. The jth Fourier mode of $(\mathcal{K}q \cdot \nabla)q$ is

$$\langle (\mathcal{K}q \cdot \nabla)q, e_j \rangle_{L^2} = \sum_{k+\ell=j} C_{k\ell} q_k q_\ell$$

where

$$C_{k\ell} := \frac{\langle k, \ell^{\perp} \rangle}{4\pi} \left(\frac{1}{|k|^2} - \frac{1}{|\ell|^2} \right) \tag{6.13}$$

with $\langle \cdot, \cdot \rangle$ the standard inner product in $\mathbb{R}^2$, $\ell^{\perp} := (\ell_2, -\ell_1)$, and $|\ell|^2 := \ell_1^2 + \ell_2^2$. Therefore

$$\sum_j \dot{q}_j e_j = \partial_t q = -(\mathcal{K}q \cdot \nabla) q = -\sum_j \left(\sum_{k+\ell=j} C_{k\ell} q_k q_\ell \right) e_j$$

and hence $\dot{q}_j = -\sum_{k+\ell=j} C_{k\ell} q_k q_\ell$. Moreover, since q is real-valued,

$$\sum_j q_j e_j = q = \overline{q} = \sum_j \overline{q}_j e_{-j}$$

which gives $q_j = \overline{q}_{-j}$. In particular,

$$\dot{q}_j = \dot{\overline{q}}_{-j} = -\sum_{j+k+\ell=0} C_{k\ell} \overline{q}_k \overline{q}_\ell \, .$$

Writing each Fourier mode $q_j = a_j + ib_j$ in terms of real and imaginary parts gives

$$\begin{aligned} \dot{a}_j + i\dot{b}_j = \dot{q}_j &= -\sum_{j+k+\ell=0} C_{k\ell}(a_k - ib_k)(a_\ell - ib_\ell) \\ &= \sum_{j+k+\ell=0} C_{k\ell}(b_k b_\ell - a_k a_\ell) + i \sum_{j+k+\ell=0} C_{k\ell}(a_k b_\ell + a_\ell b_k) \, . \end{aligned}$$

Thus the Fourier modes of solutions to the Euler equation in vorticity form satisfy

$$\begin{cases} \dot{a}_j = \displaystyle\sum_{j+k+\ell=0} C_{k\ell}(b_k b_\ell - a_k a_\ell) \\ \dot{b}_j = \displaystyle\sum_{j+k+\ell=0} C_{k\ell}(a_k b_\ell + a_\ell b_k) \end{cases} \tag{6.14}$$

for all $j \in \mathbb{Z}^2_\infty$. While (6.14) could be studied as is, notice the constraint $q_{-j} = \overline{q}_j$ implies $a_{-j} = a_j$ and $b_{-j} = -b_j$, which introduces redundancy in (6.14). Therefore we restrict to

$$\mathbb{Z}^2_+ := \{ j \in \mathbb{Z}^2 : j_2 > 0 \} \cup \{ j \in \mathbb{Z}^2 : j_2 = 0 \text{ and } j_1 > 0 \} \, .$$

Specifically, by straightforward computation together with the identities $a_{-j} = a_j$, $b_{-j} = -b_j$, and $C_{k\ell} = C_{-k,-\ell} = -C_{-k,\ell} = -C_{k,-\ell}$, the system (6.14) can be re-expressed as

$$\begin{cases} \dot{a}_j = \sum_{j+k-\ell=0} C_{k\ell}(a_k a_\ell + b_k b_\ell) + \sum_{j-k-\ell=0} C_{k\ell}(b_k b_\ell - a_k a_\ell) \\ \dot{b}_j = \sum_{j+k-\ell=0} C_{k\ell}(a_k b_\ell - b_k a_\ell) - \sum_{j-k-\ell=0} C_{k\ell}(a_k b_\ell + b_k a_\ell) \end{cases} \tag{6.15}$$

for all $j \in \mathbb{Z}^2_+$ with each sum running over all pairs $k, \ell \in \mathbb{Z}^2_+$ satisfying the specified identity. To split (6.15) note that for any $j, k, \ell \in \mathbb{Z}^2_+$ satisfying $j + k - \ell = 0$ (and hence $\ell - j - k = 0$) we can isolate from the above sums exactly 6 equations involving only these indices:

$$\begin{aligned} &\dot{a}_j = C_{k\ell}(a_k a_\ell + b_k b_\ell)\,, && \dot{a}_k = C_{j\ell}(a_j a_\ell + b_j b_\ell)\,, && \dot{a}_\ell = C_{jk}(b_j b_k - a_j a_k)\,, \\ &\dot{b}_j = C_{k\ell}(a_k b_\ell - b_k a_\ell)\,, && \dot{b}_k = C_{j\ell}(a_j b_\ell - b_j a_\ell)\,, && \dot{b}_\ell = -C_{jk}(a_j b_k + b_j a_k)\,. \end{aligned} \tag{6.16}$$

For reasons to be made clear shortly, we recombine (6.16) into 4 groups of 3 equations:

$$\begin{cases} \dot{a}_j = C_{k\ell} a_k a_\ell \\ \dot{a}_k = C_{j\ell} a_j a_\ell \\ \dot{a}_\ell = -C_{jk} a_j a_k \end{cases} \begin{cases} \dot{a}_j = C_{k\ell} b_k b_\ell \\ \dot{b}_k = C_{j\ell} a_j b_\ell \\ \dot{b}_\ell = -C_{jk} a_j b_k \end{cases} \begin{cases} \dot{b}_j = C_{k\ell} a_k b_\ell \\ \dot{a}_k = C_{j\ell} b_j b_\ell \\ \dot{b}_\ell = -C_{jk} b_j a_k \end{cases} \begin{cases} \dot{b}_j = -C_{k\ell} b_k a_\ell \\ \dot{b}_k = -C_{j\ell} b_j a_\ell \\ \dot{a}_\ell = C_{jk} b_j b_k \end{cases} . \tag{6.17}$$

Let $V_{a_j a_k a_\ell}$, $V_{a_j b_k b_\ell}$, $V_{b_j a_k b_\ell}$, and $V_{b_j b_k a_\ell}$ be the vector fields associated to the equations of (6.17) from left to right. For example, $V_{a_j a_k a_\ell}$ is the vector field on $\mathbb{R}^\infty$ mapping the a_j coordinate to $-C_{k\ell} a_k a_\ell$, the a_k coordinate to $-C_{j\ell} a_j a_\ell$, the a_ℓ coordinate to $-C_{jk} a_j a_k$, and all other coordinates to 0. These are the *splitting vector fields*. Our sought-after splitting is

$$V = \sum_{j+k-\ell=0} V_{a_j a_k a_\ell} + V_{a_j b_k b_\ell} + V_{b_j a_k b_\ell} + V_{b_j b_k a_\ell}, \tag{6.18}$$

where V is the vector field associated to (6.15). As noted earlier, our focus will be on finite truncations of the infinite-dimensional system (6.15). Thus we fix an integer $N \geqslant 2$ and define the *Nth Galerkin approximation* of (6.15) to be (6.15) with indices restricted to the set

$$\mathbb{Z}_N^2 := \left\{ j \in \mathbb{Z}_+^2 : \max\{|j_1|, |j_2|\} \leqslant N \right\}.$$

The splitting (6.18) remains valid in this finite-dimensional setting, bearing in mind that now all indices lie in $\mathbb{Z}_N^2$. By a slight abuse of notation, we denote the finite-dimensional counterpart of V by V and similarly for the splitting vector fields. Thus our family of splitting vector fields is

$$\mathcal{V} = \left\{ V_{a_j a_k a_\ell}, V_{a_j b_k b_\ell}, V_{b_j a_k b_\ell}, V_{b_j b_k a_\ell} : j, k, \ell \in \mathbb{Z}_N^2 \text{ and } j + k - \ell = 0 \right\}. \tag{6.19}$$

Since $\mathbb{Z}_N^2$ has cardinality $2N(N+1)$ and each index $j \in \mathbb{Z}_N^2$ has an associated a_j and b_j coordinate, these are all vector fields on $\mathbb{R}^n$, where throughout this section we set $n := 4N(N+1)$. We also abuse notation by conflating elements j in $\mathbb{Z}_N^2$ with elements j in $\{1, \ldots, n/2\}$, which can be formalized via any bijection between the two sets. Moreover, we denote elements of $\mathbb{R}^n$ by $q = (a_j, b_j)_{j=1}^{n/2}$. This reflects that the a_j and b_j coordinates of q in $\mathbb{R}^n$ are in one-to-one correspondence with the real and imaginary parts of the jth mode of q.

Remark 6.5. *There are many possible splittings of a given equation. For the Euler equations, we made the particular choice we have so that both energy and enstrophy are conserved but the dynamics of each splitting are still relatively easily understood. We could have further decomposed the three-dimensional dynamics in the above splitting into a number of two-dimensional dynamics, similar in spirit to the decomposition into rotations used in Lorenz-96. However, that would have necessitated only conserving either the energy or the enstrophy.*

6.2.2 *Conservation and convergence*

The conservative Lorenz-96 dynamics discussed in Section 6.1 conserves Euclidean norm (energy in that case) and therefore remains on whichever sphere it starts on. So too do the flows of each of the splitting vector fields (6.2). We now show a similar thing is true for Galerkin approximations of 2d Euler. Define the *energy* and *enstrophy* of $q = (a_j, b_j)_{j=1}^{n/2}$ by

$$E(q) := \sum_{j \in \mathbb{Z}_N^2} \frac{a_j^2 + b_j^2}{|j|^2} \qquad \text{and} \qquad \mathcal{E}(q) := \sum_{j \in \mathbb{Z}_N^2} a_j^2 + b_j^2, \tag{6.20}$$

respectively (note the aforementioned conflation of j in $\mathbb{Z}_N^2$ and $j \in \{1, \ldots, n/2\}$ in the summations). Straightforward computation shows that for all $j, k, \ell \in \mathbb{Z}_N^2$ with $j + k - \ell = 0$,

$$C_{k\ell} + C_{j\ell} - C_{jk} = \frac{C_{k\ell}}{|j|^2} + \frac{C_{j\ell}}{|k|^2} - \frac{C_{jk}}{|\ell|^2} = 0,$$

which in turn implies that under the dynamics (6.15),

$$\partial_t E(q) = \partial_t \mathcal{E}(q) = 0$$

for all $q \in \mathbb{R}^n$. That is, both energy and enstrophy are conserved by the true dynamics and

$$\mathcal{Q}_0(E, \mathcal{E}) := \big\{ q \in \mathbb{R}^n : E(q) = E,\ \mathcal{E}(q) = \mathcal{E} \big\}. \tag{6.21}$$

is invariant under (6.15). This is a well-established property of the 2d Euler equations. Moreover, if we flow by $V_{a_j a_k a_\ell}$ starting from q for any $j, k, \ell \in \mathbb{Z}_N^2$ with $j + k - \ell = 0$, then

$$\tfrac{1}{2}\partial_t E(q) = \frac{a_j \dot{a}_j}{|j|^2} + \frac{a_k \dot{a}_k}{|k|^2} + \frac{a_\ell \dot{a}_\ell}{|\ell|^2} = \left(\frac{C_{k\ell}}{|j|^2} + \frac{C_{j\ell}}{|k|^2} - \frac{C_{jk}}{|\ell|^2} \right) a_j a_k a_\ell = 0\,,$$

and similarly $\partial_t \mathcal{E}(q) = 0$. The same computation shows energy and enstrophy are conserved by *all* of the splitting vector fields in $\mathcal{V}$, which provides the motivation for recombining (6.16) as (6.17) in the first place. In particular, we have

Proposition 6.6. *All the finite time convergence results of Chapter 4 apply to the random splitting* (6.18) *of every Galerkin approximation of 2d Euler starting from any initial condition.*

Proof. The splitting vector fields are smooth and Assumption 1 is satisfied since every $\mathcal{V}$-orbit lies on a sphere, so the conclusions of Theorems 4.1 and 4.4 both hold. □

6.2.3 Ergodicity

Fix energy and enstrophy values E and $\mathcal{E}$ and set $\mathcal{Q}_0 := \mathcal{Q}_0(E, \mathcal{E})$.$\mathcal{Q}_0$ is an $n-2$-dimensional submanifold of $\mathbb{R}^n$ where, recall, $n := 4N(N+1)$; denote its volume form by λ. As with conservative Lorenz-96, the Nth Galerkin approximation of 2d Euler has points q in $\mathcal{Q}_0$ whose $\mathcal{V}$-orbits are not dense in $\mathcal{Q}_0$. For example, any q with exactly one nonzero coordinate is a fixed point of (6.15) and of all the equations (6.17). In this subsection we characterize these points and prove there is exactly one $\mathcal{V}$-orbit $\mathcal{Q}$ on $\mathcal{Q}_0$ such that $\lambda(\mathcal{Q}) = 1$. By a slight abuse of notation we denote the restriction of λ to $\mathcal{Q}$ by λ as well. We then show there exists a unique P_h-invariant measure on $\mathcal{Q}$ – and hence on $\mathcal{Q}_0$ – that is absolutely continuous with respect to λ on $\mathcal{Q}_0$.

To make the above statements precise, we begin by enumerating the coordinates of $q \in \mathbb{R}^n$ by extending the indices $j \in \mathbb{Z}_N^2$ with an element $\chi \in \{+,-\}$ which denotes the real $(+)$ or imaginary $(-)$ part of the corresponding mode. Then, for $\mathbf{j} = (j,\chi) \in \mathbb{Z}_N^2 \times \{+,-\}$, we define the *type* of such coordinates via the function $\mathrm{T}(\mathbf{j}) = \chi$ so that $q_{\mathbf{j}}$ is identified with a_j if $\mathrm{T}(\mathbf{j}) = +$ and with b_j if $\mathrm{T}(\mathbf{j}) = -$. For $q \in \mathbb{R}^n$, denote by

$$\mathcal{A}(q) := \left\{\mathbf{j} \in \mathbb{Z}_N^2 \times \{+,-\} \ : \ q_{\mathbf{j}} \neq 0\right\} \tag{6.22}$$

the set of "active" coordinates. To streamline our analysis, we define the following operation to expand the set $\mathcal{A}$:

$$\mathcal{A} \oplus \boldsymbol{\ell} := \begin{cases} \mathcal{A} \cup \{\boldsymbol{\ell}\} & \text{if } \ell \in \{j \pm k\} \cap \mathbb{Z}_N^2 \text{ for } \mathbf{j}, \mathbf{k} \in \mathcal{A}, C_{jk} \neq 0, \mathrm{T}(\mathbf{j}) \cdot \mathrm{T}(\mathbf{k}) = \mathrm{T}(\boldsymbol{\ell}) \,, \\ \mathcal{A} & \text{else} \,, \end{cases} \tag{6.23}$$

where $\mathrm{T}(\mathbf{j})\cdot\mathrm{T}(\mathbf{k})$ is $+$ if $\mathrm{T}(\mathbf{j}) = \mathrm{T}(\mathbf{k})$ and $-$ if $\mathrm{T}(\mathbf{j}) \neq \mathrm{T}(\mathbf{k})$. This operation corresponds to extending the nonzero coordinates of q from $\boldsymbol{j}, \boldsymbol{k}$ to $\boldsymbol{\ell}$ by letting a triple $\iota = \boldsymbol{jk\ell}$ interact.

We assume that the initial condition is sufficiently nondegenerate, as stated in the following assumption similar to the one made in [37, Theorem 2.1].

Definition 6.7 (Nondegenerate point)**.** *A point q in $\mathcal{Q}_0$ is* nondegenerate *if there exists $M \in \mathbb{N}$, $j^* \in \mathbb{Z}_N^2$ with $|j^*|^2 > 1$, and an ordered set of indices $(\boldsymbol{\ell}_i)_{i=1}^M$ in $\mathbb{Z}_N^2 \times \{+,-\}$ such that*

$$\big\{(1,0,+),(0,1,+),(j^*,-)\big\} \subseteq \big((\mathcal{A}(q) \oplus \boldsymbol{\ell}_1) \oplus \boldsymbol{\ell}_2\big) \cdots \oplus \boldsymbol{\ell}_M. \tag{6.24}$$

Definition 6.8 (Generic point)**.** *A point in $\mathbb{R}^n$ is* generic *if all of its coordinates are nonzero.*

Remark 6.9. *Every point with all coordinates nonzero is a nonfixed point of conservative Lorenz-96; similarly, every generic point in $\mathcal{Q}_0$ is nondegenerate. However, comparing* (6.24) *with* (6.3)*, we see the conditions defining nondegenerate points in $\mathcal{Q}_0$ are more complicated than the easily characterized nonfixed points of conservative Lorenz-96. The difference is that, unlike spheres in conservative Lorenz-96, there are proper subspaces of $\mathcal{Q}_0$ which are invariant for our splitting of the Euler dynamics but are not fixed points. One such subspace is the collection of purely real points; another is the purely imaginary points.*

The following analogs of Proposition 6.2 and Corollary 6.3 are the main results of this subsection.

Proposition 6.10. *Every nondegenerate point in $\mathcal{Q}_0$ belongs to the same $\mathcal{V}$-orbit, $\mathcal{Q}$, and for all $h > 0$ there exists a unique P_h-invariant probability measure on $\mathcal{Q}$. Furthermore, this unique invariant measure is absolutely continuous with respect to the volume form on $\mathcal{Q}$.*

Proof. By Proposition 6.13 there is a q^* in $\mathcal{Q}_0$ such that every nondegenerate point in $\mathcal{Q}_0$ belongs to the $\mathcal{V}$-orbit $\mathcal{Q} := \mathcal{Q}(q^*)$, and for every q in $\mathcal{Q}$ there is an $m \in \mathbb{N}$ and a $t \in \mathbb{R}_+^{mn}$ satisfying $\Phi^m(q,t) = q^*$. By Lemma 6.19 the splitting vector fields span the tangent space of $\mathcal{Q}$ at generic points; in particular, the Lie bracket condition holds at every generic point. Thus, since the vector fields in $\mathcal{V}$ are analytic, Corollary 3.6 implies P_h has at most one invariant probability measure on $\mathcal{Q}$, which is necessarily the one identified by Lemma 6.18. □

Corollary 6.11. *For all $h > 0$ the measure from Proposition 6.10 is the unique P_h-invariant ergodic probability measure on $\mathcal{Q}_0$ that is absolutely continuous with respect to the volume form on $\mathcal{Q}_0$.*

Proof. Let λ denote volume form on $\mathcal{Q}_0$. Since $\mathcal{Q}$ contains all generic points in $\mathcal{Q}_0$, it is an open subset of $\mathcal{Q}_0$ satisfying $\lambda(\mathcal{Q}) = 1$. In particular, the unique invariant measure on $\mathcal{Q}$ from Proposition 6.10 is an ergodic invariant measure on $\mathcal{Q}_0$. Since ergodic invariant measures are mutually singular, see e.g. [36], any other ergodic invariant measure on $\mathcal{Q}_0$ must be singular with respect to λ. □

Remark 6.12. *Continuing in the spirit of Remark 6.5, we observe* (6.17) *splits q_j into its real and imaginary parts. We could have chosen another basis of $\mathbb{C}$ and even randomized over this choice for each evolution of an interacting triple (j,k,ℓ). More explicitly, if we define $e(\vartheta) = cos(\vartheta) + i\sin(\vartheta)$ then $e(\vartheta)$ and $e(\vartheta + \frac{\pi}{2})$ form an orthonormal basis of $\mathbb{C}$ for any ϑ. Then we can drive a system analogous to* (6.17) *by setting $q_\ell = a_\ell^\vartheta e(\vartheta) + b_\ell^\vartheta e(\vartheta + \frac{\pi}{2})$. As the form is similar to* (6.17)*, the results of* *the paper extend to this system. In particular, by randomizing the choice of ϑ for each such triple (j, k, ℓ), we can relax the characterization of nondegenerate points in Definition 6.7 by destroying some of the invariant structures discussed in Remark 6.9 which obstruct controllability starting from some initial conditions.*

Controllability

We now prove controllability of the dynamics (6.17). By conservation of energy and enstrophy, the $\mathcal{V}$-orbit of an initial condition $q^{(0)}$ in $\mathcal{Q}_0$ is contained in $\mathcal{Q}_0$. Recalling the definition of extended indices in Section 6.2.3, we define the set of *interacting coordinate triples*

$$\mathcal{I} := \Big\{(\mathbf{j}, \mathbf{k}, \boldsymbol{\ell}) \in (\mathbb{Z}_N^2 \times \{+,-\})^3 \ : \ j + k = \ell,\ (C_{jk}, C_{j\ell}, C_{k\ell}) \neq (0,0,0), \\ \mathrm{T}(\mathbf{j}) \cdot \mathrm{T}(\mathbf{k}) = \mathrm{T}(\boldsymbol{\ell})\Big\}. \tag{6.25}$$

Then, for any such triple of interacting indices $\iota \in \boldsymbol{\ell}$ we denote by $\varphi_t^\iota \ : \ \mathcal{Q}_0 \to \mathcal{Q}_0$ the flow of the ODEs (6.17) evolving the corresponding coordinates. The dynamics we consider is then obtained by cycling through the set $\mathcal{I}$ in a fixed or random order. For any $\iota \in \mathcal{I}$ we denote by $\Phi_t^\iota \ : \ \mathcal{Q}_0 \to \mathcal{Q}_0$ the flow of (6.17) after one such full cycle where the flow times are chosen as

$$\tau^\xi = \begin{cases} t & \text{if } \xi = \iota\,, \\ 0 & \text{else}\,, \end{cases} \tag{6.26}$$

so that for any $q \in \mathcal{Q}_0$, $\Phi_t^\iota(q) = \varphi_t^\iota(q)$.

Let $q^* = (a_j^*, b_j^*)_{j=1}^{n/2}$ be the point in $\mathcal{Q}_0$ defined as follows:

$$q^*_{(1,0)} = q^*_{(0,1)} = (a^*, 0)\,, \qquad q^*_{(N,N)} = (0, b^*)\,, \tag{6.27}$$

for $a^*, b^* \geqslant 0$ and $q_j^* = (0,0)$ for all other $j \in \mathbb{Z}_N^2$. We show below that for any nondegenerate initial condition $q^{(0)} \in \mathcal{Q}_0$ the system can be driven to this unique point q^*.

Proposition 6.13. *For any nondegenerate point* $q^{(0)} = (a_j^{(0)}, b_j^{(0)})_{j=1}^{n/2}$ *in* $\mathcal{Q}_0$ *there exists* M *and a joint sequence of transition times and coordinate triples* $\{(\iota(m), \tau(m))\}_{m=1}^{M}$ *such that*

$$\Phi_{\tau(M)}^{\iota(M)} \circ \cdots \circ \Phi_{\tau(1)}^{\iota(1)}(q^{(0)}) = q^*. \tag{6.28}$$

Thus every nondegenerate point belongs to the same orbit, $\mathcal{Q} := \mathcal{Q}(q^*)$*. Furthermore, for every* q *in* $\mathcal{Q}$ *there is an* $m \in \mathbb{N}$ *and a* $t \in \mathbb{R}_+^{mn}$ *such that* $\Phi^m(q, t) = q^*$*.*

Recall that the only property of the exponential distribution used in this proof is the fact that it has a density around 0, allowing to choose the flow of some of the split vector fields to be the identity as, e.g., in (6.26). This comment also applies to the proof of Proposition 6.2 in the previous section. We further note that, since the trajectories of each of the $\varphi^{\iota(m)}$ in the above theorem are periodic (see Lemma 6.15 and Lemma 6.16), each of these transformations can be inverted by choosing complementary transition times to $\tau(m)$. Inverting the order of the transformations yields the converse statement:

Corollary 6.14. *For any nondegenerate point* $q^{(0)} = (a_j^{(0)}, b_j^{(0)})_{j=1}^{n/2}$ *in* $\mathcal{Q}_0$ *there exists* M *and a joint sequence of transition times and coordinate triples* $\{(\tilde{\iota}(m), \tilde{\tau}_0(m))\}_{m=1}^{M}$ *such that*

$$\Phi_{\tilde{\tau}(M)}^{\tilde{\iota}(M)} \circ \cdots \circ \Phi_{\tilde{\tau}(1)}^{\tilde{\iota}(1)}(q^*) = q^{(0)}\,. \tag{6.29}$$

While the Corollary 6.14 will not be used in the remainder of the paper, it offers an alternative to Theorem 2.7 in proving that, when applying Corollary 3.5, it is sufficient to verify that Lie bracket condition holds at *any* point in $\mathcal{Q}$, not necessarily at q^*.

Proof of Proposition 6.13. We prove the first statement by first evolving the initial condition $q^{(0)}$ into a sufficiently nondegenerate state $q^{(1)}$, and then by sequentially shrinking the set of active components of the coordinate vector q to the ones listed in

(6.27). We realize this program by following, in order, the sequence of steps described below, represented schematically in Figure 6.1:

0. If it is not the case at initialization, Lemma B.1 shows that we can "prepare" our state by evolving $q^{(0)}$ into $q^{(1)}$ such that
$$a^{(1)}_{(1,0)}, b^{(1)}_{(1,0)}, a^{(1)}_{(0,1)}, b^{(1)}_{(0,1)}, a^{(1)}_{(1,1)}, b^{(1)}_{(1,1)} \neq 0\,, \tag{6.30}$$
as represented in Figure 6.1a.

1. As shown in Lemma B.2, we can then transform $q^{(1)}$ into $q^{(2)}$ with the property
$$q^{(2)}_j = (0,0) \qquad \text{for all } j \in \mathbb{Z}^2_N \backslash \{(0,1),(1,0),(1,1),(N,N),(-N,N)\}\,, \tag{6.31}$$
as represented in Figure 6.1b, and
$$a^{(2)}_{(1,0)}, b^{(2)}_{(1,0)}, a^{(2)}_{(0,1)}, b^{(2)}_{(0,1)}, a^{(2)}_{(1,1)}, b^{(2)}_{(1,1)} \neq 0\,. \tag{6.32}$$

2. Lemma B.3 shows that we can then "transfer" the amplitude from modes $a_{(-N,N)}$, $b_{(-N,N)}$, $a_{(N,N)}$ to mode $b_{(N,N)}$ i.e. we can reach a state $q^{(3)}$ that satisfies
$$q^{(3)}_j = (0,0) \qquad \text{for all } j \in \mathbb{Z}^2_N \backslash \{(0,1),(1,0),(1,1),(N,N)\}\,, \tag{6.33}$$
$$q^{(3)}_{(N,N)} = (0, b^{(3)}_{(N,N)}) \quad \text{with } b^{(3)}_{(N,N)} \geqslant 0\,. \tag{6.34}$$
This state is represented in Figure 6.1c.

3. Finally, Lemma B.5 shows that we can "transfer" the amplitude from modes $a_{(1,1)}$, $b_{(1,1)}$, $b_{(0,1)}$ and $b_{(1,0)}$ to modes $a_{(0,1)}, a_{(1,0)}, b_{(N,N)}$ so that, after the transfer, $a_{(0,1)} = a_{(1,0)}$ and $a_{(0,1)}, a_{(1,0)}, b_{(N,N)} > 0$ i.e. we reach the unique state q^* from (6.27) (represented in Figure 6.1d).

This proves the first part of Proposition 6.13, which immediately implies nondegenerate points in $\mathcal{Q}_0$ belong to $\mathcal{Q} = \mathcal{Q}(q^*)$. Let q be *any* point in $\mathcal{Q}$. By definition there exist m and t in $\mathbb{R}^{mn}$ such that
$$\Phi^m(q,t) = \varphi^{(n)}_{t_{mn}} \circ \cdots \varphi^{(1)}_{t_1}(q) = q^*.$$

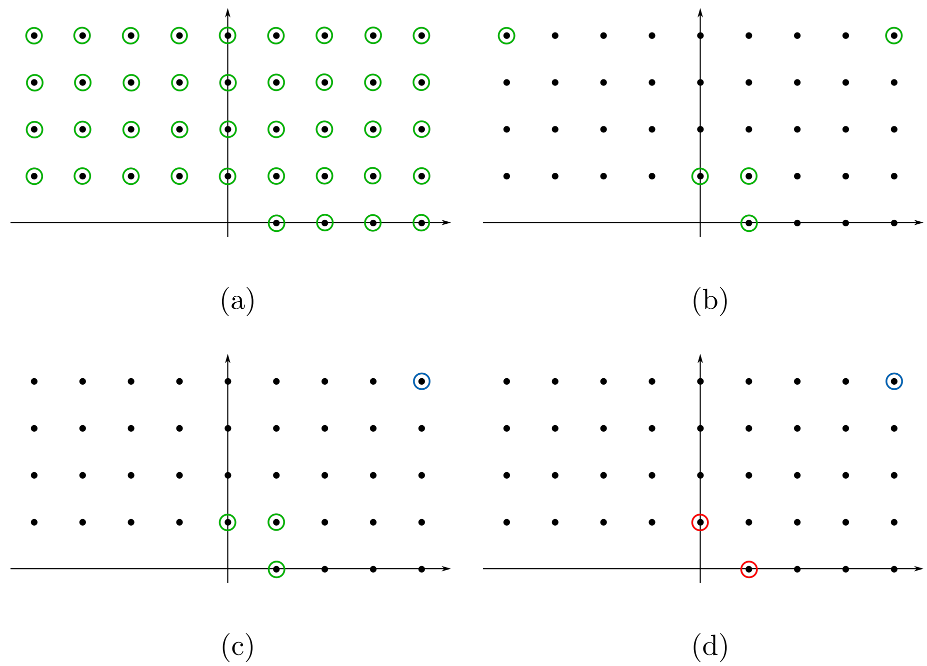


FIGURE 6.1: *Controlling Euler.* Representation of the state of the network in a generic initial state (a), after step 1 of the procedure in the proof of Proposition 6.13 (b), and after step 2 (c) and after step 3 (d) of the same procedure. In the above pictures, each point corresponds to a mode, i.e. an element of $\mathbb{Z}_N^2$ while the color of each circle represents the real/complex value of the corresponding mode: zero (white, no circle), purely imaginary (red), purely real (blue) or having both nonvanishing real and imaginary parts (green).

Note that the times t_i may be negative; however, by Lemma 6.15 each $\varphi^{(i)}$ is periodic. Thus for every $t_i \leqslant 0$ there exists a $t_i' > 0$ such that $\varphi^{(i)}_{t_i}(q') = \varphi^{(i)}_{t_i'}(q')$ for all q' in $\mathcal{Q}$. Let t' be t with all $t_i \leqslant 0$ replaced by t_i'. Then t' is in $\mathbb{R}_+^{mn}$ and $\Phi^m(q,t') = \Phi^m(q,t) = q^*$. □

Defining similarly to (6.23) the operation of removing a coordinate from the set $\mathcal{A}$

$$\mathcal{A} \ominus \boldsymbol{\ell} = \begin{cases} \mathcal{A}\backslash\{\boldsymbol{\ell}\} & \text{if } \ell \in \{j+k, j-k\} \cap \mathbb{Z}_N^2 \text{ for } \mathbf{j},\mathbf{k} \in \mathcal{A}, C_{jk} \neq 0, \mathrm{T}(\mathbf{j}) \cdot \mathrm{T}(\mathbf{k}) = \mathrm{T}(\boldsymbol{\ell}), \\ \mathcal{A} & \text{else}, \end{cases} \tag{6.35}$$

we now proceed to *construct* (sequences of) times τ and interacting triples ι such that the transformations $\Phi^{(\iota)}_\tau$ of q implement the operations $\oplus, \ominus$ from (6.23), (6.35) through the flow of (6.17), i.e. such that $\mathcal{A}(q)\oplus\boldsymbol{\ell} = \mathcal{A}(\Phi^\iota_\tau(q))$ or $\mathcal{A}(q)\ominus\boldsymbol{\ell} = \mathcal{A}(\Phi^\iota_\tau(q))$ respectively. To do so we separate the possible interactions between the modes in two types:

$$
\begin{aligned}
&a) \quad \iota = \boldsymbol{j}\boldsymbol{k}\boldsymbol{\ell} \in \mathcal{I} \;:\; |j| \neq |k| \neq |\ell|\,, \\
&b) \quad \iota = \boldsymbol{j}\boldsymbol{k}\boldsymbol{\ell} \in \mathcal{I} \;:\; |j| = |k| \neq |\ell|\,.
\end{aligned}
\tag{6.36}
$$

Note that these two types of interactions are exhaustive, since if $|j| = |k| = |\ell|$, $C_{j\ell} = C_{jk} = C_{k\ell} = 0$.

The following preparatory lemmas describe the properties of these two types of interactions that we will leverage throughout our proof. The first one shows that for interactions of type a), ordering the indices so that $|j| < |k| < |\ell|$, it is always possible to activate all modes $\boldsymbol{j}, \boldsymbol{k}, \boldsymbol{\ell}$ or to distribute the amplitude of the k-mode to the j and ℓ-modes reaching, in finite time, a state with $q_{\boldsymbol{k}} = 0$. As we show in the proof below, while such a point with $q_{\boldsymbol{k}} = 0$ always exists on the orbits of (6.17), this point is reachable in finite time for $\iota = \boldsymbol{j}\boldsymbol{k}\boldsymbol{\ell} \in \mathcal{I}$ with $|j| < |k| < |\ell|$ only if

$$
E_\iota(q) \neq |k|^2 \mathcal{E}_\iota(q)\,, \tag{6.37}
$$

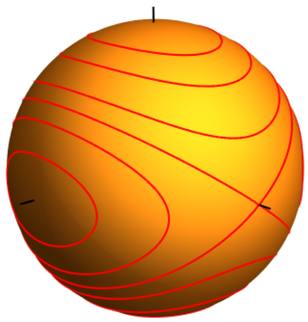
where $E_\iota(q)$ and $\mathcal{E}_\iota(q)$ denote the energy and enstrophy of the coordinates in $\iota \in \mathcal{I}$:

$$
E_\iota(q) := \sum_{\boldsymbol{\ell}\in\iota} |q_{\boldsymbol{\ell}}|^2\,, \qquad \mathcal{E}_\iota(q) := \sum_{\boldsymbol{\ell}\in\iota} \frac{|q_{\boldsymbol{\ell}}|^2}{|\ell|^2}\,. \tag{6.38}
$$

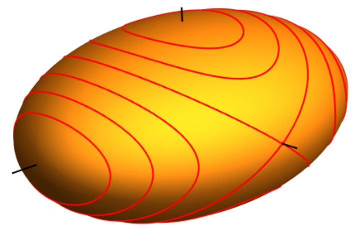
In the following lemma and throughout the section, we abuse notation slightly by defining $\operatorname{sign}(x) = +1$ for $x \in [0, \infty)$ and -1 otherwise.

Lemma 6.15. *Fix $\iota = \boldsymbol{j}\boldsymbol{k}\boldsymbol{\ell} \in \mathcal{I}$ with $|j| < |k| < |\ell|$. Let q be a nondegenerate point in $\mathcal{Q}_0$ satisfying* (6.37) *and let $q_{\boldsymbol{l}} = 0$ for at most an index $\boldsymbol{l} \in \{\boldsymbol{j}, \boldsymbol{k}, \boldsymbol{\ell}\}$. Then the orbit of V_ι is periodic and there exist $\tau^\iota_-, \tau^\iota_+ \geqslant 0$ such that*

(a) $\varphi^\iota_{\tau^\iota_-}(q) = q'$ *with* $q'_{\boldsymbol{k}} = 0$, $\operatorname{sign}(q_{\boldsymbol{j}}) = \operatorname{sign}(q'_{\boldsymbol{j}})$ *and* $\operatorname{sign}(q_{\boldsymbol{\ell}}) = \operatorname{sign}(q'_{\boldsymbol{\ell}})$,



(a)



(b)

FIGURE 6.2: *Energy and enstrophy.* Orbits $\mathcal{Q}_\iota$ of (6.17) (in red) corresponding in (A) to various values of the energy $\mathcal{E}_\iota(q)$ on the sphere of constant enstrophy $E_\iota(q)$ and in (B) to various values of the enstrophy $E_\iota(q)$ on the ellipsoid of constant energy $\mathcal{E}_\iota(q)$. The axes are, sequentially, $q_{\boldsymbol{k}}, q_{\boldsymbol{j}}, q_{\boldsymbol{\ell}}$. The orbit with a degenerate point at the pole of the sphere or ellipsoid corresponds to values of $E_\iota, \mathcal{E}_\iota$ violating (6.37).

(b) $\varphi^\iota_{\tau^\iota_+}(q) = q''$ *with* $q''_{\boldsymbol{j}}, q''_{\boldsymbol{k}}, q''_{\boldsymbol{\ell}} \neq 0$, $\operatorname{sign}(q_{\boldsymbol{j}}) = \operatorname{sign}(q''_{\boldsymbol{j}})$ *and* $\operatorname{sign}(q_{\boldsymbol{\ell}}) = \operatorname{sign}(q''_{\boldsymbol{\ell}})$.

Furthermore, if $|j|^2\mathcal{E}_\iota(q) < E_\iota(q) < |k|^2\mathcal{E}_\iota(q)$*, there exists* $\tau^\iota_= \geqslant 0$ *such that*

(c) $\varphi^\iota_{\tau^\iota_=}(q) = q'''$ *with* $q'''_{\boldsymbol{\ell}} = 0$, $\operatorname{sign}(q_{\boldsymbol{j}}) = \operatorname{sign}(q'''_{\boldsymbol{j}})$ *and* $\operatorname{sign}(q_{\boldsymbol{k}}) = \operatorname{sign}(q'''_{\boldsymbol{k}})$.

Proof. We consider the intersection between the sphere and the ellipse corresponding to the enstrophy and the energy in the coordinates $\iota = \boldsymbol{jk\ell} \in \mathcal{I}$ of interest, resulting in the set

$$\mathcal{Q}_\iota := \left\{ (q'_{\boldsymbol{j}}, q'_{\boldsymbol{k}}, q'_{\boldsymbol{\ell}}) \in \mathbb{R}^3 \ : \ |q'_{\boldsymbol{j}}|^2 + |q'_{\boldsymbol{k}}|^2 + |q'_{\boldsymbol{\ell}}|^2 = E_\iota(q), \frac{|q'_{\boldsymbol{j}}|^2}{|j|^2} + \frac{|q'_{\boldsymbol{k}}|^2}{|k|^2} + \frac{|q'_{\boldsymbol{\ell}}|^2}{|\ell|^2} = \mathcal{E}_\iota(q) \right\}. \tag{6.39}$$

This set is represented in Figure 6.2. We observe that this set has exactly 2 disjoint simply connected components when $|j|^2\mathcal{E}_\iota(q) < E_\iota(q) < |k|^2\mathcal{E}_\iota(q)$ and $|k|^2\mathcal{E}_\iota(q) < E_\iota(q) < |\ell|^2\mathcal{E}_\iota(q)$. These components are diffeomorphic to S^1. By continuity the dynamics are limited to one such component of $\mathcal{Q}_\iota$. Furthermore, $|\dot{q}|^2$ is uniformly bounded away from 0 on each such component: the fixed points of (6.17) must have at least two coordinates vanishing, which cannot be realized on the curves of interest. Therefore the dynamics on these sets are periodic.

We start by proving part (b) of the lemma. If $q_{\boldsymbol{j}}, q_{\boldsymbol{k}}, q_{\boldsymbol{\ell}} \neq 0$ the result follows by choosing $\tau_+^\iota = 0$. Else, if $q_{\boldsymbol{l}} = 0$ for $\boldsymbol{l} \in \iota$ the result follows immediately choosing τ_+^ι small enough by combining the continuity of the flow Φ_t^ι and the fact that $\dot{q}_{\boldsymbol{l}} = C_{\boldsymbol{l}'\boldsymbol{l}''} q_{\boldsymbol{l}'} q_{\boldsymbol{l}''} \neq 0$ for $\{\boldsymbol{l}', \boldsymbol{l}''\} = \iota \backslash \{\boldsymbol{l}\}$.

To prove part (a) we consider the cases where $|j|^2 \mathcal{E}_\iota(q) < E_\iota(q) < |k|^2 \mathcal{E}_\iota(q)$ and $|k|^2 \mathcal{E}_\iota(q) < E_\iota(q) < |\ell|^2 \mathcal{E}_\iota(q)$ separately. In the first case, we see that there is no point $q \in \mathcal{Q}_\iota$ with $q_{\boldsymbol{j}} = 0$: if that were the case we would have

$$E_\iota(q) = q_{\boldsymbol{k}}^2 + q_{\boldsymbol{\ell}}^2 = |k|^2 \left(\frac{q_{\boldsymbol{k}}^2}{|k|^2} + \frac{q_{\boldsymbol{\ell}}^2}{|k|^2} \right) > |k|^2 \mathcal{E}_\iota(q) \,, \tag{6.40}$$

contradicting our assumption. Consequently the points $(p_{\boldsymbol{j}}, 0, p_{\boldsymbol{\ell}}), (p_{\boldsymbol{j}}, 0, -p_{\boldsymbol{\ell}})$ with $p_{\boldsymbol{\ell}} > 0$, $\mathrm{sign}(p_{\boldsymbol{j}}) = \mathrm{sign}(q_{\boldsymbol{j}})$ and

$$p_{\boldsymbol{j}}^2 + p_{\boldsymbol{\ell}}^2 = E_\iota(q) \,, \qquad \frac{p_{\boldsymbol{j}}^2}{|j|^2} + \frac{p_{\boldsymbol{\ell}}^2}{|\ell|^2} = \mathcal{E}_\iota(q) \,, \tag{6.41}$$

belong to the same connected component as q and by the lower bound on the velocity on this connected component both these points are reachable in finite time from q. This also proves part (c) by continuity of the dynamics. The second case where $|k|^2 \mathcal{E}_\iota(q) < E_\iota(q) < |\ell|^2 \mathcal{E}_\iota(q)$ can be handled analogously: in this case we have $\mathcal{Q}_\iota \cap \{q_{\boldsymbol{\ell}} = 0\} = \varnothing$ and we can reach $(p_{\boldsymbol{j}}, 0, p_{\boldsymbol{\ell}}), (-p_{\boldsymbol{j}}, 0, p_{\boldsymbol{\ell}})$ with $p_{\boldsymbol{j}} > 0$, $\mathrm{sign}(p_{\boldsymbol{\ell}}) = \mathrm{sign}(q_{\boldsymbol{\ell}})$ in finite time. □

The following lemma considers interactions of type b) in (6.36). Recalling the definition $j^\perp := (j_2, -j_1)$ we show that interactions with $|j| = |k| \neq |\ell|$ leave component $\boldsymbol{\ell}$ fixed and move $\boldsymbol{j}, \boldsymbol{k}$ in a circle at constant angular speed.

Lemma 6.16. *Fix an unordered interacting triple $\iota = \boldsymbol{jk\ell}$ with $|k| = |j|$ and $q_{\boldsymbol{\ell}} \neq 0$. For all ϑ in $[0, 2\pi)$ there exists $t \geqslant 0$ such that $\varphi_t^\iota(q) = q'$ with $(q'_{\boldsymbol{j}}, q'_{\boldsymbol{k}}) =$*

$\sqrt{q_{\boldsymbol{j}}^2 + q_{\boldsymbol{k}}^2}(\cos(\vartheta), \sin(\vartheta))$ *and* $q'_{\boldsymbol{\ell}} = q_{\boldsymbol{\ell}}$.

Corollary 6.17. *Fix an (unordered) interacting triple* $\iota = \boldsymbol{jk\ell} \in \mathcal{I}$ *with* $|k| = |j|$ *and let* $q_{\boldsymbol{\ell}}, q_{\boldsymbol{k}} \neq 0$*. Then there exist* $\tau_+^\iota, \tau_-^\iota \geqslant 0$ *such that* $(\varphi^\iota_{\tau^\iota_+}(q))_{\mathbf{j}} > 0$ *and* $(\varphi^\iota_{\tau^\iota_-}(q))_{\mathbf{j}} = 0$.

Proof of Lemma 6.16. Recall from (6.13) that if $|j| = |k| \neq |\ell|$ we have $C_{jk} = 0$. This implies that, by our choice of $|k| = |j|$, $\dot{q}_{\boldsymbol{\ell}} = 0$ and $q'_{\boldsymbol{\ell}} = q_{\boldsymbol{\ell}}$. Again by (6.13) and since to have an interacting triple $\ell = j + k$ we must have

$$\langle k^\perp, \ell \rangle = \langle k^\perp, k + j \rangle = \langle k^\perp, j \rangle = \langle (k+j)^\perp - j^\perp, j \rangle = \langle \ell^\perp, j \rangle = -\langle j^\perp, \ell \rangle , \quad (6.42)$$

so that

$$C_{k\ell} = \frac{\langle k, \ell^\perp \rangle}{4\pi} \left(\frac{1}{|k|^2} - \frac{1}{|\ell|^2} \right) = -\frac{\langle j, \ell^\perp \rangle}{4\pi} \left(\frac{1}{|j|^2} - \frac{1}{|\ell|^2} \right) = -C_{j\ell} \, . \quad (6.43)$$

This implies that the dynamics of the vector $\tilde{q} := (q_{\boldsymbol{j}}, q_{\boldsymbol{k}})$ can be written as $\dot{\tilde{q}} = \tilde{C}\tilde{q}^\perp$ for $\tilde{C} := C_{j\ell} q_{\boldsymbol{\ell}} \neq 0$, proving the claim. □

Existence of invariant measure

As with conservative Lorenz-96, each vector field of the 2d Euler splitting is divergence free and so Lebesgue measure in $\mathbb{R}^n$ is invariant. Consequently, we have

Lemma 6.18. *Let* Leb *denote Lebesgue measure on* $\mathbb{R}^n$*. The measure obtained by conditioning* Leb *to lie on* $\mathcal{Q} \subset Q_0(E, \mathcal{E})$*, (or equivalently conditioned to lie on* $Q_0(E, \mathcal{E})$*) is* P_h*-invariant.*

Proof. As in the proof of Proposition 6.2 we have that Lebesgue measure in $\mathbb{R}^n$ is P_h-invariant. Since the vector fields V_j defined in (6.19) are divergence free, the continuity equation[2] reads

$$\partial_t \operatorname{Leb} + \operatorname{div}(V_j \operatorname{Leb}) = \partial_t \operatorname{Leb} + \nabla \operatorname{Leb} \cdot V_j = 0.$$

[2] As in the proof of Proposition 6.2, the continuity equation is intended here in the weak sense.

Because each flow $\varphi^{(j)}$ conserves energy E and enstrophy $\mathcal{E}$, we locally fiber $\mathbb{R}^n$ using coordinates $(E, \mathcal{E}, \vartheta) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^{n-2}$. In these coordinates, we have $V_j(E, \mathcal{E}, \vartheta) = 0\,\partial_E + 0\,\partial_{\mathcal{E}} + v_j(E, \mathcal{E}, \vartheta)\nabla_\vartheta$ so by a change of coordinates of the divergence operator the stationary equation becomes

$$0 = \operatorname{div}\left(V_j(x)\operatorname{Leb}(x)\right) = u(E, \mathcal{E}, \vartheta)\operatorname{div}_\vartheta(\operatorname{Leb}(E, \mathcal{E}, \vartheta)v_j(E, \mathcal{E}, \vartheta))\,, \tag{6.44}$$

where $\operatorname{div}_\vartheta$ denotes the "angular" terms of the divergence in $(E, \mathcal{E}, \vartheta)$-coordinates, and $u(E, \mathcal{E}, \vartheta)$ result from the change of variables. Hence, we can factor the solution $\operatorname{Leb}(E, \mathcal{E}, \vartheta) = \overline{\operatorname{Leb}}(\vartheta|E, \mathcal{E}) \cdot \operatorname{Leb}^\perp(E, \mathcal{E})$, where $\overline{\operatorname{Leb}}(\vartheta|E, \mathcal{E})$ is the conditional density of Lebesgue measure on a fiber, solving $u(E, \mathcal{E}, \vartheta)\operatorname{div}_\vartheta(\overline{\operatorname{Leb}}(\vartheta|E, \mathcal{E})v_j(E, \mathcal{E}, \vartheta)) = 0$ for any choice of $E/(2N^2) < \mathcal{E} < E$. This proves the invariance of $\overline{\operatorname{Leb}}(\vartheta|E, \mathcal{E})$ under the flow map for any value of the flow times τ. The stationarity of $\overline{\operatorname{Leb}}(\vartheta)$ under P_h follows immediately as in Proposition 6.2 $\square$

Spanning

For j, k, ℓ in $\mathbb{Z}_N^2$ with $j + k - \ell = 0$ define $M_{jk\ell}$ to be the matrix

$$M_{jk\ell} := \begin{pmatrix} | & | & | & | \\ V_{a_j a_k a_\ell} & V_{a_j b_k b_\ell} & V_{b_j a_k b_\ell} & V_{b_j b_k a_\ell} \\ | & | & | & | \end{pmatrix} = \begin{pmatrix} C_{k\ell}a_k a_\ell & C_{k\ell}b_k b_\ell & 0 & 0 \\ 0 & 0 & C_{k\ell}a_k b_\ell & -C_{k\ell}b_k a_\ell \\ C_{j\ell}a_j a_\ell & 0 & C_{j\ell}b_j b_\ell & 0 \\ 0 & C_{j\ell}a_j b_\ell & 0 & -C_{j\ell}b_j a_\ell \\ -C_{jk}a_j a_k & 0 & 0 & C_{jk}b_j b_k \\ 0 & -C_{jk}a_j b_k & -C_{jk}b_j a_k & 0 \end{pmatrix} \tag{6.45}$$

and let $M'_{jk\ell}$ and $M''_{jk\ell}$ be the 4-by-4 and 2-by-4 matrices consisting of the bottom four and bottom two rows of $M_{jk\ell}$, respectively. Straightforward Gaussian elimination shows that M, M', and M'' have ranks 4, 3, and 2 whenever C_{jk}, $C_{j\ell}$, $C_{k\ell}$, a_j, b_j, a_k, b_k, a_ℓ, and b_ℓ are nonzero.

Recalling that a point $q \in \mathbb{R}^n$ is *generic* if all its coordinates are nonzero, we have

Lemma 6.19. *The family of vector fields*

$$\mathcal{V} := \left\{V_{a_j a_k a_\ell}, V_{a_j b_k b_\ell}, V_{b_j a_k b_\ell}, V_{b_j b_k a_\ell} : j, k, \ell \in \mathbb{Z}_N^2 \text{ and } j + k - \ell = 0\right\}$$

span $T_q\mathcal{Q}$ at every generic point q in $\mathcal{Q}$.

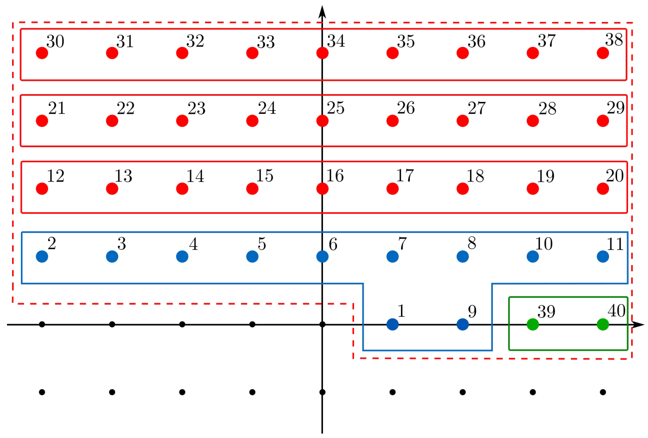


FIGURE 6.3: *Ordering indices.* Ordering of $\mathbb{Z}_N^2$ when $N = 4$.

Proof. Fix a generic point q in $\mathcal{Q}$. The main idea of the proof is to choose an enumeration of $\mathbb{Z}_N^2$ and a subset of vector fields from $\mathcal{V}$ so that the matrix made up of these vector fields evaluated at q is in a convenient form whose rank is readily deduced. Formally, the enumeration is the bijection $F : \mathbb{Z}_N^2 \to \{1, \ldots, 2N(N+1)\}$ given by

$$F(j) := \begin{cases} 1 & j = (1,0)\,, \\ 5 + N & j = (2,0)\,, \\ j_1 + N(2N+1) & j = (j_1, 0) \text{ with } j_1 > 2\,, \\ j_1 + 2 + N & j = (j_1, 1) \text{ with } j_1 < 3\,, \\ j_1 + 3 + N & j = (j_1, 1) \text{ with } j_1 \geqslant 3\,, \\ j_1 + 2 - N + (2N+1)j_2 & j = (j_1, j_2) \text{ with } j_2 > 1\,. \end{cases}$$

Figure 2 gives this enumeration in the case $N = 4$. Informally, F starts at $(1,0)$, then counts lattice points from left to right along the horizontal line $y = 1$ until the point $(2,1)$, which corresponds to $4+N$. It then assigns $5+N$ to $(2,0)$ and continues counting along the line $y = 1$. From there it moves up to the lines $y = 2$, $y = 3$, and so on, counting from left to right along each. Finally, it goes back down to the line $y = 0$ and counts the remaining indices from left to right.

The motivation for F is that all horizontally-adjacent indices (j_1, j_2) and (j_1+1, j_2) form an interacting triple together with $(1,0)$. Fix for the moment an integer $y > 1$

and consider the yth horizontal line of $\mathbb{Z}_N^2$; that is, the points with second coordinate y. These are outlined by red blocks in Figure 6.3. By the preceding remarks we can choose the vector fields corresponding to the horizontally-adjacent indices and concatenate them column-wise to get the block matrix

$$B_y := \left(\begin{array}{c|c|c|c} \widetilde{M}_y & * & * & * \\ \hline 0 & M''_{y,-N+2} & * & * \\ \hline 0 & 0 & \ddots & * \\ \hline 0 & 0 & 0 & M''_{y,N} \end{array}\right).$$

Here, slightly abusing notation, each $M''_{y,i}$ is the 2-by-4 matrix consisting of the bottom two rows of (6.45) for the indices $j = (1,0), k = (i-1,y), \ell = (j,y)$ and

$$\widetilde{M}_y := \begin{pmatrix} C_{j\ell}a_j a_\ell & 0 & C_{j\ell}b_j b_\ell & 0 & 0 & 0 \\ 0 & C_{j\ell}a_j b_\ell & 0 & -C_{j\ell}b_j a_\ell & 0 & 0 \\ -C_{jk}a_j a_k & 0 & 0 & C_{jk}b_j b_k & -C_{j'k'}a_{j'}a_{k'} & \\ 0 & -C_{jk}a_j b_k & -C_{jk}b_j a_k & 0 & 0 & -C_{j'k'}a_{j'}b_{k'} \end{pmatrix}$$

where $j = (1,0), k = (-N,y), \ell = (-N+1,y)$ and $j' = (0,1)$ and $k' = (-N+1,y-1)$. This is M' with two columns from the interacting triple $(0,1),(-N+1,y-1),(-N+1,y)$ adjoined to the end. Note that these adjoined columns contribute entries in the coordinates corresponding to $(0,1)$ and $(-N+1,y-1)$, but these come before all indices in the yth row for our ordering. By adding the latter two columns, $\widetilde{M}_y$ has rank 4 at any generic point. Further, since each $M''_{y,j}$ has rank 2, each B_y has rank $4 + 2(2N-1) = 4N+2$. This establishes spanning of the red blocks in Figure 6.3.

For the blue block we perform a similar procedure to the one above to get

$$B_1 := \left(\begin{array}{c|c|c|c|c|c} M_{123} & * & * & * & * & * \\ \hline 0 & M''_{1,-N+2} & * & * & * & * \\ \hline 0 & 0 & \ddots & * & * & * \\ \hline 0 & 0 & 0 & \widehat{M} & * & * \\ \hline 0 & 0 & 0 & 0 & \ddots & * \\ \hline 0 & 0 & 0 & 0 & 0 & M''_{1,N} \end{array}\right)$$

where M_{123} is the matrix from (6.45) for the interacting triple $(1,0),(-N,1),(-N+1,1)$, each M'' is as before, and $\widehat{M}$ is the 6-by-8 matrix

$$\widehat{M} := \left(\begin{array}{cccc|cccc} & & & & 0 & 0 & 0 & 0 \\ & M'_{1,N+3,N+4} & & & 0 & 0 & 0 & 0 \\ & & & & --- & --- & --- & --- \\ --- & --- & --- & --- & & & & \\ 0 & 0 & 0 & 0 & & M'_{N+2,N+4,N+5} & & \\ 0 & 0 & 0 & 0 & & & & \end{array}\right)$$

located at the rows corresponding to $N+3, N+4$, and $N+5$. The reason for $\widehat{M}$, and for considering the blue block separately, is that $C_{jk} = 0$ when $j = (1,0)$ and $k = (0,1)$. The matrix M has rank 6 at a generic point. Since M_{123} has rank 4, $\widehat{M}$ has rank 6, and each of the $2N-3$ remaining M'' blocks has rank 2, the matrix B_y has rank $4 + 6 + 2(2N-3) = 4N + 4$.

Finally, none of the indices of the green block interact with $(1,0)$ since the C_{jk} are all 0 in this case. However, by an entirely similar procedure to above, we can use the interactions between $(0,1),(x,0)$, and $(x,1)$ for $x > 1$ to get a rank $2(N-2)$ block matrix for the last $N-2$ coordinates of the form

$$B_{N+1} := \left(\begin{array}{c|c|c|c} \tilde{M}''_{0,2} & * & * & * \\ \hline 0 & \tilde{M}''_{0,3} & * & * \\ \hline 0 & 0 & \ddots & * \\ \hline 0 & 0 & 0 & \tilde{M}''_{0,N} \end{array}\right)$$

where $\tilde{M}''_{0,x} = M''_{(0,1),(x,0),(x,1)}$ for $M''_{jk\ell}$ consisting of the two bottom rows of (6.45). Combining the above results we observe that there is an ordering of indices and vector fields such that the matrix whose columns consist of these vector fields has the form

$$B := \left(\begin{array}{c|c|c|c} B_1 & * & * & * \\ \hline 0 & B_2 & * & * \\ \hline 0 & 0 & \ddots & * \\ \hline 0 & 0 & 0 & B_{N+1} \end{array}\right).$$

Moreover, B has rank

$$\operatorname{rank}(B) = \operatorname{rank}(B_1) + \operatorname{rank}(B_{N+1}) + \sum_{y=2}^{N} \operatorname{rank}(B_y) = 4N(N+1) - 2 = n - 2$$

at every generic point in $\mathcal{Q}$. Now since the dynamics conserve energy and enstrophy, every tangent vector to $\mathcal{Q}$ is perpendicular to the normal vectors for these two quantities which are linearly independent at every generic point. Therefore the maximum dimension of $T_q\mathcal{Q}$ is $n-2$, and by the above argument we have shown the vector fields $\mathcal{V}$ span $T_q\mathcal{Q}$ at q. □

6.2.4 Positive top Lyapunov exponent

In the previous section we obtained the splitting vector fields of 2d Euler corresponding to (6.17) by restricting to indices in $\mathbb{Z}^2_+$ via the constraint $q_{-j} = \bar{q}_j$. In this section we do not take this additional step of restricting to $\mathbb{Z}^2_+$ and instead redefine $\mathbb{Z}^2_N := \{j \in \mathbb{Z}^2 : \max\{|j_1|, |j_2|\} \leqslant N\}$. In this case the splitting of the Nth Galerkin approximation of 2d Euler becomes

$$V = \sum_{j+k+\ell=0} V_{a_j a_k a_\ell} + V_{a_j b_k b_\ell} + V_{b_j a_k b_\ell} + V_{b_j b_k a_\ell} \tag{6.46}$$

with indices now ranging over $\mathbb{Z}^2_N$ and the vector fields $V_{a_j a_k a_\ell}, V_{a_j b_k b_\ell}, V_{b_j a_k b_\ell}$, and $V_{b_j b_k a_\ell}$ now given by

$$\begin{cases} \dot{a}_j = -C_{k\ell} a_k a_\ell \\ \dot{a}_k = -C_{j\ell} a_j a_\ell \\ \dot{a}_\ell = -C_{jk} a_j a_k \end{cases} \quad \begin{cases} \dot{a}_j = C_{k\ell} b_k b_\ell \\ \dot{b}_k = C_{j\ell} a_j b_\ell \\ \dot{b}_\ell = C_{jk} a_j b_k \end{cases} \quad \begin{cases} \dot{b}_j = C_{k\ell} a_k b_\ell \\ \dot{a}_k = C_{j\ell} b_j b_\ell \\ \dot{b}_\ell = C_{jk} b_j a_k \end{cases} \quad \begin{cases} \dot{b}_j = C_{k\ell} b_k a_\ell \\ \dot{b}_k = C_{j\ell} b_j a_\ell \\ \dot{a}_\ell = C_{jk} b_j b_k, \end{cases} \tag{6.47}$$

respectively. Note this differs from the previous splitting only by a few negative signs. In particular, all results for the previous splitting apply to this slightly altered one.

Similar to conservative Lorenz-96, throughout this section we refer to the orbits from Proposition 6.10 that contain all nondegenerate points as *generic orbits*. The main result of this section is

Theorem 6.20. *The top Lyapunov exponent of the Nth Galerkin approximated Euler random splitting* (6.46) *on a generic orbit is positive for every $h > 0$ and $N \geqslant 3$.*

Proof. We saw in Proposition 6.10 that for any generic orbit $\mathcal{X}$ of the Galerkin approximated 2d Euler splittings there exists a unique P_h-invariant measure μ on $\mathcal{X}$ for every $h > 0$. Thus the Lyapunov exponents exist and are almost-surely constant on generic orbits. We also showed in Lemma 6.18 that each such μ is the disintegration of Lebesgue measure onto its respective orbit and is therefore invariant under $\Phi_{h\tau}$ for every τ. In particular, the pushforward measures μ_m defined in Section 5.2 satisfy $\mu_m = \mu$ and hence $\mathbb{E}D_{KL}(\mu_m \parallel \mu) = 0 < \infty$ for all m. Furthermore $\lambda_\Sigma = 0$ (and hence $\lambda_1 \geqslant 0$) since the splitting vector fields conserve Euclidean norm. This establishes the hypotheses of Theorem 5.2 for random splittings on generic orbits of 2d Euler. And since $\lambda_\Sigma = 0$, Theorem 5.2 says that if $\lambda_1 = 0$ then Alternative 1 or 2 must hold. So to prove Theorem 6.20, it remains to show neither alternative holds.

Simple computation shows the constants C_{jk} satisfy $C_{jk} = 0$ and $C_{j\ell} = -C_{k\ell}$ whenever $|j| = |k|$ and $j + k + \ell = 0$, e.g. when $j = (1,0)$, $k = (0,1)$, and $\ell = -(1,1)$. In this case the equation $\dot{q} = V_{a_j a_k a_\ell}(q)$ is given by

$$\begin{cases} \dot{a}_j = C_{j\ell} a_\ell a_k \\ \dot{a}_k = -C_{j\ell} a_\ell a_j \\ \dot{a}_\ell = 0, \end{cases}$$

which is equivalent to (5.28). So Proposition 5.12 rules out Alternative 1. All that remains then is to rule out Alternative 2, which we do via Proposition 5.13.

As with conservative Lorenz-96, to rule out Alternative 2 we use Proposition 5.13. Again, the main idea of the proof is to fix a point $(q^*, \eta^*) \in T\mathcal{X}$ and to check that the Lie Bracket condition holds for $\mathcal{V}$ from (6.47) at that point. In this spirit, we proceed to choose a subset of vector fields and related commutators (indexed by the related triples of interacting indices) from $\mathcal{V}$ whose spanning dimension, when evaluated at

(q^*, η^*), can be readily deduced. Concretely, this will be done by computing the rank of the matrix whose columns are given by such vector fields. To increase readability, while presenting the full idea of the computation and its results we have suppressed the lengthy algebraic manipulations such computation entails. In the interest of reproducibility, Mathematica code to reproduce such computations is available at [1]. The code's inputs and outputs are reported in Appendix C.

Focusing on a given triple $j, k, \ell \in \mathbb{Z}_N^2$, we define the vector fields in $\mathcal{X}$ which we express in 6 dimensions – corresponding to coordinates $(a_j, b_j, a_k, b_k, a_\ell, b_\ell)$ as

$$V_{jk\ell}^{(1)} = \begin{pmatrix} -C_{k\ell}a_k a_\ell \\ 0 \\ -C_{\ell j}a_j a_\ell \\ 0 \\ -C_{jk}a_j a_k \\ 0 \end{pmatrix} \quad V_{jk\ell}^{(2)} = \begin{pmatrix} C_{k\ell}b_k b_\ell \\ 0 \\ 0 \\ C_{\ell j}a_j b_\ell \\ 0 \\ C_{jk}a_j b_k \end{pmatrix} \quad V_{jk\ell}^{(3)} = \begin{pmatrix} 0 \\ C_{k\ell}a_k b_\ell \\ C_{\ell j}b_j b_\ell \\ 0 \\ 0 \\ C_{jk}b_j a_k \end{pmatrix} \quad V_{jk\ell}^{(4)} = \begin{pmatrix} 0 \\ C_{k\ell}b_k a_\ell \\ 0 \\ C_{\ell j}b_j a_\ell \\ C_{jk}b_j b_k \\ 0 \end{pmatrix} \tag{6.48}$$

The commutators of these vector fields read

$$[V_{jk\ell}^{(1)}, V_{jk\ell}^{(2)}] = \begin{pmatrix} 0 \\ 0 \\ C_{k\ell}C_{\ell j}a_\ell b_k b_\ell \\ -C_{k\ell}C_{\ell j}b_\ell a_k a_\ell \\ C_{k\ell}C_{jk}a_k b_k b_\ell \\ -C_{k\ell}C_{jk}b_k a_k a_\ell \end{pmatrix} \qquad [V_{jk\ell}^{(1)}, V_{jk\ell}^{(3)}] = \begin{pmatrix} C_{k\ell}C_{\ell j}a_\ell b_j b_\ell \\ -C_{k\ell}C_{\ell j}b_\ell a_j a_\ell \\ 0 \\ 0 \\ C_{jk}C_{\ell j}a_j b_j b_\ell \\ -C_{\ell j}C_{jk}b_k a_j a_\ell \end{pmatrix}$$

$$[V_{jk\ell}^{(1)}, V_{jk\ell}^{(4)}] = \begin{pmatrix} C_{k\ell}C_{jk}a_k b_j b_k \\ -C_{k\ell}C_{jk}a_j a_k b_k \\ C_{\ell j}C_{jk}a_j b_j b_k \\ -C_{\ell j}C_{jk}a_j a_k b_j \\ 0 \\ 0 \end{pmatrix} \qquad [V_{jk\ell}^{(2)}, V_{jk\ell}^{(3)}] = \begin{pmatrix} -C_{k\ell}C_{jk}a_k b_j b_k \\ C_{k\ell}C_{jk}a_j a_k b_k \\ C_{\ell j}C_{jk}a_j b_j b_k \\ -C_{\ell j}C_{jk}a_j a_k b_j \\ 0 \\ 0 \end{pmatrix}$$

$$[V_{jk\ell}^{(2)}, V_{jk\ell}^{(4)}] = \begin{pmatrix} -C_{k\ell}C_{\ell j}a_\ell b_j b_\ell \\ C_{k\ell}C_{\ell j}b_\ell a_j a_\ell \\ 0 \\ 0 \\ C_{jk}C_{\ell j}a_j b_j b_\ell \\ -C_{\ell j}C_{jk}b_k a_j a_\ell \end{pmatrix} \qquad [V_{jk\ell}^{(3)}, V_{jk\ell}^{(4)}] = \begin{pmatrix} 0 \\ 0 \\ -C_{k\ell}C_{\ell j}a_\ell b_k b_\ell \\ C_{k\ell}C_{\ell j}b_\ell a_k a_\ell \\ C_{k\ell}C_{jk}a_k b_k b_\ell \\ -C_{k\ell}C_{jk}b_k a_k a_\ell \end{pmatrix}$$

This yields the extended, 12-dimensional vector fields on $T\mathcal{X}$ obtained by stacking the vector fields (6.48) and their commutators with the corresponding linearizations. Throughout, we represent such vector fields reordering the coordinates of $T\mathcal{X}$ for a triple of indices $j, k, \ell \in \mathbb{Z}_N^2$ as

$$(a_j, b_j, \eta_j^a, \eta_j^b, a_k, b_k, \eta_k^a, \eta_k^b, a_\ell, b_\ell, \eta_\ell^a, \eta_\ell^b)\,.$$

where for any $j \in \mathbb{Z}_N^2$, η_j^a, η_j^b are the linearization coordinates corresponding to a_j, b_j . Then, denoting by $[f]$ the linearization of a map $f \;:\; \mathbb{R}^6 \to \mathbb{R}$, i.e., $[f](x,\eta) := Df \cdot \eta$, we write such extended vector fields for a given triple $j, k, \ell \in \mathbb{Z}_N^2$ as the columns of a matrix $M_{jk\ell}$, obtaining

$$M_{jk\ell} = \begin{pmatrix} | & | & | & | & | & | & | & | & | & | \\ V^{(1)}_{jk\ell} & V^{(2)}_{jk\ell} & V^{(3)}_{jk\ell} & V^{(4)}_{jk\ell} & [V^{(1)}_{jk\ell}, V^{(2)}_{jk\ell}] & [V^{(1)}_{jk\ell}, V^{(3)}_{jk\ell}] & [V^{(1)}_{jk\ell}, V^{(4)}_{jk\ell}] & [V^{(2)}_{jk\ell}, V^{(3)}_{jk\ell}] & [V^{(2)}_{jk\ell}, V^{(4)}_{jk\ell}] & [V^{(3)}_{jk\ell}, V^{(4)}_{jk\ell}] \\ | & | & | & | & | & | & | & | & | & | \end{pmatrix}$$

$$= \begin{pmatrix}
C_{k\ell}a_k a_\ell & C_{k\ell}b_k b_l & 0 & 0 & B^j_{k,b} & B^j_{\ell,b} & 0 & 0 & -B^j_{\ell,b} & -B^j_{k,b} \\
0 & 0 & C_{k\ell}a_k b_\ell & C_{k\ell}b_k a_\ell & -B^j_{k,a} & -B^j_{\ell,a} & 0 & 0 & B^j_{\ell,a} & B^j_{k,a} \\
C_{k\ell}[a_k a_\ell] & C_{k\ell}[b_k b_l] & 0 & 0 & [B^j_{k,b}] & [B^j_{\ell,b}] & 0 & 0 & -[B^j_{\ell,b}] & -[B^j_{k,b}] \\
0 & 0 & C_{k\ell}[a_k b_\ell] & C_{k\ell}[b_k a_\ell] & -[B^j_{k,a}] & -[B^j_{\ell,a}] & 0 & 0 & [B^j_{\ell,a}] & [B^j_{k,a}] \\
C_{\ell j}a_j a_\ell & 0 & C_{\ell j}b_j b_\ell & 0 & B^k_{j,b} & 0 & B^k_{\ell,b} & -B^k_{\ell,b} & 0 & B^k_{j,b} \\
0 & C_{\ell j}a_j b_\ell & 0 & C_{\ell j}b_j a_\ell & -B^k_{j,a} & 0 & -B^k_{\ell,a} & B^k_{\ell,a} & 0 & -B^k_{j,a} \\
C_{\ell j}[a_j a_\ell] & 0 & C_{\ell j}[b_j b_\ell] & 0 & [B^k_{j,b}] & 0 & [B^k_{\ell,b}] & -[B^k_{\ell,b}] & 0 & [B^k_{j,b}] \\
0 & C_{\ell j}[a_j b_\ell] & 0 & C_{\ell j}[b_j a_\ell] & -[B^k_{j,a}] & 0 & -[B^k_{\ell,a}] & [B^k_{\ell,a}] & 0 & -[B^k_{j,a}] \\
C_{jk}a_j a_k & 0 & 0 & C_{jk}b_k b_j & 0 & B^\ell_{j,b} & B^\ell_{k,b} & B^\ell_{k,b} & B^\ell_{j,b} & 0 \\
0 & C_{jk}a_j b_k & C_{jk}a_k b_j & 0 & 0 & -B^\ell_{j,a} & -B^\ell_{k,a} & -B^\ell_{k,a} & -B^\ell_{j,a} & 0 \\
C_{jk}[a_j a_k] & 0 & 0 & C_{jk}[b_k b_j] & 0 & [B^\ell_{j,b}] & [B^\ell_{k,b}] & [B^\ell_{k,b}] & [B^\ell_{j,b}] & 0 \\
0 & C_{jk}[a_j b_k] & C_{jk}[a_k b_j] & 0 & 0 & -[B^\ell_{j,a}] & -[B^\ell_{k,a}] & -[B^\ell_{k,a}] & -[B^\ell_{j,a}] & 0
\end{pmatrix},$$

with

$$B^\ell_{k,a} := B_k a_\ell \qquad \text{for} \qquad B_k := C_{jk}C_{k\ell}a_k b_k\,.$$

We further denote by $M'_{jk\ell}$ the last four rows intersected with the first, second, sixth and seventh column of the above matrix:

$$M'_{jk\ell} := \begin{pmatrix} C_{jk}a_j a_k & 0 & B^\ell_{j,b} & B^\ell_{k,b} \\ 0 & C_{jk}a_j b_k & -B^\ell_{j,a} & -B^\ell_{k,a} \\ C_{jk}[a_j a_k] & 0 & [B^\ell_{j,b}] & [B^\ell_{k,b}] \\ 0 & C_{jk}[a_j b_k] & -[B^\ell_{j,a}] & -[B^\ell_{k,a}] \end{pmatrix}. \tag{6.49}$$

To check spanning, we now write a subset of the vector fields introduced above as the columns of a matrix whose rank will be shown to be $\dim T\mathcal{X} = 4n - 4$. To choose such vector fields, we introduce a convenient enumeration of the index space $\mathbb{Z}_N^2$ as the bijection $F : \{0, \ldots, 2N(N+1)\} \to (\mathbb{Z}_N^2)^3$ given by

$$F(i) := \begin{cases} (1,0), (-N+i,1), (-N+i+1,1) & i \leqslant N-2 \\ (-1,1), (-3,1), (2,0) & i = N-1\,, \\ (2,0), (-2,1), (0,1) & i = N \\ (0,1), (2,0), (2,1) & i = N+1 \\ (1,0), (2,1), (1,1) & i = N+2 \\ (1,0), (-N+i-1,1), (-N+i,1) & N+3 \leqslant i \leqslant 2N \\ (0,1), (-2N+i,1), (-2N+i+1,0) & 2N+1 \leqslant i \leqslant 3N-1 \\ (1,0), (L_1(i), L_2(i)), (L_1(i)+1, L_2(i)) & 3N \leqslant i \leqslant 2N(N+1) \end{cases}$$

where $(L_1(i), L_2(i))$ is the element of $\mathbb{Z}_N^2$ obtained by starting from $(-N,1)$ (for $j = 3N$) and for each new value of j proceeding incrementally to the right (i.e., adding 1 to the first component of the two-dimensional index) until $(N,1)$, then moving to the row above at $(-N,2)$, then again moving progressively to the right until $(N,2)$ and so on until (N,N). Formally $(L_1(i), L_2(i))$ is therefore defined as

$$L_1(i) := -N + \mathrm{mod}(i - 3N, 2N+1)$$
$$L_2(i) := \lfloor (i-3N)/(2N+1) \rfloor\,,$$

where $\mathrm{mod}(\,\cdot\,,\,\cdot\,)$ and $\lfloor\,\cdot\,\rfloor$ respectively denote integer part and the modulo operation, i.e., the remainder of division of the first argument by the second. Note that this ordering of the indices is motivated by the fact that each time we let a new triple interact, two of the indices have already interacted in a previous triple, adding exactly a new one. Because of this, from this enumeration of the triples follows an enumeration of the indices: after mapping the first three interacting indices $(1,0), (-N+i,1), (-N+i+1,1)$ to $1,2,3$, we proceed incrementally assigning the number i to the third element of $F(i-3)$. We also note for future reference that for all interacting triples considered above we have nonvanishing interaction constants $C_{jk}, C_{j\ell}, C_{k\ell}$.

Using the above enumeration we can define

$$\mathcal{A}_i = \begin{pmatrix} | & | & | & | \\ V^{(1)}_{F(i)} & V^{(2)}_{F(i)} & \left[V^{(1)}_{F(i)}, V^{(3)}_{F(i)}\right] & \left[V^{(1)}_{F(i)}, V^{(4)}_{F(i)}\right] \\ | & | & | & | \end{pmatrix} = \begin{pmatrix} \bigstar \\ M'_{F(i)} \\ \mathbf{0} \end{pmatrix},$$

where $\bigstar$ indicates irrelevant entries, $\mathbf{0}$ indicates the rest of the matrix is filled with zeros. We then define the $4n \times 4n - 4$ matrix

$$\mathcal{A} := \begin{pmatrix} M_0 & | & | & \cdots & | \\ \mathbf{0} & \mathcal{A}_1 & \mathcal{A}_1 & & \mathcal{A}_{n-3} \\ & | & | & & | \end{pmatrix} = \left(\begin{array}{c|cccccc} & \star & \star & \star & \star & \star & \star \\ M_0 & \star & \star & \star & \star & \star & \star \\ & \star & \star & \star & \star & \star & \star \\ \hline 0 & \boxed{M'_{F(1)}} & \star & \star & \star & \star & \star \\ 0 & 0 & \boxed{M'_{F(2)}} & \star & \star & \star & \star \\ 0 & 0 & 0 & \boxed{M'_{F(3)}} & \star & \star & \star \\ 0 & 0 & 0 & 0 & \boxed{M'_{F(4)}} & \star & \star \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & \boxed{M'_{F(n-3)}} \end{array}\right),$$

where M_0 is a 12×8 matrix obtained by removing the eigth and the tenth column from $M_{F(0)}$. Note that the above matrix inherits its structure from the indexing of the interacting triples: each triple contains exactly one index that has not yet interacted with other modes. To prove the desired result, it therefore remains to show that for every compatible[3] choice of conserved quantities $(E, \mathcal{E})$ there exists a point (q^*, η^*) where each of the $n-3$ block elements has rank 4 and the M_0 matrix has rank 8, thereby yielding the desired total rank of $4n-4$. Proceeding with this plan, in the following we fix the vector (q^*, η^*) where we establish spanning as

$$a^*_j = \begin{cases} \alpha_1 & \text{if } j \in \{(0,1),(1,0)\} \\ \alpha_2 & \text{if } j = (N,N) \\ \beta & \text{else} \end{cases} \qquad b^*_j = \begin{cases} 2a^*_j & \text{if } j \in \{(0,1),(1,0),(-1,1),(2,0)\} \\ a^*_j & \text{else} \end{cases}$$

$$(\eta^*)^{\#}_j = \begin{cases} 1 & \text{if } \# = a \\ -1/2 & \text{if } j \in \{(0,1),(1,0),(-1,1),(2,0)\}, \# = b \\ -1 & \text{else} \end{cases}$$

[3] We call a pair of conserved quantities compatible when there exists a nondegenerate point in state space with the given energy and enstrophy. This is true if $E/(2N^2) < \mathcal{E} < E$

for values of $\alpha_1, \alpha_2 > 0$ to be chosen shortly and $\beta > 0$ a sufficiently small, free parameter. Throughout, we denote the set of indices for which the complex part is twice the real part in q^* by $J_2 := \{(0,1),(1,0),(-1,1),(2,0)\}$. Note $\eta^* \in T_{q^*}\mathcal{Q}$ since

$$\eta^* \cdot \nabla E = \eta^* \cdot x^* = 0 \qquad \text{and} \qquad \eta^* \cdot \nabla \mathcal{E} = \sum_{j \in \mathbb{Z}_N^2} \frac{1}{|j|^2}((\eta^*)_j^a a_j^* + (\eta^*)_j^b b_j^*) = 0\,.$$

Furthermore, it is easy to check that for every compatible pair $(E, \mathcal{E})$, we have $\mathcal{E}(q^*) = \mathcal{E}$, $\|q^*\| = E$ upon choosing

$$\alpha_1 := \sqrt{\frac{1}{10} G_{\mathcal{E},E}(\beta)}\,, \qquad \alpha_2 := \frac{1}{\sqrt{2}} \sqrt{E - 2\beta^2 \left[\frac{15}{8} + \sum_{k \in \mathbb{Z}_N^2 \setminus J_2} \frac{1}{|k|^2}\right] - G_{\mathcal{E},E}(\beta)}\,,$$

where

$$G_{\mathcal{E},E}(\beta) := \frac{1}{1 - \frac{1}{2N^2}} \left(\mathcal{E} - \frac{1}{2N^2} E - 2\beta^2 \left[\left(\frac{15}{8} - \frac{1}{N^2}\right) + \sum_{k \in \mathbb{Z}_N^2 \setminus J_2} \left(\frac{1}{|k|^2} - \frac{1}{2N^2}\right)\right]\right)$$

for $\beta \in (0, Z_+(\mathcal{E}, E))$, $Z_+(\mathcal{E}, E)$ being the positive zero of $G_{\mathcal{E},E}(\beta)$. Here, the expression $2\beta^2 \frac{15}{8}$ results from $\frac{a_k^2}{|k^2|} + \frac{b_k^2}{|k^2|}$ for $k \in \{(-1,1),(0,2)\}$.

We further note that for any triple $(j,k,\ell) \in \{F(i)\}_{i \in \{1,\dots,n-3\} \setminus \{N+1\}}$, since $j \in J_2$, $k \notin J_2 \cup \{(N,N)\}$, the vector (q^*, η^*) projected on the interacting coordinates can be written

$$(a_j, b_j, \eta_j^a, \eta_j^b, a_k, b_k, \eta_k^a, \eta_k^b, a_\ell, b_\ell, \eta_\ell^a, \eta_\ell^b) = (\gamma_1, 2\gamma_1, 1, -\tfrac{1}{2}, \beta, \beta, , 1, -1, \gamma_2, m\gamma_2, 1, -\tfrac{1}{m}),$$

with $\gamma_1 \in \{\alpha_1, \beta\}$, $\gamma_1 \in \{\alpha_1, \alpha_2, \beta\}$, $m = 2$ for $\ell \in \{(0,1),(1,0),(-1,1),(0,2)\}$ and $m = 1$ otherwise. Evaluating $M'_{jk\ell}$ at (q^*, η^*) for such interacting triples one therefore obtains a matrix of the form

$$M'_{jk\ell} = \begin{pmatrix} \beta\gamma_1 C_{jk} & 0 & 2C_{jk}C_{j\ell}\gamma_1^2\gamma_2 m & -\beta^2\gamma_2 m C_{jk}C_{k\ell} \\ 0 & \beta\gamma_1 C_{jk} & -2\gamma_1^2\gamma_2 C_{jk}C_{j\ell} & \beta^2\gamma_2 C_{jk}C_{k\ell} \\ \beta C_{jk} + \gamma_1 C_{jk} & 0 & C_{jk}C_{j\ell}\left(\frac{3}{2}\gamma_1\gamma_2 m - \frac{2\gamma_1^2}{m}\right) & \frac{\beta^2 C_{jk}C_{k\ell}}{m} \\ 0 & \beta C_{jk} - \gamma_1 C_{jk} & C_{jk}C_{j\ell}\left(-2\gamma_1^2 - \frac{3}{2}\gamma_2\gamma_1\right) & \beta^2 C_{jk}C_{k\ell} \end{pmatrix},$$

whose determinant is given by

$$\det(M'_{jk\ell}) = \frac{3\beta^3\gamma_1^3\gamma_2 C_{jk}^4 C_{j\ell} C_{k\ell}\left(\beta + \beta m^2 + 2\gamma_2 m^2\right)}{2m}\,. \tag{6.50}$$

For any choice of $\gamma_1 \in \{\alpha_1, \beta\}$, $\gamma_1 \in \{\alpha_1, \alpha_2, \beta\}$ and $m \in \{1, 2\}$, this determinant is a polynomial in $\beta \in (0, Z_+(\mathcal{E}, E))$ and as such is nonzero ouside on a set of measure 0. This establishes that for any pair $(E, \mathcal{E})$, $M'_{F(i)}$ for $i \in \{4, \dots\}$ has full rank for almost every value of $\beta \in (0, Z_+(\mathcal{E}, E))$.

The only interactions that were not considered above are the ones corresponding to triples $F(N+1)$ and $F(0)$. A similar computation to the one carried out in the above paragraph, also carried out in [1] and in Appendix C, shows defining M'_0 as M_0 without its first, second, eleventh and twelfth row, we have

$$\det(M'_0) = -96\alpha_1^5\beta^{11}C_{jk}^4 C_{j\ell}^5 C_{k\ell}^3 \text{ and } \det(M'_{F(N+1)}) = \frac{39}{2}\alpha_1^3\beta^4\left(\beta - \alpha_1\right)C_{jk}^4 C_{j\ell} C_{k\ell}\,. \tag{6.51}$$

Again, these determinants are analytic functions of $\beta \in (0, Z_+(\mathcal{E}, E))$ and as such are nonzero except on a set of measure 0. Combining (6.50) and (6.51) we see that the matrix $\mathcal{A}$ has rank $8 + 4 \cdot (n-3) = 4n - 4 = \dim(T\mathcal{X})$. Since upon changing α_1, α_2 this holds for any compatible choice of the conserved quantities $(E, \mathcal{E})$ we have shown the desired result. □

6.3 Related work

It is natural to compare our work with [10, 11] which consider the related Lorenz-96 and Galerkin approximated 2d Euler models with Brownian forcing and balancing dissipation. Here we work directly on the conservative equations rather than removing dissipation and stochastic forcing through a limiting procedure to approach the conservative dynamics. We expect this direct approach to allow us to say more about the conservative dynamics. Here the randomness is injected through the random splitting and used mainly to make the dynamics generic. We also expect this separation of the different roles of the forcing to be useful.

Our choice of how to inject randomness leads to a more “elliptic” dynamics in that noise directly affects most of the model’s building blocks, though the mechanism is arguably less disruptive than elliptic additive Brownian forcing. While possibly more disruptive than the “minimally” hypoelliptic[4] forcing considered in [11, 37], the random splitting dynamics is likely more analytically tractable. For example, here we are able to directly verify that the lifted dynamics on the projective space satisfies the Lie bracket condition while in [11] (impressive) computer assisted algebra is needed to verify the conditions. We are hopeful that this class of models will lead to analyses in directions currently unfeasible for the more ubiquitous models with Brownian forcing.

[4] The terms *elliptic* and *hypoelliptic* used here are analogous to the identical terms used in the partial differential equations literature.

7

Nonconservative systems

In this chapter we add dissipation and fixed body forcing to both conservative Lorenz-96 and Galerkin approximations of 2d Euler by introducing a new vector field

$$V_0(x) = -\nu\Lambda x + F \tag{7.1}$$

to the splittings constructed in Sections 6.1 and 6.2, where $\nu > 0$, F is a fixed nonzero vector with nonnegative entries, and Λ a linear operator satisfying

$$\Lambda x \cdot x \geqslant \alpha \|x\|^2 \tag{7.2}$$

for some $\alpha > 0$. For the remainder of this chapter we consider random splittings associated to families of complete, smooth vector fields $\mathcal{V} = \{V_j\}_{j=0}^n$ on $\mathbb{R}^d$ satisfying

Assumption 2. *V_0 is as in* (7.1) *with Λ satisfying* (7.2) *and the flows of the other V_j conserve Euclidean norm.*

Fix $h > 0$ and let P_h be the transition kernel of a random splitting satisfying Assumption 2. When Λ is the identity, the addition of V_0 to the splitting of conservative Lorenz-96 gives a splitting of the Lorenz-96 model introduced in Section 1.3, while for 2d Euler the resulting V_0 corresponds to a friction or drag term sometimes called

Ekman damping. When Λ is diagonal with diagonal entry $|j|^2$ in the positions associated to[1] a_j and b_j, which corresponds to the Laplacian in Fourier space, the addition of V_0 to the splitting of 2d Euler gives a splitting of 2d Navier-Stokes.

Note that the dissipative part of V_0 in (7.1) depends linearly on x whereas the forcing is constant. Thus dissipation dominates forcing for sufficiently large x and, since the remaining vector fields are conservative, the splitting dynamics cannot grow too large. Specifically, letting $\Phi_{h\tau}$ be as before but with the flow $\varphi^{(0)}$ of V_0 appended to the beginning of each cycle, i.e. $\Phi_{h\tau} = \varphi^{(n)}_{h\tau_n} \circ \cdots \circ \varphi^{(0)}_{h\tau_0}$, we have

Lemma 7.1. *Under Assumption 2 for any initial x and $m > 0$,*

$$\|\Phi^m_{h\tau}(x)\|^2 \leqslant \|x\|^2 e^{-\nu\alpha h \sum_{k=0}^m \tau_{k(n+1)}} + \frac{1}{\nu^2\alpha^2}\|F\|^2 \left(1 - e^{-\nu\alpha h \sum_{k=0}^m \tau_{k(n+1)}}\right). \tag{7.3}$$

Proof. Letting $\varphi = \varphi^{(0)}$, we have

$$\partial_t \|\varphi_t\|^2 = 2\langle F, \varphi_t\rangle - 2\nu\langle \Lambda\varphi_t, \varphi_t\rangle \leqslant \frac{1}{\nu\alpha}\|F\|^2 + \nu\alpha\|\varphi_t\|^2 - 2\nu\alpha\|\varphi_t\|^2$$

$$= \frac{1}{\nu\alpha}\|F\|^2 - \nu\alpha\|\varphi_t\|^2,$$

where the inequality follows from (7.2) and $2\langle F, \varphi_t\rangle \leqslant (\nu\alpha)^{-1}\|F\|^2 + \nu\alpha\|\varphi_t\|^2$. Solving

$$\dot{y} = \frac{1}{\nu\alpha}\|F\|^2 - \nu\alpha y$$

from $y(0) = \|x\|$ together with the comparison theorem for ODEs [55] then gives

$$\|\varphi_t(x)\|^2 \leqslant \|x\|^2 e^{-\nu\alpha t} + \frac{1}{\nu^2\alpha^2}\|F\|^2 \left(1 - e^{-\nu\alpha t}\right)$$

for all time. Furthermore, since $\varphi^{(k)}$ conserves norm for $1 \leqslant k \leqslant n$, the above implies

$$\|\Phi_{h\tau}(x)\|^2 = \|\varphi^{(n)}_{h\tau_n} \circ \cdots \circ \varphi^{(0)}_{h\tau_0}(x)\|^2 = \|\varphi^{(0)}_{h\tau_0}(x)\|^2$$

$$\leqslant \|x\|^2 e^{-\nu\alpha\tau_0} + \frac{1}{\nu^2\alpha^2}\|F\|^2 \left(1 - e^{-\nu\alpha\tau_0}\right).$$

The result follows by straightforward induction on the number of cycles, m. ☐

[1] Recall that for each index $j \in \mathbb{Z}^2_N$, we have two real coordinates a_j and b_j.

Remark 7.2. *The convergence results of Chapter 4 do not directly apply to Lorenz-96 and Galerkin approximations of 2d Navier-Stokes since $\mathcal{V}$-orbits are generally unbounded in both models. However, Lemma 7.1 implies that any splitting starting from x whose vector fields satisfy Assumption 2 will lie inside the ball of radius $\|x\|^2 + (\nu\alpha)^{-2}\|F\|^2$ centered at the origin for all nonnegative times. In particular, since the splitting vector fields are smooth, a bound analogous to* (4.2) *holds for all x in the ball $B_r(0)$ of radius r centered at the origin in the ambient Euclidean space. Thus all convergence results of Chapter 4 hold for these random splittings when $\mathcal{C}^k(\mathcal{X})$ is replaced by $\mathcal{C}^k_r(\mathcal{X})$, the space of k-times continuously differentiable functions that vanish outside $B_r(0)$. Intuitively, this says that for any initial condition x, the trajectories of a random splitting satisfying Assumption 2 will converge on average and almost surely as $h \to 0$ to the trajectory of the true dynamics starting from x.*

Corollary 7.3. *The Euclidean norm is a Lyapunov function for P_h. That is, there exist constants $K \geqslant 0$ and $\gamma \in (0,1)$ such that for all $x \in \mathbb{R}^d$,*

$$(P_h\|\cdot\|)(x) \leqslant \gamma\|x\| + K.$$

Proof. By Lemma 7.1, specifically $\|\Phi_{ht}(x)\| \leqslant \|x\|e^{-\frac{1}{2}\nu\alpha t_0} + (\nu\alpha)^{-1}\|F\|$, we have

$$(P_h\|\cdot\|)(x) = \int_{\mathbb{R}^{n+1}_+} \|\Phi_{ht}(x)\|e^{-\sum t_k}dt \leqslant \frac{1}{1+\frac{1}{2}\nu\alpha h}\|x\| + \frac{1}{\nu\alpha}\|F\|$$

for any x. The result follows with $K = (\nu\alpha)^{-1}\|F\|$ and $\gamma = (1 + \frac{1}{2}\nu\alpha h)^{-1}$. □

7.1 Ergodicity

We now present a variation of Theorem 3.1, namely Theorem 7.4, which simplifies verification of ergodicity in the present setting. Recall from Sections 6.1 and 6.2 that one of the difficulties in verifying Theorem 3.1 was proving controllability, i.e. the existence of a distinguished point x_* that could be reached by the splitting dynamics in finite time from any other point. With the addition of dissipation, the fixed point $\nu^{-1}\Lambda^{-1}F$ of $\dot{x} = V_0(x)$ is a natural candidate for x_* and, as we will see, the fact that it is globally attracting obviates several technicalities associated with controllability in the conservative cases discussed above.

Theorem 7.4. *Suppose Assumption 2 holds and set $x_* = \nu^{-1}\Lambda^{-1}F$. If there exist $m \geqslant 0$ and t in $\mathbb{R}_+^{mn}$ such that the Lie bracket condition holds at $\widetilde{x} := \Phi_{ht}^m(x_*)$, then P_h has a unique invariant measure μ for all $h > 0$. Furthermore, there exist $C > 0$ and γ in $(0, 1)$ such that for all x in $\mathbb{R}^d$,*

$$\|P_h^m(x, \cdot) - \mu\| \leqslant C\gamma^m \tag{7.4}$$

where $\|\cdot\|$ is the norm on probability measures induced by the weighted supremum norm $\|f\| := \sup_x |f(x)|/(1 + \|x\|)$ on bounded measurable functions $f : \mathbb{R}^d \to \mathbb{R}$.

The proof of Theorem 7.4 uses the following lemmas. The first, due to Krylov-Bogolubov, is a standard result from the theory of Markov processes [36]. The second, which follows from Lemma 3.2 and Theorem 2.5, is from [12, Theorem 4.4]. For the statement of Lemma 7.5, recall a transition kernel P on $\mathbb{R}^d$ is *Feller* if Pf is continuous whenever $f : \mathbb{R}^d \to \mathbb{R}$ is continuous and bounded. Also, a sequence of probability measures $\{\mu_m\}$ on $\mathbb{R}^d$ is *tight* if for every $\varepsilon > 0$ there exists a compact subset K of $\mathbb{R}^d$ such that $\mu_m(K) \geqslant 1 - \varepsilon$ for all m.

Lemma 7.5. *Let P be a Feller probability transition kernel on $\mathbb{R}^d$. If there exists x in $\mathbb{R}^d$ such that $\{P^m(x, \cdot)\}_{m=0}^{\infty}$ is tight, then P has an invariant probability measure.*

Lemma 7.6. *Suppose $\Phi_{ht}^m(x) = \widetilde{x}$ and the Lie bracket condition holds at $\widetilde{x}$. Then there exists a $c > 0$, an $\widetilde{m}$, and neighborhoods U_x of x and $\widetilde{U}$ of $\widetilde{x}$ such that for all y in U_x and B in $\mathcal{B}(\mathcal{X})$,*

$$P_h^{\widetilde{m}}(y, B) \geqslant c\lambda\left(B \cap \widetilde{U}\right).$$

The following proof is another instance of the rather classical idea, dating at least to the split chains of Nummelin [66] and work of Meyn and Tweedie [57], that the existence of a globally accessible point at which the dynamics is continuous in the right sense implies the transition densities converge to a unique equilibrium measure. If the return to the globally accessible point has finite expectation, then mixing is exponential. The same basic structure of the SDE version of our system was leveraged in [26] to prove exponential mixing (see also [54]). In the closely related PDMP setting, analogous results are found in [47] in a specific example and [13] in a more general context.

Proof of Theorem 7.4. We first prove existence. Continuity of Φ_{ht} immediately implies P_h is Feller. Furthermore, Lemma 7.1 implies that random splitting starting from any x is constrained to lie in a compact subset of $\mathbb{R}^d$, namely the closed ball of radius $\|x\|^2 + (\nu\alpha)^{-2}\|F\|^2$ centered at the origin. Thus, for any x, the sequence $\{P_h^m(x,\cdot)\}_{m=0}^{\infty}$ is tight and existence follows from Lemma 7.5.

Next we prove uniqueness. The hypothesis and Lemma 7.6 together imply the existence of $c > 0$, $\widetilde{m}$, and neighborhoods U_* of x_* and $\widetilde{U}$ of $\widetilde{x}$ such that

$$P_h^{\widetilde{m}}(x, B) \geqslant c\lambda\left(B \cap \widetilde{U}\right) \tag{7.5}$$

for all $x \in U_*$ and Borel sets B. Also, positive-definiteness of Λ implies

$$\|\varphi_t^{(0)}(x) - x_*\| \leqslant e^{-\alpha t}\|x - x_*\|$$

for any $x \in \mathbb{R}^d$ and $t \geqslant 0$. In particular, for any open ball B_r of radius r centered at the origin, there exists $T_0 > 0$ such that $\varphi_{ht}^{(0)}(B_r)$ is properly contained in U_* whenever $ht > T_0$. And since $\varphi_{ht}^{(0)}(B_r)$ is properly contained in U_* and the $\varphi^{(k)}$ are continuous, there exist $T_k > 0$ such that $\Phi_{ht} = \varphi_{ht_n}^{(n)} \circ \cdots \circ \varphi_{ht_0}^{(0)}(x) \in U_*$ for all $x \in B_r$ and $ht_k \in (0, T_k)$. So, for any $x \in B_r$,

$$P_h(x, U_*) \geqslant \int_0^{T_n} \cdots \int_0^{T_1} \int_{T_0}^{\infty} \mathbb{1}_{U_*}\left(\Phi_{ht}(x)\right) e^{-\sum t_k} dt = \frac{1}{T_0} \prod_{k=1}^{n} \left(1 - e^{-T_k}\right) > 0$$

and hence $\inf_{x \in B_r} P_h(x, U_*) > 0$.

As in the proof of Theorem 3.1, suppose toward a contradiction that μ_1 and μ_2 are distinct P_h-ergodic probability measures and that A_1 and A_2 are disjoint measurable sets partitioning $\mathbb{R}^d$ with $\mu_i(B) = \mu_i(B \cap A_i)$ for all Borel sets B. Fix x_i in the support of μ_i, let r be sufficiently large that $x_1, x_2 \in B_r$, and set $\kappa := \inf_{x \in B_r} P_h(x, U_*) > 0$. Then by (7.5) for any Borel set B,

$$\begin{aligned} \mu_i(B) = \mu_i P_h^{\tilde{m}+1}(B) &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} P_h^{\tilde{m}}(y, B) P_h(x, dy) \mu_i(dx) \\ &\geqslant \int_{B_r}\int_{U_*} P_h^{\tilde{m}}(y, B) P_h(x, dy) \mu_i(dx) \geqslant \kappa c \lambda\left(B \cap \widetilde{U}\right) \mu_i\left(B_r\right). \end{aligned} \tag{7.6}$$

In particular, $\mu_i(B) = 0$ implies $\lambda(B \cap \widetilde{U}) = 0$ since c, κ, and $\mu_i(B_r)$ are all strictly positive (the latter because B_r is an open set containing both x_1 and x_2 which were chosen to be in the supports of μ_1 and μ_2, respectively). But $\mu_1(A_2 \cap \widetilde{U}) = \mu_2(A_1 \cap \widetilde{U}) = 0$ and so we obtain the contradiction

$$0 < \lambda\left(\widetilde{U}\right) = \lambda\left(A_1 \cap \widetilde{U}\right) + \lambda\left(A_2 \cap \widetilde{U}\right) = 0,$$

which concludes the proof of uniqueness.

Finally, for the exponential convergence statement (7.4), we have from (7.6) that for any $r > 0$,

$$\inf_{x \in B_r} P_h^{\tilde{m}+1}(x, B) \geqslant \kappa c \lambda\left(B \cap \widetilde{U}\right)$$

for all Borel sets B. That is, the transition probabilities $P_h^{\tilde{m}+1}(x, \cdot)$ are minorized uniformly over B_r by the probability measure $\widetilde{\lambda} := \lambda(\widetilde{U})^{-1}\lambda(\cdot \cap \widetilde{U})$. Exponential convergence then follows from Corollary 7.3 upon taking $r > 2K/(1-\gamma)$. See for example Theorem 1.2 in [38]. □

Corollary 7.7. *Consider the random splitting of Lorenz-96 associated to the vector fields* $\{V_k\}_{k=0}^n$*, where* $V_0(x) = -\nu x + F$ *and* $\{V_k\}_{k=1}^n$ *are the splitting vector fields* *of conservative Lorenz-96 from Section 6.1. If $x_* := -\nu F$ is not a fixed point of conservative Lorenz-96, i.e. $\nu^2 \sum_{k=1}^n (F_k^2 + F_{k+1}^2)F_{k-1}^2 \neq 0$, then the random splitting has a unique invariant measure on $\mathbb{R}^n$ and the dynamics converge to this measure at an exponential rate in the sense of* (7.4).

Proof. The determinant of the n-by-n matrix

$$\begin{pmatrix} | & | & & | \\ V_0(x) & V_1(x) & \cdots & V_{n-1}(x) \\ | & | & & | \end{pmatrix} = \begin{pmatrix} -\nu x_1 + F_1 & x_2 x_n & 0 & \dots & 0 \\ -\nu x_2 + F_2 & -x_1 x_n & x_3 x_1 & \dots & 0 \\ -\nu x_3 + F_3 & 0 & -x_2 x_1 & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & x_n x_{n-2} \\ -\nu x_n + F_n & 0 & 0 & & -x_{n-1} x_{n-2} \end{pmatrix}$$

is

$$x_1 x_{n-1} x_n \left(\prod_{x=2}^{n-2} x_k^2 \right) \left(\nu^2 \|x\|^2 - \langle F, x \rangle \right).$$

So the $\{V_k\}_{k=0}^n$ span $\mathbb{R}^n$ at every x with nonzero coordinates and satisfying $\nu^2 \|x\|^2 \neq \langle F, x \rangle$. In particular, since x_* is not a fixed point of conservative Lorenz-96, we showed in the proof of Proposition 6.2 that x_* can be moved via the splitting dynamics to some $\widetilde{x}$ with nonzero coordinates. Finally, by rotating slightly more on the last step if necessary, we can also guarantee $\nu^2 \|\widetilde{x}\|^2 \neq \langle F, \widetilde{x} \rangle$. Thus the Lie bracket condition holds at $\widetilde{x}$ and the result follows by Theorem 7.4. □

Corollary 7.8. *Fix $N \geqslant 2$ and set $n = 4N(N+1)$. Consider the random splitting of the Nth Galerkin approximation of 2d Navier-Stokes associated to $\{V_k\}_{k=0}^n$, where $V_0(x) = -\nu \Lambda x + F$ with Λ the n-by-n diagonal matrix corresponding to the Laplacian discussed at the beginning of this section, and $\{V_k\}_{k=1}^n$ the splitting vector fields of 2d Euler from Section 6.2. If F is nondegenerate in the sense of Definition 6.7, then the random splitting has a unique, and hence ergodic, invariant measure and the dynamics converge to this measure at an exponential rate in the sense of* (7.4).

Proof. Recall in this case $V_0(x) = -\nu\Lambda x + F$ where Λ is the diagonal matrix with diagonal entry $|k|^2$ in the slots corresponding to the coordinates a_k and b_k. Fix $j, k, \ell \in \mathbb{Z}_N^2$ with $j + k - \ell = 0$ and let W be one of the vector fields $V_{a_j a_k a_\ell}$, $V_{a_j b_k b_\ell}$, $V_{b_j a_k b_\ell}$, or $V_{b_j b_k a_\ell}$. Letting e.g. $(x_j, x_k, x_\ell) = (a_j, a_k, a_\ell)$ when $W = V_{a_j a_k a_\ell}$ and similarly for the other cases, direct computation yields

$$\begin{aligned}
[V_0, W]_j(x) &= C_{k\ell}\left(F_k x_\ell + F_\ell x_k + \nu(|j|^2 - |k|^2 - |\ell|^2) x_k x_\ell\right), \\
[V_0, W]_k(x) &= C_{j\ell}\left(F_j x_\ell + F_\ell x_j + \nu(|k|^2 - |j|^2 - |\ell|^2) x_j x_\ell\right), \\
[V_0, W]_\ell(x) &= -C_{jk}\left(F_j x_k + F_k x_j + \nu(|\ell|^2 - |j|^2 - |k|^2) x_j x_k\right),
\end{aligned} \tag{7.7}$$

where $[V_0, W]_j(x)$ is the component of $[V_0, W]$ corresponding to the component x_j of x, and similarly for $[V_0, W]_k$ and $[V_0, W]_\ell$. As in the 2d Euler case, Gaussian elimination shows that the 6-by-6 matrix (see (6.45) for an explicit form of the middle 4 columns)

$$\begin{pmatrix} | & | & | & | & | & | \\ V_0 & V_{a_j a_k a_\ell} & V_{a_j b_k b_\ell} & V_{b_j a_k b_\ell} & V_{b_j b_k a_\ell} & [V_0, W] \\ | & | & | & | & | & | \end{pmatrix} \tag{7.8}$$

is rank 6 at every generic[2] point q in $\mathbb{R}^n$. Thus V_0 and $[V_0, W]$ add two new directions to the splitting vector fields of 2d Euler and by an entirely similar argument to the spanning argument in Section 6.2.3 we have that the Lie bracket condition holds at every such q. Furthermore, since F is nondegenerate the controllability argument of Section 6.2.3 implies x_* can be evolved via the split dynamics to a generic point. The result then follows by Theorem 7.4. □

Remark 7.9. *A very similar argument to the one above proves unique ergodicity for Ekman damping as well, i.e. when* Λ *is the identity matrix on* $\mathbb{R}^n$*. In this case* (7.7)

[2] Recall a *generic point* is one with all coordinates nonzero; see Definition 6.8.

becomes

$$[V_0, W]_j(x) = C_{k\ell}\left(F_k x_\ell + F_\ell x_k - \nu x_k x_\ell\right),$$

$$[V_0, W]_k(x) = C_{j\ell}\left(F_j x_\ell + F_\ell x_j - \nu x_j x_\ell\right),$$

$$[V_0, W]_\ell(x) = -C_{jk}\left(F_j x_k + F_k x_j - \nu x_j x_k\right),$$

and the rest of the argument goes through unchanged.

8

Conclusion

In this dissertation we introduced random splitting and gave conditions for such systems to be ergodic, to converge to their deterministic counterparts, and to be chaotic, i.e. to have a positive top Lyapunov exponent. We also applied these results to random splittings of the conservative Lorenz-96 and Lorenz-96 equations, as well as Galerkin approximations of the Euler and Navier-Stokes equations on the 2-dimensional torus. The method of random splitting invites much future work. A short list of possible directions includes constructing and studying splittings of other models, studying higher-order random splitting schemes, and proving chaos for systems whose sum of Lyapunov exponents λ_Σ is negative.

Appendix A

Convergence lemmas

In this appendix we prove Lemmas 4.2 and 4.5 from Chapter 4.

A.1 Semigroups, norms, and bounds

In this subsection we elaborate on the semigroup framework of Chapter 4. The notation and results are used extensively in the proofs of Lemmas 4.2 and 4.5, which are given in subsections A.2 and A.4, respectively.

Fix a $\mathcal{V}$-orbit $\mathcal{X}$. The $\mathcal{C}^2$ assumption implies the V_k, which act on functions f via $V_k f(x) = Df(x)V_k(x)$, are linear operators from $\mathcal{C}^2(\mathcal{X})$ to $\mathcal{C}^1(\mathcal{X})$ and from $\mathcal{C}^1(\mathcal{X})$ to $\mathcal{C}(\mathcal{X})$. It also implies the semigroups $\{S_t\}_{t\geqslant 0}$ and $\{\widetilde{S}_t^{(k)}\}_{t\geqslant 0}$ defined in (4.3) and (4.4) are linear operators on $\mathcal{C}^k(\mathcal{X})$ for $k \leqslant 2$. Our aim now is to obtain bounds on norms of compositions of these random semigroups. For $i \leqslant j$ define $\Phi_{h\tau}^{(i,j)} := \varphi_{h\tau_j}^{(j)} \circ \cdots \circ \varphi_{h\tau_i}^{(i)}$ and $\widetilde{S}_{h\tau}^{(i,j)} := \widetilde{S}_{h\tau}^{(i)} \cdots \widetilde{S}_{h\tau}^{(j)}$. Note $\widetilde{S}_{h\tau}^{(i,j)}$ acts on functions f via

$$\widetilde{S}_{h\tau}^{(i,j)} f(x) = f\left(\Phi_{h\tau}^{(i,j)}(x)\right) = f\left(\varphi_{h\tau_j}^{(j)} \circ \cdots \circ \varphi_{h\tau_i}^{(i)}(x)\right).$$

So for any $f \in \mathcal{C}(\mathcal{X})$ with $\|f\|_\infty = 1$, we have

$$\|\widetilde{S}_{h\tau}^{(i,j)} f\|_\infty = \|f(\Phi_{h\tau}^{(i,j)})\|_\infty = 1$$

and hence $\|\widetilde{S}^m_{h\tau}\|_{0\to 0} = 1$. Next, let $\varphi = \varphi^{(k)}$ for arbitrary k. Then

$$\varphi_t(x) = x + \int_0^t V(\varphi_s(x))ds$$

and so

$$D\varphi_t(x) = I + \int_0^t DV(\varphi_s(x))D\varphi_s(x)ds$$

and

$$D^2\varphi_t(x) = \int_0^t D^2V(\varphi_s(x))\left(D\varphi_s(x), D\varphi_s(x)\right) + DV(\varphi_s(x))D^2\varphi_s(x)ds.$$

In particular, $\|D\varphi_t(x)\| \leqslant 1 + C_* \int_0^t \|D\varphi_s(x)\|ds$ for all x in $\mathcal{X}$ so by Grönwall's inequality

$$\sup_{x\in\mathcal{X}} \|D\varphi_t(x)\| \leqslant e^{C_* t}, \tag{A.1}$$

where here and throughout C_* is the constant from (4.2) corresponding to $\mathcal{X}$. Similarly, since $\|D^2V\left(D\varphi, D\varphi\right)\| \leqslant \|D^2V\|\|D\varphi\|^2 \leqslant C_*\|D\varphi\|^2$,

$$\|D^2\varphi_t(x)\| \leqslant C_* \int_0^t \|D\varphi_s(x)\|^2 + \|D^2\varphi_s(x)\|ds \leqslant C_* t e^{2C_* t} + C_* \int_0^t \|D^2\varphi_s(x)\|ds$$

and Grönwall's inequality implies

$$\sup_{x\in\mathcal{X}} \|D^2\varphi_t(x)\| \leqslant C_* t e^{3C_* t}. \tag{A.2}$$

Note (A.1) and (A.2) hold uniformly over all $\varphi^{(k)}$. Thus, for $f \in C^1(\mathcal{X})$ with $\|f\|_1 = 1$,

$$\left\| D\left(\widetilde{S}^{(i,j)}_{h\tau} f\right)\right\| = \left\| Df\left(\Phi^{(i,j)}_{h\tau}\right) D\Phi^{(i,j)}_{h\tau}\right\| \leqslant \prod_{k=i}^{j} \|D\varphi^{(k)}_{h\tau_k}\| \leqslant e^{C_* h \sum_{k=i}^{j} \tau_k},$$

where the first inequality follows from submultiplicity and the second from (A.1). Similarly,

$$D^2\Phi^{(i,j)}_{h\tau} = \sum_{k=i}^{j} D\varphi^{(j)}_{h\tau_j} \cdots D\varphi^{(k+1)}_{h\tau_{k+1}} D^2\varphi^{(k)}_{h\tau_k}\left(D\Phi^{(i,k-1)}_{h\tau}, D\Phi^{(i,k-1)}_{h\tau}\right)$$

together with (A.1) and (A.2) gives

$$\left\| D^2 \Phi_{h\tau}^{(i,j)} \right\| \leqslant \sum_{k=i}^{j} \left\| D\varphi_{h\tau_j}^{(j)} \right\| \cdots \left\| D\varphi_{h\tau_{k+1}}^{(k+1)} \right\| \left\| D^2 \varphi^{(k)} \right\| \left\| D\Phi_{h\tau}^{(i,k-1)} \right\|^2$$

$$\leqslant C_* \sum_{k=i}^{j} h\tau_k e^{C_* h \sum_{k+1}^{j} \tau_\ell} e^{3C_* h\tau_k} e^{2C_* h \sum_1^{k-1} \tau_\ell} \leqslant C_* h e^{3C_* h \sum_{k=i}^{j} \tau_k} \sum_{k=i}^{j} \tau_k.$$

Therefore

$$\left\| D^2 \left(\widetilde{S}_{h\tau}^{(i,j)} f \right) \right\| = \left\| D^2 f \left(\Phi_{h\tau}^{(i,j)} \right) \left(D\Phi_{h\tau}^{(i,j)}, D\Phi_{h\tau}^{(i,j)} \right) + Df \left(\Phi_{h\tau}^{(i,j)} \right) D^2 \Phi_{h\tau}^{(i,j)} \right\|$$

$$\leqslant \left\| D\Phi_{h\tau}^{(i,j)} \right\|^2 + \left\| D^2 \Phi_{h\tau}^{(i,j)} \right\| \leqslant e^{2C_* h \sum_{k=i}^{j} \tau_k} + \left\| D^2 \Phi_{h\tau}^{(i,j)} \right\|$$

$$\leqslant e^{2C_* h \sum_{k=i}^{j} \tau_k} + C_* h e^{3C_* h \sum_{k=i}^{j} \tau_k} \sum_{k=i}^{j} \tau_k$$

$$\leqslant \left(1 + C_* h \sum_{k=i}^{j} \tau_k \right) e^{3C_* h \sum_{k=i}^{j} \tau_k}.$$

The above computations prove

Lemma A.1. *For any $h > 0$ and $i \leqslant j$, we have $\|\widetilde{S}_{h\tau}^{(i,j)}\|_{0\to 0} = 1$ as well as*

$$\left\| \widetilde{S}_{h\tau}^{(i,j)} \right\|_{1\to 1} \leqslant e^{C_* h \sum_{k=i}^{j} \tau_k} \quad \text{and} \quad \left\| \widetilde{S}_{h\tau}^{(i,j)} \right\|_{2\to 2} \leqslant \left(1 + C_* h \sum_{k=i}^{j} \tau_k \right) e^{3C_* h \sum_{k=i}^{j} \tau_k}.$$

In particular, $\left\| \widetilde{S}_{h\tau}^{(i,j)} \right\|_{\ell\to\ell} \leqslant \left(1 + C_* h \sum_{k=i}^{j} \tau_k \right) e^{3C_* h \sum_{k=i}^{j} \tau_k}$ for all $\ell \leqslant 2$.

Note that under the $\mathcal{C}^2$ assumption $\widetilde{S}_{h\tau}^{(i,j)}$ can also be regarded as a linear operator from $\mathcal{C}^2(\mathcal{X})$ to $\mathcal{C}^1(\mathcal{X})$. So since $\{f \in \mathcal{C}^2(\mathcal{X}) : \|f\|_2 = 1\}$ is a subset of $\{f \in \mathcal{C}^1(\mathcal{X}) : \|f\|_1 = 1\}$, we have

$$\left\| \widetilde{S}_{h\tau}^{(i,j)} \right\|_{2\to 1} = \sup_{\|f\|_2 = 1} \left\| \widetilde{S}_{h\tau}^{(i,j)} f \right\|_1 \leqslant \sup_{\|f\|_1 = 1} \left\| \widetilde{S}_{h\tau}^{(i,j)} f \right\|_1 = \left\| \widetilde{S}_{h\tau}^{(i,j)} \right\|_{1\to 1} \leqslant e^{C_* h \sum_{k=i}^{j} \tau_k}. \quad \text{(A.3)}$$

We also have the following corollary of Lemma A.1.

Corollary A.2. *Fix $i \leqslant j$ and set $m := j - i + 1$. For all $\ell \leqslant 2$ and polynomials $p : \mathbb{R}_+^m \to \mathbb{R}$ there exists $h_* > 0$ such that for all $h < h_*$,*

$$\mathbb{E}\|p(\tau_i, \ldots, \tau_j)\widetilde{S}_{h\tau}^{(i,j)}\|_{k\to k} < \infty. \tag{A.4}$$

Proof. Writing $t = (t_i, \ldots, t_j)$ and $dt = dt_i \cdots dt_j$, we have

$$\begin{aligned}\mathbb{E}\|p(\tau_i, \ldots, \tau_j)\widetilde{S}_{h\tau}^{(i,j)}\|_{\ell\to\ell} &= \int_{\mathbb{R}_+^m} |p(t)| \left\|\widetilde{S}_{ht}^{(i,j)}\right\|_{\ell\to\ell} e^{-\sum t_k} dt \\ &\leqslant \int_{\mathbb{R}_+^m} |p(t)| \left(1 + C_* h \sum_{k=i}^{j} t_k\right) e^{(3C_* h - 1)\sum_{k=i}^{j} t_k} dt\end{aligned}$$

which is finite for all $h < h_* := (3C_*)^{-1}$. □

A.2 Proof of Lemma 4.2

We highlight the steps of the proof with italicized font.

Variation of constants. We begin by differentiating $\widetilde{S}_{h\tau}$ in h:

$$\partial_h \widetilde{S}_{h\tau} = \sum_{k=1}^{n} \tau_k e^{h\tau_1} \cdots e^{h\tau_{k-1}} V_k e^{h\tau_k} \cdots e^{h\tau_n} = \sum_{k=1}^{n} \tau_k \widetilde{S}_{h\tau}^{(1,k-1)} V_k \widetilde{S}_{h\tau}^{(k,n)}.$$

Next, commute $\widetilde{S}_{h\tau}^{(1,k-1)}$ and V_k via $[\widetilde{S}_{h\tau}^{(1,k-1)}, V_k] := \widetilde{S}_{h\tau}^{(1,k-1)} V_k - V_k \widetilde{S}_{h\tau}^{(1,k-1)}$ to get

$$\partial_h \widetilde{S}_{h\tau} = \sum_{k=1}^{n} \tau_k V_k \widetilde{S}_{h\tau} + \sum_{k=1}^{n} \tau_k [\widetilde{S}_{h\tau}^{(1,k-1)}, V_k] \widetilde{S}_{h\tau}^{(k,n)} = V \widetilde{S}_{h\tau} + (V_\tau - V)\widetilde{S}_{h\tau} + E_{h\tau}$$

where $V_\tau := \sum_{k=1}^{n} \tau_k V_k$ and $E_{h\tau} := \sum_{k=1}^{n} \tau_k [\widetilde{S}_{h\tau}^{(1,k-1)}, V_k] \widetilde{S}_{h\tau}^{(k,n)}$. So, by variation of constants,

$$\widetilde{S}_{h\tau} - S_h = \int_0^h S_{h-r}(V_\tau - V)\widetilde{S}_{r\tau} dr + \int_0^h S_{h-r} E_{r\tau} dr. \tag{A.5}$$

Call $S_{h-r}(V_\tau - V)\widetilde{S}_{r\tau}$ *error term 1* and $S_{h-r}E_{r\tau}$ *error term 2.* These terms will be treated separately in what follows. First however, we invoke variation of constants again to get an expression for $[\widetilde{S}^{(1,k-1)}_{r\tau}, V_k]$ that will be used to control error term 2. Differentiating in r gives

$$\partial_r[\widetilde{S}^{(1,k-1)}_{r\tau}, V_k] = \sum_{j=1}^{k-1} \tau_j [\widetilde{S}^{(1,j-1)}_{r\tau} V_j \widetilde{S}^{(j,k-1)}_{r\tau}, V_k]$$

$$= \sum_{j=1}^{k-1} \tau_j \Big([V_j \widetilde{S}^{(1,k-1)}_{r\tau}, V_k] + \big[[\widetilde{S}^{(1,j-1)}_{r\tau}, V_j]\widetilde{S}^{(j,k-1)}_{r\tau}, V_k\big] \Big)$$

$$= \sum_{j=1}^{k-1} \tau_j V_j [\widetilde{S}^{(1,k-1)}_{r\tau}, V_k] + \sum_{j=1}^{k-1} \tau_j \Big([V_j, V_k]\widetilde{S}^{(1,k-1)}_{r\tau} + \big[[\widetilde{S}^{(1,j-1)}_{r\tau}, V_j]\widetilde{S}^{(j,k-1)}_{r\tau}, V_k\big] \Big).$$

The second equality follows from commuting $\widetilde{S}^{(1,j-1)}_{h\tau}$ and V_j as before, and the third follows from the identity $[XY, Z] = X[Y, Z] + [X, Z]Y$. So, by variation of constants,

$$\begin{aligned}[\widetilde{S}^{(1,k-1)}_{r\tau}, V_k] = \sum_{j=1}^{k-1} \int_0^r \tau_j e^{(r-s)\sum_{j=1}^{k-1}\tau_j V_j} [V_j, V_k] \widetilde{S}^{(1,k-1)}_{s\tau} ds \\ + \sum_{j=1}^{k-1} \int_0^r \tau_j e^{(r-s)\sum_{j=1}^{k-1}\tau_j V_j} \big[[\widetilde{S}^{(1,j-1)}_{s\tau}, V_j]\widetilde{S}^{(j,k-1)}_{s\tau}, V_k\big] ds.\end{aligned} \tag{A.6}$$

Note $\|e^{(r-s)\sum_{j=1}^{k-1}\tau_j V_j}\|_{0\to 0} = 1$. So, by Corollary A.2 the integrands above satisfy

$$\mathbb{E}\|\tau_j e^{(r-s)\sum_{j=1}^{k-1}\tau_j V_j}[V_j, V_k]\widetilde{S}^{(1,k-1)}_{s\tau}\|_{2\to 0} \leqslant \|[V_j, V_k]\|_{2\to 0} \mathbb{E}\|\tau_j \widetilde{S}^{(1,k-1)}_{s\tau}\|_{2\to 2} < C$$

and, similarly,

$$\mathbb{E}\big\|\tau_j e^{(r-s)\sum_{j=1}^{k-1}\tau_j V_j}\big[[\widetilde{S}^{(1,j-1)}_{s\tau}, V_j]\widetilde{S}^{(j,k-1)}_{s\tau}, V_k\big]\big\|_{2\to 0} < C$$

for some C. Therefore

$$\mathbb{E}\|[\widetilde{S}^{(1,k-1)}_{r\tau}, V_k]\|_{2\to 0} \leqslant 2\sum_{j=1}^{k-1} \int_0^r C ds \leqslant Cr \tag{A.7}$$

for some new constant C (we will often absorb arbitrary constants into existing ones).

Error term 1. Rewrite error term 1 as

$$\begin{aligned} S_{h-r}(V_\tau - V)\widetilde{S}_{r\tau} &= \sum_{k=1}^{n}(\tau_k - 1)S_{h-r}V_k\widetilde{S}_{r\tau} \\ &= \sum_{k=1}^{n}(\tau_k - 1)S_{h-r}V_k\widetilde{S}^{(1,k-1)}_{r\tau}\widetilde{S}^{(k+1,n)}_{r\tau} \\ &\quad + \sum_{k=1}^{n}(\tau_k - 1)S_{h-r}V_k\widetilde{S}^{(1,k-1)}_{r\tau}(e^{r\tau_k V_k} - I)\widetilde{S}^{(k+1,n)}_{r\tau} \\ &=: \mathcal{A}_1 + \mathcal{A}_2 \end{aligned} \tag{A.8}$$

where $\mathcal{A}_1$ and $\mathcal{A}_2$ are the first and second sums in the preceding expression. The second equality is obtained by adding and subtracting the identity I as follows: $\widetilde{S}_{r\tau} = \widetilde{S}^{(1,k-1)}_{r\tau}\big(e^{r\tau_k V_k} - I + I\big)\widetilde{S}^{(k+1,n)}_{r\tau} = \widetilde{S}^{(1,k-1)}_{r\tau}\widetilde{S}^{(k+1,n)}_{r\tau} + \widetilde{S}^{(1,k-1)}_{r\tau}(e^{r\tau_k V_k} - I)\widetilde{S}^{(k+1,n)}_{r\tau}$.

Notice $\widetilde{S}^{(1,k-1)}_{r\tau}\widetilde{S}^{(k+1,n)}_{r\tau}$ does not depend on τ_k. So, since the τ_i are independent with mean 1,

$$\mathbb{E}(\mathcal{A}_1) = \sum_{k=1}^{n} S_{t-r}V_k\mathbb{E}(\tau_k - 1)\mathbb{E}\big(\widetilde{S}^{(1,k-1)}_{r\tau}\widetilde{S}^{(k+1,n)}_{r\tau}\big) = 0. \tag{A.9}$$

For the second sum, Taylor expanding $r \mapsto e^{r\tau_k V_k}$ about $r = 0$ with remainder gives

$$e^{r\tau_k V_k} - I = r\tau_k V_k e^{r_* \tau_k V_k}$$

for some $r_* \in [0, r]$. Therefore

$$\mathcal{A}_2 = r\sum_{k=1}^{n}\tau_k(\tau_k - 1)S_{h-r}V_k\widetilde{S}^{(1,k-1)}_{r\tau}V_k e^{r_*\tau_k V_k}\widetilde{S}^{(k+1,n)}_{r\tau}$$

and by Lemma A.1 and Corollary A.2,

$$\|\mathbb{E}(\mathcal{A}_2)\|_{2\to 0} \leqslant Cr\sum_{k=1}^{n}\mathbb{E}\|\widetilde{S}^{(1,k-1)}_{r\tau}\|_{1\to 1}\mathbb{E}\|\tau_k(\tau_k - 1)\widetilde{S}^{(k,n)}_{r\tau}\|_{2\to 2} \leqslant Cr \tag{A.10}$$

for some $C > 0$. Combining Equations (A.8), (A.9), and (A.10) gives

$$\|\mathbb{E}(S_{h-r}(V_\tau - V)\widetilde{S}_{r\tau})\|_{2\to 0} \leqslant Cr. \tag{A.11}$$

Error term 2. Recall error term 2 is $S_{h-r}E_{r\tau} := \sum_{k=1}^{n} \tau_k S_{h-r}[\widetilde{S}_{r\tau}^{(1,k-1)}, V_k]\widetilde{S}_{r\tau}^{(k,n)}$. Hence

$$\|S_{h-r}E_{r\tau}\|_{2\to 0} \leqslant \sum_{k=1}^{n} \tau_k \|S_{h-r}\|_{0\to 0} \|[\widetilde{S}_{r\tau}^{(1,k-1)}, V_k]\|_{2\to 0} \|\tau_k \widetilde{S}_{r\tau}^{(k,n)}\|_{2\to 2} .$$

Note $[\widetilde{S}_{r\tau}^{(1,k-1)}, V_k]$ is independent of τ_k. So, by (A.7) Corollary A.2,

$$\|\mathbb{E}(S_{h-r}E_{r\tau})\|_{2\to 0} \leqslant Cr \tag{A.12}$$

for some $C > 0$.

Final step. Combining (A.5), (A.11), and (A.12) and absorbing constants into C, we have

$$\begin{aligned} \|P_h - S_h\|_{2\to 0} &= \|\mathbb{E}(\widetilde{S}_{h\tau} - S_h)\|_{2\to 0} \\ &\leqslant \int_0^h \|\mathbb{E}(S_{h-r}(V_\tau - V)\widetilde{S}_{r\tau})\|_{2\to 0} dr + \int_0^h \|\mathbb{E}(S_{h-r}E_{r\tau})\|_{2\to 0} dr \\ &\leqslant C \int_0^h r dr = \tfrac{1}{2} C h^2. \end{aligned}$$

□

A.3 Concentration of exponentials

The proof of Lemma 4.5 will itself use two lemmas.

Lemma A.3. *Let $\{\tau_k\}_{k=1}^{\infty}$ be iid exponential with mean 1. For any $m \in \mathbb{N}$, $K > 0$ and $\beta > 1$,*

$$\mathbb{P}\left(\sum_{k=1}^{m} \tau_k > K m^\beta\right) \leqslant 2^m e^{-\frac{1}{2} K m^\beta}. \tag{A.13}$$

Proof. Note if $\tau \sim \mathrm{Exp}(1)$ then $\mathbb{E}(e^{\tau/2}) = 2$. So, by Markov's inequality and independence,

$$\mathbb{P}\left(\sum_{k=1}^{m} \tau_k > K m^\beta\right) = \mathbb{P}\left(e^{\frac{1}{2}\sum_{k=1}^{m} \tau_k} > e^{\frac{1}{2} K m^\beta}\right) \leqslant e^{-\frac{1}{2} K m^\beta} \left(\mathbb{E}\left[e^{\frac{1}{2}\tau}\right]\right)^m = 2^m e^{-\frac{1}{2} K m^\beta}. \square$$

Lemma A.4. *Let $\{\tau_k\}_{k=1}^{\infty}$ be iid exponential with mean 1. For any $m \in \mathbb{N}$ and $K \in (0,1)$,*

$$\mathbb{P}\left(\left|\sum_{k=1}^{m} \tau_k - 1\right| > Km\right) < 2e^{-\frac{1}{2}K^2 m}. \tag{A.14}$$

Proof. Fix m. For any $\gamma \in (0,1)$,

$$\begin{aligned}
\mathbb{P}\left(\left|\sum_{k=1}^{m} \tau_k - 1\right| > Km\right) &= \mathbb{P}\left(\sum_{k=1}^{m} \tau_k > (1+K)m\right) + \mathbb{P}\left(-\sum_{k=1}^{m} \tau_k > -(1-K)m\right) \\
&= \mathbb{P}\left(e^{\gamma \sum_{k=1}^{m} \tau_k} > e^{(1+K)\gamma m}\right) + \mathbb{P}\left(e^{-\gamma \sum_{k=1}^{m} \tau_k} > e^{-(1-K)\gamma m}\right) \\
&\leqslant e^{-(1+K)\gamma m} \left(\mathbb{E}\left[e^{\gamma\tau}\right]\right)^m + e^{(1-K)\gamma m} \left(\mathbb{E}\left[e^{-\gamma\tau}\right]\right)^m \\
&= e^{-(1+K)\gamma m} (1-\gamma)^{-m} + e^{(1-K)\gamma m} (1+\gamma)^{-m} \\
&= \exp\left(-\gamma m \left[1 + K + \frac{\log(1-\gamma)}{\gamma}\right]\right) \\
&\quad + \exp\left(\gamma m \left[1 - K - \frac{\log(1+\gamma)}{\gamma}\right]\right).
\end{aligned}$$

The inequality is Markov's inequality and the equality right after follows from independence together with $\mathbb{E}[\exp(\alpha\tau)] = (1-\alpha)^{-1}$ for any $\alpha \in (-1,1)$. The other steps are algebraic manipulations. By Taylor's theorem with remainder there exists $\gamma_1 \in (-\gamma, 0)$ such that

$$\frac{1}{\gamma}\log(1-\gamma) = -1 - \frac{\gamma}{2(1-\gamma_1)^2} > -1 - \frac{\gamma}{2},$$

where the inequality follows since $\gamma_1 < 0$. Therefore

$$\exp\left(-\gamma m \left[1 + K + \frac{\log(1-\gamma)}{\gamma}\right]\right) \leqslant \exp\left(-\gamma m \left[K - \frac{\gamma}{2}\right]\right).$$

Similarly,

$$\exp\left(\gamma m \left[1 - K - \frac{\log(1+\gamma)}{\gamma}\right]\right) \leqslant \exp\left(-\gamma m \left[K - \frac{\gamma}{2}\right]\right).$$

So combining with the first computation of this proof and taking $\gamma = K$ gives

$$\mathbb{P}\left(\left|\sum_{k=1}^{m} \tau_k - 1\right| > Km\right) \leqslant 2\exp\left(-\gamma m\left[K - \frac{\gamma}{2}\right]\right) = 2e^{-\frac{1}{2}K^2 m}. \qquad \square$$

A.4 Proof of Lemma 4.5

Fix $t > 0$. The argument is similar to that of Lemma 4.2.

Variation of constants. Fix $m \in \mathbb{N}$. Since $\widetilde{S}^m_{h\tau} = \exp(h\tau_1 V_1)\cdots\exp(h\tau_{mn}V_{mn})$,

$$\partial_h \widetilde{S}^m_{h\tau} = \sum_{k=1}^{mn} \tau_k \widetilde{S}^{(1,k-1)}_{h\tau} V_k \widetilde{S}^{(k,mn)}_{h\tau} = \sum_{k=1}^{mn} \tau_k V_k \widetilde{S}^m_{h\tau} + \tau_k[\widetilde{S}^{(1,k-1)}h\tau, V_k]\widetilde{S}^{(k,mn)}_{h\tau}$$

$$= mV\widetilde{S}^m_{h\tau} + \sum_{k=1}^{mn}(\tau_k - 1)V_k\widetilde{S}^m_{h\tau} + \sum_{k=1}^{mn}\tau_k[\widetilde{S}^{(1,k-1)}_{h\tau}, V_k]\widetilde{S}^{(k,mn)}_{h\tau},$$

where the second equality is obtained by commuting $\widetilde{S}^{(1,k-1)}_{h\tau}$ and V_k, and the third by replacing τ_k with $\tau_k - 1 + 1$. So, setting $E^{(m)}_{h\tau} := \sum_{k=1}^{mn}\tau_k[\widetilde{S}^{(1,k-1)}_{h\tau}, V_k]\widetilde{S}^{(k,mn)}_{h\tau}$, variation of constants implies

$$\widetilde{S}^m_{h\tau} - S_{hm} = \int_0^h S_{m(h-r)}\left(\sum_{k=1}^{mn}(\tau_k - 1)V_k\right)\widetilde{S}^m_{r\tau}dr + \int_0^h S_{m(h-r)}E^{(m)}_{r\tau}dr.$$

Therefore, since $\|S_{m(h-r)}\|_{0\to 0} = 1$,

$$\|\widetilde{S}^m_{h\tau} - S_{hm}\|_{2\to 0} \leqslant \int_0^h \left\|\sum_{k=1}^{mn}(\tau_k - 1)V_k\right\|_{1\to 0}\left\|\widetilde{S}^m_{r\tau}\right\|_{2\to 1} dr + \int_0^h \|E^{(m)}_{r\tau}\|_{2\to 0}dr.$$

Let $I_1(h)$ and $I_2(h)$ denote the first and second integrals, respectively. Then for any $\varepsilon > 0$,

$$\mathbb{P}\left(\|\widetilde{S}^m_{h\tau} - S_{hm}\|_{2\to 0} > \frac{\varepsilon}{m}\right) \leqslant \mathbb{P}\left(I_1(h) > \frac{\varepsilon}{2m}\right) + \mathbb{P}\left(I_2(h) > \frac{\varepsilon}{2m}\right). \qquad \text{(A.15)}$$

Consider the two probabilities on the right, called the *first* and *second probabilities*, separately.

First probability. Note $\sum_{k=1}^{mn}(\tau_k-1)V_k = \sum_{k=1}^{n}\sum_{j=1}^{m}(\tau_j^{(k)}-1)V_k$ where $\tau_j^{(k)} := \tau_{(j-1)n+k}$. So

$$\left\|\sum_{k=1}^{mn}(\tau_k-1)V_k\right\|_{1\to 0} \leqslant C_* \sum_{k=1}^{n}\left|\sum_{j=1}^{m}\tau_j^{(k)}-1\right|,$$

and together with Lemma A.1 and Equation (A.3),

$$I_1(h) \leqslant C_* \sum_{k=1}^{n}\left|\sum_{j=1}^{m}\tau_j^{(k)}-1\right| \int_0^h \prod_{k=1}^{n} e^{C_* r\sum_{j=1}^{m}\tau_j^{(k)}}\,dr\,.$$

Therefore

$$\begin{aligned}\mathbb{P}\left(I_1(h) > \frac{\varepsilon}{2m}\right) &\leqslant \mathbb{P}\left(C_* \sum_{k=1}^{n}\left|\sum_{j=1}^{m}\tau_j^{(k)}-1\right| \int_0^h \prod_{k=1}^{n} e^{C_* r\sum_{j=1}^{m}\tau_j^{(k)}}\,dr > \frac{\varepsilon}{2m}\right)\\ &\leqslant \sum_{k=1}^{n}\mathbb{P}\left(\left|\sum_{j=1}^{m}\tau_j^{(k)}-1\right| \int_0^h \prod_{k=1}^{n} e^{C_* r\sum_{j=1}^{m}\tau_j^{(k)}}\,dr > \frac{\varepsilon}{2C_* mn}\right).\end{aligned}$$

The second inequality follows from a union bound and the fact that for any nonnegative random variables X_k and constant c, $\{\sum_{k=1}^{n} X_k > c\} \subseteq \cup_{k=1}^{n}\{X_k > c/n\}$. Set

$$A(h) := \bigcap_{k=1}^{n}\left\{h\sum_{j=1}^{m}\tau_j^{(k)} \leqslant \alpha\right\}$$

and

$$B_k(h) := \left\{\left|\sum_{j=1}^{m}\tau_j^{(k)}-1\right| \int_0^h \prod_{k=1}^{n} e^{C_* r\sum_{j=1}^{m}\tau_j^{(k)}}\,dr > \frac{\varepsilon}{2C_* mn}\right\}$$

for arbitrary $\alpha > 0$ and note that

$$A(h)\cap B_k(h) \subseteq \left\{\left|\sum_{j=1}^{m}\tau_j^{(k)}-1\right| h e^{C_* n\alpha} > \frac{\varepsilon}{2C_* mn}\right\} =: B(h).$$

Therefore

$$\begin{aligned}\mathbb{P}\left(I_1(h) > \frac{\varepsilon}{m}\right) &\leqslant \sum_{k=1}^{n}\mathbb{P}\left(B_k(h)\cap A(h)\right) + \mathbb{P}\left(B_k(h)\cap A(h)^c\right)\\ &\leqslant n\Big[\mathbb{P}\left(B(h)\right) + \mathbb{P}\left(A(h)^c\right)\Big].\end{aligned}$$

Set $h = t/m^2$. By Lemma A.4 for all $\varepsilon > 0$ such that $K := \varepsilon(2C_* tn)^{-1}e^{-C_* n\alpha} < 1$,

$$\mathbb{P}\left(B(h)\right) = \mathbb{P}\left(\left|\sum_{j=1}^{m} \tau_j^{(k)} - 1\right| > \frac{\varepsilon m}{2C_* tne^{C_* n\alpha}}\right) \leqslant 2e^{-\frac{1}{2}K^2 m}.$$

And by Lemma A.3,

$$\mathbb{P}\left(A(h)^c\right) = \mathbb{P}\left(\bigcup_{k=1}^{n}\left\{\sum_{j=1}^{m} \tau_j^{(k)} > \frac{\alpha}{h}\right\}\right) \leqslant n\mathbb{P}\left(\sum_{j=1}^{m} \tau_j > \frac{\alpha m^2}{t}\right) \leqslant n2^m e^{-\frac{1}{2}K' m^2}$$

where $K' := \alpha/t$. Therefore

$$\mathbb{P}\left(I_1(h) > \frac{\varepsilon}{2m}\right) \leqslant 2e^{-\frac{1}{2}K^2 m} + 2^m n e^{-\frac{1}{2}K' m^2} \leqslant 2^m C e^{-\frac{1}{2}Cm^2} \tag{A.16}$$

for some positive constant C independent of m.

Second probability. Recall $E_{r\tau}^{(m)} := \sum_{k=1}^{mn} \tau_k [\widetilde{S}_{r\tau}^{(1,k-1)}, V_k]\widetilde{S}_{r\tau}^{(k,mn)}$. Also, from Equation (A.6),

$$\begin{aligned}[\widetilde{S}_{r\tau}^{(1,k-1)}, V_k]\widetilde{S}_{r\tau}^{(k,mn)} &= \sum_{j=1}^{k-1}\int_0^r \tau_j e^{(r-s)\sum_{j=1}^{k-1}\tau_j V_j}[V_j, V_k]\widetilde{S}_{s\tau}^{(1,k-1)}\widetilde{S}_{r\tau}^{(k,mn)} ds \\ &\quad + \sum_{j=1}^{k-1}\int_0^r \tau_j e^{(r-s)\sum_{j=1}^{k-1}\tau_j V_j}\big[[\widetilde{S}_{s\tau}^{(1,j-1)}, V_j]\widetilde{S}_{s\tau}^{(j,k-1)}, V_k\big]\widetilde{S}_{r\tau}^{(k,mn)} ds.\end{aligned}$$

Lemma A.1 together with $\|[V_j, V_k]\|_{2\to 0} \leqslant \|V_j\|_{1\to 0}\|V_k\|_{2\to 1} + \|V_k\|_{1\to 0}\|V_j\|_{2\to 1} \leqslant 2C_*^2$ give

$$\left\|[V_j, V_k]\widetilde{S}_{s\tau}^{(1,k-1)}\widetilde{S}_{r\tau}^{(k,mn)}\right\|_{2\to 0} \leqslant 2C_*^2\left(1 + C_* r\sum_{j=1}^{mn}\tau_j\right)e^{3C_* r\sum_1^{mn}\tau_j}.$$

Also,

$$\begin{aligned}\big[[\widetilde{S}_{s\tau}^{(1,j-1)}, V_j]\widetilde{S}_{s\tau}^{(j,k-1)}, V_k\big] &= \widetilde{S}_{s\tau}^{(1,j-1)}V_j\widetilde{S}_{s\tau}^{(j,k-1)}V_k - V_k\widetilde{S}_{s\tau}^{(1,j-1)}V_j\widetilde{S}_{s\tau}^{(j,k-1)} \\ &\quad - V_j\widetilde{S}_{s\tau}^{(1,k-1)}V_k + V_kV_j\widetilde{S}_{s\tau}^{(1,k-1)}\end{aligned}$$

together with Lemma A.1 gives

$$\left\|\left[[\widetilde{S}_{s\tau}^{(1,j-1)}, V_j]\widetilde{S}_{s\tau}^{(j,k-1)}, V_k\right]\widetilde{S}_{r\tau}^{(k,mn)}\right\|_{2\to 0} \leqslant 4C_*^2\left(1 + C_* r\sum_{j=1}^{mn}\tau_j\right)e^{3C_* r\sum_1^{mn}\tau_j}.$$

Therefore for any $0 \leqslant r \leqslant h$,

$$\begin{aligned}\left\|E_{r\tau}^{(m)}\right\|_{2\to 0} &\leqslant \sum_{k=1}^{mn}\sum_{j=1}^{k-1}\tau_k\tau_j\int_0^r \left\|[V_j, V_k]\widetilde{S}_{s\tau}^{(1,k-1)}\widetilde{S}_{r\tau}^{(k,mn)}\right\|_{2\to 0} \\ &\qquad + \left\|\left[[\widetilde{S}_{s\tau}^{(1,j-1)}, V_j]\widetilde{S}_{s\tau}^{(j,k-1)}, V_k\right]\widetilde{S}_{r\tau}^{(k,mn)}\right\|_{2\to 0} ds \\ &\leqslant 6C_*^2 r\left(1 + C_* r\sum_{\ell=1}^{mn}\tau_\ell\right)e^{3C_* r\sum_1^{mn}\tau_\ell}\sum_{k=1}^{mn}\sum_{j=1}^{k-1}\tau_k\tau_j \\ &\leqslant Ch\left(1 + Ch\sum_{\ell=1}^{mn}\tau_\ell\right)e^{Ch\sum_1^{mn}\tau_\ell}\left(\sum_{k=1}^{mn}\tau_k\right)^2\end{aligned}$$

for some $C > 0$. So, we have that

$$I_2(h) = \int_0^h \|E_{r\tau}^{(m)}\|_{2\to 0}dr \leqslant Ch^2\left(1 + Ch\sum_{\ell=1}^{mn}\tau_\ell\right)e^{Ch\sum_1^{mn}\tau_\ell}\left(\sum_{k=1}^{mn}\tau_k\right)^2.$$

For arbitrary $\alpha > 0$, set

$$A(h) := \left\{h\sum_{k=1}^{mn}\tau_k \leqslant \alpha\right\}$$

and

$$B(h) := \left\{Ch^2\left(1 + Ch\sum_{\ell=1}^{mn}\tau_\ell\right)e^{Ch\sum_1^{mn}\tau_\ell}\left(\sum_{k=1}^{mn}\tau_k\right)^2 > \frac{\varepsilon}{2m}\right\}.$$

Then taking $h = t/m^2$ as before,

$$\begin{aligned}
\mathbb{P}\left(I_2(h) > \frac{\varepsilon}{2m}\right) &= \mathbb{P}\left(A(h) \cap B(h)\right) + \mathbb{P}\left(A(h)^c \cap B(h)\right) \\
&\leqslant \mathbb{P}\left(Ch^2\left(1 + C\alpha\right)e^{C\alpha}\left(\sum_{k=1}^{mn}\tau_k\right)^2 > \frac{\varepsilon}{2m}\right) + \mathbb{P}\left(h\sum_{k=1}^{mn}\tau_k > \alpha\right) \\
&= \mathbb{P}\left(\sum_{k=1}^{mn}\tau_k > Km^{\frac{3}{2}}\right) + \mathbb{P}\left(\sum_{k=1}^{mn}\tau_k > \frac{\alpha m^2}{t}\right) \\
&\leqslant n\left[\mathbb{P}\left(\sum_{k=1}^{m}\tau_k > K'm^{\frac{3}{2}}\right) + \mathbb{P}\left(\sum_{k=1}^{m}\tau_k > \frac{\alpha m^2}{nt}\right)\right] \\
&\leqslant n\left(2^m e^{-\frac{1}{2}K'm^{3/2}} + 2^m e^{-\frac{1}{2}K''m^2}\right) \leqslant 2^m C' e^{-\frac{1}{2}C'm^{3/2}}
\end{aligned} \tag{A.17}$$

for some $C' > 0$ where $K = (\varepsilon(2t^2C(1 + C\alpha)e^{C\alpha})^{-1})^{1/2}$, $K' = Kn^{-1}$, $K'' = \alpha(nt)^{-1}$, and the second-to-last last inequality follows from Lemma A.3. Combining (A.15), (A.16), and (A.17) and taking $h = t/m^2$ we therefore have that for all ε sufficiently small,

$$\mathbb{P}\left(\|\widetilde{S}^m_{t\tau/m^2} - S_{t/m}\|_{2\to 0} > \tfrac{\varepsilon}{m}\right) \leqslant 2^m C'' e^{-\frac{1}{2}C''m^{3/2}}$$

for some constant $C'' > 0$ independent of m. So, we have that

$$\sum_{m=1}^{\infty}\mathbb{P}\left(\|\widetilde{S}^m_{t\tau/m^2} - S_{t/m}\|_{2\to 0} > \tfrac{\varepsilon}{m}\right) \leqslant \sum_{m=1}^{\infty} 2^m C'' e^{-\frac{1}{2}C''m^{3/2}} < \infty. \qquad \square$$

Appendix B

Controllability lemmas

Combining the partial results obtained above we show the existence of transformations implementing the steps listed at the beginning of the section:

Lemma B.1. *If $q^{(0)}$ in $\mathcal{Q}_0$ is nondegenerate, then there exists M_1 and a sequence of transition times and interaction triples $\{\iota(m), \tau(m)\}_{m=1}^{M_1}$ such that $\Phi_{\tau(M_1)}^{\iota(M_1)} \circ \cdots \circ \Phi_{\tau(1)}^{\iota(1)}(q^{(0)}) = q^{(1)}$ as in* (6.30).

Proof. If (6.30) is satisfied by $q^{(0)}$ we simply set $M_1 = 0$, $q^{(1)} = q^{(0)}$. If not, by nondegeneracy there exists a sequence of triples $\{\iota(m)\}_{m=1}^{M}$ with $\iota(m) = \boldsymbol{j}(m)\boldsymbol{k}(m)\boldsymbol{\ell}(m)$ such that $\mathcal{A}_0 := \mathcal{A}(q^{(0)})$ and $\mathcal{A}_m = \mathcal{A}_{m-1} \oplus \boldsymbol{\ell}(m)$ with $\{(0,1,+),(1,0,+),(j^*,-)\} \subset \mathcal{A}_M$. We notice that all steps of this procedure satisfy, upon possibly reordering the indices within each triple, either the conditions of Lemma 6.15 (b) or of Lemma 6.16, so we sequentially choose $\tau(m) = \tau_+^{\iota(m)}$ from those lemmas.

To activate coordinate $(1,1,-)$ – if this was not already done in the previous procedure – we start with component $b_{j^*} \neq 0$ for $|j^*| \neq 1$ and consider a nearest neighbors path $\{\ell(n)\}_{n=1}^{M'}$ in $\mathbb{Z}_N^2$ connecting j^* to $(1,1)$ without performing any step on the axes. It is easy to see that such path can be realized through repeated application of Lemma 6.15 (b) by choosing for the n-th step the triples $\iota(n) = (0,1,+)(\ell(n),-)(\ell(n) \pm (0,1),-)$ or $\iota(n) = (1,0,+)(\ell(n),-)(\ell(n) \pm (1,0),-)$ for vertical and horizontal steps respectively.

Finally, coordinates $(1,0,-)$ and $(0,1,-)$ can be activated by applying Corollary 6.17 to the triples $(1,0,-)(0,1,+)(1,1,-)$ and $(1,0,+)(0,1,-)(1,1,-)$ respectively, while $(1,1,+)$ is activated by (b) by interchanging the type of modes $(1,1,-)$ and $(1,0,+)$ (or $(0,1,+)$) in $\iota(M')$ from the previous paragraph to $(1,1,+)$ and $(1,0,-)$ (or $(0,1,-)$). $\square$

Lemma B.2. *Let $q^{(1)}$ be a nondegenerate point in $\mathcal{Q}_0$ satisfying* (6.30)*. Then there exists M_2 and a sequence of interacting triples and transition times $\{\iota(m),\tau(m)\}_{m=1}^{M_2}$ such that $\Phi^{\iota(M_2)}_{\tau(M_2)} \circ \cdots \circ \Phi^{\iota(1)}_{\tau(1)}(q^{(1)}) = q^{(2)}$ is a nondegenerate point in $\mathcal{Q}_0$ satisfying* (6.31) *and* (6.32)*.*

Proof. In this part of the proof, we only consider interactions involving triples $\iota(m)$ of the form

$$\Big\{(0,1)(l,h)(l,h\pm 1) \text{ or } (1,0)(l,h)(l\pm 1,h) :\ |l|,|h| \leqslant N,\ |(l,h)| \neq 1\Big\}. \qquad \text{(B.1)}$$

By Lemma 6.15 (a), if $|j| < |k| < |\ell|$ and $(0,1),(l,h) \in \mathcal{A}(q)$ there exists $\tau(m) = \tau_-^{\iota(m)}$ such that defining $\mathcal{A}_m = \mathcal{A}(\varphi^{\iota(m)}_{\tau(m)}(q))$ we have $(l,h) \notin \mathcal{A}_m$ and $(0,1) \in \mathcal{A}_m$ (and similarly for $(1,0)$) [1]. Note that while a triple as above satisfies by assumption that $|j| < |k| < |\ell|$ and at least two of its coordinates are nonvanishing, it does not, in general, satisfy (6.37). However, assuming that q does not satisfy (6.37), by Lemma 6.16 and setting $\iota' = (1,0)(0,1)(1,1)$, there exists $\tau^{\iota'}$ such that $|q_{(1,0)}| \neq |(\Phi^{\iota'}_{\tau^{\iota'}}(q))_{(1,0)}| > 0$. Since none of the coordinates in $\mathbb{Z}^2_N \backslash \{(1,0)(0,1)(1,1)\}$ are affected

[1] Note that the same result can trivially be obtained if $(l,h) \notin \mathcal{A}(q)$ setting $\tau_-^{\iota(m)} = 0$

by this operation, $(\Phi^{\iota'}_{\tau^{\iota'}}(q))$ satisfies (6.37) and Lemma 6.15 can be applied to this state.

To conclude the proof we identify a sequence of triples $\iota(m) = (j(m), k(m), \ell(m)) \in \mathcal{I}$ of the form (B.1) such that for $\mathcal{A}_0 = \mathcal{A}(q^{(1)}) \subseteq \mathbb{Z}_N^2 \times \{+,-\}$

$$\begin{aligned} &(((\mathcal{A}_0 \ominus k(1)) \ominus k(2)) \ominus \dots) \ominus k(M_2) \\ &\quad = \{(1,0,\chi),(0,1,\chi),(1,1\chi),(N,N\chi),(-N,N\chi)\,,\chi \in \{+,-\}\}\,. \end{aligned} \tag{B.2}$$

A possible such sequence is given by triples of the form

$$\Big\{(1,0,+)(l,h,\chi)(l+1,h,\chi) \;:\; (l,h) \in \{(0,2),\dots,(0,N)\}\,,\chi \in \{+,-\}\Big\}$$

to remove the vertical column of $\mathbb{Z}_N^2$ (which cannot interact with $(0,1)$), followed by

$$\begin{aligned} \Big\{\Big((0,1,+)(l,h,\chi)(l,h+1,\chi) \;:\; (l,h) \in \big\{(l,0),\dots,(l,N) : \\ |l| \in (1,\dots,N-1)\big\}\backslash\{(1,1)\}\Big)\,,\chi \in \{+,-\}\Big\}, \end{aligned} \tag{B.3}$$

where importantly the set of transitions for each l is ordered. The above transformation zeroes all coefficients except those in the set $\{(1,1),(0,1),(1,0)\} \cup \{(l,N) \;:\; l \in (-N,\dots,N)\}$. We further remove the coefficients from $\{(l,N) \;:\; l \in (-N+1,\dots,N-1)\}$ by sequentially applying Lemma 6.15 to the ordered sequence of interacting triples

$$\Big((1,0,+)(l,h,\chi)(l+1,h,\chi) \;:\; (l,h) \in \{(0,N),\dots,(N-1,N)\,,\chi \in \{+,-\}\}\Big)\,, \tag{B.4}$$

and then

$$\Big((1,0,+)(l,h,\chi)(l-1,h,\chi) \;:\; (l,h) \in \{(-1,N),\dots,(-N+1,N)\}\,,\chi \in \{+,-\}\Big)\,. \tag{B.5}$$

It is easy to check that each transition in the above construction sequentially satisfies the assumptions of Lemma 6.15 (a), and that once a mode has been removed from $\mathcal{A}$ it will not interact again in this procedure. The fact that (6.32) holds follows from

(6.30) and that in an interacting triple $\iota = \boldsymbol{jk\ell}$ with $|j| < |k| < |l|$ both modes $\boldsymbol{j}$ and $\boldsymbol{\ell}$ are in $\mathcal{A}$ at the end of the interaction by τ_-^ι. □

Lemma B.3. *Let $q^{(2)}$ be a nondegenerate point in $\mathcal{Q}_0$ satisfying* (6.31) *and* (6.32). *Then there exists M_3 and a sequence of interacting triples and transition times $\{\iota(m), \tau(m)\}_{m=1}^{M_3}$ such that $\Phi_{\tau(M_3)}^{\iota(M_3)} \circ \cdots \circ \Phi_{\tau(1)}^{\iota(1)}(q^{(2)}) = q^{(3)}$ is a nondegenerate point in $\mathcal{Q}_0$ satisfying* (6.33) *and* (6.34).

Since it may not be possible to "transfer" the content of *e.g.*, mode $(-N, N)$ to $(-N+1, N)$ through one single interaction with mode $(1, 0)$ – and therefore it won't be possible to transfer the amplitude of mode $(-N, N)$ to (N, N) in one single "pass" – we proceed to prove that, through a sequence of interactions, we can transfer a *finite* and $q_{(-N,N)}$-independent amount of energy from mode $(-N, N)$ to (N, N). Therefore, the transfer of amplitude from mode $(-N, N)$ to (N, N) may be accomplished by repeating this sequence of interactions sufficiently many times.

The following corollary of Lemma 6.16 will be instrumental for the proof of Lemma B.3:

Corollary B.4. *Let $q_{(1,1)}, b_{(1,1)} \neq 0$ then for any q, q' with $q_{\boldsymbol{j}} = q'_{\boldsymbol{j}}$ for all $|j| > 1$ there exist a sequence $\{\iota(m), \tau(m)\}_{m=1}^{4}$ such that $\Phi_{\tau(4)}^{\iota(4)} \circ \cdots \circ \Phi_{\tau(1)}^{\iota(1)}(q) = q'$.*

Proof of Lemma B.3. The desired result follows upon showing that for any $i \in \{-N, \ldots, N\}$, setting $\boldsymbol{\ell} = (-i, N, \chi), \boldsymbol{\ell}' = (i, N, \chi')$ for $\chi, \chi' \in \{-, +\}$ there exists $M_{\boldsymbol{\ell},\boldsymbol{\ell}'}$ and a sequence of triples and interaction times $\{\iota(m),\, \tau(m)\}_{m=1}^{M_{\boldsymbol{\ell},\boldsymbol{\ell}'}}$ such that for any q satisfying $\bigcup_{|i'|<i}\{(i', N, +), (i', N, -)\} \cap \mathcal{A}(q) = \varnothing$ and $q' = \Phi_{\tau(M_{\boldsymbol{\ell},\boldsymbol{\ell}'})}^{\iota(M_{\boldsymbol{\ell},\boldsymbol{\ell}'})} \circ \cdots \circ \Phi_{\tau(1)}^{\iota(1)}(q)$ we have

$$q'_{\boldsymbol{j}} = \begin{cases} q_{\boldsymbol{j}} & \text{for } \boldsymbol{j} \in \mathbb{Z}_N^2 \backslash \{\boldsymbol{\ell}, \boldsymbol{\ell}'\}\,, \\ 0 & \text{for } \boldsymbol{j} = \boldsymbol{\ell} \text{ if } \boldsymbol{\ell} \neq \boldsymbol{\ell}'\,, \end{cases} \tag{B.6}$$

and for $\boldsymbol{k} \in \{\boldsymbol{\ell}, \boldsymbol{\ell}'\}$, $\mathrm{sign}(q_{\boldsymbol{k}}) = \mathrm{sign}(q'_{\boldsymbol{k}})$ holds if $q'_{\boldsymbol{k}} \neq 0$ (recalling our choice of notation $\mathrm{sign}(0) = +1$). Indeed, if $\mathrm{sign}(b_{(N,N)}) \geqslant 0$ we sequentially apply the above result to the pairs

$$(\boldsymbol{\ell}, \boldsymbol{\ell}') = ((N, N, +), (-N, N, +)), ((-N, N, +), (N, N, -)), ((-N, N, -), (N, N, -))\,. \tag{B.7}$$

Otherwise, when $\mathrm{sign}(b_{(N,N)}) = -1$ we first apply the above result to $\boldsymbol{\ell} = (N, N, -)$, $\boldsymbol{\ell}' = (-N, N, -)$ and then proceed as in the previous case.

We prove the result above by induction on $i \in \{0, \ldots, N\}$. The proof for $i \leqslant 0$ is analogous.

Base case ($i = 0 : (0, N, \chi) \to (0, N, \chi')$): If $\boldsymbol{\ell} = \boldsymbol{\ell}'$ there is nothing to show. We proceed to consider the case $\boldsymbol{\ell} = (0, N, +)$, $\boldsymbol{\ell}' = (0, N, -)$, as the converse follows by analogous arguments. In this case, for a sufficiently small $\varepsilon > 0$ we consider the interactions $\iota = (1, 0, +)(0, N, +)(1, N, +)$ and $\iota' = (1, 0, -)(0, N, -)(1, N, +)$, running the corresponding flow maps by a small amount of time $\tau(\varepsilon)$, $\tau'(\varepsilon)$ such that $(\Phi^{\iota'}_{\tau'(\varepsilon)} \circ \Phi^{\iota}_{\tau(\varepsilon)}(q)_{(0,N,-)})^2 = b^2_{(0,N)} + \varepsilon$. We then apply Corollary B.4 to the coordinates $(1, 0, +), (1, 0, -)$ to return them in the initial configuration. Note that the existence of a uniform $\varepsilon > 0$ such that the transitions above can be performed in a single pair of interactions (and therefore the finiteness of the total number of interactions required to perform the desired transformation) follows from the fact that $b_{(0,N)}$ is nondecreasing and the continuity of the dynamics together with Lemma 6.15.

Induction step ($i > 0 : (-i, N, \chi) \to (i, N, \chi')$): We consider two possibilities for q: a) there exists q'' with $|a''_{(1,0)}| \in [|a_{(1,0)}|/2, |a_{(1,0)}|]$, $q''_{(-i,N,\chi)} = 0$ and for $\iota'' = (1, 0, +)(-i + 1, N, \chi)(-i, N, \chi)$

$$E_{\iota''}(q) = E_{\iota''}(q'')\,, \qquad \mathcal{E}_{\iota''}(q) = \mathcal{E}_{\iota''}(q'')\,,$$

or b) such q'' does not exist.

In case a) the state q'' can be reached by letting $\iota = (1, 0, +)(-i+1, N, \chi)(-i, N, \chi)$ interact for a finite amount of time τ from Lemma 6.15 (c). Then, by the induction assumption there is a sequence of triples and interaction times allowing to reach a state q''' with $q'''_{(-i+1,N,\chi)} = 0$, $q'''_{(i-1,N,\chi')} = q''_{(-i+1,N,\chi)}$ and $q'''_{\boldsymbol{j}} = q''_{\boldsymbol{j}}$ for all other $j \in \mathbb{Z}^2_N$. The desired state can then be reached by application of Lemma 6.15 (a) to the triple $\iota = (1,0,+)(i-1,N,\chi')(i,N,\chi')$. We proceed to check that the final state satisfies (B.6). Because modes $j \notin \{(-i,N),\ldots,(i,N),(1,0)\}$ did not interact in the procedure above for such $\boldsymbol{j}$ we must have that $q_{\boldsymbol{j}} = q'_{\boldsymbol{j}}$. The fact that for $j \in \{(-i,N),\ldots,(i-1,N)\}$ $q'_{\boldsymbol{j}} = 0$ follows by construction and the induction assumption. It remains to check that $|a'_{(1,0)}| = |a_{(1,0)}|$. Since the only modes affected by the above transformation are $(-i,N,\chi),(i,N\chi'),(1,0,+)$, this follows directly by conservation of energy and enstrophy:

$$(q_{(-i,N,\chi)})^2 + (q_{(i,N,\chi')})^2 + (q_{(1,0,+)})^2 = (q'_{(i,N,\chi')})^2 + (q'_{(1,0,+)})^2\,,$$

$$\frac{(q_{(-i,N,\chi)})^2}{N^2+i^2} + \frac{(q_{(i,N,\chi')})^2}{N^2+i^2} + (q_{(1,0,+)})^2 = \frac{(q'_{(-i,N,\chi)})^2}{N^2+i^2} + (q'_{(1,0,+)})^2\,.$$

In case b) we proceed to show that case a) can be reached with a finite number of interactions. More specifically if condition a) is not satisfied we let the triple $\iota'' = (-i,N,\chi)(-i+1,N,\chi)(1,0,+)$ for $\chi \in \{+,-\}$ interact as described by Lemma 6.15 for a time τ'' to reach a nondegenerate point q'' in $\mathcal{Q}_0$ with $q''_{\boldsymbol{j}} = q_{\boldsymbol{j}}$ for $\boldsymbol{j} \notin \{(-i,N,\chi),(-i+1,N,\chi),(1,0,+)\}$, $a''_{(1,0)} = a_{(1,0)}/2$ and $q''_{(-i,N,\chi)}, q''_{(-i+1,N,\chi)}$ satisfying the conservation laws

$$(q_{(-i,N,\chi)})^2 + (q_{(1,0,+)})^2 = (q''_{(-i,N,\chi)})^2 + (q''_{(-i+1,N,\chi)})^2 + (q_{(1,0,+)}/2)^2\,,$$

$$\frac{(q_{(-i,N,\chi)})^2}{N^2+i^2} + (q_{(1,0,+)})^2 = \frac{(q''_{(-i,N,\chi)})^2}{N^2+i^2} + \frac{(q''_{(-i+1,N,\chi)})^2}{N^2+(i-1)^2} + (q_{(1,0,+)}/2)^2\,,$$

so that $(q''_{(-i,N,\chi)})^2 = (q_{(-i,N,\chi)})^2 - C_{N,i}(q_{(1,0)})^2$ for $C_{N,i} = \frac{3}{4}\frac{N^2+i^2}{i^2-(i-1)^2}(N^2+(i-1)^2-1)$. We see that a positive, $q_{(1,0,+)}$-dependent amplitude is removed from $(q_{(-i,N,\chi)})^2$. Again applying the induction step and Lemma 6.15 (a) to transfer, respectively, the amplitude from $(-i+1,N,\chi)$ to $(i-1,N,\chi')$ and from $(i-1,N,\chi')$ to (i,N,χ') we reach the state q' with $q_{\boldsymbol{j}} = q'_{\boldsymbol{j}}$ for modes $j \notin \{(-i,N),\dots,(i,N),(1,0)\}$ (since these modes either vanish in both cases or they did not interact). Further, by conservation of energy and enstrophy, we have that

$$(q_{(-i,N,\chi)})^2 + (q_{(i,N,\chi')})^2 + (q_{(1,0,+)})^2 = (q''_{(-i,N,\chi)})^2 + (q''_{(i,N,\chi')})^2 + (q''_{(1,0,+)})^2 \,,$$

$$\frac{(q_{(-i,N,\chi)})^2}{N^2+i^2} + \frac{(q_{(i,N,\chi')})^2}{N^2+i^2} + (q_{(1,0,+)})^2 = \frac{(q''_{(-i,N,\chi)})^2}{N^2+i^2} + \frac{(q''_{(i,N,\chi')})^2}{N^2+i^2} + (q''_{(1,0,+)})^2 \,,$$

so that $|q''_{(1,0,+)}| = |q_{(1,0,+)}|$. This shows that the amplitude $C_{N,i}(q_{(1,0,+)})^2$ subtracted to $q_{(-i,N,\chi)}$ is constant at each cycle, showing by boundedness of $q_{(-i,N,\chi)}$ that with a finite number of iterations as the one described above we can reach state a), concluding the proof. □

Lemma B.5. *Let $q^{(3)}$ be a nondegenerate point in $\mathcal{Q}_0$ satisfying* (6.33) *and* (6.34). *Then there exists M_4 and a sequence of interacting triples and transition times $\{\iota(m),\tau(m)\}_{m=1}^{M_4}$ such that $\Phi_{\tau(M_4)}^{\iota(M_4)} \circ \dots \circ \Phi_{\tau(1)}^{\iota(1)}(q^{(3)}) = q^*$ is a nondegenerate point in $\mathcal{Q}_0$ satisfying* (6.27).

Proof. We start the proof by applying Corollary B.4 to transform the state $q^{(3)}$ into $q = \Phi_{\tau(1)}(q^{(3)})$ satisfying $q_{\boldsymbol{j}}^{(3)} = q_{\boldsymbol{j}}$ for all $|j| > 1$ and $a_{(0,1)} = b_{(0,1)} = b_{(1,0)} = a_{(1,0)} > 0$. Throughout this proof, we refer to states q such that $q_{(i,i',\chi)} = q_{(i',i,\chi)}$ for all $i, i' \in (0,\dots,N)$, $\chi \in \{+,-\}$ as *symmetric.*

We then proceed to transfer the amplitude from $a_{(1,1)}$ to $b_{(2,1)}, b_{(1,2)}$ by transforming q into another symmetric state q' with $(2,1,-),(1,2,-) \in \mathcal{A}(q')$ and $(1,1,+) \notin \mathcal{A}(q')$. This can be done by letting triples $\iota(2) = (1,0,-)(1,1,+)(2,1,-) \in \mathcal{I}$ and $\iota(3) = (0,1,-)(1,1,+)(1,2,-) \in \mathcal{I}$ interact, and choosing the interaction times $\tau, \tau'(\tau)$ such that $\Phi_{\tau'(\tau)}^{\iota(3)} \circ \Phi_{\tau}^{\iota(2)}(q)_{(1,1,+)} = 0$. Further, we note that the difference $b'_{(1,2)} - b'_{(2,1)}$ is negative for $\tau = 0$, positive for $\tau'(\tau) = 0$ and is continuous in τ, so there must exist τ^* such that $b'_{(1,2)} = b'_{(2,1)}$. To show that q' is symmetric it only remains to show that

$b'_{(1,0)} = b'_{(0,1)}$. This follows from the conservation laws:

$$B_{(1,0)(1,1)}\left((b'_{(1,0)})^2 - (b_{(1,0)})^2\right) = B_{(2,1)(1,1)}(b'_{(2,1)})^2$$
$$= B_{(1,2)(1,1)}(b'_{(1,2)})^2 = B_{(0,1)(1,1)}\left((b'_{(0,1)})^2 - (b_{(0,1)})^2\right) \tag{B.8}$$

where

$$B_{jk} := \frac{1}{|j|^2} - \frac{1}{|k|^2}\,. \tag{B.9}$$

Next, let $\iota(4) = (1,0,-)(0,1,+)(1,1,-)$ and $\iota(5) = (0,1,-)(1,0,+)(1,1,-)$ interact. By Lemma 6.16 there exists an interaction time such that the initial state q' is mapped to q'' with $b''_{(1,0)} = b''_{(0,1)} = 0$ and $a''_{(1,0)} = a''_{(0,1)} > 0$, so that $(1,0,-),(0,1,-) \notin \mathcal{A}(q'')$.

We then proceed to transfer the amplitude from modes $(1,2,-)$ and $(2,1,-)$ to $(2,2,-)$. This is done letting triples $\iota(6) = (1,0,+)(1,2,-)(2,2,-)$ and $\iota(7) = (0,1,+)(2,1,-)(2,2,-)$ interact until the modes $(2,1,-),(1,2,-)$ are depleted, as proved in Lemma 6.15. The symmetry of the final state q''' is again a consequence of the conservation laws:

$$B_{(1,0)(2,2)}\left((a'''_{(1,0)})^2 - (a''_{(1,0)})^2\right) = B_{(2,1)(2,2)}(b''_{(2,1)})^2 = B_{(1,2)(2,2)}(b''_{(1,2)})^2 \tag{B.10}$$

$$= B_{(0,1)(2,2)}\left((a'''_{(0,1)})^2 - (a''_{(0,1)})^2\right). \tag{B.11}$$

Summarizing, we have reached a symmetric state $q''' = \Phi^{\iota(7)}_{\tau(7)} \circ \cdots \circ \Phi^{\iota(2)}_{\tau(2)}(q)$ with

$$\mathcal{A}(q''') = \{(1,0,+),(0,1,+),(2,2,-),(1,1,-),(N,N,-)\}\,. \tag{B.12}$$

The desired result then follows immediately if we can show that we can transfer the amplitude of mode $(i-1,i-1,-)$ to $(i,i,-)$ for $i \in (2,\ldots,N)$ while preserving the fact that $a'_{(1,0)} = a'_{(0,1)}$. We show this by considering, sequentially, the interaction triples

$$\iota(4i) = (1,0,+)(i-1,i-1,-)(i,i-1,-)\,,$$

$$\iota(4i+1) = (0,1,+)(i-1,i-1,-)(i-1,i,-)\,,$$

$$\iota(4i+2) = (0,1,+)(i,i-1,-)(i,i,-)\,,$$

$$\iota(4i+3) = (1,0,+)(i-1,i,-)(i,i,-)\,.$$

More specifically, we consider the family of endpoints

$$q''(t) = \Phi^{\iota(4i+3)}_{\tau_-^{\iota(4i+3)}} \circ \Phi^{\iota(4i+2)}_{\tau_-^{\iota(4i+2)}} \circ \Phi^{\iota(4i+1)}_{\tau_-^{\iota(4i+1)}} \circ \Phi^{\iota(4i)}_{t}(q')\,, \tag{B.13}$$

where τ_-^{ι} is defined in Lemma 6.15 (a). By construction, this sequence implies that $a''_{(i-1,i-1)} = a''_{(i-1,i)} = a''_{(i,i-1)} = 0$ and $a''_{(i,i)} \neq 0$. It remains to prove $a''_{(1,0)} = a''_{(0,1)}$. As a composition of continuous functions, $q''(t)$ is continuous in t so $\Delta q(t) = a''_{(1,0)}(t) - a''_{(0,1)}(t)$ is as well. Further, since by symmetry $a''_{(1,0)}(0) = a''_{(0,1)}(\tau_-^{\iota(4i)})$, we must have $\mathrm{sign}(\Delta q(0)) = -\mathrm{sign}(\Delta q(\tau_-^{\iota_1}))$. This implies the existence of $\tau(4i) \in [0, \tau_-^{\iota_1}]$ with $\Delta q(0) = 0$, concluding the proof. $\square$

Appendix C

Euler spanning

What follows are the Mathematica commands needed to reproduce some of the more tedious calcluations from Section 6.2.4. We have reformated some of the outputs in LaTeXto improve readability. The source Mathematica file can be found at [1].

```
In[1]:= V1={C_kl a_k a_l,0,C_jl a_j a_l,0,C_jk a_j a_k,0};
V2={C_kl b_k b_l,0,0,C_jl a_j b_l,0,C_jk a_j b_k};
V3={0,C_kl a_k b_l,C_jl b_j b_l,0,0,C_jk b_j a_k};
V4={0,C_kl b_k a_l,0,C_jl b_j a_l,C_jk b_j b_k,0};

W1=∂_{{a_j,b_j,a_k,b_k,a_l,b_l}}V1.{η_j^a,η_j^b,η_k^a,η_k^b,η_l^a,η_l^b}
W2=∂_{{a_j,b_j,a_k,b_k,a_l,b_l}}V2.{η_j^a,η_j^b,η_k^a,η_k^b,η_l^a,η_l^b}
W3=∂_{{a_j,b_j,a_k,b_k,a_l,b_l}}V3.{η_j^a,η_j^b,η_k^a,η_k^b,η_l^a,η_l^b}
W4=∂_{{a_j,b_j,a_k,b_k,a_l,b_l}}V4.{η_j^a,η_j^b,η_k^a,η_k^b,η_l^a,η_l^b};

V12=∂_{{a_j,b_j,a_k,b_k,a_l,b_l}}V1.V2-∂_{{a_j,b_j,a_k,b_k,a_l,b_l}}V2.V1;
V13=∂_{{a_j,b_j,a_k,b_k,a_l,b_l}}V1.V3-∂_{{a_j,b_j,a_k,b_k,a_l,b_l}}V3.V1;
```

```
V14=∂{{aj,bj,ak,bk,al,bl}}V1.V4-∂{{aj,bj,ak,bk,al,bl}}V4.V1;
V23=∂{{aj,bj,ak,bk,al,bl}}V2.V3-∂{{aj,bj,ak,bk,al,bl}}V3.V2;
V24=∂{{aj,bj,ak,bk,al,bl}}V2.V4-∂{{aj,bj,ak,bk,al,bl}}V4.V2;
V34=∂{{aj,bj,ak,bk,al,bl}}V3.V4-∂{{aj,bj,ak,bk,al,bl}}V4.V3;

W12=∂{{aj,bj,ak,bk,al,bl}}V12.{ηj^a,ηj^b,ηk^a,ηk^b,ηl^a,ηl^b};
W13=∂{{aj,bj,ak,bk,al,bl}}V13.{ηj^a,ηj^b,ηk^a,ηk^b,ηl^a,ηl^b};
W14=∂{{aj,bj,ak,bk,al,bl}}V14.{ηj^a,ηj^b,ηk^a,ηk^b,ηl^a,ηl^b};
W23=∂{{aj,bj,ak,bk,al,bl}}V23.{ηj^a,ηj^b,ηk^a,ηk^b,ηl^a,ηl^b};
W24=∂{{aj,bj,ak,bk,al,bl}}V24.{ηj^a,ηj^b,ηk^a,ηk^b,ηl^a,ηl^b};
W34=∂{{aj,bj,ak,bk,al,bl}}V34.{ηj^a,ηj^b,ηk^a,ηk^b,ηl^a,ηl^b};
```

The computation for the determinant of $M'_{F(i)}$ for $i \in \{1, \ldots, n-3\} \backslash \{N+1\}$ is

```
In[2]:= A0 = Join[Transpose[{V1,V2,V13,V34}][[5;;]],Transpose[{W1,W2,W13,W34}][[5;;]]]
          /.{aj->α1 ,ak->β,al->α2,bj->2*α1,bk->β,bl->α2,
             ηj^a->1,ηk^a->1,ηl^a->1,ηj^b-> -1/2, ηk^b->-1,ηl b->-1};;
        A0//MatrixForm
        Simplify[Det[A0]]
```

The matrix output is

$$M'_{F(i)} = \begin{pmatrix} \beta\gamma_1 C_{jk} & 0 & \gamma_1^2\gamma_2 j C_{jk} C_{j\ell} & -j\beta^2\gamma_2 C_{jk} C_{k\ell} \\ 0 & \beta\gamma_1 C_{jk} & -\gamma_1^2\gamma_2 C_{jk} C_{j\ell} & \beta^2\gamma_2 C_{jk} C_{k\ell} \\ \beta C_{jk} + \gamma_1 C_{jk} & 0 & C_{jk} C_{j\ell}\left(-\frac{\gamma_1^2}{j} + \gamma_2\gamma_1 j - \gamma_2\gamma_1 j\right) & \frac{\beta^2 C_{jk} C_{k\ell}}{j} \\ 0 & \beta C_{jk} - \gamma_1 C_{jk} & C_{jk} C_{j\ell}\left(-\gamma_1^2 - \gamma_2\gamma_1 k + \gamma_2\gamma_1\right) & \beta^2 C_{jk} C_{k\ell} \end{pmatrix},$$

and the corresponding determinant

$$\det M'_{F(i)} = \frac{3}{2j} C_{jk}^4 C_{jl} C_{kl} \beta^3 \gamma_1^3 \gamma_2 \left(\beta + \beta j^2 + 2\gamma_2 j^2\right) .$$

The computation for the determinant of the matrix $M'_{F(N+1)}$ is given by

```
In[3]:= A1 = Join[Transpose[{V1,V2,V13,V34}][[5;;]],Transpose[{W1,W2,W13,W34}][[5;;]]]
   /.{aj->α1 ,ak->β,al->β,bj->2*α1,bk->2*β,bl->β,
     ηj^a->1,ηk^a->1,ηl^a->1,ηj^b->-1/2, ηk^b->-1/2,ηl^b->-1};;
A1//MatrixForm
Simplify[Det[A1]]
```

The matrix output is

$$M'_{F(N+1)} = \begin{pmatrix} \alpha_1\beta C_{jk} & 0 & 2\alpha_1^2\beta C_{jk}C_{j\ell} & -2\beta^3 C_{jk}C_{k\ell} \\ 0 & 2\alpha_1\beta C_{jk} & -2\alpha_1^2\beta C_{jk}C_{j\ell} & 2\beta^3 C_{jk}C_{k\ell} \\ \alpha_1 C_{jk} + \beta C_{jk} & 0 & \frac{3}{2}\alpha_1\beta C_{jk}C_{j\ell} - 2\alpha_1^2 C_{jk}C_{j\ell} & \frac{1}{2}\beta^2 C_{jk}C_{k\ell} \\ 0 & 2\beta C_{jk} - \frac{\alpha_1 C_{jk}}{2} & -\frac{3}{2}\alpha_1\beta C_{jk}C_{j\ell} - 2\alpha_1^2 C_{jk}C_{j\ell} & \frac{7}{2}\beta^2 C_{jk}C_{k\ell} \end{pmatrix},$$

and the corresponding determinant

$$\det M'_{F(N+1)} = \frac{39}{2}\alpha_1^3\beta^4(\beta - \alpha_1)C_{jk}^4 C_{j\ell}C_{k\ell}\,.$$

The computation for the determinant of the matrix M'_0 is given by

```
In[8]:= A2 = Join[Transpose[{V1,V2,V3,V4,V12,V13,V14,V23}][[2;;5]],
   Transpose[{W1,W2,W3,W4,W12,W13,W14,W23}][[2;;5]]]
    /.{aj→ α1,ak→β,al→β,bj→2*α1,bk→β,bl→β,
      ηj^a→1, ηk^a→1,ηl^a→1,η j^b→-1/2,ηk^b→-1,ηl^b→-1};
A2//MatrixForm
Simplify[Det[A2]]
```

The first five columns of the matrix output are

$$\begin{pmatrix} 0 & 0 & \beta^2 C_{k\ell} & \beta^2 C_{k\ell} & 0 \\ \alpha_1\beta C_{j\ell} & 0 & 2\alpha_1\beta C_{j\ell} & 0 & \beta^3 C_{j\ell}C_{k\ell} \\ 0 & \alpha_1\beta C_{j\ell} & 0 & 2\alpha_1\beta C_{j\ell} & -\beta^3 C_{j\ell}C_{k\ell} \\ \alpha_1\beta C_{jk} & 0 & 0 & 2\alpha_1\beta C_{jk} & \beta^3 C_{jk}C_{k\ell} \\ 0 & 0 & 0 & 0 & 0 \\ \alpha_1 C_{j\ell} + \beta C_{j\ell} & 0 & -2\alpha_1 C_{j\ell} - \frac{\beta C_{j\ell}}{2} & 0 & -\beta^2 C_{j\ell}C_{k\ell} \\ 0 & \beta C_{j\ell} - \alpha_1 C_{j\ell} & 0 & 2\alpha_1 C_{j\ell} - \frac{\beta C_{j\ell}}{2} & -\beta^2 C_{j\ell}C_{k\ell} \\ \alpha_1 C_{jk} + \beta C_{jk} & 0 & 0 & -2\alpha_1 C_{jk} - \frac{\beta C_{jk}}{2} & -\beta^2 C_{jk}C_{k\ell} \end{pmatrix},$$

while the last three columns read

$$\begin{pmatrix} \alpha_1\left(-\beta^2\right)C_{j\ell}C_{k\ell} & \alpha_1\left(-\beta^2\right)C_{jk}C_{k\ell} & \alpha_1\left(-\beta^2\right)C_{jk}C_{k\ell} \\ 0 & 2\alpha_1^2\beta C_{jk}C_{j\ell} & -2\alpha_1^2\beta C_{jk}C_{j\ell} \\ 0 & -2\alpha_1^2\beta C_{jk}C_{j\ell} & 2\alpha_1^2\beta C_{jk}C_{j\ell} \\ 2\alpha_1^2\beta C_{jk}C_{j\ell} & 0 & 0 \\ -\beta^2 C_{j\ell}C_{k\ell} & -\beta^2 C_{jk}C_{k\ell} & -\beta^2 C_{jk}C_{k\ell} \\ 0 & \frac{3}{2}\alpha_1\beta C_{jk}C_{j\ell} - 2\alpha_1^2 C_{jk}C_{j\ell} & 2\alpha_1^2 C_{jk}C_{j\ell} - \frac{3}{2}\alpha_1\beta C_{jk}C_{j\ell} \\ 0 & -\frac{3}{2}\alpha_1\beta C_{jk}C_{j\ell} - 2\alpha_1^2 C_{jk}C_{j\ell} & \frac{3}{2}\alpha_1\beta C_{jk}C_{j\ell} + 2\alpha_1^2 C_{jk}C_{j\ell} \\ \frac{3}{2}\alpha_1\beta C_{jk}C_{j\ell} - 2\alpha_1^2 C_{jk}C_{j\ell} & 0 & 0 \end{pmatrix},$$

and the corresponding determinant is

$$\det M_0' = -96\alpha_1^5\beta^{11}C_{jk}^4C_{j\ell}^5C_{k\ell}^3\,.$$

Bibliography

[1] A. Agazzi, J. C. Mattingly, and O. Melikechi. `https://github.com/agazzian/RandomSplittingLyapExp.git`, 2022.

[2] A. Agazzi, J. C. Mattingly, and O. Melikechi. Random splitting of fluid models: Ergodicity and convergence, 2022.

[3] A. Agazzi, J. C. Mattingly, and O. Melikechi. Random splitting of fluid models: Positive lyapunov exponents, 2022.

[4] A. Arnold and C. Ringhofer. An operator splitting method for the Wigner-Poisson problem. *SIAM J. Numer. Anal.*, 33(4):1622–1643, 1996.

[5] L. Arnold. *Random dynamical systems*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 1998.

[6] Y. Bakhtin and T. Hurth. Invariant densities for dynamical systems with random switching. *Nonlinearity*, 25:2937–2952, 2012.

[7] Y. Bakhtin, T. Hurth, S. D. Lawley, and J. C. Mattingly. Smooth invariant densities for random switching on the torus. *Nonlinearity*, 31(4):1331–1350, 2018.

[8] Y. Bakhtin, T. Hurth, and J. C. Mattingly. Regularity of invariant densities for 1D systems with random switching. *Nonlinearity*, 28(11):3755–3787, 2015.

[9] P. H. Baxendale. Lyapunov exponents and relative entropy for a stochastic flow of diffeomorphisms. *Probability Theory and Related Fields*, 81:521–554, 1989.

[10] J. Bedrossian, A. Blumenthal, and S. Punshon-Smith. A regularity method for lower bounds on the lyapunov exponent for stochastic differential equations. *Inventiones mathematicae*, 227(2):429–516, sep 2021.

[11] J. Bedrossian and S. Punshon-Smith. Chaos in stochastic 2d galerkin-navier-stokes, 2021.

[12] M. Benaïm, S. L. Borgne, F. Malrieu, and P.-A. Zitt. Qualitative properties of certain piecewise deterministic Markov processes. *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 51(3):1040 – 1075, 2015.

[13] M. Benaïm, T. Hurth, and E. Strickler. A user-friendly condition for exponential ergodicity in randomly switched environments. *Electronic Communications in Probability*, 23(none), Jan 2018.

[14] B. Bermúdez, A. Nikolás, and F. J. Sánchez. On operator splitting methods with upwinding for the unsteady Navier-Stokes equations. *East-West J. Numer. Math.*, 4(2):83–98, 1996.

[15] J. Bierkens, G. O. Roberts, and P.-A. Zitt. Ergodicity of the zigzag process. *Ann. Appl. Probab.*, 29(4):2266–2301, 2019.

[16] G. D. Birkhoff. Proof of the ergodic theorem. *Proceedings of the National Academy of Sciences*, 17:656 – 660, 1931.

[17] P. Bougerol. Comparaison des exposants de Lyapounov des processus markoviens multiplicatifs. *Ann. Inst. H. Poincaré Probab. Statist.*, 24(4):439–489, 1988.

[18] A. M. Childs, A. Ostrander, and Y. Su. Faster quantum simulation by randomization. *Quantum*, 3:182, Sep 2019.

[19] A. M. Childs, Y. Su, M. C. Tran, N. Wiebe, and S. Zhu. Theory of trotter error with commutator scaling. *Physical Review X*, 11(1):011020, Feb 2021.

[20] D. Chouliara, Y. Gong, S. He, A. Kiselev, J. Lim, O. Melikechi, and K. Powers. Hitting time of brownian motion subject to shear flow. *Involve, a Journal of Mathematics*, 15(1).

[21] I. P. Cornfeld, S. V. Fomin, and Y. G. Sinaĭ. *Ergodic theory*, volume 245 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, New York, 1982. Translated from the Russian by A. B. Sosinskiĭ.

[22] G. Da Prato and A. Debussche. Two-dimensional Navier-Stokes equations driven by a space-time white noise. *J. Funct. Anal.*, 196(1):180–210, 2002.

[23] R. L. Devaney. *An Introduction To Chaotic Dynamical Systems*. CRC Press, 3nd edition edition, 2022.

[24] P. Diaconis and D. Freedman. Iterated random functions. *SIAM Rev.*, 41(1):45–76, 1999.

[25] A. Durmus, A. Guillin, and P. Monmarché. Piecewise deterministic Markov processes and their invariant measures. *Ann. Inst. Henri Poincaré Probab. Stat.*, 57(3):1442–1475, 2021.

[26] W. E and J. C. Mattingly. Ergodicity for the Navier-Stokes equation with degenerate random forcing: finite-dimensional approximation. *Comm. Pure Appl. Math.*, 54(11):1386–1402, 2001.

[27] H. Federer. *Geometric measure theory.* Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York, Inc., New York, 1969.

[28] F. Flandoli and B. Maslowski. Ergodicity of the 2-D Navier-Stokes equation under random perturbations. *Comm. Math. Phys.*, 172(1):119–141, 1995.

[29] G. Folland. *Real Analysis: Modern Techniques and Their Applications.* Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 2013.

[30] M. I. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems*, volume 260 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, New York, second edition, 1998. Translated from the 1979 Russian original by Joseph Szücs.

[31] H. Furstenberg and H. Kesten. Products of random matrices. *Ann. Math. Statist.*, 31:457–469, 1960.

[32] H. Furstenberg and Y. Kifer. Random matrix products and measures on projective spaces. *Israel J. Math.*, 46(1-2):12–32, 1983.

[33] R. Glowinski, S. Osher, and W. Yin. *Splitting Methods in Communication, Imaging, Science and Engineering.* 01 2016.

[34] D. Goldman and T. J. Kaper. Nth-order operator splitting schemes and nonreversible systems. *SIAM J. Numer. Anal.*, 33(1):349–367, 1996.

[35] Y. Guivarc'h. Quelques propriétés asymptotiques des produits de matrices aléatoires. In *Eighth Saint Flour Probability Summer School—1978 (Saint Flour, 1978)*, volume 774 of *Lecture Notes in Math.*, pages 177–250. Springer, Berlin, 1980.

[36] M. Hairer. Convergence of markov processes. *Lecture notes*, 2010.

[37] M. Hairer and J. C. Mattingly. Ergodicity of the 2d navier-stokes equations with degenerate stochastic forcing. *Annals of Mathematics*, pages 993–1032, 2006.

[38] M. Hairer and J. C. Mattingly. Yet another look at Harris' ergodic theorem for Markov chains. In *Seminar on Stochastic Analysis, Random Fields and Applications VI*, volume 63 of *Progr. Probab.*, pages 109–117. Birkhäuser/Springer Basel AG, Basel, 2011.

[39] L. Hörmander. Hypoelliptic second order differential equations. *Acta Math.*, 119:147–171, 1967.

[40] V. Jurdjevic. *Geometric Control Theory.* Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1996.

[41] A. Katok and B. Hasselblatt. *Introduction to the modern theory of dynamical systems*, volume 54 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1995. With a supplementary chapter by Katok and Leonardo Mendoza.

[42] Y. Kifer. *Ergodic theory of random transformations*, volume 10 of *Progress in Probability and Statistics*. Birkhäuser Boston, Inc., Boston, MA, 1986.

[43] S. Kuksin, V. Nersesyan, and A. Shirikyan. Exponential mixing for a class of dissipative PDEs with bounded degenerate noise. *Geom. Funct. Anal.*, 30(1):126–187, 2020.

[44] S. Kusuoka. *Approximation of expectation of diffusion processes based on Lie algebra and Malliavin calculus*, volume 6 of *Adv. Math. Econ.*, page 69–83. Springer, 2004.

[45] S. D. Lawley, J. C. Mattingly, and M. C. Reed. Sensitivity to switching rates in stochastically switched ODEs. *Commun. Math. Sci.*, 12(7):1343–1352, 2014.

[46] J. M. Lee. *Introduction to Smooth Manifolds*. 2000.

[47] D. Li, S. Liu, and J. Cui. Threshold dynamics and ergodicity of an sirs epidemic model with markovian switching. *Journal of Differential Equations*, 263(12):8873–8915, Dec 2017.

[48] E. Lorenz. Predictability: a problem partly solved. In *Seminar on Predictability, 4-8 September 1995*, volume 1, pages 1–18, Shinfield Park, Reading, 1995. ECMWF, ECMWF.

[49] E. N. Lorenz. Deterministic nonperiodic flow. *Journal of Atmospheric Sciences*, 20(2):130 – 141, 1963.

[50] S. Luzzatto, I. Melbourne, and F. Paccaut. The Lorenz Attractor is Mixing. *Communications in Mathematical Physics*, 260(2):393–401, Dec. 2005.

[51] S. MacNamara and G. Strang. Operator splitting. In *Splitting methods in communication, imaging, science, and engineering*, Sci. Comput., pages 95–114. Springer, Cham, 2016.

[52] F. Maggi. *Sets of Finite Perimeter and Geometric Variational Problems: An Introduction to Geometric Measure Theory*. Cambridge University Press, 1 edition, Aug 2012.

[53] J. C. Mattingly. On recent progress for the stochastic Navier Stokes equations. In *Journées "Équations aux Dérivées Partielles"*, pages Exp. No. XI, 52. Univ. Nantes, Nantes, 2003.

[54] J. C. Mattingly, A. M. Stuart, and D. J. Higham. Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stochastic Process. Appl.*, 101(2):185–232, 2002.

[55] A. McNabb. Comparison theorems for differential equations. *Journal of mathematical analysis and applications*, 119(1-2):417–428, 1986.

[56] O. Melikechi, A. Young, T. Tang, T. Bowman, D. Dunson, and J. Johndrow. Limits of epidemic prediction using sir models. *Journal of Mathematical Biology*, 85, 09 2022.

[57] S. Meyn and R. L. Tweedie. *Markov chains and stochastic stability*. Cambridge University Press, Cambridge, second edition, 2009. With a prologue by Peter W. Glynn.

[58] B. Mityagin. The zero set of a real analytic function, 2015.

[59] P. Monmarché. On $\mathcal{H}^1$ and entropic convergence for contractive PDMP. *Electron. J. Probab.*, 20:Paper No. 128, 30, 2015.

[60] F. Morgan. *Geometric measure theory*. Elsevier/Academic Press, Amsterdam, fifth edition, 2016. A beginner's guide, Illustrated by James F. Bredt.

[61] T. Nagano. Linear differential systems with singularities and an application to transitive Lie algebras. *Journal of the Mathematical Society of Japan*, 18(4):398 – 404, 1966.

[62] L. I. Nicolaescu. The coarea formula. 2011. `https://www3.nd.edu/~lnicolae/Coarea.pdf`.

[63] M. Ninomiya and S. Ninomiya. A new higher-order weak approximation scheme for stochastic differential equations and the runge-kutta method. *Finance and Stochastics*, 13(3):415–443, 2009.

[64] S. Ninomiya and N. Victoir. Weak approximation of stochastic differential equations and application to derivative pricing. *Applied Mathematical Finance*, 15(2):107–121, 2008.

[65] E. Nummelin. The discrete skeleton method and a total variation limit theorem for continuous-time Markov processes. *Math. Scand.*, 42(1):150–160, 1978.

[66] E. Nummelin. *General irreducible Markov chains and nonnegative operators*, volume 83 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1984.

[67] S. H. Strogatz. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering*. Westview Press, 2000.

[68] H. J. Sussmann. Orbits of families of vector fields and integrability of distributions. *Transactions of the American Mathematical Society*, 180:171–188, 1973.

[69] B. Williamson. *On SDEs with Partial Damping Inspired by the Navier-Stokes Equations*. PhD thesis, Duke University, May 2019. https://hdl.handle.net/10161/18773.

Biography

Omar Melikechi graduated from Dover High School in 2009, earned a B.A. in Government from Dartmouth College in 2013, and studied in the Department of Mathematics at the University of Arizona from 2015 to 2017. In August 2017 he started his Ph.D. in Mathematics at Duke University. Omar will be a Postdoctoral Fellow in the Department of Statistics at Duke University beginning in January 2023 and then in the Department of Biostatistics at Harvard University beginning in August 2023. He is the coauthor of two published works [20, 56] and two preprints [2, 3].